Chromoneurodynamics, Ambiguity and Creative Cognition: A new approach of study in the realm of Statistical and Quantum Physics

SOUPARNO ROY

Chromoneurodynamics, Ambiguity and Creative Cognition: A new approach of study in the realm of Statistical and Quantum Physics

Thesis submitted for
Degree of Doctor of Philosophy (Science)
Jadavpur University

SOUPARNO ROY
Department Of Physics
Sir C.V. Raman Centre for Physics & Music
Jadavpur University

JADAVPUR UNIVERSITY
KOLKATA – 700 032

CERTIFICATE FROM THE SUPERVISOR(S)

This is to certify that the thesis entitled "Chromoneurodynamics, Ambiguity and Creative Cognition: A new approach of study in the realm of Statistical and Quantum Physics" submitted by Sri Souparno Roy, who got his name registered on 22nd September 2015 (22.09.2015) for the award of Ph.D. (Science) degree of Jadavpur University, is absolutely based upon his own work under the Supervision of Prof. Dipak Ghosh and Dr. Ranjan Sengupta and that neither this thesis nor any part of it has been submitted for either any degree/diploma or any other academic award anywhere before.

PROFESSOR DIPAK GHOSH
EMERITUS PROFESSOR
SIR C.V. RAMAN CENTRE
for PHYSICS and MUSIC
JADAVPUR UNIVERSITY
Kolkata - 700032

Dipak Ghosh 28/09/2022

1. Prof. Dipak Ghosh

SIR CV RAMAN CENTRE
for PHYSICS and MUSIC
Jadavpur University
Kolkata - 32. India

Ranjan Sengupta 28/09/2022

2. Dr. Ranjan Sengupta

(Signature of the supervisors & date with official seal)

dedicated to…

that small passage of time between being amazed at a trick and figuring out how it's done, for we, the students of science, are perpetually stuck there…

TABLE OF CONTENTS

CHAPTER 5

A NOVEL APPROACH TO CHARACTERIZE THE COMPLEXITY OF RAGA AND ITS RENDITIONS USING STATISTICAL METHODS

CHAPTER 7

AN IN-DEPTH QUANTITATIVE EXPLORATION OF ABSTRACTNESS OF IMPROVISATION INHERENT IN INDIAN CLASSICAL MUSIC USING STATISTICAL PHYSICS APPROACH

CHAPTER 8

NONCLASSICALITY IN THE COGNITION OF AUDITORY STIMULUS WITH AMBIGUOUS MUSICAL PROPERTIES

ACKNOWLEDGEMENT

First and foremost, the individual whom I would like to express my deepest and most sincere gratitude to is Prof. Dipak Ghosh, my supervisor and my guide in this long academic journey. Championing for the inherent harmony in nature and natural laws, he travelled from the depths of Nuclear Physics to the cliffs of Neurocognition, inspiring multiple generations of students, including me. He, continually and convincingly, conveyed a spirit of adventure in his teaching and his ideas. Prof. Ghosh spearheaded pioneering research in neuro-cognitive aspects of visual and auditory stimulus from the perspective of hard-core physics in the excellent laboratory of Sir C. V. Raman Centre for Physics and Music which I am privileged to be a part of. Without his constant guidance and help, this thesis would not have been possible. I can’t imagine having anyone else as my friend, philosopher and guide to start my academic career. Thank you, Sir, for being you.

My Co-supervisor, Dr. Ranjan Sengupta, an accomplished scientist in the fields of physics as well as music, has been a continuous source of encouragement. His knowledge and experiences have enriched me and my research in a number of ways for which I will remain forever grateful to him.

I feel deeply indebted to my seniors and co-researchers, Dr. Shankha Sanyal and Dr. Archi Banerjee, for their perpetual help and support that made the studies described in the thesis possible. From discussing ideas at length to computing statistical analysis, their contributions enhanced the quality of research in more ways than one. They were, and remain, more than friends, inside the laboratory as well as outside.

I would like to extend my sincere thanks to senior and fellow researchers of the laboratory including, but not limited to, Dr. Anirban Patranabis, Dr. Kaushik Banerjee, Dr. Tarit Guhathakurata, Dr. Alakananda Ghosh and Ms. Chandrima Roy. Their helpful advice and assistance in writing the papers and collecting experimental data are deeply appreciated. I also thank Mr. Sadananda Pramanik who was a saviour in multiple occasions for laboratory related issues. I would like to acknowledge the academic enthusiasm that was provided by a bunch of extremely talented young group of researchers that includes Sayan Nag, Uddalok Sarkar, Chirayata Bhattacharya, Ratul Ghosh, Pijush Kanti Gayen and many more. I found it extremely enriching to work and discuss several ideas with them, basking in their infectious exuberance.

I had the privilege to cross paths with nationally and internationally respected scientists and eminent personalities like Prof. Partha Ghose, Prof. Subhendu Ghosh, Late Prof. Asok Datta, Prof. Bikash Sinha, Prof. Hima Bindu Maringanti and Pandit Ajoy Chakraborty. Watching them up close and getting to interact with them were some of the highlights of my tenure as a PhD scholar. I would like to take this opportunity to express my sincere regards for letting me pick their brains whenever I approached them. I am also grateful to Prof. Elio Conte, Director of School of Advanced International Studies on Applied Theoretical and non-Linear Methodologies of Physics, Bari, Italy and Prof. Charles Spence, Head of the Crossmodal

Research Laboratory, Oxford University for taking an interest in our research work and being keen to share their expertise on the subjects.

I would like to thank some of the Professors from my very own Department of Physics, Prof. Argha Deb, Prof. Sukhen Das and Prof. Sanjay Kumar who offered their support whenever I needed them. Though this space falls short to name many of my family, my friends and colleagues from my university days with whom I discussed research ideas and knowledge, I would like to recognize their contribution to this thesis as well.

None of the work would have been possible without the funding agencies. I would like to acknowledge Jadavpur University for the UGC Non-NET fellowship and Department of science and Technology, West Bengal, for providing the research fellowship under the scheme 'Neuro cognitive Physics – Emotion using Bio-sensors' which granted the necessary funding to carry out the research.

Last, but not the least, I reserve this space to specially acknowledge my parents, Aruna Roy and Late Tapas Kumar Roy, whose unfathomable love and support have always been my source of strength and inspiration. My father, unfortunately losing after a long battle with his ailments while I was pursuing this degree, and my mother have lent their unconditional support in all my endeavours throughout the journey and although words can never do justice in describing their contributions to my life, this passage remains as a small token of offering to them. Finally, to my father, who I know would have been immensely delighted today despite trying to hide it under his otherwise stoic demeanour, using the words of the great Bard of Bengal Rabindranath Tagore – *"Death is not extinguishing the light; it is only putting out the lamp because the dawn has come."* This is for you, Baba.

Souparno Roy

September, 2022.

THESIS ORGANIZATION

Chapter 1 is the preamble to the interesting journey that lies ahead, through the complex world of neuro cognition and also the stimuli that provokes the action. Mankind's continuous struggle to understand '*Mind*' and its relationship to the brain dates back to the days of Pythagoras and Plato. From the philosophical abstraction, to Descartes' notion of self-aware system, to the 19^{th} century's neurobiological revolution that embedded the psychological attributes into physiological origins – the road has spawned countless debates. Numerous answers were found and even more questions got raised. Cognition, perception, emotion, language, memory, sleep and above all, consciousness – multiple facets of the 'mind problem' has emerged. Some claimed that mind is a function of the brain and an offshoot of its neuronal activity. To this, some doubted that the structural intricacy of 10^{12} neurons might not be sufficient to account for functional aspects of cognition. For some, cognition is a self-aware trait of an organic system sharing a complex symbiosis with its environment. This chapter welcomes the reader into a voyage that explores the nooks and crannies of several of the historic as well as modern takes regarding cognition and its various aspects.

But whatever theory one likes to subscribe to, it can't be denied that sensory information, visual, auditory, tactile or olfactory – shapes the sense of self via cognition. This fact remains true irrespective of whether it can be explained by either structural approach or functional approach, or both. Now, each sensory modality is a separate information channel, carrying separate kind of messages. Some evoke memories, some may trigger feelings and emotions. Hence, along with respective effects on the cognitive domain, also analysing the nature of sensory input may well improve the chances of further understanding the perceptual process. This, in turn, moves our focus to the binding factor that helps engrave the inputs: Emotion. Largely an outcast in cognitive discussions up until recent times, emotion has found a new paradigm when modern psychological and evolutionary studies have shown that preferential processing of environmental stimuli, an evolutionary advantage, is achieved via emotional enhancement of attention or memory. In this chapter, the role of emotion in different perceptual processes is explored, along with the theories regarding it.

One of the best stimuli that induces emotion is music, which itself is self-similar and scale-free. The process of human cognition facilitates scales and similarity. Its impact on the neurological level can be investigated using neuro-scientific biosensors like EEG, EMG, GSR, fMRI etc. the bio-signals emanating from these also exhibit self-similar nature. Physical and statistical methods that are put to study these complex systems are also discussed in the chapter.

Lastly, reducing the psychological into structural bounds inside the brain has a major limitation - the natural laws motivating this idea has been proven to be limited in the light of the theoretical advancements in basic Physics in the later part of the 20th century. Contemporary physical theory differs profoundly from classical physics, contradicting the older idea that local mechanical processes alone can account for the structure of all observed empirical data. This development has provided an alternative conceptual framework to analyse and describe neural and perceptual processes, which is founded on one of the most empirically successful theories of physics, i.e., quantum theory, and has introduced subjective and objective aspects of psychological constructs, tied by rules that directly specify the causal effects upon the subject's brain of the choices made by the subject, without needing to specify how these choices came about. The final sections of this chapter are dedicated to review such ideas and their potential in explaining cognitive conundrums.

Overall, this chapter provides an extensive inspection of the concepts mentioned and attempts to bind multitude of ideas into one seamless thread, working as a conceptual base on which the rest of the thesis builds itself.

Chapter 2 deals with various techniques associated with the analysis of the structures of music signals as well as bio-signals obtained from Encephalography data. A detailed analysis on the following tools of complex data analysis have been presented here which is used later in the thesis for different studies:

i) Empirical mode decomposition (EMD)

ii) Detrended fluctuation analysis (DFA).

iii) Multifractal detrended fluctuation analysis (MFDFA)

iv) Multifractal cross correlation analysis (MFDXA)

v) Maxwell-Boltzmann (MB) analysis

vi) Bose-Einstein (BE) analysis

Four of these techniques (i-iv) make use of Fractal Dimension (FD) or multifractal spectral width (obtained as an output of the MFDFA technique) as an important parameter with which the affective arousal corresponding to certain cognitive task (visual or auditory) can be quantified. EMD is a decomposition method for non-stationary and nonlinear signals such as EEG. Detrended Fluctuation Analysis (DFA) is used to analyze the long-range temporal correlations (LRTC) of the observed fluctuations in the signal. Advancing from DFA, MFDFA is a fractal analysis technique which is highly scale dependent. Here, signal components of different scaling ratios are analysed together in the form of a multifractal spectral width which indicates the degree of complexity of the signal. Lastly, MFDXA is an important tool with which the degree of cross correlation between two non-linear EEG signals originating from different lobes of brain can be quantitatively measured during higher order cognitive tasks. With this, we can have a quantitative assessment of how the different lobes is cross-correlated during higher order thinking tasks and the perception of external stimulus. A higher degree of cross-correlation would imply similarity between the signals in certain aspects. This in turn can be used to obtain a cue for the informational connectivity human brain displays while processing real world stimuli.

The rest of the two methods discussed here (v-vi) rely on the fact that the distribution of basic units among the complex signal under observation follows power law distribution pattern. Analogies of the distribution of said units with ensembles of distinguishable (for MB) and indistinguishable (for BE) gas particles in a container enables us to use the physical properties of MB and BE distributions in describing their dynamics. This approach is novel in the field of acoustic stimuli, especially in Indian Classical Music scenario. The emergent parameters could prove to be of great importance in categorization and classification of structured data regarding acoustic and bio signals.

Chapter 3 discusses the relationship between color and music as part of the complex system consisting of visual and auditory domain. Influences of one on the other, across modalities, has been investigated in this chapter. Though both the stimulus forms are processed in different parts of the brain, there have been numerous reports about their correlation. It will be really interesting to examine whether they share a similarity in perception. The interesting incident of sensory arousal due to multi-sensory stimulus is scientifically identified as synesthesia. But Studies have shown that non-synesthetic people also have visual-to-auditory associations. These are termed as cross-modal correspondences. Needless to say, that color and music both have strong impact on emotion and feelings & also a few studies have been reported in literature to explore causal relationship between color and emotion.

The work detailed in this chapter reports a neuro-cognitive study on response of brain to two different stimulus and their cross-modal associations. Here, the correlation between emotional arousal and the effect of audio and visual stimuli has been studied from a new perspective. 93 participants were asked to hear 6 different music pieces (each of 30 second duration). The type of emotion elicited by different music pieces were identified by the participants from a given collection of possible emotional responses, constructed from the *Rasa* module of Indian school of aesthetics. Then they are asked to assign a color associating the emotion from a given color wheel (structured according to Munsell color system/RGB color space). Each color, associated with a particular music piece, is a mixture of specific Red, Green and Blue values (RGB triplet) and has a specific HEX number (hexadecimal representation), which is recorded for each response. Then, the musical pieces used were further zoomed with the help of fractal technique to identify different emotions related to music in a quantitative measure. Here, to analyze the complexity of the sound signal (which are nonstationary and scale varying in nature), we have used Multifractal detrended fluctuation analysis (MFDFA), which is capable of determining multifractal scaling behavior of non-stationary time series. From the experimental data, it was found that the color-music association is not domain-specific, i.e., they connect across modalities. Also, it was interesting to see that such association was mediated by the emotional arousal invoked due to the presence of long-range correlations in the music. The affective responses correspond to the multifractal width (a measure of the degree of complexity present in the signal) in a specific manner. So does the R, G and B components. Higher stimulus complexity corresponds to higher R component, and lower complexity tends to favour higher B component. In fact, a transition phase from one component to another has also been identified. The fact that the visual and emotional response to the auditory stimulus follows a specific trend which is directly related to the stimulus complexity, is a novel and unprecedented finding. This also consolidates the effectiveness of fractal approach to cross-modal researches which will certainly enrich our knowledge on stimulus perception further.

Chapter 4 delves deep into color perception which is a major guiding factor in the evolutionary process of human civilization. Yet, most of the neurological background of the same remain unknown. In this chapter, we attempt to address this area with an EEG based neuro-cognitive study on response of brain to different color stimuli. Our aim is to investigate the neuronal complexities that arise in the brain when it receives various colors as a visual stimulus. With respect to a Grey baseline seven colors of the VIBGYOR (Violet, Indigo, Blue, Green, Yellow, Orange, Red) were exposed to 16 participants with normal color vision. Grey sets a baseline for comparison. The corresponding EEG signals from different lobes (Frontal, Occipital & Parietal) were recorded. Five Frontal electrodes (F3, F4, F7, F8, Fz), two Parietal (P3, P4) and two Occipital (O1, O2) electrodes were analysed since these areas are mostly reported to be associated with cognition and perceptions of visual stimulus. For the quantitative analysis of the collected EEG data, we have applied two fractal-based non-linear techniques MFDFA and MFDXA (Multifractal Detrended Fluctuation Analysis and Mutifractal Detrended Cross-correlation Analysis). MFDFA assesses the degree of complexity present in the signal using multifractal width as a parameter. Higher the width, higher the long-range cross-correlations present in the series, implying higher complexity. MFDXA, on the other hand, measures the degree of how much correlation is present between various inter and intra lobe electrodes in the EEG signals using a cross correlation co-efficient (γ_x). MFDFA revealed that for all the participants the spectral width, and hence the complexity of the EEG signals, reaches a maximum when exposed to color Blue, followed by colors Red and Green in all the brain lobes. MFDXA, on the other hand, suggests a lower degree of inter and intra lobe correlation while watching the VIBGYOR colors compared to baseline Grey, hinting towards a post processing of visual information. We hope that along with the novelty of methodologies, the unique outcomes of this in-depth study may leave a long-term impact in the domain of color perception research.

Chapter 5 marks the initiation of an attempt to categorize a complex information assemblage in terms of its constituent basic units using tools of equilibrium statistical mechanics. Now, majority of systems found in nature is usually made up of repetitive arrangements of basic patterns. The presence of such sequences contributes to the unique style/behaviour the information congregation it represents. Various such congregations like music, language, biological signals like heartbeat rhythm, nitrogen bases of DNA and RNA exhibit this kind of repetitive patterns of symbolic sequences. The problem of categorizing such information collection is that the usual methods used are mostly non-quantifiable. In this study, we try to quantify such abstractness using measurable parameters. For that, we introduce methods based on well-known concepts used in Statistical Physics (especially thermodynamics), namely Maxwell-Boltzmann (MB) statistics and Bose-Einstein (BE) distribution, in an attempt of categorization and classification of musical information present in Hindustani classical music. Both MB and BE statistics have wide applications outside the realm of describing the energy level occupation of elementary particles. For example: the domains of linguistics or social sciences. Here, statistical methods based on these distributions have been applied to find new parameters (equivalent to 'temperature' in physical systems) to distinguish between different features of different ragas in Hindustani classical music.

To apply MB statistics to music, it is assumed that different notes (combined with their durations) with different occurrence frequencies are at different energy levels, the distribution of which follows the MB distribution pattern. This approach produces several parameters, one of which is so called Maxwell-Boltzmann temperature or T_{MB}. The notion of such temperature has been brought up by Benoit Mandelbrot while dealing with literary examples. In this case, the emerging 'temperature' parameter is compared with a reference temperature set using all the note-duration combinations present in all the music samples under study, our working 'musical corpus'. This comparison shows how close the said rendition is with the traditional grammatical structure of the said raga (or the amount of improvisations/creativity present in the rendition). The MB approach recognizes each of the constituent units as individual molecules, so to speak, denoting their distinguishability. On the other hand, BE distribution deals with low frequency data. In case of BE statistics, a rank-frequency distribution of the time durations of various notes occurring in different ragas is studied. Then it was reassembled by comparing the occurrence number with an energy level and the occupant number with the particles sharing the same energy level. It sheds the individual identities of the constituent units and considers them indistinguishable. From this, we obtain 'temperature-alike' (T_{BE}) statistical parameter that denotes the degree of diversity (regarding a certain property) present in the sample. This novel method of analysis is applied on different renditions of the same Raga. Music clips chosen were the *Vandish* part of the same raga, *Marwa*, sung by three legendary classical music maestros. The resulting analysis gives a number of parameters (they come from the analogy between the rank-frequency distribution and the respective statistical distribution) that help categorize the singing styles of the three artists and parameters which are indicative of abstract ideas such as individual improvisation pattern. The results found shed new lights on the classification of artists' style and the difference in their renditions of the same raga. Such an approach to harvest information from an acoustic stimulus (Indian or otherwise) is unprecedented in the global scenario. Also, the outcomes were interesting enough to deserve to be applied to more specific tasks in the classical music genre which we elaborate in the upcoming chapters.

Chapter 6 extends our endeavour in categorizing musical data using statistical physics tools to new horizons. In this chapter, we focus our attention on the most important ingredient that is the backbone of North Indian (or Hindustani) Classical Music – Raga. Raga, a tonal multifarious module, is basically a well-defined melodic structure that consists of a series of four/five (or more) musical notes. For every Raga, the tonic 'Sa' is the most important note. Other fundamental elements include: a pair of standing notes (*Vad-Samvad* notes, usually in a fifth harmonic relationship), steady and longer notes that are sung at the end of a phrase (*Nyas*), presence of ascending/descending note patterns (*arohana/avarohana*), different kinds of transitions between the notes (*Meend, Andolan, Gamak* etc.) – also known as ornamentation. Also, the performer is allowed to improvise liberally, unlike western classical music where artist has to stick to specific compositions. The goal of a raga performance is to perform it such a way so it evokes and consolidates an emotional state in the mind of the listener. Hence, it remains a question that how do a raga invoke emotions and how closely its acoustic framework and structural features contribute to said induction. Existing works mainly focus on raga identification, which has an application in AI domain and music digitization. In this chapter, we will focus not only on characterizing ragas based on their elemental units, but to quantitively analyse some structural features which were rendered abstract until now.

Our approach is based on the information-related energy. The central idea is that different elemental units with different occurrence frequency has different 'energies' and their energy distribution follows Maxwell-Boltzmann (MB) distribution. The MB distribution is related to the relevant power law distribution (Zipf's law) via direct correlation of the Hamiltonian with the rank of the element (explained in chapter 2). This treatment gives rise to certain parameters, one of which is equivalent to temperature (T_{MB}). Also, apart from MB distribution which deals with higher frequency elements, lower frequency ones can be compared to boson gas under grand canonical approach. The basis of the idea is reflected in the fact that the rank-frequency distribution of these elements show similarity with Bose-Einstein (BE) distribution pattern. A number of parameters also spawns out of this analysis including temperature-like parameter T_{BE}. We use these approaches on music samples from the 'Alap' part of three different ragas (*Marwa*, *Puriya*, *Sohini*) sung by Pandit Ajay Chakrabarty, a legendary classical music maestro. All of the chosen three ragas are based on the following note structure: Sa, *komal* Re, *shuddha* Ga, *tivra* Ma, *shuddha* Dha, *shuddha* Ni. Despite that, they evoke very different emotions among the listeners. Our objective is to identify structural cues that help explain it. From the resulting parameters, we were able to parametrize kinetic natures of the three ragas, complexity in their note usage pattern, and how closely they follow the characteristic notes described above. Also, the degree of ornamentation and 'musical analyticity' (the tendency to resist using inflective movements) can also be quantified. Using another parameter, τ, the ragas can be categorized according to their analytic levels. Such properties are directly related to their emotion-inducing effects and remained mostly qualitative in previous literature. The methods studied here are novel in the music research field and can prove to be useful in the fields of music and speech as quantifiers in categorization problems.

Chapter 7 attempts to address one more avenue, very specific to both Indian Classical Music and its affective response. It is the practice of improvisation. Improvisation usually refers to those features or details which are extemporaneous, added by the performer without changing the composition's identity. It is a pivotal part of Indian classical Music whose expression relies upon the imaginative insight of a particular artist. Improvisation is also of paramount importance in invoking any aesthetic/emotional experience. But unlike other acoustic features, it remains an abstraction and usually described in a qualitative manner. The signature of improvisation can be detected in various parts of a raga rendition, such as, melodic expansion (*Vistar*), Rhythmic intensification, Pitch permutation, sequential transposition (*Alankar*) etc. Existing literature on improvisation in ICM are rare and scarce. Musicological roots of the techniques are studied but unlike the amount of hardcore scientific approach used in improvisation of western Jazz and folk traditions, same in the case of Indian classical music is rarely explored. With this background, we wish to study improvisational characteristics of ICM using the statistical physics approach. As explained in the previous chapters, our methods start from the point that the musical structure is analogous to a physical system and different elemental units with different occurrence frequency has different 'energies', whose distribution follows Maxwell-Boltzmann (MB) statistics. This is used for the data with high frequency of occurrence. In case of low-frequent data, their rank-frequency distribution is compared with a boson gas and Bose-Einstein distribution is applied. These methods help in quantifying the necessary properties which could help parametrize signatures of improvisation. Since the improvisation is rendered more accessible when a piece of performance is compared to another rendition of the same piece performed by the same artist, we have decided to analyze multiple performances of the same raga performed by the same artist. Our chosen samples belonged to raga '*Sur Malhar'*, performed by Pandit Kumar Gandharva, one of the most renowned vocalists of ICM. We used six different performances of the *Vilambit Bandish* segment of the raga sung by the same artist in six different occasions to examine their improvisational differences. *Bandish* holds the compositional structure and characteristic notes specific to the raga which is why we chose this part as the note usage in all the renditions would remain the same. Therefore, the consequent changes in the structural patterns would be because of the lyrical or melodic improvisations included by the artist each time he performed the raga. The resulting analysis proved that such subtleties exist in each raga rendition that makes each of them different from the other. It can be distinguished via the resulting parameters. To confirm our results, we also consulted a classical music expert whose analysis was compared to the ones found from the statistical methods. It was seen that the accuracy of the parameters was satisfactory in quantifying various musical properties and successfully categorizing the samples according to their improvisational traits. We could also identify possible correlations between a transitional characteristic called '*Meend*' and corresponding improvisation. The results are promising and the emerging parameters have immense potential to be used as quantifiers in studying improvisational characteristics of raga renditions.

Chapter 8 presents an interesting take on auditory stimulus perception showing how an induced ambiguity can result in the breakdown of classical Boolean logic. We start with the hypothesis that cognitive phenomena regarding ambiguous auditory stimuli can't be described using classical theories. In classical physics, the well-known formula of total probability is represented by the sum of discrete probability values of discrete events including the respective conditional probabilities of A and B, if they are two dichotomous random variables. Here we seek to examine whether an induced dichotomy in auditory signal makes the cognitive properties adhere to this classical notion of total probability. For this study, as auditory stimuli we took two pairs of contrast emotion (happy-sad) music clips (each clip of 30 seconds duration and absolute tempo or BPM fixed for each pair) and for the perceptual experiment, we took total 100 participants, divided them into two equal groups (say Group 1 and Group 2), and posed two different experimental conditions. Group 1 was exposed to a pair of music clips (having similar tempo but contrasting musical structure, separated by five seconds) and were asked whether the two clips had the same tempo. Group 2 was made to listen to a different such pair first, then followed by the previous pair after 15 seconds and was asked the same question for each pair. To eradicate musician/non-musician bias, participants were asked not to use any means of tempo measurements, including finger/foot-tapping. The answer 'yes' to the posed question on Pair A gave p(A+), and 'no', gave p(A-). Responses from Group 2 gave us probabilities p(B+), p(B-), p(A+/B+), p(A-/B+), p(A-/B+), p(A-/B-). If classical theories hold true, the collected data should match the law of total probability. But our analysis clearly revealed that for Group 2, the calculated values of p(A+) and p(A-) using the law of total probability, i.e., p(A±) = p(B+). p (A±/B+) + p(B-). p(A±/B-), featured significant difference from the p(A+) and p(A-) values calculated for Group 1. Here, we have found an intriguing result of violation of classical probability laws in case of ambiguous auditory signals. This data agrees with previous reports made on visual stimuli and hence, registers a significant claim of the presence of non-classicality in cognitive properties.

Chapter 9 being the concluding chapter of the thesis, it expands itself into horizons hitherto unexplored. Inspired by the ideas discussed in chapters 1 and 8, about the mind vs brain, structural vs functional debate, this chapter presents a novel hypothesis on an aspect of cognition that resonates with other facets like perception, emotion and aesthetics closely. It is the creative cognition, or in general, creativity. Creativity is defined as 'the tendency to generate or recognize new ideas or alternatives and to make connections between seemingly unrelated phenomena'. It is a pillar upon which human civilization is built and further advanced. Several theories over the years have tried to encapsulate this important aspect of cognition. One general consensus is found in the 'Free association theory'. It says, the more freely a person's conceptual 'node's are connected, the more divergent thinker he or she is. Divergent thinking, loosely called 'lateral thinking', is the spontaneous and free-flowing process by which unexpected conceptual connections are made to overcome a problem in a novel way. It is often related to creative cognition. Also, tolerance of ambiguity is found to be related to divergent thinking. The resolution of ambiguity is an inseparable part of our everyday interactions with the world: lexical ambiguity, semantic ambiguity, ambiguous pictures and facial expressions etc. It is also been associated with aesthetics and creativity. The idea of deviational aesthetics is based on taking conceptual elements out of their usual context and reassemble in a new and unexpected way.

In this chapter, we approach the problem of creativity from a theoretical physics standpoint. Theoretically, for the initial conceptual state, the next 'jump' to any other node is equally probable and non-deterministic. And to study such a probabilistic system, Quantum theory has been proven the most successful, time and again. We suggest that this collection of nodes form a system which is likely to be governed by quantum physics and specify the transformations which could help explain the conceptual jump between states. Our argument, from the point of view of physics is that the initial evolution of the 'creative process' is identical, person or field independent. To answer the next obvious question about individual creativity, we hypothesize that the quantum system, under continuous measurements (in the form of external stimuli) evolves with chaotic dynamics, hence separating a painter from a musician. This is the gist of our hypothesis, Quantum Leap Interpretation, which is elaborated in this chapter. Also, its advantages and relation to other prevalent theories are discussed.

CHAPTER 1

PREAMBLE:

‘COGITO ERGO SUM’

“Biology gives you a brain. Life turns it into a mind.”

Jeffrey Eugenides

ABSTRACT

Humans interpret nature through their sensory modalities and cognitive capabilities. Nature is a machinery that has an incredibly complex working algorithm. But, at the same time, it presents an elegant face that makes the reality around us worth surviving for. And it is our sensory organs, working in synchrony with the cognitive core, that provide us the quintessential vessel to travel this sea of complexity. From gargantuan celestial bodies to esoteric music-art-poetry, everything makes sense because our intrinsic system makes it sensible. It works across scales – from macro to micro – and across time – from polypeptide vibration (hundreds of thousand times per second) to genetic mutation (0.5×10^{-9} per base pair per year) – and even across tangibility – from complex models of architecture to complex models of emotion. Yet, the framework of our guiding system remains elusive to us. Little has been known, even less agreed upon. How does cognition emerge from a biological hot soup called brain? How is visual stimulus perceived – fragmented, or integrated? What happens when different sensory modalities cross paths – does it help or hinder sensory perception? What is emotion and is it actually subjective? How does perceptual ambiguity and uncertainty bode on aesthetic processes and/or creativity? Questions like these are yet to be met with answers conclusive enough. In this introductory chapter, we will discuss the prevailing ideas and debates surrounding both the psychological and physiological correlates of Mind, perception, emotion, cognition, creative aesthetics and ponder upon their complexity. Also, we will discuss different statistical methodologies that will help us in multiple occasions in the thesis to analyze complex systems like bio-signals and acoustic signals – from power law and related distributions to fractal-based scaling analysis techniques. Indeed, these techniques have been very useful in extracting features from the face of complexity. But, to discuss the emergence of it in cognitive domain, one needs a theory that is forged in the depths of psychology and involves the same in its interpretations – the quantum theory. Hence, the chapter's concluding notes review all the quantum approaches made in regard to resolve various cognitive conundrums. In due course of the thesis, we shall put these methodological aspects - both statistical and quantum – in extensive use to explore emotion's role in multimodal perception, to identify complexity features of visual perception, to categorize as well as classify acoustic information, to recognize non-classical attributes of auditory perception and to hypothesize the emergence of creative cognition. This introductory chapter acts as a preface to the works that follow.

Keywords: Cognition, Color perception, Emotion, Crossmodal perception, Mind, Ambiguity, Creativity, Quantum theory, Chaos, Fractals.

1.1. NATURE'S CREATION AND HUMAN APPRECIATION THROUGH MIND, BRAIN AND OTHER ORGANS

- "What is mind?"

- "No matter."

- "What is matter?"

- "Never mind."

Do you see the mind or does the mind see you? Mankind's continuous struggle to understand '*Mind*' and its relationship to the brain dates back to some six hundred years before the birth of Christ. Early Greek philosophers - from Pythagoras to Plato, separated by few hundred years - have returned back the subject time and again only to be restricted inside the philosophical boundaries of 'life-forces', 'humours' and 'immortal souls' (Russell, 2022). Hippocrates (460 B.C.), the father of modern medicine, was the first one who made the associations between brain, emotions and thoughts. But even he didn't address the question of 'Mind' directly. The issue remained largely elusive until Rene Descartes, in 17th century, announced the notion of primacy of consciousness, in which the mind knows itself by stating '*cogito ergo sum*' ('I think therefore I am') (Hansotia, 2003; Durant, 1961). With this, the emphasis of ancient Greeks on human's unique capacity of thought came to a full circle. On the other hand, it took some time for the structural knowledge of the brain to catch up to the developments of the mind problem. By late 19th century, using electrical stimulation and other techniques, several researchers showed that changes in parts of the brain affect specific cognitive functions, hence furthering the possibility of interrelations between brain and mind. In 1870s, John Hughlings Jackson, a London neurologist, proposed the revolutionary idea that brain was the 'organ of the mind', as in, cognition, emotion, memory, language – things that are associated with the concept of mind are found and regulated at various parts of cerebral lobes. He reasoned that the structure of the mind should strictly follow the structure of the nervous system, and hence, exhibit evolutionary levels connected by representation as the human brain structure does (York & Steinberg, 2011). Later, Russian scientist Ivan Pavlov, through his seminal experiments on behaviour conditioning, established that appropriate stimulation could influence behaviour positively or negatively. His work demonstrated that specific sensory stimulation could induce specific responses and hence, modify behavioural patterns. Psychologist John B Watson used this conditioned-reflex mechanism as the basis of all forms of learning, and framed the now famous theory of behaviourism.

The 'mind question' comprises of several separate (but interrelated) segments such as consciousness, cognition, language, perception, emotion, memory, sleep and dreaming – most of whose details are still obscured. Over the years it has been found that several of these functions face impairment or disintegrate progressively with the damage of specific cerebral areas (or shrinkage of the brain, in general) (Hansotia, 2003). This goes to show that the neural connections that form networks and circuits in the brain are essential to the existence of the mind. According to Schwartz and Begley (2009), this viewpoint forms the basis of the most supported schools of thought regarding the mind-brain question - Functionalism and Epiphenomenalism – both of which regard mind as a function of the brain and a product (or epiphenomenon) of its neuronal activity. To quote neurophysiologist Sir Charles Sherrington, brain is “an enchanted loom” where “millions of flashing shuttles weave a dissolving pattern, always a meaningful pattern, though never an abiding one” (Pinchot & Gersten, 1930).

Now, the neurobiological and neuroanatomical explorations of brain come at a cost. It begets scepticisms which Churchland (1980) has distinguished into two categories: ‘principled’ and ‘boggled’. The former argues that the neurological explanations will always, given the nature of the case, be inappropriate to our understanding of things such as perception, memory, and consciousness, whereas, the latter, boggled by the scale of intricacy inside the human brain, doubts that the structural complexity of 10^{12} neurons might lead to wrong generalisations. For the most part of the 20th century brain research, the prevailing notion has been that that particular part of the brain specialises in processing particular functions. This notion, fuelling the above-mentioned scepticisms, posit questions that how a collection of specialised functions alone could give rise to a coherent perception of the reality and consciousness. Without different parts of the brain communicating with each other to share and to combine information, this seems quite improbable. During the last decades a richer perspective of the brain has been introduced in which networks of segregated but interacting processes govern neural dynamics on top of the processing of the specialised regions. This capacity of processing information of different modalities separately (e.g., visual, auditory or olfactory) and simultaneously combining that information is necessary for the cognitive aspect to flourish. With the advent of neuroimaging instruments like fMRI and the advancement of classical EEGs, it is now possible to see the brain ‘at work’ and vigorously analyse its dynamics.

Still, the structural vs. functional debate have continued to hover over the modern brain research. Because cognitive states and processes are properly understood as functional states and processes, hence psychology in contrast to neuroscience, some believe, should be able to determine the functional organization that accounts for cognitive activity (Churchland, 1980). They argue that these processes cannot be reduced solely to physiological roots. Neuroscientist Antonio Damasio has reasoned that the brain, operating in an organism which interacts with its environment like an open system, processes sensory information and forms ‘mental images’ or ‘concepts’, assessing which gives rise to the sense of self in that organism-environment complex entity (Damasio, 2000).

The fact that sensory information – be it visual, auditory or olfactory – shapes the sense of self via cognition is undeniable, irrespective of whether it can be explained by neurobiological or psychological approach. Now, each human sense is sensitive to a different type of stimulation. Since each sensory modality may be considered as a separate information channel, not all of the incoming sensory information will necessarily communicate the same message. For instance, some information evokes memories or associations, some may trigger feelings and emotions (Schifferstein & Spence, 2008). Hence, along with their effect in brain, also analysing the nature of sensory input may well improve the chances of further understanding the perceptual process.

Also, emotion, a fundamental part of human evolution, is another important topic which has received attention. It is described as "...neurochemical processes by which the body monitors changes in its internal situation and is thereby alerted to the need for certain responses to its ever-changing states" (Johnson, 2001). There are sufficient evidences that emotion and cognition are interrelated (Pessoa, 2013; Dolcos et al., 2014; Okon-singer et al., 2015). What role emotion plays in the cognitive binding of sensory inputs is also a widely debated field of research.

Lastly, the idea that all causal mechanisms in relevance to neuroscience can be formulated in terms of the material particles inside the brain (thus, forsaking all intrinsically psychological contents) has one restriction - the natural laws motivating this idea has been proven to be limited in the light of the theoretical advancements in basic Physics in the later part of the 20th century. Contemporary basic physical theory differs profoundly from classic physics on the important matter of how the consciousness of human agents enters into the structure of empirical phenomena, contradicting the older idea that local mechanical processes alone can account for the structure of all observed empirical data (Schwartz et al., 2005). This development has provided the neuroscientists and psychologists an alternative conceptual framework to analyse and describe neural and perceptual processes to various degree of success (Busemeyer et al., 2006, 2011; Khrennikov, 2009; Aerts, 2012). This non-classical framework is founded on one of the most empirically successful theories of physics, i.e., quantum theory, and has introduced subjective and objective aspects of psychological constructs, tied by rules that directly specify the causal effects upon the subject's brain of the choices made by the subject, without needing to specify how these choices came about. For example, in the context provided in Ochsner et al. (2002), this intention-induced modulation of cerebral mechanisms is the key factor necessary for the emotional self-regulation seen in the active cognitive reappraisal condition.

In view of this, this thesis tries to study the brain dynamics under the influence of external visual and auditory stimulus in tandem with the complexity and structural significance of the stimuli themselves in provoking emotional response in the observing mind. We try to highlight following three facets: In the first part, we analyse the effects of auditory and visual stimuli in the brain in both neurobiological and cross-modal perspective using robust non-linear analysis of both the stimuli and the brain electrical activity. In the second, we attempt to develop advanced structural knowledge of the music signal (as a representative auditory stimulus) using fundamental statistical approach to predict the improvisational and dynamic nature (since they are instrumental in the possibility of invoking a range of emotional responses) of the said musical structure. In the final part, the psychological framework of cognition is explored through the non-classical approach and an attempt was made to interpret creative cognition process using quantum theory. All of these approaches are novel and have potential in advancing the boundaries of multiple fields like neuroscience, psychology and signal analysis.

1.2. CONCEPT OF MIND: INDIAN AND WESTERN APPROACHES

It might be fairly difficult to find any philosophers and free thinkers, both in Oriental and Occidental civilizations, who has not debated or discussed about the topic of mind and brain.

The terminology used may vary depending on person or context, but the central questions remain kindred nonetheless.

Indian approach:

In the Orthodox or *Vedic* section of Indian philosophy, mind (*Manas*) is one of the basic structural units of man, along with *Atman* (soul), *Indriyas* (sensory organs), *Sarira* (body) and, in some school of thoughts, *Jnana* (consciousness). Mind bridges the *Indriyas* with the *Atman* by carrying the sensory information to the soul. Also, it is the recurrent tendency of mind to attend different information in a continuous manner that gives rise to the 'stream of thought' (Vidyabhusana & Basu, 1913). Some schools of philosophy designate consciousness as an unintended product of the conjunction between the 'soul' and the 'mind'. According to *Vaisesika* school, mind is a substance which is atomic, unconscious and capable of action, helping in perception, cognition, pleasure, volitions (Atreya, 1985). In addition to this, it performs two fundamental functions, i.e., *Samkalpa* (desire to do an action) and *Vikalpa* (doubting and registering the sensory information) (Bala, 2009). Contrarily, *Mimamsa* school defines mind as intangible, dimensionless substance which is all-pervading and all-perceiving but incapable of causing or affecting any physical changes (Radhakrishnan, 1953).

The Upanishads described mind using terminologies such as *Manas*, *Prajna*, *Citta* etc. Here also, mind is a subtle form of matter and made up of physical substances like the body. The physical and non-physical world includes five sensory organs, five motor organs (hand, feet, speech organs, sexual organs and anus), mind (*Manas*), intellect (*Buddhi*), self-ego (*Ahamkara*) and subconscious (*Citta*). The hierarchy starts with the functions of sensory and motor organs, followed by mind coordinating their perceptions; intellect, the higher organ of thought, discriminates the self-ego, and finally, the subconscious which stores past impressions (Bala, 2009). Upanishad also describes a cognitive 'highest state': "When the five knowledges (*jnāna*) are settled (*ava-sthã*), together with the mind, and the intellect (*buddhi*) does not move about (*vi-ceṣṭ*), that they call the highest state (*gati*)" (Whitney, 1890).

There is another stream in Indian philosophy known as Heterodox or *Non-Vedic* philosophy. The only materialistic school of thought named *Carvaka* school fall under this category. *Carvaka* philosophy is based on a generally materialist metaphysic and an empiricist epistemology. The key claim of all materialism is that mind results from a particular and very special organization of matter. Carvaka school subscribes to this idea wholeheartedly: 'consciousness emerges when the material elements are combined in a certain way' (Chakrabarti, 1999). The non-materialistic argument against it says since the body is made up of atoms and particles lacking consciousness, then consciousness must be the result of something besides the body. *Carvaka* argues that fermented liquid is intoxicating even though it is made up of materials which aren't intoxicating (Radhakrishnan, 2009). According to them, mind merely regulates sensory perceptions, whose continuity causes the false impression of the mind as a substance. On the other hand, Buddhist school of thought, spanning some fifteen centuries, share similarities with the larger philosophical context that mental phenomena include various interconnected biological and cognitive processes. But what differ them from other philosophical schools is the fact that instead of associating mind with an independent self, Buddhism practices the doctrine of no-self. This postulates the denial of a permanent self as the agent of sensory and mental activity. The perceptual stream changes its contents every moment and leads to a dualistic conundrum of objective (external and empirically accessible) and subjective (internal, accessible only via consciousness) perspectives. Buddhism argues that the sense of self is basically an imaginary coping mechanism, constructed to compete against such conundrum (Giles, 1993).

Jainism, the final school belonging to the Heterodox section, doesn't emphasize on mind as much as the other school of thoughts do. Quite unlike Buddhist philosophy, Jainism describes the self or *Jiva* as the abode of conscience and limits the mind as only an internal tool which mediates sensory perception with the self and helps in performing intrinsic actions such as pleasure, pain, love etc (Atreya, 1985).

Western approach:

The western approach to tackle the problem of mind has been based upon the different interpretations of the relationship between mind and body. Believers of the idea that mind and body are not different from each other (mind is just an aspect of the body, located in or identical to the brain) are called *Monists*. And those who consider mind and body to be significantly separate entities (brain is not equivalent to mind) are known as *Dualists*.

Monism or the idea that mind and body are indistinguishable has branched into various forms. One such idea is *Physicalism* (sometimes branded as *Materialism*), which simply states that everything – biological, psychological, moral, social – are at the end of the day either physical or supervene on the physical (Lange, 1865; Armstrong, 1968). The corollary that follows is: since everything is physical, hence physical laws can lead us to the answers. Physicalism also introduces an interesting concept of 'supervenience'. When a painting is reduced to examine the singular brushstrokes it is made up of, one loses the wholesomeness of the painting. Hence, although it does consist of those strokes, it is not identical with them, i.e., the painting *supervenes* on the strokes. Similarly, while reducing mental to physical, it can be said that the mental *supervenes* on the physical (Horgan, 1982; Mclaughlin, 1984; Davidson, 2001).

Identity theory (Reichenbach, 1938; Feigl, 1958, Place, 1970), another form of Monism, posits that mental states are identical to specific brain states. That is to say, identification of mind (and mental states) can be achieved through identifying physical states in the brain. One of the main arguments against this proposition is the multiple realizability of mental states (Kripke, 1972), i.e., various different physical states could cause the realisation of same mental state (e.g.: similar pain could be realised due to various different physical states or events). To avoid such criticisms, theory of *Functionalism* was developed. The central idea of this theory is that mental states (beliefs, desires, etc.) are constituted solely to serve their practical role. So, instead of 'What is the mind?' one seeks to answer 'What does it do?' Functionalists claim that mental states can be identified with the function they have externally (Jackson, 1982, 1988).

Eliminative materialism (or *Eliminativism*) is a radical and much debated idea which denies specific types of mental states (common-sense states/folk psychology) which could be over or above physical states while considering the theory of mind. According to eliminativists, these states are poorly defined and have no neurobiological basis (Feyerabend, 1963; Churchland, 1981). They argue that the psychological concepts of behaviour and experience should be judged by reducing these to the biological level.

One form of Monism that stands out because of its converse interpretation of mind as immaterial and the ultimate foundation of all reality (physical concepts are explained in terms of mental states) is called *Idealism*. Idealists solve the mind-body problem easily, as they consider material realism a construction of mental states and hence eliminate the question of mind-body interaction. There are two variations of this idea: Absolute Idealism (that reality consists of one vast all-encompassing mind) and Subjective Idealism (that reality consists of plurality of minds). Renowned philosophers from German philosophical school such as Immanuel Kant (1998), JG Fichte (1993), FWJ Von Schelling (1988) and GWF Hegel (1977) has advocated for the former whereas George Berkeley (1948) argued in favour of the latter.

In contrast to the above forms of Monisms, i.e., the theory of mind/body being one entity, Dualism suggests that the mental and the physical – the mind and the body/brain – are, in some sense, radically different kinds of things. The most definitive statement of dualism is found in the philosophy of Descartes according to whom mind and matter are two separate and distinct sorts of substances, absolutely opposed in their natures and each capable of existing entirely independent of the other (Descartes, 1988). Hence, the question of how these two substances interact has spun off various forms of Dualisms over the years.

Interactionism is the view that mind and body—or mental and physical events—causally influence each other.The physical world influences one's experience through his senses and they react behaviourally to those experiences. This interpretation, although, has been criticised and debated upon its fallacy to conserve basic physical laws (Larmer, 1986). To escape the criticism regarding interactionism, *Epiphenomenalism* allows the physical to causally influence the mental but denies the opposite (Green, 2003). But apart from sounding counterintuitive (mental states do affect various physical events), this idea seems flawed since it can't explain the evolution of mental states. A solution was offered by Leibniz in terms of *Parallelism*: that both realms exist, but don't causally interact at all (Lopston, 2006). Instead, they run parallel, in harmony with each other. This idea roots for a pre-existing harmony, and hence, has been of little use outside theological perspectives.

To sum up, although the central questions have remained similar in terms of the mind problem, the approaches used in Indian and Western school of thought are quite different. The Indian psychological works, mainly of the ancient and medieval times, have not separated themselves from philosophical treatises. As a result, the theories on the nature and functioning of mind have remained partially (or entirely) correlated with the nature of reality and experience. On the other hand, the Western approach keeps the psychological enquiries detached from the ontological aspect of mind. There, it is common practice to exclusively handle the problem in an analytical manner and regard the conclusions thus arrived as the final outcome. The synthetic approach of Indian schools (of trying to interrelate and harmonise the analytical findings using common underlying principles) and the detached analytical approach of Western schools has remained their principal characteristics toward the concept of mind question.

1.3. CONCEPT OF EMOTION

Quite similar to the mind (and mind vs. brain) problem, the question regarding the nature of emotions has also been dividing opinions among thinkers throughout the history. Since the pursuit of reason flourished in ancient Greece, mostly due to Socrates and his disciple Plato, theories about human emotions have thrived as an antithesis. Although Plato and his school of thought had famously campaigned for reason and neglected emotions, Aristotle through his writings put certain emphasis on it, even discussing few (such as anger) at length. His discussion on anger and fear has cognitive, social and behavioural components, along with recognition of physical arousal (Fortenbaugh, 1975). Aristotle placed emotions in a central role for the good life, hence acknowledging its psychological effects as well. On the other hand, in ancient Rome, Stoic philosophers like Seneca studied emotion (conjoined with ethics) and their part in human lives in detail. For them, emotions are erroneous judgements leading to misery – a reflection of Roman societal failure (Nussbaum, 1996). An early version of cognitive theory of emotions, developed by Stoics, has found relevance in terms of modern contemporary psychology.

After the ethical concomitance practiced by the dominant Christian psychology in the Middle Ages (hence the introduction of 'sins' and 'virtues', often interconnected with emotions), it was René Descartes and his ideas of Dualism that made it essential to address emotions as part of the mind-body interaction in an undeniable way. He argued that the mind and body comes in contact in the pineal gland and this interaction brings about the emotional changes and their physical effects on the body parts (Descartes, 1989). For Descartes, emotions are a type of 'passion', formed as 'agitations of the animal spirits' due to the interaction and they are, most importantly, necessary modules of human life.

Although recognised as an important part of life and society, none of the early developers of the concept of emotion regarded it as superior to 'reason', only a 'passive influencer' at best. It was the eighteenth-century thinker David Hume who, in his highly influential book *A treatise of human nature* (2003), challenged this notion and famously proclaimed "reason is, and ought to be, the slave of the passions." Hume, like many of his contemporaries and predecessors, defined an emotion as a certain kind of sensation, or 'impression', which (as in Descartes) is physically stimulated by the movement of the 'animal spirits' in the blood. Though not exactly ground-breaking on the surface, Hume's theory advocated that emotions are a result of a causal network of other 'impression's or ideas and hence, can be identified only when the entire complex of ideas are considered. This paved the way for later theories like cognitive dimension of emotions (check Solomon (1993) for more). In fact, 'experimental philosophy', is the term he used to refer to a field of study that eventually came to know as psychology.

With the advent of psychology as an academic discipline in the second half of the nineteenth century, study of emotion followed three main paths (there are numerous overlapping and variations between these ideas, see Scarantino (2014), Gendron & Barrett (2009) for detailed accounts):

1) Feeling tradition – discusses the way emotions ‘feel’, and defines emotions as distinct conscious experiences.

2) Evaluative tradition – considers emotions as type of cognitions; defines them as distinctive evaluations of the eliciting circumstances.

3) Motivational Tradition – says emotions are distinct motivational states aimed at some goal.

Feeling tradition encompasses the early philosophical ideas that are already mentioned above – that emotions are primitive feelings, resulting from one’s conscious experiences. William James, in 1884, proposed an alternate idea to this. He, although considering emotions as ‘feelings’ much like his predecessors, said that emotions are constituted by perceptions of changes in physiological conditions relating to the autonomic and motor functions: “our feeling of (bodily) changes as they occur IS the emotion” (James, 1984). The main counterargument against this idea was that the physiological changes are too insignificant and, more often than not, too indistinguishable to tell two different emotions apart (Cannon, 1927). In the recent times, James’ theory has been revived and found prominence again due to the ‘psychological constructionist’ movement in emotion study. This approach argues that emotions are mental states that are constructed from the superposition of several psychological components (or building blocks) that are not specific to emotions themselves. Various models have been proposed in partial or complete support of constructionist approach to varying degrees of acceptance (Russell, 2003; Barrett, 2009, 2017; Ghose, 2012).

Evaluative tradition considers emotions as object-directed or endowed with intentionality, or in other words, emotions are targeted towards an object. And such intentionality emerges when emotions are either cognitive evaluations themselves or caused by them (Bedford, 1956; Kenny, 2003). One of the offshoots of this idea is identifying emotions with judgements, since judgements precede the process of directedness – the theory called ‘*judgementalism*’ (Solomon, 1980; Nussbaum, 2001). A parallel tradition to the evaluative approach from Arnold in 1960s (Arnold, 1960) pioneered the now famous ‘Appraisal theory’ which attempts to explain how emotions are elicited. Appraisal theory posits the idea that rather than the stimulus itself, it is the cognitive appraisal of the stimulus (as ‘good’ or ‘bad’) that urges the emotion elicitation. This is a dynamic process that extracts significance from the stimulus and can differentiate between various emotions. Appraisal theory has later blossomed into a huge genre of literature in emotion study thanks to subsequent research works expanding the initial paradigm of Arnold (Lazarus, 1991; Scherer et al., 2001). Altogether along with the evaluative tradition, these cognitivist approaches have argued for the emotions to be essentially a cognitive construction.

Virtually, all of the emotions result in behavioural actions, however minimal. Explaining this relation between emotion and action led to the third aspect of emotion research: the Motivation tradition. Researchers following this tradition view emotion as a mechanism that affects the change in one’s readiness of action (Dewey, 1894, 1895). One variant of this theory, called facial feedback hypothesis, proposes that it is possible to induce emotions with the perception of facial changes, which in turn produces an emotional experience (Tomkins, 2008). Tomkins also proposed that nine emotions are innate and inherited as instincts - interest, enjoyment, surprise, fear, anger, distress, shame, contempt and disgust. They invoke similar patterns of behavioural and bodily changes, due to the sharing a common biological cause. Tomkins's theory inspired the modern day 'basic emotion theory' (comprising the universality of facial expression), later expanded by Ekman (1977, 1980, 1999a, 1999b) and Izard (1971, 1992). Basing on Darwin's works (Darwin & Prodger, 1998) that helped shape the techniques used in cross-cultural studies on facial expression, empirical evidence in favour of the basic emotion theory indicated towards the universality of certain emotional states across cultures.

One noteworthy dispute that sits in the heart of emotion research today is: whether the emotional experience is a subjective or an objective feature. Neuroscientists like LeDoux (2014) and Hofmann (LeDoux & Hofmann, 2018) says that emotions are essentially subjective and verbal reports are the only reliable source of finding that out. The instances where it can't be expressed (infants, animals or brain-damaged adults), the authors argue that they are 'reacting emotionally' but not 'feeling the emotion'. Reacting without feeling amounts to implicitness (non-conscious) regarding the emotional state. LF Barrett (2017) also shared similar arguments, although her perspective drives from a language-based reasoning. On the other hand, a number of noted psychologists and neuroscientists opposed the above view and backed the idea of emotion being an unconscious as well as subjective process (Morris et al., 1998; Damasio, 2004; Wienkielman, 2005; Greenwald & Banaji, 2017). In fact, the evolution of emotional brain predates the human cognitive development – the neural mechanisms of subjectivity have evolved long ago and is hard-wired across species - indicating that human emotions are shared with animals (Frijda, 2007; Panksepp, 2011; Anderson & Adolphs, 2014). A number of evidences suggest that emotions are not exclusively subjective. Emotional reaction can occur without conscious feelings and can leave significantly objective consequences detectable via physiological means (Berridge, 2018). It is, therefore, necessary to distinguish subjective emotional feeling from objective emotional consequence in order to further the understanding of the nature of emotion and its cognitive as well as neural processing.

1.4. EMOTION, PERCEPTION, COGNITION

For a long time, cognitive phenomena like perception, attention, decision-making, memory were kept separated from the domain of emotion. This is due to the long-held belief that 'feelings' and 'reason' were mutually exclusive events. The idea that emotion can influence perception or cognition was not welcomed for a large part of the previous century, even after the advancement of psychology and neurosciences. In recent times, a number of researches have shed light on how emotional processing can affect the functioning of human cognition. Neuropsychological studies showed that many socially relevant cognitive skills suffer due to emotional dysfunction. Similarly, Neuroimaging studies have demonstrated that the brain regions previously regarded as controlling only one of either emotion or cognition interact substantially in complex manner. Psychological and evolutionary studies indicate that preferential processing of environmental stimuli, an evolutionary advantage, is achieved via emotional enhancement of attention or memory (Dolan, 2002). These researches have put emphasis on how emotion and human mental functioning is highly interrelated and, in most cases, arguably inseparable.

Human's ability to categorize the environmental events according to desirability is an indicator of neural evaluation process that comes with the developed nervous system (Scherer & Wallbott, 1994). Emotions represent complex psychological and physiological states that, to a greater or lesser degree, index such occurrences of value. Brosch *et al* has defined emotion as 'an event-focused process consisting of (a) specific elicitation mechanisms based on the relevance of a stimulus that (b) shape an emotional response instantaneously across several organismic subsystems, including motivational changes (changes in action tendency, such as approach versus withdrawal), physiological changes (e.g., heart rate, skin conductance), changes in motor expression (in face, voice, and body), and changes in subjective feeling' (Brosch et al., 2013; Sander, 2013). Hence, the emotional process (and the elicitation of it) depends on the 'memory' and the evaluation of the specific stimulus, linking its experience with the possible response. This take is similar to the 'psychological appraisal theory', discussed in the previous section. Such appraisal of the stimulus, and the response to it, is both subjective (for two different people) and contextual (for the same person) (Siemer et al., 2007).

Emotion and perception and attention:

In our everyday environment, we are constantly bombarded with sensory information in a large amount. Human body, through various senses, send about 11 million bits of data every second to the brain (Fan, 2014). Our conscious brain cannot process all this information and have to prioritise a relatively smaller portion to process using cognitive control - a set of processes that guides adaptive responses and reduces information processing according to the goal and homeostatic demands. Distinct and integrated attentional network systems dedicatedly continue this selection process granting access to conscious awareness (Driver, 2001). Various studies, behavioural and neurobiological, have attempted to understand the processes involved in attention (Ohman et al., 2001; Armony & Dolan, 2001; Mogg et al., 1997; Corbetta & Shulman, 2002). Low-level properties such as the physical intensity of a stimulus may trigger an automatic, reflexive orienting, referred to as *exogenous attention*. In contrast, stimuli that are important to the current behaviour of the organism are selected by a voluntary top-down deployment of *endogenous attention*, driven by implicit or explicit expectations for a specific object or location (Brosch et al., 2013). Corbetta and Shulman (2002) suggest that both endogenous and exogenous attention primarily implicate frontoparietal networks of cortical regions, with endogenous attention control being exerted by interactions of dorsal regions such as the intraparietal sulcus and the frontal eye fields, and the exogenous reorienting of the attentional focus mediated by ventral regions in the right hemisphere such as the right ventral frontal cortex and temporoparietal junction.

Stimulus prioritisation also depends on the emotional relevance of it. Behavioural findings across many different tasks and paradigms indicate that emotional stimuli may draw attention and facilitate perception more rapidly and for a longer duration than neutral stimuli (Phelps et al., 2006; Brosch et al., 2010). For example, in visual search tasks, the detection of a target among distractors is faster when the target is emotional, as opposed to neutral (Ohman et al., 2001). Emotion processing can directly shape the content of our percepts and awareness. These effects operate not only in the visual, but also in the auditory modality (Grandjean et al., 2005) and even across sensory modalities, suggesting that the prioritisation of emotionally relevant stimuli is organised supramodally across multiple sensory channels (Gerdes et al., 2014). Brain imaging studies have also indicated that emotional stimulus carry more prominence in neural signatures than neutral ones. Studies using fMRI have consistently revealed increased neural responses to many different emotional stimuli (Grandjean et al., 2005; Vullieumier et al., 2001). Hence, emotional relevance to external stimulus could be a great contributing factor in preferential access to cognitive processing, behaviour control and awareness.

It has been suggested that the prioritisation of emotional information is executed mainly in the amygdala area which is a limbic region critically involved in the processing of emotional information (Vullieumier, 2005; Dolan, 2002). Once it has determined the relevance of sensory stimuli, amygdala then modulates the processing of these stimuli through direct feedback projections to sensory cortex and biasing signals to fronto-parietal attention regions. Patients with medial temporal lobe sclerosis leading to amygdala damage do not show the cortical increases to emotional information, which supports a causal role of the amygdala in the boosting of the neural representation of emotional information (Vullieumier et al., 2004). Thus, emotion modulates our perception and attention by privileging stimuli that are especially emotionally relevant. This mechanism may help us organise the perception of our environment depending on our current needs, goals and values. Our evolutionary neural mechanisms allow emotionally relevant events to be noticed readily and, once detected, become the focus of attention, evaluation, and action.

Emotion and visual perception:

To illustrate the effect of emotion even in the most mundane of visual perception, let’s look at the study by Reiner et al (2011). Participants influenced by negative/sad mood reported a slant steeper than those with positive/happy mood. This goes to show emotion not only changes one’s interpretations of social events, but can also change the perceptions of the scale of the world. In fact, emotions routinely affect how and what our visual senses decide to see. In this section, we shall gloss over the effects emotion has on our visual perception. For detailed accounts, it is advised to study works such as Zadra & Clore (2011) and Niedenthal & Wood (2019).

As far back as 1943, research efforts based on Gestalt psychology argued that emotions attached during experiencing an object fundamentally altered perceptual representations of it (Schafer & Murphy, 1943). In more recent works, Niedenthal et al. (2000), exploring the exogenous levels of visual perception, argued that the perceiver’s current emotional state affects and his efficiency of processing visual cues involving emotional stimuli. In a similar vein, Siegel et al. (2018) shows that visual percepts are infused with affective feelings, indicating the inseparability of affect from cognition and perception. Using CFS (Continuous Flash Suppression) paradigm, Gayet et al. (2016) provides evidence that fear-induced stimulus potentiates associated visual percepts. Now, not only emotional arousal guides attention to influence perceptual processing, but it also can influence pre-attentive processes. As demonstrated by Phelps et al. (2006), contrast sensitivity (process of discriminating between emotion-induced and neutral stimuli) increases following visuals of fearful faces. Again, while investigating the influence of any top-down motivational process, one important issue is to separate between perceptual and post-perceptual (or judgemental) responses. In such an attempt with ambiguous color stimulus, the results revealed that even low-level processing of basic perceptual features can be open to motivational top-down influences (Voss et al., 2008). In other words, the superordinate states (goals, tasks, actions, motives, emotional states, beliefs, expectations etc.) can have a wide-ranging influence on many aspects of information processing.

These studies indicate that emotions not only influence our experiences, all the way down to perception, but in fact these influences are quite frequent and possibly, multi-modal.

Emotion and auditory perception:

It is well established that sounds evoke emotions and can provide affective information, perhaps more effectively than many other forms of information channels (Juslin & Västfjäll, 2008). Auditory percepts often help elicit emotional reactions in our everyday world - startled by the sudden sound of a door slamming, annoyed by the noise of horns honking in the street, pleased by the sound of a water stream in the forest etc (Tajadura-Jiménez & Västfjäll, 2008). Although humans are exposed to sounds continuously, researches on how humans respond to affective auditory stimulus is still underreported. Frequently, emotional experience has been defined as either discrete feeling states or states that can be placed along dimensions of experience (Levenson, 1999). This approach, following the ‘basic emotion theory’, assumes that there exist universal fundamental emotions, which are spontaneous and uncontrollable, and other emotions are derived from their complex combinations. Other approach characterise emotions in a dimensional model and emotional responses can be placed in affective space (Osgood et al., 1957).

Most researches on affective response to auditory stimuli is centred around its physical determinants (Tajadura-Jiménez & Västfjäll, 2008), i.e., focused on the affective state of *annoyance*, quantified in decibel level, or in loudness, sharpness etc (Guski, 1999). It has also been argued that intensity of sound and arousal might be analogous, since increasing loudness results in an increase in the orienting response. This assumption holds until a level where the intensity of the sound becomes highly aversive (85-90 dB) (Lang et al., 1990).

Physical properties of sound are no doubt an important factor in their affective appearances, but it is the psychological attributes that moderates the subjective interpretations. Many studies attempted to characterize natural acoustic stimuli on the basis of their affective quality (Bjork, 1985; Bradley & Lang, 2000). In these works, Physiological changes showed to be highly correlated with self-reported emotional reactions in terms of valence, arousal and dominance, indicating that acoustic cues activate the appetitive and defensive motivational circuits underlying emotional expression in ways similar to visual stimulus. In human social interaction, vocal cues are as important as visual cues, and human speech contains many features (melody, rate, pauses, intonation, etc.) which inform us about the speaker’s affective state. Hietanen et al. (1998) explored how tendencies to approach or withdraw varied in response to vocal affective expressions. Results suggested that emotional expressions may be contained in vocal cues. Also, music and its ability to both express and produce emotions is an important field of study in auditory perception domain. A number of studies over the years have investigated and provided evidence of music's importance in eliciting affective responses (musical emotions are innate and pre-attentive perceptual processes in most cases) (Juslin & Sloboda, 2001; Ghosh et al., 2018).

Emotion and cross-modal perception:

As we have seen in the previous sections, emotion is significantly interrelated in affecting both the visual and auditory perception and subsequent cognition, separately. Usually, the studies on emotion have traditionally focused on restricted sensory aspects such as visual, auditory, tactile etc (Krishna, 2012). In real world, a wide variety of emotional cues reach different sensory channels simultaneously and interactions between different sensory modalities are usually nonlinear, i.e., the 'whole' is not a simple algebraic sum of its components (Bresciani et al., 2005; Seigneuric et al., 2010; Shimojo & Shams, 2001; Driver & Noesselt, 2008). Although neurobiological studies have shown that emotional signals delivered via different sensory modalities interact at multiple processing levels in the brain, influence each other, and form holistic percepts, involving a variety of brain structures from unisensory cortices to high-level association areas (Klasen et al., 2014), it is still not clear how multisensory input interacts with emotion and affects behaviour. In this section, we review briefly on the present state of research on emotional responses due to multisensory stimuli (interested readers should go through Schreuder et al. (2016) for a detailed review).

In a multisensory environment, each modality affects the other modalities, which can lead to the sensory cues to be multiplied, inhibited or even produce some novel effect as well (Schreuder et al., 2016). When Emotional stimuli across modalities are congruent, then the perceptions are enhanced and positively facilitated (De Gelder et al., 2005), whereas incongruent stimuli increase activation in regions involved in cognitive control and generates surprise (Ludden et al., 2009). The amygdala is activated when the stimuli is arousing, and the activity is significantly attenuated when a neutral stimulus is paired with an emotional one compared to conditions where emotional stimuli were present in both channels (Muller et al., 2012). Also, emotional information from one modality can automatically and unconsciously influence emotion processing in another. This is why sad (happy) faces are perceived sadder (happier) in combination with music that evokes a sad (happy) emotion. Interestingly, fear recognition in the voice is modulated by unconsciously recognized facial expressions but not by unconsciously recognized affective pictures, indicating that processing of facial expressions is more exogenous than pictures (De Gelder, 2002). In addition, this study suggests that negative or fearful stimuli are more likely to bias the perception than positive or happy stimuli (since the former is evolutionarily relevant to immediate survival).

In audio-visual context, it is seen that congruency among the modalities increase positive appraisal and requires less cognitive effort to accept the environment compared to incongruent stimuli (Russell, 2002; Baumgartner et al., 2006). Furthermore, when presented together, it is the visual cue that dominates the appraisal over auditory cues (Kuwano et al., 2001). The reverse is also argued by some (Vines et al., 2006). In our recent study, we showed how structural complexity of the auditory stimulus is consistent with visual and emotional correspondence related to it (Roy et al., 2020). Hence, it can be said that which modality would contribution majorly or influence the other that depends on the task parameters and related contexts. Altogether, knowledge on emotion and multi-sensory stimulus remains considerably limited since studies on this topic are still scarce and haphazard. Any holistic framework seems difficult to achieve given the variations in experimental conditions or methodological diversities.

To summarise, the growing interest in the neurobiology and psychology of emotion parallels a wider recognition of its importance to human experience. Intertwined with cognition, emotion determines our holistic view towards the environment surrounding us and our decisions regarding every aspect of it. Complex human behaviour, also, emerges from dynamic interactions between multiple processes and brain networks. Hence, emotion remains a guiding factor in navigating through the complex sea of perception and information in our everyday world.

1.5. COLOR, MUSIC, *RASA*S

Color and music form the very core of human sensory experiences. Although being two separate sensory perceptions, the awareness of the relationship between color and music has existed since ancient times. Aristotle wrote – "It is possible that colors may stand in relationship to each other in the same manner as concords in music." (Von Goethe, 2006). In eighteenth century, Sir Isaac Newton (1952, originally published in 1704) argued whether the harmony and discord of colors arise from the vibrational proportions propagated through optic nerves into the brain, much like sound arising from the vibrations of the air. This idea was furthered by Jesuits like Athanasius Kircher and Pere Louis Bertrand Castel in the next century - the latter even built a color-clavichord (ocular harpsichord) that linked music with color (Franssen, 1991). It was an attempt, albeit not well-accepted, to make the deaf experience the beauty of music by means of color. In the next century the German literary giant Goethe gave a rich analysis of colors in his seminal book 'Theory of colors'. He differed from the mathematical approach of Newton and his students and talked about the association of colors with moods, attitudes and emotional states. He believed that color and music, though not comparable with one another, stems from the same 'higher law' (Von Goethe, 2006).

The exploration of the interrelation between color and music blossomed in the twentieth century through some of the synaesthetic artists, musicians and architects like Alexander Scriabin (1872-1915) – creator of a color keyboard named *Tastiera per luce* (Ballard & Bengtson, 2017), Wassily Kandinsky (1866-1944) – wrote at length about the emotional qualities of colors and musical instruments that corresponds those emotions (Walz, 2013), Ira Belmont (1877-1964) – started the 'color-music expressionism' in painting, Gyorgy Doczi (1909-1995) – argued about shared harmonies in colors and musical chords (Doczi, 1981), Olivier Messiaen (1908-1992) and Johannes Itten (1888-1967), among others (Beattie, 1998). Their theories and analytical deductions formed the core of later developed concepts of color-music association, consistent with modern psychological ideas. Swiss expressionist painter Johannes Itten (1970) designed and elaborated a basic course on color in his book 'The elements of color' using inspirations from basic practice of music scales. In this book he also gave a chart of seven basic VIBGYOR color hues and their corresponding wavelengths and frequencies. Something similar was done by Gyorgy Doczi. From their works, it was seen that the musical tone C-B# was compared to the color green. Such correspondences of wavelengths and frequencies between other tonal combinations and colors were reported by Itten and they, together, formed a 12-pointed color star (Itten 1970, p. 48). This worked as a guideline of creating colors from musical scores. Later, in the last two decades on the previous century, systematic visual-auditory synaesthesia studies were conducted by Polzella and Biers (1987; Polzella et al., 1982; Polzella & Hassen, 1997). They concluded that "..*the link between chromesthetic responses and music are mediated by attributes common to both visual and auditory experience*."

Modern psychology has proposed two possible explanations regarding such mediating attributes. First one is direct correspondence of the two perceptions (direct link hypothesis). For example, Caivano (1994) says that properties of the musical scale maps properties of color (luminosity = loudness, saturation = timbre and size = duration). In similar manner, Pridmore (1992) argued that due to the similarity in the cyclic nature of both the octave and hue cycles, amplitude connects loudness to brightness/lightness and wavelength connects musical tone to hue. Effect of audio pitch on perceived brightness is also documented in some studies (Collier & Hubbard, 2004; Ward et al., 2006). In contrast to such directness, the alternate explanation involves an indirect relation between color and music via emotional association (emotional mediation hypothesis). In this view, colors are associated with music based on shared emotional content. It is seen in various studies that happy/sad music has a strong correlation with happy/sad color when mediated with an emotional component, even in non-synesthetes (Bresin, 2005; Palmer et al., 2013, 2016; Lindborg & Freiberg, 2015). Between these two possibilities, evidence supporting the latter has been in abundance (Isbilen & Krumhansl, 2016; Whiteford et al., 2018; Roy et al., 2020), that is to say, the color-music association is strongly influenced by emotional or affective mediation (as Whiteford et al., (2018) suggests, this doesn't negate the relevancy of the directness of music's perceptual properties in such association, since they carry information which helps determine the emotional character of the music).

As we have discussed in section 1.3, the Western view of emotion (when perceived as mental states) is dominated by the basic emotion theory which includes: happiness, sadness, fear, anger, surprise, disgust (Ekman, 1992; Izard, 1992). However, there do exist concepts of some background and social emotion sets such as pain-pleasure, envy, empathy, guilt etc. (Tripathi et al., 2018). There are many arguments that have come up against the discreet states approach. Foremost of which is the psychological constructionist movement (see section 1.3) that says emotions are not fixed observable entities or natural kinds, rather products of various psychological processes interacting through neural mechanisms. Other arguments include lack of dynamism in considering emotional expression (which is crucial for certain emotional states), lack of functionalism (Campos et al., 1994), and lack of diversity in emotional taxonomy across cultures (Russell, 1991). Another glaring omission in dominant Western outlook of emotion is its experiential component, loosely termed as 'feeling'. The closest feature to it probably is 'Affect'. This constitutes the subjective experiential aspects of emotions as against the physiological changes and behavior accompanying emotions. The Indian traditional point of view towards emotion is, on the other hand, centred on this very idea and treats ‘Affect’ as the focal point (Ramaprasad, 2013). The major source that guides the Indian approach of emotion classification is the *Natyashastra* by Bharatha, written in the 2nd century AD (Kavi, 1934). This is probably the oldest surviving treatise on performing arts in the world and hosts a detailed compendium on the theory of emotions in a manner quite different than the Western view: it articulates the *Rasa* theory in light of the performer's pragmatic goal of conveying emotional states to the audience. *Natyashastra* discusses the practical means for creating a distinct mood through the performance, which evokes a certain ‘sentiment’ or ‘Rasa’, the aesthetic equivalent of an emotional tone, in the viewer. It is evident that this certainly corresponds to the experiential aspect of emotion in a direct manner, akin to Immanuel Kant's cognitive ‘free play of imagination and understanding’ or David Hume's generic aesthetic ‘sentiment’ (Higgins, 1907). *Rasa* is a conjunction of three components: (1) permanent dominant emotional states or *sthayibhava*: 8 in number (Erotic Love/*Rati*, Mirth/*Hasya*, Sorrow/*Soka*, Anger/*Krodha*, Enthusiasm/*Utsaha*, Fear/*Bhaya*, Disgust/*Jugupsa*, Astonishment/*Vismaya*), (2) transitory emotional states or *vyabhicaribhava*: 33 in number, feeds the dominant emotions but incapable of existing independently and (3) temperamental states or *sattvikabhava*: 8 in number, basically are the involuntary and spontaneous physiological changes while experiencing the emotion. Each of the *sthayibhava* serves as a basic affective tone to produce one sentiment or *Rasa*, each. They are: *Sringara* (The erotic, comes from *Rati*), *Hasya*(The comic, from *Hasya*), *Karuna* (The pathetic, from *Soka*), *Raudra* (The furious, from *Krodha*), *Veera* (The heroic, from *Utsaha*), *Bhayanaka* (The terrible, from *Bhaya*), *Bibhatsa* (The odious, from *Jugupsa*) and *Adbhuta* (The marvellous, from *Vismaya*) (Higgins, 1907). Another Rasa, called *Santa* (The peace) was later added and together, these are known as *Navarasa* or the ‘Nine sentiments’. This Rasa theory of emotions laid down by Bharatha is the strong foundation that provides a perennial vision of the rich and complex world of emotions in Indian aesthetics.

The western taxonomy of emotions is more familiar with the negative emotions depicted in *Natyashastra* than its positive ones. Anger, Fear, Sadness/Sorrow and Disgust are well-documented in Western approach but concepts of Enthusiasm, Erotic Love, Astonishment or Peace are not (Hejmadi et al., 2000). Additionally, *Natyashastra* puts huge emphasis on the gestures and expressions unlike the western theories. It contains rich descriptions of how the emotions are to be expressed. Schechner (2003) considered the expressive treatments of *Natyashastra* to be comparable to the facial expression photographs from Ekman et al (1987). The main advantage that the Rasa theory enjoys over the Western ideas is the incorporation of conscious experience, which promises therapeutic implications.

Although mainly a discourse on dramaturgy, the emotional theory described in *Natyashastra* has expanded into the music and color perceptions as a guiding light, especially in Indian aesthetic perspective. Rasa, perceived as a sentiment, could also be described as an aesthetic experience, focused on two very basic characteristic: firstly, the disinterestedness or the feeling of exclusion, i.e., the experience makes the viewer/listener gradually grow unconscious of his/her existence. And secondly, the conscious realisations of joy of the experience, even when painful events are represented through it. As Wieczorkowska et al. (2010) puts it – ‘the Eastern approach to emotional aesthetics and intelligence treats rasa as a multi-dimensional principle that explains thoroughly the relation between a sentiment, a mood, the creative process and its transpersonal qualities’.Indian musicological treatment, mainly based on affective cues in *Natyashastra*, says that even individual notes bear the potential of producing emotional effects.In north Indian classical vocal music (also known as Hindusthani vocal music) *Raga* forms the musical basis. Ragas are congregations of various notes and musical properties (scale type, transilience, tessitura, characteristic motion etc.), and they elicit some specific emotions described in Rasa theory, although the same Raga could express more than one rasa (Kaufmann, 1965). There is a long tradition in assigning Ragas for evoking emotions in Hindustani and Carnatic systems (for example, Subhapantuvarali is said to evoke penitential emotions). And since Indian music is far less strict on structures, mostly not even having written scores, the characteristic phrases that identify the Raga can be elaborated multiple times, often with different ornamentations. Ragas are also assigned to a particular time of the day and particular season. Such associative elements, along with the complex relation between the composition, tempo, rhythm, and improvisational ornamentation correspond to a richly textured music with one (or more) deep emotional experiences. That is why various performers and scholars have agreed that the essence of Raga, that is the emotion it conveys, is not entirely reducible into these structural features. It is rather knowable only through the immersive experience of listening/performing. Clayton (2001) has categorised the following features of Rasa theory in regard to musical meaning:

- It highlights the role of the perceiving subject and the importance of empathy.
- As a successor to the above point, this theory asserts on experience and affect; a piece of art is evaluated by its ability to invoke emotion among the perceivers.
- Rasa, literally standing for juice or essence, expresses the affective experience as a fluid and dynamic entity unlike Western static approach.
- Dissociates affect with the stimulus structure (structure is important, but not absolute).
- And of course, the emphasis on a holistic ‘tasting’, resisting reductionism.

He concludes, saying, that perhaps it is the nature of unwrittenness and abstraction of the pure auditory experience that “theorists allowed the affective power of music to remain within the irreducible domain of sound – something from which Western musicology could certainly learn.”

Color perception and the emotional responses attached to it have strong psychological and physiological attributes - something which was incorporated in the Rasa theory efficiently, that too even long before the advent of any of the above as a serious scientific discipline. Generally, the emotional experience of color is elicited through association. The superficial association of a certain color with a certain deeper lying memory explains the affective responses (Cheriyan, 2017). The *Natyashastra* contains extremely detailed descriptions of garments and ornaments and their color combinations. According to Bharatha, *bhavanukarana* (or the ‘imitation of the emotional states’) becomes clear by the impact of colors. He also assigned a color corresponding to each Rasa, as given in Table 1.5.1.

Rasa	Color
Shringara (Erotic)	Green
Hasya (Comic)	White
Karuna (Pathetic)	Dove-colored Grey
Raudra (Furious)	Red
Veera (Heroic)	Wheat Brown/ Pale Orange
Bhayanaka (Terrible)	Black
Bibhatsa (Odious)	Blue
Adbhuta (Wondrous)	Yellow

Table 1.5.1. Rasas and their corresponding colors, described in Chapter 21 of *Natyashastra* **(**Kavi, 1934)

The *Natyashastra* explains each color and the rationale behind choosing it to associate with said emotional state. For a brief review one should refer to Cheriyan (2017). Later studies on various aspects and properties of color and related psychological attributes have corroborated some of these ancient ideas. For example, studies on color and their association with mood-tones show that red is associated with 'exciting' and 'stimulating', both of which imply pleasure and high arousal. For Blue it is 'secure/comfortable' and 'tender/soothing', which imply pleasure and low arousal. On the other hand, Orange, associated with 'disturbing/distressed/upset', implies displeasure and high arousal and Black was associated with 'powerful/strong/masterful', implying high dominance (Wexner, 1954). Lakoff & Johnson (1999) proposed the conceptual metaphor theory of color where he argues that we think about abstract concepts like color in terms of metaphors to help understand and connect them with perceptual experiences. Hence, anger, which entails reddening of the face, is metaphorically described as 'seeing red'. Such metaphors drive moral judgements and stereotyping as well (White = pure and Black = evil). In his book (1991), Davidoff not only acknowledges the effect of warm or cool colors on physiological or emotional states, but believes that they affect the endocrine system directly via the pituitary gland. In another work with short wavelength lights, blue light in particular, is reported to activate the melanopsin photoreceptor system which, in turn, activates the brain structures involved in sub-cortical arousal and higher-order attentional processing (Cajochen et al., 2005). Also, psychological studies on combat fliers indicate that bodily reactions to anxiety and fear (pounding of the heart, tenseness of muscles, dryness of mouth, sickness in the stomach etc.) are quite similar to the ones described by Bharatha (Shaffer, 1947; Rachman, 1983). The treatise in *Natyashastra* is not only well founded in affective responses to colors, but it also described the primary colors as red , blue and yellow and the rest of the colors as derived compounds from the primary ones. According to Bharatha, appropriately painted body along with appropriately colored costumes are essentials in propagating the affective experience.

To summarise, investigating color and music perceptions and their association posit several challenges in both the fields of psychology and neurobiology. So does the abstractness of emotional experience. Despite that, one thing could be said with certainty - that to find connections between different perceptual events, it is emotions that are forming the cognitive bridges. Emotions are so much innate to our lives that they are central not only to how we interpret incoming information, but also to how we respond to them. Although the question remains on the best possible approach to explore emotions – Western stoicism vs. Eastern dynamism – but given the myriad connections from perceptions to emotions and from emotions to actions, it seems quite natural that emotions emerge so strongly, and perhaps, unconsciously.

1.6. A SUMMARY ON STUDIES REGARDING COLOR PERCEPTION

Humans encounter the world primarily through color. Color is perceived on essentially every object that we view in daily life. A person with normal color vision experiences a vast and rich chromatic palette, i.e., up to 2.3 million discernible colors (Linhares et al., 2008). Considerations about color or color combinations emerge regularly in our everyday decision making and conversations, from the color of clothes to wear, or the color for a new car, to passing judgements on the color of someone’s skin, hair, or makeup. Popular opinions abound on the nature of color associations and on presumed influences of color on our feelings, aesthetic judgments, and beyond.

Empirical research on color perception has a long history (Elliot & Maier, 2014; Elliot, 2019). Goethe, back in 1810, gave observation-based accounts of color perception and its association with emotional states (or ‘feelings’) in his book *Theory of colors* (Von Goethe, 2006). Some colors were assigned to induce positive feelings and some induced negative ones. According to him, Red is associated with both gravity and lightness, Blue with coldness and negativity, Green with calmness, Yellow with warmth and Orange with energy and power. Later in the same century, psychologist and writers like Wilhelm Wundt and Grant Allen addressed this issue in a similar manner and associated color with emotions based on experience and observation. They shared similar ideas such as: White is linked to positive emotions (pleasant, stimulating, cheerful), Black to negative emotions (unpleasant, gravity), Yellow and Red are exciting and arousing, Blue is depressing/relaxing, and Green is linked to tranquillity and calmness (Wundt, 1874; Allen, 1877). Subsequent quantitative researches from Charles Féré (1887) and Hugo Münsterberg et al. (1894), involving color stimulus and physiological signatures – pulse rate, respiratory pattern, head and eye movement – demonstrated that Red (and Yellow, in the latter) had a comparatively higher arousal and attention capturing effect whereas Blue (and Green, to some degrees) is calming/relaxing. Also, in some color preference studies, Blue was reported to be the most preferred color, followed by Red (Cohn, 1894; Jastrow, 1897).

These early empirical works, for obvious reasons, show various methodological weaknesses, such as: not maintaining participant naiveté and confidentiality (they, often, knew hypotheses beforehand and were individually identifiable), participants' color vision deficiency was not assessed, sample sizes were small, color stimuli and background color (also ambient illumination) were not precisely specified etc.

In the next century, color perception studies blossomed thanks to the advancement of science and gradual understanding of the underlying mechanism of light and vision. With that, the focus shifted towards the change in behavioral or psychological manifestation with changing color wavelengths (Nakshian, 1964). But even after studies spanning a good part of the century, the results are not always conclusive. For example, a study by Hill and Barton (2005) suggests that the color Red can be associated with dominance and aggression in both human and non-humans. Another one associates similar responses in human with Black, in competitive sports' perspective (Frank & Gilovich, 1988). Numerous studies on color and its association with different psychological attributes show that the longer wavelength colors like Red have an enhanced effect on such attributes. For example: viewing Red and feeling excited or stimulated (Goldstein, 1942; Ou et al., 2004; Buechner & Maier, 2016), wearing Red and exhibiting increased strength during combat sports (Hill & Burton, 2005; Dreiskaemper et al., 2013), enhanced attraction towards opposite sex when they wear Red (Elliot & Niesta, 2008) etc. On the other hand, shorter wavelength color such as Blue is reported to have increased effect on alertness (Vandewalle et al., 2007). Blue has also been found to elevate perception of trustworthiness in marketing research (Lee & Rao, 2010). When compared during cognitive task performances, it was seen that Red encourages avoidance motivation and analytical execution. Contrarily, Blue exhibit approach motivation and encourages creative performance (Mehta & Zhu, 2009). Majority of studies compare effects of colors Red and Blue because their wavelength differences put them on the opposite end of the spectrum. Few studies have ventured into other colors and linked Green with calmness (Suk & Irtel, 2010) and Orange/Yellow to excitement (AL-Ayash et al., 2016) but other studies disagreed (Ainsworth et al., 1993; Briki & Hue, 2016; Wilms & Oberfeld, 2018). The general trend in the literature available shows that the association between Red and excitement is the most reported scenario. The association of calmness/relaxation has been divided majorly between colors Blue and Green.

Another avenue of color perception research is studying the physiological response due to color stimulus exposure. Studies done in this area are mostly motivated by the hypothesis that long-wavelength colors (red/yellow) are more arousing than short-wavelength colors (blue/green). Although the fundamental question that remains unanswered in this field is whether the response is direct (i.e., stimulus evokes the response directly without cognitive intermediation) or indirect (cognition acts as an intermediary) (Kaiser, 1984). To measure physiological signals, varieties of parameters have been used – Galvanic Skin Resistance, Heart rate, respiration, blood-pressure etc. EEG remains the mostly used form of technique among them. Gerard (1958) used EEG to report high cortical arousal (the lower prominence of alpha waves) under the effect of Red color. Since then, brain activity during color exposure is studied by measuring the alpha waves. These EEG studies have given contradictory reports too: some argue that Red is associated with higher physiological arousal (Deutsch, 1937; Wilson, 1966; Ali, 1972), some disagree or remain unconvinced (Brown, 1966; Choi et al., 2011; Sakuragi & Sugiyama, 2011). During cognitive tasks, Blue light is reported to induce higher arousal and brain activity as well (Klimesch, 1999). Similar effects were attributed to Orange, Yellow and Violet (when compared to Green) by some (Erwin et al., 1961; Nourse & Welch, 1971; AL-Ayash et al., 2016).

The psychological investigations of color perception have relatively higher agreement in terms of generality than its physiological counterpart where the general outcome remains largely inconclusive. Also, majority of these studies explore the applicative side of the problem, i.e., establishing relationship between one particular color and one behavioural trait. The underlying neural mechanism is not addressed or explained. Evidently, the numbers of works on color induced EEG are in short supply; detailed studies on the physiological manifestations (especially that on the brain) is required to further our understanding about the problem. Taking this into account, an attempt was made in this thesis to investigate the brain response (via EEG) to color stimuli using novel nonlinear analysis techniques which are far more robust and efficient in detailing the underlying dynamics than any of the linear methods previously used. The obtained results not only report some unprecedented trends regarding color perception, but also analyses the existing ones, to a good degree of success, in the light of state-of-the-art nonlinear chaos-based tools and parameters.

1.7. CROSS MODAL CORRESPONDENCE: AN OVERVIEW

Human senses are separate entities designed to perform separate specific duties. By definition, they are not supposed to poke their noses (pun intended) into each other’s responsibilities. For example, eyes don’t go tasting food or skin doesn’t try to visualise objects. Or so it was thought. Counter-intuitively, it is not the distinctness of the senses that shapes the human perceptual world, but it is their power of combining. More often than not one sensual modality crosses path with the other. Depending on whether such ‘superposition’s are constructive or destructive, human experiences about the external environment either thrive or thwart. At the outset, the cognitive neuroscience research on multisensory integration have examined the temporal and spatial aspects of multimodal perception, i.e., how closer the different stimuli (presented to different modalities) are on a time scale (Jones et al., 2006; Shore et al., 2006), or spatially in some cases (Frens et al., 1995; Slutsky & Recanzone, 2001). Based on the nature of the stimulus attributes, multisensory perception can be either semantically congruent (Chen & Spence, 2010; Grassi & Casco, 2010; Doehrmann & Naumer, 2008) - meaning the paired-up stimuli have different identity and meant to stimulate two different senses – or they can be synaesthetically congruent (Gallace & Spence, 2006; Evans & Treisman, 2010) – correspondences between basic stimulus features shared in different modalities – or, a bound effect of both of these together, known as the “unity effect” (Vatakis et al., 2008). There are a number of different ways in which sensory information can be associated. Due to the basic amodal redundant features of the stimulus, or seemingly unrelated modal features, or abstract features like pleasantness, user arousal potential or potential to change affective states like emotion etc. - reasonable evidences have supported the claim that crossmodal correspondences can occur from such low-level amodal properties to high level cognitive properties. For a detailed review on this topic, see Spence (2011).

The initial researches on multimodal association involved sound symbolism – associating a sound (usually nonsense) with a shape. It was seen that subjects associate a specific phonetic with a specific shape (for example, in Sapir (1929), "mal" and "mil" were linked to large and small rounded objects, respectively) (Köhler, 1929). In recent years, interest in this research has rekindled, due to the advances in neuropsychological studies and related investigations (Ramachandran & Hubbard, 2001, 2003; Ramachandran & Oberman, 2006; Nuckolls, 2003). The *bouba/kiki effect*, as it is called, shows that 'bouba' is associated with rounded edged star whereas kiki with sharp edged star. Curiously enough, this effect doesn't hold in children with autism spectrum disorder, which is suggestive of high-level neural mechanisms behind such phonetic symbolism. Besides this stream of works, there are a lot of studies in semantic congruency which reported many non-arbitrary crossmodal correspondences between a variety of auditory and visual stimulus features. For instance, between brightness and loudness (Bond & Stevens, 1969), high-pitched tones with bright surfaces (Marks, 1974), loud sounds with high contrasting visuals (Wickers, 1968) and association of high pitch with high elevations (Roffler & Butler, 1968). Developmental researches showed that the perception and association between multiple senses is exhibited in children under 2 years of age: loudness and large objects (Smith & Sera, 1992), auditory pitch and visual elevation (Walker et al., 2010). Sensory integration is documented in other modalities as well, such as – vision and touch (Martino & Marks, 2000), flavours and auditory stimulus (Crisinel & Spence, 2009), colors and odour (Kemp & Gilbert, 1997), smells and shapes (Seo et al., 2010), and even shapes and flavours (Gal et al., 2007). From these studies, it appears likely that crossmodal correspondences exist between all possible pairings of sensory modalities.

The obvious question that follows is – whether the effect of simultaneous sensory stimulus can help or hinder sensory information integration in one certain modality. In their seminal paper, Bernstein and Edelstein (1971) showed that subjects respond more slowly to visual stimuli when their elevation happens to be inconsistent with the relative pitch of a task-irrelevant sound. The association between auditory pitch and visual elevation has turned out to be one of the most reported and replicated event in later researches. Since then, such speeded classification studies have been conducted involving various modal features, and it was seen that people find it harder (slower and less accurate) to classify the target stimuli presented in one sensory modality (e.g., vision) when the distractor stimuli presented in a task-irrelevant modality (e.g., audition) happen to vary incongruously with the target dimension (Marks, 2004). Later, Gallace & Spence (2006) reported similar findings in a visual size discrimination task. Additionally, their results indicated that such effects were more relative than absolute, dependent on experimental conditions and stimulus dimensions under consideration (Spence, 2011). Evans & Treisman (2010) suggested that bidirectional correspondence exist between auditory pitch and visual elevation, size and spatial frequency. Although their methodological treatment shows that these effects were fundamentally perceptual and not some form of response compatibility effect, despite that, most of the speeded classification studies fall under the disputed umbrella on whether the observed effects are decisional/response related or actually perceptive (Marks, 2004).

Among various models of cross-modal associations, one of the recent and most interesting involves the Bayesian integration theory (Ernst, 2006). This theory suggests that humans combine the sensory information in a statistically optimal manner by combining prior knowledge regarding them and weighting each of them by their relative reliabilities. Hence, the strength of crossmodal combination of two stimuli depends on the prior knowledge of our sensory system that those two can couple effectively: this knowledge is expressed using 'coupling prior', a representation of the expected joint distributions of the signals. According to such a Bayesian model, 'the reliability of a person's estimate regarding intersensory conflict is proportional to the strength of the coupling between the signals being integrated' (Ernst et al., 2007). Meaning, stronger the coupling between the unimodal stimuli, better the chances that they would be blended together to form one multimodal percept. This effect can be interpreted in terms of differences in the strength (i.e., variance) of the coupling prior: for stimuli which are congruent in cross-modal aspect, the variance is small; and it is opposite in case of incongruent ones. Studies have reported that the strength of the crossmodal association can be modulated by meddling with or influencing the variance of the coupling prior (the knowledge about the modal stimuli) (Ernst, 2007; Helbig & Ernst, 2007). Also, a number of previous studies on multimodal correspondence can be explained using the Bayesian approach (Spence, 2011). That being said, it is possible that some of the reports result from the abnormal traits of the neural mechanisms responsible in integrating modal information.

When discussing about the neural mechanisms behind cross-modal correspondence, the first issue that is needed to be addressed is the representation of such effects in the neuroanatomical structure of the brain. Areas in the temporal cortex like superior temporal sulcus are shown to respond to cross-modal stimuli (Beauchamp et al., 2004a, 2004b). Also, damaged angular gyrus can interfere with a person's ability to match stimuli cross-modally (Ramachandran & Hubbard, 2003). Moving to the next concern, which is the change on a neurobiological level during new cross-modal associations, enhancement and suppression of activity in the task-relevant and task-irrelevant sensory systems (thalamic structures), respectively, were found when concurrently presented auditory and visual stimuli are paired by chance. But when information in the two modalities is reliably associated, activity is enhanced in both systems regardless of which modality is task-relevant (Baier et al., 2006). This indicates that when multisensory information is expected to integrate, the neural mechanism facilitates the process. And conversely, it suppresses the process when information is judged as a distraction. Another study also reported such implicit associations in Primary visual and auditory cortex using PET scan (Zangenehpour & Zatorre, 2010), something that was achieved within 45 minutes of exposure to congruent stimuli and lasted for more than 24 hours. Reports from these neuroimaging studies suggest that cross-modal connectivity is evident in the neurobiological level, which facilitate multimodal correspondence even in the presence of a unimodal component of a multisensory pair.

Another growing body of work had emerged in recent times where they dealt with cross-modal correspondence regarding complex auditory and visual stimuli, that is, music and color or painting (Barbiere et al., 2007; Palmer et al., 2013, 2016; Whiteford et al., 2018; Albertazzi et al., 2015). The main difference this stream of literature has from the main body of cross-modal studies is that their advocacy of the emotional mediation, typically illustrated by the participant's response to the stimulus on a number of semantic differential scales (Parrott, 1982). Spence (2019) has detailed a chronological review of this field. This review centres on a large number of studies, most of which used short classical music pieces (except Lindborg & Freiberg (2015) who used film music) and color patches and asked participants to associate the music with their choice of colors. The major findings in those works indicate that Grey is usually associated with sad music whereas lighter and saturated colors are associated with happier music. Also, Palmer et al. (2013) linked faster and slower music with saturated (and lighter) and unsaturated (and darker) colors. Isbilen & Krumhansl (2016) found color correspondence with various musical attributes like tempo, mode, pitch etc. Highly significant correlations were reported between colour–music matches and the similarity of the emotional content by Whiteford et al. (2018). Although the cross-modal correspondence results in these studies are interesting, but Spence questions whether emotional mediation is based on the emotion experienced in response to the music, or rather, on a more cognitive assessment of the emotion. With the definition regarding emotion not clearly addressed, it is often difficult to distinguish between perceptual and induced emotion. Hence, doubts remain on emotion and its affective correspondence during musical intervention (Marin & Bhattacharya, 2010).

Rapid growth in the multisensory research in the last few decades have demonstrated that cross-modal correspondence, in both perceptual and decisional levels, has influenced people's performance in a wide range of different paradigms, including direct cross-modal matching, speeded classification tasks, speeded simple detection tasks, perceptual discrimination tasks and the likes. Various classes of structural, statistical and semantic correspondences highlight their importance in human information processing and its perceptual view of the world around.

In light of this, we have attempted to address a relatively newer approach in the study of multimodality (Music and color associations, specifically) involving emotion as a mediator in this thesis. Our argument follows from the fact that both color and music have similar emotional qualities that inspire arousal in a similar manner, even from a young age (Cutietta & Haggerty, 1987). We hypothesise that the structural complexities of music and the basic components of compound color share a possible correspondence, and it is facilitated by the emotional cues shared by both of these stimuli. Applying latest fractal-based methodologies, well documented in scrutinizing the complex nonlinear nature of non-stationary signals, we seek to obtain a quantitative correlation between music, color and perceived emotion.

1.8. STUDYING COMPLEX SYSTEMS: CHAOS THEORY AND BUTTERFLY EFFECT

Clouds are not spheres, mountains are not cones … Nature exhibits not simply a higher degree but an altogether different level of complexity.

Benoît Mandelbröt

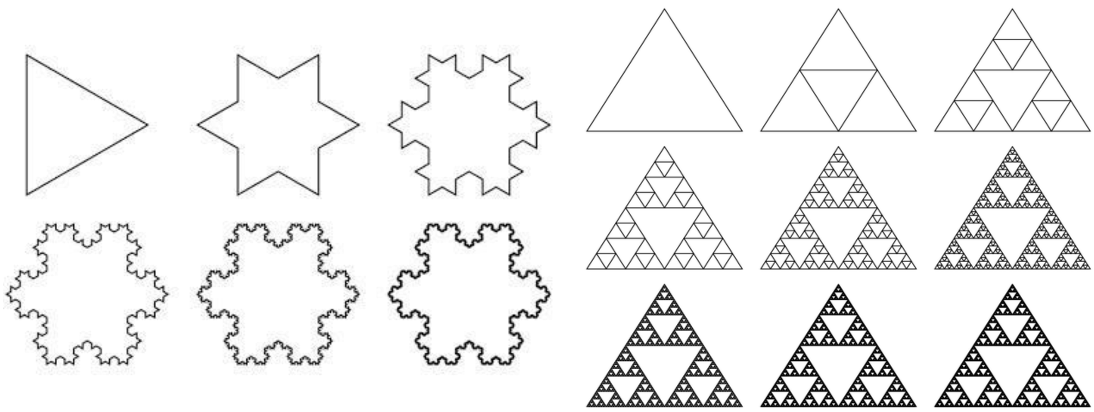
Chaos theory is a delicious mix of science, contradiction and unpredictability. And if one intends to extract hidden ordered patterns among nature's apparent disorganization and randomness, this scientific tool will certainly be his most trustworthy aid. Chaos theory was synthesized theoretically by Prigogine & Stengers (1984) and popularized by James Gleick & Michael Berry (1987). Unlike The traditional scientific approach which deals with supposedly deterministic physical phenomena like gravity, electricity, or chemical reactions, Chaos, a simplification of the theoretical construct, can be defined as an event, behavior, or process which is variable, nonlinear, and unpredictable, like turbulence, weather, stock market, human brain and cognitive states etc. What was, through observation and experimentation, considered random and unpredictable and, therefore, categorized as error or divergence, Chaos theory helps those to be understood as representative of patterned behavior. These patterns as well as the boundaries are flexible and non-deterministic, changing unpredictably (Pool, 1989). Such phenomena are often described by *fractal* mathematics, which captures the infinite complexity of nature. Almost all natural objects surrounding us exhibit fractal properties and complex chaotic behavior, including landscapes, clouds, trees, organs, rivers etc. To gain an insight into the inner workings of both the physical and biological world, one needs to recognize their chaotic nature. The importance of chaos theory lies in its explanatory power to understand the behavior of diverse complex systems.

In chaotic systems, an attractor is not a point or a simple, smooth, continuous curve, like in case of Moon's orbit around the earth. The attractor — often called a *strange attractor* or sometimes a *fractal* — could be an infinite set of unconnected points (e.g., a Cantor dust), or a smooth curve with mathematical discontinuities, or a curve that is fully connected but discontinuous everywhere. The term fractal in this sense refers to a space of fractional dimensions: not 1-dimensional or 2- dimensional or n-dimensional, but 1.23, or 2.78 or π, or any positive non-integer number. The attractor cannot fit in X dimensions, but also cannot fill up X+1 dimensions. Any system that tries to conform to such a curve will behave chaotically, for the same reason that trying to roll a tennis ball on a pebbly beach ends up with the ball bouncing and jumping all over the place; the pebbly surface is something more than 2 dimensional and something less than 3.

Another one of the hallmarks of a chaotic dynamical system is sensitive dependence on initial conditions, more commonly known as the 'Butterfly Effect' - the idea that a small change in initial conditions can lead to a large change in the behavior of a system. *"Does the flap of a butterfly's wings in Brazil set off a tornado in Texas?"* Edward Lorenz, an MIT meteorologist, accidentally discovered (and coined the term) Butterfly effect while trying to model the weather in 1961 (Lorenz, 1963). Lorenz truncated his data in one run-through, entering 0.506 instead of 0.506127. That tiny alteration in a single variable among more than a dozen representing atmospheric conditions in his model dramatically changed the long-term forecast he was working on. Hence, introduction of the Butterfly effect: the flap of a butterfly's wings, a miniscule perturbation in the atmosphere, might be the small change necessary at the beginning of a very long and convoluted chain of events that leads to the formation of a tornado somewhere else.

1.9. STUDYING COMPLEX SYSTEMS: FRACTALS AND MULTIFRACTALS

"Fractal", a term coined by Mandelbröt himself (1967), is a rough or fragmented geometrical object that can be subdivided in parts, each of which is (at least approximately) a reduced-size copy of the whole. This is the most important feature of fractals – they are self-similar across different scales, i.e., the smallest of structures look exactly the same as the whole, when zoomed in. Fractals are created by repeating a simple process over and over in anon-going feedback loop. Some of the well-known fractal systems including Koch snowflake and Sierpinski triangle (**Fig. 1.9.1**) demonstrate this process. Driven by recursion, fractals are images of dynamic systems – the pictures of Chaos. Geometrically, they exist somewhere in between our familiar concept of dimensions and these fractional dimensions (Fractal Dimension or FD) represent the degree of brokenness or irregularity of the object.



a. The Koch Snowflake **b.** The Sierpinski triangle

Fig. 1.9.1 (a, b). The Koch Snowflake and the Sierpinski triangle: Examples of fractal systems

The measured metric properties of the fractal systems, such as length or area, are a function of the scale of measurement. For example, while discussing 'How long is the coastline of Britain?', Mandelbrot answers that it depends on the length of your measuring-stick (**Fig. 1.9.2**). When it is measured at a given spatial scale d, the total length of a crooked coastline L(d) is estimated as a set of N straight line segments of length d. Since the small intricate details of the coastline are not measured in lower resolution, the length L(d) of the coastline keeps on increasing with the increase of measurement scale of 'd'.

The concept of fractal dimension works similarly: the smaller the scale of measurement the more intricate geometry of the signals come into focus. In this regard, the fractal techniques act as a mathematical microscope magnifying fine and complex details of the object with efficiency. On the other hand, a multifractal is a set of intertwined fractals, each having their own separate FDs. Self-similarity of multifractals is also scale dependent (resulting in a spectrum of dimensions).

Fig. 1.9.2. The coastline of Britain in different scales: larger (left) and smaller (right)

Chaos has been studied and discovered in a wide range of natural phenomena such as the weather, population cycles of animals, the structure of coastlines and trees and leaves, bubble-fields and the dripping of water, biological systems such as rates of heartbeat, acoustical systems such as that of woodwind multiphonics and interestingly, music in general (Salter, 2009). Chaos is approached and modelled through the use of nonlinear dynamic systems, which are the mathematical equations whose evolution is unpredictable and whose behavior can show both orderly and/or chaotic conditions depending on the values of initial parameters. What have attracted the non-science community to these dynamic systems are their fractal properties and, thus, patterns of self-similarity on many levels. Discoveries of math and science usually have an effect on art and music. As Pressing (1988) says: although "we cannot say that the music of J.S. Bach is great because it is the aural equivalent of Cartesian geometry... we can hardly deny that it arises from the same *Zeitgeist* or whatever one chooses to call the nexus of intellectual, cultural and aesthetic currents that influence an artist". So, "new music models will undoubtedly arise from the intellectual milieu that includes fractal geometry and chaotic nonlinear systems". One can see the influences in Steve Reich’s compositions — small changes in rhythmic structure (e.g., slight phase shift) lead to big changes in the music (e.g., Clapping Music, Violin Phase, Six Marimbas, etc.) (Lochhead, 2001). Hungarian-Austrian composer GyörgyLigeti, influenced by Chaos theory, composed a set of 18 technically challenging piano pieces titled *Études* where a small musical idea is continuously repeated, transformed, deformed, and expanded (Ligeti, 2003). Etude No. 1, “Désordre”, begins with an orderly and simultaneous melody in both hands but reduces by one-eighth note in the right hand every four measures. As a result, the sequence of accents in the left hand falls behind the right. The right hand plays solely the white keys (C major) while the left hand only the black keys (a pentatonic scale) (Lee, 2015). So, musicians, too, are starting to realize the potential of nonlinear dynamic systems, and its fractal properties are being utilized in a manner that allows musical patterns with chaotic/unpredictable variation to be automatically composed with the computer. Consequently, nonlinear dynamical modelling for source indicates the relevance of non-deterministic/chaotic approaches in understanding the speech/music signals (Kumar & Mullick, 1996; Bigerelle & Iost, 2000; Hsü & Hsü, 1990; Sengupta et al., 2010). Fractal analysis of the signal, in this context, assumes significance since it reveals the geometry embedded. The first step towards such an idea was taken by Voss & Clarke (1975), who analyzed amplitude spectra of audio signals to find out a characteristic frequency *fc*, which separates white noise (which is a measure of flatness of the power spectrum) at frequencies much lower than *fc* from very correlated behavior ($\sim 1/f^2$) at frequencies much higher than *fc*. However, it is well-established that naturally evolving geometries and phenomena are seldom characterized by one single scaling ratio; more often than not different parts of a system exhibit different scaling behaviour. That is, the clustering pattern is not uniform over the whole system. Such systems are characterized as 'multifractal' (Lopes & Betrouni, 2009). Multifractals are complex sets constructed from sub-sets with different local fractal dimensions. Most of the experiential reality, including music, has multifractal features (Su & Wu, 2006).

The human brain, one of the most complex systems in the Universe, is also organized by chaos. Hence, fractal dimension (FD) is a key characteristic of the brain dynamics which indicates the level of complexity on which the neuronal regions function or interact and quantifies the associated brain processes on a scale ranging from fully deterministic to fully random. Kiselev Studied the organizational geometry of human cortical grey matter and showed that the self-similarity is displayed down to the spatial scale of 2.5 mm (Kiselev et al., 2003). Also, the folding of the brain shows fractality for the largest spatial scales with fractal dimension ranging between 2 and 3. As a whole, it can be said that fractality exists in both area and volume of the brain. Hence, human brain functional networks demonstrate a fractal small-world architecture which supports critical dynamics and task-related spatial reconfiguration while preserving global topological parameters. It involves billions of interacting physiological and chemical processes that give rise to experimentally observed neuro-electrical activity, which is exhibited via an electroencephalogram (EEG). The scalp EEG arises from the interactions, generally nonlinear, of a large number of neurons. Thus, linear decomposition is ill-fitted to describe these fluctuations. On the other hand, classical nonlinear dynamics methods such as correlation dimension and Lyapunov exponents are very sensitive to noise and require the stationary condition, while EEG signals often are highly non-stationary. But, the use of nonlinear parameters to determine the music induced emotional states from EEG data is very scarce in literature. Natarajan et.al (2004) evaluated linear and nonlinear parameters like Correlation Dimension (CD), Largest Lyapunov Exponent (LLE), Hurst Exponent (H) and Approximate Entropy (ApEn) from the EEG signals under different mental states. The results obtained show that EEG complexity is significantly lower when the subjects are under music or reflexologic stimulation as compared to the normal state. The measures increase with the degree of the cognitive activity. This suggests that when the subjects are under sound or reflexologic stimuli, the number of active parallel functional processes in the brain decrease resulting in a more relaxed state. One of the earliest studies that used the advanced scaling techniques like Detrended Fluctuation Analysis (DFA) on EEG signals was from Gao et.al (2007) who obtained two scaling exponents' β_1 and β_2 corresponding to high and low alpha band respectively which had direct correlations with the varying emotional intensity of the musical stimulus. Albeit structurally, but several other studies have also corroborated the presence of scale-invariance and self-organised criticality in temporal dynamics of brain mechanisms (Linkenkaer-Hansen et al., 2001; Hwa & Ferree, 2002; Lee et al., 2002). Also, in order to understand many neurological and cognitive/behavioral diseases and disorders (Alzheimer's disease, autism spectrum disorder, Major depressive disorder), this approach of analysis has been applied to a good success (Ahmadlou et al., 2010, 2012; Ahmadlou & Adeli, 2011). Studies on various cognitive tasks and functions like visual processing (Tong et al., 2005), creative thinking (Mölle et al., 1996) or sleep state (Weiss et al., 2011) have explored the possibility of fractal structures in brain processes. In recent years, fractal and multifractal technique has been applied in a number of studies to assess change of brain state when subjected to audio stimuli in the form of tanpura drone (Maity et.al, 2015), studying neural hysteresis effects (Banerjee et.al, 2016) etc. Although a convincing demonstration of chaos has only been obtained at the level of neurons, acting as coupled oscillators, some scientists still believe that there could be considerable benefits for the brain to operate in chaotic regimes due to rich range of behaviors. Regardless of the presence of chaos in brain activity, it is obvious that the neuroscience could benefit from methods developed for the analysis of nonlinear and chaotic behavior.

1.10. STUDYING COMPLEX SYSTEMS: POWER LAWS

Theory of Complex systems represents an umbrella term for various fields of research that include dynamical systems, discrete dynamical systems and cellular automata, game theory, information theory, networks, computational complexity, numerical methods and many more (Newman, 2011). Organizational clusters of complex systems hold within themselves dozens of thematic frameworks, since each of these covers several of such themes. They include topics that are not readily identified with complex systems and not always studied within mainstream scientific disciplines. Some of these topics are: structural scaling of Johann Sebastian Bach's Cello Suite No. 3 (Brothers, 2007), fractal properties of player's positional dependence in a soccer game (Kim, 2006), measuring temperature of texts (Miyazima & Yamamoto, 2008), complexity perturbations in tanpura signals (Sengupta et al., 2005), risk evaluation of diabetes mellitus by relation of chaotic globals to HRV (De Souza et al., 2015) among others. The diversity of the topics illustrates the impact as well as the width complexity garnered in recent scientific research.

Among the fundamental tools in the theory of complex systems, some of the most important have been the physical ideas of scaling, phase transitions, and criticality. One startling phenomenon observed in a number of complex systems is the appearance of 'power-law' distributions of measured quantities (Newman, 2011). Power-law distributions are said to 'scale' or exhibit 'scaling' behaviour, because they retain their shape even when the measured quantity is 'rescaled', i.e., multiplied by a constant. The observation and origin of power laws and scaling in complex systems has been a subject of discussion and research for many decades (Mitzenmacher, 2004), going back as far as the 1890s. Mechanisms of such power law behaviours are studied with interest. The ubiquitous nature of power laws in the natural world has led to claims that advocate for a single underlying mathematical mechanism responsible for all power laws, essentially indicating towards grand unification of the complex systems theory. 'Self-organised criticality' is one such candidate (Bak et al., 1987).

In this context, the basics of scaling and power law distributions mentioned above, needs to be addressed. Most of the quantities measured in scientific investigations have a 'scale', a typical value around which individual measurements are centred (Bell curve). For example: distribution of heights (in centimetres) of American males or distribution of speeds (in miles per hour) of cars on UK motorways (Newman, 2005). But not always this idea holds. Some cases vary over an enormous dynamic range, sometimes many orders of magnitude. Those distributions show high skewness, that is, asymmetry or distortion from the bell-like mean. Interestingly, when they are replotted in logarithmic scales (log-log scale), the distribution closely follows a straight line with negative slope. Hence, the equation describing the line looks like: $\ln y(x) = -\alpha \ln(x) + c$, where α and c are constants. Taking exponentials on both sides, it changes into:

$$y(x) = Cx^{-\alpha}, \text{ where } C = e^{c} \tag{1}$$

Distributions of the form (1) are said to follow a power law, with α being the exponent (Newman, 2005).

Power law distributions can be observed in a diverse range of fields – city population (Gabaix, 1999), seismological frequency (Gutenberg & Richter, 1944), solar flare distributions (Lu & Hamilton, 1991), internet network traffic (Crovella & Bestavros, 1997), word usage frequency in human languages (Zipf, 1949), distribution of family names (Zanette & Manrubia, 2001), distribution of scientific papers (Lotka, 1926), number of hits on web pages (Adamic, 2000), recording sales (Cox et al., 1995), geographical distribution of plants and animals (Willis & Yule, 1922), annual incomes (Pareto, 1964) and in human cognitive functioning as well (Kello et al., 2010). Newman in his paper (2005) illustrated few examples of cumulative frequency distributions (rank vs. frequency plots), all of whom follows power law characteristics (**Fig. 1.10.1**).

In linguistic behaviour, a power law distribution that has been studied for almost 70 years is Zipf's law, named after G.K. Zipf who described it in his pioneering study in 1949. Zipf's law, as originally formulated, states that the frequency of a word (f) in a given corpus is proportional to the inverse of its frequency rank (r), $f \sim 1/r$ (which is a distribution of the form (1), with the exponent α being 1). Zipf originally explained his law in terms of the principle of least effort, that is, the distribution appears as a result of the optimization between speakers' preference of high-frequency words and listeners' preference of low-frequency words. Same basic principle can also be applied at other linguistic scales, which suggests Zipf's law is an adaptive property of communication. Evidently, this law has been found to be universal and nontrivial in the field of human languages (i Cancho & Solé, 2003; i Cancho, 2010). Over the years, Zipf's law, its modifications and altercations gave a plethora of distribution laws, which are applied to different complex systems. These include Zipf-Mandelbröt law (Mandelbröt, 1965), Yule-Simon distribution (Simon, 1955), Menzerath-Altmann law (Cramer, 2005), and Pareto law (Pareto, 1964) among others. These Zipf and Zipf-like distributions have been claimed to be present in a diverse range of physical phenomena (Li, 2002): word usage in human languages, city population, Webpage visits and various bibliometric data.

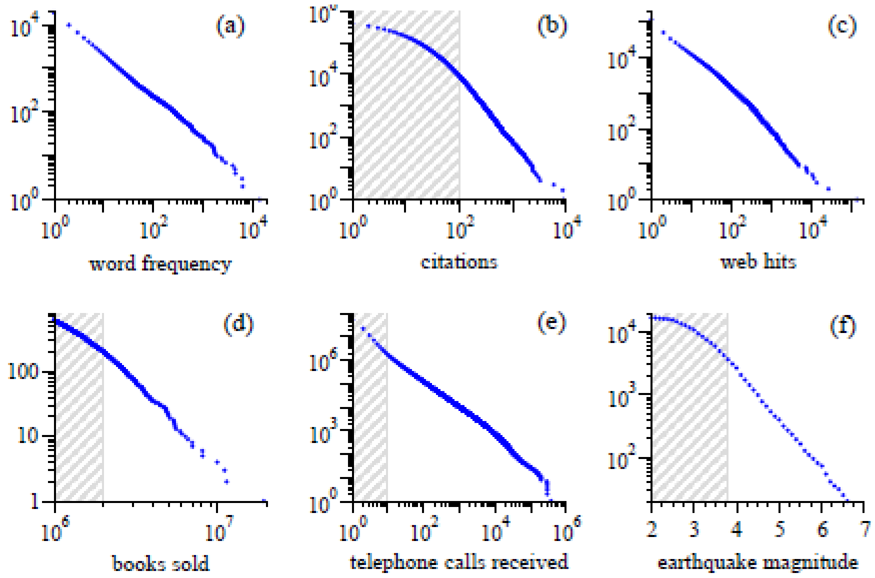


Fig. 1.10.1. Rank-frequency plots of various quantities following power law given in Newman (2005). (a) Numbers of occurrences of words in the novel Moby Dick by Hermann Melville. (b) Numbers of citations to scientific papers published in 1981. (c) Numbers of hits on web sites by 60000 users of the America Online Internet service for the day of 1 December 1997. (d) Numbers of copies of bestselling books sold in the US between 1895 and 1965. (e) Number of calls received by AT&T telephone customers in the US for a single day. (f) Magnitude of earthquakes in California between January 1910 and May 1992

In recent years, one interesting use of Zipf's law in the field of linguistics has been the development of thermodynamic variables like Energy, Temperature and Specific heat to categorize and quantify complex system using their information content. The concept of 'temperature' was given long back by Mandelbröt (1953), which was reiterated and extended in subsequent works by a number of authors to be used in measuring communicative ability, comparing vocabulary complexity levels, assessing readership suitability or evaluating author's writing performances (De Campos & Tolman, 1982; Kosmidis et al., 2006; Miyazima & Yamamoto, 2008; Reˆgo et al., 2014; Chang et al., 2017). The basis of this concept is the assumption that human language can be described as a physical system within the framework of equilibrium statistical mechanics. In Mandelbrot's interpretation, the text's informational temperature (θ) is reciprocal to a state variable B, as in, $\theta = 1/B$ and whose value, barring some very rare occasions, is always < 1. Nearer the text temperature to 1, greater the 'wealth of the vocabulary'. On the other hand, low temperatures indicate 'badly employed' words. Vocabulary of James Joyce and vocabulary of children are the examples of said two types, respectively. Kosmidis et al. connected the Zipf exponent to the temperature and suggested that this parameter can be used to measure the text's communicative ability. Miyazima-Yamamoto and later Reˆgo et al. associated words with energies based on a general standard Maxwell–Boltzmann (MB) distribution. It is found that, the linguistic relative temperature of a book can be determined by measuring the deviation from a standard Maxwell–Boltzmann distribution of a corpus of English words. This relative temperature parameter can be used to measure vocabulary complexity level (Miyazima & Yamamoto, 2008) or author's writing capacity and repertory of thoughts (Chang et al., 2017). In a different take, Rovenchak and Buk (2011) focused on the behaviour of low-frequency words and mapped word rank–frequency distributions onto the Bose-Einstein (BE) distribution within the grand-canonical approach. The respective physical analogues are the power of the excitation spectrum α and the temperature T (Rovenchak, 2014). The analogue of the fugacity z is determined from the number of words occurring only once (so-called '*hapax legomena*'). It was seen that the calculated parameters have a correlation with the language structure (the level of analyticity). Lower parameter values correspond to higher analyticity, indicating lesser word inflection.

Since music and language both are embedded in human social construct for communicative as well as entertainment purposes, it is only natural to attempt to seek commonalities between the two. Language and music share more than one aspect, cognitively and structurally (Jackendoff, 2009). Both require memory capacity for storing representations which is later integrated in a combinatorial manner by means of a set of structural rules. The processing of both language and music involves creating expectations of what is to come. Fine-scale voluntary vocal production and the ability to imitate it are involved in the learning and producing of both. And most importantly, they share the characteristic of individual's ability to create and improvise, in order to be emotionally meaningful and also, distinct. Structurally, both phonology and music share similar metric system, based on a hierarchical metrical grid (although the grid appears far more isochronous for latter). In some languages (West African, Chinese) a discrete tonal feature is found, as in, tones do not drift downward throughout the utterance. Patel argues that such languages provide a closer parallel to musical pitch (Patel, 2008). Syntactic structure of language and prolongational structure in music could also provide a parallel since they are both recursive headed hierarchies. Neuroimaging studies have indicated that Broca's area in brain is responsible for both the syntactic comprehension and music perception (Patel, 2003).

In view of this, the structural complexity of music, a congregation of informative symbolic sequences similar to the human language, is analysed in this thesis using the derived statistical distributions. The parameters developed helps to categorise Indian Classical Music (ICM) according to their information content, dynamic nature and improvisation patterns, something that could provide a stepping stone towards understanding their cognitive distinction due to hidden emotional attributes, at least in ICM scenario.

1.11. EMOTIONS, AMBIGUITY AND CREATIVE AESTHETICS

We discussed in section 1.3 about the subjective vs. objective debate on emotion. The subjective experience corresponding to an affective event has always been a widely researched topic. Since the experience only involves the subject, the affective responding has largely relied on two-dimensional quantification of said experience. That is: Valence, the quality of acceptation or aversion to the stimulus and Arousal, the degree of emotion the stimulus can invoke (positive or negative) (Russell, 1980; Lang et al., 1998). This view has dominated the research on subjective response to affective stimulus for long (Kensinger, 2004). Over time, majority of works in affective psychology has adopted experimental designs and theoretical discussions around two contrasting approaches to measure the subjective ratings based on valence-arousal dimensions. These two approaches are called bipolar and bivariate models (although Kuppens et al. (2013) discussed six, but these two have managed the highest participation in empirical findings among them). The bipolar model says that valence and arousal are independent of each other, but higher levels of perceived positivity mean lower levels of negativity and vice versa. One might put it as a zero-sum game of arousal. On the other hand, bivariate model suggests that the perceived positivity and negativity are distinct, but changes in arousal are a function of changes in valence (Brainerd, 2018). It is evident that these approaches differ quite a lot and as one might guess, their findings are often contradictory (Kuppens et al., 2013, Mattek et al., 2017). In response to this, recent years has seen some studies asking for a higher dimensional model to replace the valence-arousal duopoly. For instance, Yik et al. (2011) offered a 12-point Circumplex model to describe 'Core Affect', which according to them "is the heart of, but not the whole of, mood and emotion". Similarly, Cowen et al. (2020) used music as stimulus and constructed a complex, high-dimensional (at least 13) space of subjective experience associated with music in multiple cultures. In a different vein, the valence-arousal relationship should be analysed by correlating them with a third variable, namely – valence ambiguity, proposed Mattek et al. (2017). Instead of the usual notion of a third variable separately being correlated with each of the previous two, Mattek et al. suggest that authentic correlations between the primary two variables appear and disappear, or even reverse direction, depending on levels of the third. An unmeasured ambiguity variable, which is the degree of indefiniteness or uncertainty in people's subjective impressions, directly controls the perceived valence and perceived arousal (Brainerd, 2018).

Indeed, the resolution of ambiguity is an inseparable part of our everyday interactions with the world, with ambiguous information being presented to all our sensory modalities. A touch on skin could both be an innocent strand of hair or a potentially poisonous insect. A sentence could both be literal or figurative. Generally, the ability to interpret ambiguous signs correctly in favour of the responder is crucial for adaptive functioning. A reasonable part of emotion research has been about how emotional states could influence or resolve these ambiguities. Some research has examined the processing of ambiguous emotional information, and what processes are involved in resolution in all individuals, whereas other research has focused on individual differences in these interpretations (this includes the ones mentioned earlier). Many forms of verbal and non-verbal ambiguity, including lexical ambiguity in form of homophones (e.g., *brews* and *bruise*), semantic ambiguity in the form of homographs (e.g., stroke is both brain haemorrhage and caress), ambiguous pictures and facial expressions, ambiguous musical forms etc. has been studied on and with [see Blanchette & Richards (2010) for a comprehensive account]. Studies on anxiety-congruent interpretation have found that anxiety has robust effects on stimulus interpretation: anxious individuals are prone to interpret ambiguous stimulus more negatively than neutrals. Also, contextuality plays an important role in ambiguity resolution. Recent studies have shown that participants resolve emotionally ambiguous information using contextual references. The interpretive effects, although, mostly exhibited by anxiety but research on judgement reveals that it is affected by a wide range of emotions including anger, sadness, anxiety and positive ones. Blanchette suggest that this is due to the difference in underlying mechanism - attention for interpretation and memory for judgement (Blanchette & Richards, 2010). Similarly, affective ambiguity has influenced other cognitive functions like decision making and reasoning, too.

In her seminal works in 1949 and 1951, Frenkel-Brunswik argued that tolerance of ambiguity or TA is an "emotional and perceptual personality variable" and generalises to the various aspects of emotional and cognitive functioning of the individual, characterising cognitive style, belief and attitude systems, interpersonal and social functioning and problem-solving behaviour. She also related TA to other personality variables, predicting a positive relationship with the authoritarian family of personality traits. Since then, the topic has attracted considerable research and remains a well-used variable to this day (check out Furnham & Marks (2013) for a review of the same). Debated on whether it's a context-specific construct or a personality trait, TA is found to be linked with the approach-vs.-aversion feature and it is found that individuals with low TA perceive ambiguous situations as threat due to lack of information and expresses aversion. Additionally, various studies have associated TA with different measures and behaviours like - ethical decision making, openness, challenge and threat appraisals, Entrepreneurial performance, Schizotypal personality disorder, identity conflict, thinking style, risk-taking propensity, equivocality to name a few (Furnham & Marks, 2013).

Ambiguity has influenced another field of study in great detail, namely the aestheticism of modernist art and literature. 20th century models of poetical or literary language based on deviational aesthetics have practically attached highest importance to the phenomenon of ambiguity, although not always intentionally. From the structuralist model of Roman Jakobson to the ideas of the Soviet semiotician Jurij Lotman or the aesthetics of the post-structuralist Roland Barthes - all acknowledge that ambiguity, central to poetry and literature is an unavoidable and necessary effect emanating from the restructuring of the language used (Bode, 1991). The fundamental idea of all deviational aesthetics is to take conceptual elements out of their usual context and reassemble in a new and unexpected way. The aesthetic experience is based on the viewer or reader's effort in tackling the artwork in a secondary plane trying to decipher what these once familiar but now strangely recontextualized elements mean beyond their everyday referential use. This is where ambiguity comes in crucial. Since no secondary structuring of such conceptualities can totally obliterate the deeply ingrained referential meanings but can only, by various devices, loosen their formal ties. Hence, these elements now characteristically oscillate between what they usually mean and the new meaning they are striving to constitute. Bode argues that the high degree of ambiguity in modernist literature and art can be identified as an unavoidable spin-off effect of a superordinate tendency or evolution towards higher 'auto-referentiality' - i.e., the effect of simultaneously slowing down the recognisability of usual elements by unfamiliar use and appealing to the imagination and creativity – which is discernible in all arts, but more pronounced in literature.

TA and creativity share a long history too. Vernon (1970) considered tolerance of ambiguity to be an essential factor of creativity since it enables individuals to not be satisfied by partial or non-optimal solutions to complex problems. Other studies have linked creativity with TA based on the hypothesis that tolerance of ambiguity allows individuals to continue to remain open, and increase the probability of finding a novel solution to any complex problem (Sternberg & Lubart, 1995; Urban, 2003). Since ambiguous stimulus or events invoke anxiety and general psychological discomfort, features which are similar to complex problems requiring out-of-the-box solutions, individuals who can handle the former well are more likely to generate the latter (Zenasni et al., 2008). Bhusan and Amal (1986) reviewed the existing research to concentrate the reactions of people with low TA into three separate types: a) cognitive reactions: where individuals perceive the ambiguous situations as either entirely black or white, b) emotional reactions: where individuals respond to ambiguity with expressions of anger, dislike, uneasiness, anxiety, and discomfort and c) behavioral reactions: where individuals attempt to avoid or even reject ambiguous situations. On the other hand, people with higher tolerance to ambiguity have found ambiguous situations interesting and enjoyable and produced more creative output than their lower TA counterparts (Furnham & Ribchester, 1995). The empirical research on this front has been scarce but the limited existent works support theoretical ideas to great extent. Tegano (1990) used a Tolerance of Ambiguity Scale to report a positive correlation between tolerance of ambiguity and creative style score. Students with higher tolerance to ambiguity were found to produce more ideas and showed positive attitude towards the task in a study (Comadena, 1984). Similar results were found by Stoycheva (1998, 2003). Stoycheva argues that the tolerance of ambiguity is an important motivation for creativity because firstly, creative work demands this ability since during the process one needs to learn to cope with the intuitive reluctance towards ambiguous situations. Secondly, qualities related to creativity like risk-taking, non-conformism, and openness to experiences are integrated into tolerance to ambiguity, qualities that balances the resistance adaptation and furthers creative process. Thirdly, when faced with ambiguous situations, it is important to control the tendency to jump directly to easy, simple, and unambiguous solutions since "Resistance to premature closure and psychological openness are beneficial to the creative process, allowing time and space for a free and flexible exploration of the incoming information". However, Stoycheva also points out that tolerance of ambiguity is necessary for creativity but not enough. Intellectual competencies, domain-specific knowledge and skills, creativity-relevant abilities, task commitment, motivation and other personality traits are also important (Liu, 2015).

To sum up, it is more than evident that ambiguity and acceptance or rejection to it serves an important purpose in the perceptual space. From the manipulation of the emotional states and moods and feelings, to the higher-level cognitive aspects like judgement and decision making, to the development of personality and creative potential – ambiguity affects them all in various degrees. But compared with other variables, such as anxiety, empathy or motivation which have been well discussed and investigated by researchers and theorists, the study of tolerance of ambiguity is a relatively under-cultivated area, although it seems to bear no less importance. Moreover, studies on the psychological domain have relied excessively on methods that have become increasingly outdated in physical systems research. This makes one wonder that when the governing rules of micro and macro worlds continued to evolve, why shouldn't the same happen with methods to investigate high level functioning of human brain? Properties like contextuality, superposition and interference are frequent and crucial in explaining human cognitive domain and behavioural phenomena and can prove to be detrimental in developing successful models of mind and emotion. In view of this, in the third and final part of the thesis, we have attempted to use one of the most important physical theories developed in the last century, namely Quantum theory, in order to hypothesize a possible model regarding the creative process and its relation to stimulus ambiguity. First, our objective is to show that during the perception of an ambiguous stimulus, classical Bayesian theories of cognition does not hold true and there exists an element of non-classicality which can't be explained via classical approach. Then, we proceed to propose the 'Quantum Leap interpretation' of creative process, based on the free association theory, that we believe could help in explaining the role of ambiguity in the domain of creativity and also provide a platform for future studies in this field that incorporates advanced ideas of physics.

1.12. QUANTUM APPROACHES TO MIND AND BRAIN

In this section we shall briefly review the existing ideas and approaches related to the quantum theory in the field of human brain and cognition (Atmanspacher, 2017). Although the starting point of the mind-brain problem in scientific history has been through disciplines like philosophy and psychology, but they were later joined by behavioral science, cognitive science, and neuroscience. In addition, the physics of complex systems and quantum theory have played stimulating roles in the discussion as well. Since the brain evidently is one of the most complex systems observable, the study of neural complexity and their relation to the operation of several brain mechanisms and complex cognitive models have been quite profitable using complex systems approaches. Similarly, quantum events also occur and are effective in the brain as elsewhere in the material world – including biological systems (see Huelga & Plenio (2013) for a review). But it is greatly debated whether these events can cause or influence or even are relevant in the discussion of mind and cognitive aspects.

The initial motivation of applying quantum properties to explain psychological domain was the philosophical stance regarding 'free will' and its intuitive non-determinism. It was suggested that quantum randomness, as opposed to epistemic and deterministic classical randomness, could open up novel possibilities to account for free will. Other features of quantum theory, like complementarity and entanglement, were also considered to be key in resolving the physical determinism vs. conscious free will by Bohr, Schrödinger, Pauli (among others) - the founding fathers of quantum physics. Using Bohr's proposal of extraphysical relevance of complementarity, in recent years a number of studies have been developed with respect to psychology and cognitive science. These include addressing quantum-like behavior in non-quantum systems by Aerts et al (1993), application of quantum mechanics in cognitive spaces with non-classical probabilities (Khrennikov, 1999) and outlining an algebraic framework with non-commuting operations (Atmanspacher et al., 2002). There are two main avenues that theories of quantum-brain associations follow. Firstly, the usage of quantum theory in explaining the structural aspect of the brain and emergence of the non-physical from it (or the 'Quantum Brain' approach) and secondly, using quantum concepts to describe various cognitive phenomena with an agnostic view of the former approach (the 'Quantum Mind' approach).

A substantial part of quantum brain approach stems from the age-old debate on the nature of observation and measurement. In 1955, von Neumann described measurement as a discontinuous, non-causal, instantaneous and irreversible act through which the wave-function collapses. He also discussed the conceptual distinction between observed and observing system and concluded that the nature of the observer (human or non-human) is redundant (von Neumann, 2018; originally published in 1955). In contrast, first London and Bauer (1939) and then Wigner (1967) put emphasis that it is indeed human consciousness which completes the quantum measurement process (though Wigner later renounced this claim). More recently, Stapp used Heisenberg's distinction between the *potential* and *actual* events and attached an experiential aspect to the actual that 'gives it its status as an intrinsic actuality' (Stapp, 2004). Furthermore, Stapp also advocates the idea that conscious intentions can influence the material brain activities. According to him, each conscious experience leads to a quantum state reduction which is the 'neural correlate of that conscious experience' (Stapp, 1999). Stapp, needless to say, is one of the most vocal supporters of the influence of observer in quantum physical processes and argues that the randomness of individual quantum events originates from the positive or negative biases in the minds of the observers (Stapp, 2015).

A different approach to describe the brain states comes from Ricciardi and Umezawa by means of quantum field theory. They address the brain as a many-particle system (where particles are basically neurons and neuronal assemblies correlate directly with mental activity) and provide a quantum field theoretical derivation of ordered states in many-body systems (Ricciardi & Umezawa, 1967). These ordered states represent coherent activity in neuronal assemblies. Later, Vitiello (1995) advanced this idea with adding the concept of dissipation and showed how system-environment interactions lead to multiplication of such ordered states with finite lifetimes. He also points out that the emergence of (self-similar, fractal) power-law distributions in general is intimately related to dissipative quantum coherent states (Vitiello, 2012). Overall, the main role of the quantum field theory model is to provide the explanation of the classical behaviour at the level of brain activity from the inherent quantumness.

One of the most debated and discussed ideas in the field of quantum brain is the Orch OR theory proposed by Hameroff & Penrose (1996). This theory argues that argued that elementary acts of consciousness are non-computable, and their neurophysiological manifestation is the gravitation-induced reductions of coherent superposition states in brain microtubuli. Penrose (1994) proposes that quantum state reduction should describe an objective physical process and can only be facilitated by gravitational effects; hence, it is a 'gravitation-induced objective state reduction'. The physical place that allows this event is the brain microtubules, special locations where the quantum states can live long enough to be collapsed and form consciousness. Though criticised on the possible decoherence duration being too short to be significant for neurophysiological processes (Tegmark, 2000), studies have also emerged supporting this revolutionary idea. For example, there were evidences that entangled states can be maintained in noisy open quantum systems at high temperatures (Hartmann et al., 2006). Also, features similar to macroscopic quantum systems were found in microtubules (Sahu et al., 2013).

The quantum mind approach sits quite aloof from the methodological attempts of explaining brain in terms of quantum theory. It incorporates quantum behavior in mental systems, without referring to quantum brain activity explicitly, by describing mental systems based on well-constructed state spaces. As Atmanspacher (2017) puts it: "it is likely that mental states and observables show features that resemble quantum behavior although the correlated brain activity may be entirely classical–quantum mind without quantum brain." Various psychological features are analysed using quantum framework including decision processes, sequence effects, bistable perception, learning, semantic networks and so on.

The basis of the usage of quantum theory in decision and judgment processes is defining the outcomes in terms of quantum probability amplitudes. It is seen that the applicability includes not only general decision making (Busemeyer et al., 2006), but in conjunction and disjunction effects (Pothos & Busemeyer, 2009) and asymmetry of similarity judgments (Pothos et al., 2013) as well. Similarly, sequence and order effects in polls can be well understood with contextual quantum effects (Aerts & Aerts, 1995). Bistable perception, a form reaction to ambiguous stimulus, is addressed by different quantum approaches. For instance, Manousakis (2009) proposed a model for the subjective psychophysical dynamics during binocular rivalry using orthodox quantum theory of measurement, where the observer plays a key role. On the other hand, applying quantum zeno features, Atmanspacher and Filk (2013) forwarded a model of bistable perception, known as Necker–Zeno model. Additionally, they also proposed a conjecture that particular distinguished states in bistable perception may violate temporal Bell inequalities – if proven, this would indicate definite presence of quantum behaviour (Atmanspacher & Filk, 2010). Similar non-commutative behaviour is shown to be observed in numerical and analytical learning processes too (Atmanspacher et al., 2006). For the interesting case of meaning in natural languages, Gabora & Aerts (2002) used state context property formalism, originally developed as a generalization of quantum mechanics, and described the contextual manner in which concepts are evoked, used and combined to generate meaning. Bruza et al. (2015) framed the principle of semantic compositionality in terms of probabilistic models to associate meaning with entanglement-style features, for which they provided empirical results as well.

One more approach that exists parallelly in the realm of quantum approaches to mind and brain is the 'double-aspect theory' which considers the mental and physical to be two aspects of the same underlying substance or reality. Depending on the explanations on how these two aspects get separate entities, there are several ideas that belong to this domain such as: Chalmers' ideas on 'consciousness and information', Bohm and Hiley's 'implicit and explicit order', Pauli-Jung conjecture of 'dual-aspect monism', Hans Primas' 'time-entanglement of mind and matter' etc. (check (Atmanspacher, 2017) for a condensed overview).

The following chapters of this thesis give an account of the diverse variety of unique experiments in the domain of neuro-cognition of various chromatic and acoustic stimulus performed by me for the last 5 years at the laboratory of Sir C.V. Raman Centre for Physics and Music, Jadavpur University, India. Using state-of-the-art robust scientific tools for biological and acoustic signal analysis, an attempt was made towards the quantification, along with categorization, of emotion and emotion-related perceptual signatures. The concept of color-music correspondence has been dealt with from a nonlinear approach in this thesis. Additionally, visual perception of chromatic cues is analyzed in terms of the complexities they produce in the cognitive level. We have also attempted to build a method of acoustic categorization based on scaling laws to introduce novel parameters which promises to uncover hidden structural patterns responsible for its perceived emotional association. Also, a study on the non-classical nature of the cognition of ambiguous acoustic stimuli has been attempted for the first time. The thesis ends with an ambitious proposition of a novel 3-step model of creative cognition using the concept of orthodox quantum theory of measurement and the Wigner distribution of the mental state function; it tries to explain the chaotic classical evolution of mental states emanating from quantum processes in terms of ambiguity of the stimulus and continuous system measurement. This thesis is expected to provide a platform as well as a direction for researchers in complex systems, neuroscience, cognitive science and psychology to explore further based on these interesting premises.

REFERENCES

1.1. Nature's creation and human appreciation through mind, brain and other organs

Aerts, D. (2012). Quantum interference and superposition in cognition: A theory for the disjunction of concepts. In *Worldviews, science and us: Bridging knowledge and its implications for our perspectives of the world* (pp. 169-211).

Busemeyer, J. R., Pothos, E. M., Franco, R., & Trueblood, J. S. (2011). A quantum theoretical explanation for probability judgment errors. *Psychological review*, *118*(2), 193.

Busemeyer, J. R., Wang, Z., & Townsend, J. T. (2006). Quantum dynamics of human decision-making. *Journal of Mathematical Psychology*, *50*(3), 220-241.

Churchland, P. S. (1980). A perspective on mind-brain research. *The Journal of Philosophy*, *77*(4), 185-207.

Damasio, A. R. (2000). A second chance for emotion. *Cognitive neuroscience of emotion*, *12*, 23.

Dolcos, F., Wang, L., & Mather, M. (2014). Current research and emerging directions in emotion-cognition interactions. *Frontiers in Integrative Neuroscience*, *8*, 83.

Durant, W. (1961). *Story of philosophy*. Simon and Schuster. New York: Washington Square Press Publication of Pocket Books.

Hansotia, P. (2003). A neurologist looks at mind and brain:"The enchanted loom". *Clinical medicine & research*, *1*(4), 327-332.

Johnson, M. L. (2001). The feeling of what happens: body and emotion in the making of consciousness.

Khrennikov, A. (2009). Quantum-like model of cognitive decision making and information processing. *Biosystems*, *95*(3), 179-187.

Ochsner, K. N., Bunge, S. A., Gross, J. J., & Gabrieli, J. D. (2002). Rethinking feelings: an FMRI study of the cognitive regulation of emotion. *Journal of cognitive neuroscience*, *14*(8), 1215-1229.

Okon-Singer, H., Hendler, T., Pessoa, L., & Shackman, A. J. (2015). The neurobiology of emotion–cognition interactions: fundamental questions and strategies for future research. *Frontiers in human neuroscience*, *9*, 58.

Pessoa, L. (2013). *The cognitive-emotional brain: From interactions to integration*. MIT press.

Pinchot, R. B., & Gersten, J. (1930). *The brain: mystery of matter and mind.* Washington DC: US New Books.

Russell, B. (2022). *My philosophical development*. Taylor & Francis. (Originally published by London: George Allen and Unwin, Ltd., 1959.)

Schifferstein, H. N., & Spence, C. (2008). Multisensory product experience. In *Product experience* (pp. 133-161). Elsevier.

Schwartz, J. M., & Begley, S. (2009). *The mind and the brain*. Springer Science & Business Media.

Schwartz, J. M., Stapp, H. P., & Beauregard, M. (2005). Quantum physics in neuroscience and psychology: a neurophysical model of mind–brain interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1458), 1309-1327.

York III, G. K., & Steinberg, D. A. (2011). Hughlings Jackson's neurological ideas. *Brain*, *134*(10), 3106-3113.

1.2. Concept of mind: Indian and Western approaches

Armstrong, A. H. (1968). Plotinus. *The Classical Review*, *18*(02).

Atreya, J. P. (1985). *Mind and its function in Indian thought*. Classical Publ. Company.

Bala, S. (2009). *The Concept of Mind* (Doctoral dissertation, Guru Nanak Dev University Amritsar).

Berkeley, G. (1948). An essay toward a new theory of vision, 1709.

Chakrabarti, K. K. (1999). *Classical Indian philosophy of mind: The Nyaya dualist tradition*. SUNY Press.

Churchland, P. M. (1981). Eliminative materialism and propositional attitudes. *the Journal of Philosophy*, *78*(2), 67-90.

Davidson, D. (2001). *Essays on Actions and Events: Philosophical Essays Volume 1*. Clarendon Press.

Descartes, R. (1988). *Descartes: Selected philosophical writings*. Cambridge University Press.

Feigl, H. (1958). The 'mental'and the 'physical'. *Minnesota studies in the philosophy of science*, *2*(2), 370-497.

Feyerabend, P. (1963). Mental Events and the Brain. *Journal of Philosophy*, 40, 295–296.

Fichte, J. G. (1993). *Fichte: Early philosophical writings*. Cornell University Press.

Giles, J. (1993). The no-self theory: Hume, Buddhism, and personal identity. *Philosophy East and West*, *43*(2), 175-200.

Green, C. D. (2003). Early psychological thought: Ancient accounts of mind and soul.

Hegel, G. W. F. (1977). *Difference Between Fichte's and Schelling's System of Philosophy, The: An English Translation of GWF Hegel s Differenz des Fichte schen und Schelling schen Systems der Philosophie*. Suny Press.

Horgan, T. (1982). 1 982: Supervenience and micro-physics. *Pacific Philosophical Quarterly*, *63*, 27-43.

Jackson, F. (1982). Epiphenomenal qualia. The Philosophical Quarterly 32 (127), 127–136.

Jackson, F. (1998). Postscript on Qualia, in his mind, method and conditionals (pp. 76–79).

Kant, I. (1998). Critique of Pure Reason. *Trans. and ed. P. Guyer and AW Wood. Cambridge: Cambridge University Press*.

Kripke, S. A. (1972). Naming and necessity. In *Semantics of natural language* (pp. 253-355). Springer, Dordrecht.

Lange, F. A. (1865). The History of Materialism and Criticism of its Present Importance, trans. E. Thomas. *New York: The Humanities Press*, *75*, 1950.

Larmer, R. (1986). Mind-body interaction and the conservation of energy. *International philosophical quarterly*, *26*(3), 277-285.

Loptson, P. (2006). Leibniz's body realism: Two interpretations. *The Leibniz Review*, *16*, 1-42.

McLaughlin, B. P. (1984). Perception, causation, and supervenience. *Midwest Studies in Philosophy*, *9*(1), 569-591.

Place, U. T. (1970). Materialism as a scientific hypothesis. In *The Mind-Brain Identity Theory* (pp. 83-86). Palgrave, London.

Radhakrishnan, M. (2009). Indian Philosophy of Mind: A Comparative Study. فلسفه, *37*(1), 1-19.

Radhakrishnan, S. (1953). *History of Philosophy, Eastern and Western*. Geoge Allen & Unwin, London.

Reichenbach, H. (1938). Experience and prediction: An analysis of the foundations and the structure of knowledge.

Vidyabhusana, S., Basu, B. (1913). *The Nyaya Sutras of Gotama: The sacred book of Hindus*. In Basu, M.D. (Ed.), The sacred books of the Hindus (pp. 1–175). Allahabad: Panini office, Bhuvaneswari Asrama.

Von Schelling, F. W. J. (1988). *Ideas for a Philosophy of Nature*. Cambridge University Press.

Whitney, W. D. (1890). Translation of the Katha-Upanishad. *Transactions of the American Philological Association (1869-1896)*, *21*, 88-112.

1.3. Concept of Emotion

Anderson, D. J., & Adolphs, R. (2014). A framework for studying emotions across species. *Cell*, *157*(1), 187-200.

Arnold, M. B. (1960). *Emotion and personality.* Columbia University Press.

Barrett, L. F. (2009). The future of psychology: Connecting mind to brain. *Perspectives on psychological science*, *4*(4), 326-339.

Barrett, L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Social cognitive and affective neuroscience*, *12*(1), 1-23.

Bedford, E. (1956, January). Emotions. In *Proceedings of the Aristotelian Society* (Vol. 57, pp. 281-304). Aristotelian Society, Wiley.

Berridge, K. C. (2018). Evolving concepts of emotion and motivation. *Frontiers in Psychology*, 1647.

Cannon, W. B. (1927). The James-Lange theory of emotions: A critical examination and an alternative theory. *The American journal of psychology*, *39*(1/4), 106-124.

Damasio, A. R. (2004, April). Emotions and feelings. In *Feelings and emotions: The Amsterdam symposium* (pp. 49-57). Cambridge: Cambridge University Press.

Darwin, C., & Prodger, P. (1998). *The expression of the emotions in man and animals*. Oxford University Press, USA.

Descartes, R. (1989). The passions of the soul (S. Voss, Trans.). Indianapolis, IN: Hackett Publishing. (Original work published 1649).

Dewey, J. (1894). The theory of emotion: I: Emotional attitudes. *Psychological review*, *1*(6), 553.

Dewey, J. (1895). The theory of emotion. *Psychological review*, *2*(1), 13.

Ekman, P. (1977). Biological and cultural contributions to body and facial movement. *The anthropology of the body*.

Ekman, P. (1999a). Basic emotions. *Handbook of cognition and emotion*, *98*(45-60), 16.

Ekman, P. (1999b). Facial expressions. *Handbook of cognition and emotion*, *16*(301), e320.

Ekman, P., Freisen, W. V., & Ancoli, S. (1980). Facial signs of emotional experience. *Journal of personality and social psychology*, *39*(6), 1125.

Fortenbaugh, W. W. (1975). Aristotle on emotion: a contribution to philosophical psychology, rhetoric, poetics, politics, and ethics.

Frijda, N. H. (2007). What might emotions be? Comments on the Comments. *Social Science Information*, *46*(3), 433-443.

Gendron, M., & Feldman Barrett, L. (2009). Reconstructing the past: A century of ideas about emotion in psychology. *Emotion review*, *1*(4), 316-339.

Ghose, P. (2012). Nonquantum Cognition. *arXiv preprint arXiv:1210.4473*.

Greenwald, A. G., & Banaji, M. R. (2017). The implicit revolution: Reconceiving the relation between conscious and unconscious. *American Psychologist*, *72*(9), 861.

Hume, D. (2003). *A treatise of human nature*. Courier Corporation.

Izard, C. E. (1971). *The face of emotion.* Appleton-Century-Crofts.

Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion-cognition relations. *Psychological Review*, *99*(3), 561–565.

James, W. (1984). *Psychology, briefer course* (Vol. 14). Harvard University Press.

Kenny, A. (2003). *Action, emotion and will*. Routledge.

Lazarus, R. S. (1991). Cognition and motivation in emotion. *American psychologist*, *46*(4), 352.

LeDoux, J. E. (2014). Coming to terms with fear. *Proceedings of the National Academy of Sciences*, *111*(8), 2871-2878.

LeDoux, J. E., & Hofmann, S. G. (2018). The subjective experience of emotion: a fearful view. *Current Opinion in Behavioral Sciences*, *19*, 67-72.

Morris, J. S., Öhman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature*, *393*(6684), 467-470.

Nussbaum, M. (1996). Compassion: The basic social emotion. *Social Philosophy and Policy*, *13*(1), 27-58.

Nussbaum, M. (2001). Upheavals of thought: the intelligence of emotions Cambridge University Press.

Panksepp, J. (2011). The basic emotional circuits of mammalian brains: do animals have affective lives?. *Neuroscience & Biobehavioral Reviews*, *35*(9), 1791-1804.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological review*, *110*(1), 145.

Scarantino, A. (2014). The motivational theory of emotions. In J. D'Arms & D. Jacobson (Eds.), *Moral psychology and human agency: Philosophical essays on the science of ethics* (pp. 156–185). Oxford University Press.

Scherer, K. R., Schorr, A., & Johnstone, T. (Eds.). (2001). *Appraisal processes in emotion: Theory, methods, research*. Oxford University Press.

Solomon, R. C. (1993). *The passions: Emotions and the meaning of life*. Hackett Publishing.

Solomon, R. L. (1980). The opponent-process theory of acquired motivation: the costs of pleasure and the benefits of pain. *American psychologist*, *35*(8), 691.

Tomkins, S. S. (2008). *Affect imagery consciousness: the complete edition: two volumes*. Springer publishing company.

Winkielman, P., Berridge, K. C., & Wilbarger, J. L. (2005). Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Personality and social psychology bulletin*, *31*(1), 121-135.

1.4. Emotion, Perception, Cognition

Armony, J. L., & Dolan, R. J. (2001). Modulation of auditory neural responses by a visual context in human fear conditioning. *Neuroreport*, *12*(15), 3407-3411.

Baumgartner, T., Esslen, M., & Jäncke, L. (2006). From emotion perception to emotion experience: Emotions evoked by pictures and classical music. *International journal of psychophysiology*, *60*(1), 34-43.

Björk, E. A. (1985). The perceived quality of natural sounds. *Acustica*, *57*(3), 185-188.

Bradley, M. M., & Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology*, *37*(2), 204-215.

Bresciani, J. P., Ernst, M. O., Drewing, K., Bouyer, G., Maury, V., & Kheddar, A. (2005). Feeling what you hear: auditory signals can modulate tactile tap perception. *Experimental brain research*, *162*(2), 172-180.

Brosch, T., Pourtois, G., & Sander, D. (2010). *The perception and categorisation of emotional stimuli: A review* (pp. 76-108). Psychology Press.

Brosch, T., Scherer, K. R., Grandjean, D. M., & Sander, D. (2013). The impact of emotion on perception, attention, memory, and decision-making. *Swiss medical weekly*, *143*, w13786.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience*, *3*(3), 201-215.

De Gelder, B., Morris, J. S., & Dolan, R. J. (2005). Unconscious fear influences emotional awareness of faces and voices. *Proceedings of the National Academy of Sciences*, *102*(51), 18682-18687.

De Gelder, B., Pourtois, G., & Weiskrantz, L. (2002). Fear recognition in the voice is modulated by unconsciously recognized facial expressions but not by unconsciously recognized affective pictures. *Proceedings of the National Academy of Sciences*, *99*(6), 4121-4126.

Dolan, R. J. (2002). Emotion, cognition, and behavior. *science*, *298*(5596), 1191-1194.

Driver, J. (2001). A selective review of selective attention research from the past century. *British journal of psychology*, *92*(1), 53-78.

Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron*, *57*(1), 11-23.

Fan, J. (2014). An information theory account of cognitive control. *Frontiers in human neuroscience*, *8*, 680.

Gayet, S., Paffen, C. L., Belopolsky, A. V., Theeuwes, J., & Van der Stigchel, S. (2016). Visual input signaling threat gains preferential access to awareness in a breaking continuous flash suppression paradigm. *Cognition*, *149*, 77-83.

Gerdes, A. B., Wieser, M. J., & Alpers, G. W. (2014). Emotional pictures and sounds: a review of multimodal interactions of emotion cues in multiple domains. *Frontiers in psychology*, *5*, 1351.

Ghosh, D., Sengupta, R., Sanyal, S., & Banerjee, A. (2018). *Musicality of human brain through fractal analytics*. Springer Singapore.

Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nature neuroscience*, *8*(2), 145-146.

Guski, R. (1999). Personal and social variables as co-determinants of noise annoyance. *Noise and health*, *1*(3), 45.

Hietanen, J. K., Surakka, V., & Linnankoski, I. (1998). Facial electromyographic responses to vocal affect expressions. *Psychophysiology*, *35*(5), 530-536.

Juslin, P. N., & Sloboda, J. A. (2001). Music and emotion. *Theory and research*.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and brain sciences*, *31*(5), 559-575.

Klasen, M., Kreifelts, B., Chen, Y. H., Seubert, J., & Mathiak, K. (2014). Neural processing of emotion in multimodal settings. *Frontiers in human neuroscience*, *8*, 822.

Krishna, A. (2012). An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. *Journal of consumer psychology*, *22*(3), 332-351.

Kuwano, S., Namba, S., Hayashi, Y., Komatsu, M., & Kato, T. (2001). Auditory and visual interaction in the aesthetic evaluation of environment. *Empirical Studies of the Arts*, *19*(2), 191-200.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological review*, *97*(3), 377.

Levenson, R. W. (1999). The intrapersonal functions of emotion. *Cognition & Emotion*, *13*(5), 481-504.

Ludden, G. D., Schifferstein, H. N., & Hekkert, P. (2009). Visual–tactual incongruities in products as sources of surprise. *Empirical Studies of the Arts*, *27*(1), 61-87.

Mogg, K., Bradley, B. P., De Bono, J. O., & Painter, M. (1997). Time course of attentional bias for threat information in non-clinical anxiety. *Behaviour research and therapy*, *35*(4), 297-303.

Müller, V. I., Cieslik, E. C., Turetsky, B. I., & Eickhoff, S. B. (2012). Crossmodal interactions in audiovisual emotion processing. *Neuroimage*, *60*(1), 553-561.

Niedenthal, P. M., & Wood, A. (2019). Does emotion influence visual perception? Depends on how you look at it. *Cognition and Emotion*, *33*(1), 77-84.

Niedenthal, P. M., Halberstadt, J. B., Margolin, J., & Innes-Ker, Å. H. (2000). Emotional state and the detection of change in facial expression of emotion. *European journal of social psychology*, *30*(2), 211-222.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *Journal of experimental psychology: general*, *130*(3), 466.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning* (No. 47). University of Illinois press.

Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological science*, *17*(4), 292-299.

Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological science*, *17*(4), 292-299.

Riener, C. R., Stefanucci, J. K., Proffitt, D. R., & Clore, G. (2011). An effect of mood on the perception of geographical slant. *Cognition and Emotion*, *25*(1), 174-182.

Roy, S., Roy, C., Nag, S., Banerjee, A., Sengupta, R., & Ghosh, D. (2020). Chaos based non-linear cognitive study of different stimulus in the cross-modal perspective. *Physica A: Statistical Mechanics and its Applications*, *546*, 122842.

Russell, C. A. (2002). Investigating the effectiveness of product placements in television shows: The role of modality and plot connection congruence on brand memory and attitude. *Journal of consumer research*, *29*(3), 306-318.

Sander, D. (2013). Models of emotion: The affective neuroscience approach. In J. Armony & P. Vuilleumier (Eds.), *The Cambridge handbook of human affective neuroscience* (pp. 5–53). Cambridge University Press.

Schafer, R., & Murphy, G. (1943). The role of autism in a visual figure-ground relationship. *Journal of Experimental Psychology*, *32*(4), 335.

Scherer, K. R., & Wallbott, H. G. (1994). Evidence for universality and cultural variation of differential emotion response patterning. *Journal of personality and social psychology*, *66*(2), 310.

Schreuder, E., Van Erp, J., Toet, A., & Kallen, V. L. (2016). Emotional responses to multisensory environmental stimuli: A conceptual framework and literature review. *Sage Open*, *6*(1), 2158244016630591.

Seigneuric, A., Durand, K., Jiang, T., Baudouin, J. Y., & Schaal, B. (2010). The nose tells it to the eyes: crossmodal associations between olfaction and vision. *Perception*, *39*(11), 1541-1554.

Shimojo, S., & Shams, L. (2001). Sensory modalities are not separate modalities: plasticity and interactions. *Current opinion in neurobiology*, *11*(4), 505-509.

Siegel, E. H., Wormwood, J. B., Quigley, K. S., & Barrett, L. F. (2018). Seeing what you feel: Affect drives visual perception of structurally neutral faces. *Psychological science*, *29*(4), 496-503.

Siemer, M., Mauss, I., & Gross, J. J. (2007). Same situation--different emotions: how appraisals shape our emotions. *Emotion*, *7*(3), 592.

Tajadura-Jiménez, A., & Västfjäll, D. (2008). Auditory-induced emotion: A neglected channel for communication in human-computer interaction. In *Affect and Emotion in Human-Computer Interaction* (pp. 63-74). Springer, Berlin, Heidelberg.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, *101*(1), 80-113.

Voss, A., Rothermund, K., & Brandtstädter, J. (2008). Interpreting ambiguous stimuli: Separating perceptual and judgmental biases. *Journal of Experimental Social Psychology*, *44*(4), 1048-1056.

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in cognitive sciences*, *9*(12), 585-594.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron*, *30*(3), 829-841.

Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., & Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature neuroscience*, *7*(11), 1271-1278.

Zadra, J. R., & Clore, G. L. (2011). Emotion and perception: The role of affective information. *Wiley interdisciplinary reviews: cognitive science*, *2*(6), 676-685.

1.5. Color, Music, *Rasa*s

Ballard, L., & Bengtson, M. (2017). *The Alexander Scriabin companion: History, performance, and lore*. Rowman & Littlefield.

Beattie, L. A. (1998). *Color and sound interrelated*. San Jose State University.

Bresin, R. (2005, September). What is the color of that music performance?. In *ICMC*.

Caivano, J. L. (1994). Color and sound: Physical and psychophysical relations. *Color Research & Application*, *19*(2), 126-133.

Cajochen, C., Munch, M., Kobialka, S., Krauchi, K., Steiner, R., Oelhafen, P., ... & Wirz-Justice, A. (2005). High sensitivity of human melatonin, alertness, thermoregulation, and heart rate to short wavelength light. *The journal of clinical endocrinology & metabolism*, *90*(3), 1311-1316.

Campos, J. J., Mumme, D., Kermoian, R., & Campos, R. G. (1994). A functionalist perspective on the nature of emotion. *Japanese Journal of Research on Emotions*, *2*(1), 1-20.

Cheriyan, L. B. (2017). The Psychological and Physiological Relevance of Bharathamuni's Theory of Rasa: A Study on the Basis of Colorgenics.

Clayton, M. (2001). Introduction: Towards a theory of musical meaning (in India and elsewhere). *British Journal of Ethnomusicology*, *10*(1), 1-17.

Collier, W. G., & Hubbard, T. L. (2004). Musical scales and brightness evaluations: Effects of pitch, direction, and scale mode. *Musicae Scientiae*, *8*(2), 151-173.

Davidoff, J. (1991). *Cognition through color*. The MIT Press.

Doczi, G. (1981). *The power of limits: Proportional harmonies in nature, art, and architecture* (p. 150). Boulder, CO: Shambhala.

Ekman, P. (1992). An argument for basic emotions. *Cognition & emotion*, *6*(3-4), 169-200.

Ekman, P., Friesen, W. V., O'sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., ... & Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of personality and social psychology*, *53*(4), 712.

Franssen, M. (1991). The ocular harpsichord of louis-bertrand castel. *Tractrix*, *3*(1991), 15-77.

Hejmadi, A., Davidson, R. J., & Rozin, P. (2000). Exploring Hindu Indian emotion expressions: Evidence for accurate recognition by Americans and Indians. *Psychological Science*, *11*(3), 183-187.

Higgins, K. M. (2007). An alchemy of emotion: Rasa and aesthetic breakthroughs. *The Journal of Aesthetics and Art Criticism*, *65*(1), 43-54.

Isbilen, E. S., & Krumhansl, C. L. (2016). The color of music: Emotion-mediated associations to Bach's Well-tempered Clavier. *Psychomusicology: Music, Mind, and Brain*, *26*(2), 149.

Itten, J. (1970). *The elements of color* (Vol. 4). John Wiley & Sons.

Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion-cognition relations.

Kaufmann, W. (1965). Rasa, rāga-mālā and performance times in North Indian rāgas. *Ethnomusicology*, 272-291.

Kavi, R. (1934). *NATYASHASTRA with Commentary of ABHINAVAGUPTA*. Рипол Классик.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The cognitive unconscious and the embodied mind: How the embodied mind creates philosophy*. Basic Books.

Lindborg, P., & Friberg, A. K. (2015). Colour association with music is mediated by emotion: Evidence from an experiment using a CIE Lab interface and interviews. *PloS one*, *10*(12), e0144013.

Newton, I. (1952). *Opticks, or, a treatise of the reflections, refractions, inflections & colours of light*. Courier Corporation.

Palmer, S. E., Langlois, T. A., & Schloss, K. B. (2016). Music-to-color associations of single-line piano melodies in non-synesthetes. *Multisensory Research*, *29*(1-3), 157-193.

Palmer, S. E., Schloss, K. B., Xu, Z., & Prado-León, L. R. (2013). Music–color associations are mediated by emotion. *Proceedings of the National Academy of Sciences*, *110*(22), 8836-8841.

Polzella, D. J., & Biers, D. W. (1987). Chromesthetic responses to music: Replication and extension. *Perceptual and Motor Skills, 65*(2), 439-443.

Polzella, D. J., & Hassen, J. L. (1997). Aesthetic preferences for combinations of color and music. *Perceptual and motor skills*, *85*(3), 960-962.

Polzella, D. J., Kuna, A. M., Biers, D. W., & DaPolito, F. (1982). Hemispheric asymmetry in musically-induced color imagery. *Psychomusicology: A Journal of Research in Music Cognition*, *2*(2), 64.

Pridmore, R. W. (1992). Music and color: Relations in the psychophysical perspective. *Color Research & Application*, *17*(1), 57-61.

Rachman, S. (1983). Fear and courage among military bomb-disposal operators. A*dvances in Behaviour Research and Therapy*, 4(3), 97-175.

Ramaprasad, D. (2013). Emotions: an Indian perspective. *Indian Journal of Psychiatry*, *55*(Suppl 2), S153.

Roy, S., Roy, C., Nag, S., Banerjee, A., Sengupta, R., & Ghosh, D. (2020). Chaos based non-linear cognitive study of different stimulus in the cross-modal perspective. *Physica A: Statistical Mechanics and its Applications*, *546*, 122842.

Russell, J. A. (1991). In defense of a prototype approach to emotion concepts. *Journal of personality and social psychology*, *60*(1), 37.

Schechner, R. (2003). *Performance theory*. Routledge.

Shaffer, L. F. (1947). Fear and courage in aerial combat. *Journal of consulting psychology*, *11*(3), 137.

Tripathi, R., Mukhopadhyay, D., Singh, C. K., Miyapuram, K. P., & Jolad, S. (2018). Characterizing functional brain networks and emotional centers based on Rasa theory of Indian aesthetics. *arXiv preprint arXiv:1809.05336*.

von Goethe, J. W. (2006). Theory of colours (trans: Eastlake, CL). *Mineola: Dover (Original work published 1810)*.

Walz, R. (2013). Wassily Kandinsky, Concerning the spiritual in art. In *Modernism* (pp. 146-147). Routledge.

Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all?. *Cortex*, *42*(2), 264-280.

Wexner, L. B. (1954). The degree to which colors (hues) are associated with mood-tones. *Journal of applied psychology*, *38*(6), 432.

Whiteford, K. L., Schloss, K. B., Helwig, N. E., & Palmer, S. E. (2018). Color, music, and emotion: Bach to the blues. *i-Perception*, *9*(6), 2041669518808535.

Wieczorkowska, A. A., Datta, A. K., Sengupta, R., Dey, N., & Mukherjee, B. (2010). On search for emotion in Hindusthani vocal music. In *Advances in music information retrieval* (pp. 285-304). Springer, Berlin, Heidelberg.

1.6. A summary on studies regarding Color Perception

Ainsworth, R. A., Simpson, L., & Cassell, D. (1993). Effects of three colors in an office interior on mood and performance. *Perceptual and motor skills*, *76*(1), 235-241.

AL-Ayash, A., Kane, R. T., Smith, D., & Green-Armytage, P. (2016). The influence of color on student emotion, heart rate, and performance in learning environments. *Color Research & Application*, *41*(2), 196-205.

Ali, M. R. (1972). Pattern of EEG recovery under photic stimulation by light of different colors. *Electroencephalography and Clinical Neurophysiology*, *33*(3), 332-335.

Allen, G. (1877). *Physiological aesthetics*. New York: D. Appleton.

Briki, W., & Hue, O. (2016). How red, blue, and green are affectively judged. *Applied Cognitive Psychology*, *30*(2), 301-304.

Brown, B. (1966). Specificity of EEG photic flicker responses to color as related to visual imagery ability. *Psychophysiology*, *2*(3), 197-207.

Buechner, V. L., & Maier, M. A. (2016). Not always a matter of context: direct effects of red on arousal but context-dependent moderations on valence. *PeerJ*, *4*, e2515.

Choi, Y., Kim, M., & Chun, C. (2015). Measurement of occupants' stress based on electroencephalograms (EEG) in twelve combined environments. *Building and Environment*, *88*, 65-72.

Cohn, J. (1894). *Experimentelle Untersuchungen über die Gefühlsbetonung der Farben, Helligkeiten und ihrer Combinationen*. Engelmann.

Deutsch, F. (1937). Psycho-physical Reactions of the Vascular System to Influence of Light and Impressions Gained through Light. *Folia Clinica Orientalia*, *1*(3).

Dreiskaemper, D., Strauss, B., Hagemann, N., & Büsch, D. (2013). Influence of red jersey color on physical parameters in combat sports. *Journal of Sport and Exercise Psychology*, *35*(1), 44-49.

Elliot, A. J. (2019). A historically based review of empirical work on color and psychological functioning: Content, methods, and recommendations for future research. *Review of General Psychology*, *23*(2), 177-200.

Elliot, A. J., & Maier, M. A. (2014). Color psychology: Effects of perceiving color on psychological functioning in humans. *Annual review of psychology*, *65*(1), 95-120.

Elliot, A. J., & Niesta, D. (2008). Romantic red: red enhances men's attraction to women. *Journal of personality and social psychology*, *95*(5), 1150.

Erwin, C. W., Lerner, M., Wilson, N. J., & Wilson, W. P. (1961). Some further observations on the photically elicited arousal response. *Electroencephalography and Clinical Neurophysiology*, *13*(3), 391-394.

Féré, C. (1887). Note sur les conditions physiologiques des émotions [Note on the physiological conditions of emotions]. *Revue Philosophique de la France et de l'Étranger*, *24*, 561-581.

Frank, M. G., & Gilovich, T. (1988). The dark side of self-and social perception: black uniforms and aggression in professional sports. *Journal of personality and social psychology*, *54*(1), 74.

Gerard, K. (1958). Differential effects of colored lights on psycho-physiological functions. *Unpublished doctoral dissertation, University of California, Los Angeles.*

Goldstein, K. (1942). Some experimental observations concerning the influence of colors on the function of the organism. *Occupational Therapy*.

Hill, R. A., & Barton, R. A. (2005). Red enhances human performance in contests. *Nature*, *435*(7040), 293-293.

Jastrow, J. (1897). The popular aesthetics of color. *Popular Science Monthly, 50*, 361–368.

Kaiser, P. K. (1984). Physiological response to color: A critical review. *Color research & application*, *9*(1), 29-36.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, *29*(2-3), 169-195.

Lee, S., & Rao, V. S. (2010). Color and store choice in electronic commerce: The explanatory role of trust. *Journal of Electronic Commerce Research*, *11*(2).

Linhares, J. M. M., Pinto, P. D., & Nascimento, S. M. C. (2008). The number of discernible colors in natural scenes. *JOSA A*, *25*(12), 2918-2924.

Mehta, R., & Zhu, R. (2009). Blue or red? Exploring the effect of color on cognitive task performances. *Science*, *323*(5918), 1226-1229.

Münsterberg, H., Campbell, W. W., Bigham, J., Pierce, A. H., Calkins, M. W., & Pierce, E. (1894). Studies from the Harvard Psychological Laboratory (II). *Psychological Review*, *1*(5), 441.

Nakshian, J. S. (1964). The effects of red and green surroundings on behavior. *The Journal of General Psychology*, *70*(1), 143-161.

Nourse, J. C., & Welch, R. B. (1971). Emotional attributes of color: A comparison of violet and green. *Perceptual and motor skills*, *32*(2), 403-406.

Ou, L. C., Luo, M. R., Woodcock, A., & Wright, A. (2004). A study of colour emotion and colour preference. Part I: Colour emotions for single colours. *Color Research & Application*, *29*(3), 232-240.

Sakuragi, S., & Sugiyama, Y. (2011). Effect of partition board color on mood and autonomic nervous function. *Perceptual and motor skills*, *113*(3), 941-956.

Suk, H. J., & Irtel, H. (2010). Emotional response to color across media. *Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The*

Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur, *35*(1), 64-77.

Vandewalle, G., Schmidt, C., Albouy, G., Sterpenich, V., Darsaud, A., Rauchs, G., ... & Dijk, D. J. (2007). Brain responses to violet, blue, and green monochromatic light exposures in humans: prominent role of blue light and the brainstem. *PloS one*, *2*(11), e1247.

von Goethe, J. W. (2006). Theory of colours (trans: Eastlake, CL). *Mineola: Dover (Original work published 1810)*.

Wilms, L., & Oberfeld, D. (2018). Color and emotion: effects of hue, saturation, and brightness. *Psychological research*, *82*(5), 896-914.

Wilson, G. D. (1966). Arousal properties of red versus green. *Perceptual and motor skills*.

Wundt, W. (1874). *Grundzuge der physiologischen Psychologie*. [Main features of physiological psychology]. Leipzig, Germany: Wilhelm Engelmann.

1.7. Cross modal correspondence: An overview

Albertazzi, L., Canal, L., & Micciolo, R. (2015). Cross-modal associations between materic painting and classical Spanish music. *Frontiers in Psychology*, *6*, 424.

Baier, B., Kleinschmidt, A., & Müller, N. G. (2006). Cross-modal processing in early visual and auditory cortices depends on expected statistical relationship of multisensory information. *Journal of Neuroscience*, *26*(47), 12260-12265.

Barbiere, J. M., Vidal, A., & Zellner, D. A. (2007). The color of music: Correspondence through emotion. *Empirical studies of the arts*, *25*(2), 193-208.

Beauchamp, M. S., Argall, B. D., Bodurka, J., Duyn, J. H., & Martin, A. (2004b). Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nature neuroscience*, *7*(11), 1190-1192.

Beauchamp, M. S., Lee, K. E., Argall, B. D., & Martin, A. (2004a). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron*, *41*(5), 809-823.

Bernstein, I. H., & Edelstein, B. A. (1971). Effects of some variations in auditory input upon visual choice reaction time. *Journal of experimental psychology*, *87*(2), 241.

Bond, B., & Stevens, S. S. (1969). Cross-modality matching of brightness to loudness by 5-year-olds. *Perception & Psychophysics*, *6*(6), 337-339.

Chen, Y. C., & Spence, C. (2010). When hearing the bark helps to identify the dog: Semantically-congruent sounds modulate the identification of masked pictures. *Cognition*, *114*(3), 389-404.

Crisinel, A. S., & Spence, C. (2009). Implicit association between basic tastes and pitch. *Neuroscience letters*, *464*(1), 39-42.

Cutietta, R. A., & Haggerty, K. J. (1987). A comparative study of color association with music at various age levels. *Journal of research in music education*, *35*(2), 78-91.

Doehrmann, O., & Naumer, M. J. (2008). Semantics and the multisensory brain: how meaning modulates processes of audio-visual integration. *Brain research*, *1242*, 136-150.

Ernst, M. O. (2006). A Bayesian View on Multimodal. *Human body perception from the inside out*, 105.

Ernst, M. O., Lange, C., & Newell, F. N. (2007). Multisensory recognition of actively explored objects. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *61*(3), 242.

Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of vision*, *10*(1), 6-6.

Frens, M. A., Van Opstal, A. J., & Van der Willigen, R. F. (1995). Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. *Perception & psychophysics*, *57*(6), 802-816.

Gal, D., Wheeler, S. C., & Shiv, B. (2007). Cross-modal influences on gustatory perception. *Available at SSRN 1030197*.

Gallace, A., & Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Perception & psychophysics*, *68*(7), 1191-1203.

Grassi, M., & Casco, C. (2010). Audiovisual bounce-inducing effect: When sound congruence affects grouping in vision. *Attention, Perception, & Psychophysics*, *72*(2), 378-386.

Helbig, H. B., & Ernst, M. O. (2007). Optimal integration of shape information from vision and touch. *Experimental brain research*, *179*(4), 595-606.

Isbilen, E. S., & Krumhansl, C. L. (2016). The color of music: Emotion-mediated associations to Bach's Well-tempered Clavier. *Psychomusicology: Music, Mind, and Brain*, *26*(2), 149.

Jones, M. R., Johnston, H. M., & Puente, J. (2006). Effects of auditory pattern structure on anticipatory and reactive attending. *Cognitive psychology*, *53*(1), 59-96.

Kemp, S. E., & Gilbert, A. N. (1997). Odor Intensity and Color Lightness Are Correlated Sensory Dimensions. *The American Journal of Psychology*, 35-46.

Köhler, W. (1929). *Gestalt psychology*. New York: Liveright.

Lindborg, P., & Friberg, A. K. (2015). Colour association with music is mediated by emotion: Evidence from an experiment using a CIE Lab interface and interviews. *PloS one*, *10*(12), e0144013.

Marin, M. M., & Bhattacharya, J. (2010). Music induced emotions: some current issues and cross-modal comparisons. *Music education*, 1-38.

Marks, L. E. (1974). On associations of light and sound: The mediation of brightness, pitch, and loudness. *The American journal of psychology*, 173-188.

Marks, L. E. (2004). 6 Cross-Modal Interactions. *The handbook of multisensory processes*, 85.

Martino, G., & Marks, L. E. (2000). Cross-modal interaction between vision and touch: the role of synesthetic correspondence. *Perception*, *29*(6), 745-754.

Nuckolls, J. B. (2003, April). To be or not to be ideophonically impoverished. In *Proceedings of the Eleventh Annual Symposium about Language and Society—Austin*.

Palmer, S. E., Langlois, T. A., & Schloss, K. B. (2016). Music-to-color associations of single-line piano melodies in non-synesthetes. *Multisensory Research*, *29*(1-3), 157-193.

Palmer, S. E., Schloss, K. B., Xu, Z., & Prado-León, L. R. (2013). Music–color associations are mediated by emotion. *Proceedings of the National Academy of Sciences*, *110*(22), 8836-8841.

Parrott, A. C. (1982). Effects of paintings and music, both alone and in combination, on emotional judgments. *Perceptual and Motor Skills*, *54*(2), 635-641.

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia--a window into perception, thought and language. *Journal of consciousness studies*, *8*(12), 3-34.

Ramachandran, V. S., & Hubbard, E. M. (2003). Hearing colors, tasting shapes. *Scientific American*, *288*(5), 52-59.

Ramachandran, V. S., & Hubbard, E. M. (2003). Hearing colors, tasting shapes. *Scientific American*, *288*(5), 52-59.

Ramachandran, V. S., & Oberman, L. M. (2006). Broken mirrors. *Scientific American*, *295*(5), 62-69.

Roffler, S. K., & Butler, R. A. (1968). Factors that influence the localization of sound in the vertical plane. *The Journal of the Acoustical Society of America*, *43*(6), 1255-1259.

Sapir, E. (1929). A study in phonetic symbolism. *Journal of experimental psychology, 12*(3), 225.

Seo, H. S., Arshamian, A., Schemmer, K., Scheer, I., Sander, T., Ritter, G., & Hummel, T. (2010). Cross-modal integration between odors and abstract symbols. *Neuroscience letters, 478*(3), 175-178.

Shore, D. I., Barnes, M. E., & Spence, C. (2006). Temporal aspects of the visuotactile congruency effect. *Neuroscience letters*, *392*(1-2), 96-100.

Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neuroreport*, *12*(1), 7-10.

Smith, L. B., & Sera, M. D. (1992). A developmental analysis of the polar structure of dimensions. *Cognitive psychology*, *24*(1), 99-142.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*(4), 971-995.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*(4), 971-995.

Spence, C. (2019). On the relative nature of (pitch-based) crossmodal correspondences. *Multisensory Research*, *32*(3), 235-265.

Vatakis, A., Ghazanfar, A. A., & Spence, C. (2008). Facilitation of multisensory integration by the "unity effect" reveals that speech is special. *Journal of Vision*, *8*(9), 14-14.

Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science*, *21*(1), 21-25.

Whiteford, K. L., Schloss, K. B., Helwig, N. E., & Palmer, S. E. (2018). Color, music, and emotion: Bach to the blues. *i-Perception*, *9*(6), 2041669518808535.

Wicker, F. W. (1968). Mapping the intersensory regions of perceptual space. *The American Journal of Psychology*, *81*(2), 178-188.

Zangenehpour, S., & Zatorre, R. J. (2010). Crossmodal recruitment of primary visual cortex following brief exposure to bimodal audiovisual stimuli. *Neuropsychologia*, *48*(2), 591-600.

1.8. Studying complex systems: Chaos theory and Butterfly effect

Gleick, J., & Berry, M. (1987). Chaos-making a new science. *Nature*, *330*, 293.

Ilya, P., & Stengers, I. (1984). Order out of chaos: man's new dialogue with nature. *New York, Bantam*.

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of atmospheric sciences*, *20*(2), 130-141.

Pool, R. (1989). Is It Healthy to Be Chaotic? Chaos may provide a healthy flexibility to the heart, brain, and other parts of the body. Conversely, many ailments may be associated with a loss of this chaotic flexibility. *Science*, *243*(4891), 604-607.

1.9. Studying complex systems: Fractals and Multifractals

Ahmadlou, M., & Adeli, H. (2011). Functional community analysis of brain: A new approach for EEG-based investigation of the brain pathology. *Neuroimage*, *58*(2), 401-408.

Ahmadlou, M., Adeli, H., & Adeli, A. (2010). New diagnostic EEG markers of the Alzheimer's disease using visibility graph. *Journal of neural transmission*, *117*(9), 1099-1109.

Ahmadlou, M., Adeli, H., & Adeli, A. (2012). Fractality analysis of frontal brain in major depressive disorder. *International Journal of Psychophysiology*, *85*(2), 206-211.

Banerjee, A., Sanyal, S., Patranabis, A., Banerjee, K., Guhathakurta, T., Sengupta, R., ... & Ghose, P. (2016). Study on brain dynamics by non linear analysis of music induced EEG signals. *Physica A: Statistical Mechanics and its Applications*, *444*, 110-120.

Bigerelle, M., & Iost, A. (2000). Fractal dimension and classification of music. *Chaos, Solitons & Fractals*, *11*(14), 2179-2192.

Gao, T. T., Wu, D., Huang, Y. L., & Yao, D. Z. (2007). Detrended fluctuation analysis of the human EEG during listening to emotional music. *Journal of Electronic Science and Technology*, *5*(3), 272-277.

Hsü, K. J., & Hsü, A. J. (1990). Fractal geometry of music. *Proceedings of the National Academy of Sciences*, *87*(3), 938-941.

Hwa, R. C., & Ferree, T. C. (2002). Scaling properties of fluctuations in the human electroencephalogram. *Physical Review E*, *66*(2), 021901.

Kiselev, V. G., Hahn, K. R., & Auer, D. P. (2003). Is the brain cortex a fractal?. *Neuroimage*, *20*(3), 1765-1774.

Kumar, A., & Mullick, S. K. (1996). Nonlinear dynamical analysis of speech. *The Journal of the Acoustical Society of America*, *100*(1), 615-629.

Lee, J. M., Kim, D. J., Kim, I. Y., Park, K. S., & Kim, S. I. (2002). Detrended fluctuation analysis of EEG in sleep apnea using MIT/BIH polysomnography data. *Computers in biology and medicine*, *32*(1), 37-47.

Lee, M. (2015). *György Ligeti's" Désordre" and" L'escalier du Diable": Compositional style and method of practice*. The Ohio State University.

Ligeti, G. (2003). *Etudes, books I and II*. Naxos.

Linkenkaer-Hansen, K., Nikouline, V. V., Palva, J. M., & Ilmoniemi, R. J. (2001). Long-range temporal correlations and scaling behavior in human brain oscillations. *Journal of Neuroscience*, *21*(4), 1370-1377.

Lochhead, J. (2001). Hearing Chaos. *American Music*, 210-246.

Lopes, R., & Betrouni, N. (2009). Fractal and multifractal analysis: a review. *Medical image analysis*, *13*(4), 634-649.

Maity, A. K., Pratihar, R., Mitra, A., Dey, S., Agrawal, V., Sanyal, S., ... & Ghosh, D. (2015). Multifractal detrended fluctuation analysis of alpha and theta EEG rhythms with musical stimuli. *Chaos, Solitons & Fractals*, *81*, 52-67.

Mandelbrot, B. (1967). How long is the coast of Britain? Statistical self-similarity and fractional dimension. *science*, *156*(3775), 636-638.

Mölle, M., Marshall, L., Lutzenberger, W., Pietrowsky, R., Fehm, H. L., & Born, J. (1996). Enhanced dynamic complexity in the human EEG during creative thinking. *Neuroscience Letters*, *208*(1), 61-64.

Natarajan, K., Acharya U, R., Alias, F., Tiboleng, T., & Puthusserypady, S. K. (2004). Nonlinear analysis of EEG signals at different mental states. *Biomedical engineering online*, *3*(1), 1-11.

Pressing, J. (1988). Improvisation: methods and models. *Generative processes in music*, 129-178.

Salter, J. R. (2009). *Chaos in music: Historical developments and applications to music theory and composition*. The University of North Carolina at Greensboro.

Sengupta, R., Dey, N., Datta, A. K., Ghosh, D., & Patranabis, A. (2010). Analysis of the signal complexity in sitar performances. *Fractals*, *18*(02), 265-270.

Su, Z. Y., & Wu, T. (2006). Multifractal analyses of music sequences. *Physica D: Nonlinear Phenomena*, *221*(2), 188-194.

Tong, S., Huang, H., Luan, J., & Chen, H. (2005). Dimensional complexity analysis of human EEG in visual processing. *Neurocomputing*, *67*, 297-305.

Voss, R. F., & Clarke, J. (1975). '1/f noise in music and speech. *Nature*, *258*(5533), 317-318.

Weiss, B., Clemens, Z., Bódizs, R., & Halász, P. (2011). Comparison of fractal and power spectral EEG features: Effects of topography and sleep stages. *Brain research bulletin*, *84*(6), 359-375.

1.10. Studying complex systems: Power laws

Adamic, L. A. (2000). Zipf, power-laws, and pareto-a ranking tutorial. *Xerox Palo Alto Research Center, Palo Alto, CA, http://ginger. hpl. hp. com/shl/papers/ranking/ranking. html*.

Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the 1/f noise. *Physical review letters*, *59*(4), 381.

Brothers, H. J. (2007). Structural scaling in Bach's Cello Suite no. 3. *Fractals*, *15*(01), 89-95.

Chang, M. C., Yang, A. C. C., Stanley, H. E., & Peng, C. K. (2017). Measuring information-based energy and temperature of literary texts. *Physica A: Statistical Mechanics and its Applications*, *468*, 783-789.

Cox, R. A., Felton, J. M., & Chung, K. H. (1995). The concentration of commercial success in popular music: an analysis of the distribution of gold records. *Journal of cultural economics*, *19*(4), 333-340.

Cramer, I. (2005). The parameters of the Altmann-Menzerath law. *Journal of Quantitative Linguistics*, *12*(1), 41-52.

Crovella, M. E., & Bestavros, A. (1997). Self-similarity in World Wide Web traffic: Evidence and possible causes. *IEEE/ACM Transactions on networking*, *5*(6), 835-846.

De Campos, H., & Tolman, J. M. (1982). The informational temperature of the text. *Poetics Today*, *3*(3), 177-187.

De Souza, N. M., Vanderlei, L. C. M., & Garner, D. M. (2015). Risk evaluation of diabetes mellitus by relation of chaotic globals to HRV. *Complexity*, *20*(3), 84-92.

Gabaix, X. (1999). Zipf's law for cities: an explanation. *The Quarterly journal of economics*, *114*(3), 739-767.

Gutenberg, B., & Richter, C. F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological society of America*, *34*(4), 185-188.

i Cancho, R. F. (2010). Network theory. In *The Cambridge encyclopedia of the language sciences* (pp. 555-557). Cambridge University Press.

i Cancho, R. F., & Solé, R. V. (2003). Statistical mechanics of complex networks. *Lecture notes in Physics*, *625*, 114-125.

Jackendoff, R. (2009). Parallels and nonparallels between language and music. *Music perception*, *26*(3), 195-204.

Kello, C. T., Brown, G. D., Ferrer-i-Cancho, R., Holden, J. G., Linkenkaer-Hansen, K., Rhodes, T., & Van Orden, G. C. (2010). Scaling laws in cognitive sciences. *Trends in cognitive sciences*, *14*(5), 223-232.

Kim, S. (2006). Player's positional dependence of fractal behaviors in a soccer game. *Fractals*, *14*(01), 71-76.

Kosmidis, K., Kalampokis, A., & Argyrakis, P. (2006). Language time series analysis. *Physica A: Statistical Mechanics and its Applications*, *370*(2), 808-816.

Li, W. (2002). Zipf's Law everywhere. *Glottometrics*, *5*(2002), 14-21.

Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington academy of sciences*, *16*(12), 317-323.

Lu, E. T., & Hamilton, R. J. (1991). Avalanches and the distribution of solar flares. *The astrophysical journal*, *380*, L89-L92.

Mandelbrot, B. (1953). An informational theory of the statistical structure of language. *Communication theory*, *84*, 486-502.

Mandelbrot, B. (1965). Self-similar error clusters in communication systems and the concept of conditional stationarity. *IEEE Transactions on Communication Technology*, *13*(1), 71-90.

Mitzenmacher, M. (2004). A brief history of generative models for power law and lognormal distributions. *Internet mathematics*, *1*(2), 226-251.

Miyazima, S., & Yamamoto, K. (2008). Measuring the temperature of texts. *Fractals*, *16*(01), 25-32.

Newman, M. E. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary physics*, *46*(5), 323-351.

Newman, M. E. (2011). Complex systems: A survey. *arXiv preprint arXiv:1112.1440*.

Pareto, V. (1964). *Cours d'économie politique* (Vol. 1). Librairie Droz.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature neuroscience*, *6*(7), 674-681.

Patel, A. D. (2008). *Music, language, and the brain*. Oxford University Press.

Reˆgo, H. H. A., Braunstein, L. A., D′ Agostino, G., Stanley, H. E., & Miyazima, S. (2014). When a text is translated does the complexity of its vocabulary change? Translations and target readerships. *PloS one*, *9*(10), e110213.

Rovenchak, A. (2014). Trends in language evolution found from the frequency structure of texts mapped against the Bose-distribution. *Journal of Quantitative Linguistics*, *21*(3), 281-294.

Rovenchak, A., & Buk, S. (2011). Application of a quantum ensemble model to linguistic analysis. *Physica A: Statistical Mechanics and its Applications*, *390*(7), 1326-1331.

Sengupta, R., Dey, N., Datta, A. K., & Ghosh, D. (2005). Assessment of musical quality of tanpura by fractal-dimensional analysis. *Fractals*, *13*(03), 245-252.

Simon, H. A. (1955). On a class of skew distribution functions. *Biometrika*, *42*(3/4), 425-440.

Willis, J. C., & Yule, G. U. (1922). Some statistics of evolution and geographical distribution in plants and animals, and their significance. *Nature*, *109*(2728), 177-179.

Zanette, D. H., & Manrubia, S. C. (2001). Vertical transmission of culture and the distribution of family names. Physica A: Statistical Mechanics and its Applications, 295(1-2), 1-8.

Zipf, G. K. (1949). Human behavior and the principle of least effort: an introduction to human ecology.

1.11. Emotions, Ambiguity and Creative aesthetics

Bhushan, L. I., & Amal, S. B. (1986). A situational test of intolerance of ambiguity. *Psychologia: An International Journal of Psychology in the Orient*.

Blanchette, I., & Richards, A. (2010). The influence of affect on higher level cognition: A review of research on interpretation, judgment, decision making and reasoning. *Cognition & Emotion*, *24*(4), 561-595.

Bode, C. (1991). The aesthetics of ambiguity.

Brainerd, C. J. (2018). The emotional-ambiguity hypothesis: A large-scale test. *Psychological science*, *29*(10), 1706-1715.

Comadena, M. E. (1984). Brainstorming groups: Ambiguity tolerance, communication apprehension, task attraction, and individual productivity. *Small Group Behavior*, *15*(2), 251-264.

Cowen, A. S., Fang, X., Sauter, D., & Keltner, D. (2020). What music makes us feel: At least 13 dimensions organize subjective experiences associated with music across different cultures. *Proceedings of the National Academy of Sciences*, *117*(4), 1924-1934.

Frenkel-Brunswick, E. (1949). Intolerance of ambiguity as emotional and perceptual personality disposition. *Journal of Personality*, *18*, 103-143.

Frenkel-Brunswick, E. (1951). Personality theory and perception. Perception. An approach to personality Blake, R., Ramsay, C.

Furnham, A., & Marks, J. (2013). Tolerance of ambiguity: A review of the recent literature. *Psychology*, *4*(09), 717-728.

Furnham, A., & Ribchester, T. (1995). Tolerance of ambiguity: A review of the concept, its measurement and applications. *Current psychology*, *14*(3), 179-199.

Kensinger, E. A. (2004). Remembering emotional experiences: The contribution of valence and arousal. *Reviews in the Neurosciences*, *15*(4), 241-252.

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological bulletin*, *139*(4), 917.

Lang, P. J., Bradley, M. M., Fitzsimmons, J. R., Cuthbert, B. N., Scott, J. D., Moulder, B., & Nangia, V. (1998). Emotional arousal and activation of the visual cortex: an fMRI analysis. *Psychophysiology*, *35*(2), 199-210.

Liu, C. (2015). Relevant researches on tolerance of ambiguity. *Theory and Practice in Language Studies*, *5*(9), 1874.

Mattek, A. M., Wolford, G. L., & Whalen, P. J. (2017). A mathematical model captures the structure of subjective affect. *Perspectives on Psychological Science*, *12*(3), 508-526.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161.

Sternberg, R. J., & Lubart, T. I. (1995). *Defying the crowd: Cultivating creativity in a culture of conformity*. Free press.

Stoycheva, K. (1998). Ambiguity Tolerance: Adolescents' Responses to Uncertainty in Life. Research Report, September 1996-December 1997.

Stoycheva, K. (2003). Talent, science and education: How do we cope with uncertainty and ambiguities?

Tegano, D. W. (1990). Relationship of tolerance of ambiguity and playfulness to creativity. *Psychological reports*, *66*(3), 1047-1056.

Urban, K. K. (2003). Toward a componential model of creativity,[w:]: Creative Intelligence. Toward Theoretic Integration,(red.) D. Ambros, LM Cohen, AJ Tennenbaum.

Vernon, P. E. (1970). Creativity. *Modern Psychology Readings. First edn. Baltimore: Penguin*.

Yik, M., Russell, J. A., & Steiger, J. H. (2011). A 12-point circumplex structure of core affect. *Emotion*, *11*(4), 705.

Zenasni, F., Besançon, M., & Lubart, T. (2008). Creativity and tolerance of ambiguity: An empirical study. *The Journal of Creative Behavior*, *42*(1), 61-73.

1.12. Quantum approaches to Mind and Brain

Aerts, D., & Aerts, S. (1995). Applications of quantum statistics in psychological studies of decision processes. *Foundations of Science*, *1*(1), 85-97.

Aerts, D., Durt, T., Grib, A. A., Van Bogaert, B., & Zapatrin, R. R. (1993). Quantum structures in macroscopic reality. *International Journal of Theoretical Physics*, *32*(3), 489-498.

Atmanspacher, H. (2017). Quantum approaches to brain and mind. *The Blackwell companion to consciousness*, 298-313.

Atmanspacher, H., & Filk, T. (2010). A proposed test of temporal nonlocality in bistable perception. *Journal of Mathematical Psychology*, *54*(3), 314-321.

Atmanspacher, H., & Filk, T. (2013). The Necker–Zeno model for bistable perception. *Topics in cognitive science*, *5*(4), 800-817.

Atmanspacher, H., Filk, T., & Foundation, P. (2006). Complexity and non-commutativity of learning operations on graphs. *BioSystems*, *85*(1), 84-93.

Atmanspacher, H., Römer, H., & Walach, H. (2002). Weak quantum theory: Complementarity and entanglement in physics and beyond. *Foundations of physics*, *32*(3), 379-406.

Bruza, P. D., Wang, Z., & Busemeyer, J. R. (2015). Quantum cognition: a new theoretical approach to psychology. *Trends in cognitive sciences*, *19*(7), 383-393.

Busemeyer, J. R., Wang, Z., & Townsend, J. T. (2006). Quantum dynamics of human decision-making. *Journal of Mathematical Psychology*, *50*(3), 220-241.

Gabora, L., & Aerts, D. (2002). Contextualizing concepts using a mathematical generalization of the quantum formalism. *Journal of Experimental & Theoretical Artificial Intelligence*, *14*(4), 327-358.

Hameroff, S., & Penrose, R. (1996). Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness. *Mathematics and computers in simulation*, *40*(3-4), 453-480.

Hartmann, L., Dür, W., & Briegel, H. J. (2006). Steady-state entanglement in open and noisy quantum systems. *Physical Review A*, *74*(5), 052304.

Huelga, S. F., & Plenio, M. B. (2013). Vibrations, quanta and biology. *Contemporary Physics*, *54*(4), 181-207.

Khrennivov, A. (1999). Classical and quantum mechanics on information spaces with applications to cognitive, psychological, social, and anomalous phenomena. *Foundations of Physics*, *29*(7), 1065-1098.

London, F., & Bauer, E. (1939). La theorie de l'observation en mecanique quantique. Hermann, Paris. English translation in Wheeler, JA and Zurek, WH eds. 1983. Quantum Theory and Measurement.

Manousakis, E. (2009). Quantum formalism to describe binocular rivalry. *Biosystems*, *98*(2), 57-66.

Penrose, R. (1994). *Shadows of the Mind* (Vol. 4). Oxford: Oxford University Press.

Pothos, E. M., & Busemeyer, J. R. (2009). A quantum probability explanation for violations of 'rational'decision theory. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1665), 2171-2178.

Pothos, E. M., Busemeyer, J. R., & Trueblood, J. S. (2013). A quantum geometric model of similarity. *Psychological Review*, *120*(3), 679.

Ricciardi, L. M., & Umezawa, H. (1967). Brain and physics of many-body problems. *Kybernetik*, *4*(2), 44-48.

Sahu, S., Ghosh, S., Hirata, K., Fujita, D., & Bandyopadhyay, A. (2013). Multi-level memory-switching properties of a single brain microtubule. *Applied Physics Letters*, *102*(12), 123701.

Stapp, H. P. (1999). Attention, intention, and will in quantum physics. *Journal of Consciousness studies*, *6*(8-9), 143-143.

Stapp, H. P. (2004). Mind, matter, and quantum mechanics. In *Mind, matter and quantum mechanics* (pp. 81-118). Springer, Berlin, Heidelberg.

Stapp, H. P. (2015). A quantum-mechanical theory of the mind-brain connection. *Beyond Physicalism*, 157-193.

Tegmark, M. (2000). Importance of quantum decoherence in brain processes. *Physical review E*, *61*(4), 4194.

Vitiello, G. (1995). Dissipation and memory capacity in the quantum brain model. *International Journal of Modern Physics B*, *9*(08), 973-989.

Vitiello, G. (2012). Fractals, coherent states and self-similarity induced noncommutative geometry. *Physics Letters A*, *376*(37), 2527-2532.

Von Neumann, J. (2018). *Mathematical foundations of quantum mechanics*. Princeton university press.

Wigner, E. P. (1967). Explaining Consciousness: Of Molecules and Men. Francis Crick. University of Washington Press, Seattle, 1966. 115 pp. $3.95. John Danz Lecture Series. *Science*, *156*(3776), 798-799.

CHAPTER 2

RESEARCH METHODOLOGY

"The scientific method ... is nothing but the exclusion of subjective opinions as far as possible, by the devising of experiments where observation can give objective answers, yes or no, to questions whether events are causally connected."

Gavin de Beer

ABSTRACT

This Chapter deals with the various techniques associated with the analysis of the structures of music signals as well as bio-signals obtained from EEG data. A detailed analysis on the following tools of complex data analysis have been presented here which is used later in a number of different studies:

i) Empirical mode decomposition (EMD)
ii) Detrended fluctuation analysis (DFA).
iii) Multifractal detrended fluctuation analysis (MFDFA)
iv) Multifractal cross correlation analysis (MFDXA)
v) Maxwell-Boltzmann (MB) analysis
vi) Bose-Einstein (BE) analysis

Four of these techniques (i-iv) make use of Fractal Dimension (FD) or multifractal spectral width (obtained as an output of the MFDFA technique) as an important parameter with which the emotional arousal corresponding to certain cognitive task (visual or auditory) can be quantified. EMD is a decomposition method for non-stationary and nonlinear signals such as EEG. Detrended Fluctuation Analysis (DFA) is used to analyze the long-range temporal correlations (LRTC) of the observed fluctuations in the signal. Advancing from DFA, MFDFA is a fractal analysis technique which is highly scale dependent. Here, signal components of different scaling ratios are analysed together in the form of a multifractal spectral width which indicates the degree of complexity of the signal. Lastly, MFDXA is an important tool with which the degree of cross correlation between two non-linear EEG signals originating from different lobes of brain can be quantitatively measured during higher order cognitive tasks. With this, we can have a quantitative assessment of how the different lobes is cross-correlated during higher order thinking tasks and the perception of external stimulus. A higher degree of cross-correlation would imply similarity between the signals in certain aspects. This in turn can be used to obtain a cue for the informational connectivity human brain displays while processing real world stimuli.

The rest of the two methods discussed here (v-vi) rely on the fact that the distribution of basic units among the complex signal under observation follows power law distribution pattern. Analogies of the complex acoustic signal with ensembles of distinguished (for MB) and indistinguished (for BE) gas particles in a container enables us to use the physical properties of MB and BE distributions in describing their dynamics. This approach is novel in the field of acoustic stimuli, especially in Indian Classical Music scenario. The emergent parameters could prove to be of great importance in categorization and classification of structured data regarding acoustic and bio signals.

Keywords: EMD, DFA, MFDFA, MFDXA, Maxwell-Boltzmann, Bose-Einstein

2.1. INTRODUCTION

Research in the human neurophysiological and cognitive domains have gone through a technological revolution in the last few decades. Biological signals, one of the major gateways to the treasure hunt that is human physiology, exhibit both ordered as well as disordered behaviour. They are complex, often irregular and devoid of discernible patterns to the general observation. Not only bio-signals, such chaos can be found everywhere in nature – from music to visual arts to natural phenomena like weather patterns to even finance (Mantegna & Stanley, 1999) and culture (Gao et al., 2012). Edward Lorentz, in a seminal paper in 1963, first introduced deterministic chaos and explored the possibility of weather prediction (Lorentz, 1963). He saw that the nonperiodic solutions to the nonlinear differential equations designed to represent dissipative nature of a complex system are highly unstable and even the slightest change in initial conditions could result in considerably different outcomes. Thus, the Chaos theory was born. It studied how systems that follow simple, straightforward, deterministic laws can exhibit very complicated and seemingly random long-term behaviour. Soon, Chaos theory was found to be ubiquitous across the fields like DNA sequence, earthquake, turbulent water flows, weather changes, financial time series, faults in bone etc. and it led the researches on the mechanisms describing such complex unpredictable systems. But, to call an observed system chaotic, some strict criteria need to be met. These include the criteria of determinism, stationarity and exponential divergence (Kantz & Schreiber, 2004; Gao et al., 2007). Also, the signal needs to be relatively noise-free. Such conditions make application of Chaos theory in experimental bio-signals complicated because often such signals tend to be noisy and non-stationary. As a result, the countertheory that began to emerge is: random fractal theory, championed by Benoit Mandelbrot (1982). This theory presupposes that the signal has intrinsic randomness - unlike Chaos theory which deals with irregular behaviours generated by nonlinear deterministic interactions with only a few degrees of freedom but excludes noise or randomness majorly. The epicentre of random fractal theory is, of course, the fractals (and fractal dimension). In the words of Mandelbrot, a fractal is defined as: "a rough or fragmented geometric shape that can be subdivided in parts, each of which is (at least approximately) a reduced/size copy of the whole" (Mandelbrot & Mandelbrot, 1982). Quantitative categorization of complex structures could be done using fractal dimensions, a dimension that corresponds in a unique fashion to the geometrical shape under study and that often is not an integer (Stanley & Meakin, 1988). The use of fractal dimension has opened a whole new plethora of studies dealing with complex dynamics in a broad range of natural phenomena. This chapter is essentially a detailed description of the different algorithms used in various sections of this thesis. First, we compare conventional Fourier decomposition methods with the non-linear techniques and then continue with specific non-linear analysis methods which were used on various EEG and music signal data later on.

Fourier decomposition is used often in EEG analysis; a method whose main purpose is to decompose the original signal from time domain to frequency domain. Such Fourier decomposition methods are linear, i.e., Fourier Transform of a sum of functions, is the sum of the Fourier Transforms of the functions. But EEG signals (and most of the signals occurring in nature) are essentially complex and rarely linear. The scalp EEG arises from a large number of neurons interacting with the neighbouring as well as remote neurons via electrical pulses. In such a complex neuronal network, nonlinearity stems from the cellular level, because threshold, integration and saturation phenomena control the dynamics of every individual neuron.

Moreover, the hypothesis of an entirely stochastic brain can be rejected due to its ability to perform sophisticated cognitive tasks (Kannathal et al., 2005). Hence, emerging dynamic patterns are ought to be nonlinear and nonstationary and thus they can generate fluctuations that are not best described by linear decomposition. To analyse these signals, it would be more suitable to use techniques that are robust against inherent non-stationarities. Classical methods of signal analysis, though work well on stationary signals, are extremely limited in dealing with bio-signals in an adequate manner. Fast Fourier Transform (FFT), Wavelet Transform (WT), Matching Pursuit (MP) are some of the examples. The results obtained could very well be misleading. For example, an observed signal having frequency of 12 Hz (with amplitude modulation of 1 Hz frequency) will decompose into two harmonic components having frequencies 11 and 13 Hz after Fourier decomposition, leading to complete disappearance of the original 12 Hz signal (Klonowski, 2009).

This highlights the need for a new solution and new methods of analysis. Non-linear dynamical analysis has emerged as a novel method for the study of complex systems in the past few decades. The non-linear analysis method is effectively applied to electroencephalogram (EEG) data to study the dynamics of the complex underlying behaviour. The idea of ‘fractal dimension’s being a measure of the signal complexity introduces newer horizons which were unattainable using classical techniques. Calculated directly in the time domain using a moving window, the fractal analysis could be applied to signals that are nonstationary, noisy, stochastic or even deterministic. Principles of non-linear dynamics (also, elements of deterministic chaos) govern this approach and involves the characterization of the system attractors with its invariant parameters. For these reasons, EEG data appears to be an appropriate area for the application of nonlinear time series analysis techniques.

Nonlinear dynamics determine the scaling exponent of the signal which indicates the presence or absence of fractal properties (self-similarity). The fractal dimension, or FD, of a waveform represents a powerful tool for transient detection. It is recognised that many systems with apparent random structure possess something called a scale symmetry, implying that the structures remain same under different scales of observation. However, in recent years, quite a number of natural phenomena has been studied which require multiple scaling exponents in describing them. This feature has been used in the analysis of ECG and EEG to identify and distinguish specific states of physiological function. The fractal tool could be described as a mathematical microscope zooming its way into the inherent complex patterns of the signal and helping to decipher complex scaling exponents from the apparent random pattern. Many robust algorithms and methodological approaches are available to examine the fractality present in the signal such as - Correlation dimension, Lyapunov exponent, Box counting method etc. These are very sensitive to noise and require the stationary condition while EEG signals are highly non stationary. To tackle this issue, a nonlinear method named Detrended Fluctuation Analysis (DFA), followed by its advanced form Multifractal Detrended Fluctuation Analysis (MFDFA) has been developed that has the ability to capture scale varying nature of different naturally occurring time-series signals. To further elucidate how the internal dynamics of one signal affects the other, i.e., what is the degree of cross-correlation among the two, we take the help of Multifractal Detrended Cross Correlation analysis (MFDXA) which produces the cross-correlation coefficient as an indicator of correlation.

Also, in the context of stimulus cognition the scientific analysis of the source also assumes importance as the perception of any external stimulus, especially visual and auditory, has a heavy emotional aspect attached to it. The structural complexity of such a stimulus obviously impacts the arousal due to it. Hence, it is important to find structural cues that are associated with universal or individual emotional experience and influences the cognition process. Now, coming to music cognition, music signals are said to possess a chaotic but self-similar structure in different scales and using nonlinear methods are proven to be relevant (Hsu & Hsu, 1990; Bigerelle & Iost, 2000; Sengupta et al., 2005). They are far more complex in their dynamics as compared to the EEG signals due to the superposition of a large number of frequency components. At every instant, various musical segments (in micro and macro scale: pitch, timbre, accent, duration, phrase, melody etc.) are closely linked to each other (Di Lorenzo, 2003). Therefore, studying music using conventional linear frameworks like power spectral analysis seems not to be useful. Instead, fractal analysis of the signal, which reveals the embedded geometry, assumes more significance. In case of acoustic/music signals, MFDFA technique is more accurate than the DFA because of the fact that they include segments with extremely large variation mixed with segments of very small variations, indicating their multifractal nature. Therefore, the normal distribution considering second order RMS variation cannot be applied and all the q-order moments need to be considered. The music signals were initially processed with music analysis software Wavesurfer (Sjölander & Beskow, 2000) and Cool Edit (Johnston, 1999) before being used for data acquisition. The post-hoc statistics of the results include ANOVA and Tukey Kramer multiple comparison tests which have proven to be significant so as to deliver reasonable conclusion.

In the next sections, we present algorithms of the various techniques utilized later in this thesis, some of which have been used widely while some are novel and specific to the studies conducted. Firstly, working with any signals require de-contamination of external artifacts. Bio-signals like EEG in its raw form contains high contaminations due to eye blinks, muscular movement etc., mainly characterized by frequency of less than 4Hz and high amplitude. For analysis, it is essential to get an EEG data free from these artifacts which may induce considerable error in the final results. A novel data-driven noise removal technique called Empirical Mode Decomposition (EMD) is used in our study which helps in generation of noise/artifact-free EEG data in a few steps.

2.2. EMPIRICAL MODE DECOMPOSITION (EMD)

EMD is a decomposition method for non-stationary and nonlinear signals (Huang et.al., 1998). The EMD technique decomposes a signal into a number of intrinsic mode functions (IMFs) that represent fast to slow oscillations. An IMF is a function that satisfies two conditions:

(1) the number of extrema and the number of zero crossings must either be equal or differ by at most one; and (2) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. To obtain an IMF from the original signal x, a sifting process is performed (Huang et.al., 1998) as follows:

First, all extrema of the original signal x need to be identified. All local maximum points are connected by a cubic spline line to form the upper envelope e_u. All local minima points are connected likewise to form the lower envelope e_l. The mean of e_u and e_l, a_1, is calculated as:

$$a_1 = \frac{(e_u + e_l)}{2} \tag{1}$$

The difference between the original signal and the mean is defined as the first component h_1:

$$h_1 = x - a_1 \tag{2}$$

In the next sifting process, h_1 is treated as the signal, and the mean a_{11} of its local maxima and local minima is found. Thus, we have:

$$h_{11} = h_1 - a_{11} \tag{3}$$

Subsequently, we can repeat this sifting procedure k times until h_{1k} is an IMF, with:

$$h_{1k} = h_{1(k-1)} - a_{1k} \tag{4}$$

Therefore, the first IMF component derived from the original signal is designated as:

$$c_1 = h_{1k} \tag{5}$$

The sifting process has been stopped when an IMF has been established by limiting the size of the standard deviation (SD), calculated from the two consecutive sifting sequences as below:

$$SD = \sum_{t=0}^{T} \frac{[h_{1(k-1)}(t) - h_{1k}(t)]^2}{h^2_{1(k-1)}(t)} \tag{6}$$

A typical value for SD can be set between 0.2 and 0.3 (Huang et.al., 1998). In our case the value was set to 0.25. To extract the 2nd IMF component, we remove c_1 from the original signal x:

$$r_1 = x - c_1 \tag{7}$$

The residual r_1 is treated as a new signal, and the same sifting process is applied to obtain the 2nd IMF component c_2 and the residual:

$$r_2 = r_1 - c_2 \tag{8}$$

This procedure is repeated on the subsequent residuals r_j's, until the final residual r_j no longer contains any oscillation information,

$$r_j = r_{j-1} - c_j \tag{9}$$

By summing up Equations (7)–(9), we can obtain:

$$x = \sum_{j=0}^{J} c_j + r_j \tag{10}$$

Thus, original signal x is decomposed into J empirical modes c_j's and a residue r_j.

Since, the artifacts lie in the low frequency regions (<3.5 Hz) (Bizopoulos et. al., 2013; Jung & Saikiran, 2016), the IMFs that appear in this band are rejected. Thus, the filtered signal is the sum of the remaining IMFs and more specifically, only the first few IMFs including the residue were kept (Bizopoulos et. al., 2013). We have obtained noise free EEG data for all the electrodes using the EMD technique and used this data for further analysis and classification of acoustic stimuli induced EEG features.

Fig. 2.2.1 (a-k) shows a representative figure of the F3 electrode in 10 second duration which was subjected to EMD technique and the noise-free EEG data. The sifting process was continued until the final residue is a constant, a monotonic function, i.e., a function with only one maxima or minima from which no more IMF's can be derived. The value of SD was set to be 0.25 after which the sifting process has been stopped.

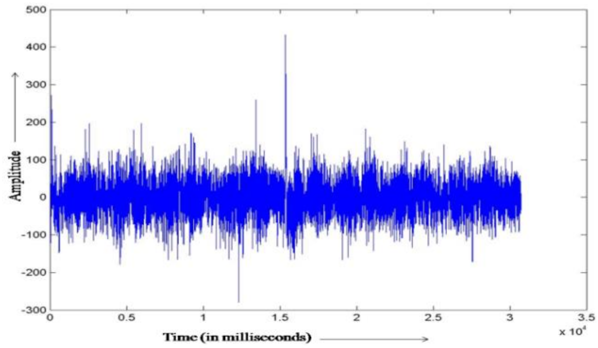


Fig. 2.2.1 (a): Raw EEG Signal

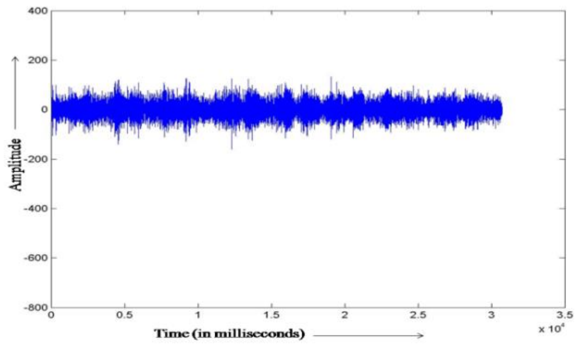


Fig. 2.2.1(b): IMF 1

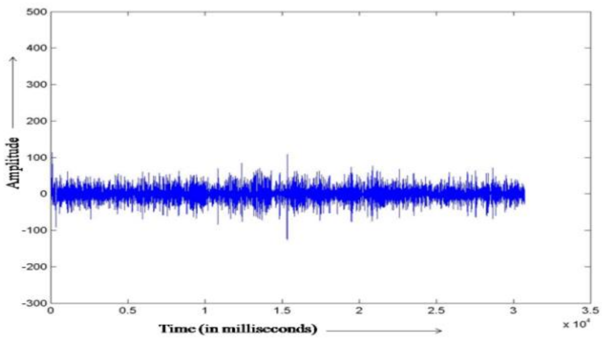


Fig. 2.2.1(c): IMF 2

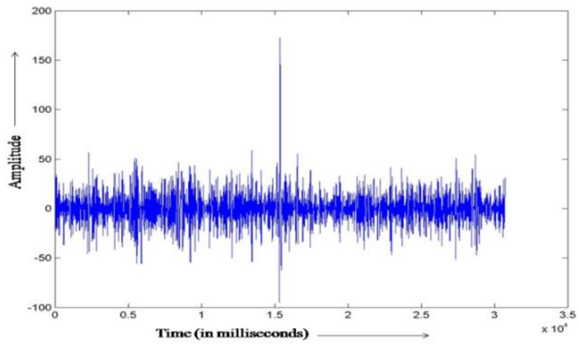


Fig. 2.2.1(d): IMF 3

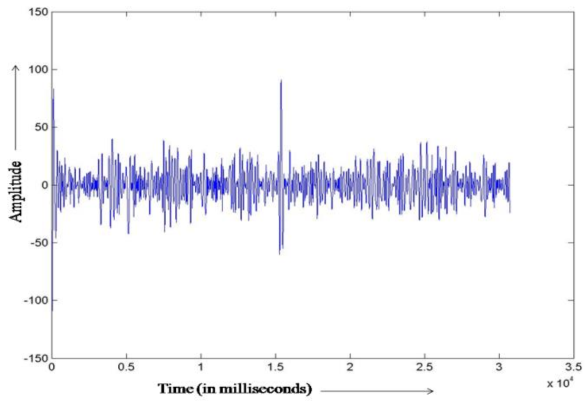


Fig. 2.2.1(e): IMF 4

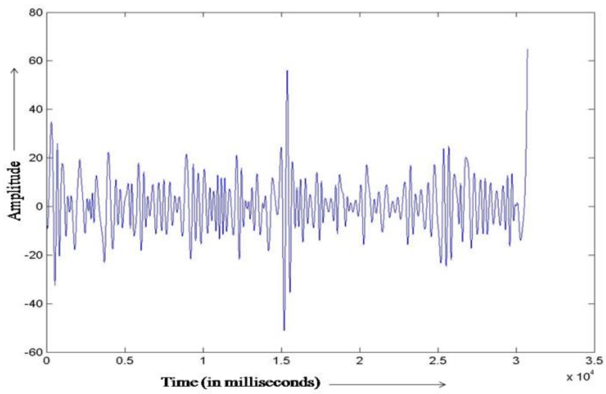


Fig. 2.2.1(f): IMF 5

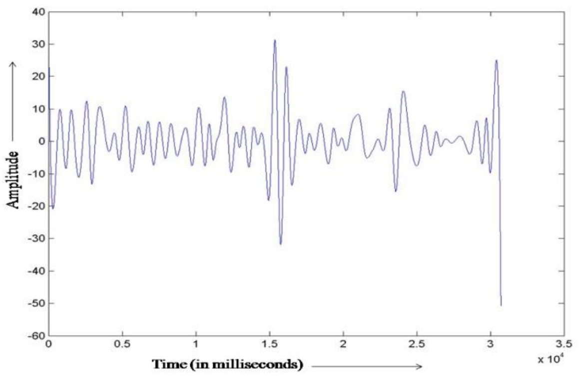


Fig. 2.2.1(g): IMF 6

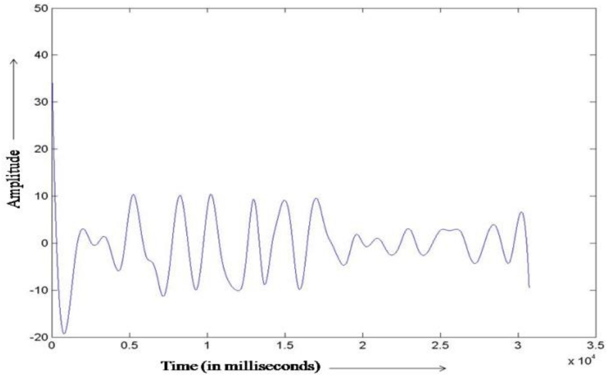


Fig. 2.2.1(h): IMF 7

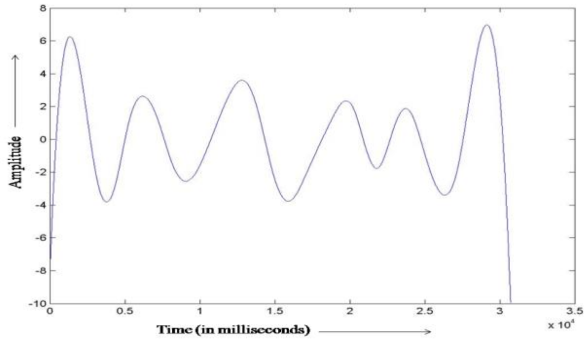


Fig. 2.2.1(i): IMF 8

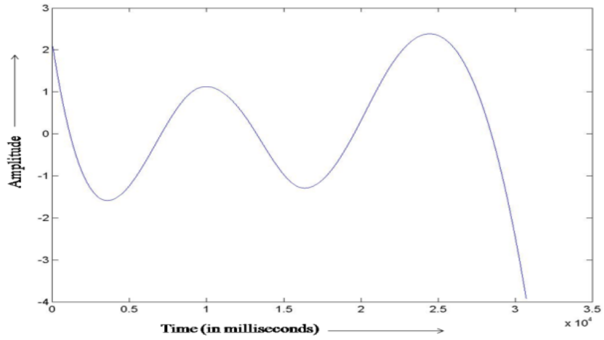


Fig. 2.2.1(j): IMF 9

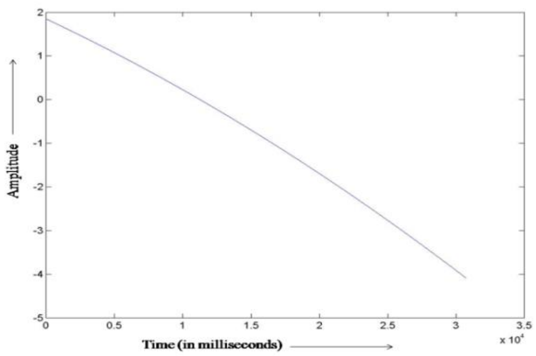


Fig. 2.2.1(k): Residue

Fig.2.2.1: Empirical Mode decomposition of a 10s EEG signal of F3 electrode

The noise-free signal obtained after the removal of muscular and blink artifacts has been used for subsequent analysis for DFA/MFDFA/MFDXA (**Fig. 2.2.2**).

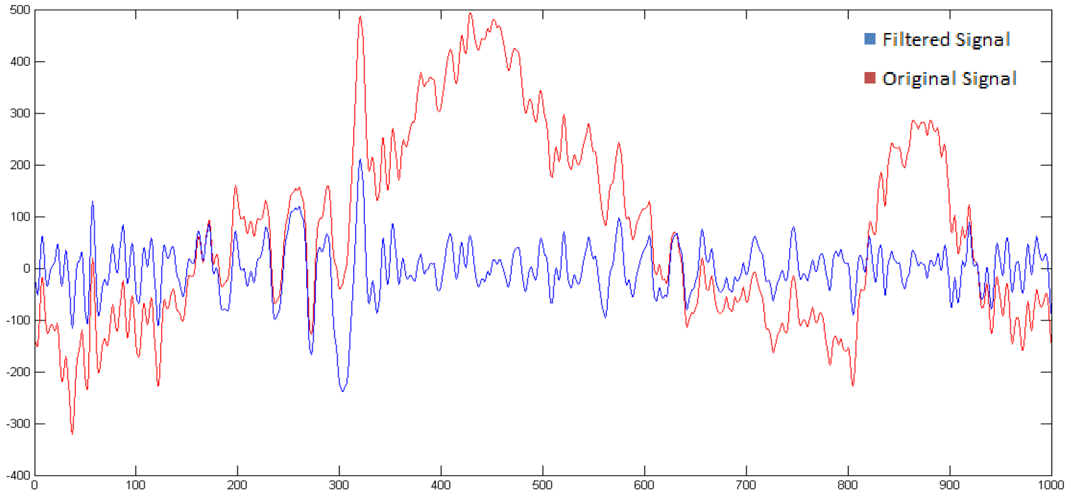


Fig. 2.2.2: Raw EEG signal and Artifact free EEG signal of 10 seconds

2.3. DETRENDED FLUCTUATION ANALYSIS (DFA)

Detrended Fluctuation Analysis (DFA) is used to analyze the long-range temporal correlations (LRTC) of the observed fluctuations in EEG. In the realm of complex cognition, scaling analysis technique was used to confirm the presence of universality and scale invariance in spontaneous EEG signals (Linkenkaer-Hansen et.al., 2001; Peng et.al, 1994). In stochastic processes, chaos theory and time series analysis, DFA is a method for determining the statistical self-affinity of a signal. It is useful for analyzing time series that appear to be long-memory processes (diverging correlation time, e.g., power-law decaying autocorrelation function) or 1/f noise. The obtained exponent is similar to the Hurst exponent, except that DFA may also be applied to signals whose underlying statistics (such as mean and variance) or dynamics are non-stationary (changing with time). DFA method was applied in (Karkare, Saha & Bhattacharya, 2009) to show that scale-free long-range correlation properties of the brain electrical activity are modulated by a task of complex visual perception, and further, such modulations also occur during the mental imagery of the same task. In case of music induced emotions, DFA was applied to analyze the scaling pattern of EEG signals in emotional music (Gao et.al, 2007) and particularly Indian music (Banerjee et.al, 2016). The DFA of a time series $[x_1, x_2,...,x_N]$ are as follows.

Step 1: Converting the noise like structure of the signal into a random walk like signal. It can be represented as:

$$\mathrm{Y}(i) = \sum (x_k - \bar{x}) \tag{11}$$

where $\bar{x}$ is the mean value of the signal.

Step 2: The whole length of the signal is divided into Ns number of segments consisting of certain no. of samples. For s as sample size and N the total length of the signal the segments are

$$Ns = \mathrm{int}\left(\frac{N}{s}\right) \tag{12}$$

The original signal with the extracted trends has been shown in the Figs. 2.3.1-2.3.2 given at the end of this section.

Step 3: The local RMS variation for any sample size *s* is the function *F(s,v)*. This function can be written as follows:

$$F^2(s,v) = \frac{1}{s}\sum_{i=1}^{s}\{Y[(v-1)s+i]-y_v(i)\}^2 \tag{13}$$

Step 4: The q-order overall RMS variation for various scale sizes can be obtained by the use of following equation

$$F_q(s) = \left\{\frac{1}{Ns}\sum_{v=1}^{Ns}[F^2(s,v)]^{\frac{q}{2}}\right\}^{\left(\frac{1}{q}\right)} \tag{14}$$

Step 5: The scaling behavior of the fluctuation function is obtained by drawing the log-log plot of $F_q(s)$ vs. s for each value of q.

$$F_q(s) \sim s^{h(q)} \tag{15}$$

where h(q) is called the generalized Hurst exponent expressed as the slope of a double logarithmic plot. For q=2, we obtain the monofractal scaling exponent or α. A monofractal time series is characterized by unique h(q) for all values of q. The parameter α (scaling exponent, autocorrelation exponent, self-similarity parameter etc.) represents the autocorrelation properties of the signal. DFA technique was applied following the NBT algorithm used in Hardstone et.al (2012). The scaling exponent provides a quantitative measure of long-range temporal correlation (LRTC) that exists in the EEG. When the EEG is completely uncorrelated (Gaussian or non-Gaussian probability distribution), the calculation of the scaling exponent yields 0.5, also called "white noise".

When applied to EEG data with LRTC, power-law behavior will generate scaling exponents with greater than 0.5 and less than 1. As the scaling exponent increases from 0.5 to 1, the LRTC in the EEG are more persistent (decaying more slowly with time). If a scaling exponent is greater than 1, the LRTC no longer exhibits power law behavior. Finally, if the scaling exponent = 1.5, this indicates Brownian noise, which is the integration of white noise. It can be converted into the Hurst exponent H = α –1 and the estimated FD accordingly as

$$D_{DFA} = 3 - \alpha. \tag{16}$$

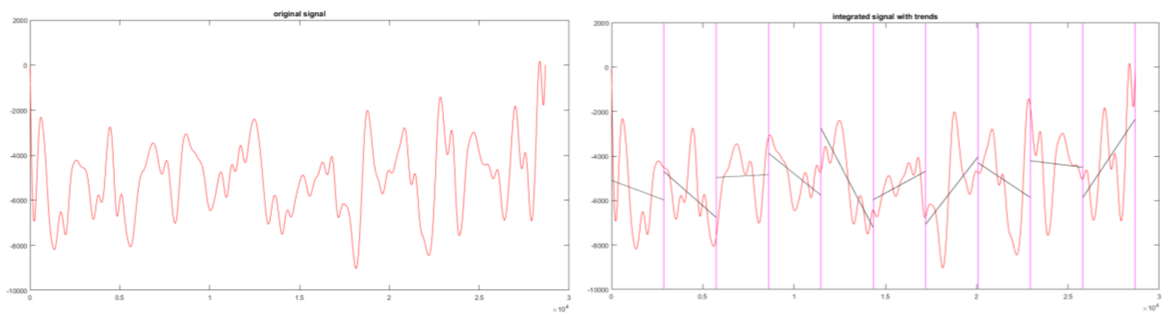


Fig. 2.3.1: Raw signal with the polyfit trends as found in Steps 1 and 2

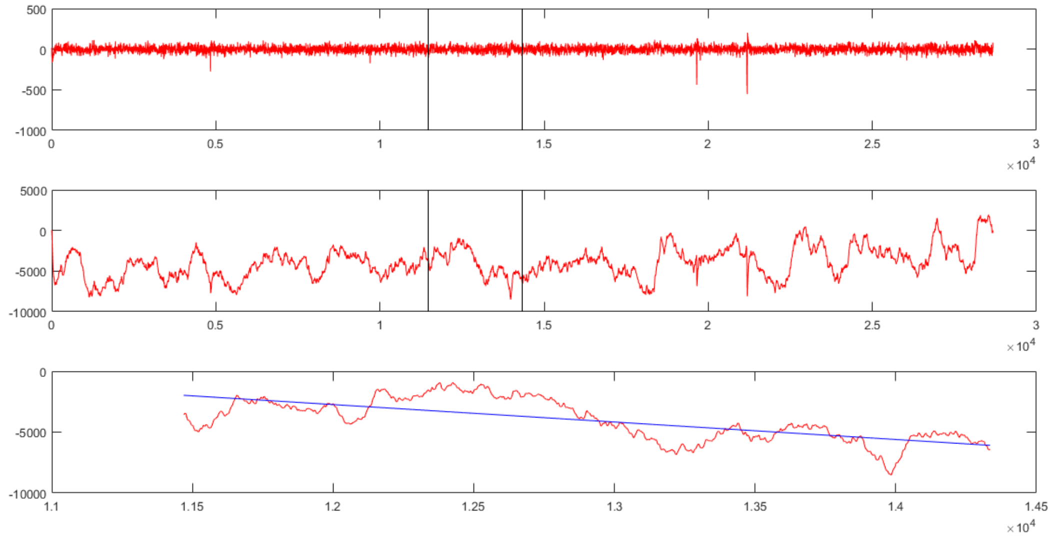


Fig. 2.3.2: Complete Signal with poly-fit trends

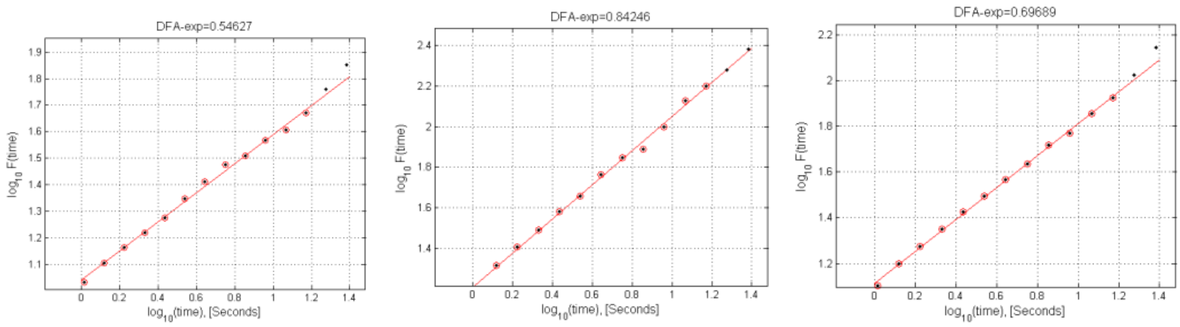


Fig. 2.3.3 (a-c): A log-log plot of an EEG signal in three different experimental conditions

2.4. MULTIFRACTAL DETRENDED FLUCTUATION ANALYSIS (MFDFA)

The real-life fractal patterns that we see hardly scale according to a single scaling exponent, rather there should be multiple scaling laws to capture their growth or variation over time. These spatial and temporal scale variations indicate a multifractal structure of a particular signal that is defined by a multifractal spectrum of power law exponents. For these more practical cases, Kantelhardt et.al (2002) formulated the MFDFA algorithm which is essentially a generalization of the DFA algorithm as given before but takes into account different scaling ratios. For eqn. (15), putting q = 2, the standard DFA procedure is retrieved. We are interested in how the generalized q dependent fluctuation functions $F_q(s)$ depend on the time scale s for different values of q. Hence, we must repeat steps 2 to 4 for several time scales s. It is apparent that $F_q(s)$ will increase with increasing s. Of course, $F_q(s)$ depends on the DFA order m. By construction, $F_q(s)$ is only defined for $s \geq m+2$. Again Step 5 is repeated with different values of q;

Step 5: Determination of the scaling behavior of the fluctuation functions by analyzing log-log plots $F_q(s)$ versus s for each value of q. If the series x_i are long-range power-law correlated, $F_q(s)$ increases, for large values of s, as a power-law,

$$F_q(s) \sim s^{h(q)} \ldots \tag{17}$$

In general, the exponent h(q) may depend on q. For stationary time series, h(2) is identical to the well-known Hurst exponent H. Thus, we will call the function h(q) generalized Hurst exponent.

The generalized Hurst exponent h(q) of MFDFA is related to the classical scaling exponent τ(q) by the relation

$$\tau(q) = qh(q) - 1 \tag{18}$$

A monofractal series with long range correlation is characterized by linearly dependent q order exponent τ(q) with a single Hurst exponent H. Multifractal signal on the other hand, possess multiple Hurst exponent and in this case, τ(q) depends non-linearly on q (Ashkenazy et.al, 2003).

The singularity spectrum f(α) is related to h(q) by

$$\alpha = h(q) + qh'(q) \tag{19}$$

$$f(\alpha) = q[\alpha - h(q)] + 1 \tag{20}$$

Where α denoting the singularity strength and *f(α)*, the dimension of subset series that is characterized by α. The width of the multifractal spectrum essentially denotes the range of exponents. The spectra can be characterized quantitatively by fitting a quadratic function with the help of least square method (Figliola et.al, 2007) in the neighbourhood of maximum α_0,

$$f(\alpha) = A(\alpha - \alpha_0)^2 + B(\alpha - \alpha_0) + C \tag{21}$$

Here C is an additive constant C = f(α_0) = 1and B is a measure of asymmetry of the spectrum. So obviously it is zero for a perfectly symmetric spectrum. We can obtain the width of the spectrum very easily by extrapolating the fitted quadratic curve to zero.

Width W is defined as,

$$W = \alpha_1 - \alpha_2 \,,$$

$$\text{with,} \quad f(\alpha_1) = f(\alpha_2) = 0 \tag{22}$$

The width of the spectrum gives a measure of the multifractality of the spectrum. Greater is the value of the width W greater will be the multifractality of the spectrum. For a monofractal time series, the width will be zero as h(q) is independent of q.

The origin of multifractality in a EEG time series can be verified by randomly shuffling the original time series data (Figliola et.al, 2007). In general, two different types of multifractality are present in a time series data: (i) Multifractality due to a broad probability density function for the values of the time series. Here, the multifractality of the time series cannot be removed by random shuffling and the shuffled data has the same variation of h(q) as the original data (ii) Multifractality due to a variety of long-range correlations due to the small and large fluctuations. In this case, the probability density function of the values can be a regular distribution with finite moments, for e. g. a Gaussian distribution. The corresponding shuffled series will exhibit non-multifractal scaling, since all long-range correlations are destroyed by the shuffling procedure. All long-range correlations that existed in the original data are removed by this random shuffling and what remains is a totally uncorrelated sequence. Hence, if the multifractality of the original data was due to long range correlation, the shuffled data will show non-fractal scaling. If any series has multifractality both due to long range correlation as well as due to probability density function, then the shuffled series will have smaller width W and hence weaker multifractality than the original time series.

The qth order fluctuation function $Fq(s)$ for 10 points of q in between −5 to +5 was obtained. The time series values of both the waves have been randomly shuffled to destroy all the long-range correlations present in the data, and what remained is a totally uncorrelated sequence. The regression plot of ln ($Fq(s)$) vs. ln(s) averaged for different values of q (q= -3 to q= + 3 is shown in the plot for scales varying from 16 to 1024) for a sample electrode F3 is given in Fig. 2.4.1 (a,b). The slope of the best fit line thus obtained from ln($Fq(s)$) vs. ln(s) plot gives the values of $h(q)$. It is seen from Fig. **2.4.1** that the shuffled values do not change with the values of q, and thus has a fixed slope h(q)=H, which is the conventional Hurst exponent for monofractal time series.

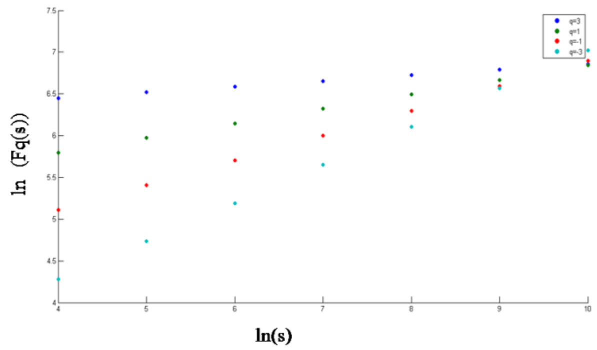


Fig. 2.4.1(a): ln(Fq(s)) vs. ln(s) for F3 electrode (original)

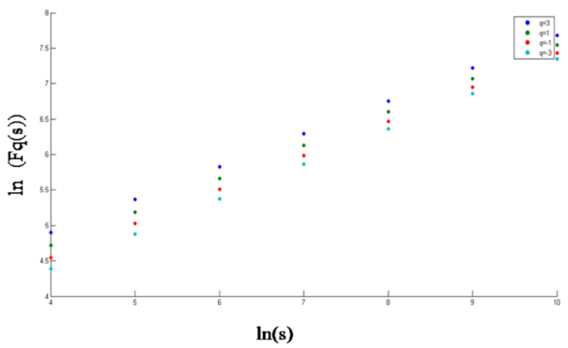


Fig. 2.4.1(b): ln(Fq(s)) vs. ln(s) for F3 electrode (shuffled)

Fig. 2.4.1(a, b): Plot of ln($Fq(s)$) vs. ln(s) showing different h(q) corresponding to each q for original and shuffled EEG signals

For monofractal time series, h(q) is independent of q, since the scaling behavior of the variances $F^2(s, v)$ is identical for all segments v, and the averaging procedure in Eq. (15) will give just this identical scaling behavior for all values of q, only if small and large fluctuations scale differently, there will be a significant dependence of h(q) on q. For positive values of q, h(q) describes the scaling behavior of the segments with large fluctuations. Usually, the large fluctuations are characterized by a smaller scaling exponent h(q) for multifractal series. On the contrary, for negative values of q, the segments v with small variance $F^2(s,v)$ will dominate the average $F_q(s)$. Hence, for negative values of q, h(q) describes the scaling behavior of the segments with small fluctuations, which are usually characterized by a larger scaling exponent.

A representative figure for variation of *h* (*q*) with *q* for two different time-series is shown in **Fig. 2.4.2 (a-b)**. It is clearly evident from the figures that the values of h (q) decrease with the increase of q, showing multifractal scaling in both the signals. For monofractal signals, a single value of Hurst exponent is obtained corresponding to different values of q, like the shuffled value of *h* (*q*) as seen in both the figures, where *h*(*q*) remains almost constant with the change of q. The amount of multifractality can be determined quantitatively in each of the windows of each signal from the width of the multifractal spectrum [f(α) vs α]. The shuffled width obtained, is found to be always smaller than the original width of the signal (**Fig. 2.4.3**). This ascertains the fact that multifractality in the signals is both due to long range correlations as well as broad probability density function. In the ideal case, the shuffled data should behave as a monofractal signal with no multifractal scaling. Thus, in the plot of Hurst exponent, it is seen that the shuffled values of h (q) do not change in general with q, and in the f(α) vs. α plot, the shuffled series will show a peak at α_0 close to 0.5.

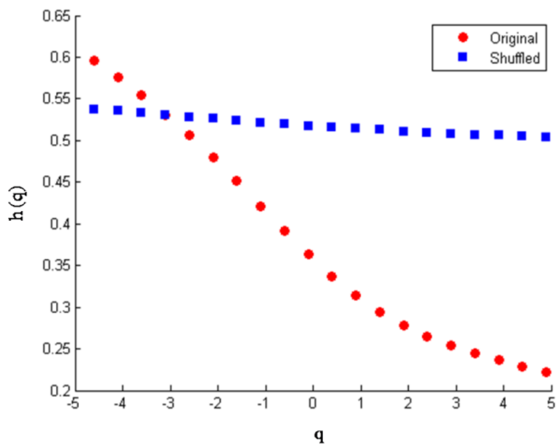


Fig. 2.4.2 (a)

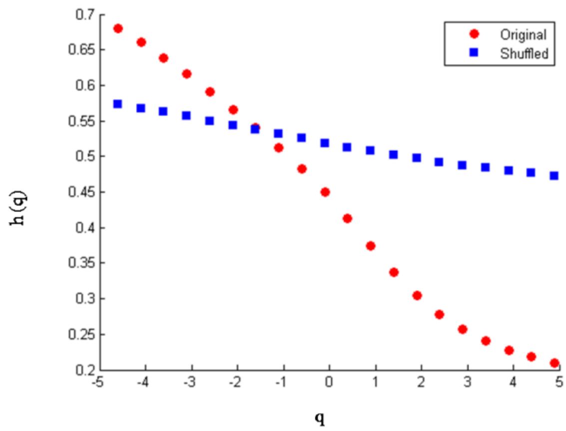


Fig. 2.4.2 (b)

Fig. 2.4.2 (a-b): Variation of h(q) with q for original and shuffled series

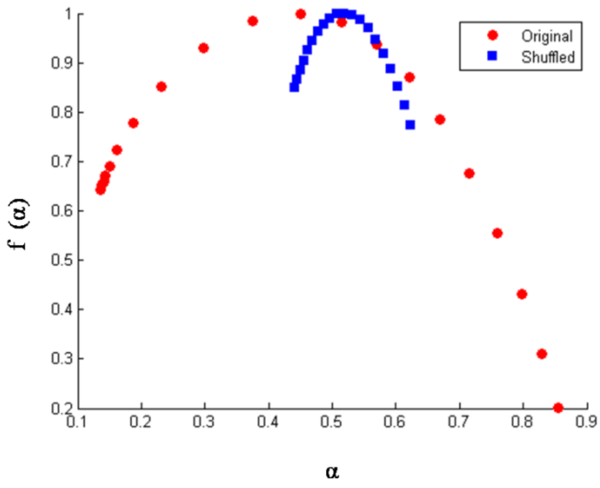


Fig. 2.4.3: f(α) vs α curve for original and shuffled series

As a generalization of the DFA method, the detrended cross-correlation analysis (DCCA) is proposed to investigate the long-term cross-correlations between two non-stationary time series (Podobnik et.al, 2008; 2009; Podobnik & Stanley, 2008; Xu, Shang & Kamae, 2010), and Multifractal Detrended Cross-Correlation Analysis (MFDXA) can unveil the multifractal features of two cross-correlated signals (He & Chen, 2011; Jiang & Zhou, 2011; Wang, Liao, Zhou & Shi, 2013; Ghosh, Dutta & Chakraborty, 2014).

2.5. MULTIFRACTAL DETRENDED CROSS-CORRELATION ANALYSIS (MFDXA)

MFDXA method was first used by Zhou (2008) and is an offshoot of the generalized MFDFA method. Here, we compute the profiles of the underlying data series x(i) and y(i) as:

$$X(i) \equiv [\textstyle\sum_{k=1}^{i} x(k) - x_{avg}] \text{ for } i = 1 \ldots N$$

$$Y(i) \equiv [\textstyle\sum_{k=1}^{i} x(k) - x_{avg}] \text{ for } i = 1 \ldots N \qquad (23)$$

The next steps proceed in the same way as the MFDFA method, with the only difference being we have to take $2N_s$ bins here. The qth order detrended covariance $F_q(s)$ is obtained after averaging over 2Ns bins.

$$F_q(s) = \{1/2N_s \textstyle\sum_{v=1}^{2Ns} [F(s,v)]^{q/2}\}^{1/q} \qquad (24)$$

where q is an index which can take all possible values except zero because in that case the factor 1/q blows up. The procedure can be repeated by varying the value of s. Fq(s) increases with increase in value of s. If the series is long range power correlated, then Fq(s) will show power law behavior

$$F_q(s) \sim s^{\lambda(q)} \qquad (25)$$

If such a scaling exists ln F_q will depend linearly on ln s, with λ(q) as the slope. Scaling exponent λ(q) represents the degree of the cross-correlation between the two time series. In general, the exponent λ(q) depends on q. We cannot obtain the value of λ(0) directly because

F_q blows up at q = 0. F_q cannot be obtained by the normal averaging procedure; instead, a logarithmic averaging procedure is applied

$$F_0(s) = \{1/4N_s \sum_{v=1}^{2Ns}[F(s,v)]\} \sim s^{\lambda(0)}. \quad (26)$$

For q = 2 the method reduces to standard DCCA. If scaling exponent λ(q) is independent of q, the cross-correlations between two time series are monofractal. If scaling exponent λ(q) is dependent on q, the cross-correlations between two time series are multifractal. Furthermore, for positive q, λ(q) describes the scaling behavior of the segments with large fluctuations and for negative q, λ(q) describes the scaling behavior of the segments with small fluctuations. Scaling exponent λ(q) represents the degree of the cross-correlation between the two time series x(i) and y(i). The value λ(q) = 0.5 denotes the absence of cross-correlation. λ(q) > 0.5 indicates persistent long-range cross-correlations where a large value in one variable is likely to be followed by a large value in another variable, while the value λ(q) < 0.5 indicates anti-persistent cross-correlations where a large value in one variable is likely to be followed by a small value in another variable, and vice versa (Movahed & Hermanis, 2008).

Zhou (2008) found that for two time series constructed by binomial measure from p-model, there exists the following relationship:

$$\lambda\,(q = 2) \approx [h_x\,(q = 2) + h_y\,(q = 2)]/2. \quad (27)$$

Podobnik and Stanley have studied this relation when q = 2 for monofractal Autoregressive Fractional Moving Average (ARFIMA) signals and EEG time series (Podonik & Stanley, 2008).

In case of two time series generated by using two uncoupled ARFIMA processes, each of both is autocorrelated, but there is no power-law cross correlation with a specific exponent (Movahed & Hermanis, 2008). According to auto-correlation function given by:

$$C\,(\tau) = \langle[x(i + \tau) - \langle x\rangle][x(i) - \langle x\rangle]\rangle \sim \tau^{-\gamma}. \quad (28)$$

The cross-correlation function can be written as:

$$C_x\,(\tau) = \langle[x(i + \tau) - \langle x\rangle][y(i) - \langle y\rangle]\rangle \sim \tau^{-\gamma_x} \quad (29)$$

where γ and γ_x are the auto-correlation and cross-correlation exponents, respectively. Due to the non-stationarities and trends superimposed on the collected data, direct calculation of these exponents is usually not recommended; rather the reliable method to calculate auto-correlation exponent is the DFA method, namely $\gamma = 2 - 2h\,(q = 2)$ (Movahed & Hermanis, 2008). Recently, Podobnik et al. (2011), have demonstrated the relation between cross-correlation exponent, γ_x and scaling exponent λ(q) derived by Eq. (27) according to $\gamma_x = 2 - 2\lambda(q = 2)$. For uncorrelated data, γ_x has a value 1 and the lower the value of γ and γ_x more correlated is the data. In general, λ(q) depends on q, indicating the presence of multifractality. In other words, we want to point out how two non-linear signals are cross-correlated in various time scales.

The qth order detrended covariance Fq(s) was obtained from relations 3 and 4 for values of q from −5 to +5 in steps of 1 just like the MFDFA part. Power law scaling of $F_q(s)$ with s is observed for all values of q as is seen from **Fig. 2.4.1(a-b)** same as those found for MFDFA. We have also shown variation of h (q) with q by means of MFDFA in **Fig. 2.4.2**. The plot depicts multifractal behavior of cross-correlations because for different q, there are different exponents; that is, for different q, there are different power-law cross-correlations. Further from the same figure we can see that the value of H(q) depends on q. We know that H(q) = 0.5 indicates that the series is an independent random process, and for H(q) < 0.5 it is characterized by long-range anti-correlations while for 0.5 < H(q)< 1, it is featured by long-term correlations. In this case the signal is stationary. The exponent H (q = 2) is equivalent with the well-known Hurst index. A representative figure (**Fig. 2.5.1**) reports the variation of cross correlation exponent λ (q) with q for two random sample signals, also the variation of h (q) with q for those two samples obtained from MFDFA technique are also shown in the same figure for comparison.

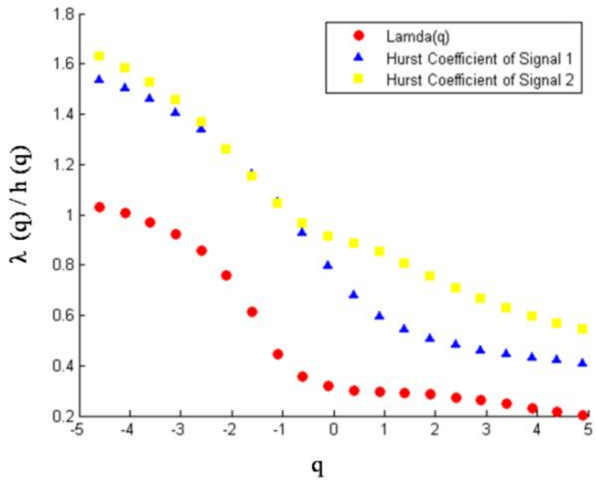


Fig. 2.5.1: Variation of λ (q) and h (q) for two sample signals

The variation of λ (q) with q for the two cross correlated signals show that they are multifractal in nature. To illustrate further the presence of multifractality in the cross-correlated music signals, i.e., to have information about the distribution of degree of cross-correlation in various time scales, a representative multifractal spectrum was plotted for the two signals in **Fig. 2.5.2**. The way to characterize multifractality of cross correlation between two samples is to relate via a λ(q) Legendre Transform as in the case of single series (Feder, 2013). The growth of the width of f(α) or equivalently Δα shows the increase in degree of multifractality of the coupled signals. Again, it becomes evident from the spectrum that the cross correlated signals are multifractal in various time scales.

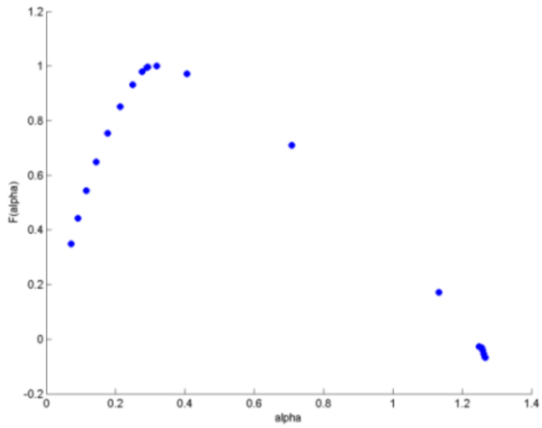


Fig. 2.5.2: Multifractal cross-correlated Spectrum of two Sample signals

Jones and Kaul (1996) were the first to reveal a stable negative cross-correlation between oil prices and stock prices. The negative cross-correlations were also found in a number of previous works (Berument, Ceylan & Dogan, 2010; Reboredo, Rivera-Castro & Zebende, 2014). A negative value of cross correlation is an indication of strong cross-correlation between the two samples for which the cross correlation is being carried out. Using the MFDXA technique we have estimated the degree of cross-correlations between neuronal potentials originating from different lobes of human brain when subjected to external stimulus.

In the next section, we proceed to propose novel statistical mechanisms for feature extraction from acoustic signals which may open up new vistas in categorizing signals based on patterns found in its informational content and also in developing automated classification algorithms in the field of Indian Classical Music (ICM), and by extension, could find an applicative domain in bio-signals as well in a not-so-distant future (Roy et al., 2019, 2021).

2.6. STATISTICAL PHYSICS APPROACH TO CATEGORIZE DATA

2.6.1 Maxwell-Boltzmann Analysis

In recent years, quite a few statistical distributions have been used in the analysis of structured and unstructured experimental data. Among them, a specific kind of distribution has found particular prominence: distributions that exhibit power-law characteristics, i.e., Cx^{-a} for $x \to \infty$ (a is the scaling exponent). Distributions having power-law tails are observed in numerous natural phenomenon –Physics (Lemoine & Sigl, 2001; Walton & Rafelski, 2000; Schlesinger et al., 1993), Seismology (Kashahara, 1981), Weather patterns (Ausloos & Ivanova, 2001), Genomics (Nacher & Ochiai, 2008), Economics (Gabaix et al., 2003; Figueira et al., 2011), City population (Blank & Solomon, 2000), E-mail networks (Ebel et al., 2002), Linguistics (Kosmidis et al., 2006), frequency of terrorist attacks (Clauset et al., 2007), Scientific paper citation (Redner, 1998), among other examples [check Newman (2005) or Kaniadakis (2009) for more]. An analytic form, known as Zipf's law, for the tails of this distribution is used regularly in data analysis purposes.

Zipf's law (Zipf, 2016, Originally published in 1949) states that if the most frequently occurring element of a distribution is assigned a rank m = 1, second most frequently occurring element is assigned rank m = 2, and so on, then the frequency of occurrence P_m of the m^{th} ranked element is:

$$P_m \sim m^{-\alpha} \tag{30}$$

It is experimentally verified that the exponent α has a value close to 1. An alternative way to represent Zipf's law is saying that the proportion of elements p_f whose frequency is f ($0 < f < 1$) in a given sample is modelled by a power function $p_f \sim f^{-\beta}$, where β is related to the exponent α of Eq. (30) by (Kosmidis et al., 2006):

$$\beta - 1 = 1/\alpha \tag{31}$$

Now, it can be seen that Zipf's law could be controlled by the Maxwell-Boltzmann (M-B) distribution which is generally associated with statistical properties of particle ensembles in physical world. For that, it is considered that complex signals, be it biological, acoustic or physical, are essentially carriers of information which has direct correlation with the dynamical state of the system under observation. Such a signal is made up of numerous basic units that are distinct and commensurate with different microstates related to the system.

Let us consider a dynamical system having l number of microstates, and correspondingly, the signal generated from this system consists of l distinct patterns: $\{s_1, s_2,\ldots,s_l\}$ (Peng et al., 2007). In this signal, any random pattern s_i will be manifested n_i times if the corresponding microstate is repeated n_i number of times. Let us also consider that there exist energy values $\{\varepsilon_1, \varepsilon_1,\ldots, \varepsilon_l\}$ associated with every corresponding microstate, i.e., pattern s_i has energy ε_i. The introduction of energy helps describe the problem in the realm of statistical physics, thus making it easier to deal with analytically. Low-energy states, relatively more stable, occurs often and higher-energy states (more energy needed for transition) are far less frequent. Also, we assume that each microstate is context-independent and is equally likely to occur. Next, we will look for the distribution of the system visiting all the microstates.

The total number of microstates, N, and the total enegy of said microstates, $\mathcal{E}$, is given by:

$$N = \sum_{i=1}^{l} n_i$$
$$\text{and } \mathcal{E} = \sum_{i=1}^{l} \varepsilon_i \tag{32}$$

For the set $\{n_1, n_2,\ldots,n_l\}$, total combination of possible microstate occurrences:

$$\Omega = \frac{N!}{\prod(n_i!)} \tag{33}$$

Now, higher the value of **Ω**, higher the probability of the set $\{n_1, n_2,\ldots,n_l\}$ occurring. So, the value of n_i is dependent on the type of distribution of $\{s_i\}$ that maximises the number of configurations **Ω**. For that, we find the maximal value of ln Ω since it is a monotonic function of Ω. Using two Lagrange multipliers α and β to consider the boundary conditions given in Eq. (32), the condition for maximum Ω, therefore, is:

$$\frac{\partial \ln \Omega}{\partial n_i} - \alpha \frac{\partial \sum_i n_i}{\partial n_i} - \beta \frac{\partial \sum_i n_i \varepsilon_i}{\partial n_i} = 0 \tag{34}$$

Using Eq. (33) and putting ln $N! \approx N \ln N - N$, for large N (Stirling formula), it becomes,

$$\ln n_i = -\alpha - \beta\varepsilon_i$$

$$\text{or, } n_i = \exp(-\alpha - \beta\varepsilon_i) \tag{35}$$

Eq. (35) indicates that the maximal Ω value is given by Boltzmann distribution as the probability of finding pattern s_i with set n_i is:

$$p(s_i) = \frac{n_i}{N} = \frac{1}{Z}\exp(-\beta\varepsilon_i) \tag{36}$$

Where Z is the partition function of the system. However, Eq. (36) isn't Zipf's law. Notable that Zipf's law deals with frequency and ranks but this approach deals with frequency and energy, which are entirely statistical parameters. The association of Zipf's law with M-B distribution could be ascertained by the interrelation between rank and energy. Kosmidis argues that the these two are related via the Hamiltonian of the system, i.e.,

$$H(k) = \varepsilon \ln k \tag{37}$$

Where, k is the 'usefulness' of a word (in linguistic analysis) and it is 'more or less proportional to its rank m' (Kosmidis et al., 2006).

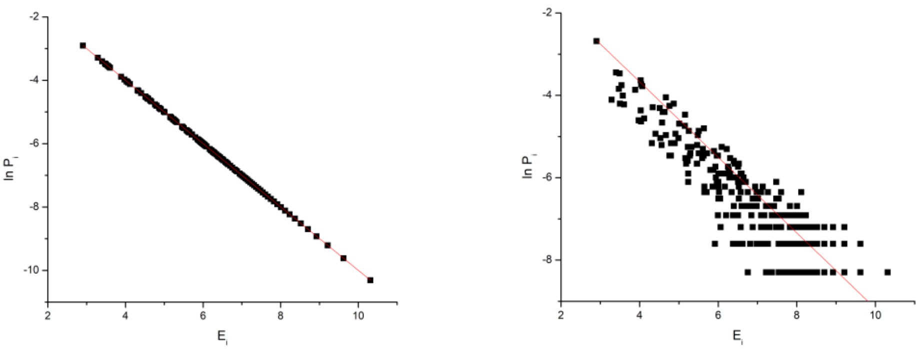
In the confinement of equilibrium statistical physics, it is well known that the probability of occurrence of a given state with a given energy is proportional to the exponential of (energy of the state)/kT, where k is Boltzmann constant (1.38 x 10^{-23} J/K) and T is the temperature, very similar to Eq. (36). T is the 'measure' of the interaction between the system and the environment (Kosmidis et al., 2006). The occurrence probability of the states is related to temperature as: high T indicates the existence of different states are equally probable and low T favours the probability of lower energy states to occur more than the higher ones. Based upon this idea, the concept of 'text temperature' has been put to use in the last few years to linguistic analysis with the rationale that human language could be described and analysed as a physical system consisting of ensemble of particles, of course under statistical boundary conditions. The basic elements of such a system (words, in this case) represent different energy states and their respective energies ('word energy') and occurrence probability can be computed with respect to a model system: for example, in linguistic analysis, it is the comprehensive corpus of the texts that are being analysed. This field of research has generated a number of novel and noteworthy ideas regarding aesthetic information (De Campos & Tolman, 1982), communicative ability (Kosmidis et al., 2006), temperature of textbooks (Miyazima & Yamamoto, 2008), size of text (Rovenchak & Buk, 2011), vocabulary complexity in different languages (Rego et al., 2014) and author's writing capacity (Chang et al., 2017).

We apply the M-B distribution to the acoustic samples obtained from various Indian Classical Music (ICM) performances in order to extract underlying dynamics of the note distribution pattern. In our approach, the basic units of the complex signal consisted of distinct combinations of different notes and their respective durations. Now, modifying Eq. (36) very slightly, we get:

$$p(s_i) \sim \exp(-\beta\varepsilon_i) \tag{38}$$

Here onwards, for simplicity, we consider k = 1, irrespective of its unit. We assume that each element has a corresponding energy value in the M-B distribution given by Eq. (38). Now, from their occurrence probability, it is not possible to calculate the distinct ε_i for each one, only the expression ($\beta\varepsilon_i$) together. However, since $\beta = 1/kT$, if we assume a 1K standard temperature for a working corpus, we can find out the respective energies by comparison. When we calculate the occurrence distribution of the elements present in an acoustic sample, we assume it will deviate from the occurrence distribution of the elements in the working musical corpus. This deviation is further used to determine the 'energy' of the elements and subsequently, the temperature of the signal as a whole. Fitting $p(s_i)$ vs ε_i for the corpus, we find the R^2 to be 1 and a slope of -1 in a semi-log plot of $\ln p(s_i)$ vs ε_i, i.e., a perfect fit which reflects the standard M-B distribution. Using similar procedure separately for each of the acoustic samples, we find that the fitting (and the slope) is slightly higher or lower than the standard distribution. The energies of the elements are determined by equating occurrence probabilities to the Maxwell-Boltzmann distribution function. Then, fitting them in the model equation y0 + A1.exp(-E/t1), the respective temperatures are obtained with respect to the corpus temperature which is assumed to be 1K (Miyazima & Yamamoto, 2008).

Figure 2.7.1 gives such an example. Here, the working corpus consists of all the note distributions of six renditions of Raga Sur Malhar sung by ICM Maestro Pandit Kumar Gandharva. The semi-log plot displays a perfect fit ($R^2 = 1$) with a slope of -1 (N = 30,069). On the other hand, the best fit slope for the semi-log plot of one individual rendition among the six, sample 5, is -0.91 (N = 4015). The respective temperatures are 1K (as assumed) and 1.03 K. Similar procedures were applied for different renditions as well.



(a) (b)

Fig. 2.7.1 (a,b). Comparison of ln (Probability) vs Energy for **(a)** a corpus containing six renditions of note distributions of Raga *Sur Malhar* sung by Pandit Kumar Gandharva, having slope -1 and **(b)** one of the individual renditions, having slope -0.91

The key idea of the analysis described in this section is to correlate the dynamics of the occurrence patterns of elements in the complex data ensemble with the underlying dynamics of microstates. However, the knowledge could only get optimised if the patterns representing the microstates could be selected in an optimal manner. This remains abstruse without prior information of the dynamical system. Current approach of assuming the microstates to be equally likely and independent of each other are of course the first order approximation which requires further refinements like incorporating patterns of different lengths as microstates or introduction of transition probabilities (Peng et al., 2007).

2.6.2. Bose-Einstein Analysis

In statistical physics, contrary to Maxwell-Boltzmann distribution which deals with classical distinguishable particles, Bose-Einstein (B-E) distribution describes the dynamics of quantum particles which are identical, indistinguishable and have integer spin, also known as Bosons (Borelli, 2009). One of the major characteristics of said particles is that they do not obey Pauli's exclusion principle, i.e., identical Bosons can exist in the same quantum state (Pauli, 1940). The general formula for B-E distribution is given by:

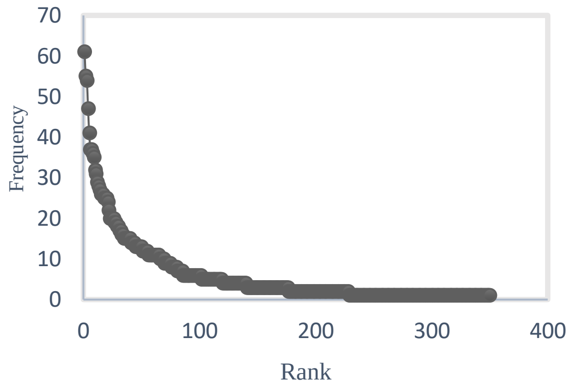
$$f(E) = \frac{1}{Ae^{E/kT} - 1} \tag{39}$$

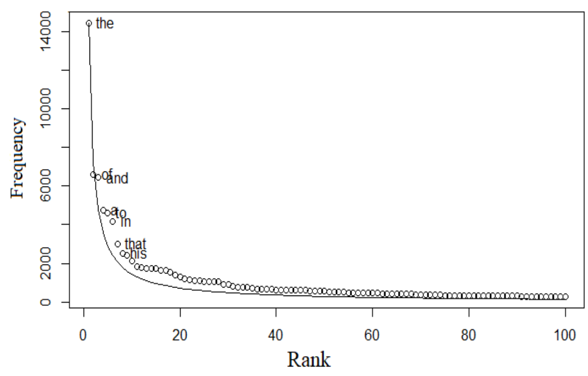
Where f(E) is the probability density function for a particle to exist in energy state E, A is normalisation constant, k = Boltzmann constant, and T is the absolute temperature. The value of A is devised such that the sum of all the probabilities is 1. The -1 factor is included to account for the indistinguishability of the particles (Beltran, 2021).

It is also worthy of note that the B-E distribution can be approximated to Maxwell-Boltzmann distribution, if the effects of the quantum particles in question are negligible. For example, the energy distribution of a gas of air molecules at room temperature can be described adequately using M-B distribution although the particles themselves are quantum. This is because the de Broglie waves of individual particles do not overlap, resulting in negligible quantum effects of a particle on its surrounding ones. In such a case, Maxwell-Boltzmann distribution is sufficient as the particles behave effectively like classical ones.

The usage of B-E distribution in this thesis is applied on the distribution of note occurrences in Indian classical music (ICM) along with their duration. This note-duration combination constructs the basis elements the repetitive patterns of which finally completes the skeleton of the auditory signal. The analysis is based on the approach used in recent years to study linguistic data (Rovenchak & Buk, 2011; Rovenchak, 2014). It relies on the analogy between rank-frequency distribution of the musical elements and the B-E distribution in the grand canonical formulation. To apply it, first, a rank-frequency distribution of the combinations present is computed, which follows the Zipfian distribution, as given in Eq. (30).

Fig. 2.7.2 demonstrates the similarity in Zipfian distribution of an ICM sample with a linguistic one. The large plateaus at the lower frequencies represent high population occurring once (or twice). In linguistic terms, such elements are called *hapax legomena* (or *hapaxes*) in Greek, which translates to 'something that occurs once/twice' (Rovenchak & Buk, 2011). It is seen that in case of text examples, about 40-60% of the words are hapaxes. In various applications in the thesis, we have found that in case of music signals, too, the number of 'musical hapaxes', i.e., elements occurring once/twice is quite high. This approach focuses mainly on such low frequency elements and their dynamics, as they are often crucial in ascertaining the melodic structure of the Raga in ICM.





(a) (b)

Fig. 2.7.2 (a,b). Comparison of Zipfian distributions of Frequency (vertical axis) and Rank of the element in the frequency table (horizontal axis): **(a)** Rank-Frequency distribution of a 1-minute *Bandish* of Raga *Shree* and **(b)** Rank-frequency distribution of the first hundred words of the book Moby Dick (Lestrade, 2017)

In the next step, the rank vs frequency distribution is 'inverted', which gives us the relation between number of elements N_j having absolute frequency j. This is computed assuming the ranks indicate a certain energy level j and the frequency of elements corresponding to that rank is the number that can be found to have that specific energy state, N_j. The energy levels are identified with occurrence frequencies. Hence, elements that have frequency 1 belong to the energy level j = 1, elements occurring twice belong to j = 2, and so on. This way, our approach makes use of the two important characteristics of Bose-Einstein distribution. Firstly, B-E statistics deal with identical and indistinguishable particles. Similarly, once we invert the rank-frequency table into energy state and its occupants, the individual elements hold no distinct identity. However different the note-duration combination can be, here onwards they behave like indistinguishable elements that can only be characterized as what j (energy state) do they belong. Second, similar to how B-E distribution doesn't obey Pauli's exclusion principle, in our case, the number of N_j that can occupy a certain j-level is not governed by any particular rule. The lowest level of such a distribution containing musical hapaxes corresponds to the famed B-E condensate in physical world (at, T = 0). Now, the j-th level occupation is given by (Rovenchak & Buk, 2011):

$$N_j = \frac{1}{z^{-1}e^{\varepsilon_j/T}-1} \tag{40}$$

Where, z is the analogue of fugacity, ε_j is the energy of j-th level and T is analogous to temperature. The z and T parameters are helpful in characterising and categorizing the dynamical nature of the samples used.

For lower-level energy spectra, the relation of ε_j and j is given as:

$$\varepsilon_j = (j-1)^{\alpha}, \qquad \text{with } 1 < \alpha < 2 \tag{41}$$

The unity is subtracted to ensure that the lower most level (j = 1) has zero energy ($\varepsilon_1 = 0$).

The parameters that are of our interest are z, α and T from Eq. (40). The 'fugacity' parameter, z, is determined from the occupancy of the lowermost state (i.e., the number of musical hapaxes) N_1:

$$N_{hapax} = N_1 = \frac{z}{1-z}$$

$$\text{or, } z = \frac{N_1}{N_1+1} \tag{42}$$

Eq. (42) gives the value of z. Parameter α is found by replacing ε_j in Eq. (40) with Eq. (41):

$$N_j = \frac{1}{z^{-1}e^{(j-1)^{\alpha}/T}-1} \tag{43}$$

Fitting the Eq. (43) for various N_j and j (≥2) by curve-fitting technique, we obtain both α and T parameters. It is of note that both T and ε_j are dimensionless quantities here, unlike the case in M-B distribution where the temperature has a unit (Kelvin) as it is computed by comparing its deviation from a standard working corpus (whose temperature is assumed as 1K).

In our B-E approach, the lower j values are the ones that contribute significantly as N_j falls rapidly with increasing j. The upper limit of j is ideally kept between j_{max} = 10 to j_{max} = 20 as the relation between ε_j and j is most likely to be different for higher j's (Rovenchak & Buk, 2011).

It is seen that the temperature T grows as the element number N increases, unlike the physical idea of temperature that doesn't depend on the number of particles. This is most likely due to the absence of thermodynamic limiting conditions (Rovenchak, 2014). It is also seen that the ratio (ln T/ln N) has a very weak dependence on the size of the sample N and hence, could be used for the comparative analysis of different experimental samples. The τ-N plane is used for categorization purposes, where τ stands for:

$$\tau = \frac{\ln T}{\ln N} \tag{44}$$

To sum up, it is interesting to see that the analytical and statistical approaches described in the previous sections, despite them belonging to the realm of traditional physical world, are applied into a vastly different domain and yet they churn out results which indicate the underlying dynamics of the systems efficiently. Such novel approaches often lack the clarity of taxonomy, hence resorting to analogy. Nonetheless, the fact that they have proven to be empirically emphatic in a number of scenarios speaks volume for the symmetry and consistency nature and natural laws hold across both tangible and intangible spheres of life. The observed systems - be it bio-signals or acoustic signals - are complex and subsequently require robust analytical tools, as described in this chapter, to be explored explicitly. In subsequent chapters, such efforts were made to bring about an analytical discipline under which a variety of complex signals could be analysed and parametrized.

REFERENCES

Ashkenazy, Y., Baker, D. R., Gildor, H., & Havlin, S. (2003). Nonlinearity and multifractality of climate change in the past 420,000 years. *Geophysical research letters*, *30*(22).

Ausloos, M., & Ivanova, K. (2001). Power-law correlations in the southern-oscillation-index fluctuations characterizing El Ni ñ o. *Physical Review E*, *63*(4), 047201.

Banerjee, A., Sanyal, S., Patranabis, A., Banerjee, K., Guhathakurta, T., Sengupta, R., ... & Ghose, P. (2016). Study on brain dynamics by non linear analysis of music induced EEG signals. *Physica A: Statistical Mechanics and its Applications*, *444*, 110-120.

Beltran, L. (2021). Quantum Bose–Einstein Statistics for Indistinguishable Concepts in Human Language. *Foundations of Science*, 1-13.

Berument, H., Ceylan, N. B., & Dogan, N. (2010). The impact of oil price shocks on the economic growth of selected MENA countries. *Energy Journal*, *31*(1), 149.

Bigerelle, M., & Iost, A. (2000). Fractal dimension and classification of music. *Chaos, Solitons & Fractals*, *11*(14), 2179-2192.

Bizopoulos, P. A., Al-Ani, T., Tsalikakis, D. G., Tzallas, A. T., Koutsouris, D. D., & Fotiadis, D. I. (2013, July). An automatic electroencephalography blinking artefact detection and removal method based on template matching and ensemble empirical mode decomposition. In *2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC)* (pp. 5853-5856). IEEE.

Blank, A., & Solomon, S. (2000). Power laws in cities population, financial markets and internet sites (scaling in systems with a variable number of components). *Physica A: Statistical Mechanics and its Applications*, *287*(1-2), 279-288.

Borrelli, A. (2009). Bose—einstein statistics. In *Compendium of Quantum Physics* (pp. 74-78). Springer, Berlin, Heidelberg.

Chang, M. C., Yang, A. C. C., Stanley, H. E., & Peng, C. K. (2017). Measuring information-based energy and temperature of literary texts. *Physica A: Statistical Mechanics and its Applications*, *468*, 783-789.

Clauset, A., Young, M., & Gleditsch, K. S. (2007). On the frequency of severe terrorist events. *Journal of Conflict Resolution*, *51*(1), 58-87.

De Campos, H., & Tolman, J. M. (1982). The informational temperature of the text. *Poetics Today*, *3*(3), 177-187.

Di Lorenzo, P. (2003, December). Chaos hidden in recorded monodic music: what is the role of timbre. In *Proccedings of" 2nd International Conference" Understanding and creating music", Caserta* (pp. 11-15).

Ebel, H., Mielsch, L. I., & Bornholdt, S. (2002). Scale-free topology of e-mail networks. *Physical review E*, *66*(3), 035103.

Feder, J. (2013). *Fractals*. Springer Science & Business Media.

Figliola, A., Serrano, E., Rostas, J. A. P., Hunter, M., & Rosso, O. A. (2007, May). Study of EEG brain maturation signals with multifractal detrended fluctuation analysis. In *AIP Conference Proceedings* (Vol. 913, No. 1, pp. 190-195). AIP.

Figueira, F. C., Moura Jr, N. J., & Ribeiro, M. B. (2011). The gompertz–pareto income distribution. *Physica A: Statistical Mechanics and its Applications*, *390*(4), 689-698.

Gabaix, X., Gopikrishnan, P., Plerou, V., & Stanley, H. E. (2003). A theory of power-law distributions in financial market fluctuations. *Nature*, *423*(6937), 267-270.

Gao, J., Cao, Y., Tung, W. W., & Hu, J. (2007). *Multiscale analysis of complex time series: integration of chaos and random fractal theory, and beyond*. John Wiley & Sons.

Gao, J., Hu, J., Mao, X., & Perc, M. (2012). Culturomics meets random fractal theory: insights into long-range correlations of social and natural phenomena over the past two centuries. *Journal of The Royal Society Interface*, *9*(73), 1956-1964.

Gao, T. T., Wu, D., Huang, Y. L., & Yao, D. Z. (2007). Detrended fluctuation analysis of the human EEG during listening to emotional music. *Journal of Electronic Science and Technology*, *5*(3), 272-277.

Ghosh, D., Dutta, S., & Chakraborty, S. (2014). Multifractal detrended cross-correlation analysis for epileptic patient in seizure and seizure free status. *Chaos, Solitons & Fractals*, *67*, 1-10.

Hardstone, R., Poil, S. S., Schiavone, G., Jansen, R., Nikulin, V. V., Mansvelder, H. D., & Linkenkaer-Hansen, K. (2012). Detrended fluctuation analysis: a scale-free view on neuronal oscillations. *Frontiers in physiology*, *3*, 450.

He, L. Y., & Chen, S. P. (2011). Multifractal detrended cross-correlation analysis of agricultural futures markets. *Chaos, Solitons & Fractals*, *44*(6), 355-361.

Hsü, K. J., & Hsü, A. J. (1990). Fractal geometry of music. *Proceedings of the National Academy of Sciences*, *87*(3), 938-941.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., ... & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences*, *454*(1971), 903-995.

Jiang, Z. Q., & Zhou, W. X. (2011). Multifractal detrending moving-average cross-correlation analysis. *Physical Review E*, *84*(1), 016106.

Johnston, D. (1999). Cool edit 2000. *Syntrillium Software Corporation, Phoenix, Arizona, EEUU*.

Jones, C. M., & Kaul, G. (1996). Oil and the stock markets. *The Journal of Finance*, *51*(2), 463-491

Jung CY, Saikiran SS (2016). A review on EEG artifacts and its different removal technique. *Asia-Pacific J Converg Res Interchang 2*(4):43–60

Kaniadakis, G. (2009). Maximum entropy principle and power-law tailed distributions. *The European Physical Journal B*, *70*(1), 3-13.

Kannathal, N., Acharya, U. R., Lim, C. M., & Sadasivan, P. K. (2005). Characterization of EEG—a comparative study. *Computer methods and Programs in Biomedicine*, *80*(1), 17-23.

Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., & Stanley, H. E. (2002). Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, *316*(1-4), 87-114.

Kantz, H., & Schreiber, T. (2004). *Nonlinear time series analysis* (Vol. 7). Cambridge university press.

Karkare, S., Saha, G., & Bhattacharya, J. (2009). Investigating long-range correlation properties in EEG during complex cognitive tasks. *Chaos, Solitons & Fractals*, *42*(4), 2067-2073.

Kasahara, K. (1981). *Earthquake mechanics* (Vol. 248). Cambridge: Cambridge university press.

Klonowski, W. (2009). Everything you wanted to ask about EEG but were afraid to get the right answer. *Nonlinear biomedical physics*, *3*(1), 1-5.

Kosmidis, K., Kalampokis, A., & Argyrakis, P. (2006). Statistical mechanical approach to human language. *Physica A: Statistical Mechanics and its Applications*, *366*, 495-502.

Lemoine, M., & Sigl, G. (Eds.). (2001). *Physics and astrophysics of ultra high energy cosmic rays* (Vol. 576). Springer Science & Business Media.

Lestrade, S. (2017). Unzipping Zipf's law. *PloS one*, *12*(8), e0181987.

Linkenkaer-Hansen, K., Nikouline, V. V., Palva, J. M., & Ilmoniemi, R. J. (2001). Long-range temporal correlations and scaling behavior in human brain oscillations. *Journal of Neuroscience*, *21*(4), 1370-1377.

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of atmospheric sciences*, *20*(2), 130-141.

Mandelbrot, B. B., & Mandelbrot, B. B. (1982). *The fractal geometry of nature* (Vol. 1). New York: WH freeman.

Mantegna, R. N., & Stanley, H. E. (1999). *Introduction to econophysics: correlations and complexity in finance*. Cambridge university press.

Miyazima, S., & Yamamoto, K. (2008). Measuring the temperature of texts. *Fractals*, *16*(01), 25-32.

Movahed, M. S., & Hermanis, E. (2008). Fractal analysis of river flow fluctuations. *Physica A: Statistical Mechanics and its Applications*, *387*(4), 915-932

Nacher, J. C., & Ochiai, T. (2008). Power-law distribution of gene expression fluctuations. *Physics Letters A*, *372*(40), 6202-6206.

Newman, M. E. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary physics*, *46*(5), 323-351.

Pauli, W. (1940). The connection between spin and statistics. *Physical Review*, *58*(8), 716.

Peng, C. K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., & Goldberger, A. L. (1994). Mosaic organization of DNA nucleotides. *Physical review e*, *49*(2), 1685.

Peng, C. K., Yang, A. C. C., & Goldberger, A. L. (2007). Statistical physics approach to categorize biologic signals: from heart rate dynamics to DNA sequences. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *17*(1), 015115.

Podobnik, B., & Stanley, H. E. (2008). Detrended cross-correlation analysis: a new method for analyzing two nonstationary time series. *Physical review letters*, *100*(8), 084102.

Podobnik, B., Grosse, I., Horvatić, D., Ilic, S., Ivanov, P. C., & Stanley, H. E. (2009). Quantifying cross-correlations using local and global detrending approaches. *The European Physical Journal B*, *71*(2), 243-250.

Podobnik, B., Horvatic, D., Ng, A. L., Stanley, H. E., & Ivanov, P. C. (2008). Modeling long-range cross-correlations in two-component ARFIMA and FIARCH processes. *Physica A: Statistical Mechanics and its Applications*, *387*(15), 3954-3959.

Podobnik, B., Jiang, Z. Q., Zhou, W. X., & Stanley, H. E. (2011). Statistical tests for power-law cross-correlated processes. *Physical Review E*, *84*(6), 066118

Reˆgo, H. H. A., Braunstein, L. A., D′ Agostino, G., Stanley, H. E., & Miyazima, S. (2014). When a text is translated does the complexity of its vocabulary change? Translations and target readerships. *PloS one*, *9*(10), e110213.

Reboredo, J. C., Rivera-Castro, M. A., & Zebende, G. F. (2014). Oil and US dollar exchange rate dependence: A detrended cross-correlation approach. *Energy Economics*, *42*, 132-139

Redner, S. (1998). How popular is your paper? An empirical study of the citation distribution. *The European Physical Journal B-Condensed Matter and Complex Systems*, *4*(2), 131-134.

Rovenchak, A. (2014). Trends in language evolution found from the frequency structure of texts mapped against the Bose-distribution. *Journal of Quantitative Linguistics*, *21*(3), 281-294.

Rovenchak, A., & Buk, S. (2011). Application of a quantum ensemble model to linguistic analysis. *Physica A: Statistical Mechanics and its Applications*, *390*(7), 1326-1331.

Roy, S., Banerjee, A., Sanyal, S., Ghosh, D., & Sengupta, R. (2019). Categorization of Indian Classical Music Using MB-BE Distributions. *Journal of Software Engineering Tools & Technology Trends*. *6*(3), 9–15.

Roy, S., Banerjee, A., Sanyal, S., Ghosh, D., & Sengupta, R. (2021, April). A study on Raga characterization in Indian classical music in the light of MB and BE distribution. In *Journal of Physics: Conference Series* (Vol. 1896, No. 1, p. 012007). IOP Publishing.

Sengupta, R., Dey, N., Datta, A. K., & Ghosh, D. (2005). Assessment of musical quality of tanpura by fractal-dimensional analysis. *Fractals*, *13*(03), 245-252.

Shlesinger, M. F., Zaslavsky, G. M., & Klafter, J. (1993). Strange kinetics. *Nature*, *363*(6424), 31-37.

Sjölander, K., & Beskow, J. (2000). Wavesurfer-an open source speech tool. In *Sixth International Conference on Spoken Language Processing*.

Stanley, H. E., & Meakin, P. (1988). Multifractal phenomena in physics and chemistry. *Nature*, *335*(6189), 405-409.

Walton, D. B., & Rafelski, J. (2000). Equilibrium distribution of heavy quarks in Fokker-Planck dynamics. *Physical review letters*, *84*(1), 31.

Wang, F., Liao, G. P., Zhou, X. Y., & Shi, W. (2013). Multifractal detrended cross-correlation analysis for power markets. *Nonlinear Dynamics*, *72*(1), 353-363.

Xu, N., Shang, P., & Kamae, S. (2010). Modeling traffic flow correlation using DFA and DCCA. *Nonlinear Dynamics*, *61*(1), 207-216.

Zhou, W. X. (2008). Multifractal detrended cross-correlation analysis for two nonstationary signals. *Physical Review E*, *77*(6), 066211

Zipf, G. K. (2016). *Human behavior and the principle of least effort: An introduction to human ecology*. Ravenio Books.

CHAPTER 3

COLOR OF THE MUSIC: A CHAOS-BASED NONLINEAR STUDY OF COGNITIVE CROSS-MODALITIES

"The world speaks to me in colors, my soul answers in music."

Rabindranath Tagore

ABSTRACT

The relationship between color and music as part of the complex system consisting of visual and auditory domain has been investigated in this study. As both the stimulus forms are processed in the same part of the human body, i.e., the brain, it will be really interesting to examine whether they share a similarity in perception. Needless to say, that color and music both have strong impact on emotion and feelings & also a few studies have been reported in literature to explore causal relationship between color and emotion. This work reports a neuro-cognitive study on response of brain to two different stimulus and their cross-modal associations. In this study the correlation between emotional arousal and the effect of audio and visual stimuli has been studied from a new perspective. 93 participants were asked to hear 6 different music pieces (each of 30 second duration). The type of emotion elicited by different music pieces were identified by the participants from a given collection of possible emotional responses. Then they are asked to assign a color associating the emotion from a given color wheel (structured according to Munsell color system/RGB color space). Each color, associated with a particular music piece, is a mixture of specific Red, Green and Blue values (RGB triplet) and has a specific HEX number (hexadecimal representation), which is recorded for each response. Then, the musical pieces used were further zoomed with the help of fractal technique to identify different emotions related to music in a quantitative measure. Here, to analyze the complexity of the sound signal (which are nonstationary and scale varying in nature), we have used Multifractal detrended fluctuation analysis (MFDFA), which is capable of determining multifractal scaling behavior of non-stationary time series. From the experimental data, it is seen that the visual and emotional response to the auditory stimulus follows a specific trend which is directly related to the stimulus complexity.

Keywords: Music, Color, MFDFA, RGB triplet, Hexadecimal representation, Cross-modal correspondence

3.1. INTRODUCTION

The correlation between color and music with effect causing emotional arousal has always been an integral part of interest for researchers. It was this thought that made none other than Sir Issac Newton curious enough to propose such a correspondence in his book '*Opticks*' back in eighteenth century (Newton, 1952). Since then, many attempts have been made to identify systematic links between music and color. The interesting incident of sensory arousal due to multi-sensory stimulus, scientifically identified as synesthesia, seems to intrigue the researchers for long (Galton, 1883; Marks, 1975; Ramachandran & Hubbard, 2001; Sagiv & Robertson, 2005; Zamm et al., 2013). Studies have shown that non-synesthetic people also have visual-to-auditory associations. These are termed as cross-modal correspondences (Spence, 2011). But what is the mechanism that gives rise to such correspondence? The role of the neurobiological structure of brain and the neuronal activities behind the phenomenon has been argued by some (Stein & Wallace, 1996). Some researches dealt with semantic mediation of the cross-modal association in cognitive level (Spence, 2001, 2011; Tuuri & Eerola, 2012). A relatively newer approach has highlighted the role of emotion as a mediator. The idea that emotion has a role to play in the multi-sensory association is not new (Odbert et al., 1942; Marks, 1975). Though the abstract nature of emotion made the concept harder to work with, it gained momentum in the past decade due to a plethora of work on music and associated emotions (Bresin, 2005; Barbiere et al., 2007; Palmer et al., 2013). In our study on the correspondence of the auditory and visual stimulus, we will focus on this mechanism.

The core argument we wish to present is that emotion plays a key role in mediating these two stimuli in the brain. That is to say, both color (visual stimulus) and music (auditory stimulus) has similar emotional qualities that inspire arousal in a similar manner. Why so? The reason is: music and color have been found to instigate emotional arousal time and again (Cutietta & Haggerty, 1987). The strong relation between music and emotion has been repeatedly reported in various studies (Blood & Zatorre, 2001). Again, music has been shown to affect the emotional state across age, culture and language boundaries (Juslin & Sloboda, 2001; Scherer, 2005). Similarly, association of color with emotions is also reported in previous literatures (D'Andrade & Egan, 1974; Terwogt & Hoeksma, 1995; Kaya & Epps, 2004). Not only this, in one of our past works, the complexity in bio-signals using color as a visual stimulus has been found to be higher than that of music (Roy C. et al., 2015). In light of these, this study is conducted to find whether a quantitative correlation can be found between music, color and emotion since both the stimuli is connected to emotion in a consistent manner.

3.2. UNIQUENESS OF THE STUDY AND METHODOLOGIES USED

Previous works in the audio-visual correspondence demonstrate an array of different methodologies and subsequent conclusions. Giannakis & Smith (2001) pointed out relationships between the loudness and pitch of pure tones with saturation and light intensity respectively. The study by Lipscomb & Kim (2004), primarily looking for an algorithm to translate auditory stimulus in visual animation, reported the existence of a few paired-up attributes of two stimuli: pitch of the tone and location of the visual counterpart, loudness (auditory) to size (visual) and timbre (auditory) to shape (visual). Interesting works on color associating with different musical genre is also worth noting (Holm et al., 2009). Although all of these works have identified the correspondence between specific features of two different sensory inputs, they do not include any investigation whether emotion might have a role in the said correspondences. Bresin (2005) held the emotional expressiveness of the stimulus into account while experimenting. His work indicated that the previous notion of associating brighter colors to positive emotions and darker colors to negative emotions was indeed a reality. This result is in sync with the study by Barbiere et al. (2007) where the general association of brighter colors (red, yellow) to 'joyful' music and darker colors (grey) to 'sad' music is agreed upon (although instead of actual color, the color-words are considered). In recent times, Palmer et al. (2013) and Lindborg & Freiberg (2015) used colors instead of color-words and demonstrated that robust cross modal matches between music and colors are mediated by emotion as an active component.

In light of these works, our experiment has been designed using features which were not previously used in any of the works reported. Except for Lindborg & Freiberg (2015), the visual domains used in previous studies are either color-word or limited choices of colors or specific color patches. In our experiment, to give the participants a more perceptual freedom to express their color preference, we have used a cylindrical color wheel in the RGB/HSL physical space. Similar to cylindrical co-ordinates (ρ, φ, z), this color wheel has three corresponding components, as given in Fig. 3.1:

1. The azimuthal angle (φ) corresponds to the feature *Hue*.
2. The radial distance (ρ) corresponds to *Saturation*.
3. The height of the column (z) gives the *Lightness* or *Luminance* (or *Brightness*).

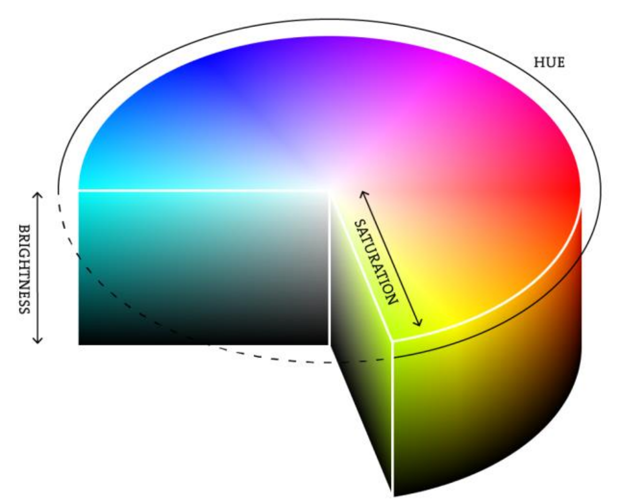


Fig. 3.1. HSL color space. Three components Hue, Saturation and Brightness are represented by angles φ, ρ and height of the column z, respectively.

Each point on this wheel is unique since it is specified by different RGB and HSL values, quite similar to the physical co-ordinate system, thus giving the full freedom to make choices. Using this wheel, the participant can specify a point expressing the color of their choice according to the auditory stimulus they hear. This point gives us the necessary details on RGB and HSL values. The color selection interface is given in Fig. 3.2.

In this study, we have chosen the RGB component for analysis. Although the color space could be described in various three-dimensional approaches (RGB, HSL, HSV, xyY amongst others), the RGB space is the one which is widely used in digital color interfaces like computer monitors and other graphic designs. Human visual perception is equipped to handle these three primary colors effectively and can use various ratios of these additive primary colors to have a wholesome color experience. In previous conventional researches including the visual stimulus, the RGB (or HSL, since it can be transformed from the former and vice versa) space has mostly been used. The reason, we believe, is partly because of the technology those works rely on constructs the color space according to the RGB model and partly because the taxonomy of human color perception, much like the human emotional experience, is underdeveloped. But unlike the emotion counterpart, the basic variables or co-ordinates in the space of color perception is more or less defined and used for long. Hence, to analyse color experience in this study, we decided to resort to the conventional color space representation of RGB, since this approach seemed the easiest to convey our understandings properly.

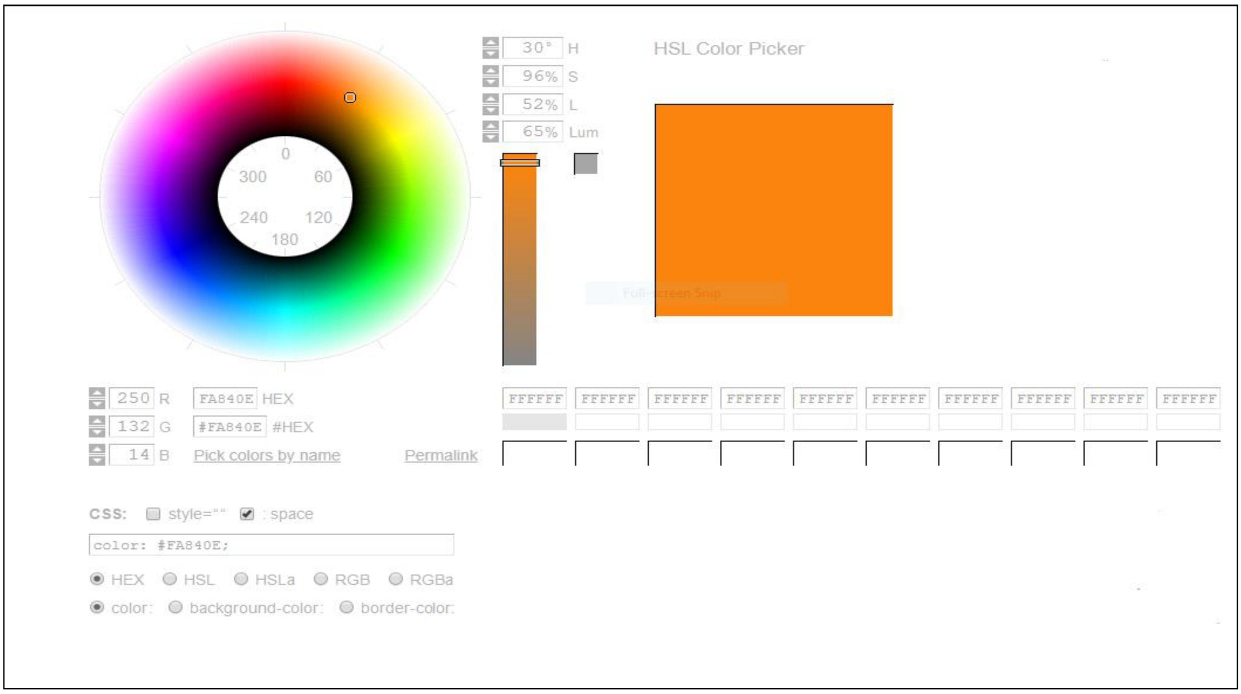


Fig. 3.2. Color Selection Interface. Any chosen point on the wheel corresponds to a unique RGB and HSL value. The circle represents the hue and saturation. Luminance or Lightness value z can be changed by shifting the slider on the column. Each point corresponds to a unique #HEX number (*Image courtesy*: workwithcolor.com, *URL*: http://www.workwithcolor.com/hsl-color-picker-01.htm**).**

Next, to ensure the emotional expressiveness, we have then recorded the participant's response by asking him/her to mark the emotion conveyed by the auditory stimulus from an array of emotion-indicating options which is inspired by the *Navarasa* (Nine *Rasa*s) theory of emotions found in Indian philosophy (from the nine discrete emotions described, we have used eight in our study (Pollock, 2016). This way, the color choice and the emotional response can be cross-checked from empirical data, rather than pre-conceived notions of emotional spectrum associated with a specific stimulus.

Thirdly, the most important part of the study consists of analyzing the auditory stimulus. No other studies, in this field, have ever done this kind of analysis previously. Before delving further into the uniqueness of the analysis, we wish to address the question that what is the need to vivisect the auditory stimuli (which, in our case are music clips of various emotions). From the perspective of human brain, the cognitive processing of music simultaneously engages most of the perceptual and emotional processes. A wide range of human response based psychological studies were conducted over the last century to know the exact modality in which emotional appraisal takes place due to various features of music. Although an extensive review is beyond the scope of this chapter, these studies reveal that specific features of music such as intensity (loudness), tempo, dissonance, and pitch, are strongly associated with emotional expressions. Small changes in any of these features result in considerable change of emotional expressions and affective experience (Ilie & Thompson, 2006, 2011; Husain et al., 2002). There have been studies in recent years that points towards music to be a complex and multidimensional nonlinear system because every instant, components (in micro and macro structural scale: pitch, timbre, accent, duration, phrase, melody etc.) are close linked to each other (González-Espinoza et al., 2017). All these properties are features of systems with chaotic, self-organized, and generally, nonlinear behavior. Hence, considering the complex nature of music, the question that follows naturally: Can analyzing the stimulus encourage a way towards the quantitative analysis of the emotional arousal? This study investigates the answers in a novel direction.

After acknowledging the need of the stimulus analysis, it is now important to address the obvious problem of the appropriate tools to do that, given the chaotic and self-similar nature of music. Unfortunately, the scarcity of robust scientific methods does us no help in his regard. Music signals are too complex in their dynamics (due to the superposition of a large number of frequency components) to be treated with conventional linear tools like Fourier transform or power spectrum techniques etc., among others. This is where we need to resort to non-linear analysis, namely: fractal technique such as MFDFA (Multifractal Detrended Fluctuation Analysis).

3.3. METHODOLIGIES

Multifractal detrended fluctuation analysis (MFDFA)

Originated from Chaos theory, fractal techniques are essential to underline the complex details hidden in an otherwise random or chaotic process. In many natural processes which are chaotic in nature, fractals help to scale the nature of chaos to an accessible level. From the structure of viruses to the distribution of earthquakes, fractal patterns are inherent in nature. These techniques determine the scaling exponent of the signal or structure in question to indicate the presence or absence of fractal properties (self-similarity). The essence of the technique hides into finding the Fractal Dimension (FD) which proves to be a powerful tool to detect self-similarity. Multifractals, a step further, are sets of intertwined fractals. The real-life fractal patterns that we see hardly scale according to a single scaling exponent, rather there should be multiple scaling laws to capture their growth or variation over time. These spatial and temporal scale variations indicate a multifractal structure of a particular signal. For these more practical cases, Kantelhardt et al. (2002) formulated the MFDFA algorithm. Since its inception, it has been applied in diverse fields starting from turbulence analysis (Telesca & Lovallo, 2011), traffic movements (Shang et al., 2008), blood flow oscillations (Liao & Jan, 2011) to stock exchange (Yuan et al., 2009; Gu et al., 2010), and even for the prognosis of diseases (Dutta et al., 2013; Ghosh et al., 2014).

The analysis of the music signals is done using MATLAB in this chapter (Ihlen, 2012) and for each step an equivalent mathematical representation is given which is taken from the prescription of Kantelhardt et al (2002). The detailed algorithm for MFDFA technique is given in the Methodology chapter (Chapter 2).

It is well-established that naturally evolving geometries and phenomena are rarely characterized by a single scaling ratio- different parts of a system may be scaled differently. That is, the clustering pattern is not uniform over the whole system (Horvatic et al., 2011). Similarly, music has been shown to have non-uniform properties in its movements as well as structural components (Su & Wu, 2006; Jafari et al., 2007; Pareyon, 2011; Das & Das, 2006; Zlatintsi & Maragos, 2012; Rankin et al., 2014; Jennings et al., 2004). The multifractal nature of music signals is well established and its analysis could prove to be an important tool (by taking the entire signal as a time series and considering all of its properties as a whole) in determining the emotional cue in the cross-modal associations (Ghosh et al., 2018).

3.4. EXPERIMENTAL DETAILS

3.4.1. Audio Sample Preparation

Total 6 instrumental clips each of 30 s duration were chosen for our study, given in Table 3.1. Both the signals were normalized to within 0 dB and hence intensity or loudness and attack cue are not being considered. Each of these sound signals was digitized at the sample rate of 16 kHz (for analysis only; for live hearing the sample rate was kept at 44.1 kHz), 16-bit resolution and in stereo channel. The music was presented with the computer-sound system. Logitech Z-4 speakers with very low S/N ratio was used in the measurement room for giving music input to the participants with a volume of 45–60 dB. A representative image of the waveform is given in Fig. 3.3.

3.4.2. Self-Responses of Emotions

The specimen six pieces of music were standardized with a set of emotions varying in the range of joy, sorrow, serenity, anger, heroic, romantic, devotion, anxiety and freedom. Table 3.2 represents a sample response table.

The participants were asked to mark the emotion they deemed appropriate for each corresponding music piece. If they felt more than one emotion suitable for a particular clip, they are asked to mark all of them for the said clip. If the participants were not emoted by any clip, they were asked not to mark and were given an option to categorise the clip's corresponding emotion in the 'Others' category.

Clip No.	Clip Name	Artist	Instruments used
1	Amelie Road Crossing	Yann Tiersen	Accordion
2	Raga Bhairavi	Ustad Amjad Ali Khan	Sarod
3	Toccata and Fugue in d minor	J.S. Bach	Organ
4	Earthquake	Ustad Zakir Hussain	Tabla
5	Hachiko Soundtrack	Jan A.P. Kaczmarek	Piano
6	Water Dewdrops	Pt. Shivkumar Sharma	Santoor

Table 3.1. Details of the music samples used.

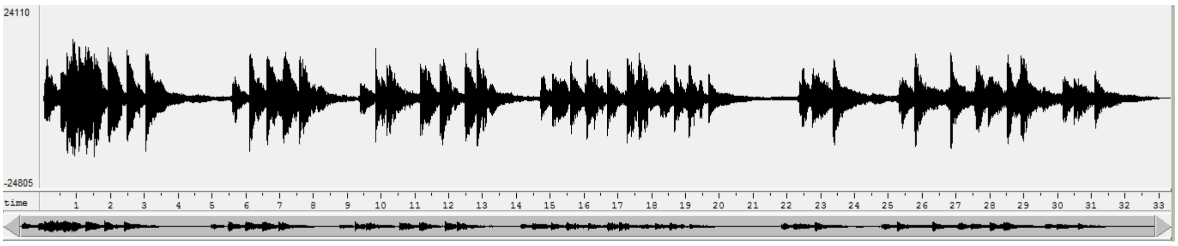


Fig. 3.3. A representative waveform of the music stimulus used, digitized at 16.1 kHz. For live hearing the digitization rate was kept 44.1 kHz.

3.4.3. Color selection interface

The participants were instructed to select the color pattern they seemed appropriate regarding each clip from the software interface given in Fig. 3.2. The circle represents the hue and saturation and the column next to it represents lightness. With the point on the circle and the slider on the column, all the possible color combinations are represented. The selected color is displayed on the large square and it can be changed until the participant makes up his/her mind. Once they finish selecting the color, it can be recorded by clicking the rectangular boxes below it.

3.4.4. Nonlinear chaos-based assessment of music samples

In the last stage, the selected music clips were analyzed using a robust analysis method called Multifractal Detrended Fluctuation Analysis (MFDFA) based on chaos and fractals. The multifractal spectrum identifies the deviations in fractal structure within time periods with large and small fluctuations. The multifractals are fundamentally more complex and describe time series featured by very irregular dynamics, with sudden and intense bursts of high-frequency fluctuations.

EMOTION	Clip 1	Clip2	Clip 3	Clip 4	Clip 5	Clip 6
JOY						
SORROW						
ANGER						
HEROIC						
ROMANTIC						
DEVOTION						
SERENITY						
ANXIETY						
OTHER						

Table 3.2. A representative table of emotion response given to the participants

3.4.5. Participant summary and experimental protocol

93 participants comprising of both male and female people were selected for the test. They belong to varied educational as well as socio-cultural background. The choice included both musician and non-musicians as well. The participants were asked to hear 6 different music pieces (each of 30 second duration).

The type of emotion elicited by different music pieces were identified by the participants from a given collection of possible emotional responses. Then they are asked to assign a color associating the emotion from a given color wheel (structured according to RGB/HSL color space). Each color, associated with a particular music piece, is a mixture of specific Red, Green and Blue values (RGB triplet) and has a specific HEX number (hexadecimal representation), which is recorded for each response. Then, the musical pieces used were further zoomed with the help of fractal technique to identify different emotions related to music in a quantitative measure. Here, to analyze the complexity of the sound signal (which are nonstationary and scale varying in nature), we have used multifractal detrended fluctuation analysis (MFDFA), which is capable of determining multifractal scaling behavior of non-stationary time series. Hence, with the data collected, we can correlate color, emotion and music quantitatively.

The experimental set up was established in a laboratory of Jadavpur University in CV Raman Centre for Physics and Music. The audio clips were played for 30 seconds via a standard music player in a room of ambient temperature. Each clip was played with a span of gap of around 2 minutes and the Color selection interface was preceded by a grey landscape before starting each experiment. The time of the experiments is a normal working hour of any weekday. The participants were expected to be mentally poised, stable while labeling the emotions and choosing the colors so that the emotions can be identified neutrally.

The work in this chapter reports a neuro-cognitive study on response of brain to different stimuli and their associations by utilizing multi-fractal methodology to access the degree of complexity with the help of quantitative parameter. The main interesting findings are summarized in the result and discussion.

3.5. RESULTS AND DISCUSSIONS

As explained in the experimental setup, each participant reacted to the different music stimuli by choosing the emotion(s) they deemed most appropriate in the given table by putting 1 as mark of reaction. A sample response sheet is given in Table 3.3.

The participants, then, recorded the color they thought suits the auditory stimulus the most using the color selection interface. Example of such a randomly chosen color selection panel is given in Fig. 3.4.

EMOTION	Clip 1	Clip2	Clip 3	Clip 4	Clip 5	Clip 6
JOY	1					1
SORROW						
ANGER						
HEROIC				1		
ROMANTIC	1					
DEVOTION						
SERENITY		1			1	
ANXIETY			1	1		
OTHER						

Table 3.3. Choice of emotions marked for different audio clips (for a randomly chosen participant)

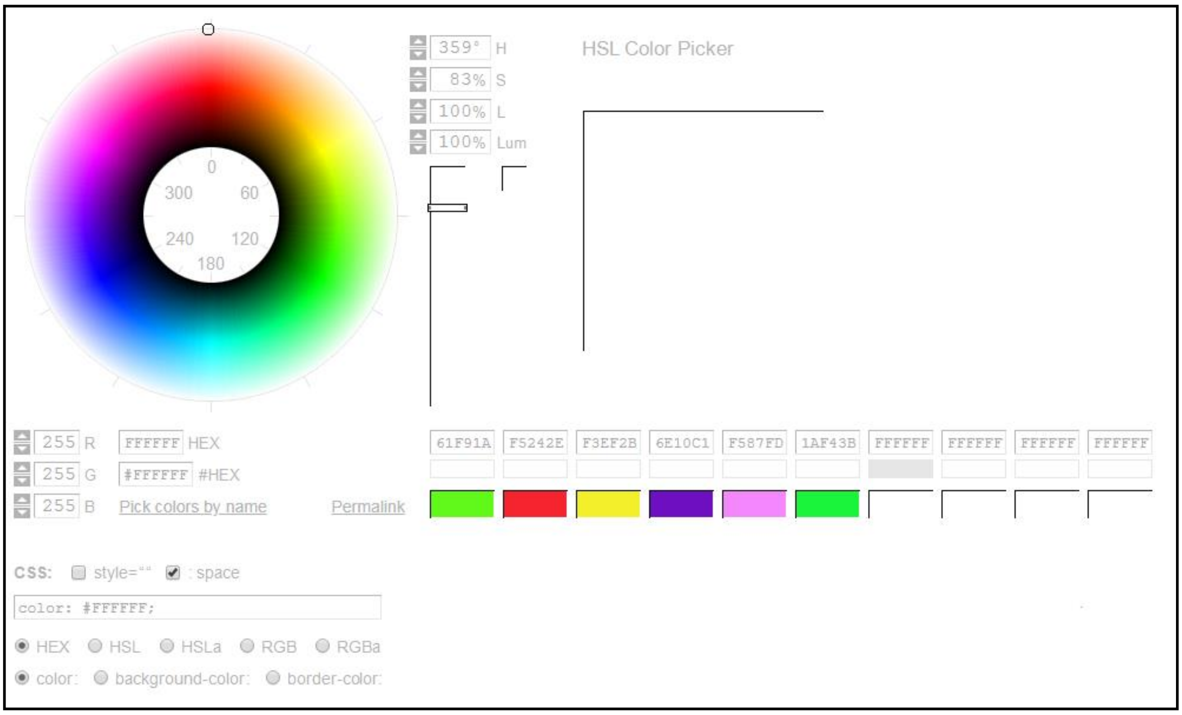


Fig. 3.4. Selection of color according to auditory stimulus (for a randomly chosen participant). Rectangular boxes record the selection corresponding to the color chosen and also indicates their #HEX numbers.

The little rectangular boxes, records the respective responses by specifying the HEX code (Hexadecimal triplets, which represents three separate values that specify the levels of the component primary colors in the chosen compound color, for example- #61F91A in Fig. 3.5):

Fig. 3.5. Hex codes, randomly chosen, corresponding to each selection of colors. They represent all the properties of the color. For example, code FFFFFF represents the color white whose RGB values are R: 255, G: 255 and B: 255.

Each of the recorded color selections are unique and can be dismantled into a unique combination of three primary colors (RGB). The choices of their color indexing the appropriate RGB values are associated with every participant and every clip. The following tables, Table(s) 3.4 and 3.5 show one such specimen data of one out of the 93 dataset. The choice of participant is random, unspecific and regardless to reaction. The percentage of the RGB constituents have also been calculated. It is given in Fig. 3.6.

COLOR	Clip 1	Clip 2	Clip 3	Clip 4	Clip 5	Clip 6
R	52	87	220	240	63	185
G	219	205	139	36	171	246
B	234	244	132	53	238	253

Table 2.4. Values of RGB corresponding to the color chosen for the six experimental audio clips (for a randomly chosen participant)

COLOR	Clip 1	Clip 2	Clip 3	Clip 4	Clip 5	Clip 6
R	0.103	0.16	0.448	0.73	0.13	0.27
G	0.434	0.38	0.283	0.11	0.36	0.36
B	0.463	0.46	0.269	0.16	0.5	0.37

Table 3.5. Proportion of RGB coefficients of the colors for each audio clip, for a random participant

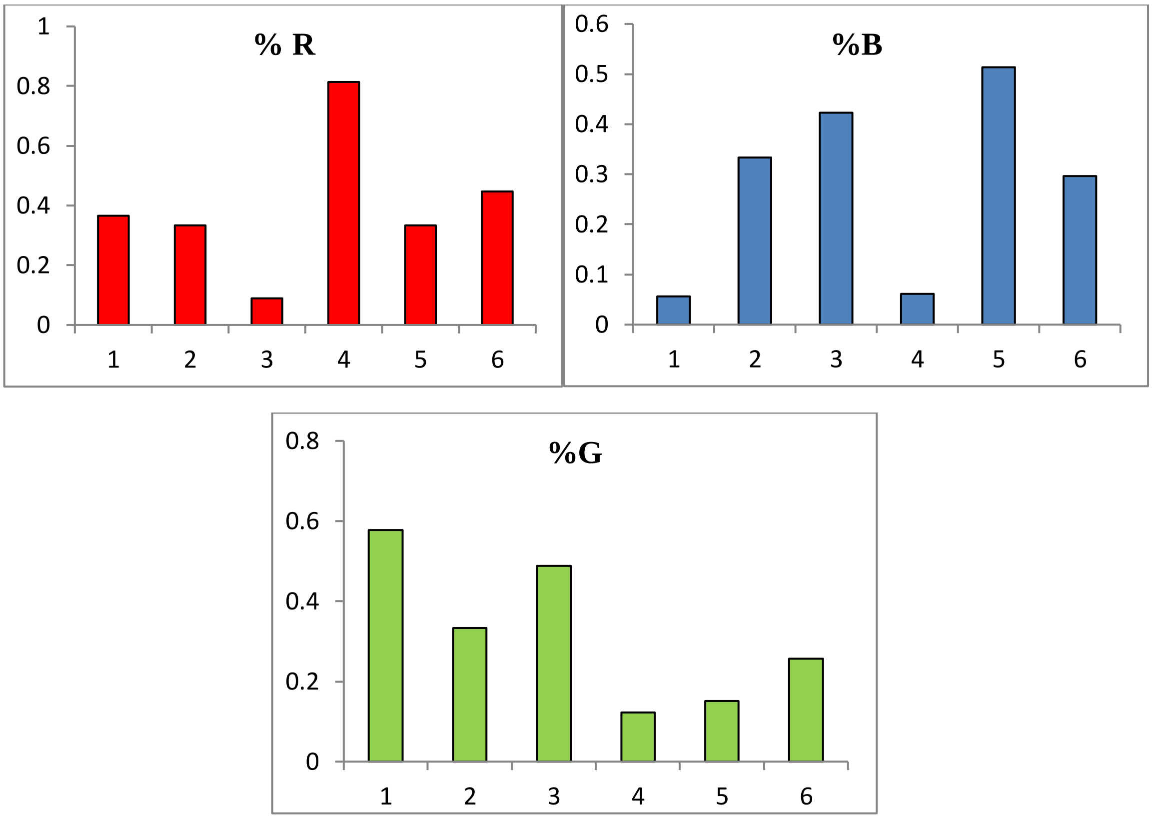


Fig. 3.6. Plot of RGB percentage values in six different audio clips (for the chosen participant)

Clip no.	Joy	Sorrow	Anger	Heroic	Romantic	Devotion	Serenity	Anxiety	Others
1	33.66	0	0	4.33	7.66	0.5	0.83	2.5	1
2	0	14.33	1	0	1	21.16	13.33	1.33	0
3	0.5	0	8.5	16	1.5	2	2.5	20.5	2
4	0	0.83	18.99	4.66	0	1	1	25.32	2.66
5	3.99	2.33	0	2	21.83	0	22.49	0	1
6	14.32	1.5	0	0.5	15	10.82	9.82	0	1

Table 3.6. Weighted average of different emotional responses given by the participants.

The HEX code associated with each chosen color has resulted in the mentioned RGB values as tabulated in Table 3.4. The proportion of their contribution is also identified in Table 3.5. We know that Red, Green and Blue are the three primary colors which can result in any other color by proportionate ratio of mixing. In the further execution, each clip has been marked by the level of their emotion snad associated values of RGB for 93 diferrent participants. It is to be mentioned that the data set reported in the literature is a specimen data set which is acconpanied with 92 more such data sets as the above one whose contribution in totality lead to the tables and figures following.

The weighted average of different emotions has been found out from the responses of the participants, listed below in Table 3.6. It represents how many participants, among the total, have reported the respective emotion corresponding to the said music clip. The more participants chose the particular emotion, the higher its weightage. The series of figures in Fig. 3.7 indicate the clip wise variation of the weighted average as per the labeled emotions.

The visual representation of Table 3.6 is given in the plots of Fig. 3.7.

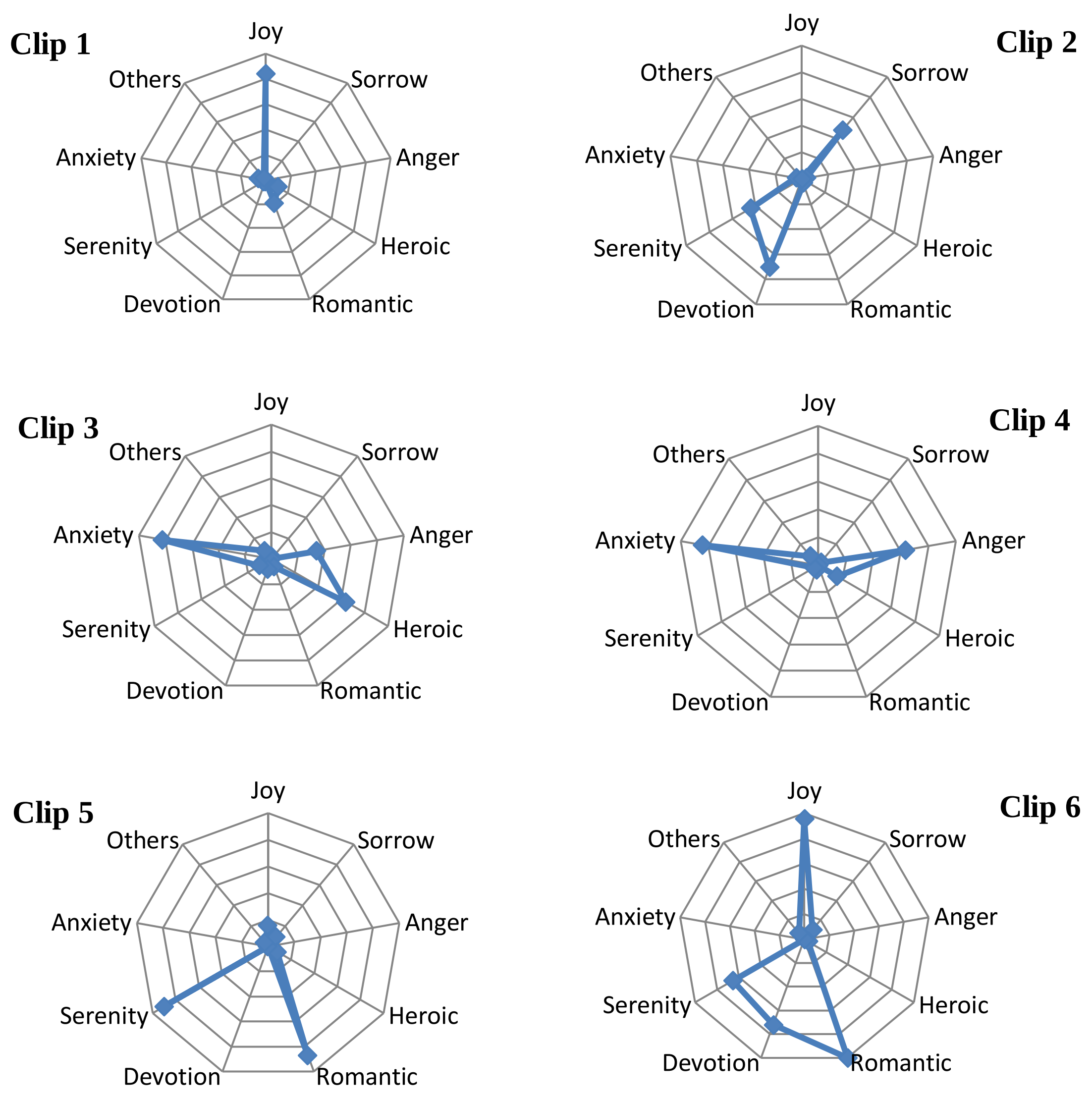


Fig. 3.7. Plots obtained from of the weighted average data of different emotional responses in respective clips. This provides the dominant emotion corresponding the clip.

From Fig. 3.7, the observations that could be made are:

- The dominant emotion in case of Clip 1 is Joy, whereas for Clip 2 it it Devotion.
- Clip 3 and 4 show similar patterns- the dominant emotion is Anxiety, followed by Anger, though the strengths are different (Clip 3 shows a slight inclination towards Heroic as well).
- In Clip 5, Serenity is the dominant emotion, but Romantic emotion also emerges strongly.
- In clip 6, Romantic emotion is the dominant one, followed by Joy, devotion and Serenity.

For the next part of the experiment, where participants chose a color corresponding to each clip, we have calculated the average R,G and B values using the formula in Eq. (1):

$$\mathrm{Ravg} = \frac{\sum(\mathrm{Wi\ X\ Ri})}{\sum \mathrm{S}} \tag{1}$$

Where, R_{avg} = Average R-value for a certain clip;

W_i= Weighted average of the emotion chosen corresponding to the clip by *i*-th participant;

R_i= R-value of the color chosen corresponding to the clip by *i*-th participant; and

$\sum S$ = Total No. of participants

Similarly, B_{avg} and G_{avg} for each clip were computed as well. Finally, the percentage of R, G and B in each clip was found by Eq. (2):

$$\text{Percentage of R} = \frac{\mathrm{Ravg}}{(\mathrm{Ravg+Gavg+Bavg})} \mathrm{X}\ 100 \tag{2}$$

From the tables and datas recorded for the six different music clips for all the participants, the average values of RGB has been calculated clipwise. These datas are tabulated in Table 3.7 and their variations are shown in Fig. 3.8.

Clip no.	R	G	B
1	30.76	39.85	29.39
2	31.17	41.93	26.74
3	41.28	24.31	32.54
4	51.09	20.15	26.91
5	34.09	30.52	34.94
6	26.60	33.72	39.22

Table 3.7. Clip wise average values of RGB. The graphical representation is given in Fig. 3.8.

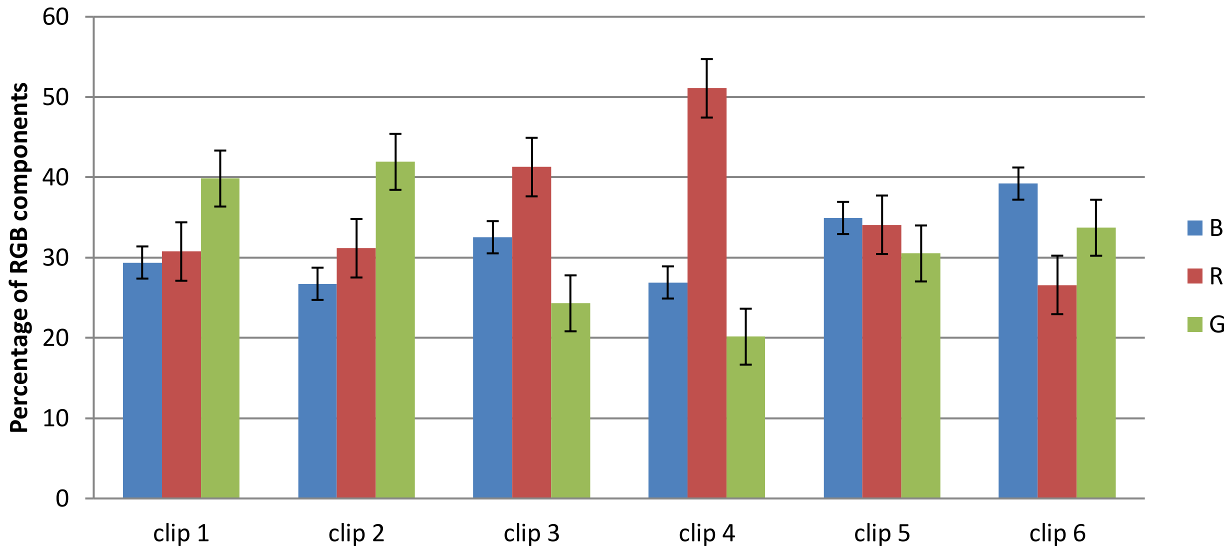


Fig. 3.8. Clip wise variation of average values of R, G, B

From Fig. 3.8, the observations can be summarised as below:

- Both clip 1 and 2 show significantly higher G values (strength is slightly higher in 2)
- R values are distinguishingly higher in clips 3 and 4 than G and B.
- Clip 6 has higher B percentage compared to the other two.
- Interestingly, in clip 5, the distribution is most even. Here, both B and R components are seen to have almost equal weightage, closely followed by G component.

From the previous two figures (Figs. 3.7 and 3.8), some curious pictures emerge. The prevalent emotions in clips 1 and 2 are Joy and Devotion respectively. The RGB analysis in both these cases show a similar trend,i.e., the higher presence of G component than the other two. In clips 3 and 4, the weighted average of emotion 'Anxiety' is higher. The R component of the color choices is seen to be very high for these two clips. Participants reported 'Romantic' emotion is strongest in clip 6. Also, color choices for this clip put B component as the most preferred one. Clip 5, mostly identified with Serenity (closely followed by Joy, Devotion and Romantic), has an even contest of RGBs: R ~ 34%, G ~ 30%, B ~ 35%.

After these results, which reveals a trend involving the emotional categorization and color choices, our aim is to investigate whether a correlation exists between the intrinsic properties of the auditory stimulus and the emotion or its visual correspondance. So, in the next phase of the experiment, i.e., acoustic analysis of the music clips is done with non-linear tools since linear techniques have been found to be insufficient for the complex fractal nature of music time and again. Considering the music piece as a continuous time-series, the multifractal width (w) of the series provides a measure of the degree of complexity of it. Higher (or lower) value of the parameter w indicates more (or less) numbers of local fluctuations present in the temporal scale. This parameter, highly sensitive and predictive in its measurement, can categorize and quantify the auditory signal such a way which can't be done by conventional tools in this domain. The fact that the music signals possess multifractal nature can be confirmed by the Hurst exponent *h(q)*. A monofractal time series is characterized by unique h(q) for all values of q, whereas if the small and large fluctuations scale differently, then the variation h(q) will depend on q – this indicates that the time series is multifractal. Representative figures of h(q) vs q for six music clips are presented in Fig. 3.9.

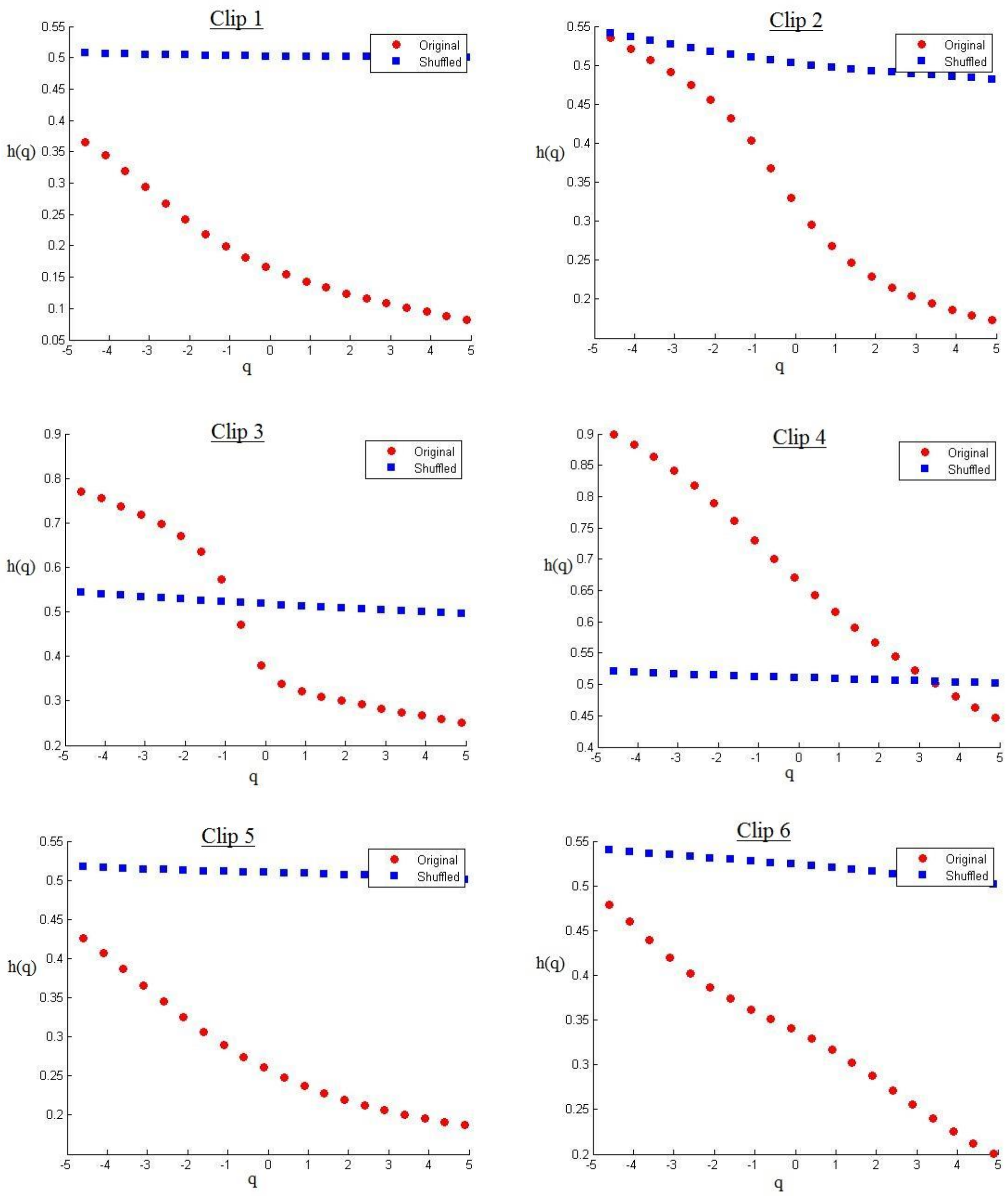
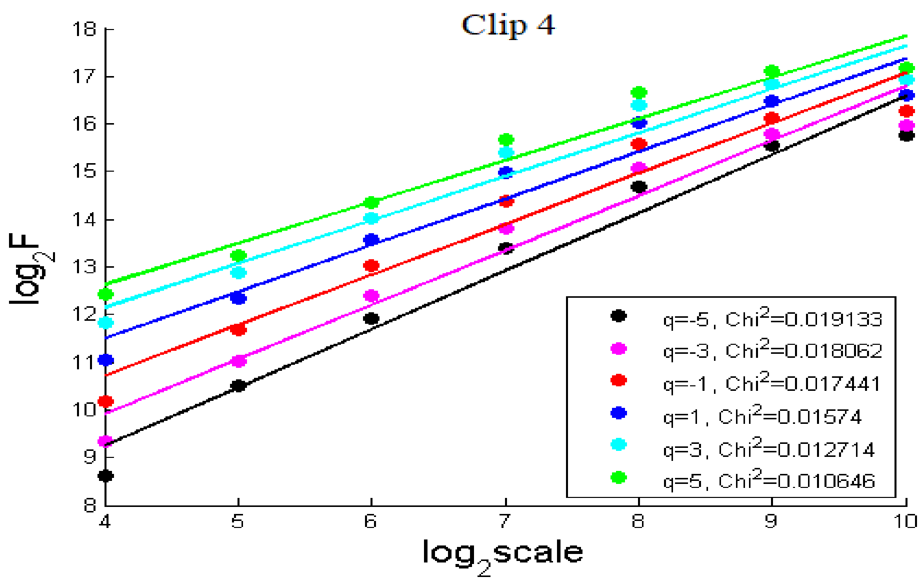


Fig. 3.9. The variation of h(q) vs q for six music samples

In the above figures, along with the variations of h(q), the shuffled values have also been shown. The variation of h(q) with q clearly indicates multifractal nature of the signal. The shuffled values show remarkable difference from the original values, because once the data is shuffled, all the long range correlations break down and multifractality is lost. Hence, as discussed earlier, h(q) takes a consant value with changing q's.

The dependency of h(q) on q is expressed in the q-order RMS fluctuations for positive and negative q values. Fig. 3.10 is a representative figure of q-order RMS (i.e., log F vs log scale graph) is presented for two randomly chosen clips (4 and 6). The scaling range is 16 to 1024 (scales used are: 16, 32, 64, 128, 256, 512, 1024):



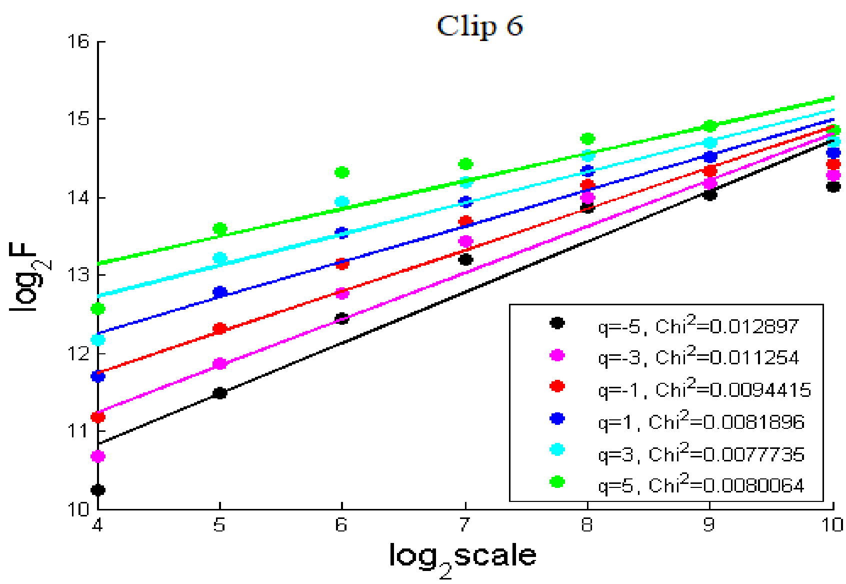


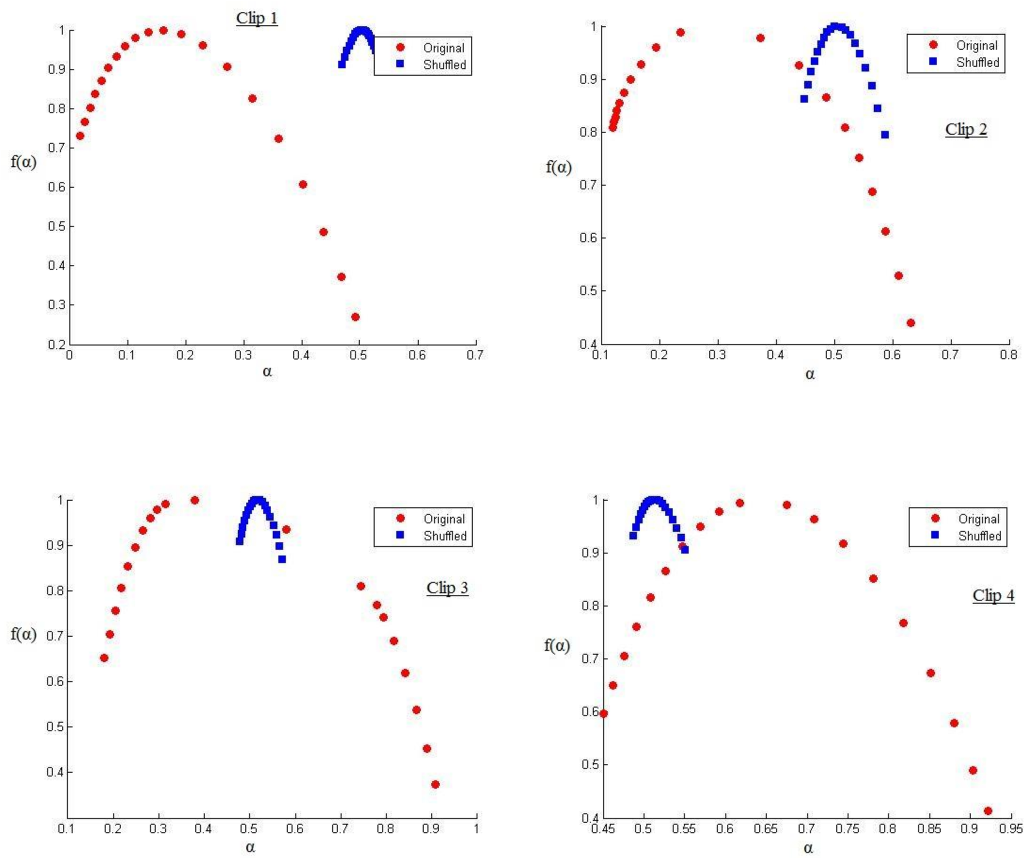
Fig. 3.10. q-order RMS and corresponding regression line for two randomly chosen clips (Clip 4 and Clip 6)

The q-order varies from q = -5 to +5 (During computation, the interval between q-values are taken to be 0.5. whereas, in Fig. 3.10, the interval is increased to +2 to make the graphical representation more facile). The corresponding chi-square values for the q-order RMS for six clips are given in Table 3.8.

Clip No.	q-order values					
	q = -5	q = -3	q = -1	q = +1	q = +3	q = +5
1	0.0077	0.0074	0.0079	0.0071	0.0056	0.0047
2	0.0077	0.0061	0.0064	0.0177	0.0237	0.0249
3	0.0012	0.0008	0.0004	0.0012	0.0012	0.0009
4	0.0191	0.0181	0.0174	0.0157	0.0127	0.0106
5	0.0555	0.0437	0.025	0.0124	0.0052	0.0023
6	0.0129	0.0112	0.0094	0.0082	0.0078	0.008

Table 3.8. Results of Chi-squared test on six music clips for various q's

The amount of multifractality can be determined quantitatively in each of the music signals from the width of the multifractal spectrum (w) from the f(α) vs α curve. Fig. 3.11 denotes the f(α) vs α plots for six clips under study.



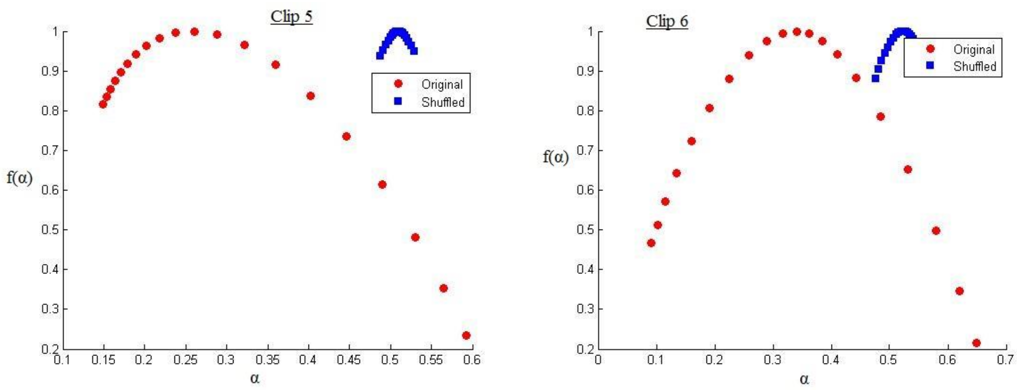


Fig. 3.11. The variation of f(α) vs α for six music samples along with shuffled data. The difference in original vs shuffled data indicates that origin of multifractality is the long-range correlations present in the signal.

The randomly shuffled series clearly shows weaker multifractality. This indicates that the origin of multifractality is due to the long range correlations present in the original signal. Ideally, for a sufficiently long series, the shuffled data would exhibit monofractal properties and f(α) would be of constant value, independent of α.

The different values of the multifractal width corresponding to each clip is given in Table 3.9 and the graphical representation of Table 3.9 is given in Fig. 3.12.

Clip No.	Clip 1	Clip 2	Clip 3	Clip 4	Clip 5	Clip 6
Multifractal spectrum width (w)	0.571 ± 0.032	0.524 ± 0.028	1.031 ± 0.015	0.761 ± 0.013	0.471 ± 0.025	0.383 ± 0.031

Table 3.9. Multifractal width (w) for the six music clips under study

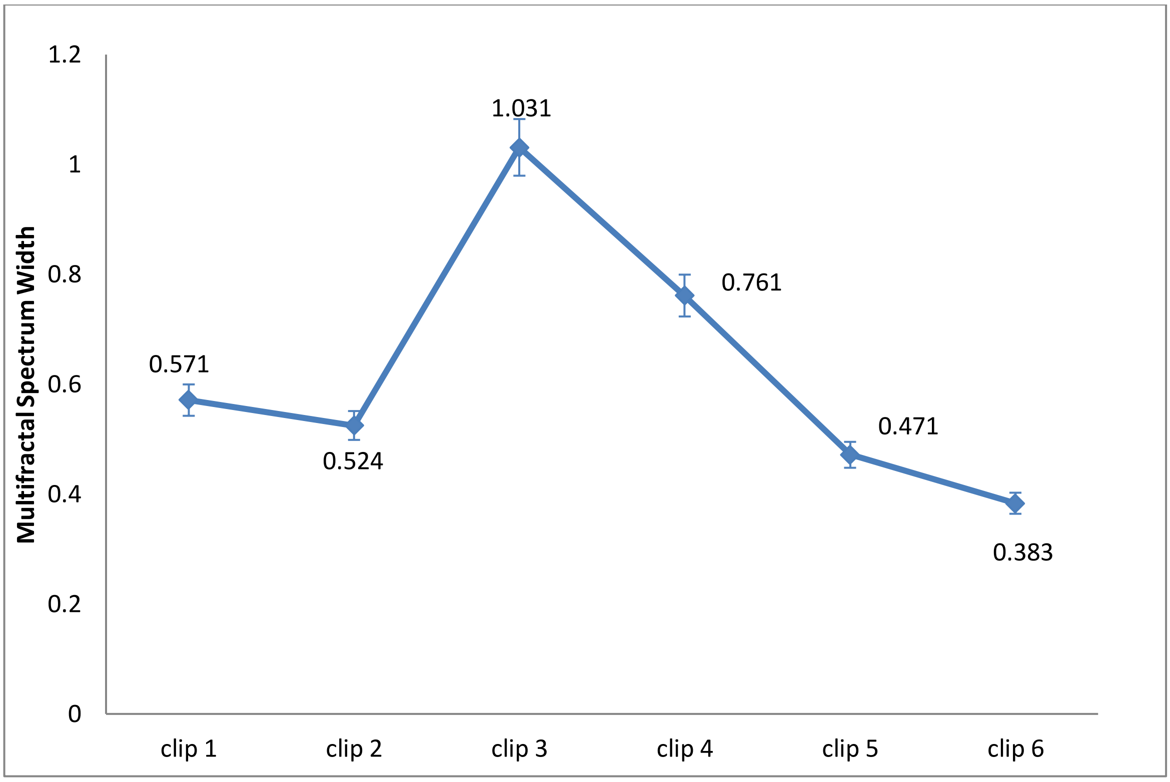


Fig. 3.12. Variation of multifractal width for six music clips. Clips 1 and 2 have similar values of w, whereas clips 3 and 4 have higher multifractal width. Clip 6 has the lowest value. Clip 5 lies between the cluster 1 & 2 and 6.

The first noteworthy observation from the variation of multifractal width is that the first two clips both have similar values of w, indicating similar amount of complexity. From Fig. 3.7, it was found that the dominant emotion corresponding to these two clips were Joy and Devotion, respectively. Fig. 3.8 indicates that the G-component of the color choices made by the participants for these two particular clips are significantly higher. The consistency of the complexity of the auditory signal and the consistency of color choices indicate towards a correspondence of these two sensory modalities. This can be further confirmed by analysing the rest of the pieces too. In clips 3 and 4, the complexity is clearly higher than the rest. In terms of emotions chosen, Anxiety and Anger seem to be the concensus regarding both of the two. The RGB analysis confirms the presence of very high R-component for clips 3 and 4. Here too, the the complexity values and the color choices show a clear trend. Clip 6, emotionally chosen as Romantic, and associated with a high B-component, has the least complexity amongst the six. Another interesting observation to make is the curious case of clip 5. This music sample is identified with emotion Serenity, closely followed by Romantic and the RGB analysis reveals a very close distribution (B > R > G). From Fig. 3.12, the complexity corresponding clip 5 lies in the middle of the zones belonging to clips 1 and 2 (high G) and clip 6 (high B). We re-arranged the clips in ascending order according to their complexity measures. With that, the respective dominant emotional response and RGB component is also added, which should help us look at the final picture as a whole. This is given in Table 3.10.

Clip	Spectrum Width	Dominant Emotion	Dominant RGB component
Clip 6	0.383 ± 0.031	ROMANTIC	B
Clip 5	0.471 ± 0.025	SERENITY	B>R>G (values are close to each other)
Clip 2	0.524 ± 0.028	DEVOTION	G
Clip 1	0.571 ± 0.032	JOY	G
Clip 4	0.761 ± 0.013	ANXIETY	R
Clip 3	1.031 ± 0.015	ANXIETY	R

Table 3.10. Arranging the clips in the ascending order with respect to their signal complexity. The corresponding emotional response and dominant RGB components are also given.

Table 3.10 shows that the complexity of the auditory stimulus has a clear association with the chosen visual counterpart. The higher the complexity rises, the more dominant the R-component seems to become. Lower complexity indicates the abundance of B. Mid-levels of complexity corresponds to higher G-component. Clip 5, interestingly, indicates a transition zone, which is why the RGB preferences are distributed evenly. This possibly indicates the transition phase from where the visual choice starts to shift towards G from B. Now, the question of how this visual-auditory association could be made in human mind arises naturally. The possible answer is hidden in the third column in the above table: the emotions. One can take note of the variation of emotions associated with the clips of higher complexity-Anxiety/Anger (conventionally known as the negative emotions and are always associated with the color RED, having highest R-component) and the emotions associated with lower complexities- Romantic/Serenity/Devotion/Joy (conventionally touted as positive emotions). The shift from emotions with less arousal (traditionally) to the ones with higher arousal (traditionally) according to the change in the auditory stimulus is evident. Hence, the data strongly suggests that the auditory stimulus prompts the visual association via the emotion as a mediator. Changes in stimulus complexity influence the changes in emotional experience and results in the variations of other sensory modalities.

3.6. SURROGATE DATA ANALYSIS USING PHASE RANDOMIZED FOURIER TRANSFORM

To confirm the statistical significance of the multifractal analysis, surrogate data analysis based on phase randomised Fourier Transform is done. The main idea behind the surrogate data analysis is that we want to compare our particular nonlinear metric estimated from the available time series to the distribution of the same metric obtained from a large number of time series (i.e, surrogate data) that satisfying some null hypothesis (Theiler et al., 1992; Schreiber &

Schmitz, 2000). This null hypothesis could be that the time series are generated by some linear stochastic process. Then, we can estimate the probability that our metric is real or obtained purely by chance by observing if and where it falls in the obtained distribution obtained from the surrogates. Then, if the probability of it being true is high enough, we can reject the null hypothesis. Here, we apply Phase randomized surrogates by randomising the phases of the time series while preserving its amplitudes (Raeth & Monetti, 2009). Hence, the surrogate that is generated is the same as the original, in the sense that it retains the dynamics, but is different, since a stochastic element has been written into it from the null hypothesis. This way, after generating a number of surrogates, the same statistical analysis is applied on them. Following which they are ranked along with the statistical analysis of the original series in order of numerical size (Unsworth et al., 2001). If the original series is placed in the main body of ranked statistic, we accept the null hypothesis. On the contrary, we reject the null hypothesis if it falls outside the body on either end.

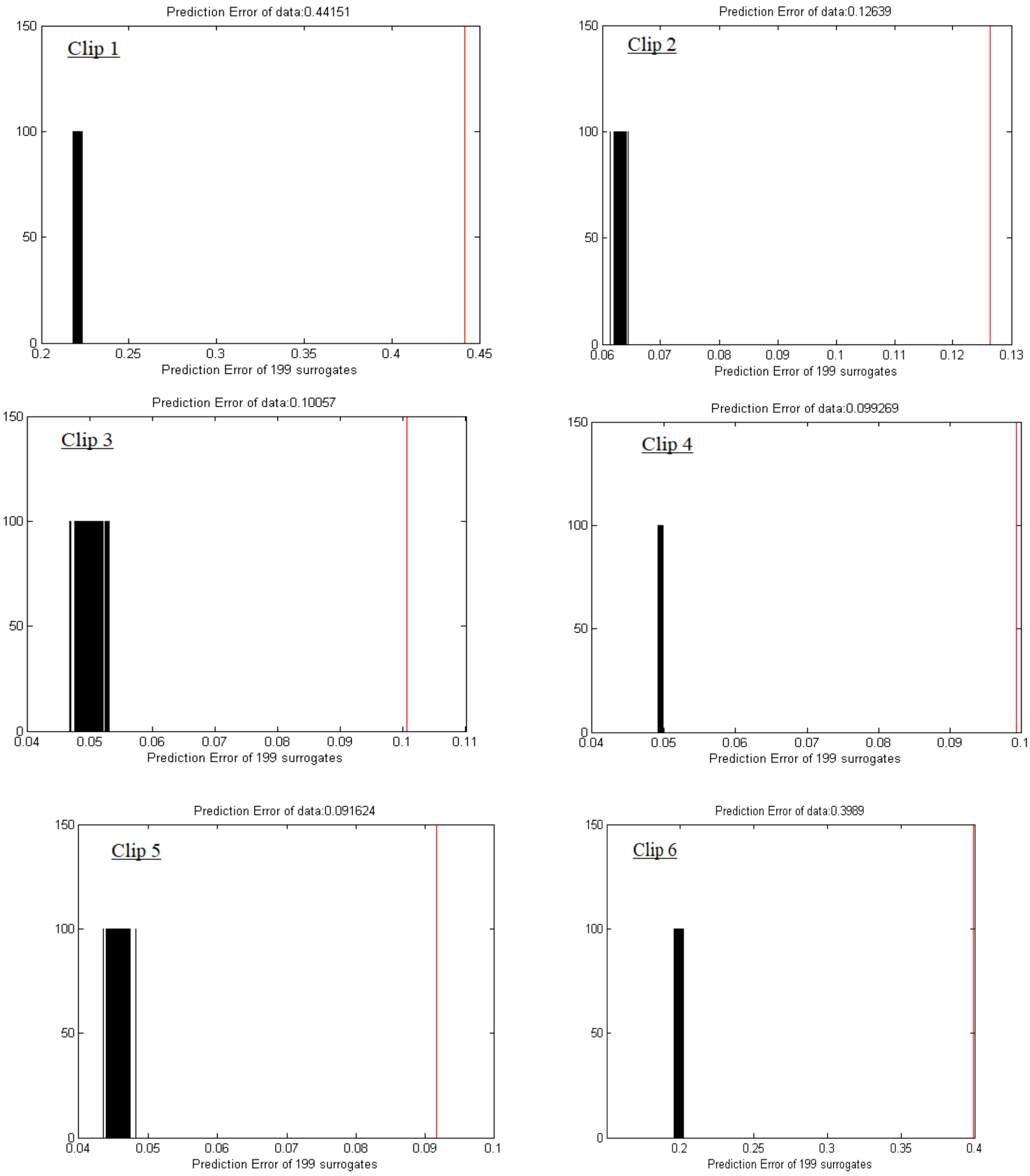


Fig. 3.13. Time series prediction error for six clips versus 199 surrogates

Here, we set the null hypothesis as: the data we work with is a linear stochastic process of unknown distribution. Now, to apply the method of surrogates, we generate 199 surrogates required (for a double sided statistical test with 99% confidence) by the Phase Randomizing scheme. Figure 3.13 shows the prediction error results for each of the clips (y axis being the normalised index used to compare the serieses).

The red line in Figure 3.13 indicates the original statistical operation on the time series and the black lines correspond to the body of synthesized surrogates. Clearly, the observed statistic lie outside the body of distributed statistics of surrogates. Hence, the null hypothesis that the signal was from a linear stochastic process can be rejected with 99% confidence.

3.7. CONCLUSION

Association of visual and auditory modalities is an important aspect of cross-modal information processing. This can propel us further towards the understanding of how the human brain simultaneously respond to and process two separate sensory inputs. Associating color and music can depend on various factors because of the complex nature of both inputs and the way they are perceived. Adding emotion to the mix further makes it difficult since it is very abstract and subjective, too. Previous studies dealt with the problem in a manner that never focused on quantifying how exactly music affects the color choices and why so (However, in a recent paper, Colley & Dean (2019) reports findings which will certainly enrich our understanding on 1/f noise in music which may be a function of small range autocorrelation in the musical structure). In that regard, this study is first of its kind where we try to locate the effect music has on color via emotion mediation using robust non-linear methods. This work provides new data in regard to the categorization and quantification of emotional responses of auditory stimulus and their corresponding color perception. The main outcomes may be summarized as follows:

1. Firstly, even though the color choices of different individuals appear different for a specific auditory stimulus, it is seen that they have an inherent similarity in terms of its structural components. This is interesting since previous studies have highlighted that visual-auditory correspondences are mainly amodal correspondences, i.e., not domain-specific (Zamm et al., 2013). This work shows that in case of modal correspondences, too, there might be possible associations present.

2. The agreement in RGB analysis might have its roots in the fact that the emotional response selection is also agreed upon previously. That is to say, the emotion content in respective clips has led the color preferences accordingly. Thus, the cross-modal correspondence owes a lot of its presence to the perception of emotion that it evokes.

3. Previous works used a limited number of color variations, thus restricting the emotional component considerably. Using a color wheel and array of color hues and saturation, this work, on the other hand, gives freedom to associate any color to emotion.

4. Correlating the auditory signal complexity with that of the visual interpretation (and emotion) is a novel approach. The results show that this approach is promising and could be of profound help in this area of research.

5. The multifractal width, a measure of complexity of the auditory stimulus, has been used as a decisive parameter to categorize and further quantify the emotional response as well as RGB analysis. The multifractal nature of music is the effect of the presence of long-range temporal correlations. These long-range correlations might also be the reason behind the emotional response while music perception.

6. The variation of multifractal width of the clips and their respective emotional responses show a clear trend. It is seen that the lowest complexities churn out the most pleasant emotions like Romantic, Serenity and Devotion. The mid phases give birth to somewhat 'positive' emotions like Joy. Higher ends of the complexity spectrum indicate mostly 'negative' emotions such as Anxiety or Anger. This kind of correspondence between signal complexity and emotional responses has never been reported before.

7. The correlation between multifractal width and the preferred RGB component is also noteworthy. Clips with the lowest complexity are associated with high B values whereas clips having highest complexities are associated with colors having dominant R component. Clips with complexities in the middle region have a high G value in their visual correspondence. While analyzing clip 5, we have also identified a transition phase in the audiovisual correspondence where the signal complexity and color preference falls midway between lower and middle levels and exhibits B to G transition. This way, an attempt was made to quantify a threshold for the change in the emotional spectrum: which was unheard of till now.

Color perception and music perceptions are difficult to understand as it is. On top of that, association of these two sensory modalities can only be more arduous. Since sensory inputs are complex in structure, decoding them might guide us to the path of solving the mysteries of perception. This is the philosophy that we have followed before this work. And thus, using the latest chaos based robust non-linear techniques like the Multifractal Detrended Fluctuation Analysis (MFDFA), we have attempted to correlate the visual and auditory stimulus response with emotion as a mediator. This pioneer study, based on non-linear chaos approach, hopes to show a new direction in the field of researches regarding cross-modal associations, both in synesthetic and non-synesthetic alike.

REFERENCES

Barbiere, J. M., Vidal, A., & Zellner, D. A. (2007). The color of music: Correspondence through emotion. *Empirical studies of the arts*, *25*(2), 193-208.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the national academy of sciences*, *98*(20), 11818-11823.

Bresin, R. (2005, September). What is the color of that music performance?. In *ICMC*.

Colley, I. D., & Dean, R. T. (2019). Origins of 1/f noise in human music performance from short-range autocorrelations related to rhythmic structures. *PloS one*, *14*(5), e0216088.

Cutietta, R. A., & Haggerty, K. J. (1987). A comparative study of color association with music at various age levels. *Journal of research in music education*, *35*(2), 78-91.

D'Andrade, R., & Egan, M. (1974). The colors of emotion 1. *American ethnologist*, *1*(1), 49-63.

Das, A., & Das, P. (2006). Fractal analysis of different eastern and western musical instruments. *Fractals*, *14*(03), 165-170.

Dutta, S., Ghosh, D., & Chatterjee, S. (2013). Multifractal detrended fluctuation analysis of human gait diseases. *Frontiers in physiology*, *4*, 274.

Galton, F. (1883). *Inquiries into human faculty and its development*. Macmillan.

Ghosh, D., Dutta, S., & Chakraborty, S. (2014). Multifractal detrended cross-correlation analysis for epileptic patient in seizure and seizure free status. *Chaos, Solitons & Fractals*, *67*, 1-10.

Ghosh, D., Sengupta, R., Sanyal, S., & Banerjee, A. (2018). *Musicality of human brain through fractal analytics*. Springer Singapore.

Giannakis, K., & Smith, M. (2001). Imaging soundscapes: Identifying cognitive associations between auditory and visual dimensions. *Musical imagery*, 161-179.

González-Espinoza, A., Larralde, H., Martínez-Mekler, G., & Müller, M. (2017). Multiple scaling behaviour and nonlinear traits in music scores. *Royal Society open science*, *4*(12), 171282.

Gu, R., Chen, H., & Wang, Y. (2010). Multifractal analysis on international crude oil markets based on the multifractal detrended fluctuation analysis. *Physica A: Statistical Mechanics and its Applications*, *389*(14), 2805-2815.

Holm, J., Aaltonen, A., & Siirtola, H. (2009). Associating colours with musical genres. *Journal of New Music Research*, *38*(1), 87-100.

Horvatic, D., Stanley, H. E., & Podobnik, B. (2011). Detrended cross-correlation analysis for non-stationary time series with periodic trends. *EPL (Europhysics Letters)*, *94*(1), 18007.

Husain, G., Thompson, W. F., & Schellenberg, E. G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music perception*, *20*(2), 151-171.

Ihlen, E. A. (2012). Introduction to multifractal detrended fluctuation analysis in Matlab. *Frontiers in physiology*, *3*, 141.

Ilie, G., & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception*, *23*(4), 319-330.

Ilie, G., & Thompson, W. F. (2011). Experiential and cognitive changes following seven minutes exposure to music and speech. *Music Perception*, *28*(3), 247-264.

Jafari, G. R., Pedram, P., & Hedayatifar, L. (2007). Long-range correlation and multifractality in Bach's inventions pitches. *Journal of Statistical Mechanics: Theory and Experiment*, *2007*(04), P04012.

Jennings, H. D., Ivanov, P. C., Martins, A. D. M., da Silva, P. C., & Viswanathan, G. M. (2004). Variance fluctuations in nonstationary time series: a comparative study of music genres. *Physica A: Statistical Mechanics and its Applications*, *336*(3-4), 585-594.

Juslin, P. N., & Sloboda, J. A. (2001). Music and emotion. *Theory and research*. Oxford University Press.

Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., & Stanley, H. E. (2002). Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, *316*(1-4), 87-114.

Kaya, N., & Epps, H. H. (2004). Relationship between color and emotion: A study of college students. *College student journal*, *38*(3), 396-405.

Liao, F., & Jan, Y. K. (2011). Using multifractal detrended fluctuation analysis to assess sacral skin blood flow oscillations in people with spinal cord injury. *J Rehabil Res Dev*, *48*(7), 787-800.

Lindborg, P., & Friberg, A. K. (2015). Colour association with music is mediated by emotion: Evidence from an experiment using a CIE Lab interface and interviews. *PloS one*, *10*(12), e0144013.

Lipscomb, S. D., & Kim, E. M. (2004, August). Perceived match between visual parameters and auditory correlates: An experimental multimedia investigation. In *Proceedings of the 8th International Conference on Music Perception and Cognition* (pp. 72-75).

Marks, L. E. (1975). On colored-hearing synesthesia: cross-modal translations of sensory dimensions. *Psychological bulletin*, *82*(3), 303.

Newton, I. (1952). *Opticks, or, a treatise of the reflections, refractions, inflections & colours of light*. Courier Corporation.

Odbert, H. S., Karwoski, T. F., & Eckerson, A. B. (1942). Studies in synesthetic thinking: I. Musical and verbal associations of color and mood. *The journal of general psychology*, *26*(1), 153-173.

Palmer, S. E., Schloss, K. B., Xu, Z., & Prado-León, L. R. (2013). Music–color associations are mediated by emotion. *Proceedings of the National Academy of Sciences*, *110*(22), 8836-8841.

Pareyon, G. (2011). *On Musical Self-Similarity: Intersemiosis as Synecdoche and Analogy*. Gabriel Pareyon.

Pollock, S. (2016). *A Rasa reader: classical Indian aesthetics*. Columbia University Press.

Raeth, C., & Monetti, R. (2009). Surrogates with random Fourier phases. In *Topics on Chaotic Systems: Selected Papers from Chaos 2008 International Conference* (pp. 274-285).

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia--a window into perception, thought and language. *Journal of consciousness studies*, *8*(12), 3-34.

Rankin, S. K., Fink, P. W., & Large, E. W. (2014). Fractal structure enables temporal prediction in music. *The Journal of the Acoustical Society of America*, *136*(4), EL256-EL262.

Roy, C., Roy, S., Sengupta, R., & Ghosh, D. (2015, November 23-24). Categorization & Quantifying Effect of Ragas and Colors on Human Brain—A Neurocognitive Nonlinear Study. *In Proceedings of International Symposium "Frontiers of Research on Speech and Music (FRSM)".* Indian Institute of Technology, Kharagpur, West Bengal, India.

Sagiv, N., & Robertson, L. C. (2005). Synesthesia and the binding problem. *Synesthesia: Perspectives from cognitive neuroscience*, 90-107.

Scherer, K. R. (2005). What are emotions? And how can they be measured?. *Social science information*, *44*(4), 695-729.

Schreiber, T., & Schmitz, A. (2000). Surrogate time series. *Physica D: Nonlinear Phenomena*, *142*(3-4), 346-382.

Shang, P., Lu, Y., & Kamae, S. (2008). Detecting long-range correlations of traffic time series with multifractal detrended fluctuation analysis. *Chaos, Solitons & Fractals*, *36*(1), 82-90.

Spence, C. (2001). Crossmodal attentional capture: A controversy resolved. *Attention, distraction and action: Multiple perspectives on attentional capture*, 231-262.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*(4), 971-995.

Stein, B. E., & Wallace, M. T. (1996). Comparisons of cross-modality integration in midbrain and cortex. *Progress in brain research*, *112*, 289-299.

Su, Z. Y., & Wu, T. (2006). Multifractal analyses of music sequences. *Physica D: Nonlinear Phenomena*, *221*(2), 188-194.

Telesca, L., & Lovallo, M. (2011). Analysis of the time dynamics in wind records by means of multifractal detrended fluctuation analysis and the Fisher–Shannon information plane. *Journal of Statistical Mechanics: Theory and Experiment*, *2011*(07), P07001.

Terwogt, M. M., & Hoeksma, J. B. (1995). Colors and emotions: Preferences and combinations. *The Journal of general psychology*, *122*(1), 5-17.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., & Farmer, J. D. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena*, *58*(1-4), 77-94.

Tuuri, K., & Eerola, T. (2012). Formulating a revised taxonomy for modes of listening. *Journal of new music research*, *41*(2), 137-152.

Unsworth, C. P., Cowper, M. R., McLaughlin, S., & Mulgrew, B. (2001). A new method to detect nonlinearity in a time-series: synthesizing surrogate data using a Kolmogorov–Smirnoff tested, hidden Markov model. *Physica D: Nonlinear Phenomena*, *155*(1-2), 51-68.

Yuan, Y., Zhuang, X. T., & Jin, X. (2009). Measuring multifractality of stock price fluctuation using multifractal detrended fluctuation analysis. *Physica A: Statistical Mechanics and its Applications*, *388*(11), 2189-2197.

Zamm, A., Schlaug, G., Eagleman, D. M., & Loui, P. (2013). Pathways to seeing music: enhanced structural connectivity in colored-music synesthesia. *Neuroimage*, *74*, 359-366.

Zlatintsi, A., & Maragos, P. (2012). Multiscale fractal analysis of musical instrument signals with application to recognition. *IEEE Transactions on Audio, Speech, and Language Processing*, *21*(4), 737-748.

CHAPTER 4

NEURODYNAMICS OF COLOR PERCEPTION: A COMPREHENSIVE NONLINEAR STUDY USING FRACTAL ANALYTICS

"Color is the keyboard, the eyes are the hammers, the soul is the piano with many strings. The artist is the hand that plays, touching one key or another purposively, to cause vibrations in the soul."

Wassily Kandinsky

ABSTRACT

Color perception is a major guiding factor in the evolutionary process of human civilization, but most of the neurological background of the same are yet unknown. This work attempts to address this area with an EEG based neuro-cognitive study on response of brain to different color stimuli. With respect to a Grey baseline seven colors of the VIBGYOR were shown to 16 participants with normal color vision and corresponding EEG signals from different lobes (Frontal, Occipital & Parietal) were recorded. In an attempt to quantify the brain response while watching these colors, the corresponding EEG signals were analyzed using two of the latest state of the art non-linear techniques (MFDFA and MFDXA) of dealing complex time series. MFDFA revealed that for all the participants the spectral width, and hence the complexity of the EEG signals, reaches a maximum while viewing color Blue, followed by colors Red and Green in all the brain lobes. MFDXA, on the other hand, suggests a lower degree of inter and intra lobe correlation while watching the VIBGYOR colors compared to baseline Grey, hinting towards a post processing of visual information. We hope that along with the novelty of methodologies, the unique outcomes of this study may leave a long-term impact in the domain of color perception research.

Keywords: Color perception, EEG, Nonlinear study, MFDFA, MFDXA, VIBGYOR

4.1. INTRODUCTION

From the advent of human civilization, color and perception of color has been intimately involved with it. For survival or evolutionary purposes such as choosing safe foods, finding safe routes to navigate or perception of time during the day, for aesthetic purposes such as variations in artistic expressions of different era, for changing range of emotional experiences to various stimuli, even in the modern world for corporate branding – color reshapes the richness of complex visual information (Hanson, 2012). And that is precisely what both helps and hinders research on the effect of color on humans: the sheer volume of research done on visual than any other sensory modality is due to the fact that our interaction with the world has historically depended more on the vision and processing visual information (Hutmacher, 2019; Pike et al., 2012). And the hindrance stems from the fact that the experience of color is very subjective and to some extent, context dependent (Lotto & Purves, 2002; Elliot & Maier, 2012). Nevertheless, the study of color perception and its effects in human brain is fascinating as well as important because it entertains both practical and theoretical concerns.

The existing literature on this component of visual perception highlights two main aspects: psychological and physiological.

4.1.1. Color and Psychological functioning

The theory that colors can cause psychological arousal dates back to Nineteenth century when Goethe (1810) first mentioned the connection between colors and emotional responses. Since then, with the advancement of science and gradual understanding of the underlying mechanism of light and vision, the focus shifted towards the change in behavioral or psychological manifestation with changing color wavelengths (Nakshian, 1964). But after half a century of research on this domain, no concrete conclusion could be drawn as of yet. For example, a study by Hill & Burton (2005) suggests that the color red can be associated with dominance and aggression in both human and non-humans. Another one associates similar responses in human with black (Frank & Gilovich, 1988). A number of empirical works on association of color with different psychological attributes indicated that colors with longer wavelengths (like red) have enhanced the arousal level of the component in consideration (see Elliot (2015) for a more detailed review): From showing attentional advantage in studies regarding color and selective attention (Buechner et al., 2014) to being a performance enhancing factor in sports (Hill & Burton, 2005; Greenlees et al., 2013; Caldwell & Burger, 2011). Again, despite showing a restricting effect in intellectual performance (Elliot et al., 2007; Shi et al., 2015), Red has been found to enhance attraction when worn by the opposite sex (Elliot & Niesta, 2008; Stephen & McKeegan, 2010). On the other hand, shorter wavelength colors such as Blue increase alertness (Lockley et al., 2006; Vandewalle et al., 2007) and perception of quality and trustworthiness, found in marketing evaluation studies (Lee & Rao, 2010; Labrecque & Milne, 2012). In some studies, conducted on the effect of color on cognitive task performances, shows that Red activates avoidance motivation and enhances performance in detail oriented analytic tasks (Mehta & Zhu, 2009; Elliot et al., 2009). Whereas, color Blue (activates approach motivation) is ideal for creative task performances. Contrary to this, experiments done by Olsen (2010), Bakker et al. (2013) or von Castell et al. (2018) could not replicate such results. It seems that most of the studies use mainly two or three colors (Red, Blue, Green in few) as the experimental setup. Although, very few studies associated Green to calmness (Suk & Irtel, 2010; Hanada,

2018) and orange/yellow to excitement (Ridgway & Myers, 2014; AL-Ayash et al., 2016) but a number of other studies didn't agree (Briki & Hue, 2016; Wilms & Oberfeld, 2018; Costa et al., 2018). The general trend in the literatures available show that the association between Red and excitement is the most reported scenario. The association of calmness/relaxation has been divided majorly between colors Blue and Green. In addition to that, some studies like Labrecque & Milne (2012), AL-Ayash et al. (2016) and Wilms & Oberfeld (2018) reported correlation between high saturation and excitement as well.

4.1.2. Color and Physiological responses

Physiological response to color stimulus is another direction that has been investigated in color perception research. Studies done in this area are mostly motivated by the hypothesis that long-wavelength colors (red/yellow) are more arousing than short-wavelength colors (blue/green) (Valdez & Mehrabian, 1994). Although the fundamental question that remains unanswered in this field is whether the response is direct, i.e., stimulus evokes the response directly without cognitive intermediation or indirect - cognition acts as an intermediary (Kaiser, 1984). Some of the studies used different means like GSR (Galvanic Skin Response), EEG, Heart rate and Respiration, Oxiometry, Blood pressure etc. to measure physiological signals against color stimulus. Among these experimental techniques, EEG or electroencephalograms remain the most used one. Various forms of EEG driven cortical activations have been used so far. In 1958, probably the earliest work describing EEG effects of color, lower prominence of alpha waves under red was reported, indicating higher cortical arousal (Gerard, 1958). Since then, analyzing the changes in alpha waves under different experimental conditions has been the usual form of investigation (Elliot, 2019). Though few of those studies supported the driving hypothesis of red color associating with arousal (Ali, 1972; Shen et al., 1999), majority either disagreed or remained inconclusive (Erwin et al., 1961; Caldwell & Jones, 1985; Mikellides, 1990; Yoto et al., 2007). Apart from red, the color that has been experimented with the most is blue, because of its shorter wavelength. Higher arousal and brain activity during cognitive tasks is found in the presence of blue light as well (Klimesch, 1999; Cabeza & Nyberg, 2000; Baek & Min, 2015). As for the other colors, Orange and yellow had been reported to cause enhanced physiological arousal in some of the studies mentioned (Erwin et al., 1961; AL-Ayash et al., 2016). Also, shorter wavelength color like violet has induced more pronounced arousal than higher wavelength color like green in atleast one study (Nourse & Welch, 1971).

To sum up, research on the physiological response is relatively less and relatively sparse than that on psychological effects of it. And in both cases, finding any conclusive pattern is rather difficult (although in psychological studies the agreement between different results seems a bit more consistent). Also, most of the existing work has focused on the applied part of the problem, as in they have sought to establish relationships between a specific color and a psychological attribute or a behavioral pattern, for practical purposes. Hardly had they cared to explain the underlying reasons behind it (Elliot et al., 2007). Detailed studies on the physiological manifestations (especially that on the brain) will undoubtedly help address the issue. It is more than evident that the number of studies in the field of color induced EEG is inadequate. In view of this, in the present study, we have made an attempt to assess in depth the effect of color stimulus on EEG pattern in humans quantitatively using chaos based novel non-linear methodology.

4.2. EEG, FRACTALITY AND MULTIFRACTALITY

The brain constantly carries out information transfer and processing via the neural system, making it extremely complex. It works through the interactions between large assemblies of neurons in the central nervous system (CNS) and the peripheral neural system. Neurons transfer and process the information via the action potentials and neural firing (also known as spikes). When this kind of electrical activity transfers to the surface of the cortex and to the surface of the scalp, we can record it as the EEG. The properties of the EEG signal are very complex and display qualities such as (Paluš, 1996; Thakor & Tong, 2004):

a) Noisy and stochastic with high degree of randomness

b) Time-varying and non-stationary (for any signal more than ~3.5 s duration)

c) High nonlinearity

Quantifying such a system using linear methods like FFT or power spectral density leads to coarse approximation and overlooking of underlying intricacies. As numerous studies like Pritchard & Duke (1992, 1995), Franka et al (2018) suggest, the highly nonlinear and chaotic dynamics of human brain need to be addressed via tools which are useful for quantifying such a system, namely, Fractals.

Fractals are said to be the visual identity of Chaos. Chaotic systems appear seemingly random and pattern-less on the surface. But when investigated using 'mathematical microscopes' i.e., fractals, shows a hidden order among them. Fractal is a rough or fragmented geometrical object that can be subdivided in parts, each of which is (at least approximately) a reduced-size copy of the whole. They possess some unique properties such as fractional dimension (called fractal dimensions or FD) and scale invariance, indicating that their nature remains same at many different scales. Also, in other words, this is called self-similarity: consisting of parts that are similar to the whole (Mandelbrot, 1983). These distinctive properties of fractals make them ideal to analyse complex systems with greater precision. Fractals are found throughout nature -- in coastlines, seashells, rivers, clouds, snowflakes, musical compositions and even in biological systems - heart rhythms, lungs, blood vessels etc. Fractal geometry has been applied to human brain dynamics for various measures, healthy and non-healthy (Pereda et al., 1998; Eke et al., 2002; Linkenkaer-Hansen et al., 2001; Gong et al., 2003; Esteller et al., 1999). In recent past, the fractal-based analysis method that had been instrumental in addressing the fractal scaling properties and long-range correlations in EEG related studies is Detrended Fluctuation Analysis or DFA. With the help of a scaling exponent, DFA quantifies correlation properties of a signal, indicative of its self-similar nature. Using this method, existence of scale invariance and the investigation on the long-range correlations present in EEG signals was studied successfully, during various cognitive (visual and auditory) tasks (Bhattacharya, 2009; Karkare et al., 2009; Banerjee et al., 2016).

Now, one major shortcoming of DFA is that it only uses single scaling ratio to examine the whole system under observation. Usually, in nature, complex systems feature different scaling patterns in different parts of the system, that is to say, the measure of self-similarity can be of multiple nature. Hence, more often than not, fractal technique with single scaling ratio (also known as 'Monofractals') is not adequate. To study such systems more accurately, one needs to use a more robust technique having multiple scaling ratios. These are 'Multifractals' (Stanley et al., 1999). Analogous to a string made of beads, Multifractals are made up of parts which have their own distinct FDs and hence, it is often expressed in a multifractal spectrum with a unique Spectral Width. To analyse complex natural systems, multifractal DFA or MFDFA has been developed by Kantelhardt et al (2002). Interestingly, it has been found that along with various natural phenomenon (explained later), human physiological signatures are also multifractal: from heartbeat dynamics (Ivanov et al., 1999) to actigraphy (Franca et al., 2019). Also, considerable amount of literature exists that indicate the human brain dynamics exhibits multifractal nature (Suckling et al., 2008; Ihlen & Vereijken, 2010; Zorick & Mandelkern, 2013). In the past few years, we have also found evidence that support the idea of multifractality in human brain and have used the spectral width parameter in the quantification of wide range of cognitive properties to a good success (Maity et al., 2015; Roy et al., 2016; Ghosh et al., 2018; Sanyal et al., 2019).

Another important part of this work is the cross-correlation analysis of the EEG data. In signal processing, cross-correlation is used broadly to provide the quantitative measure of similarity between two time series. This method has been applied in EEG signals as well. But unlike the assumptions in cross-correlation studies until the recent past, the time series in consideration here are non-stationary. To counter the problem, Podobnik and Stanley (2008) proposed Detrended Cross-Correlation Analysis (DCCA) to investigate power-law cross-correlations between two non-stationary time series. Zhou (2008) took a step further and developed the method of Multifractal Detrended Cross-Correlation Analysis or MFDXA, which is a technique that originates from MFDFA and investigates the multifractal features of two cross-correlated signals. It uses a cross correlation coefficient (γ_x) which gives the degree of correlation between two categories of signals. For uncorrelated data, γ_x has a value 1; the lower the value of γ_x the more correlated is the data. Negative value of γ_x signifies very high degree of correlation between the signals, i.e., a large increment in one would more likely to follow a large increment of the other. In recent times, there are multiple cross-correlation studies on EEG signals using DCCA or MFDXA which have argued the existence of power law cross-correlation (Jun & Da-Qing, 2012) and have been instrumental in revealing underlying dynamics in the brain (Ghosh et al., 2014, 2018; Chen et al., 2018).

4.3. OVERVIEW OF THE WORK

The principal aim of this work is to study, with state-of-the-art robust chaos-based non-linear methodologies, the different levels of neuronal complexities that arise in the brain when it receives various colors as a visual stimulus. We took the EEG data of 16 participants while they were exposed to seven colors of VIBGYOR (Violet, Indigo, Blue, Green, Yellow, Orange, Red) in that order; each separated from the next by a neutral color (grey), to set a baseline for comparison. Unlike previous works which have studied mostly two or three colours and their comparisons, our experiment consisted of the whole spectrum of natural colours. For the analysis of the collected EEG data, we have applied two fractal-based non-linear techniques MFDFA and MFDXA. These high precision tools have been proven to work on non-stationary EEG data very accurately to reveal the underlying self-similar patterns and complexity measures by quantifying them via different parameters. MFDFA assesses the degree of complexity present in the signal using multifractal width as a parameter. Higher the width, higher the long-range cross-correlations present in the series, implying higher complexity. MFDXA, on the other hand, measures the degree of how much correlation is present between various inter and intra lobe electrodes in the EEG signals using a cross correlation co-efficient (γ_x). Five Frontal electrodes (F3, F4, F7, F8, Fz), two Parietal (P3, P4) and two Occipital (O1, O2) electrodes were analysed since these areas are mostly reported to be associated with cognition and perceptions of visual stimulus (Ganis et al., 2004; Siok et al., 2009; Spillmann & Werner, 2012). The degrees of complexity corresponding to each color and their respective changes from the baseline are studied. This work, along with fulfilling its primary goal of reporting the changes in brain activity during color perception, also hopes to establish a novel investigatory paradigm in EEG based visual perception studies that will include advanced physical tools to magnify underlying mechanisms beyond the realm of conventional methods.

4.4. MATERIALS AND METHODS

4.4.1. Participant summary

16 participants, age ranging from 20 to 59 (7 females; mean age = 27.51, Standard Deviation = ±5.92), voluntarily took part in the experiment. None of the participants reported any history of neurological or psychiatric diseases (e.g.: epilepsy, anxiety etc.) or colour blindness (confirmed by Ishihara test) and they all had normal/corrected to normal vision. Informed consent about the testing procedure was obtained from each participant according to the ethical guidelines of the Ethical Committee of Jadavpur University. The participants were uninformed about the experimental hypotheses. The experiment was conducted at the Sir C.V. Raman Centre for Physics and Music, Jadavpur University, Kolkata.

4.4.2. Experimental Details

Participants were seated in a comfortable chair with back and elbow rests in a dark room with normal temperature of 25° C. The visual stimulus was displayed in a 21-inch LCD monitor (screen resolution 1920x1080, 24-bit color depth, 75 Hz refresh rate), kept 1.2 meters above ground. The distance between the monitor and the participant's eyes was 91 centimetres. The luminosity of each color and the grey in between was kept constant to factor out brightness issues. The participants were asked to focus on a point marked '+' at the centre of the screen (subtended at 2° visual angle).

The colorimetric values of the colors used as stimuli are shown in Table 4.1. Here, X, Y, and Z display the CIE XYZ tristimulus values according to the 2° CIE 1931 standard observer (Cie, 1932) and columns L^*, C^* and h^* display the lightness, chroma and hue values according to the CIE LCh 1976 system (2007).

The EEG experiments were conducted in the afternoon in a normal temperature room with the participants sitting in a comfortable chair in a normal diet condition. EEG data was acquired with an EEG recording cap with 19 electrodes (Ag/AgCl sintered ring electrodes) placed in the international 10/20 system. Fig 4.1 depicts the positions of the electrodes. Impedances were kept below 5 k Ohms. The EEG recording system (Recorders and Medicare Systems) was operated at 256 samples/sec recording on customized software of RMS. Same reference electrodes, ear electrodes A1 and A2, are used for all the channels. The ear electrodes were linked, and the average of A1 and A2 was used as reference. The forehead electrode, FPz has been used as the ground.

Color	***Hex triplet***	***sRGB [r,g,b]***	***X***	***Y***	***Z***	***L****	***C****	***h****
Violet	#7F00FF	(127, 0, 255)	26.79	11.73	95.44	40.79	127.09	311.63
Indigo	#3F00FF	(63, 0, 255)	20.09	08.28	95.13	34.55	131.23	307.55
Blue	#0000FF	(0, 0, 255)	18.04	7.22	95.03	32.30	133.81	306.29
Green	#00FF00	(0, 255, 0)	35.76	71.52	11.92	87.74	119.78	136.02
Yellow	#FFFF00	(255, 255, 0)	77.00	92.78	13.85	97.14	96.91	102.85
Orange	#FF7F00	(255, 127, 0)	48.84	36.45	4.46	66.86	85.66	59.62
Red	#FF0000	(255, 0, 0)	41.25	21.27	1.93	53.24	104.55	39.99
Grey	#808080	(128, 128, 128)	20.52	21.59	23.50	53.59	0.00	270.00

Table 4.3. Colorimetric values of color stimuli used

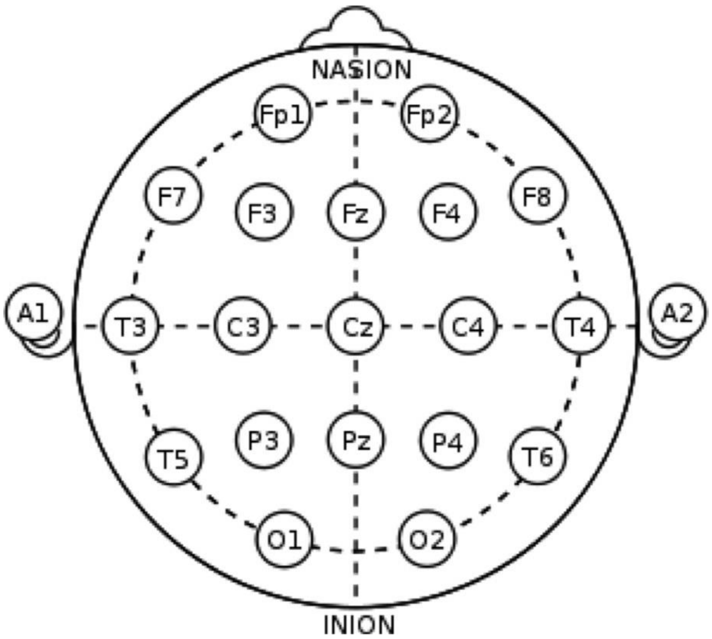


Fig. 4.1. The position of electrodes according to the 10–20 international system. Ear electrodes A1 and A2 are used as references.

4.4.3. Experimental Protocol

The visual stimulus consisted of seven VIBGYOR colors in that order (Violet, Indigo, Blue, Green, Yellow, Orange, Red), each separated from the next by a uniform grey background. The VIBGYOR colors featured for 10 sec durations each and the neutral grey persisted for 60 seconds, intended to neutralise the effect of one color on the others. Before and after the whole experiment protocol, participants were asked to keep their eyes closed for a period of 2 minutes. Order of the protocol has been illustrated in Fig 4.2.

EEG was recorded during the entire protocol, all 13 min. 10 seconds. The obtained EEG data, after cleaning out the noise portions, have been analysed using two non-linear techniques-MFDFA and MFDXA.

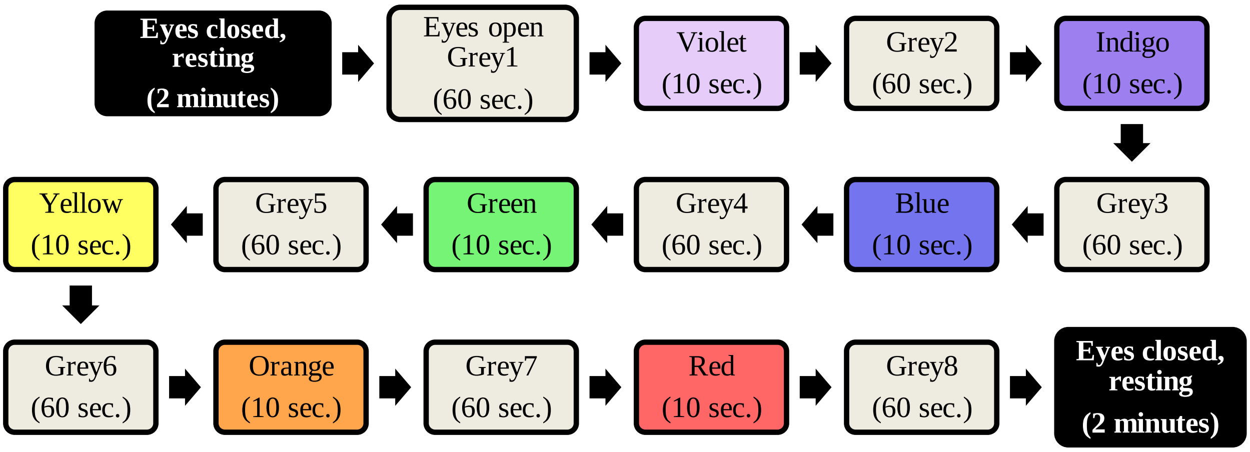


Fig. 4.2. Flow chart of experimental protocol

4.5. METHODOLOGIES

4.5.1. Pre-processing of EEG signals

Raw EEG signals were filtered using a low and high pass filter with cut-off frequencies of 0.5 to 70 Hz. The electrical interference noise (50 Hz) was eliminated using notch filter. High frequency muscle artifacts were removed in the pre-processing stage by selecting the inbuilt EMG filter, which is a second order low pass filter with cut off frequency 35 Hz (35 Hz double-pole). Before proceeding to the analysis stage, the EEG data need to be further cleaned of the low frequency artifacts such as minuscule muscular movements and eye blinks. For this, a method called Empirical Mode Decomposition (EMD) was applied. EMD decomposes the signal into various artifact free components preserving its non-linear and non-stationary features. This processing technique has been detailed in the Methodologies chapter (Chapter 2). The noise-free EEG data, hereafter, will be the main component of further analysis. Signals corresponding to nine electrodes from different lobes of the brain [five Frontal (F3, F4, F7, F8, Fz), two Parietal (P3, P4) and two Occipital (O1, O2)] are obtained and analysed.

4.5.2. Decomposing EEG waveform: Frequency bands vs. broadband EEG signal

The traditional approach in EEG analysis consists of decomposing the original signal into different frequency bands alpha, beta, theta and the likes. The necessity and methodological approach that went into such divisions were constrained by mechanical and computational limitations of 1930s and 40s (see Bladin (2006) for the historical development). Since Fourier Transform, a technique to decompose the signal into its composite frequencies, was available and used successfully in fields of engineering and communications, the practice found its application in EEG analysis as well. This tradition has continued in modern day EEG researches, despite the concern forwarded by one of its inventors in the early days itself (Walter, 1938). Should the EEG signal be a superposition of only the combining waves such analysis would be sensible. But in reality, the signal contains high level of complexity beyond simple associative and distributive intuitions. In fact, the power spectrum scaling of EEG shows 1/f like relations, which is indicative of a complex chaotic system (Pritchard, 1992). Various ‘bumps’ in this frequency structure is used to segregate the delta, theta, alpha, beta and gamma parts of the wave, from lower to higher frequency respectively. Now, as some recent researches suggest, the definition of such bands varies in different studies (Newson & Thiagarajan, 2019). Considerable inconsistency in marking the start and the end of the bands is prevalent, making the results harder to compare. Moreover, the nature of the 1/f noise displays change with factors such as age (Voytek et al., 2015). So, the bands and related frequencies have a huge degree of variability to begin with - which again, resorts to approximation and makes them unreliable neural markers (although a lot of literature and clinical methodologies continue to use the traditional divisions).

Returning to the complexity argument, the oscillatory archetype tends to miss one of the most important factors which needs addressing – the amount of information loss. While breaking the whole signal down into several ‘wave’ parts, one tends to gross out the complex non-stationary nature of the series neglecting important complexity features. Such information deficit, in turn, could end up portraying an incomplete illustration of the system in question. As put eloquently by Thiagarajan (2018), spectral decomposition of a complex EEG signal is analogous to describing an artwork by reducing it to its basic color components and discussing how much red, green or blue it has, while throwing away the data on the pixel’s relative positioning. Though this might churn out occasional accuracy (like higher green and blue is more likely to be a landscape), but to identify the scene more appropriately, one should approach the complex structure as a whole instead of in parts.

Considering these arguments, the entire artifact-free EEG signal was used for complexity analysis in this study.

4.5.3. Multifractal Detrended Fluctuation Analysis (MFDFA)

Originated from Chaos theory, fractal techniques are essential to underline the complex details hidden in an otherwise random or chaotic process. In many natural processes which are chaotic in nature, fractals help to scale the nature of chaos to an accessible level. From the structure of viruses to the distribution of earthquakes, fractal patterns are inherent in nature. These techniques determine the scaling exponent of the signal or structure in question to indicate the presence or absence of fractal properties (self-similarity). The essence of the technique hides into finding the Fractal Dimension (FD) which proves to be a powerful tool to detect self-similarity. Multifractals, a step further, are sets of intertwined fractals. The real-life fractal patterns that we see hardly scale according to a single scaling exponent, rather there should be multiple scaling laws to capture their growth or variation over time. These spatial and temporal scale variations indicate a multifractal structure of a particular signal. For these more practical cases, Kantelhardt et al. (2002) formulated the MFDFA algorithm. Since its inception, it has been applied in diverse fields starting from turbulence analysis (Telesca & Lovallo, 2011), traffic movements (Shang et al., 2008), blood flow oscillations (Liao & Jan, 2011) to stock exchange (Yuan et al., 2009), and prognosis of diseases (Dutta et al., 2013). Also, in a recent work, which involves human participation, using the multifractal nature of acoustic signal we have shown a correlation between preferred emotional states, signal complexity and RGB values of self-reported colors (Roy et al., 2020).

The analysis of the EEG signals is done using MATLAB in this chapter, as described by Ihlen (2012) and for each step an equivalent mathematical representation is given which is taken from the prescription of Kantelhardt et al (2002). The methodology is described in Chapter 2 in greater detail.

The origin of multifractality in an EEG time series can be verified by randomly shuffling the phases in the original data and producing a randomised shuffled series. Most of the long-range correlations that existed in the original data are removed by this random shuffling and what remains is a completely uncorrelated sequence. Hence, if the multifractality of the original data was due to long range correlation, the shuffled data will show non-fractal scaling. To corroborate the findings by comparison, besides phase randomised shuffling a set of surrogate series produced from the original using iAAFT (Iterative Amplitude Adjusted Fourier Transform) method is also used (Schreiber & Schmitz, 1996).

4.5.4. Multifractal Detrended Cross-Correlation Analysis (MFDXA)

MFDXA method was first used by Zhou (2008). It is an offshoot of the generalized MFDFA method and is used to study the degree of correlation between two non-stationary time series having multifractal features. The methodology has been detailed in Chapter 2.

Our desired parameter here is the cross-correlation coefficient γ_x. For uncorrelated data, γ_x has a value 1 and the lower the value of γ and γ_x more correlated is the data. In general, the scaling exponent $\lambda(q)$ depends on q, indicating the presence of multifractality. Using cross-correlation co-efficient γ_x, we want to point out how two non-linear signals are cross-correlated in various time scales.

4.5.5. Methodological approaches

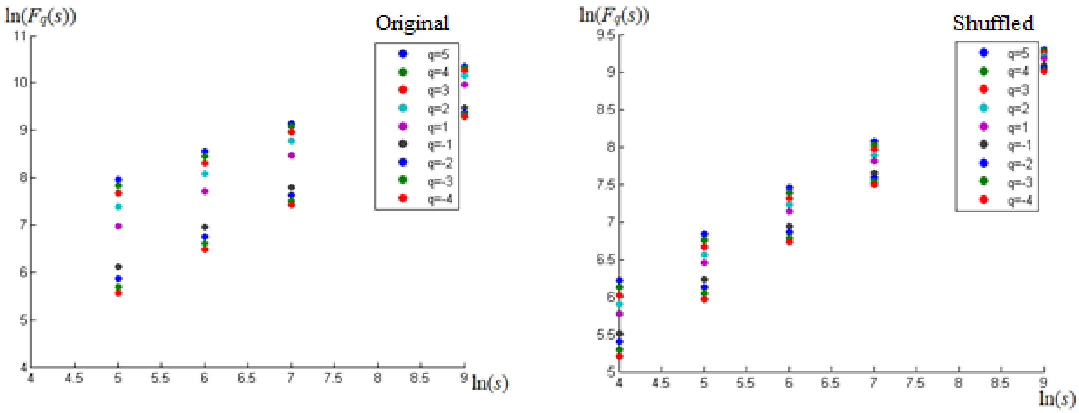
Using the methodologies discussed above, this chapter attempts to explore the neural responses of the participants from two different comparative approaches. To study the brain response change corresponding to each individual color of the VIBGYOR, a comparative analysis of the multifractal spectral width as well as multifractal cross correlation coefficient was done for different pairs of experimental conditions, where each pair consists of a color from VIBGYOR and the adjacent gray just appearing before that particular color (for example, Violet – Grey1 or Green – Grey4). Similarly, to identify the changes among the response from different electrodes corresponding to a particular color, a comparative study of spectral widths and cross-correlation was done for different electrode pairs. Among all the electrode pair combinations some were from the homologous brain regions which in turn reflected the hemispheric differences in the neuronal responses for a particular experimental condition, while the other electrode combinations indicated the nature of connectivity or correlations between different lobes of human brain during viewing a color.

4.6. RESULT AND DISCUSSIONS

For the analysis of the EEG data, we studied 9 electrodes, namely: F3, F4, F7, F8, Fz (Frontal), O1, O2 (Occipital) and P3, P4 (Parietal). Although the areas in the brain that are traditionally related to visual perception are Frontal and Occipital lobes, but researches in the recent past have indicated that Parietal lobe too plays key role in visual information processing (Avillac et al., 2005).

With the noise cleaned EEG signal using EMD process, first we performed the MFDFA methodology mentioned in the previous section. To emphasize on the unique properties pertaining to such a nonstationary nonlinear time series, features of the original time series is compared with a shuffled series constructed by reorganising the phases in the original data in a randomised fashion, in every step of the process (also in case of MFDXA). The qth order fluctuation function $F_q(s)$ for 10 values of q in between −5 and +5 was obtained in the first step. The time series values of the EEG waves have been randomly shuffled to destroy all the long-range correlations present in the data, and what remained is a completely uncorrelated sequence. The regression plot of $\ln(F_q(s))$ vs. $\ln(s)$ averaged for different values of q (q=−5 to q=+5), i.e., the q-order RMS fluctuations, for a sample electrode F4 is given in Fig. 4.3 (in the graph, q = -4 to q = +5 is shown).

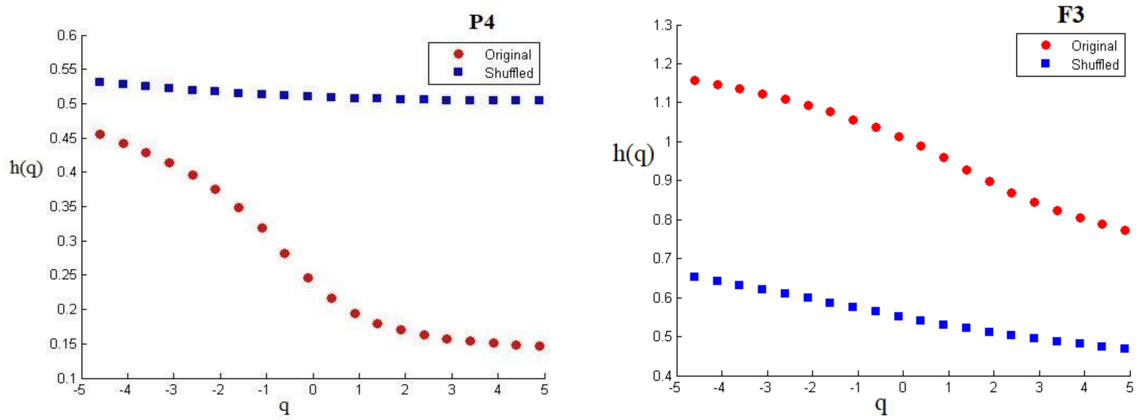
The scaling range is 16 to 1024 (scales used are: 16, 32, 64, 128, 256, 512 and 1024). Again, we see that Hurst exponent h(q) is obtained from the slope of the best fit line in the $\ln(Fq(s))$ vs. $\ln(s)$ plot. It can be seen from Fig. 4.3a that for the original data, the slopes change (the points gradually converge) with changing q's. But they remain same with different values of q, for the shuffled data (Fig. 4.3b). Thus, they have a fixed slope h(q) = H (~2, generally), which is the conventional Hurst exponent for monofractal time series.



(a) $\ln(F_q(s))$ vs. $\ln(s)$ for original F4 data (b) $\ln(F_q(s))$ vs. $\ln(s)$ for shuffled F4 data

Fig. 4.3 (a, b). Plot of $\ln(F_q(s))$ vs. $\ln(s)$ showing different $h(q)$ corresponding to each q; the scaling range being 16 to 1024. For the original series (a), the scaling function F_q and regression slope h(q) is dependent on q, unlike the shuffled series (b) which has a fixed slope H, indicating

Fig. 4.4 contains two representative figures of h(q) vs q plots from two randomly chosen electrodes F3 and P4. It is evident from the figures that for original series (in red), h(q) decreases with increasing q, showing multifractal scaling in both the electrodes. The shuffled series (in blue), very similar to a monofractal signal, has almost constant values of h(q) for different q's.



(a) P4 (b) F3

Fig. 4.4. The variations of h(q) vs. q for (a) P4 and (b) F3 electrodes. For the original series (red dots) h(q) changes with q (higher for -ve q, lower for +ve). The shuffled series (blue dots) has almost a constant h(q) ~ 0.5. Variance of h(q) shows the presence of long-range correlations in the original series.

The EEG time series corresponding to each experimental condition was shuffled and the hurst values were calculated and compared with the same for the original series. The results suggest that in the original time series, the presence of multifractality can be observed. Next, to reassure the same findings, we calculated 6 phase randomized surrogates for each EEG time series segments (corresponding to each electrode during each experimental condition) following the iAAFT method and calculated the hurst values for all the surrogate series to compare them with the original series. The results clearly demonstrate that hurst values of the surrogate series were much lower than that of the original ones (≤ 0.5). This indicates that the multifractality present in the original signal gets destroyed when the surrogates are generated through phase randomization. Two such sample plots are given in Fig.4.5a (F3, Blue) and Fig. 4.5b (O2, Yellow) for two random participants.

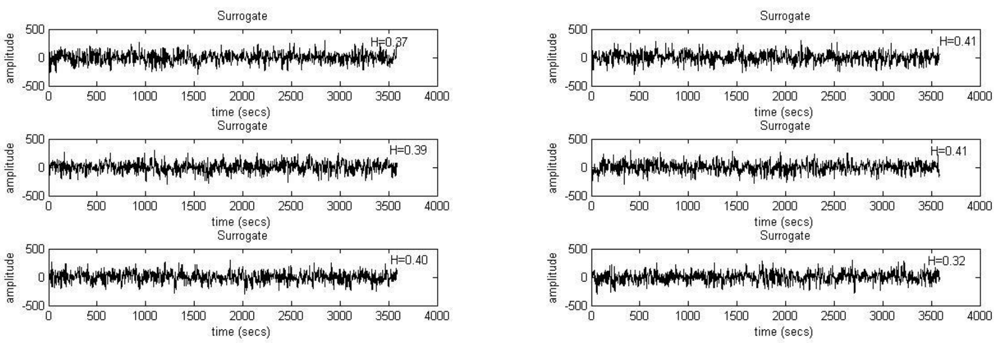


Fig. 4.5a. Six surrogates generated for F3 electrode under Blue color exposure for a random participant using iAAFT. The H stands for respective Hurst values which are far lower than 0.5, pointing at loss of multifractal properties when the data is shuffled randomly

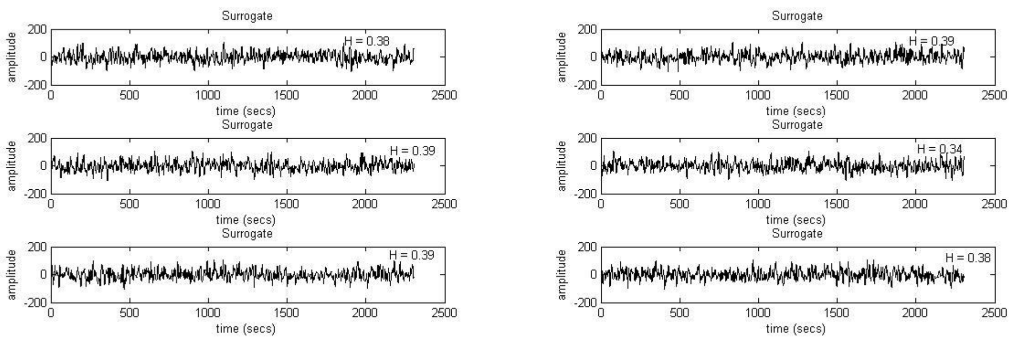
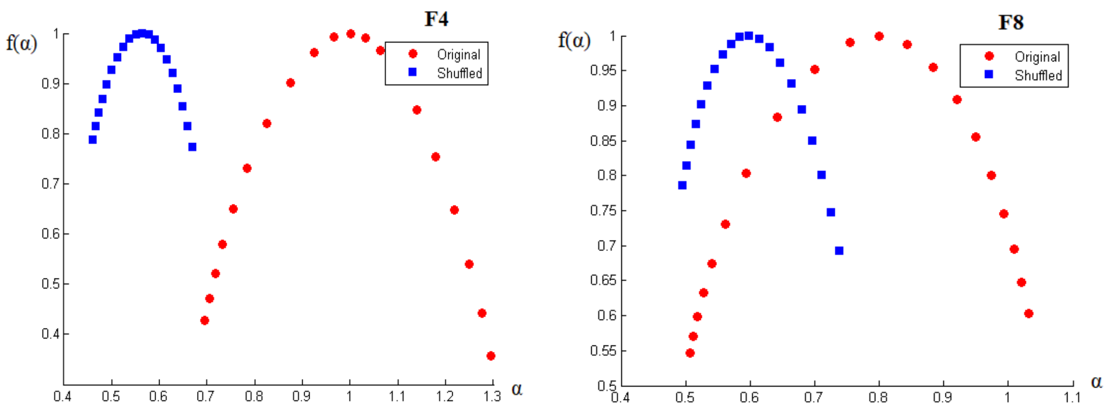


Fig. 4.5b. Six surrogates generated for O2 electrode under Yellow color exposure for a random participant using iAAFT. H stands for respective Hurst values which are far lower than 0.5, pointing at the loss of multifractal properties when the data is shuffled randomly

Now, after confirming the presence, next step is the quantitative analysis of multifractality. The amount of multifractality can be determined for each of the experimental window for every signal from the width of the multifractal spectrum (w) from the f(α) vs. α curve. Two representative figures of such curves for two electrodes along with their shuffled series are shown in Fig. 4.6. As seen from the equation, the nature of the curve is parabolic. Also, the shuffled width is found to be smaller than the width of the original signal; which tells us that the long-range correlations are present in the signal which gives rise to the multifractality. Ideally, for a sufficiently long series, the shuffled data exhibit monofractal properties (no mulifractal scaling) and f(α) would be of constant value, independent of α. Thus, as discussed previously, Hurst exponent remains independent of q and in the f(α) vs. α curve, the shuffled width has a constant f(α) peaked around $\alpha_0 \sim 0.5$.



(a) F4

(b) F8

Fig. 4.6 (a,b). The variation of f(α) vs. α for two randomly chosen electrodes (a) F4 and (b) F8, along with shuffled data. This parabolic curve represents the multifractal spectrum where α is *q*-order singularity exponent and f(α) are its dimensions. $\Delta\alpha$ (= $\alpha_{max} - \alpha_{min}$) is the spectral width, which, for original series (red) is higher than its shuffled form (blue) - indicates multifractal nature

The values of the spectral width of EEG data corresponding to each electrode for all the experimental conditions were computed. Spectral width values and their standard deviations are averaged for 16 participants. The data are presented both color-wise and electrode-wise in Tables 4.2. (a,b), for better interpretation purposes.

	Multifractal width (w) corresponding to the electrodes with SD								
Colors	**F3**	**F4**	**F7**	**F8**	**Fz**	**O1**	**O2**	**P3**	**P4**
Violet	0.516 ±0.048	0.508 ±0.038	0.511 ±0.048	0.505 ±0.029	0.511 ±0.018	0.518 ±0.038	0.508 ±0.016	0.506 ±0.017	0.512 ±0.037
Indigo	0.501 ±0.050	0.492 ±0.036	0.494 ±0.015	0.501 ±0.051	0.493 ±0.024	0.497 ±0.006	0.502 ±0.006	0.501 ±0.018	0.498 ±0.020
Blue	0.648 ±0.049	0.654 ±0.044	0.651 ±0.038	0.669 ±0.050	0.622 ±0.025	0.645 ±0.012	0.682 ±0.024	0.636 ±0.033	0.664 ±0.020
Green	0.559 ±0.038	0.547 ±0.049	0.550 ±0.039	0.563 ±0.031	0.539 ±0.015	0.547 ±0.034	0.558 ±0.014	0.547 ±0.013	0.565 ±0.021
Yellow	0.492 ±0.033	0.507 ±0.014	0.504 ±0.028	0.508 ±0.026	0.499 ±0.011	0.509 ±0.014	0.526 ±0.034	0.499 ±0.049	0.511 ±0.040
Orange	0.542 ±0.048	0.542 ±0.026	0.534 ±0.039	0.526 ±0.019	0.521 ±0.033	0.544 ±0.014	0.547 ±0.030	0.554 ±0.038	0.548 ±0.028
Red	0.585 ±0.022	0.587 ± 0.026	0.580 ±0.026	0.606 ±0.025	0.592 ±0.020	0.585 ±0.027	0.589 ±0.026	0.583 ±0.024	0.598 ±0.025

Table 4.2a. Multifractal Spectral width for different electrodes (Color-wise); averaged over n=16

Multifractal width (w) corresponding to the colors with SD

	F3	F4	F7	F8	Fz	O1	O2	P3	P4
Grey1	0.364 ±0.038	0.350 ±0.032	0.360 ±0.046	0.360 ±0.048	0.378 ±0.010	0.387 ±0.15	0.385 ±0.010	0.374 ±0.035	0.404 ±0.029
Violet	0.516 ±0.048	0.508 ±0.038	0.511 ±0.048	0.505 ±0.029	0.511 ±0.018	0.518 ±0.038	0.508 ±0.016	0.506 ±0.017	0.512 ±0.037
Grey2	0.393 ±0.046	0.347 ±0.044	0.397 ±0.041	0.363 ±0.050	0.403 ±0.040	0.389 ±0.013	0.389 ±0.033	0.382 ±0.022	0.408 ±0.038
Indigo	0.501 ±0.050	0.492 ±0.036	0.494 ±0.015	0.501 ±0.051	0.493 ±0.024	0.497 ±0.006	0.502 ±0.006	0.501 ±0.018	0.498 ±0.020
Grey3	0.351 ±0.049	0.371 ±0.048	0.359 ±0.026	0.340 ±0.047	0.375 ±0.016	0.381 ±0.026	0.384 ±0.015	0.360 ±0.038	0.403 ±0.046
Blue	0.648 ±0.049	0.654 ±0.044	0.651 ±0.038	0.669 ±0.050	0.622 ±0.025	0.645 ±0.012	0.682 ±0.024	0.636 ±0.033	0.664 ±0.020
Grey4	0.371 ±0.017	0.380 ±0.026	0.391 ±0.028	0.391 ±0.037	0.376 ±0.038	0.398 ±0.009	0.397 ±0.029	0.387 ±0.027	0.406 ±0.031
Green	0.559 ±0.038	0.547 ±0.049	0.550 ±0.039	0.563 ±0.031	0.539 ±0.015	0.547 ±0.034	0.558 ±0.014	0.547 ±0.013	0.565 ±0.021
Grey5	0.381 ±0.043	0.376 ±0.038	0.383 ±0.057	0.381 ±0.030	0.385 ±0.022	0.404 ±0.036	0.428 ±0.039	0.394 ±0.021	0.415 ±0.044
Yellow	0.492 ±0.033	0.507 ±0.014	0.504 ±0.028	0.508 ±0.026	0.499 ±0.011	0.509 ±0.014	0.526 ±0.034	0.499 ±0.049	0.511 ±0.040
Grey6	0.373 ±0.037	0.402 ±0.036	0.392 ±0.046	0.397 ±0.046	0.397 ±0.006	0.412 ±0.004	0.408 ±0.020	0.399 ±0.015	0.413 ±0.048
Orange	0.542 ±0.048	0.542 ±0.026	0.534 ±0.039	0.526 ±0.019	0.521 ±0.033	0.544 ±0.014	0.547 ±0.030	0.554 ±0.038	0.548 ±0.028
Grey7	0.403 ±0.049	0.434 ±0.041	0.390 ±0.018	0.412 ±0.021	0.415 ±0.009	0.416 ±0.010	0.416 ±0.012	0.408 ±0.030	0.434 ±0.019
Red	0.585 ±0.022	0.587 ±0.026	0.580 ±0.026	0.606 ±0.025	0.592 ±0.020	0.585 ±0.027	0.589 ±0.026	0.583 ±0.024	0.598 ±0.025
Grey8	0.389 ±0.028	0.406 ±0.039	0.411 ±0.031	0.384 ±0.018	0.422 ±0.012	0.428 ±0.007	0.433 ±0.019	0.412 ±0.010	0.441 ±0.051

Table 4.2b. Multifractal Spectral width for different colors (Electrode-wise); averaged over n=16

The graphical representation of the tables 4.2a and 4.2b is given in Figs. 4.8 and 4.7, respectively.

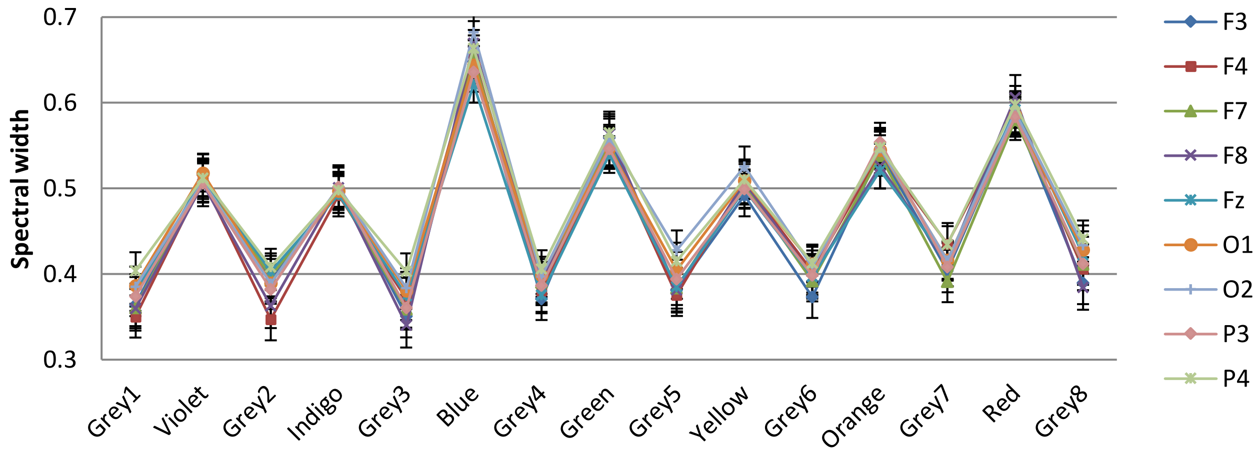
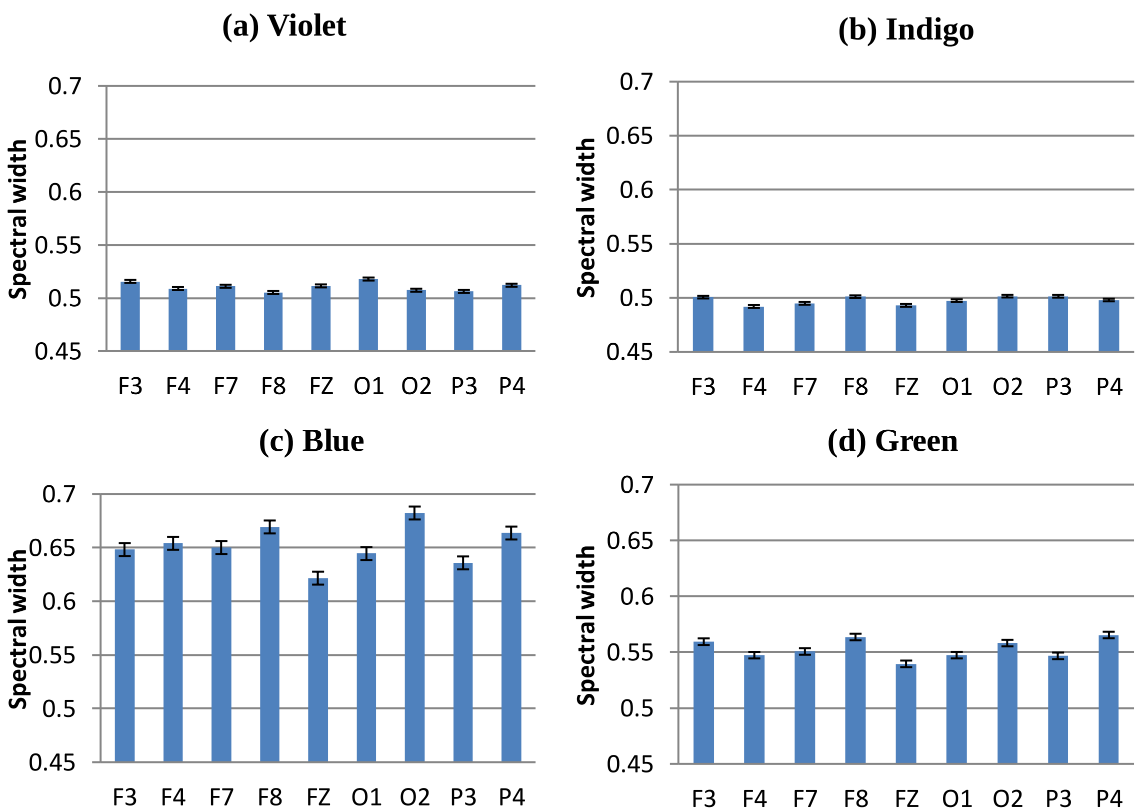


Fig. 4.7. Graphical representation of Table 4.2b. Variation in Multifractal spectral width due to different colors in nine chosen electrodes: the complexity pattern stays similar across electrodes. Three primary colors Blue, Green and Red exhibit highest multifractal widths, baseline Gray stays almost the same. Shorter wavelengths have higher width than longer ones.



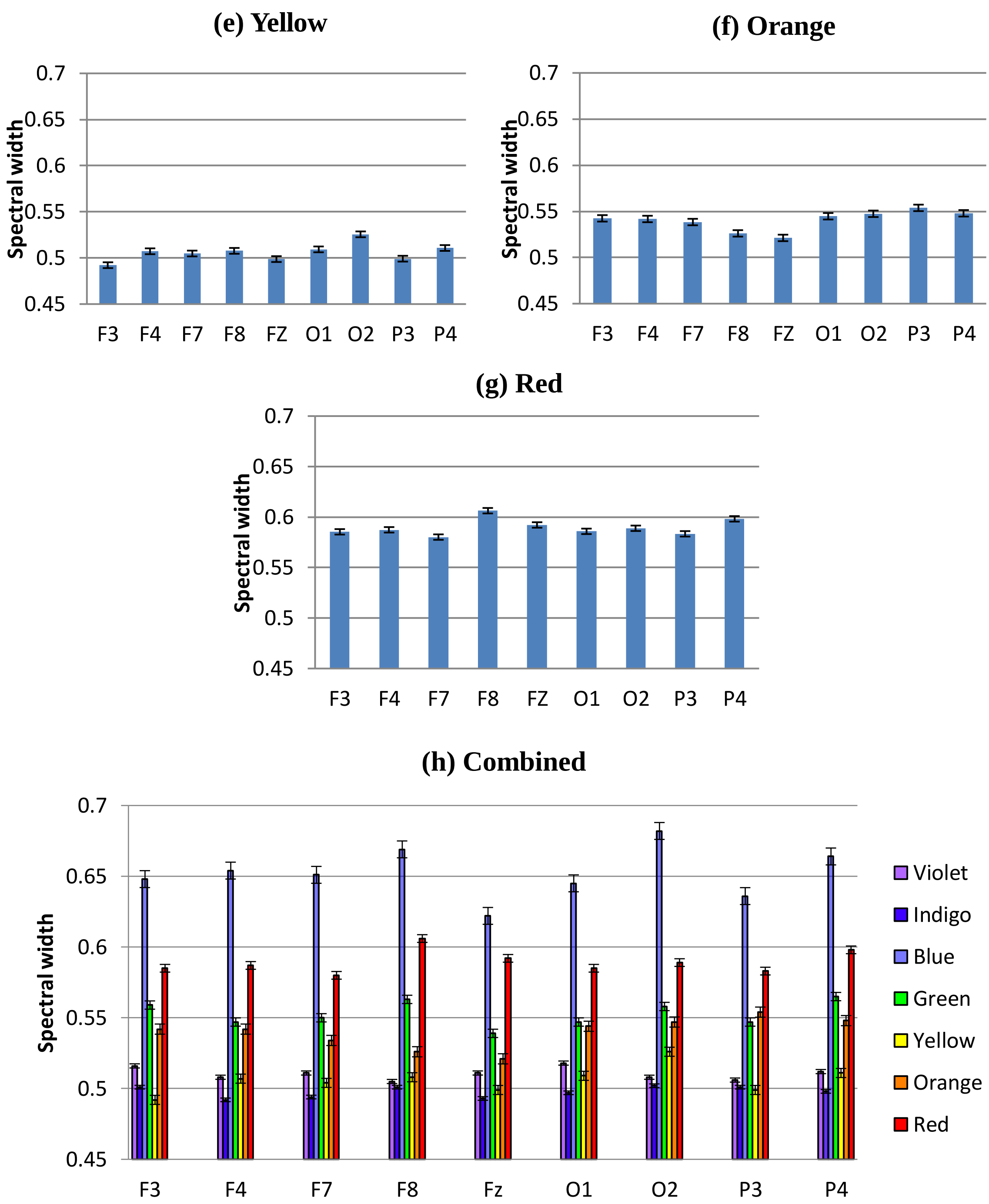


Fig. 4.8(a-h). Graphical representation of Table 4.2a. Variation in average Multifractal Spectral width in different electrodes, color-wise: Blue-Green end of the spectrum (c-d) has higher width than Orange-Red end (f-g), Indigo (b) and Yellow (e) being the lowest. Even electrodes O2, F8 and P4 register highest width (h), suggestive of right brain arousal preference

Fig. 4.7 describes the electrode-wise changes due to various color stimuli. It is clearly evident from the figure that the change of multifractality is present in all the electrodes under consideration. The usual participation of Frontal and Occipital lobes in visual perception is confirmed, with the addition that Parietal lobes, to exhibit appreciable changes in arousal level. All electrodes, interestingly, exhibit similar pattern, too – multifractal width increases with the exposure to VIBGYOR colors and decrease with Grey baseline. This behavior is consistent throughout the electrodes, during the whole experimental procedure. Another major trend which is similar across all 16 participants is the nature of variation in spectral width in presence of some specific colors. The value of spectral width is maximum in case of Blue, followed by Red and then Green. This trend has also been exhibited by all of the electrodes in general, which is a unique and important observation. In fact, the shorter wavelength part of the spectrum has higher width than the longer wavelength end, indicating that changes in signal complexity are more pronounced in former than the latter (to demonstrate this, the difference between the width during colors and the width during baseline Grey is necessary, which has been done later). This observation is novel since previous studies involved with colors are done to correlate them with some specific psychological attributes using task based experimental setup and no study, specially not one with non-linear techniques, has compared the colors solely on their influence on brain activity. From this perspective, the fact that color Blue induces the highest long-range correlations followed by Red and then Green might help explain some of the results obtained in prior researches. That being said, it is noteworthy that, three of the primary colors exhibit clearer complexity changes than the other colors in the spectrum. This supports the reasoning behind the usage of Red, Blue and Green in most of the studies in this field.

Lastly, it can be seen that the complexity for baseline Grey over the experiment has changed as well, although not that prominent. This could be because of the existence of the color exactly prior to it, since the long-range correlations present during that may not have perished completely, i.e., a residual effect might be present. Future experiments, focusing on this very aspect, are needed for detailed explanations.

Fig. 4.8, representing the information in Table 4.2a, shows multifractal width variations in different electrodes, color-wise. The observations made from the previous figure are more evident here. Complexity changes in shorter wavelengths like Blue or Green is higher than Red (and Orange). Width is lowest in case of Indigo and Yellow, whereas Violet and Orange is comparatively close. Another noteworthy observation is, in most cases (considering absolute values of multifractal width) F8, O2 and P4 has the highest complexity among the Frontal, Occipital and Parietal electrodes, respectively. So, the even electrodes show higher complexity than odd electrodes, which is an indicator that in our experimental setup, the long-range correlations found during color perceptions are higher in the right hemisphere in the brain.

To compare this result with the pattern for change in complexity (i.e., the complexity value of the grey baseline subtracted from the absolute complexity value corresponding to the immediately next color), a graph similar to Fig 4.8(h) has been computed. It is given in Fig. 4.9.

Change in complexity (Color - Grey)

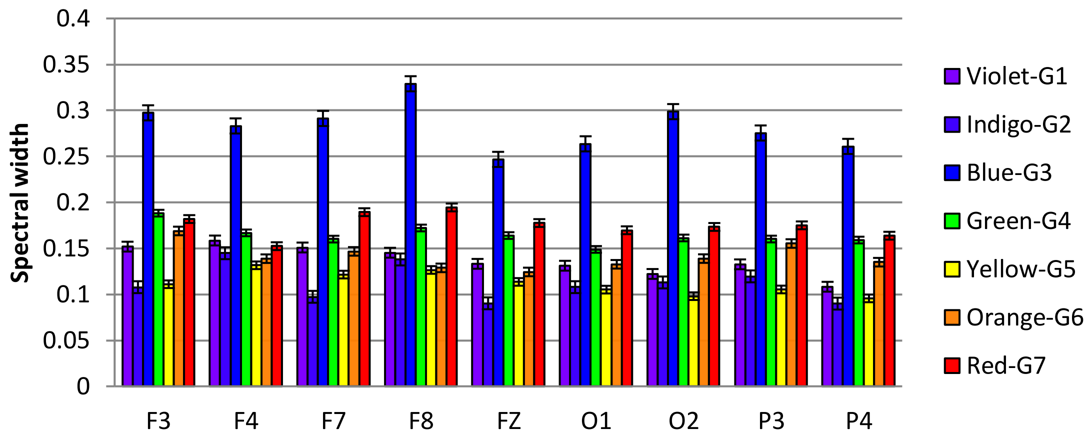


Fig. 4.9. Increase in Multifractal width from baseline Grey in nine elctrodes: the relative change of multifractal width is highest in F8 and O2, showing right hemispere bias in color perception. Also, change in width is highest in Frontal > Occipital > Parietal.

Similar to the absolute values, the changes in multifractal width are also highest in case of F8 and O2 (Violet being the only exception). Additionally, significant increase in F3 is noticed. For Parietal lobe, it is P3 which shows higher change in complexity instead of P4. Now, lateralization of color perception is a hotly debated issue in neuroscience literatures. Evidences have credited the bias to left (Franklin et al., 2008) as well as to right hemisphere (Njemanze et al., 1992). Some have found a balanced opinion that both hemispheres contribute in same extent (Witzel & Gegenfurtner, 2011). According to the lateralized category effect of colors, in the color naming tasks, left hemisphere advantage is dominant since the language lateralization also favors the left. On the other hand, the right hemisphere bias is seen during color working memory and discriminating properties like hue or saturation (Davidoff, 1976). In this light, aforementioned finding of our study, of course not related to color naming but color perception, the presence of increased complexity in mostly the right hemisphere electrodes could be very significant.

There is one more takeaway from Figures 4.8 and 4.9 worth mentioning. When the absolute values of complexity are considered, the Occipital lobe is seen to have the highest values followed by Frontal and then Parietal. But in the case of relative changes to baseline Grey, the complexity measures changes to Frontal > Occipital > Parietal. This indicates that during Grey viewing, Occipital and Parietal lobes display higher complexities than the Frontal lobe. Knowing the fact that Occipital lobes are primarily responsible for visual perception, this result doesn't come as a surprise. But the interesting part is - Parietal lobe also takes part in the visual process actively. Previous studies have reported such involvements, though in a different context (Battelli et al., 2009). This study, through measuring brain complexity via electrical activities, renders support to these claims.

Lastly, from both plots of Fig. 4.8 and Fig. 4.9, it is seen that signal corresponding to the frontal midline electrode Fz has recorded lower complexity compared to other frontal electrodes. Frontal midline power has usually been associated with emotional processing and positive emotional state (Suetsugi et al., 2000; Aftanas & Golocheikine, 2001; McFarland et al., 2016).

Lower complexity, and therefore, lower activation of Fz might be an indication that in this case, the color perception didn't involve any emotional arousal among the participants.

For the next part of the analysis using MFDXA, the combinations with the electrodes that were studied are: Left and Right hemisphere in Frontal (F3-F4, F7-F8, F3-F8, F4-F7) Occipital (O1-O2) and Parietal (P3-P4) lobes, Intra left (F3-F7) and right (F4-F8) hemisphere, Left Frontal and Occipital lobes (F3-O1, F3-O2, F7-O1, F7-O2), Right Frontal and Occipital lobes (F4-O1, F4-O2, F8-O1, F8-O2), Left Frontal and Parietal lobes (F3-P3, F3-P4, F7-P3, F7-P4), Right Frontal and Parietal lobes (F4-P3, F4-P4, F8-P3, F8-P4) and finally, Occipital and Parietal lobes (O1-P3, O1-P4, O2-P3, O2-P4). The total 28 combinations of electrodes were studied for all the experimental conditions. First, the noise cleaned EEG data were divided into N_s bins where N_s = int (N/s), N is the length of the series. The q-th order detrended covariance $F_q(s)$ was obtained for values of q from - 5 to + 5 in steps of 1. Power law scaling of $F_q(s)$ with s is observed for all values of q. The slope of this scaling λ(q) is the desired scaling exponent, which depends on q. A representative figure of the variation of λ(q) with changing q is given in Fig. 4.10 for F4-O2 electrodes during Violet color stimulus.

For comparison, the variation of H(q) with q individually for the same two electrodes F4 and O2 using MFDFA is shown in the same figure. The scaling exponent should have a constant value for a monofractal series, otherwise it implies multifractality. The plot indicates multifractal behavior of the cross-correlated time series, as for q = 2 the cross-correlation scaling exponent λ(q) is greater than 0.5 which is a confirmation of persistent long-range cross-correlation between the two electrodes. In a similar manner, λ(q) was evaluated for all the electrode combinations under consideration. The q-dependence of the classical multifractal scaling exponent τ(q) is shown in Fig. 4.11 for the electrodes F4 and O2. From the figure, it can be seen that the dependence of τ(q) on q is non-linear, which is another evidence of multifractality of the series.

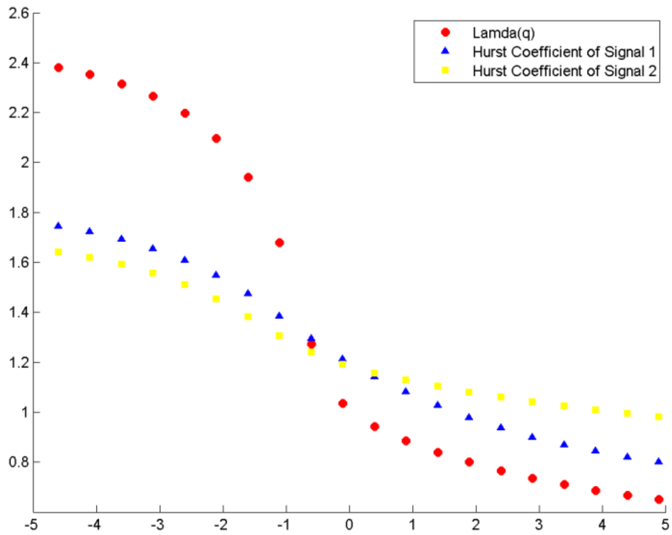


Fig. 4.10. Variation of scaling exponent λ(q) with q for F4-O2 electrode combination for Violet color exposure. The combination (red) is plotted against h(q) vs q of the individual signals F4 (blue) and O2 (yellow). The series has monofractal properties if the scaling exponent stays ~0.5 at q = 2. Here, λ(q) > 0.5, showing multifractality.

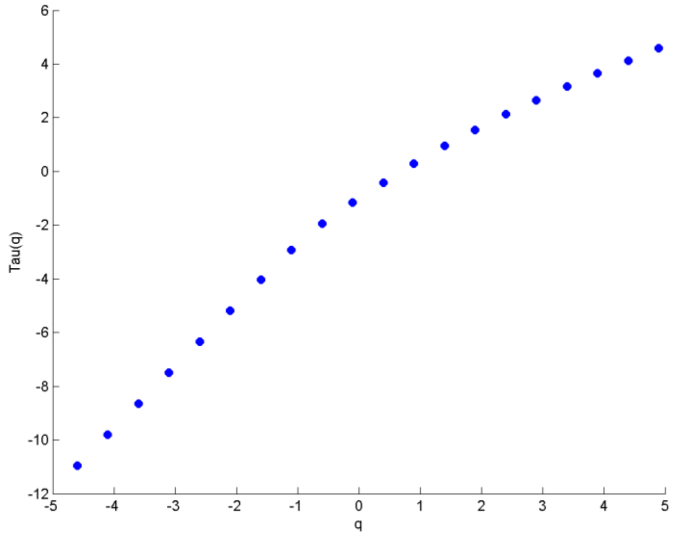


Fig. 4.11. Variation of classical multifractal scaling exponent τ(q) with q for F4-O2 electrode combination. In this case, τ(q) is a non-linear function of q, which suggests the presence of multifractality

Now, the multifractal width of the cross correlated signals of F4 and O2 is given in Fig. 4.12. The presence of spectral width in cross correlation data confirms the presence of multifractality in the correlated signal yet again.

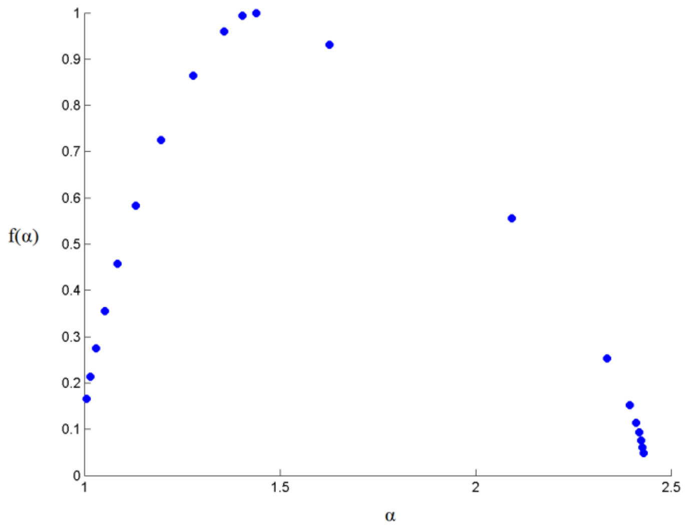


Fig. 4.12. Multifractal spectral width for cross-correlated electrodes F4 and O2. The relation of α (*q*-order singularity exponent) and f(α) (dimension of the series) represents the multifractal spectrum. Δα (= $\alpha_{max} - \alpha_{min}$) is the spectral width. The parabolic nature and the peak at α > 0.5 indicate presence of multifractal properties (Ihlen 2012)

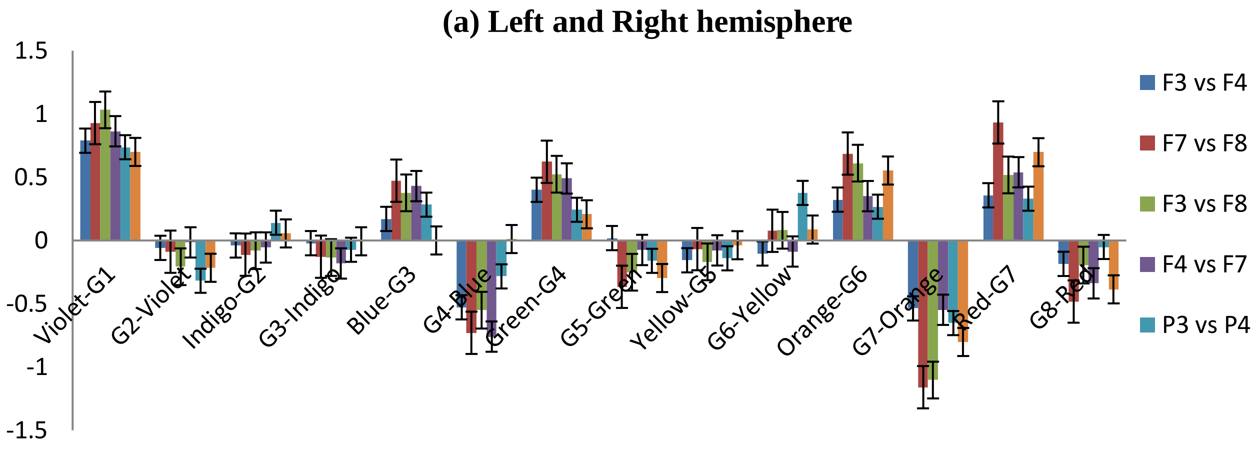
The fact that correlated signals from two electrodes in experimental conditions showing multifractality is an important observation for colour perception studies using nonlinear EEG analysis. The same analysis was done for the rest of the electrode combinations as well to find out the cross-correlation coefficient (γ_x) for all 16 participants. After that, similarly as MFDFA analysis, the averaged change in γ_x due to the color stimulus and the baseline Grey has been computed by calculating their differences. This gives us the change in cross-correlation co-efficient ($\Delta\gamma_x$) which is an effective tool for spotting the increase/decrease in cross-correlation pattern with the subsequent changes in experimental conditions (Sanyal et al., 2019). The data and the plots depicting the variations of $\Delta\gamma_x$ during different stages of the experiment are shown in Table 4.3 and Fig. 4.13. Increase and decrease in the values of $\Delta\gamma_x$ corresponds to reduced and enhanced degree of cross-correlations, respectively.

Average Difference in γ_x in experimental conditions (color - Grey) and (Grey - color)							
Electrode combinations	**Violet - G1**	**G2 - Violet**	**Indigo - G2**	**G3 - Indigo**	**Blue - G3**	**G4 - Blue**	**Green - G4**
F3-F4	0.788	-0.058	-0.039	-0.021	0.170	-0.528	0.401
F7-F8	0.928	-0.089	-0.112	-0.128	0.473	-0.730	0.622
F3-F8	1.032	-0.208	-0.080	-0.135	0.377	-0.551	0.524
F4-F7	0.863	-0.015	-0.055	-0.182	0.430	-0.760	0.490
P3-P4	0.737	-0.318	0.140	-0.073	0.283	-0.284	0.243
O1-O2	0.700	-0.216	0.056	-0.006	0.000	0.011	0.207
F3-F7	1.003	-0.084	-0.168	-0.075	0.372	-0.612	0.578
F4-F8	0.773	-0.101	-0.046	-0.044	0.272	-0.588	0.482
F3-O1	0.653	-0.261	0.036	0.048	-0.032	-0.106	0.380
F3-O2	0.720	0.125	-0.440	0.098	0.094	-0.437	0.705
F7-O1	0.529	-0.101	0.090	-0.107	0.192	-0.147	0.337
F7-O2	0.678	0.262	-0.553	0.102	0.232	-0.513	0.783
F4-O1	0.345	-0.046	-0.131	0.163	0.054	-0.215	0.258
F4-O2	0.489	0.129	-0.229	0.053	-0.030	-0.368	0.504
F8-O1	0.540	-0.223	0.114	-0.062	0.127	-0.208	0.341
F8-O2	0.581	0.224	-0.450	0.091	0.185	-0.547	0.732
F3-P3	0.840	-0.155	-0.063	-0.075	0.286	-0.570	0.546
F3-P4	1.006	-0.481	0.221	-0.179	0.385	-0.293	0.281
F7-P3	0.794	-0.043	-0.180	-0.045	0.414	-0.660	0.639
F7-P4	0.901	-0.265	0.017	-0.083	0.426	-0.346	0.241
F4-P3	0.646	-0.065	-0.049	0.021	0.159	-0.572	0.347
F4-P4	0.741	-0.343	0.194	-0.005	0.238	-0.268	0.239
F8-P3	0.715	-0.123	-0.025	-0.068	0.298	-0.584	0.545
F8-P4	0.854	-0.351	0.181	-0.147	0.431	-0.391	0.286
O1-P3	0.662	-0.250	0.121	-0.015	0.072	-0.169	0.276
O1-P4	0.768	-0.483	0.198	-0.054	0.108	-0.005	0.063
O2-P3	0.601	0.104	-0.466	0.254	0.051	-0.471	0.640
O2-P4	0.673	-0.170	-0.357	0.241	0.200	-0.147	0.338

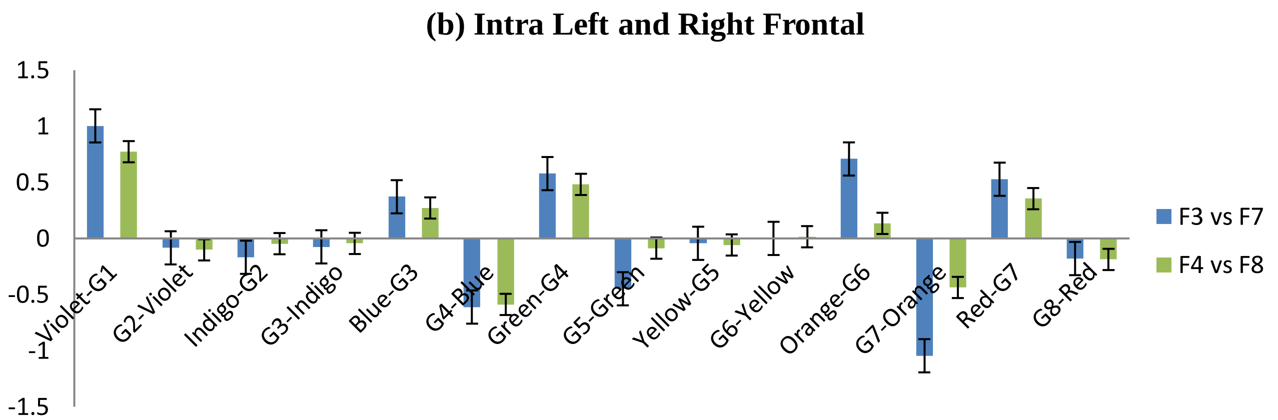
Average Difference in γ_x in experimental conditions (color - Grey) and (Grey - color) (continued..)							
Electrode combi-nations	**G5 - Green**	**Yellow - G5**	**G6 - Yellow**	**Orange - G6**	**G7 - Orange**	**Red - G7**	**G8 - Red**
F3-F4	0.019	-0.156	-0.103	0.323	-0.537	0.357	-0.185
F7-F8	-0.366	-0.068	0.077	0.687	-1.160	0.933	-0.483
F3-F8	-0.251	-0.169	0.081	0.611	-1.103	0.518	-0.194
F4-F7	-0.075	-0.078	-0.087	0.351	-0.548	0.539	-0.339
P3-P4	-0.161	-0.141	0.376	0.267	-0.653	0.330	-0.051
O1-O2	-0.297	-0.037	0.087	0.553	-0.802	0.698	-0.386
F3-F7	-0.448	-0.043	0.000	0.709	-1.046	0.528	-0.179
F4-F8	-0.086	-0.058	0.016	0.134	-0.437	0.355	-0.187
F3-O1	-0.395	-0.061	0.098	0.672	-0.920	0.479	-0.117
F3-O2	-0.351	-0.239	0.051	0.448	-0.950	0.513	-0.201
F7-O1	-0.595	0.181	0.013	0.674	-0.907	0.595	-0.224
F7-O2	-0.525	-0.112	-0.040	0.517	-0.933	0.608	-0.279
F4-O1	-0.201	-0.036	0.116	0.354	-0.502	0.467	-0.248
F4-O2	-0.041	-0.292	-0.047	0.203	-0.404	0.354	-0.201
F8-O1	-0.239	-0.110	0.165	0.589	-0.955	0.571	-0.244
F8-O2	-0.249	-0.335	0.181	0.320	-0.906	0.506	-0.235
F3-P3	-0.292	-0.038	-0.010	0.602	-0.945	0.412	-0.163
F3-P4	-0.284	-0.165	0.094	0.678	-0.661	0.121	0.075
F7-P3	-0.490	0.040	-0.048	0.763	-1.029	0.586	-0.287
F7-P4	-0.328	-0.031	0.038	0.695	-0.684	0.331	-0.032
F4-P3	0.119	-0.165	0.023	0.245	-0.463	0.391	-0.261
F4-P4	-0.088	-0.245	-0.028	0.420	-0.254	0.171	0.012
F8-P3	-0.228	-0.126	0.122	0.509	-0.950	0.621	-0.344
F8-P4	-0.228	-0.157	0.124	0.547	-0.586	0.380	-0.201
O1-P3	-0.384	-0.088	0.424	0.350	-0.773	0.549	-0.291
O1-P4	-0.209	-0.023	0.377	0.344	-0.677	0.272	-0.063
O2-P3	-0.240	-0.206	0.195	0.280	-0.896	0.514	-0.253
O2-P4	-0.319	-0.157	-0.024	0.633	-0.571	0.179	-0.003

Table 4.3. Changes in cross-correlation coefficient γ_x in electrode combinations in both (color – Grey) and (Grey - color) conditions; for n = 16. The highlighted ones belong to the combinations representing homologous brain areas (Frontal, parietal and Occipital)

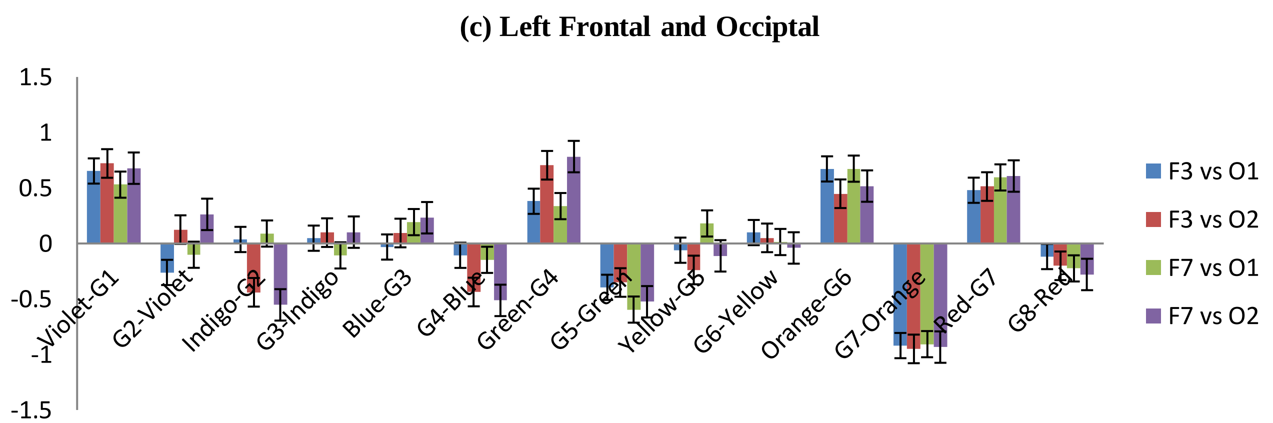
The graphical representation of Table 4.3 is given in Fig. 4.13 (a-g) (Standard deviations are included as error bars).



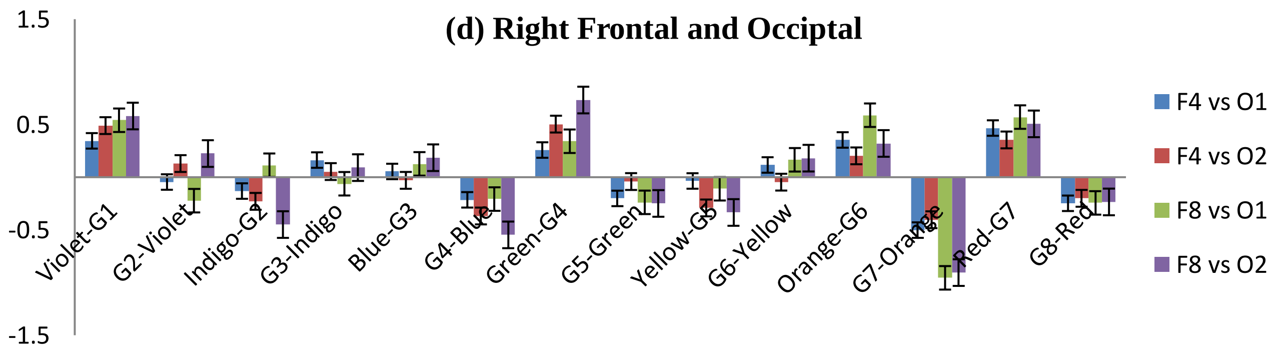
(a) Left and Right hemisphere
1.5
1
0.5
0
-0.5
-1
-1.5
Violet-G1
G2-Violet
Indigo-G2
G3-Indigo
Blue-G3
G4-Blue
Green-G4
G5-Green
Yellow-G5
G6-Yellow
Orange-G6
G7-Orange
Red-G7
G8-Red
F3 vs F4
F7 vs F8
F3 vs F8
F4 vs F7
P3 vs P4



(b) Intra Left and Right Frontal
1.5
1
0.5
0
-0.5
-1
-1.5
Violet-G1
G2-Violet
Indigo-G2
G3-Indigo
Blue-G3
G4-Blue
Green-G4
G5-Green
Yellow-G5
G6-Yellow
Orange-G6
G7-Orange
Red-G7
G8-Red
F3 vs F7
F4 vs F8



(c) Left Frontal and Occiptal
1.5
1
0.5
0
-0.5
-1
-1.5
Violet-G1
G2-Violet
Indigo-G2
G3-Indigo
Blue-G3
G4-Blue
Green-G4
G5-Green
Yellow-G5
G6-Yellow
Orange-G6
G7-Orange
Red-G7
G8-Red
F3 vs O1
F3 vs O2
F7 vs O1
F7 vs O2



(d) Right Frontal and Occiptal
1.5
0.5
-0.5
-1.5
Violet-G1
G2-Violet
Indigo-G2
G3-Indigo
Blue-G3
G4-Blue
Green-G4
G5-Green
Yellow-G5
G6-Yellow
Orange-G6
G7-Orange
Red-G7
G8-Red
F4 vs O1
F4 vs O2
F8 vs O1
F8 vs O2

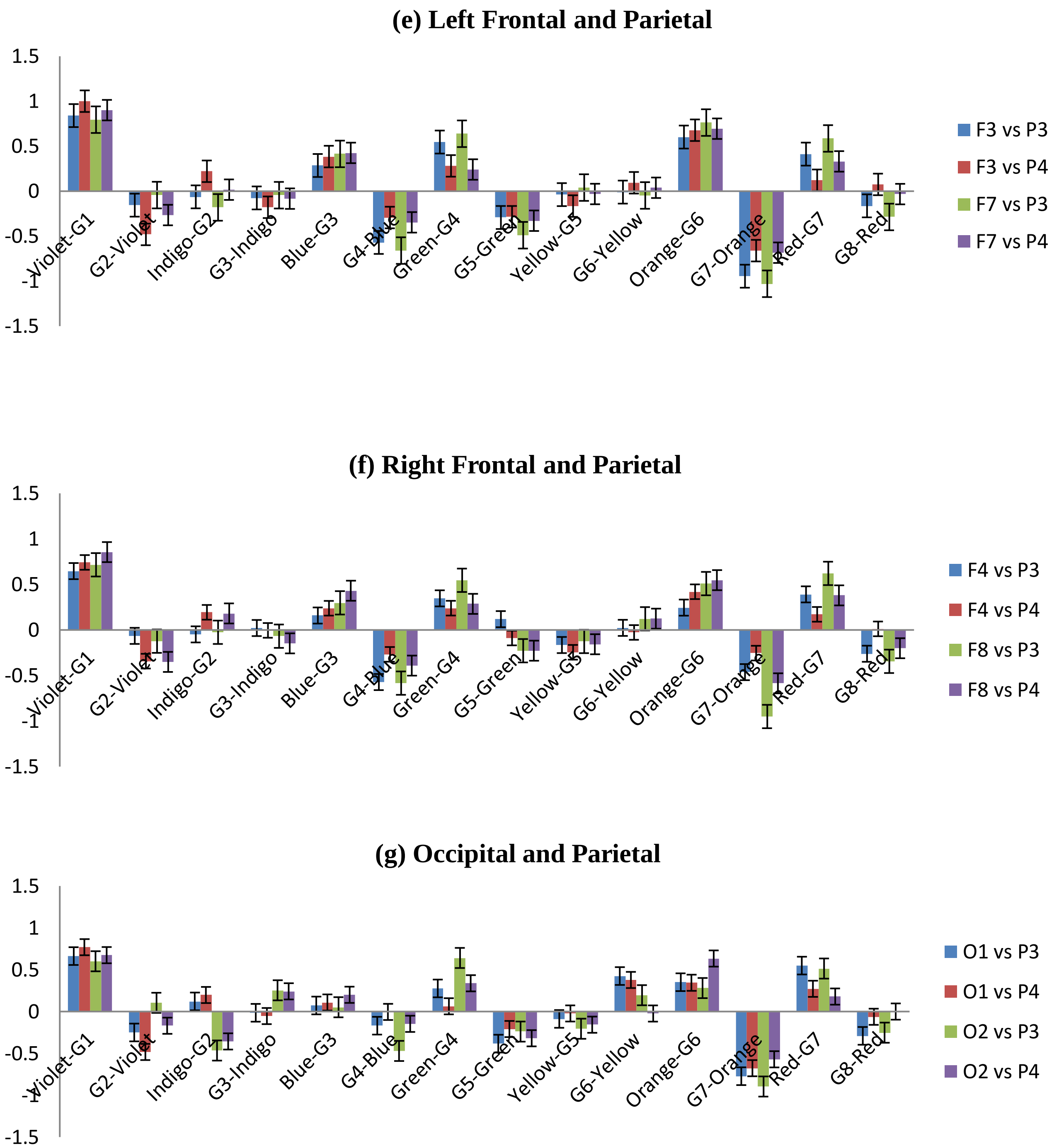
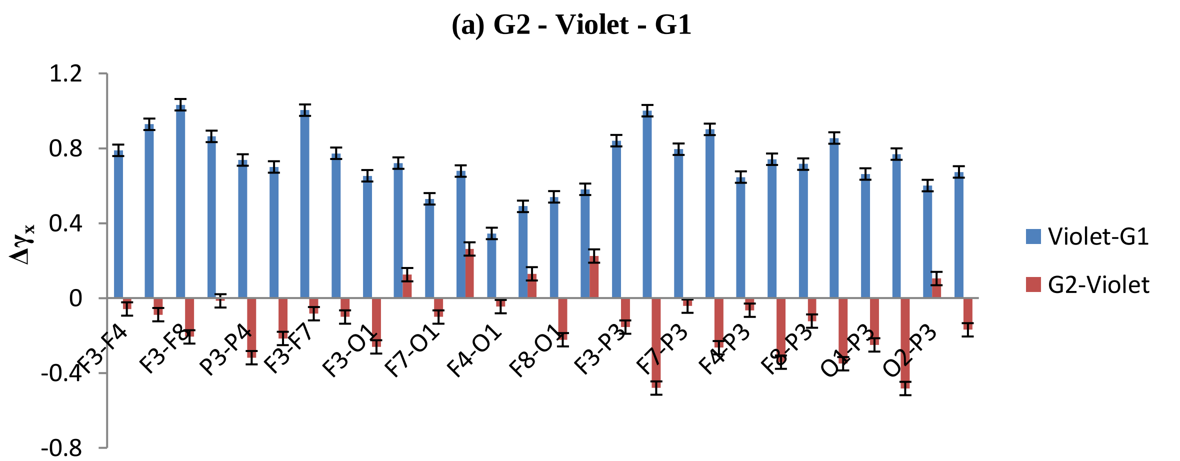
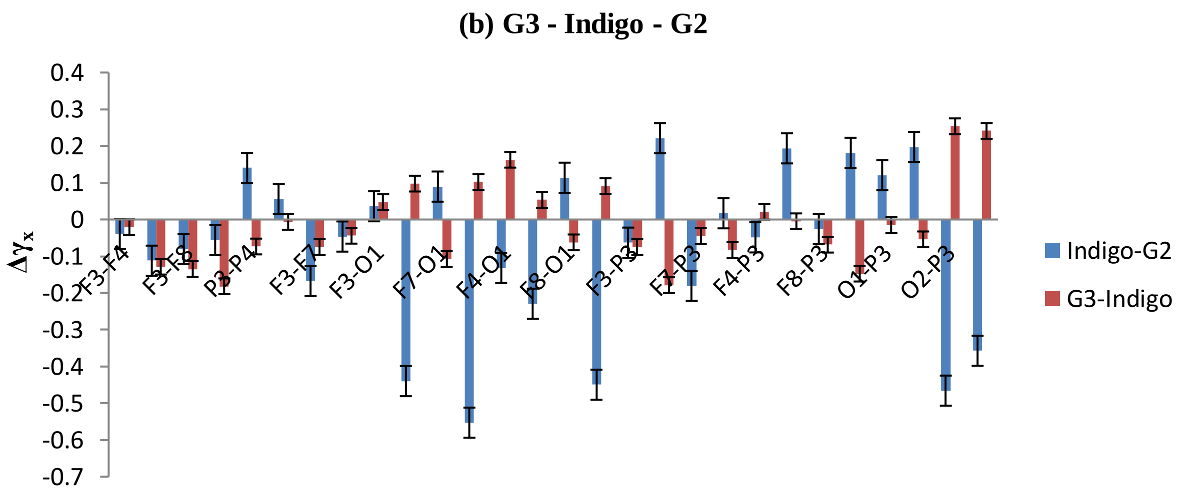


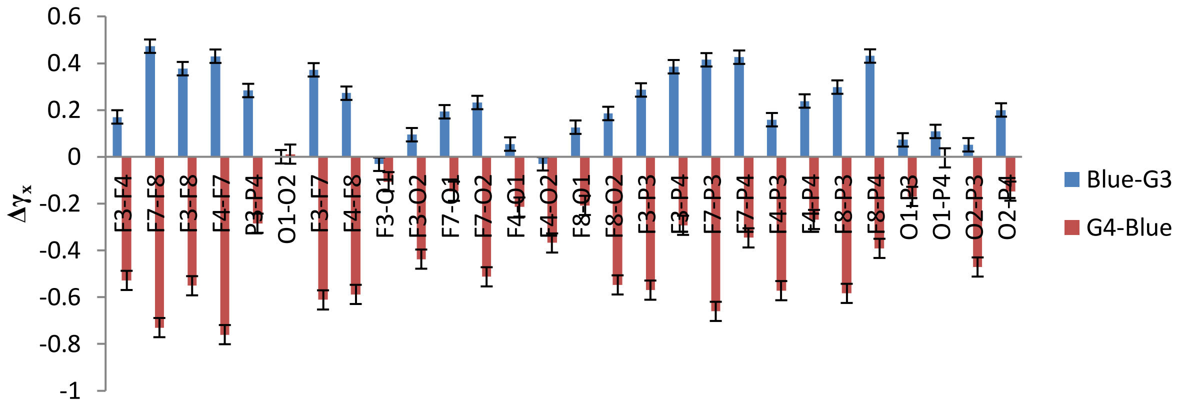
Fig. 4.13 (a-g). Changes in cross-correlation coefficient ($\Delta\gamma_x$) in electrode combinations in both (color – Grey) and (Grey - color) conditions. Negative values indicate higher cross-correlation. Pattern of change is consistent across different brain areas, remarkably higher correlation after stimulus color removal than during exposure.

From Fig. 4.13, it is seen that although the degree of cross-correlation has varied throughout the experiment, the nature of the change is somewhat similar in electrode combinations. One of the most remarkable features seen in this data is the existence of high cross-correlation in the electrodes during the baseline Grey period, in various occasions. This is fascinating considering the fact that it happened in almost all the inter/intra lobe combinations, which indicates that the responses across different lobes of the brain have stayed correlated irrespective of their spatial distribution. Moreover, in most cases, colour exposure has increased the $\Delta\gamma_x$, meaning that introduction of color stimulus has reduced the cross-correlation. This observation is unique and hitherto unseen from the point of view of color perception. The highest $\Delta\gamma_x$, and least correlations, is observed during the introduction of Violet from Grey1. On the other hand, shift to Grey7 from color Orange displayed lowest $\Delta\gamma_x$ and most enhanced degree of cross-correlations. Interestingly, exposure to colors that exhibited highest complexity in MFDFA - Blue, Red and Green – resulted in reduced cross-correlation whereas the Greys before or after them expressed enhanced effects of the same. To study the correlations in more spatial manner, we rearranged the plots highlighting each experimental condition. They are given in Fig. 4.14.

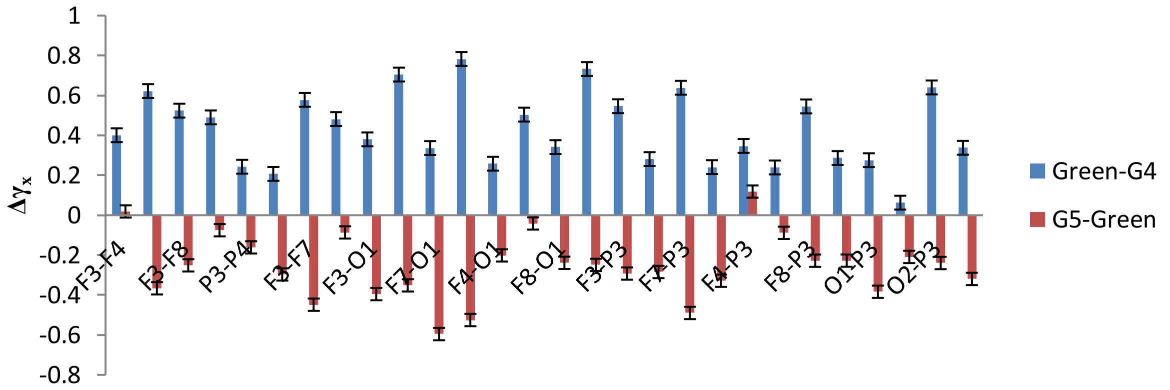




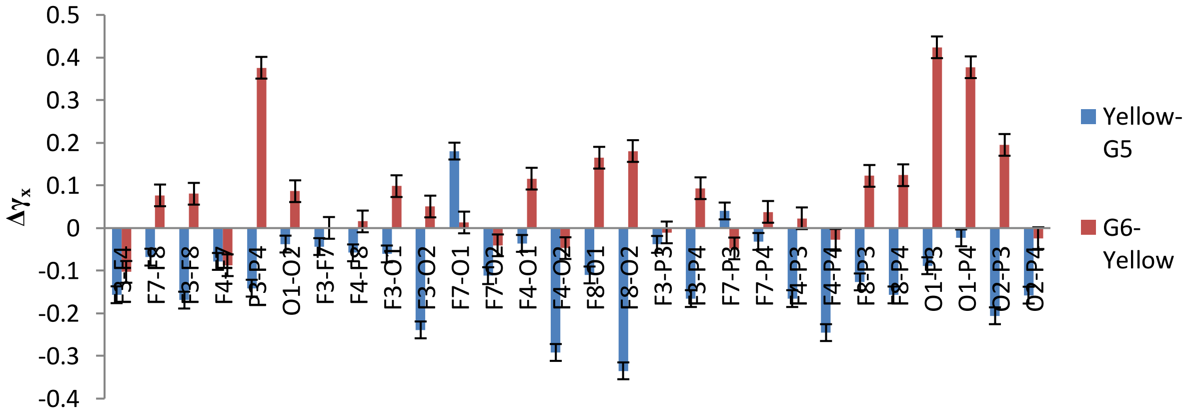
(c) G4 - Blue - G3



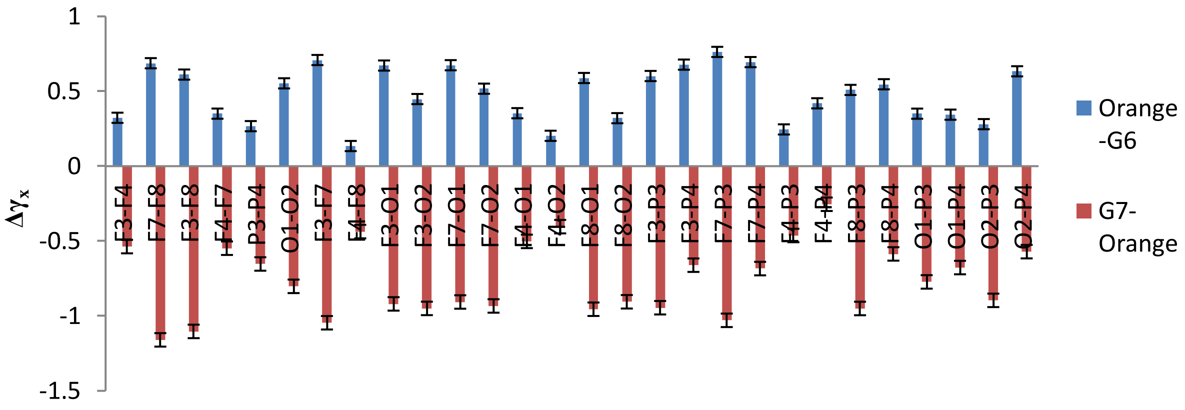
(d) G5 - Green - G4



(e) G6 - Yellow - G5



(f) G7 - Orange - G6



(g) G8 - Red - G7

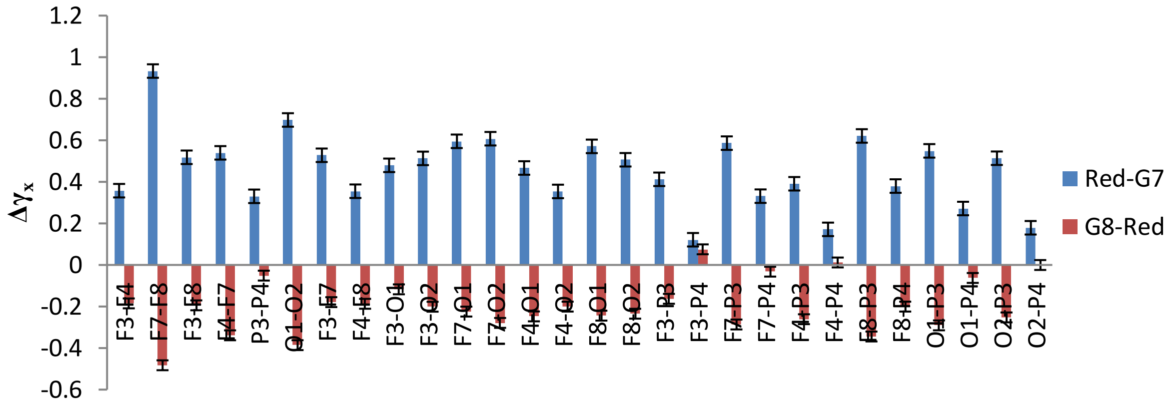


Fig. 4.14 (a-g). Color-wise distribution of $\Delta\gamma_x$ in specific (color – Grey) and (Grey - color) conditions. Negative values indicate higher correlation. Except for Indigo (b) and Yellow (e), correlation decreases on color exposure and increases on removal; magnitude of $\Delta\gamma_x$ is highest

Figure 4.14(a-g) shows the changes in cross-correlation co-efficient in specific experimental conditions in all the electrode combinations (Standard deviations are included as error bars). In case of Violet and related Greys, $\Delta\gamma_x$ increases in all the electrodes when stimulus is changed from Grey1 to violet, implying lower cross-correlation. During Violet to G2, correlation increases mainly in Occipital and Parietal electrodes, slightly in F3 and F8 combinations. In the next colour, this pattern is reversed. When exposed to Indigo, high correlations are seen in O2 electrode combinations and in left Frontal electrodes F3 and F7. Change to G3 results in reduced correlation in most of the electrodes, except the Frontal ones. For Blue, the pattern follows Violet once again. While G3 to Blue, correlations decrease throughout with higher $\Delta\gamma_x$ and Blue to G4 sees significant increase in cross-correlation throughout the brain. Green, too, shows the similar trend – correlation vanishes during its presence and increases hugely after going to next Grey, i.e., G5. Yellow, like Indigo, breaks the norm with high correlations in general (except in F7-O1) which gets destroyed with its removal. The next color shows striking consistency in all the participants. Orange follows the Blue and Green pattern, but with higher amplitude than any other colors. The cross-correlation is as high during the stimulus removal as the decrease while the color is on. This trend continues for the last color stimulus, Red, albeit in lower volumes.

To sum up, the figures indicate that except for Indigo and Yellow (interestingly the colors with lowest complexities), rest of the colors show similar patterns in the change of degree of cross-correlation ($\Delta\gamma_x$) – during the color viewing, $\Delta\gamma_x$ increases which results in reduced correlations. Once the stimulus is removed, the correlations spike up. We argue that this ebb and flow of cross-correlations are linked with the processing of visual information - which might be manifested well after the removal of the stimulus. Exposure to the color stimulus affects the bio signals emanating from the areas directly involved in its perception, thus providing the increase in signal complexity, demonstrable via MFDFA. In the post-stimulus period, the sensory information collected during the color-viewing window gets processed and integrated via various inter/intra lobe exchanges until the information is sufficiently segregated. These connections and exchanges manifest themselves via the cross-correlation parameter. Such perceptual retention of information has been studied in neuroscience for a long time (some have called it 'perceptual hysteresis', analogous to magnetic hysteresis). There have been reports of hysteresis or retention of stimulus perception in visual (Kleinschmidt et al., 2002) and auditory stimulus (Banerjee et al., 2016). Although the purpose or the experimental design in this work is not intended to find out hysteresis of color perception, but our findings could advocate for possible investigations towards it. These findings on the neuronal activity are evidently novel in this field of study and provide a strong argument for the future of robust non-linear methodologies in EEG based color perception research.

4.7. STATISTICAL ANALYSIS

To test the statistical significance of our results 2-way ANOVA was performed on the multifractal spectral width values considering the colors and the channels as factors and the detailed result of the same has been presented in Table 4.4.

At the 0.05 level (i.e., 95% confidence level), the population means of colors are **significantly** different.
At the 0.05 level, the population means of Electrodes are **not significantly** different.
At the 0.05 level, the interaction between color and Electrode is **not significant.**

Tukey test was done only for the variables with significance. Table 4.5 includes the post-hoc analysis.

Tukey test results, calculated over the population means of the multifractal spectral widths of the color wise EEG responses of the 16 participants also confirm that population mean of spectral width in response to color Blue is significantly different from all other colors. The response to Red and Green also show significant differences from others. But rest of the colors (Indigo, Violet, Yellow and Orange) does not yield such significant changes in the EEG complexities in different lobes of the brain. For the other factors, like electrode channels or interaction between electrode and color, this significance may not appear to be so significant at our predetermined 95% confidence level, but enhancement of sample size, in future, is expected to give further support.

Overall ANOVA	*df*	Sum of Squares (*SS*)	Mean Square	F-value	p-value (significance level < .05)
Color	6	2.63599	0.43933	23.30954	9.15E-26
Electrode	8	0.02985	0.00373	0.19795	0.99116
Interaction	48	0.06001	0.00125	0.06634	1
Model	62	2.72585	0.04397	2.33266	8.30E-08
Error	945	17.81107	0.01885	--	--
Corrected Total	1007	20.53692	--	--	--

Table 4.4. Detailed result of 2-way ANOVA on multifractal spectral width

Means comparisons								
Tukey Test Colors	Mean Diff	SEM	q Value	Prob	Alpha	Sig	LCL	UCL
Indigo - Violet	-0.01307	0.01618	1.1421	9.84E-01	0.05	0	-0.06087	0.03474
Blue - Violet	0.14135	0.01618	12.35531	3.53E-09	0.05	1	0.09355	0.18916
Blue - Indigo	0.15442	0.01618	13.49741	2.65E-09	0.05	1	0.10661	0.20222
Green - Violet	0.04233	0.01618	3.69984	1.22E-01	0.05	0	-0.00548	0.09013
Green - Indigo	0.05539	0.01618	4.84195	0.01145	0.05	1	0.00759	0.1032
Green - Blue	-0.09902	0.01618	8.65546	4.12E-08	0.05	1	-0.14683	-0.05122
Yellow - Violet	-0.00461	0.01618	0.40339	0.99996	0.05	0	-0.05242	0.04319
Yellow - Indigo	0.00845	0.01618	0.73871	0.99854	0.05	0	-0.03935	0.05626
Yellow - Blue	-0.14597	0.01618	12.75869	3.21E-09	0.05	1	-0.19377	-0.09816
Yellow - Green	-0.04694	0.01618	4.10323	0.05812	0.05	0	-0.09475	8.63E-04
Orange - Violet	0.02963	0.01618	2.5903	0.52686	0.05	0	-0.01817	0.07744
Orange - Indigo	0.0427	0.01618	3.7324	0.11554	0.05	0	-0.00511	0.09051
Orange - Blue	-0.11172	0.01618	9.76501	1.04E-08	0.05	1	-0.15952	-0.06391
Orange - Green	-0.01269	0.01618	1.10954	0.98644	0.05	0	-0.0605	0.03511
Orange - Yellow	0.03425	0.01618	2.99369	0.34343	0.05	0	-0.01356	0.08206
Red - Violet	0.07896	0.01618	6.90194	2.56E-05	0.05	1	0.03116	0.12677
Red - Indigo	0.09203	0.01618	8.04404	3.62E-07	0.05	1	0.04422	0.13983
Red - Blue	-0.06239	0.01618	5.45337	0.00235	0.05	1	-0.1102	-0.01458
Red - Green	0.03663	0.01618	3.2021	0.26267	0.05	0	-0.01117	0.08444
Red - Yellow	0.08358	0.01618	7.30533	6.07E-06	0.05	1	0.03577	0.13138
Red - Orange	0.04933	0.01618	4.31164	0.038	0.05	1	0.00152	0.09713

Table 4.5. Post hoc analysis (Tukey test) for colors

4.8. CONCLUSION AND GENERAL DISCUSSION

The question of how colors affect human beings is a long, much-debated one and has remained so despite years of work. The present literatures concentrate more on the applicative potential of colors in psychological perspective. They report divided opinions on the effects since color perception is often likely to be contextual and overlaps with cultural and linguistic dependency. Moreover, comprehensive studies on the physiological responses are sparse and due to the analysis technique, limited by severe approximations. With this backdrop, our work had set out in an exhaustive investigation of neuronal activities in brain during color perception via its physiological manifestation in EEG. The uniqueness our work offered was twofold – methodological and analysis related. Most of the studies use two or three colors together for comparison and to study their roles in specific psychological attributes. Usage of the whole color spectrum is unconventional otherwise. We have used it in our work to explore the effects of the whole wavelength range of visible light altogether instead of comparing some of them. This, we believe, could demonstrate how brain responds to color in a more extensive manner. And for the analysis part, no other studies have used such rigorous non-linear tools like MFDFA and MFDXA in the domain of color perception. Over the course of the study, we have seen that this novel approach provides interesting new data in regard to the color perception process which has not been reported ever before. Findings of the MFDFA analysis may be summarized in the following:

1. The presence of fractality in the color induced bio-signals indicates towards their complex non-linear nature. Tackling such systems with linear analysis methods like FFT or power spectral density is not sufficient in understanding the intricacies of color perception. They approximate various parameters and can lead to misleading results. Rigorous statistical tools such as MFDFA are necessary, considering they can identify parameters directly related to the complexities and quantify them, in due course.

2. MFDFA analysis of the color induced EEG shows the presence of multifractality in all the brain areas under consideration i.e., electrodes in Frontal, Occipital and Parietal lobes. Multifractality is quantified by multifractal spectral width, which is a measure of degree of complexity or randomness. The fact that complexity is observed in these brain areas simultaneously suggests that they participate actively during color perception. Occipital and Frontal lobes, being the visual and cognitive centers of the sensory perception, are expected to be involved in the process. But MFDFA analysis additionally point towards Parietal lobe activation during color perception as well.

3. The increase in complexity during color viewing and decrease during baseline Grey - such change in the complexity pattern is similar throughout the electrodes for the whole experimental duration. This indicates that the process of color perception includes the ability of separating a color from the set baseline by means of the change in respective degrees of long-range correlations; which makes the multifractal spectral width an efficient marker in bio-signal analysis for color perception studies in future.

4. A novel and interesting finding that must be mentioned is the nature of spectral width with respect to color stimulus. The values of the multifractal width are found highest for color Blue, followed by Red and then Green. Yellow recorded the lowest width, followed by Indigo. Multifractal spectral width measures the long-range correlations present in the signal. So, it can be said without doubt that such correlations are higher in Blue than Red, indicating higher arousal. Overall, Blue-Green part (shorter wavelengths) of the spectrum showed higher arousal than Red/Orange part (higher wavelength). This result is remarkable in terms of offering support to previous ideas. Though studies have previously reported the arousal during Blue (Lockley et al., 2006; Vandewalle et al.,

2007; Yoto et al., 2007), but any consensus is yet to be reported, much less the how's and why's of the perceptual detail. Our study, backed by robust non-linear tools, could embolden the validity of these claims.

5. An offshoot of the previous observation is the fact that the highest complexity is recorded for the three primary colors. This could be due to the fact that the photoreceptors which are directly responsible for color vision consist of the three types – red, green and blue – and perceives these colors more actively than others. The manifestation of this activation is displayed via spectral width. When the arousal due to colors other than these three is reported, it is mostly in light of some cognitive task-based study. Hence, cognition plays a role in those scenarios which might favor other colors. This work, designed specifically to address the electrical activity due to color vision, doesn't factor in such involvements. We reckon this is why the primary colors show high activation in EEG data.

6. For the color-wise breakdown of the absolute values of multifractal width, the electrodes that recorded the highest complexity were F8, O2 and P4, all of which belong to the right hemisphere. This suggests that the long-range correlations in the right part of the brain is more prominent while color viewing than the left. Data for the relative increase of width, i.e., increase from the immediately prior baseline color repeats this trend for the Occipital lobe. But P3 and F3 electrodes in Parietal and Frontal lobe show higher relative increase. This result is an indicator of hemispherical asymmetry in color perception in at least Occipital and Parietal lobes. Frontal lobes, too, could be considered for such lateralization (right brain bias) as the amplitude of increase in F8 is significantly higher than F3.

7. Lastly, substantial activation of Parietal electrodes is seen not only during color stimulus, but even for baseline Grey too. The visual process is related directly to Occipital area. Our results hint at Parietal participation in the process as well. Since part of the Parietal lobe is involved in the integration of sensory information, the complexity change seems justified. Although targeted future investigations could help reveal specific details of this involvement.

From the next part of the experiment, Multifractal detrended cross-correlation analysis (MFDXA) provides few more important conclusions regarding the cross-correlation pattern in different parts of the brain during color perception. Those are given below:

1. The cross-correlated series between all the electrode combinations shows the presence of multifractality implying long range correlations between not only in the signals themselves, but in the cross-correlated data as well. Hence, MFDXA analysis definitely has the potential to offer necessary parameters for quantifying cross-correlation process between different areas of the human brain.

2. Considering lobe-wise cross-correlation pattern, it is seen that the electrodes in general share very similar trends barring few exceptions. That is to say, during each experimental condition, the change in γ_x for most of the inter/intra lobe combinations was by and large similar. This advocates for the cross-correlation pattern to be very consistent throughout, irrespective of their spatial distribution. It is a definitive illustration of the inter/ intra lobe dependency of color vision.

3. The most significant observation in this segment of the experiment is the overall trend of reduced cross-correlation among electrodes with the introduction of color stimulus and its increase when the color is removed. Most of the colors – Blue, Green, Red, Orange and Violet – resulted in cross-correlation reduction when applied (Violet being the highest). On the other hand, enhanced cross-correlations were observed when they were changed to the next baseline Grey (Orange to G7 showing the highest magnitude). We argue that this pattern emerges due to the processing of visual data well after the stimulus retrieval. The cross-correlation coefficient γ_x is the manifestation of the enhanced connection between various parts of the brain during this phase. The two colors that fall outside this norm are Indigo and Yellow, interestingly both had low complexity compared to others. This fuels the argument further; since in these two cases, the arousal levels were far lower which points at having to process less sensory information content. Hence, the need for the correlation during and after being exposed to Indigo and Yellow appears to be far limited (apart from Right Frontal-Occipital combinations like F4-O2 and F8-O2, the magnitude of cross-correlation is quite small, too).
4. Destruction of cross-correlation was highest in case of G1 to Violet and lowest during Orange to G7. In both of the cases, Inter-frontal or Fronto-Parietal combinations have registered optimum changes.

From these vital and newfound results emerging from both the MFDFA and MFDXA analysis, it can be said with certainty that the principal aim of this study has been met successfully. We have studied the changes in several brain activities in detail and were able to challenge or consolidate the existing ideas from the stout platform set by the aforementioned methods. Also, the correlation study of the brainwaves indicates that the correlation between various lobes is significantly higher during specific color stimulus exposure and more importantly, even after the stimulus retrieval. The obtained data may be of immense importance when it comes to studying the neuro-cognitive basis of color perception. Previous researches on color perception have presented an array of psychological and physiological aspects of it, as discussed before. The points below encapsulate how this study could contribute to the growing literature:

Firstly, we would like to refer to the necessary methodological treatment. Studies in this field of research have shown some methodological shortcomings that has been eloquently summarised by Elliot (2019). Based on those weaknesses, some of the points are recommended to increase reliability and reproducibility of the work. They include confirming participant naiveté and confidentiality, checking color vision deficiency among them, adequate sample size of the trial, reporting color specification and controls during experiment, monitoring ambient illumination etc. In this work, we tried to comply with these to the best of our efforts. The participants were ensured of their anonymity and kept uniformed about anything but the procedure required for them to follow. They were asked to take an online variation of the Ishihara test (Ishihara, 1987) to ensure they don't have color-blindness or color vision deficiency. Consulting other EEG studies on sensory perceptions (Ghosh et al., 2018), adequate sample size was also ensured. Specifications of the experimental instruments and the color notations are given as per the recommendations. One important point which needs to be emphasised here is that the investigation was focused to study the complexity patterns of cortical electrical activity and to analyse it quantitatively. Hence the color control was not set to compare the effect of one/two specific colors or their colorimetric properties like hue or chroma. Although we found the change in complexity of color-induced EEG patterns, more effect-specific experimental design are necessary to comment on which properties of the colors cause said changes in a more definitive manner. Going back to the experimental process, another suggestion of keeping the ambient illumination to a minimum was also met as the participants were kept in a dark room during the procedure. This way, we have followed the recommendations to have the quality of the overall methodological arrangement be of higher standards.

Secondly, as mentioned in the above point, analysis of color-induced bio-signals was the idea of the work. Hence, the whole gamut of visual spectrum was displayed instead of few colors (as is done in majority studies). The complexity changes due to various wavelengths could be examined this way and with the emerging pattern, further explorations can be pinpointed.

Thirdly, the most important novelty that this study offers is the usage of non-linear tools for analyzing the EEG signals. Previous literatures were limited in this area since they haven't factored in the chaotic aspect of the biological manifestations. Complexity and long-range correlations are inherent properties of such series. Hence, without considering these properties, complete scrutiny of them couldn't be carried out. Establishing this fact is an important point that this study hopes to achieve; something which might be the next step in future investigations of color perception.

Fourth and finally, the parameters this study has resulted (multifractal width w, cross-correlation coefficient γ_x) are unprecedented in this field of research and of high significance. These parameters emerge from the non-linear dynamics of brain signals under color stimuli and quantifying them is a step towards understanding the signatures of such dynamics more elaborately and with more rigour. Another byproduct of this analysis is obtaining such quantifiable parameters which are reproducible under different experimental circumstances. Hence, even with altered aims, the dynamics in the brain can be well understood.

Having said the abovementioned contributions, our study eventually is a stepping stone to a larger horizon of color perception and cognition research that can be traversed by extending the work further. This study was conducted with 16 participants, which may limit us from global interpretation of the findings presented for now. To attain the same, we plan to further extend this study with a much larger pool of participants. In future, for pinpointing the most relevant parameters/leads that contribute to the major findings of this work, machine learning techniques can be used during analysis. This will, in turn, improve the overall interpretability of the parameters. This work aimed to explore the changes in the broadband EEG signal with exposure to different colors of the VIBGYOR. For further validation study, the results of this work can be compared to the results of frequency dependent analysis of the same EEG data corresponding to specific experimental conditions (to investigate if any particular EEG frequency band gets more pronounced while viewing any particular color). Also, further classifications of the participants based on their age, sex, color preference can be made before interpreting the findings of this study to reach a more generalized conclusion in the global scenario. One more interesting extension of our report could be in the direction of non-random order presentation of the visual stimulus and its subsequent influences on the results found. To best of our knowledge, reported studies on order effects in visual or color perception tasks are very scarce. Yet it wouldn't be entirely accurate, theoretically at the least, to dismiss the possibility of its presence in this case. For example: interference effect of one experimental color on the perception of the next, given the short exposure time (although the visual perception of color takes only ~150-200 ms, reported by Amano et al. in 2006). It, therefore, remains another avenue which could be delved into in future.

Before concluding this chapter, it is worth calling out that human color perception is a domain vast enough to induce interest from a plethora of disciplines like physics, neuroscience, cognitive science, psychology, psychophysics, marketing industry and even visual arts and linguistics. Needless to say, that to advance our knowledge on how we perceive and process colors the requirement of rigorous scientific tools is absolutely crucial. This work takes such a step to tackle the electrophysiological aspect of the problem. The novel approach described attempts to study the color perception by analyzing its physiological signatures meticulously using state-of-the-art non-linear tools. We demonstrate that quantifying parameters like degree of complexity and degree of cross-correlation can categorically reveal exclusive information about neuronal dynamics and nature of their interdependency. Our work, therefore, argues heavily in favor of using such advanced tools in color perception studies and hopes to initiate a modern paradigm of research in this field.

REFERENCES

Aftanas, L. I., & Golocheikine, S. A. (2001). Human anterior and frontal midline theta and lower alpha reflect emotionally positive state and internalized attention: high-resolution EEG investigation of meditation. *Neuroscience letters*, *310*(1), 57-60.

AL-Ayash, A., Kane, R. T., Smith, D., & Green-Armytage, P. (2016). The influence of color on student emotion, heart rate, and performance in learning environments. *Color Research & Application*, *41*(2), 196-205.

Ali, M. R. (1972). Pattern of EEG recovery under photic stimulation by light of different colors. *Electroencephalography and Clinical Neurophysiology*, *33*(3), 332-335.

Amano, K., Goda, N., Nishida, S. Y., Ejima, Y., Takeda, T., & Ohtani, Y. (2006). Estimation of the timing of human visual perception from magnetoencephalography. *Journal of Neuroscience*, *26*(15), 3981-3991.

Avillac, M., Deneve, S., Olivier, E., Pouget, A., & Duhamel, J. R. (2005). Reference frames for representing visual and tactile locations in parietal cortex. *Nature neuroscience*, *8*(7), 941-949.

Baek, H., & Min, B. K. (2015). Blue light aids in coping with the post-lunch dip: an EEG study. *Ergonomics*, *58*(5), 803-810.

Bakker, I., van der Voordt, T. J., de Boon, J., & Vink, P. (2013). Red or blue meeting rooms: does it matter?. *Facilities*.

Banerjee, A., Sanyal, S., Patranabis, A., Banerjee, K., Guhathakurta, T., Sengupta, R., ... & Ghose, P. (2016). Study on brain dynamics by non linear analysis of music induced EEG signals. *Physica A: Statistical Mechanics and its Applications*, *444*, 110-120.

Battelli, L., Alvarez, G. A., Carlson, T., & Pascual-Leone, A. (2009). The role of the parietal lobe in visual extinction studied with transcranial magnetic stimulation. *Journal of Cognitive Neuroscience*, *21*(10), 1946-1955.

Bhattacharya, J. (2009). Increase of universality in human brain during mental imagery from visual perception. *PloS one*, *4*(1).

Bladin, P. F. (2006). W. Grey Walter, pioneer in the electroencephalogram, robotics, cybernetics, artificial intelligence. *Journal of clinical neuroscience*, *13*(2), 170-177.

Briki, W., & Hue, O. (2016). How red, blue, and green are affectively judged. *Applied Cognitive Psychology*, *30*(2), 301-304.

Buechner, V. L., Maier, M. A., Lichtenfeld, S., & Schwarz, S. (2014). Red-take a closer look. *PLoS One*, *9*(9).

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of cognitive neuroscience*, *12*(1), 1-47.

Caldwell, D. F., & Burger, J. M. (2011). On thin ice: Does uniform color really affect aggression in professional hockey?. *Social Psychological and Personality Science*, *2*(3), 306-310.

Caldwell, J. A., & Jones, G. E. (1985). The effects of exposure to red and blue light on physiological indices and time estimation. *Perception*, *14*(1), 19-29.

Chen, Y., Cai, L., Wang, R., Song, Z., Deng, B., Wang, J., & Yu, H. (2018). DCCA cross-correlation coefficients reveals the change of both synchronization and oscillation in EEG of Alzheimer disease patients. *Physica A: Statistical Mechanics and its Applications*, *490*, 171-184.

Cie, C. (1932). Commission internationale de l'eclairage proceedings, 1931. *Cambridge University*, *Cambridge*.

Commission Internationale de l'Éclairage. Vol. CIE S 014-4/E:2007 (ISO 11664-4:2008) (2007).

Costa, M., Frumento, S., Nese, M., & Predieri, I. (2018). Interior color and psychological functioning in a university residence hall. *Frontiers in psychology*, *9*, 1580.

Davidoff, J. (1976). Hemispheric sensitivity differences in the perception of colour. *The Quarterly journal of experimental psychology*, *28*(3), 387-394.

Dutta, S., Ghosh, D., & Chatterjee, S. (2013). Multifractal detrended fluctuation analysis of human gait diseases. *Frontiers in physiology*, *4*, 274.

Eke, A., Herman, P., Kocsis, L., & Kozak, L. R. (2002). Fractal characterization of complexity in temporal physiological signals. *Physiological measurement*, *23*(1), R1.

Elliot, A. J. (2015). Color and psychological functioning: a review of theoretical and empirical work. *Frontiers in psychology*, *6*, 368.

Elliot, A. J. (2019). A historically based review of empirical work on color and psychological functioning: Content, methods, and recommendations for future research. *Review of General Psychology*, *23*(2), 177-200.

Elliot, A. J., & Maier, M. A. (2012). Color-in-context theory. In *Advances in experimental social psychology* (Vol. 45, pp. 61-125). Academic Press.

Elliot, A. J., & Niesta, D. (2008). Romantic red: Red enhances men's attraction to women. *Journal of personality and social psychology*, *95*(5), 1150

Elliot, A. J., Maier, M. A., Binser, M. J., Friedman, R., & Pekrun, R. (2009). The effect of red on avoidance behavior in achievement contexts. *Personality and Social Psychology Bulletin*, *35*(3), 365-375.

Elliot, A. J., Maier, M. A., Moller, A. C., Friedman, R., & Meinhardt, J. (2007). Color and psychological functioning: the effect of red on performance attainment. *Journal of experimental psychology: General*, *136*(1), 154.

Erwin, C. W., Lerner, M., Wilson, N. J., & Wilson, W. P. (1961). Some further observations on the photically elicited arousal response. *Electroencephalography and Clinical Neurophysiology*, *13*(3), 391-394.

Esteller, R., Vachtsevanos, G., Echauz, J., Henry, T., Pennell, P., Epstein, C., ... & Litt, B. (1999, March). Fractal dimension characterizes seizure onset in epileptic patients. In *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No. 99CH36258)* (Vol. 4, pp. 2343-2346). IEEE.

França, L. G. S., Miranda, J. G. V., Leite, M., Sharma, N. K., Walker, M. C., Lemieux, L., & Wang, Y. (2018). Fractal and multifractal properties of electrographic recordings of human brain activity: toward its use as a signal feature for machine learning in clinical applications. *Frontiers in physiology*, *9*, 1767.

França, L. G. S., Montoya, P., & Miranda, J. G. V. (2019). On multifractals: a non-linear study of actigraphy data. *Physica A: Statistical Mechanics and its Applications*, *514*, 612-619.

Frank, M. G., & Gilovich, T. (1988). The dark side of self-and social perception: Black uniforms and aggression in professional sports. *Journal of personality and social psychology*, *54*(1), 74.

Franklin, A., Drivonikou, G. V., Bevis, L., Davies, I. R., Kay, P., & Regier, T. (2008). Categorical perception of color is lateralized to the right hemisphere in infants, but to the left hemisphere in adults. *Proceedings of the National Academy of Sciences*, *105*(9), 3221-3225.

Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: an fMRI study. *Cognitive Brain Research*, *20*(2), 226-241.

Gerard, R. M. (1958). *Differential effects of colored lights on psychophysiological functions* (Doctoral dissertation, University of California, Los Angeles.).

Ghosh, D., Dutta, S., & Chakraborty, S. (2014). Multifractal detrended cross-correlation analysis for epileptic patient in seizure and seizure free status. *Chaos, Solitons & Fractals*, *67*, 1-10.

Ghosh, D., Sengupta, R., Sanyal, S., & Banerjee, A. (2018). *Musicality of human brain through fractal analytics*. Springer Singapore.

Gong, P., Nikolaev, A. R., & van Leeuwen, C. (2003). Scale-invariant fluctuations of the dynamical synchronization in human brain electrical activity. *Neuroscience letters*, *336*(1), 33-36.

Greenlees, I. A., Eynon, M., & Thelwell, R. C. (2013). Color of soccer goalkeepers' uniforms influences the outcome of penalty kicks. *Perceptual and Motor Skills*, *117*(1), 1-10.

Hanada, M. (2018). Correspondence analysis of color–emotion associations. *Color Research & Application*, *43*(2), 224-237.

Hanson, A. R. (2012). What is colour?. In *Colour Design* (pp. 3-23). Woodhead Publishing.

Hill, R. A., & Barton, R. A. (2005). Red enhances human performance in contests. *Nature*, *435* (7040), 293-293.

Hutmacher, F. (2019). Why Is There So Much More Research on Vision Than on Any Other Sensory Modality?. *Frontiers in psychology*, *10*, 2246.

Ihlen, E. A. F. E. (2012). Introduction to multifractal detrended fluctuation analysis in Matlab. *Frontiers in physiology*, *3*, 141.

Ihlen, E. A., & Vereijken, B. (2010). Interaction-dominant dynamics in human cognition: Beyond $1/f\alpha$ fluctuation. *Journal of Experimental Psychology: General*, *139*(3), 436.

Ishihara, S. (1987). *Test for colour-blindness*. Tokyo, Japan: Kanehara.

Ivanov, P. C., Amaral, L. A. N., Goldberger, A. L., Havlin, S., Rosenblum, M. G., Struzik, Z. R., & Stanley, H. E. (1999). Multifractality in human heartbeat dynamics. *Nature*, *399*(6735), 461-465.

Jun, W., & Da-Qing, Z. (2012). Detrended cross-correlation analysis of electroencephalogram. *Chinese Physics B*, *21*(2), 028703.

Kaiser, P. K. (1984). Physiological response to color: A critical review. *Color research & application*, *9*(1), 29-36.

Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., & Stanley, H. E. (2002). Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, *316*(1-4), 87-114.

Karkare, S., Saha, G., & Bhattacharya, J. (2009). Investigating long-range correlation properties in EEG during complex cognitive tasks. *Chaos, Solitons & Fractals*, *42*(4), 2067-2073.

Kleinschmidt, A., Büchel, C., Hutton, C., Friston, K. J., & Frackowiak, R. S. (2002). The neural structures expressing perceptual hysteresis in visual letter recognition. *Neuron*, *34*(4), 659-666.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, *29*(2-3), 169-195.

Labrecque, L. I., & Milne, G. R. (2012). Exciting red and competent blue: the importance of color in marketing. *Journal of the Academy of Marketing Science*, *40*(5), 711-727.

Lee, S., & Rao, V. S. (2010). Color and store choice in electronic commerce: The explanatory role of trust. *Journal of Electronic Commerce Research*, *11*(2).

Liao, F., & Jan, Y. K. (2011). Using multifractal detrended fluctuation analysis to assess sacral skin blood flow oscillations in people with spinal cord injury. *J Rehabil Res Dev*, *48*(7), 787-800.

Linkenkaer-Hansen, K., Nikouline, V. V., Palva, J. M., & Ilmoniemi, R. J. (2001). Long-range temporal correlations and scaling behavior in human brain oscillations. *Journal of Neuroscience*, *21*(4), 1370-1377.

Lockley, S. W., Evans, E. E., Scheer, F. A., Brainard, G. C., Czeisler, C. A., & Aeschbach, D. (2006). Short-wavelength sensitivity for the direct effects of light on alertness, vigilance, and the waking electroencephalogram in humans. *Sleep*, *29*(2), 161-168.

Lotto, R. B., & Purves, D. (2002). The empirical basis of color perception. *Consciousness and Cognition*, *11*(4), 609-629.

Maity, A. K., Pratihar, R., Mitra, A., Dey, S., Agrawal, V., Sanyal, S., ... & Ghosh, D. (2015a). Multifractal detrended fluctuation analysis of alpha and theta EEG rhythms with musical stimuli. *Chaos, Solitons & Fractals*, *81*, 52-67.

Mandelbrot, B. B. (1983). *The fractal geometry of nature* (Vol. 173, p. 51). New York: WH freeman.

McFarland, D. J., Parvaz, M. A., Sarnacki, W. A., Goldstein, R. Z., & Wolpaw, J. R. (2016). Prediction of subjective ratings of emotional pictures by EEG features. *Journal of neural engineering*, *14*(1), 016009.

Mehta, R., & Zhu, R. J. (2009). Blue or red? Exploring the effect of color on cognitive task performances. *Science*, *323*(5918), 1226-1229.

Mikellides, B. (1990). Color and physiological arousal. *Journal of Architectural and Planning Research*, 13-20.

Nakshian, J. S. (1964). The effects of red and green surroundings on behavior. *The Journal of General Psychology*, *70*(1), 143-161.

Newson, J. J., & Thiagarajan, T. C. (2019). EEG frequency bands in psychiatric disorders: a review of resting state studies. *Frontiers in human neuroscience*, *12*, 521.

Njemanze, P. C., Gomez, C. R., & Horenstein, S. (1992). Cerebral lateralization and color perception: a transcranial Doppler study. *Cortex*, *28*(1), 69-75.

Nourse, J. C., & Welch, R. B. (1971). Emotional attributes of color: A comparison of violet and green. *Perceptual and motor skills*, *32*(2), 403-406.

Olsen, J. (2010). The effect of color on conscious and unconscious cognition. *Dietrich College Honors Theses*, *72*, 1-31

Paluš, M. (1996). Nonlinearity in normal human EEG: cycles, temporal asymmetry, nonstationarity and randomness, not chaos. *Biological cybernetics*, *75*(5), 389-396.

Pereda, E., Gamundi, A., Rial, R., & González, J. (1998). Non-linear behaviour of human EEG: fractal exponent versus correlation dimension in awake and sleep stages. *Neuroscience letters*, *250*(2), 91-94

Pike, G., Edgar, G., and Edgar, H. (2012). "Perception," in *Cognitive Psychology*, eds N. Braisby and A. Gellatly (Oxford: Oxford University Press), 65–99.

Podobnik, B., & Stanley, H. E. (2008). Detrended cross-correlation analysis: a new method for analyzing two nonstationary time series. *Physical review letters*, *100*(8), 084102.

Pritchard, W. S. (1992). The brain in fractal time: 1/f-like power spectrum scaling of the human electroencephalogram. *International Journal of Neuroscience*, *66*(1-2), 119-129.

Pritchard, W. S., & Duke, D. W. (1992). Measuring chaos in the brain: a tutorial review of nonlinear dynamical EEG analysis. *International Journal of Neuroscience*, *67*(1-4), 31-80.

Pritchard, W. S., & Duke, D. W. (1995). Measuring chaos in the brain-a tutorial review of EEG dimension estimation. *Brain and cognition*, *27*(3), 353-397.

Ridgway, J., & Myers, B. (2014). A study on brand personality: consumers' perceptions of colours used in fashion brand logos. *International Journal of Fashion Design, Technology and Education*, *7*(1), 50-57.

Roy, C., Roy, S., & Ghosh, D. (2016). Chaos Based Study on Association of Color with Music in the Perspective of Cross-Modal Bias of the Brain. *CIC 2016 Special Issue, International Journal of Computer Science and Information Security, Volume 14, 2016, pages 36-42, ISSN: 1947-5500*

Roy, S., Roy, C., Nag, S., Banerjee, A., Sengupta, R., & Ghosh, D. (2020). Chaos based non-linear cognitive study of different stimulus in the cross-modal perspective. *Physica A: Statistical Mechanics and its Applications*, *546*, 122842.

Sanyal, S., Nag, S., Banerjee, A., Sengupta, R., & Ghosh, D. (2019). Music of brain and music on brain: a novel EEG sonification approach. *Cognitive neurodynamics*, *13*(1), 13-31.

Schreiber, T., & Schmitz, A. (1996). Improved surrogate data for nonlinearity tests. *Physical review letters*, *77*(4), 635.

Shang, P., Lu, Y., & Kamae, S. (2008). Detecting long-range correlations of traffic time series with multifractal detrended fluctuation analysis. *Chaos, Solitons & Fractals*, *36*(1), 82-90.

Shen, Z., Tone, A., & Asayama, M. (1999). The effects of viewing different colors on EEG and skin temperature in humans. *Journal of International Society of Life Information Science*, *17*(1), 105-117.

Shi, J., Zhang, C., & Jiang, F. (2015). Does red undermine individuals' intellectual performance? A test in China. *International Journal of Psychology*, *50*(1), 81-84.

Siok, W. T., Kay, P., Wang, W. S., Chan, A. H., Chen, L., Luke, K. K., & Tan, L. H. (2009). Language regions of brain are operative in color perception. *Proceedings of the National Academy of Sciences*, *106*(20), 8140-8145.

Spillmann, L., & Werner, J. S. (Eds.). (2012). *Visual perception: The neurophysiological foundations*. Elsevier.

Stanley, H. E., Amaral, L. N., Goldberger, A. L., Havlin, S., Ivanov, P. C., & Peng, C. K. (1999). Statistical physics and physiology: monofractal and multifractal approaches. *Physica A: Statistical Mechanics and its Applications*, *270*(1-2), 309-324.

Stephen, I. D., & McKeegan, A. M. (2010). Lip colour affects perceived sex typicality and attractiveness of human faces. *Perception*, *39*(8), 1104-1110.

Suckling, J., Wink, A. M., Bernard, F. A., Barnes, A., & Bullmore, E. (2008). Endogenous multifractal brain dynamics are modulated by age, cholinergic blockade and cognitive performance. *Journal of neuroscience methods*, *174*(2), 292-300.

Suetsugi, M., Mizuki, Y., Ushijima, I., Kobayashi, T., Tsuchiya, K., Aoki, T., & Watanabe, Y. (2000). Appearance of frontal midline theta activity in patients with generalized anxiety disorder. *Neuropsychobiology*, *41*(2), 108-112.

Suk, H., & Irtel, H. (2010). Emotional response to color across media.*Color Research and Application, 35,* 64–77. http://dx.doi.org/10.1002/col.20554

Telesca, L., & Lovallo, M. (2011). Analysis of the time dynamics in wind records by means of multifractal detrended fluctuation analysis and the Fisher–Shannon information plane. *Journal of Statistical Mechanics: Theory and Experiment*, *2011*(07), P07001.

Thakor, N. V., & Tong, S. (2004). Advances in quantitative electroencephalogram analysis methods. *Annu. Rev. Biomed. Eng.*, *6*, 453-495.

Thiagarajan, T. (2018). Interpreting electrical signals from the brain. *Acta Physica Polonica B*, *49*(12), 2095-2125.

Valdez, P., & Mehrabian, A. (1994). Effects of color on emotions. *Journal of Experimental Psychology. General*, *123*, 394-409.

Vandewalle, G., Schmidt, C., Albouy, G., Sterpenich, V., Darsaud, A., Rauchs, G., ... & Maquet, P. (2007). Brain responses to violet, blue, and green monochromatic light exposures in humans: prominent role of blue light and the brainstem. *PloS one*, *2*(11).

von Castell, C., Stelzmann, D., Oberfeld, D., Welsch, R., & Hecht, H. (2018). Cognitive performance and emotion are indifferent to ambient color. *Color Research & Application*, *43*(1), 65-74.

Von Goethe, J. W. (1810). Theory of Colours (CL Eastlake, Trans.). *London: Frank Cass & Co.(Original work published 1967)*.

Voytek, B., Kramer, M. A., Case, J., Lepage, K. Q., Tempesta, Z. R., Knight, R. T., & Gazzaley, A. (2015). Age-related changes in 1/f neural electrophysiological noise. *Journal of Neuroscience, 35*(38), 13257-13265.

Walter, W. G. (1938). Critical review: the technique and application of electro-encephalography. *Journal of neurology and psychiatry*, *1*(4), 359.

Wilms, L., & Oberfeld, D. (2018). Color and emotion: effects of hue, saturation, and brightness. *Psychological research*, *82*(5), 896-914.

Witzel, C., & Gegenfurtner, K. R. (2011). Is there a lateralized category effect for color?. *Journal of vision*, *11*(12), 16-16.

Yoto, A., Katsuura, T., Iwanaga, K., & Shimomura, Y. (2007). Effects of object color stimuli on human brain activities in perception and attention referred to EEG alpha band response. *Journal of physiological anthropology*, *26*(3), 373-379.

Yuan, Y., Zhuang, X. T., & Jin, X. (2009). Measuring multifractality of stock price fluctuation using multifractal detrended fluctuation analysis. *Physica A: Statistical Mechanics and its Applications*, *388*(11), 2189-2197.

Zhou, W. X. (2008). Multifractal detrended cross-correlation analysis for two nonstationary signals. *Physical Review E*, *77*(6), 066211.

Zorick, T., & Mandelkern, M. A. (2013). Multifractal detrended fluctuation analysis of human EEG: preliminary investigation and comparison with the wavelet transform modulus maxima technique. *PloS one*, *8*(7).

CHAPTER 5

A NOVEL APPROACH TO CHARACTERIZE THE COMPLEXITY OF RAGA AND ITS RENDITIONS USING STATISTICAL METHODS

"The pleasure we obtain from music comes from counting, but counting unconsciously. Music is nothing but unconscious arithmetic."

Gottfried Leibniz

ABSTRACT

Information is usually made up of repetitive arrangements of basic patterns. The presence of such sequences contributes to the unique style/behaviour the information congregation represents. Various such congregations like music, language, biological signals exhibit this kind of repetitive patterns of symbolic sequences. The problem of categorizing such information collection is that the usual methods used are mostly non-quantifiable. In this study, we try to quantify such abstractness using measurable parameters. For that, we introduce methods based on well-known concepts used in Statistical Physics (especially thermodynamics), namely Maxwell-Boltzmann (MB) statistics and Bose-Einstein (BE) distribution, in an attempt of categorization and classification of musical information present in Hindustani classical music. Both MB and BE statistics have wide applications outside the realm of describing the energy level occupation of elementary particles. For example: the usage of statistical physics in the domain of linguistics or social sciences. Here, statistical methods based on these distributions have been applied to find new parameters (equivalent to 'temperature' in physical systems) to distinguish between different features of different ragas in Hindustani classical music. To apply MB statistics to music, it is assumed that different notes (combined with their durations) with different occurrence frequencies are at different energy levels, the distribution of which follows the MB distribution pattern. Emerging 'temperature' parameter shows how close the said rendition is with the traditional grammatical structure of the said raga (or the amount of improvisations/creativity present in the rendition). In case of BE statistics, a rank-frequency distribution of the time durations of various notes occurring in different ragas is studied to obtain 'temperature-alike' statistical parameters, unprecedented in any previous studies. This novel method of analysis is then applied on different renditions of the same Raga. Music clips chosen were the *Vandish* part of the same raga sung by three legendary classical music maestros. The resulting analysis gives a number of parameters (they come from the analogy between the rank-frequency distribution and the respective statistical distribution) that help categorize the singing styles of the three renditions and parameters which are indicative of abstract ideas such as individual improvisation pattern. The results found shed new lights on the classification of artists' style and the difference in their renditions of the same raga.

Keywords: Maxwell-Boltzmann distribution, Bose-Einstein distribution, Indian classical music, Raga, temperature

5.1. INTRODUCTION: BRIEF NOTE ON COMPLEXITY

Humans and information have shared an intimate relationship throughout the course of life and history. Since birth, the world around is perceived to us via sensory channels in the form of information, a barrage of which helps gain knowledge of the environment and aids in the necessary evolution of life and its principles. Information is rather easy to describe in a colloquial approach, i.e., it is an answer to any posed question. This establishes the connection between information and the acts of observation, interrogation and subsequently, measurement. This is where causality takes a front seat as well. In fact, the idea of a 'natural law' is built upon the fact in natural systems, that causal processes should correspond implying it into an appropriate and deducible system. Newtonian Mechanics, a flagbearer of such paradigm, described natural phenomenon using simple fundamental laws that bore information about the cause and the effect, thus cementing the idea of causality as a fundamental part of nature. Furthermore, Newtonian mechanics established the boundaries of the system and its environment via two kinds of parameters: *intrinsic*, which is used to describe the system and its evolution and *extrinsic*, used in describing the environment. Robert Rosen have renamed these two as 'genome' and 'environment', respectively, to establish a biological dynamism (he also categorises the state variables as 'phenotypes') and argued that the genotype-phenotype duality exceeds the organic existence to have a universal representation (Rosen, 2012). Systems that follow the Newtonian paradigm are regarded as simple systems. Complex systems are those that fall outside this realm. Because of this paradigm's insistence on reductionism and determinism, stalwarts like Kant and Laplace have argued that a large body of naturally occurring systems, mainly biological ones, reside outside the Newtonian canon. Einstein, summarising this view aptly, wrote in a letter: "One can best appreciate, from a study of Living things, how primitive physics still is". Complex systems, generally unpredictable, counterintuitive and highly susceptible to both internal and external changes, exhibit features such as:

- Complexity is an intrinsic property of the system, not unlike any other internal properties.
- Complex systems can be approximated by simple Newtonian systems, but only locally and temporarily, implying that the former is devoid of the reductionist nature of the latter.
- Unlike simple systems and their single 'master descriptor' approach, complex systems consist of multiple dynamic descriptions that requires to be updated for specifying the nature of the system appropriately.
- Causality works very differently for complex systems. Addition to the fact that they lack fixed state parameters to address their state, any changes in the manifestation can independently cause changes in different layers in the system, thereby acting both as a cause and an effect simultaneously.
- The degree of complexity can be specified by one or more sets of numbers like dimensionality of space or length of algorithm etc.

Hence, the traditional Newtonian paradigm is not sufficient in adequately chronicle the complex world which demands special constraints and conditions or, to return to the starting point, special nature of information.

Since our environment, and the world around us is mostly observed to be of such complex nature, it becomes necessary for science to tackle these complex systems and complex inputs in ways that are more than approximations. Complexity is sometimes referred to as the intermediate state between order and disorder. This leads to a confusion between randomness and complexity. Grassberger (1986) have used three diagrams similar to Fig 5.1 to describe the systems and their intuitive distinguishability, or rather a lack of thereof. In the figure, 5.1c is the most disordered pattern with highest complexity as it is generated from a random number generator and 5.1a is the lowest. But apparently for most, 5.1b is the most complex pattern as it has a 'structure' more prominent than the one on its right. By this, Grassberger argues, that the idea of complexity, intuitively, is not a monotonically increasing function of disorderness. Instead, the peak lies somewhere in between complete order and complete chaos, quite similar to a living being that is more complex than either perfect crystal and random glass.

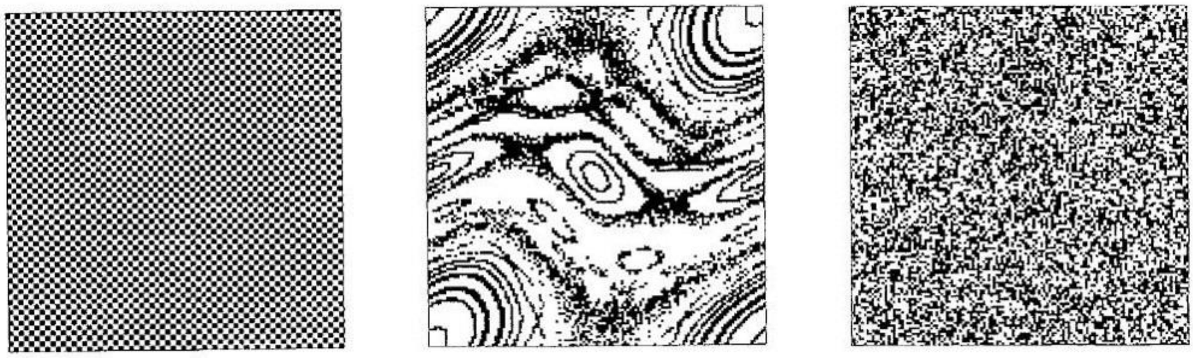


Fig. 5.1(a-c). Three figures used by Grassberger (1986) to illustrate the intuitive difference between order, complexity and randomness. Here, 5.1a has the most order and lowest complexity, followed by 5.1b and 5.1c has the highest complexity and lowest order/most randomness. But for majority, 5.1b is the most complex pattern among the three.

To address such confusions, Ladyman et al have attempted to pinpoint a list of properties that are associated with the concept of complexity (Ladyman et al. 2013):

1. Nonlinearity: It is when the principle of superposition is not applicable to a system. Although there are examples indicating that nonlinearity (and chaos) is neither necessary nor sufficient for complex systems, the authors admit that nonlinear dynamics do play a role in complexity nonetheless.
2. Feedback: Feedback involves interaction of a part of a system with its neighbours in future subjected to the nature of interaction the said part did with them in past. This is a necessary but not sufficient condition for complex dynamics.
3. Spontaneous order: Since complexity doesn't correspond to either complete order or complete disorder, it is necessary for it to exhibit spontaneous order of some kind.
4. Robustness and lack of central control: Complex phenomena is usually robust as perturbations have no effect on its order since it doesn't have any. Similarly, lack of central control is also a part of it since it remains stable under perturbation by self-correction of errors. Both these features are necessary but not sufficient.
5. Emergence: Emergence of properties, objects or processes is also an important feature of complexity. Although its definition remains confusing, but authors argue that precise epistemological characterization of emergence is a necessary condition but not sufficient as simple systems can also exhibit emergence.

6. Hierarchical organisation: Another important feature whose existence guarantees robust interaction, order, symmetries and periodicity in the system, resulting in emergence. However, it must be noted that this is subjected to strict conditions.
7. Numerosity: The above conditions are satisfied when the system is made up of large parts and they share many interactions among different levels.

The authors summarize their argument in a sentence characterizing complex system as it "..is an ensemble of many elements which are interacting in a disordered way, resulting in robust organisation and memory". However, the qualitative notion of it is given as the ability to generate high 'statistical complexity'. Besides above discussion, numerous other works have spawned that have discussed the notion as well (Crutchfield & Young, 1989; Feldman & Crutchfield, 1998; Shiner et al., 1999; Mitchell, 2009; Zurek, 2018) and most of them agreed that complexity should reflect some hidden order of the system and it is closely related to the degree of randomness (Ebeling et al., 2002).

The practical exploration of complexity has unfurled in various fields of science, including obvious ones like thermodynamics (Lloyd & Pagels, 1988; Crutchfield & Young, 1989; Shiner et al., 1999) and dynamical systems theory (Witt et al., 1997; Boccaletti et al., 2002), to Economic theory (Mantegna & Stanley, 1999), Time series analysis (Zou et al., 2019), Information processing (Chaitin, 1966), Astrophysics (Schwarz et al., 1993), Geophysics (Witt et al., 1994), diagnosis of medical ailments (Kurths et al., 1995; Saparin et al., 1998; Marwan et al., 2002) and so on. Another interesting field of application, which is of our particular interest in this chapter, is analysis of symbolic sequences (Ebeling & Nicolis, 1991, 1992; Ebeling & Pöschel, 1994; Ebeling et al., 2002).

5.2. SYMBOLIC SEQUENCES AND ITS IMPLICATIONS

In the study of complex systems, symbolic sequences have an important role to play as it is believed that most complex systems can be reduced to such sequences, meaning they are mostly an intergral part of the system under observation. In nature, biological examples like genetic codes or neural spike activities are some of the excellent examples of symbolic sequences. Heart rate dynamics, human speech and language are also said to exhibit such sequences in their core structures. There are generally two connotations of these symbolic sequences. The first case includes sequences that are used to carry and transmit informations themselves. Neural spike activities propagating neuron to neuron is such an example of information transmission. In some alternate cases, instead of carrying information and aiding in a dynamical process, a symbolic sequence could be found to be an end product of such a process (also called biomarkers). An example of this kind would be the sleep stage trnsitions. Here, sequencial patterns of sleep stages are a by-product of the sleep dynamics during the entire period under consideration. Besides the above two cases, there exist a of a third possibility, where both these factors can combine and influence the symbolic sequences involved. DNA sequencing as an illustration of this. They are a modified by-product of various stochastic mutation processes while simultaneously carrying important biological information on basis of which they get prioritised and survive via natural selection. Although the naturally occuring symbolic sequences belong to three different scenarios above, however, a common thread binds all of them together. That is, they tend to exhibit characteristic 'styles' or 'signatures' that are identifiable by certain complex but repetitive patterns. Patterns that almost always shed light on the dynamical processes associated with the sequences.

From the perspective of the system that ceaselessly generates these information carrying sequences, they are usually strings on an infinite alphabet paradigm. For instance, a book may be considered as a regular linear string of alphabets and letters. Actually any trajectory of a dynamic system could be mapped or coded to (with) a string of letters on a chosen alphabet module using the methods of symbolic dynamics. Same approach could be used to describe other sequential information carriers like amino acids, nucleotides, human speech and music, digital informations, computers programs as well. Now, let's consider speech and music. Any form of speech and music are processes that is generated from the continuous interplay of frequencies and amplitutes. Although, in naturally occurring phenomenons, or atleast in the parts of that phenomen carrying most bit of information, the sequences are coded on finite number of alphabets. Hence, language (English) can be reduced to a finite number of 'alphabets', i.e., 2^5, with the pauses and punctuations counted as ones (Ebeling & Nicolis, 1991). Similarly, considering notes as alphabets, a music piece can be coded with a string of alphabet numbers ranging 2^5-2^6. For biological information, this value ranges from 4 for DNA/RNA (2^2) to 20 for proteins (~2^4), whereas only 2 (2^1) for digital data, 0 and 1. Now, usually the strings generated this way was only known to have short range correlations. But in various instances, it has been found that they exhibit long-range correlations too. The examples range from nucleotide sequences (Peng et al., 1992), human writing (Schenkel et al., 1993), music (Jafari et al., 2007), seismic data (Martin-Montoya et al., 2015) to stock market (Costa & Vasconcelos, 2003). Clearly, the structure of the sequences and their statistical properties indicate towards the dynamics hidden in the larger complexity of the system involved. In view of this, it is justified to seek whether there are relevant patterns of symbolic sequences present in other fields that remain unexplored till now and whether those patterns, if present, could be identified and quantitatively characterized to access the dynamical processes embedded in the heart of said systems.

It has been found that linguistic like processes that involves information string sequences show nonlinearity, self-similarity and long-range correlations (Ebeling et al., 2002). Now, let us see how the sequences might evolve. Let us consider a string

$$S = A_1\ A_2\ A_3\ A_4 \ldots\ldots.. A_L \tag{1}$$

Here, the sting S is generated by a stationary dynamical process which could be a text or a music composition and A_i are letters describing the string. Also, $L \rightarrow \infty$. We assume that the letters A_i belong to a system of alphabets which has λ letters in it.

$$A_i \subset (\ A^{(1)}\ A^{(2)} \ldots\ldots.. A^{(\lambda)}) \tag{2}$$

Now, a subsequence or 'subword' of length n is defined as a substring of length n collected from the string S. The total number of existing subwords of length n in this string is (Ebeling & Nicolis, 1992):

$$N_{n\lambda} = \lambda^n \tag{3}$$

This is an astronomical number for high n's ($n > 50$). It suggests that although the possibility of naturally ocurring substring patterns are endless, nature or natural processes deploy a very strict restrictive conditioning in choosing their existence. Hence, the 'effective subwords' number, $N^*_{n\lambda}$, subwords of length n of λ letters, is counted as the total number of frequent subwords of length n. A subword or subsequence is frequent, when it exists more than once in the string or sequence and its relative occurrence exceeds the Bernoulli probability $p^{(n)} \geq 1/\lambda^n$

(single trial binomial distribution, e.g., coin toss where $\lambda = 2$, $n = 1$). Another way to define $N^*_{n\lambda}$ could be the using Eq. (4):

$$\log_\lambda N^*_{n\lambda} = < \log_\lambda (1/p^{(n)}) > \quad (4)$$

This definition is related to entropy (Ebeling & Nicolis, 1992) and it assumes the same value as Eq. (3) in case of an equal distribution where each value of a random variable is equally probable to occur. In various examples of language-like texts, the L.H.S of eq. (4) scales in a sublinear manner which is attributed to the long-range correlations present in them (Ebeling & Nicolis, 1991). As *n* grows larger, $N^*_{n\lambda}$ also grows in a way which is of importance, specially regarding this chapter:

$$N^*_{n\lambda} \rightarrow C.\, n^{\alpha}, \text{ when } n \rightarrow \infty \quad (5)$$

The relation in eq. (5) is in a form of a power law equation (with exponent α), something that is observed in a diverse range of natural processes throughout the scientific history: from moon craters (Neukum & Ivanov, 1994) to distribution of biological species (Willis & Yule, 1922). Even the network of numbers of citations received by scientific papers shows power law like behaviour (Price de Solla, 1965). In the next few sections, we shall use this idea to examine the dynamics of a musical piece, whose note sequence patterns also exhibit power law relations.

5.3. POWER LAWS AND THEIR APPLICATION IN MUSIC

Statistically put, the distribution of a random variable is said to have power law characteristics if it has a probability density function of the form of eq. (6):

$$p(x) = C.x^{-\alpha} \quad (6)$$

Here, p(x) is the probability of finding the variable x to have a certain value. C and α are two real-valued constants, latter of which is called the exponent of the power law. Compared to the other terms, C is largely of ornamental value and is computed by equating the sum of the distribution of p(x) to 1. The significance of a power law resides in the fact that it is indicative of the self-similar nature of the experimental data. That is to say, the statistical characteristics of the data fraction under observation takes after the entire data of which it is a part, hence making the fraction so called 'scale-free' (Faloutsos et al., 1999). Also, it is crucial to mention that self-similarity is reported to have been observed in musical compositions, under temporal transition (Foote, 1999) as well as in scaling (Pareyon, 2011). In most of the power law distributions that are found in nature, the scaling exponent α has a value between 2 and 3, i.e., $2 \leq \alpha \leq 3$.

Zipf's law is one variation that comes under the huge umbrella of power laws. The principle given by George Zipf implies that any efforts made by a 'self-adapting agent' (humans) to interact with its surroundings would tend to be minimized, that is, the overall effort to carry out such interactions remains very close to a global optimum level (Zipf, 1949). When the logarithm of the frequencies of the interactions are plotted against the logarithm of ranks of the same, the slope of the corresponding straight line almost always is measured to be approximately -1. Although Zipf demonstrated this result using linguistic examples, it has been applied to many natural and man-made phenomena since with remarkable accuracy (Li, 2002). Zipf's law differs from most of the other members of the power law family on two accounts: firstly, the data is interpreted in terms of rank-frequency plots, equivalent to the cumulative distribution of the observed variable and secondly, the value of the exponent is found to be between 1 and 2; hence, $1 \leq \alpha \leq 2$, for Zipf distribution. It is of note that detailed discussions regarding Zipf's law (and power laws in general) has been done already in Chapters 1 and 2. So, for this chapter, we will restrict ourselves in its application on music related fields only.

An ideal case of Zipf's law appears when the exponent value is exactly 1. It is defined as 1/f noise or 'pink noise'. Zipf studied the existence of 1/f distributions in melodic intervals of several western classical pieces, including Mozart (Bassoon Concerto, B-flat) and Chopin (Etude, F minor). He also studied the distance between repeating notes. Later, in their seminal works, Voss & Clarke studied a wide range of music genres and found Zipf distributions in pitch and loudness fluctuations (Voss & Clarke, 1975, 1978). Their pioneering work included the aesthetic astuteness of generated music. The music generated from pink noise source was found to be more pleasant than non-pink noise sources. Following their footsteps, scale free nature of frequency interval distributions in Bach and Mozart compositions was observed by Hsu and Hsu (1991). Pitch of notes and duration also follow Zipf's law, Zanette discovered, along with the fact that note progressions create a musical 'context' not unlike a linguistic framework (Zanette, 2006). Manaris et al. worked with different genres and composers and considered variety of parameters like pitch, duration, melodic intervals, harmonic consonances etc to find Zipf like distributions in majority of samples across the data set. Moreover, they used Zipf-like metrics as the basis of tasks like author attribution, style identification and assessment of pleasantness to various degrees of success (Manaris et al., 2002, 2005). Also, in a recent work, Liu et al. (2013) found scale-freeness in the distribution of pitch fluctuations of several western classical composers. Taking a different route from the studies above, usage of a Beta rank law to improve upon Zipf's law were attempted by del Rio et al (2008). They argue that such model can be used as a source of testing music aesthetics. Very recently, Perotti & Billoni (2020) set out to investigate the emergence of Zipf's law from the structure of a composition and found that it emerges when a combination of chords and notes are chosen as so called 'Zipfian units'. Additionally, they suggest that emergence of Zipf's law is a consequence of human language evolving into a more complex but efficient form of communication.

Based on this premise, we attempt our work with a two-fold objective. First, to examine the scope of applying statistical models and power law distributions in the field of Indian classical music which by its nature is significantly different from the musical practices in the west. And secondly, to advance a step further in the analysis process by introducing hardcore statistical methods like Maxwell-Boltzmann (MB) and Bose-Einstein (BE) distributions to categorize and quantify the symbolic sequences from the intrinsic structural patterns. Hence, it seems necessary to discuss both the Indian music scenario and the existing research regarding the MB-BE domain before moving on to our work.

5.4. BRIEF NOTE ON INDIAN CLASSICAL MUSIC

Indian classical music, largely divided into two categories: Hindustani Classical Music (HCM) or the northern Indian music practice and Carnatic Classical Music (CCM) or the southern Indian music practice, has a rich history whose evolution covers three general periods: Ancient, Medieval and Modern.

Ancient Period signifies the Vedic Age (upto around 2000 B.C), chronicled mostly in Vedic literatures. It was a common practice for the sages to sing and their wives to play instruments like *Veena*. Amongst the four Vedic scriptures, "*Saam Veda*" is believed to be the origin of musical chanting in India. The religious chantings in *Saam Veda*, known as '*Saam gaan*', were recited vocally. Here, three types of *swaras* (i.e tones) were used – *Anudatta* (low pitch), *Udatta* (high pitch) and *Swarit* (between low and high pitches).

The golden age of Gandharva music came in later Vedic period, during days of Ramayana and Mahabharata. *Gandharva Veda* is the study of all art forms including music, dance and poetry. It is believed that during this era, the seven *swaras* (notes) (as in *Sa, Re, Ga, Ma, Pa, Dha, Ni*) came into existence. In the post Vedic era, music remained quite popular according to Kautilya's *Arthashastra* (Rangarajan, 1992). Vatsayan during second century A.D. asserted Indian music to be a confluence of three aspects: *Geetam* (vocal music), *Vadyam* (instrumental music) and *Nrityam* (Dance).

Medieval period started from 7^{th} century and continued till 18th century AD, Indian music played a key role in India and outside. In the start of this period, Indian music was a symbol of Hindu philosophy and religious ideas. Scholarly books, like "*Geetgobindo*" by Jaidev and "*Sangeet Ratnakar*" by Sarangdev, on music were an important contribution of this era. Also, with the rise of Sri Chaitanya, the devotional genre of *Kirtan* became very popular in eastern parts of India. Other kinds of devotional music genres (*Bhajan)* were also introduced by famous saints like Meerabai, Kabir, Tulsidas, Ramdas among others. Parallel to the religious influence, music blossomed culturally, influenced by the lifestyles and philosophy of life, birthing several folk genres across the country as an outcome.

In the latter half of medieval period, between the 9th and 12th centuries, Indian classical music saw revolutionary qualitative changes as Muslim rulers from middle-east invaded and ruled India and these influences made an everlasting impact. From here onwards, North and South Indian Classical Music evolved as separate streams. During Alauddin Khiljee's reign (1296-1316 AD), the musical genius Amir Khusro used Tabla and Sitar as instruments for the first time and created new *Ragas*. He also introduced vocal music genres like *Kawali* and *Tarana*. King Akbar (1556-1605 AD), also an ardent lover of music, patronized several maestros like Nayak Bakru, Mian Tansen, Tantarang Gopal etc in his court. Mian Tansen is credited for *Ragas* like *Darbari Kanada, Mian ki Sarang, Mian ki Malhar*. Another way the Mughal period contributed substantially to the development of Classical music is by originating Khayal, Dhrupad and Thumri during this era.

Modern period of Indian classical music is generally marked since the end of the eighteenth century. This period saw the decline of Mughal empire and the rise of British colonialism, mostly impervious in the cultural front. Since the days of court sponsored musicians were now over, the musical masters started confining their musical knowledge only among the family and close disciples. This brought about the practice of *Gurukul* system. The beginning of 20th century saw the revival of Indian classical music with Pt. Bishnu Digambar Paluskar and Pt. Vishnu Narayan Bhatkhananda in the forefront. The latter invented the process of notation which is followed till date; though unlike Western Classical Music, any efficient way of writing silence or note transitions is yet to exist.

The 20^{th} century witnessed a barrage of excellence in Indian Classical music with names like Ustad Amir Khan, Ustad Alauddin Khan, Ustad Ali Akbar Khan, Pandit Onkarnath Tahkur, Tarapada Chakravarty, Pandit Ravi Shankar, Pandit Bhimsen Joshi, Pandit Hariprasad Chaurasia, Pandit Shivkumar Sharma, Pandit Ajoy Chakraborty, Ustad Zakir Hussain etc. Also, superlative influences of poets like Nobel laureate Rabindranath Tagore followed by Kazi

Nazrul Islam and Atulprasad Sen must be mentioned as trendsetters in the musical discourse. Along with the advancement of technology and invention of different instruments, composers and performers tried to mingle different genres of music. As a result, composed music evolved with lesser restrictions imposed by the *Raga* structure. Legends like Lata Mangeshkar, Manna Dey, Kishore Kumar and many other prolific singers ruled the hearts of millions of people in the latter half of the 20th century.

Comparison with Western Classical Music

Although Indian classical music is considered to be one of the oldest musical traditions in the world but compared to Western music very little work has been done in the areas of genre recognition, classification, automatic tagging, comparative studies etc. The following are broad points in which Western Classical and Indian Classical (both Hindustani and Carnatic) music differ from one another:

Homophony vs. Polyphony / Melody vs. Harmony

Indian classical music is primarily homophonic, which means its focus is on melodies created using a sequence of notes. The musical fluidity is primarily experienced with different melodies constructed within the framework of the ragas, while Western classical music's magic lies to a great extent in polyphonic composition, where counterpoint, harmony, and the texture created using multiple voices is very complex. Melody exists in Western classical music too, but from a broad perspective, is not the singular or defining focus of most of Western classical music works.

Composed vs Improvised

Western classical music is composed, Indian classical music is improvised. All Western classical music compositions are formally written using the staff notation, and performers have very little scope for improvisation. The converse is the case with Indian classical music, where no 'work' is ever written down, and the teacher-student tradition of learning Indian classical music leads to each performance being an improvised one.

Group vs Individual dynamics

In Indian classical music performances, the individual performer shines through his/her improvisations. In any recital or performance, there is a lead vocalist or instrumentalist who expounds the raga, while others provide accompaniment. In Western classical orchestrations, the composer and conductor shine as individuals, but the performance is largely a group effort. It is only in solo works and solo concertos that individual performers are under the spotlight.

Rhythm

In Indian classical music tempo is a basic component of any rendering. It is maintained through different *'Taal'*s. A *Taal* is a cycle of specific (odd or even) number of beats centered around 'Sam' that repeats itself. In this genre of music, keeping the *taal* fixed, rhythmic variations are done by performers as a part of musical improvisations. Though variations in tempo are observed in a complete *raga* performance but the change does not occur naturally as part of the music piece. Western classical music does not use the concept of *'Taal'* or any complex beat cycles, but it certainly allows a wide variation in rhythmic patterns, which make use of syncopations, or stresses of the upbeat, changes in note values etc. Basically, western classical music uses constant rhythm variation to express different emotions through the entire piece.

'Shruti' / Microtones

Indian classical music makes extensive use of quarter-tones and microtones, usually referred to as '*Shruti*'. Western classical music has a few microtonal pioneers in recent times, but has largely been restricted to using semitones.

Musical transitions between notes

In Indian classical music, wide variety of transitions is possible between the notes. Like Western Music, the jump note transition is also used in Indian music. These transitions basically act as the surprise elements in the music pieces. But, the unique feature of Indian Classical Music is the abundance of continuous transitions (*Meend*), which is conventionally responsible for strong evocation of emotions like sorrow, serenity etc. Few other varieties of transitions like *Andolan, Alangkar* etc. are also specifically found in Indian Classical Music.

With this formal introduction being set, let us gloss over the research present in the statistical distribution, specifically MB-BE, to analyze complex domains like music and linguistics.

5.5. EXISTING LITERATURE ON MB-BE DISTRIBUTIONS

In recent years, statistical distributions found a new horizon in the field of linguistics with the development of thermodynamic variables like Energy, Temperature and Specific heat to categorize and quantify complex system using their information content. The concept of 'temperature' was given long back by Mandelbröt (1953), which was reiterated and extended in subsequent works by a number of authors to be used in measuring communicative ability, comparing vocabulary complexity levels, assessing readership suitability or evaluating author's writing performances (De Campos & Tolman, 1982; Kosmidis et al., 2006; Miyazima & Yamamoto, 2008; Rego et al., 2014; Chang et al., 2017). The basis of this concept is the assumption that human language can be described as a physical system within the framework of equilibrium statistical mechanics. In Mandelbrot's interpretation, the text's informational temperature (θ) is reciprocal to a state variable B, as in, $\theta = 1/B$ and whose value, barring some very rare occasions, is always < 1. Nearer the text temperature to 1, greater the 'wealth of the vocabulary'. On the other hand, low temperatures indicate 'badly employed' words. Vocabulary of James Joyce and vocabulary of children are the examples of said two types, respectively. Kosmidis et al. connected the Zipf exponent to the temperature and suggested that this parameter can be used to measure the text's communicative ability. Miyazima-Yamamoto and later Rego et al. associated words with energies based on a general standard Maxwell–Boltzmann (MB) distribution. It is found that, the linguistic relative temperature of a book can be determined by measuring the deviation from a standard Maxwell–Boltzmann distribution of a corpus of English words. This relative temperature parameter can be used to measure vocabulary complexity level (Miyazima & Yamamoto, 2008) or author's writing capacity and repertory of thoughts (Chang et al., 2017). In a different take, Rovenchak and Buk focused on the behaviour of low-frequency words and mapped word rank–frequency distributions onto the Bose-Einstein (BE) distribution within the grand-canonical approach (Rovenchak & Buk, 2011; Rovenchak, 2014). The respective physical analogues are the power of the excitation spectrum α and the temperature T. The analogue of the fugacity z is determined from the number of words occurring only once (so-called '*hapax legomena*'). It was seen that the calculated parameters have a correlation with the language structure (the level of analyticity). Lower parameter values correspond to higher analyticity, indicating lesser word inflection.

According to Zipf himself, music is nothing but a different form of language. And quite similar to the latter, music involves organisation of symbolic sequences which are very different perceptually but put together as a whole display regularity and meaningful identifiable patterns. The difference from language, at least on the surface, lies perhaps in the lack of visible semantics. Then again, it has been reported that music can, akin to language, determine physiological indices of semantic processing (Koelsch et al., 2004).

5.6. OVERVIEW OF THE WORK

The objective of this study is to develop a novel methodology of extracting information from acoustic (mainly music) signals on the basis of existing note sequences and their appearance patterns. Three renditions of the same Raga performed by three different Indian classical music maestros (Ustad Amir Khan, Ustad Rashid Khan, Pandit Bhimsen Joshi) were analysed for this study. The raga used in this study is Raga *Marwa*, an advanced raga that is mostly performed by experienced musicians because of its unconventional movement of notes. The part under study is from the *Bandish* part of the raga as it depicts the raga in its full glory. Music clips were of 60 seconds duration. Each of the notes and their existing durations were computed for the entirety of the music sample. This combination of notes and their durations formed the necessary structural units which was then used as the basis of statistical measurements. Considering this ensemble as a collection of particles of gas in a container, we study the related distributions and parameters emerging from such distributions. For Maxwell-Boltzmann (MB) distribution, we consider this note-duration units as identical but distinguishable and further compute their related energy spectrum. For Bose-Einstein (BE) distribution, we find the cumulative distribution by finding the rank frequency plots for each note-duration unit. Considering the units as identical and indistinguishable, we study the possible BE distribution characteristics. The findings of the study are quite interesting. They show that besides the potential of using note-duration combinations as units of a statistical ensemble, the emerging ‘temperature’ parameters from MB and BE analysis could also be used as indicators of style identification, genre categorization and improvisation indicator. Overall, this analysis does exhibit potentials to be furthered into a full-fledged classificatory algorithm on the basis of used notes and their occurrence patterns.

5.7. EXPERIMENTAL DETAILS

5.7.1: Choice of Music Samples

The experimental classical music samples we chose consisted of the same raga performed by three eminent Hindustani classical music maestros, namely - Ustad Amir khan, Ustad Rashid Khan and Pandit Bhimsen Joshi. The performances were done on one of the more popular ragas of the North Indian music canon – Raga *Marwa*. Marwa, an eponymous raga of *thaat* Marwa (one of the ten fundamental *thaat*s in HCM) is described to be of a quiet and contemplative nature, often associated with sunsets (Kaufmann, 1968). One distinguishing feature of this raga is that it is harmonically ill-defined, that is, unlike the traditional notion in Indian classical music of *Sa*, *Ma* and *Pa* being the resting notes providing a sense of closure to the melodic progression, Marwa rests on the note *Dha*, an otherwise unconventional melodic structure. Additionally, the fifth note (*Pa*) is omitted and the Vadi-Samvadi pairs are *komal Re* (Re1) and

Shuddha Dha (Dha2) which doesn't form a perfect interval (fifth note relationship), hence making the harmonic imbalance more prominent. Besides the above, *Sa* is also a prominent note in this raga. Marwa is usually performed in the middle octave. With the help of classical music experts, the experimental clip was taken from the *Vilambit Bandish* part of the raga, which is the literature equivalent of the vocal performance and describes the raga in a methodical and scholarly manner. In addition to keeping the raga structure intact, it provides the improvisations made by individual artists. Samples were of 60 seconds duration for each of the three renditions.

5.7.2: Processing of Music Signals

All the audio signals were digitized at a sampling rate of 44.1 kHz 16-bit mono channel format using Cool Edit software of Syntrellium Corporation (Johnston, 1999) and amplitudes of all the clips were normalized to 0 dB. These sound signals constitute our database for this experiment. Pitch extraction technique is defined in next section. For this, Wavesurfer software of KTH, Stockholm was used (Sjölander & Beskow, 2000).

5.7.3: Pitch Extraction and Preparation of Note Profiles

The sound signals which were analyzed in this study have all been normalized to 0 dB, and hence intensity or loudness is not being considered here. Normalization helps removing any DC offset (which causes distortion) from the track by centering the waveform on the 0 dB amplitude level (maximum amplitude level). This is done by averaging all the sample values in the selection, and then subtracting the average value from the exact value of all the samples. The extraction of notes from the obtained pitch profile of each of the acoustic signal was done with the help of an experienced musician who has profound knowledge of the tonic. A skilled musician was asked to listen to the signal files one after another to detect the position of tonic 'Sa' in the signal file. The notes were extracted from the signal file following the methodology of Datta et.al (2006). The ratio-intervals were evaluated by first dividing the smoothed pitch values for each song by the pitch value of the 'Sa' tonic of that *raga* rendition. This gives the frequency ratios for each pitch data. From the ratio data, steady state sequences were created with all consecutive pitches in a sequence, which is terminated when $| x_{i+1} – M | > M/30$, where $M = (1/i) \Sigma x_i$ where x_{i+1} is the $(i+1)^{th}$ pitch and x_i is the i^{th} pitch. If the duration of any sequence were found to be less than that of a certain minimum value then the sequence is rejected. Elements extracted from these sequences were taken as suitable candidate data for this analysis. Whenever the ratio was found to be less than 1, it is multiplied by a factor of 2 and when it is greater than 2 it was divided by 2. This exercise effectively folds all pitch data into the middle octave. All the extracted pitch values were distributed in 1200 bins of one-cent width each. The peaks of these distributions for each musical signals are expected to be indicative of the note positions for that *raga* rendition. After detecting the value of 'Sa', we have calculated the pitch value of other notes (Datta et.al, 2006), thus each note was found out along with their time duration.

5.8. GENERAL METHODOLOGY

The broader rationale behind using MB and BE distributions is detailed in Chapter 2. Hence, in this chapter, we concentrate our efforts in highlighting the parts specific to this work only.

5.8.1: Maxwell-Boltzmann (MB) Distribution

To analyse the ragas, first, we need to accumulate a 'musical corpus' of used notes in the raga renditions, similar to a literary corpus (in our case it is the compilation of notes and their respective durations used in three samples under observation) (Miyazima & Yamamoto, 2008). From the pitch profile of the music sample, an experienced classical musician determines the frequency of 'Sa'. Frequencies of the rest of the notes are found out with their respective frequency ratios with 'Sa'. Afterwards, the existence and duration of occurrence of all the notes is indicated by analysing the pitch profile of the music clip using Wavesurfer software (the process is described under 'experimental details' section). The window is taken to be 10 milliseconds each. This way we find the number of occurrences of each note and their respective durations. For example: suppose Usa_{50}, indicating the occurrence of the upper octave 'Sa' for 50 milliseconds, has occurred 10 times during a piece. Similarly, Lre_{30} (lower octave 'Re', 30 millisecond duration) appears 24 times, ga_{40} (for middle octave) for 61 times etc. Next, the probabilities of the occurrence of note-duration combination will be plotted along with their respective 'energies'. Now, according to statistical mechanics, when a system of particles is in equilibrium at constant temperature T, then it can be found in one of N states permissible. The probability p_i that it is found at a given state i with energy E_i is given in Eq. (7):

$$p_i \sim 1.\exp(-\beta E_i) \tag{7}$$

Here, $\beta = 1.(kT)^{-1}$; k is the Boltzmann constant (= 1.38×10^{-23} J/K) and T is absolute temperature, the 'measure' of the interaction of the system with the environment. Now, computing the 'energy' values of note-duration combinations by putting k = T = 1, we arrive at Fig. 5.2.

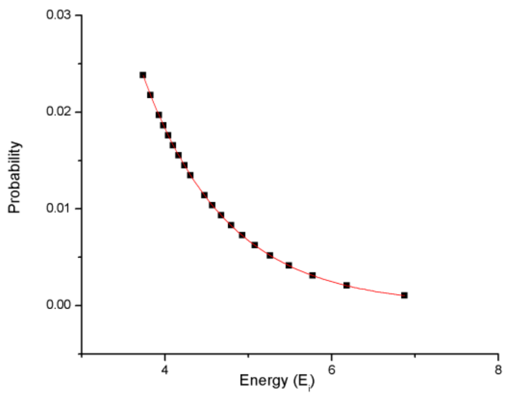


Fig. 5.2. Distribution of Probability (of occurrences of note-duration combinations) vs. 'Energy'. This constitutes the p vs E plot for the working corpus made up of every note-duration combination computed in all three renditions of Raga Marwa.

As it's clear from Fig. 5.2 that the probability vs. energy graph is perfectly exponential (as expected) and follows MB distribution pattern. The temperature of the corpus is assumed to be 1K, for comparison purposes later. The model equation we use to curve fitting is:

$$p(E) = y0 + A1.\exp(-E/t1) \tag{8}$$

The term t1 in eq. (8) is one of our desired parameters which will emerge from the curve fitting. This is Maxwell-Boltzmann distribution and the parameter is termed as Maxwell-Boltzmann temperature or T_{MB}. It has a unit of Kelvin (K). The same procedure will be followed for each of the three raga renditions individually and T_{MB} is derived in each of the cases to compare with the present 'music corpus'.

5.8.2: Bose-Einstein (BE) Distribution

5.8.2.1: Preparing Rank-Frequency distribution

To apply BE distribution, the first step is to prepare a rank-frequency distribution of the note-duration combinations. The combination having highest number of occurrences is given rank 1; the second most frequent is given rank 2, and so on. Components with the same frequency are given a consecutive range of ranks, the ordering within which can be arbitrary.

The rank-frequency distribution of a music sample is illustrated in Fig. 5.3. Horizontal plateaus in the figure having high ranks and low frequencies correspond to a large number of components having the same frequency. The longest plateau corresponds to frequency 1, i.e., elements occurring only once in the sample.

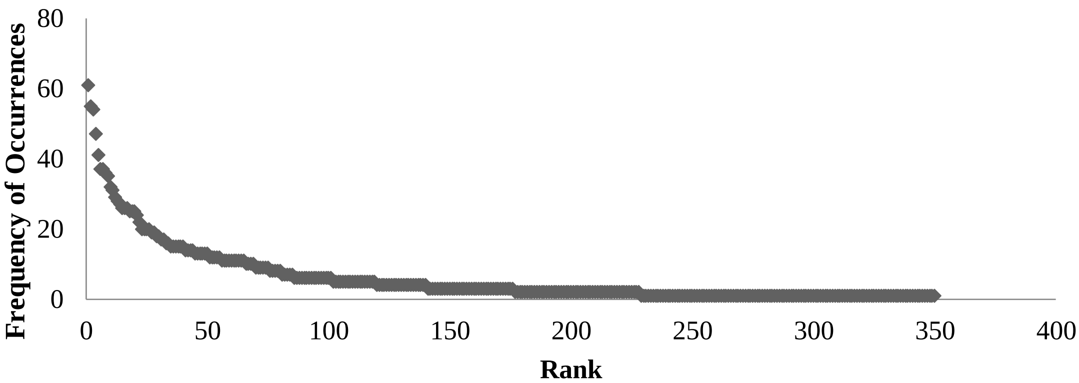


Fig. 5.3. An illustrative example of rank-frequency distribution of a music sample. It is computed from the note-duration combinations present in a Bandish of Raga *Shree* of two-minute duration.

5.8.2.2: Physical analogy of frequency structure and BE distribution

Following the treatment of Rovenchak and Buk (2011), we invert the rank-frequency distribution in a relation between number of occupants N_j vs their absolute frequencies j. We identify the energy level j with the number of occurrences of note-duration combinations. Hence, the components occurring once is situated in energy level j = 1, twice occuring components sit in energy level j = 2 etc. Each of the energy levels can have any number of occupants (N_j) without any restrictions. This idea aligns in accordance to the BE distribution, where each energy level can be occupied by any number of particles, without restricting laws like Pauli's exclusion principle. Such a plot of N_j vs j follows the BE distribution pattern. For such a distribution, the relation between occupancy number of j^{th} energy level N_j and j is:

$$N_j = \frac{1}{z^{-1}.\exp(\varepsilon_j/\mathrm{T_{BE}}) - 1} \tag{9}$$

Here, z is the fugacity, ε_j is the energy of the j^{th} level and T_{BE} is the temperature. The spectrum of ε_j is given by:

$$\varepsilon_j = (j - 1)^{\alpha} \tag{10}$$

Here, α is the exponent of the power spectrum. Unity is subtracted to make sure that the lowermost energy state, j = 1, has zero energy. The main focus of the study is on the lower frequency data (j = 2 ~ 10) since the energy spectrum relationship can look different for higher energies, i.e., states that have higher occurrence.

5.8.2.3: Parameters to be determined

First, z is calculated from the lowest N_j value, i.e., the occupancy of the lowermost occurrent state using the eq. (9), putting j = 1:

$$N_1 = \frac{1}{(z^{-1}-1)} = \frac{z}{1-z} \tag{11}$$

Also, exponent α of eq. (10) is to be determined by fitting the plot to eq. (9), along with T_{BE}, the Bose-Einstein 'temperature', another one of the desired parameters. It is important to mention that unlike T_{MB}, this is unitless as T_{BE} is not restricted by a reference corpus temperature.

5.9. RESULT AND DISCUSSIONS

5.9.1: Plots regarding MB Distribution

The probability of occurrence of the note-duration combinations for each performance is plotted against the respective energies as it is done in Fig. 5.2 and then, fitting a curve to eq. (8), we obtain the temperature T_{MB} for each of the renditions. The respective probability-energy plots are given in Fig. 5.4.

The red line indicates best fit for the data. The curve fitting data is discussed later in Table 5.1. Now we shall plot the ln (p_i) vs E_i plots and compute the slopes in each case, along with the corpus, to compare the renditions according to the 'comparative thermodynamic analysis' (Rego et al., 2014). The related plots are given in **Fig. 5.5 (a-d)**. As it is evident from figure 5.2, our working corpus of Marwa is a perfect fit. This is to be expected as we started the journey taking T = 1K for the corpus. The three Raga renditions, however, exhibit different slopes. Ustad Amir Khan's rendition display the highest slope (-0.93154). Ustad Rashid Khan's rendition remains a little behind (-0.89691) closely followed by Pandit Bhimsen Joshi's one (-0.89089). While dealing with word energy distributions of translated texts, Rego et al. (2014) argues that the reason slope value of the same text changes when translated to another language is the change in rank of the same words when it is translated, affecting the word energy and subsequently its linguistic temperature. In our case, the slope values of different renditions change although three performers performed the same raga, using same melodic structure and notations. The reason, we believe, is a result of the conscious choices made by the performer to use (or not use) a particular note or vary its duration. As it is discussed earlier, Indian classical music allows the performer enough space to improvise and adapt. Hence, an individual's improvisational characteristics remain responsible for the choices of structural units he/she takes. Thus, the same or similar note-duration combinations keep changing their ranks with different renditions of the same raga. This is an important feature to note when attempting artist classification based on the available information content. Although such changes could occur even when the same performer is performing the same raga on another occasion. Individual improvisation is also an important factor to consider, which we will be discussing in a later chapter.

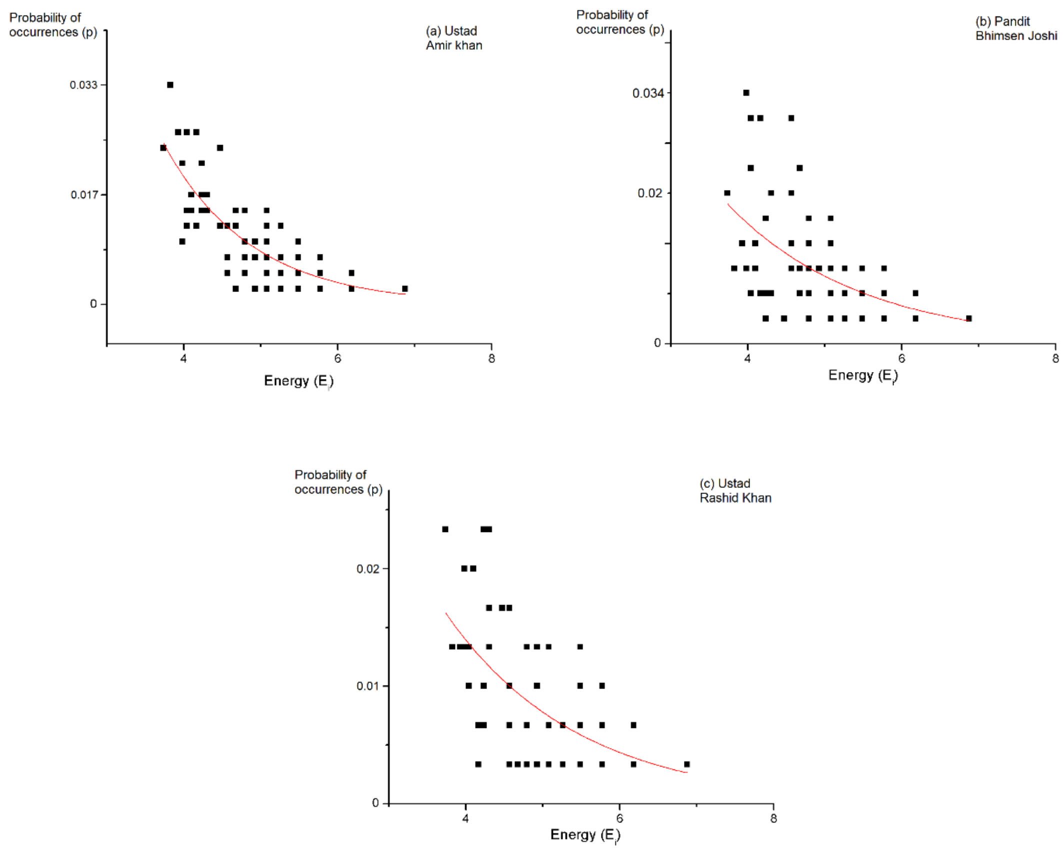


Fig. 5.4 (a-c). Occurrence probability vs Energy plots of Raga Marwa for a) Ustad Amir Khan, b) Pandit Bhimsen Joshi and c) Ustad Rashid Khan. Occurrence probability of each note-duration combination is computed with respect to all such combinations present in the corpus and the energies are calculated via eq. (7).

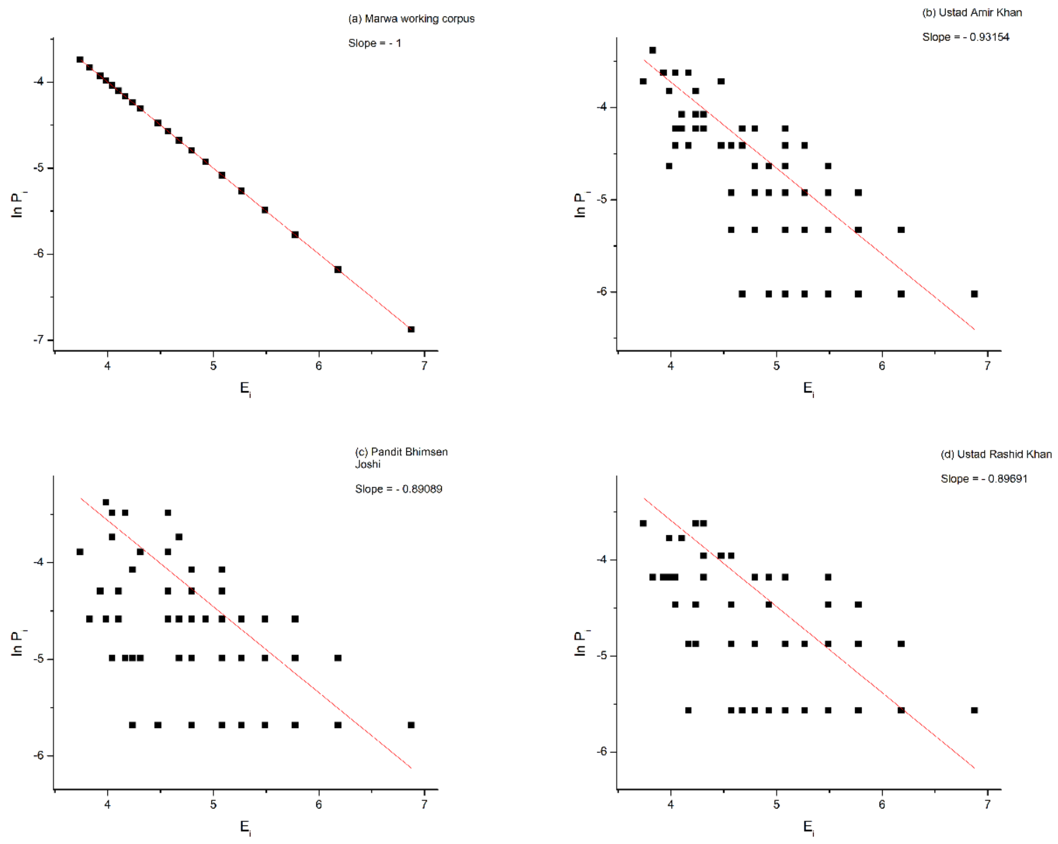


Fig. 5.5 (a-d). ln (p_i) vs E_i plots for a) Working corpus, b) Ustad Amir Khan, c) Pandit Bhimsen Joshi and d) Ustad Rashid Khan along with respective slope values, computed by taking logarithm of both sides of eq. (7). Different slopes indicate different temperatures (Rego et al., 2014).

5.9.2: Plots regarding BE Distribution

First, the rank-frequency distribution plots of the three performers are given in **Fig. 5.6 (a-c)**. It is evident from Fig. 5.3 and the plots that their nature follows the Zipf distribution, although the data points are fewer in numbers, a growth in which could significantly alter the visual representation. Similar to Fig. 5.3, the horizontal plateaus represent elements that have same occurrence frequency, with the highest number being 1.

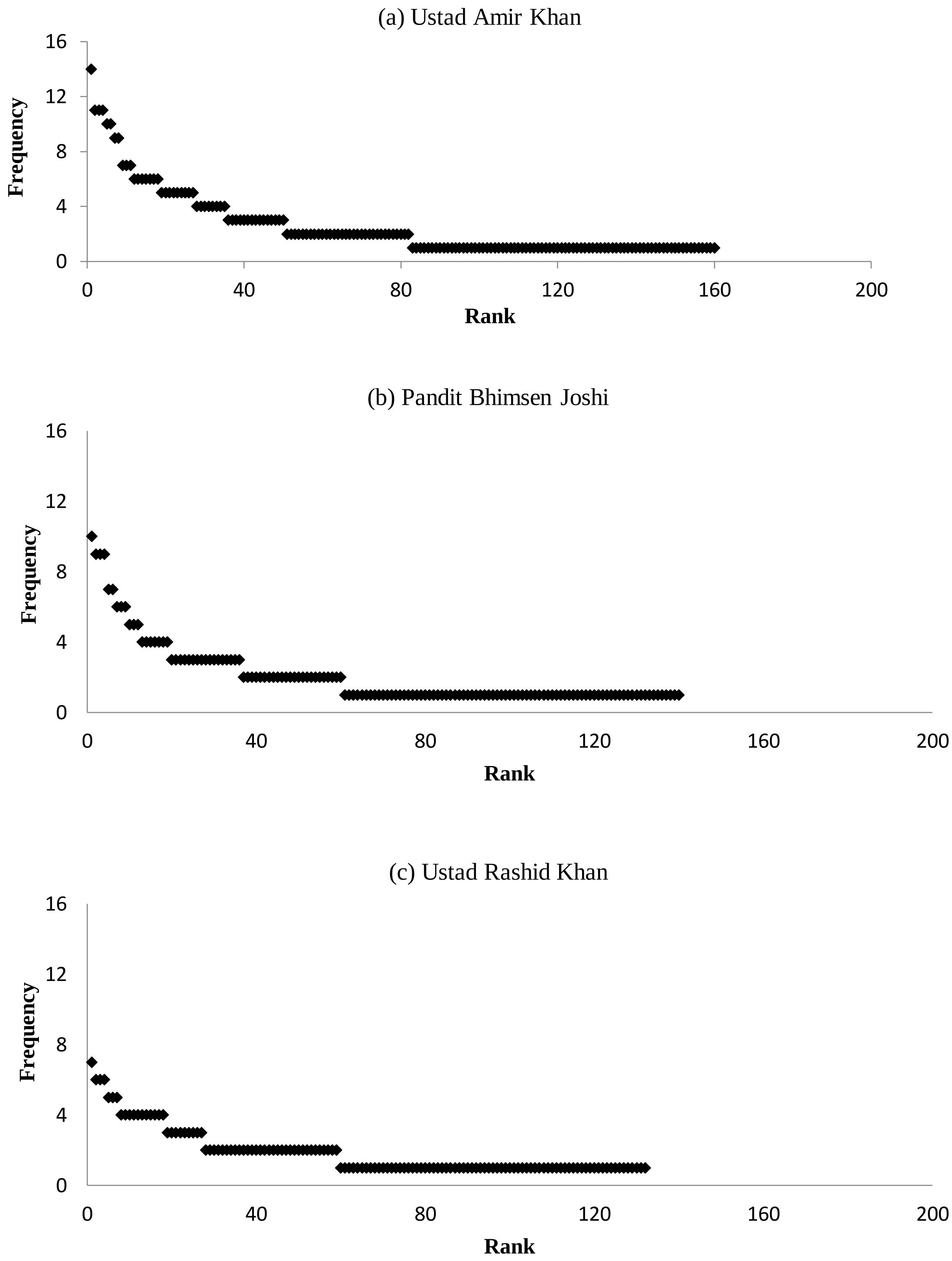


Fig. 5.6 (a-c). The rank-frequency distributions of note-duration combinations for a) Ustad Amir Khan, b) Pt. Bhimsen Joshi and c) Ustad Rashid Khan renditions.

Now, we rearrange the rank frequency data into energy level j and their occupant number N_j. In the next step, we will use this data into the fitting equations (8)-(11). Figures 5.4 and 5.6

resemble respective distribution patterns. The MB distribution is more visually evident as the number of data points illustrating the Zipf and subsequently BE distribution are relatively smaller and grouped.

5.9.3: Results from Curve Fitting

In the following step, we fit the data into eq. (8)-(11) and determine T_{MB}, z, α and T_{BE}. Results of the fitting are given in **Table 5.1**.

Music sample	N (total number of note-duration combination)	T_{MB} (K)	N_1 (total occupancy in lowermost state)	z	α	T_{BE}	τ = ln T_{BE}/ln N
Pandit Bhimsen Joshi	293	1.74 ± 0.18	80	0.9876	1.1825 ± 0.1301	23.74 ± 5.057	0.5576
Ustad Amir Khan	412	1.13 ± 0.06	78	0.9873	1.3404 ± 0.1180	48.82 ± 11.58	0.6458
Ustad Rashid Khan	261	1.72 ± 0.16	73	0.9865	2.1870 ± 0.3611	189.61 ± 29.27	0.9426

Table 5.1. Using equations (8)-(11), parameters $\mathbf{T_{MB}}$, **z**, **α** and $\mathbf{T_{BE}}$ are obtained. An additional parameter $\boldsymbol{\tau}$ = ln T_{BE}/ln N is also calculated for classification and comparison purposes.

Table 5.1 denotes the essence of the experiment as it bears the most important parameters and thereby, several important implications as well. From the table, it can be seen that among the three renditions, it was Ustad Amir Khan's one that has the highest number of note-duration combinations present in it, followed by Pandit Bhimsen Joshi. Ustad Rashid Khan has used the least number of elements in his performance. But interestingly, it was Pt. Bhimsen Joshi who has the highest N_1 value, that is, the number of once occurring elements.

One of the important parameters that we seek to find is T_{MB}, the Maxwell-Boltzmann 'temperature'. From the fitting data, it is seen that the Marwa rendition of Ustad Amir Khan has the lowest temperature value (1.13 K) whereas Pandit Joshi and Ustad Rashid's renditions have relatively higher temps, 1.74 K and 1.72 K respectively. As one can recall, the reference temperature of the working corpus was 1K. Hence, it is Ustad Amir's version of the raga that lies closest to the corpus temp. In linguistic research, such measure of temperature has been used to explore multiple avenues like vocabulary complexity level, communicative ability, evaluating author's writing capacity etc. But one thing has been generally agreed upon that swaying away from the reference temperature, usually that of a corpus, indicates deviating from the property under study. Echoing the same note, it can be said that the deviation of T_{MB} from its corpus temperature would stipulate that the performance is diverging away from the traditional singing practices of the said raga. That is, the usage of note and their durations does not closely follow the established musical grammar. Being said, one must remember how ICM handles improvisation flexibly and this deviation does not necessarily indicate the performer is ‘wrong’ in his approach. Here, it is noteworthy to mention the measure of the slope of ln p_i vs E_i plots mentioned earlier. It appears that the pattern of highest to lowest slope value is mirrored with the lowest to highest T_{MB} values. One explanation that could justify it is the fact that Amir Khan’s usage of notes are in greater agreement with the conventional Marwa singing style than the other two versions. To verify this notion, the note usage patterns of three singers are compared in **Fig. 5.7**. As it was predicted, along with the highest number of note-duration combinations used, Ustad Amir Khan’s rendition also has the highest usage of middle octave *Sa* and *komal Re* (Re1), and also *shuddha Dha* (Dha2), closely adhering to the grammar.

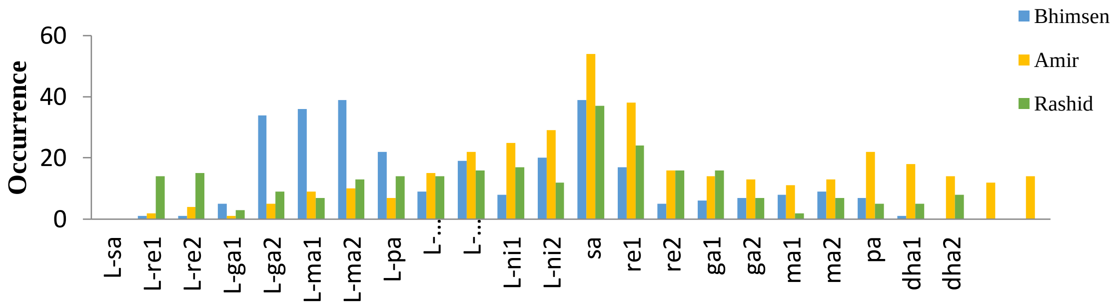


Fig. 5.7. Distribution of the occurrence frequency of notes for the three renditions of Marwa. ‘L-’ indicates the lower octave. ‘1’ and ‘2’ after the notes indicate *komal* and *shuddha* versions of the note, respectively (except for ‘*ma*’, where 1 and 2 denotes *shuddha* and *tivra*)

Turning to the other cardinal parameter, the Bose-Einstein ‘temperature’ or T_{BE}, it was Ustad Rashid Khan whose performance registered the highest value: 189.61 (± 29.27). 23.74 (± 5.057) was the lowest T_{BE} value that was observed in case of Pt. Bhimsen Joshi. Now, in linguistic framework, T_{BE} is associated with analyticity (Rovenchak & Buk, 2011). Analytic languages are those who employ strict syntactic rules and order of words to establish relationship between them. They are averse to using inflections (changing the form of definite version of a word to convey its role, for instance adding suffix or prefix). Example of a highly analytic language is Chinese. Bose-Einstein temperature has an inverse relation with analyticity - lower the former, higher the latter. Music, by nature, is devoid of inflections, so to speak. As the notes and their movements are well defined. But in Indian classical music, the melodic structure has some interesting feature (*alankar*s) that add expression and emotion to the body of a raga. These ‘alankar’s usually involve two or more notes in such a way that is reminiscent of the inflective properties of language. Some examples include: *Kan-swar* (while singing a note straight, grace notes or ‘small chunks’ of an adjacent note is borrowed), *Meend* (essentially, a form of gliding from one note to the other, a pattern of note transition that’s unique to ICM), *Gamak* (a technique of singing a note so forcefully that it creates a vibration) etc. Although not an exact replica of analyticity, but such rich and diverse classes of note usage is a distinguishing feature of ICM. Coming back to the relationship between T_{BE} and analyticity, it can thus be said that there is a scope of associating this variety and usage of ‘*alankar*’s with temperature-like parameter. This idea is facilitated once we look towards the role of temperature plays in BE distribution. Low T_{BE} implies the minimum energy variations among the particles of the system. At the lowest value of the temperature, which is zero ($T_{BE} = 0$), all atoms coalesce together to form a single quantum-mechanical entity and behave like a single atom, forming the famed Bose-Einstein condensate (BEC). Analogous to it, at lower values of T_{BE}, the diversity in note usage and existence of alankars are heavily minimised – a form of 'musical analyticity'. Hence, among the three performers of our study, Pandit Joshi's inflective assortment seemed to be the lowest whereas Ustad Rashid Khan displayed this trait heavily.

The parameter z is equivalent to fugacity. Fugacity characterizes the escaping tendency from a phase (Mackay & Paterson, 1981). Ideally, $z \approx 1$ (which corresponds to the BEC) and as mentioned by Rovenchak & Buk (2011), its value should always be closer to unity. Here, too, we see z value is sufficiently close to 1 in each of the performances.

The parameter α is the exponent of a power excitation spectrum and for lower j's, $1 < \alpha < 2$ is expected (Rovenchak, 2014) which implies good fitting of the data. In this case, barring Ustad Rashid Khan's rendition which records it at a slightly higher value (2.19 ± 0.36), the values of α remained within limits suggesting the data had not been complete misfit.

One additional parameter τ (= $\ln T_{BE}/\ln N$) has been computed as it has shown insignificant growth in case of sample size increase, something that makes it useful in comparative linguistic studies. Different planes like α-τ and N-τ has been used for classification purposes. From the data obtained in this study, the α-τ plane is illustrated in **Fig. 5.8**. Low values of α and τ are associated with higher analyticity, which means renditions in those area includes fewer range of inflexions, i.e., musical ornamentations.

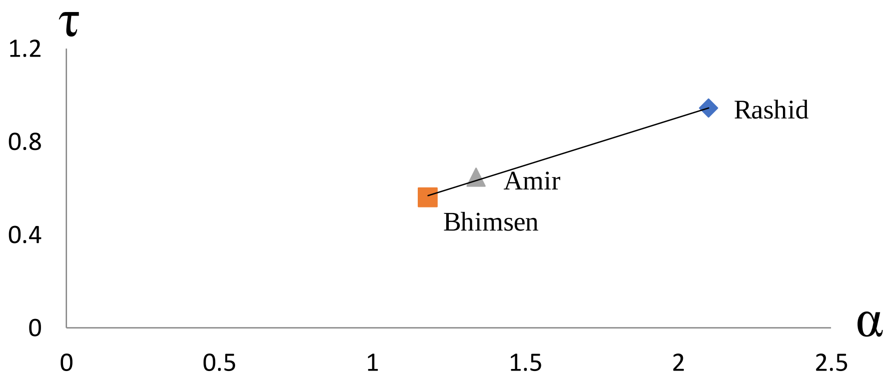


Fig. 5.8. The position of the raga renditions on the α-τ plane. It is seen that low α-low τ region hosts high 'musical analyticity'.

5.9.4: Measurement of Goodness-of-fit

Goodness-of-fit was tested using the determination co-efficient R^2 which seems most suitable (Rovenchak, 2015; Macutek & Wimmer, 2013) for the data. R^2 is given as eq. (12):

$$R^2 = 1 - \frac{\sum_i (f_i - NP_i)^2}{\sum_i (f_i - \bar{f})^2} \quad (12)$$

Here, f_i = the observed frequency of the value i, P_i = the theoretical probability of the value i, N = the sample size (the total number of the observations) and $\bar{f}$ is the mean of the observed frequencies ($\bar{f} = \sum_i^n f_i/n$). In case of BE data, the summation runs from j = 2 (since j = 1 is fixed by z), for the rest it runs from j = 1.

The R^2 value are given below in **Table 5.2**.

Raga rendition	MB distributions	BE distributions
Ustad Amir Khan	0.89	0.68
Pt. Bhimsen Joshi	0.79	0.59
Ustad Rashid Khan	0.81	0.61

Table 5.2. R^2 values for the three Raga renditions of Ustad Amir Khan, Ustad Rashid Khan and Pt. Bhimsen Joshi. The MB distribution data fits acceptably but the BE distribution data is scattered.

Generally, the R^2 value ≥ 0.9 is considered satisfactory, although ≥ 0.8 is also acceptable (Macutek & Wimmer, 2013). The data fits MB distribution satisfactorily but in case of BE distribution it appears as scattered. The probable reason for this is the reduction of data points once the rank-frequency data is converted into j-N_j plots. However, we believe, increased sample size would increase the efficiency of the fit significantly.

5.10. CONCLUSIONS

The kaleidoscope of complex system has innumerable interesting facets: from stock markets to human language, from neuronal networks to compositions of music. Studying these systems and their dynamics has gradually become well appreciated, especially in the physics community as the tools of physics have been proven instrumental in gaining ground into the world of complexities, time and again. Physical methodologies have identified and systematically characterized properties that are intrinsic to these systems. Although the canon 'complex system' typically gives shelter to a broad category of natural phenomenon, but at the same time, they often display orderly patterns which can even be categorised as, rather hopefully, simple. This might point towards an underlying system, akin to 'simple' or 'master program' in computational terms, providing a common basis for the growth and subsequent evolution of complexity throughout nature. Physics and physical theories should lead such quest as unifying distant and seemingly unrelated fields is a primary objective of theirs. Models and ideas might not find a definite home from the beginning in as arduous a journey as this. Terminology can often lack physical sense and resort to analogy. However, occasional success in establishing bridges is rather a pleasant outcome that keeps the proverbial flames alive and paves the way for future explorations into unknown horizons.

Analysing Indian classical music and its structural complexity provides such an opportunity. As we mentioned earlier, this work had set out in search of two objectives. First, to investigate the possibility of using hardcore statistical physics models in the complex domain of music and second, to advance the analysis methods by quantifying and categorising on the basis of informational patterns present in the structure.

The main conclusions that could be summarised from this study are given below.

1. The general fitting of the probability vs energy (and its semi-log counterpart) graphs looks satisfactory. This is an indicator that even with a smaller corpus, significant outcomes could be obtained. But one thing we must admit, that this conclusion comes with a cautionary take. Corpus size and its diversity has been a topic of research in linguistics for a long time. Contrary to it, music, especially Indian classical music, suffers from lack of usable corpus in a regular manner (Srinivasamurthy et al., 2014). Most of the corpuses are not accessible to large part of researchers. Digitization rate and quality, too, suffers a similar fate. Hence, in this specific work, we took resort to a working corpus made up of samples that we decided to study. Needless to say, size of the corpus plays a crucial role in the accuracy of the model. Hence, it is imperative to extend the size, quality and diversity of the corpus in future endeavours.
2. The parameter $\mathbf{T_{MB}}$ demonstrates **closeness to the traditional singing pattern** and hence, has the potential to be used in the training phase as an **indicator of correctness** (also can be used in **style identification**). However, the malleable approach of Indian classical music in terms of grammatical structure suggests that using the word 'incorrect' in performing a raga is rather a reductionist concoction than reality. Hence, instead of the literal, usage of the word 'correctness' comes from a comparative point of view that a student of music might face as his teacher's (or previous scholar's) rendition remains his frame of reference.
3. The consistency in the z-value suggests that the analogy of BE distribution and high-structured sequential data congregation such as music is surprisingly significant. The parameter z, or fugacity, is used in thermodynamics as a measure of deviation from the ideal. Its resolution can be improved by using an analog of chemical potential μ.
4. Barring one occasion, consistency in α values indicate that the assumed energy spectrum fits well for low frequency data. For higher frequencies, spectrum of ε_j needs to be investigated. It also has an inverse correlation with the analyticity level of the rendition.
5. The parameter $\mathbf{T_{BE}}$ can be used to indicate **diversity in note variations**, that is, the level of musical ornamentation and inflective movements are present in the sample. This clearly has a potential to be used as an improvisational parameter that can, besides distinguishing features between different artists, also tell how an individual's creativity differs in each rendition of the same song he performs.
6. α-τ plane can be used to categorize and compare between different genres, different ragas of the same melodic origin, different artists of the same genre or different renditions of the same performance as they vary in information content and contextual liabilities. τ remains fairly unchanged under the growth of the sample size which makes it an ideal parameter in comparative studies. It is seen that the low α- low τ region includes performances with higher analytic quality and the more inflective ones reside in high α-high τ part of the plane.

To sum up, a novel attempt was made in this study to introduce a model of harvesting musical information using robust statistical tools like Maxwell-Boltzmann and Bose-Einstein distributions whose traditional microscopic horizon has been expanded to include the complex world of time-series analysis. The yielding parameters show considerable promise in categorizing and quantifying artists based on their usage of notes and related improvisational skills that remained largely non-quantifiable till date. Usage of such statistical methodologies as a classificatory algorithm in the music domain is unique. With larger and diverse range of corpus, further correlation between the parameters and finer categorization of musical information remains in the realm of possibility. Early results are indicative that this method could add one more weapon in the quiver of music research focusing on categorization and style identification on an applicative front, thus fulfilling the objectives we set out to achieve.

REFERENCES

Rosen, R. (2012). Anticipatory systems. In *Anticipatory systems* (pp. 313-370). Springer, New York, NY.

Grassberger, P. (1986). Toward a quantitative theory of self-generated complexity. *International Journal of Theoretical Physics*, *25*(9), 907-938.

Ladyman, J., Lambert, J., & Wiesner, K. (2013). What is a complex system?. *European Journal for Philosophy of Science*, *3*(1), 33-67.

Crutchfield, J. P., & Young, K. (1989). Inferring statistical complexity. *Physical review letters*, *63*(2), 105.

Feldman, D. P., & Crutchfield, J. P. (1998). Measures of statistical complexity: Why?. *Physics Letters A*, *238*(4-5), 244-252.

Shiner, J. S., Davison, M., & Landsberg, P. T. (1999). Simple measure for complexity. *Physical review E*, *59*(2), 1459.

Mitchell, M. (2009). *Complexity: A guided tour*. Oxford university press.

Zurek, W. H. (2018). *Complexity, entropy and the physics of information*. CRC Press.

Ebeling, W., Molgedey, L., Kurths, J., & Schwarz, U. (2002). Entropy, complexity, predictability, and data analysis of time series and letter sequences. In *The science of disasters* (pp. 2-25). Springer, Berlin, Heidelberg.

Lloyd, S., & Pagels, H. (1988). Complexity as thermodynamic depth. *Annals of physics*, *188*(1), 186-213.

Witt, A., Neiman, A., & Kurths, J. (1997). Characterizing the dynamics of stochastic bistable systems by measures of complexity. *Physical Review E*, *55*(5), 5050.

Boccaletti, S., Kurths, J., Osipov, G., Valladares, D. L., & Zhou, C. S. (2002). The synchronization of chaotic systems. *Physics reports*, *366*(1-2), 1-101.

Mantegna, R. N., & Stanley, H. E. (1999). *Introduction to econophysics: correlations and complexity in finance*. Cambridge university press.

Chaitin, G. J. (1966). On the length of programs for computing finite binary sequences. *Journal of the ACM (JACM)*, *13*(4), 547-569.

Schwarz, U., Benz, A. O., Kurths, J., & Witt, A. (1993). Analysis of solar spike events by means of symbolic dynamics methods. *Astronomy and Astrophysics*, *277*, 215.

Witt, A., Kurths, J., Krause, F., & Fischer, K. (1994). On the validity of a model for the reversals of the Earth's magnetic field. *Geophysical & Astrophysical Fluid Dynamics*, *77*(1-4), 79-91.

Kurths, J., Voss, A., Saparin, P., Witt, A., Kleiner, H. J., & Wessel, N. (1995). Quantitative analysis of heart rate variability. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *5*(1), 88-94.

Saparin, P. I., Gowin, W., Kurths, J., & Felsenberg, D. (1998). Quantification of cancellous bone structure using symbolic dynamics and measures of complexity. *Physical Review E*, *58*(5), 6449.

Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., & Kurths, J. (2002). Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. *Physical review E*, *66*(2), 026702.

Zou, Y., Donner, R. V., Marwan, N., Donges, J. F., & Kurths, J. (2019). Complex network approaches to nonlinear time series analysis. *Physics Reports*, *787*, 1-97.

Ebeling, W., & Nicolis, G. (1991). Entropy of symbolic sequences: the role of correlations. *EPL (Europhysics Letters)*, *14*(3), 191.

Ebeling, W., & Nicolis, G. (1992). Word frequency and entropy of symbolic sequences: a dynamical perspective. *Chaos, Solitons & Fractals*, *2*(6), 635-650.

Ebeling, W., & Pöschel, T. (1994). Entropy and long-range correlations in literary English. *EPL (Europhysics Letters)*, *26*(4), 241.

Peng, C. K., Buldyrev, S. V., Goldberger, A. L., Havlin, S., Sciortino, F., Simons, M., & Stanley, H. E. (1992). Long-range correlations in nucleotide sequences. *Nature*, *356*(6365), 168-170.

Schenkel, A., Zhang, J., & Zhang, Y. C. (1993). Long range correlation in human writings. *Fractals*, *1*(01), 47-57.

Jafari, G. R., Pedram, P., & Hedayatifar, L. (2007). Long-range correlation and multifractality in Bach's inventions pitches. *Journal of Statistical Mechanics: Theory and Experiment*, *2007*(04), P04012.

Martin-Montoya, L. A., Aranda-Camacho, N. M., & Quimbay, C. J. (2015). Long-range correlations and trends in Colombian seismic time series. *Physica A: Statistical Mechanics and its Applications*, *421*, 124-133.

Costa, R. L., & Vasconcelos, G. L. (2003). Long-range correlations and nonstationarity in the Brazilian stock market. *Physica A: Statistical Mechanics and its Applications*, *329*(1-2), 231-248.

Neukum, G., & Ivanov, B. A. (1994). Crater size distributions and impact probabilities on earth from lunar, terrestrial planeta, and asteroid cratering data. *Hazards due to Comets and Asteroids*, 359-416.

Willis, J. C., & Yule, G. U. (1922). Some statistics of evolution and geographical distribution in plants and animals, and their significance. *Nature*, *109*(2728), 177-179.

Price de Solla, D. J. (1965). Networks of scientific papers. *Science*, *149*(3683), 510-515.

Faloutsos, M., Faloutsos, P., & Faloutsos, C. (1999). On power-law relationships of the internet topology. *ACM SIGCOMM computer communication review*, *29*(4), 251-262.

Foote, J. (1999, October). Visualizing music and audio using self-similarity. In *Proceedings of the seventh ACM international conference on Multimedia (Part 1)* (pp. 77-80).

Pareyon, G. (2011). *On Musical Self-Similarity: Intersemiosis as Synecdoche and Analogy*. Gabriel Pareyon.

Zipf, G. K. (1949). Human behaviour and the principle of least-effort. Cambridge MA edn. *Reading: Addison-Wesley*

Li, W. (2002). Zipf's Law everywhere. *Glottometrics, 5*(2002), 14-21.

Voss, R. F., & Clarke, J. (1975). '1/f noise' in music and speech. *Nature*, *258*(5533), 317-318.

Voss, R. F., & Clarke, J. (1978). ''1/f noise''in music: Music from 1/f noise. *The Journal of the Acoustical Society of America*, *63*(1), 258-263.

Hsü, K. J., & Hsü, A. (1991). Self-similarity of the" 1/f noise" called music. *Proceedings of the National Academy of Sciences*, *88*(8), 3507-3509.

Zanette, D. H. (2006). Zipf's law and the creation of musical context. *Musicae Scientiae*, *10*(1), 3-18.

Manaris, B., Purewal, T., & McCormick, C. (2002, April). Progress towards recognizing and classifying beautiful music with computers-MIDI-encoded music and the Zipf-Mandelbrot law. In *Proceedings IEEE SoutheastCon 2002 (Cat. No. 02CH37283)* (pp. 52-57). IEEE.

Manaris, B., Romero, J., Machado, P., Krehbiel, D., Hirzel, T., Pharr, W., & Davis, R. B. (2005). Zipf's law, music classification, and aesthetics. *Computer Music Journal*, *29*(1), 55-69.

Liu, L., Wei, J., Zhang, H., Xin, J., & Huang, J. (2013). A statistical physics view of pitch fluctuations in the classical music from Bach to Chopin: Evidence for scaling. *PLoS One*, *8*(3), e58710.

del Río, M. B., Cocho, G., & Naumis, G. G. (2008). Universality in the tail of musical note rank distribution. *Physica A: Statistical Mechanics and its Applications*, *387*(22), 5552-5560.

Perotti, J. I., & Billoni, O. V. (2020). On the emergence of Zipf's law in music. *Physica A: Statistical Mechanics and its Applications*, *549*, 124309.

Rangarajan, L. N. (Ed.). (1992). *The arthashastra*. Penguin Books India.

Mandelbrot, B. (1953). An informational theory of the statistical structure of language. *Communication theory*, *84*, 486-502.

De Campos, H., & Tolman, J. M. (1982). The informational temperature of the text. *Poetics Today*, *3*(3), 177-187.

Kosmidis, K., Kalampokis, A., & Argyrakis, P. (2006). Statistical mechanical approach to human language. *Physica A: Statistical Mechanics and its Applications*, *366*, 495-502.

Miyazima, S., & Yamamoto, K. (2008). Measuring the temperature of texts. *Fractals*, *16*(01), 25-32.

Reˆgo, H. H. A., Braunstein, L. A., D′ Agostino, G., Stanley, H. E., & Miyazima, S. (2014). When a text is translated does the complexity of its vocabulary change? Translations and target readerships. *PloS one*, *9*(10), e110213.

Chang, M. C., Yang, A. C. C., Stanley, H. E., & Peng, C. K. (2017). Measuring information-based energy and temperature of literary texts. *Physica A: Statistical Mechanics and its Applications*, *468*, 783-789.

Rovenchak, A., & Buk, S. (2011). Application of a quantum ensemble model to linguistic analysis. *Physica A: Statistical Mechanics and its Applications*, *390*(7), 1326-1331.

Rovenchak, A. (2014). Trends in language evolution found from the frequency structure of texts mapped against the Bose-distribution. *Journal of Quantitative Linguistics*, *21*(3), 281-294.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: brain signatures of semantic processing. *Nature neuroscience*, *7*(3), 302-307.

Kaufmann, W. (1968). *The ragas of north India* (No. 1). Indiana University Press.

Johnston, D. (1999). Cool edit 2000. *Syntrillium Software Corporation, Phoenix, Arizona, EEUU*.

Sjölander, K., & Beskow, J. (2000). Wavesurfer-an open source speech tool. In *Sixth International Conference on Spoken Language Processing*.

Datta A.K., Sengupta, R., Dey, N., Nag, D. and Mukerjee A (9-10 January, 2006), *A methodology of note extraction from the song signals.* Proc. Frontiers of Research on Speech and Music (FRSM-2006), U P Technical University, Lucknow.

Mackay, D., & Paterson, S. (1981). Calculating fugacity. *Environmental Science & Technology*, *15*(9), 1006-1014.

Rovenchak, A. (2015, May). Models of frequency spectrum in texts based on quantum distributions in fractional space dimensions. In *2015 20th International Conference on Control Systems and Computer Science* (pp. 645-649). IEEE.

Mačutek, J., & Wimmer, G. (2013). Evaluating goodness-of-fit of discrete distribution models in quantitative linguistics. *Journal of Quantitative Linguistics*, *20*(3), 227-240.

Srinivasamurthy, A., Koduri, G. K., Gulati, S., Ishwar, V., & Serra, X. (2014). Corpora for music information research in indian art music. In *Georgaki A, Kouroupetroglou G, eds. Proceedings of the 2014 International Computer Music Conference, ICMC/SMC; 2014 Sept 14-20; Athens, Greece.[Michigan]: Michigan Publishing; 2014.*. Michigan Publishing.

CHAPTER 6

EXTENDING THE APPLICATION OF STATISTICAL METHODS IN INDIAN CLASSICAL MUSIC: A STUDY ON RAGA CATEGORIZATION USING MAXWELL-BOLTZMANN AND BOSE-EINSTEIN DISTRIBUTIONS

"Bottomless wonders spring from simple rules, which are repeated without end."

Benoit Mandelbrot

ABSTRACT

Raga characterization in Indian classical music is an important aspect of music learning in this country. But the methods usually followed by the learners' perspective are mostly qualitative. In this study, we intend to quantify such abstractness using measurable parameters. To study musical information, we introduce methods based on well-known concepts used in Statistical Physics, namely Maxwell-Boltzmann (MB) and Bose-Einstein (BE) distribution. In this present study, these distributions have been applied on the chosen acoustic signals to find new parameters (equivalent to 'temperature' in physical systems) which can distinguish between various features of different ragas (containing similar notes) in Indian classical music. Music clips chosen were the 'Alap' part of these three different ragas (Marwa, Puriya, Sohini) sung by Pandit Ajay Chakrabarty, a legendary classical music maestro. All of the chosen three ragas are based on the following note structure: Sa, *komal* Re, *shuddha* Ga, *tivra* Ma, *shuddha* Dha, *shuddha* Ni. To apply MB statistics to music, it is assumed that different notes with different occurrence frequencies are at different energy levels, the distribution of which follows the MB distribution pattern. In case of BE statistics, a rank-frequency distribution of the time durations of various notes of different ragas is studied. The resulting analysis gives rise to a number of parameters that help to categorize the individual characteristics of ragas like kinetic nature, complexity of note usage pattern, 'analyticity' level. The methods studied here are novel in the music research field and can prove to be useful in the fields of music and speech as quantifying parameters in other categorization problems.

Keywords: Maxwell-Boltzmann distribution, Bose-Einstein distribution, Indian classical music, Raga characterization, temperature

6.1. INTRODUCTION: BRIEF NOTE ON *RAGA* IN INDIAN CLASSICAL MUSIC

Indian classical music (ICM) is one of the most creative art forms existing in this world and Raga is the heart of it. Raga, essentially, is a well-defined melodic structure that consists of a series of four/five (or more) musical notes upon which its melody is constructed. However, it is the approach of the performer towards handling the notes and his skills in rendering them in musical phrases and also the mood they convey are more important in defining a Raga than the notes themselves. So, abundant scope of improvisation is allowed within the structured framework. It means, every performer of this genre is effectively composing as well as performing. Performers of ICM visualize every Raga as a living existence. The word Raga finds its roots in the Sanskrit word "*Ranj*" which means 'to please' as well as 'to color' (Holroyde, 1972). Although there are a number of definitions attributed to a Raga, it is basically a tonal multifarious module. The listener needs to listen to a number of pieces of the Raga in order to recognize it. In Hindustani music, the goal of a performer, mainly, is to convey the musical structure and expression in such a way that it evokes pleasantness for the audience. For every Raga, the tonic 'Sa' is the most important note as each Raga can be identified uniquely from appropriate establishment of 'Sa'. The notes used in raga (and ICM in general) and their respective symbols are given in **Table 6.1**.

Name of the note	Short pronunciation	Symbol used
Sharaj	'Sa'	Sa
Komal Rishabh	'Re'	Re1
Shuddha Rishabh	'Re'	Re2
Komal Gandhar	'Ga'	Ga1
Shuddha Gandhar	'Ga'	Ga2
Shuddha Madhyam	'Ma'	Ma1
Tivra Madhyam	'Ma'	Ma2
Pancham	'Pa'	Pa
Komal Dhaivat	'Dha'	Dha1
Shuddha Dhaivat	'Dha'	Dha2
Komal Nishad	'Ni'	Ni1
Shuddha Nishad	'Ni'	Ni2

Table 6.1. Notes (Swaras) used in ICM and their symbols and representation

Some fundamental elements that should be traditionally present in a Raga –

- *Vad-Samvad* – Vad and Samvad are the harmonic pillars of Raga's musical structure. Vad-Samvad notes are the fifth note relationships essential to establish a Raga. At least any one of the following pairs: 'Sa (middle octave) – Pa (middle octave)' or 'Ma (middle octave) –

Sa (higher octave)' must be present. Apart from that one or two pair(s) of Vad-Samvad notes may occur in a particular Raga.

- *Nyas* – Steady notes that are sung at the end of a phrase for a longer duration than the rest of the notes and acts as the resting point. Basically, the Vad-Samvad note pairs act as the *Nyas swaras* or resting points in almost all cases.

- Transitions between notes – Different kinds of transitions between the notes are possible in a single Raga. Some examples include: *Meend, Andolan, Alangkar* (*Gamak, Murki, Khatka* and many more), *Sparsh swara* or touch notes, Jump notes etc. Depending on the dominance of a particular variety of transition, overall mood of the Raga can change as well. Ragas with more jump notes and faster movements intend to evoke happier emotion among listeners whereas Ragas having more *Meend* or *Andolan* type transitions are usually responsible for evoking emotions like sorrow and serenity.

- Other conditions include presence of atleast 5 (and maximum of 7) notes or *swara*s, presence of ascending/descending patterns (*arohana/avarohana*), presence of characteristic melodic phrases (raga specific) etc. Few examples of ragas with their characteristic ascending and descending notes are given in **Table 6.2**.

As discussed in the previous chapter, Indian Classical Music is mainly divided into two genres according to singing style – Hindustani Music (Originated in the northern parts of India) and Carnatic Music (popular in southern parts of India). The vocal presentation of a Raga in Hindustani music style can also be divided into two categories – *Khayal* and *Dhrupad*. *Khayal* singing usually starts with a short *Alap* or *Aochar* while in case of *Dhrupad* performances, *alap* usually sits for a much longer duration than the former. The *Alap* is the opening section (or, the preface) of a typical Hindustani Music (HM) performance. In the *alap* part, the raga is introduced and the paths of its development are revealed briefly using all the notes used in that particular raga and allowed transitions between them with proper distribution over time. *Alap* is usually accompanied by the tanpura drone only and performed at a slow tempo or sometimes even without tempo. Next, comes the *vilambit bandish* (in case of vocal music) where the lyrics and *taal* are introduced. *Bandish* is a song i.e., a fixed, melodic composition in Hindustani vocal or instrumental music, set in a specific raga, performed with rhythmic accompaniment by a tabla or pakhawaj, a steady drone, and melodic accompaniment by a sarangi, harmonium etc. (Neuman, 1990). *Vilambit* ('late' literally), is a type of *bandish* which is sung at a very slow tempo, or *laya*, of 10-40 beats per minute. Then the *madhyalaya* (mid-tempo) or *drut bandish* (high-tempo) are sung usually at a much higher bpm. Along with the *bandishes* the performers sing *vistara* (melodic and rhythmic variations or improvisations of the lyrical contents of the *bandish*, which help in conveying the meaning of the lyrics more elaborately to the listeners) and *taan* (melodic and rhythmic improvisations of the phrases of a raga, mostly sung without lyrics). *Dhrupads* are simpler forms of *Khayal* having higher emphasis on *Meends* and lesser emphasis on *Alangkars*. In case of instrumental representation of a Raga, after the completion of *alap, Gat, Jor* and *Jhala* are played sequentially. In HM, the existing phrases are stretched or compressed, and the same may happen to motives from the phrases; further motives may be prefixed, infixed and suffixed. Phrases may be broken up or slid in with others, and motives or phrases may be sequenced through different registers (Neuman, 1990). Thus, during a performance, a singer steadily breathes out of the chokehold of the grammatical rules in a subtle way. He does not disregard them, instead merely interprets them in a new way staying within the framework – this is a unique characteristic of Hindustani classical music and it represents the wisdom that Raga and its grammar are only means and not ends in themselves. The way a performer interprets the same raga differently, but in a subtle manner, during each of his performances, is certainly a distinctive feature and it constitutes the very essence of improvisation in Hindustani music. Raga is unpredictable, unlike symphony or a concerto. It is ever-blooming and blossoms out into new and vivid forms during each and every performance – that's the crux of improvisation (McNeil, 2007).

Ragas	**Constituent notes**	**Ascending notes (*Arohana*)**	**Descending notes (*Avarohana*)**
Bhairav	Sa Re1 Ga2 Ma1 Pa Dha1 Ni1 U-Sa	Sa Ga2 Ma1 Pa Ga2 Ma1 Dha1 U-sa	U-Sa Ni1 U-Sa Dha1 Pa Ga2 Ma1 Pa Ga2 Ma1 Re1 Sa
Bhoopali	Sa Re2 Ga2 Pa Dha2 U-Sa	Sa Re2 Ga2 Pa Dha2 U-Sa	U-Sa Dha2 Pa Ga2 Re2 Sa
Desh	Sa Re2 Ga2 Ma1 Pa Dha2 Ni1 U-Sa	Sa Re2 Ma1 Pa Ni2 U-Sa	U-Sa Ni1 Dha2 Pa Ma1 Ga2 Re2 Ga2 L-Ni2 Sa
Durga	Sa Re2 Ma1 Pa Dha2 U-Sa	Sa Re2 Ma1 Pa Dha2 U-Sa	U-Sa L-Dha2 Ma1 Pa Dha2 Ma1 L-Re2 L-Dha2 Sa

Table 6.2. Few examples of used notes and their ascending/descending patterns in Ragas; prefixes U- and L- represents said notes in upper and lower octaves respectively (Katte, 2013)

6.2. RAGAS AND EMOTIONS

In the Indian subcontinent, music has been a constant source of aesthetic gratification since the ancient times. One of the early proponents discussing music's correlation with emotion was Bharata's Natyashastra, somewhere between 200 BCE to 200 CE (Ghosh, 2002). With time, a number of treatises spoke in favor of the various *rasas* (emotional experiences) that are conveyed by the different forms of musical performances. In fact, the objective of fine arts such as music is to evoke and consolidate *rasa* in the mind of the viewer/listener. This is the central concept in Indian artistic practice since Natyashastra put forward the doctrine of rasa with its eight categories for the first time. These categories are: Love or Romance (*Śṛngāra*), Gaiety or Humor (*Hāsya*), Compassion (*Kāruṇya*), Fury (*Raudra*), Valor (*Veera*), Terrible (*bhayankara*), Loathsomeness (*bibhatsa*), and Wonder (*adbhuta*). From the third or fourth century onwards Silence or Tranquillity (*shanta*) was added as the ninth category and was considered as a crucial part of the rasa-system (Mukherjee, 1965). It is believed that the concept of *"Rasa"* is the most important and significant contribution of the Indian mind to the field of aesthetics. The study of aesthetics usually dealt with the realization of beauty in art and the awareness of joy that accompanies an experience of beauty. Science, however, had largely avoided discussion on the aesthetic experiences corresponding to a particular performance and it was kept at bay as a fringe study.

It is only from the last two decades of the 20th century that scientists began to understand the huge potential of systematic research that Indian Classical Music (ICM) has to offer in the advancement of cognitive science as well as psychological research. A number of works tried to harvest this immense potential by studying objectively the emotional experiences attributed to the different *ragas* of Indian Classical Music (Balkwill & Thompson, 1999; Chordia & Rae, 2007; Wieczorkowska et.al, 2010; Mathur et.al, 2015). The studies have revealed unlike Western Music the emotions evoked by Indian Classical Music are often more ambiguous and far more subdued. Earlier few musicologists believed that a particular emotion can be assigned to a particular *Raga* but recent studies (Wieczorkowska et.al, 2010) clearly revealed that different phrases of a particular *Raga* is capable of evoking emotions among the listeners.

Since different sets of ragas form the basis and backbone of Indian Classical Music, it is only natural to try to seek how do they invoke emotions and how closely their acoustic framework and structural features contribute to said induction.

6.3. EXISTING WORKS IN RAGA IDENTIFICATION AND CLASSIFICATION BASED ON STRUCTURAL CUES

The existing works in Raga classification and identification are based in the field of computational musicology and Music Information Retrieval (MIR). We shall gloss over them briefly as detailed discussion is out of the scope of this chapter. However, should one intend to divulge deeper, Katte (2013), Waghmare and Sonkamble (2017), Murthy & Koolagudi (2018) provide some deeper dives into this subject. Usually, the already developed features in speech processing are primary operators in dealing with music related inputs. But with time, some music-specific features are added to the cannon as well. They are generally classified into three categories – low, mid and high-level features (Fu et al., 2010). Low-level features are extracted from very small audio segments (< 100 ms). In case of music, timbre and temporal features are classified as such. Timbre is that property of an acoustic signal which differentiates between two types of instruments (or a vocal and an instrument) even when their pitch and energy, mathematically, are the same. Timbre feature techniques available in speech processing that are also used in music related works include mel-frequency cepstral coefficients (MFCCs), zero crossing rate (ZCR), root mean square (RMS) energy, spectral centroid (SC), spectral flux (SF) etc. Temporal features deal with the temporal changes of properties in a signal. Statistical parameters (mean, variance, co-variance), computed from a large number of tiny windows taken from the signal, are the features. MuVar, different multi-variate regressions, Hidden Markov Models (HMM) are some of the techniques used in this regard (Reed & Lee, 2009). Mid-level features are extracted at note level and bear identities of the entire signal. They are divided into three categories – rhythm, harmony and pitch. Rhythmic features help provide information on the repetitive tension patterns present in the signal, like beat. They are also indicators of musical emotion as rhythm is often associated with mood (Yang et al., 2008). Harmony is the combination of multiple individual notes played/sung together, also known as chords. Hence, chord detection and chord sequence factor in harmony features. They are used in detecting fundamental frequencies present in a chord (Bello, 2007). The third and the most used feature specially in case of Raga related studies, is pitch, since their detection and subsequent ratios singles out the notes and their occurrences. Pitch features like pitch histograms (PHs), pitch class profile (PCP) helps in detecting the melodic structure of the music, and also, genre and mood (Krumhansl, 2001). Other than the two, semantic information (for the retrieval of the clip) falls under high-level features. Generally, combinations of mid and low-level features are used in harvesting music-related information. Once the features retrieve information from the sample, they undergo a supervised or unsupervised classification model like support vector machines (SVM), artificial neural network (ANN) or k-nearest neighbour (KNN).

Raga identification and categorization is a subject of great research interest among the MIR community. The reason is its rich applicative potential across fields of music learning, music appreciation, similarity measures and genre classification. In majority of studies, identifying fundamental frequency F0 and its respective ratios to other tonic frequency are applied. Additional features like Pitch Class Distributions (PCD) and Pitch Class Dyad Distribution (PCDD) are also used (Chordia & Rae, 2007; Belle et al., 2009). Belle, Joshi and Rao (2009) converted polyphonic signal into mono and manually detected the tonic Sa. Then, extracted the pitch features and calculated the folded pitch distribution, and subsequently, the PCD. This produced features like peak, mean, standard deviation and probability of pitch corresponding to each note. Later, classifier algorithm (nearest neighbour) was used. In a related work, authors used First-order pitch distributions and template matching technique (Koduri et al., 2012). Probability Density Function based of pitch contours has also been attempted in a study where a feature set of 36-dimensions (12x3) is extracted from the PDF to classify ragas in Carnatic music (Suma & Koolagudi, 2015). In another interesting study, Bhattacharjee and Srinivasan (2011) used Transition Probability Matrix (TPM) of 12x12 dimension to predict ascending-descending sequences (Aaroha and Avaroha) of notes and compared it to the literature. They used manual pitch detection. Based on single step transition Markov model, TPM showed high accuracy in representing and classification of raga. Other studies included pitch related features and classifiers like random forest, SVM and ANN (Shetty & Achary, 2009; Dighe et al., 2013; Kumar et al., 2014).

6.4. OUR APPROACH BASED ON INFORMATION RELATED ENERGY

Like many natural systems, music is also a complex phenomenon formed by a large information congregation restricted by some boundary conditions that gives it its characteristic structure and specific patterns of temporal evolution. The characteristics present in any of the *Raga* structure throw a vast amount of information towards the listener and analysts alike. When looked at as a whole, this appears too complex. But, similar to every other information humanly perceivable, it also is made up of repetitive pattern or sequences of some basic elements. Fortunately, the basis of approaching the dynamics of such systems remained the same, across any field. That is, to effectively categorize the information content based on the common elements and/or their dynamic distribution. Human language, in written form, have been posing similar challenges to science. There have been many approaches used in tackling the categorization or classification issues. Occurrence of different words have been a classificatory feature on which Bayesian methods are applied (Mosteller & Wallace, 1963). To address authorship attribute problems another well debated method, ‘Delta’, has been proposed which is based on the differences in the frequencies of the most frequent words in a collection of texts (Burrows, 2002). Lexical richness, argued in some studies, is a useful feature to categorize author characteristics (Holmes, 1992; Tweedie & Baayen, 1998). Also, there have been models using statistical complexity quantifiers based on entropy in computational linguistics (Montemurro & Zanette, 2002; Rosso et al., 2009). But the one idea that endured the test of time and blossomed in new horizons is given in 1949 by George K. Zipf based on the rank of a word and its occurrence frequency (Zipf, 1949). Zipf's argument was that the speaker/writer prefers to minimize word usage, that is, few words for most meanings whereas the hearer/reader needs things maximized, that is, every meaning has a different word. The higher the degree of satisfaction of the requirements (of one of the two), the less its effort. The result of this is a brilliant empirical law, known as the 'Zipf's law'. G.K. Zipf proposed that Zipf's law emanates from this exchange of requirement between the speaker and the audience. The mathematical form of Zipf's law looks like the following: if we assign the rank j = 1 to the most frequent word of a language, j = 2 to the second one, etc., then the frequency of occurrence f(j) of a given word varies with its rank j as:

$$f(j) \sim 1.\ j^{-\alpha} \qquad (1)$$

Here, α is an exponent which is to be determined from the rank vs. frequency distribution.

Zipf's law is remarkable because it is robust. It applies to diverse systems in nature, including economics, urbanization and growing social networks, among others (Gabaix, 1999; Naldi, 2003; Zhang & Sornette, 2011). In recent years, a growing body of work has developed which is based on the assumption that human language can be analogous to a physical system (e.g., gas particles in container) within the boundaries of equilibrium statistical mechanics (Miyazima & Yamamoto, 2008; Chang et al., 2017). The theoretical fundamentals of the idea have been discussed in Chapter 2. We shall restrict ourselves into a brief overview of it for now. The central idea can be summarised as: different words with different occurrence frequency has different energies ('word energies') and their energy distribution follows Maxwell-Boltzmann (MB) distribution. It is connected to the Zipf's law via direct correlation of the Hamiltonian with the rank of the word (usefulness) (Kosmidis et al., 2006). Now, according to statistical mechanics, for a system of particles at constant temperature T, the probability p_i that it is found at a given state i with energy E_i is:

$$p_i \sim 1.\ \exp(-\beta E_i) \qquad (2)$$

Here, $\beta = 1.\ (kT)^{-1}$; k is the Boltzmann constant (= 1.38×10^{-23}J/K) and T is absolute temperature, which is regarded as the 'measure' of the interaction between system and environment. Taking k as 1, computing p_i values and fitting them into eq. (2) gives the value of T. This 'text temperature' has been used in diverse measures such as authorship disputes, changes in complexity of vocabulary etc.

Now, music and language are two attires sewn from the same piece of cloth. They share similarities in both cognitive and structural fronts. In fact, the differences in them only appear due to the fundamental difference in their building blocks (sound-meaning pairings for language, pitch-classes and pitch-class combinations for music), otherwise they are argued to be very similar in the 'Identity thesis' (Katz & Pesetsky, 2011). Indeed, the processing of both language and music involves creating expectations of what is to come. Fine-scale voluntary vocal production and the ability to imitate it are involved in the learning and producing of both. And most importantly, they share the characteristic of individual's ability to create and improvise, in order to be emotionally meaningful and also, distinct. This prompts us to suspect that the statistical models used in linguistic studies could also be of use in the categorization problems regarding music. The possibility of using power law relations in music related research has already been explored in western classical music. The studies cover music aesthetics (Manaris et al., 2005), creation of musical context (Zanette, 2006), statistical distributions of short-time timbral codings (Haro et al., 2012) and more. In one study, authors argued that the emergence of Zipf's law is a mere consequence of music being more efficient and complex form of language (Perotti & Billoni, 2020). In ICM, the *swara*s or notes are the fundamental structural units of the Raga framework. Hence, if a musical rendition was to be assumed a physical system, then the notes are the natural 'Zipfian' units. But unlike words, music has a temporal component which must be taken into account. The notes themselves are only as important as what duration they are performed or sung for. A note sung for 100 ms. is entirely different from it being sung for 200 ms, in context to the raga performance, its structure and its movements. The Zipfian units, in the case of music should have a temporal representation which will take into account the duration of the notes as well. Considering this factor, it was decided that the notes and their occurrence duration together would form the organisational units. The occurrence probability of such a combination would ascertain a corresponding 'energy' whose distribution should, theoretically, give us a characteristic temperature regarding that music sample.

Other than MB distribution which has been associated with high-frequency words, Rovenchak and Buk (2011) suggested a new set of parameters by comparing low-frequency words with a bosonic system within the grand canonical formulation (The rationale is explained in Chapter 2). The basis of the idea is reflected in the fact that the rank-frequency distribution of the words in text show similarity with Bose-Einstein (BE) distribution pattern. The result of this analysis includes a temperature like parameter which correlates with the analyticity level of a language. Recently, in unrelated studies, authors have explored the possibility of describing human language as a Boson gas and argued in favour of using 'words' as fundamental units carrying different energies. Interestingly, they too related the energy levels with Zipf rankings (Aerts & Beltran, 2020; Beltran, 2021). Low frequency data is of particular interest in context to ICM and Raga. In most cases, the Raga structure is asymmetric. That is, the number of notes used in the ascending or aarohana part of the raga are mostly fewer than the number of notes during descending or avarohana. For example, Raag Dhani uses the note *Shuddha* Re (Re2) while descending but skips it during ascending. This creates the asymmetry and it is one of the reasons of a raga's aesthetic character. Hence, the dynamics regarding the lower frequent notes are sometimes of prime importance in Indian classical renditions.

To sum up, our approach renders the music piece as a physical system with computable statistical parameters. It doesn't disregard the attributes of MIR discussed in the previous section as our bedrock is constructed upon notes/swaras, a concept that is derived from extracting a mid-level feature (pitch). Using it, we explore the dynamics of used notes (and the duration performer decides to spend on them) and attempt to characterize and categorise the Raga, along with its otherwise unquantifiable properties. Of course, a Raga is more than just the notes and their durations. But as the fundamental particles create the universe despite having different applicable rules at different scales, we hope that this approach could sow the seeds of a novel research horizon in ICM scenario. Before moving on to the study itself, it would be a good opportunity to briefly discuss the two distributions we are going to be deploying in subsequent sections.

6.5. MAXWELL-BOLTZMANN (MB) AND BOSE-EINSTEIN (BE) DISTRIBUTIONS IN BRIEF

6.5.1. Maxwell-Boltzmann distribution

The Maxwell-Boltzmann (MB) statistics is generally used for distribution of an amount of energy between identical but distinguishable particles. MB statistics predicts that the probability of finding a particle with a specific energy decreases exponentially with increasing energy, considering the system consists of a huge number of non-interacting distinguishable particles.

The distribution function has the following form:

$$f(E_i) = 1/Ae^{Ei/kT} \tag{3}$$

where $f(E_i)$ is the probability of a particle having energy E_i, A is the normalisation constant, E_i is the energy of the i-th state, k is the Boltzmann constant, and T is absolute temperature.

6.5.2. Bose-Einstein distribution

Bose-Einstein (BE) statistics describes the dynamics of an ensemble of identical and indistinguishable particles occupying discrete energy states. The distribution function indicating the energy distribution looks like:

$$f(E_i) = 1/[Ae^{Ei/kT} - 1] \tag{4}$$

where $f(E_i)$ is the probability of a particle having energy E_i, 1/A denotes the degeneracy, i.e., how many particles are having particular energy state E_i, E_i is the energy of the i-th state, k is the Boltzmann constant, and T is absolute temperature. The -1 factor in the denominator recognises the fact that the particles are indistinguishable, unlike MB statistics. BE distribution applies to a very particular kind of particles who have integer spin values, known as Bosons. They do not obey the Pauli's exclusion principle and hence unlimited number of particles can occupy the same energy state (The particles that obey the Pauli's exclusion principle are called Fermions). This unusual property of the BE distribution helps in applying this concept in various systems beyond the sub-atomic world.

Fig. 6.1 shows the MB and BE distributions on a population vs. energy graph.

6.6. MB-BE DISTRIBUTIONS IN LINGUISTIC ANALYSIS

The concept of text 'temperature' is a long-discussed subject. It was extended in a number of recent studies to investigate various linguistic measures: comparing vocabulary complexity levels, assessing readership suitability or evaluating author's writing performances (Kosmidis et al., 2006, Miyazima & Yamamoto, 2008; Rego et al., 2014, Chang et al., 2017). Kosmidis et al. connected the Zipf exponent to temperature and suggested that this parameter can be used to measure the text's communicative ability. Miyazima-Yamamoto and later Rego et al.

associated words with energies based on a general standard Maxwell–Boltzmann (MB) distribution. It is found that, the linguistic relative temperature of a book can be determined by measuring the deviation from a standard Maxwell–Boltzmann distribution of a corpus of English words. This relative temperature parameter can be used to measure vocabulary complexity level or author's writing capacity and repertory of thoughts. Rovenchak and Buk focused on the behaviour of low-frequency words and mapped word rank–frequency distributions onto the Bose-Einstein (BE) distribution within the grand-canonical approach (Rovenchak & Buk, 2011; Rovenchak, 2014). The respective physical analogues are the power of the excitation spectrum α and the temperature T. The analogue of the fugacity z is determined from the number of words occurring only once (so-called '*hapax legomena*'). It was seen that the calculated parameters have a correlation with the language structure (the level of analyticity). Lower parameter values correspond to higher analyticity, indicating lesser word inflection.

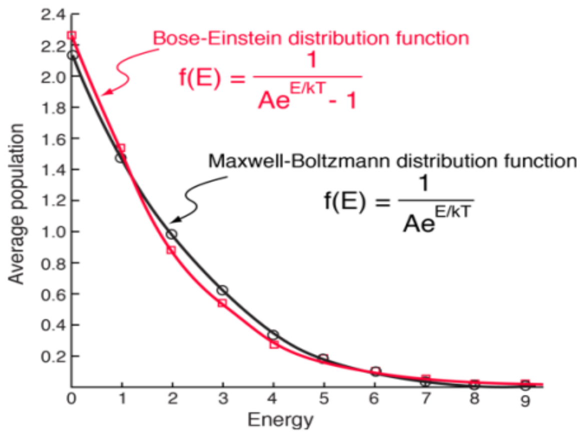


Fig. 6.2. Maxwell-Boltzmann and Bose-Einstein distributions, population vs energy plot. Mathematically, they are quite similar, except for the -1 factor in BE to factor in the indistinguishability of the particles [*Image courtesy*: http://hyperphysics.phy-astr.gsu.edu; *URL*: http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/imgqua/disbemb.png]

6.7. OVERVIEW OF THE WORK

Raga characterization in Indian classical music is an important aspect of music learning in this country. But from the learner's perspective, the methods followed are mostly qualitative. In this study, we intend to quantify such abstractness using measurable parameters. To study musical information congregation quantifiably, we introduce methods based on well-known concepts used in Statistical Physics, namely Maxwell-Boltzmann (MB) and Bose-Einstein (BE) distribution. The approach we follow in this study is based on the analogy between the rank-frequency distributions (using Zipf's law) of a note-duration combination of a music sample and the statistical distributions (both the MB distribution and the BE distribution in grand canonical formulation). To apply MB statistics to music, it is assumed that musical combination units (of notes and their durations) with different occurrence frequencies are at different energy levels, the distribution of which follows the MB distribution pattern. In case of BE statistics, a rank-frequency distribution of the time durations of various notes of different ragas is studied. The resulting analysis gives rise to a number of parameters that help to categorize the individual characteristics of ragas. One of which is the equivalent of Temperature in case of physical systems. This 'temperature' parameter is quite familiar in linguistic research and it has been used to specify the underlying dynamics of various languages, authorship disputes, changes in complexity of vocabulary and many more. When dealing with Indian classical music, such parameters help in highlighting different features of different ragas (containing the same notes). Music clips chosen were the 'Alap' part of three ragas (Marwa, Puriya, Sohini) sung by a legendary classical music maestro. All of the chosen ragas are based on the following same note structure: Sa, *komal* Re (Re1), *shuddh* Ga (Ga2), *tivra* Ma (Ma2), *shuddh* Dha (Dha2), *shuddh* Ni (Ni2). The methods described here are novel in the music research field and can prove to be useful in the study of music and speech as quantifiers for classification and categorization.

6.8. EXPERIMENTAL DETAILS

6.8.1: Choice of *Raga*s

Here, we have chosen to study three of the most popular and frequently performed ragas in North Indian classical music, namely: *Marwa*, *Puriya* and *Sohini* (all three performed by Pandit Ajay Chakrabarty, an Indian classical music maestro). The reason of choosing them is the interesting similarities that these three Ragas shares. They use the same set of notes: *Sa*, *Komal Re* (Re1), *Shuddha Ga* (Ga2), *Tivra Ma* (Ma2), *Shuddha Dha* (Dha2), *Shuddha Ni* (Ni2). All three of them belong to *Thaat* Marwa and are far too collegial in their appearance, but they are widely believed to evoke three distinctly different categories of emotions among listeners. Marwa's notional scale is said to be based on lower octave Dha (L-Dha2). Puriya, similarly, has a notional scale base in L-Ni2. Contrarily, the same for Sohini resides at *shuddha* Ga (Ga2) in the middle octave as it is performed at higher frequencies. Also, Sohini's treatment of its constituent notes is heavily ornamented and dominated by ascending phrases whereas Puriya's treatment is subtle and ascending/descending phrases are in tandem with each other. Marwa, however, employs a robust treatment with a general room for improvisations where descending phrases are highlighted. **Table 6.3** summarizes the acoustic features of the three Ragas.

Raga	Representative mood	*Vadi-Samvadi* pair	Prolonged/dominating note (*Nyas swara*)
Marwa	Quiet, contemplative	Re1 and Dha2	Sa, Re1
Puriya	Melancholic, sad	Ga2 and Ni2	Sa, Ni2
Sohini	Romantic, vivacious	Dha2 and Ga2	Sa

Table 6.3. Acoustic attributes of Ragas Marwa, Puriya and Sohini

The sample music clips were of 3-minute 30 sec duration from the *Alap* part of the raga, selected by a classical music expert. *Alap* provides the essence of the raga as it gradually exposes the characteristic phrases, *Vadi-Samvadi* pairs and other striking features of it.

6.8.2: Processing of Music Signals

All the audio signals were digitized at a sampling rate of 44.1 kHz 16-bit mono channel format using Cool Edit software of Syntrellium Corporation (Johnston, 1999) and amplitudes of all the clips were normalized to 0 dB. These sound signals constitute our database for this experiment. Pitch extraction technique is defined in next section. For this, Wavesurfer software of KTH, Stockholm was used (Sjölander & Beskow, 2000).

6.8.3: Pitch Extraction and Note Profiles

The sound signals which were analyzed in this study have all been normalized to 0 dB, and hence intensity or loudness is not being considered here. Normalization helps removing any DC offset (which causes distortion) from the track by centering the waveform on the 0 dB amplitude level (maximum amplitude level). This is done by averaging all the sample values in the selection, and then subtracting the average value from the exact value of all the samples. The extraction of notes from the obtained pitch profile of each of the acoustic signal was done with the help of an experienced musician who has profound knowledge of the tonic. A skilled musician was asked to listen to the signal files one after another to detect the position of tonic 'Sa' in the signal file. The notes were extracted from the signal file following the methodology of Datta et.al (2006). The ratio-intervals were evaluated by first dividing the smoothed pitch values for each song by the pitch value of the 'Sa' tonic of that *raga* rendition. This gives the frequency ratios for each pitch data. From the ratio data, steady state sequences were created with all consecutive pitches in a sequence, which is terminated when $| x_{i+1} – M | > M/30$, where $M = (1/i) \Sigma x_i$ where x_{i+1} is the $(i+1)^{th}$ pitch and x_i is the i^{th} pitch. If the duration of any sequence were found to be less than that of a certain minimum value then the sequence is rejected. Elements extracted from these sequences were taken as suitable candidate data for this analysis. Whenever the ratio was found to be less than 1, it is multiplied by a factor of 2 and when it is greater than 2 it was divided by 2. This exercise effectively folds all pitch data into the middle octave. All the extracted pitch values were distributed in 1200 bins of one-cent width each. The peaks of these distributions for each musical signals are expected to be indicative of the note positions for that *raga* rendition. After detecting the value of 'Sa', we have calculated the pitch value of other notes (Datta et.al, 2006), thus each note was found out along with their time duration.

6.9. GENERAL METHODOLOGY

The general methodological approach has been expounded in the Chapters 2 and 5. In search for relevancy as well as brevity, we shall only refer to the pertinent equations and corresponding parameters that are to be computed.

6.9.1: Maxwell-Boltzmann (MB) Distribution

Our working corpus, in this study, consists of all the available note-duration combinations from the three raga renditions considered. Firstly, the 'energies' of each note-duration combination is evaluated using eq. (2). Then, they are compiled and plotted against the probability of occurrence in the following four cases: the corpus and three ragas each, separately, with the help of model equation Eq. (5).

$$p(E) = y0 + A1.\ \exp(-E/t1) \tag{5}$$

Here, t1 denotes the desired Maxwell-Boltzmann temperature (T_{MB}) in Kelvins. The results from the curve fitting will also be used in comparison purposes.

6.9.2: Bose-Einstein (BE) Distribution

Here, our focus is on the low-frequency end of the data. The rank-frequency distribution plots of note-duration combinations are plotted. Followed by this, we invert the rank frequency data to find the corresponding N_j vs j plots. The relevant equation here is given by eq. (6):

$$N_j = \frac{1}{z^{-1}.\exp(\varepsilon_j/T_{BE}) - 1} \tag{6}$$

Here, N_j is occupancy number of j^{th} energy level and z is fugacity. The energy of the j^{th} level is ε_j and T_{BE} is the Bose-Einstein temperature. The spectrum of ε_j is given by eq. (7).

$$\varepsilon_j = (j - 1)^{\alpha} \tag{7}$$

Unity is subtracted to make sure that the lowermost energy state, j = 1, has zero energy. According to Rovenchak and Buk (2011), this spectrum is effective only for low-frequency data.

Value of z is found from eq. (6) using j = 1,

$$N_1 = \frac{z}{1-z} \tag{8}$$

The other parameter, T_{BE}, is determined from eq. (6) using curve fitting.

6.10. RESULT AND DISCUSSIONS

6.10.1. Plots regarding MB Distribution

The probability of occurrence of the note-duration combinations is plotted against the respective energies, for four cases: the working corpus, Raga Marwa, Raga Puriya, Raga Sohini. The plot for probability vs energy for the working corpus is given in fig. 6.2.

The plot is perfectly exponential and perfectly fits the data which is to be expected as we decided to take k = T = 1 in eq. (2). Of course, it follows the MB distribution. We assumed the corpus temperature to be 1K for comparison purposes, as this will serve as the reference temperature. Now, the individual p vs E plots for three of the ragas are given in fig. 6.3 (a-c).

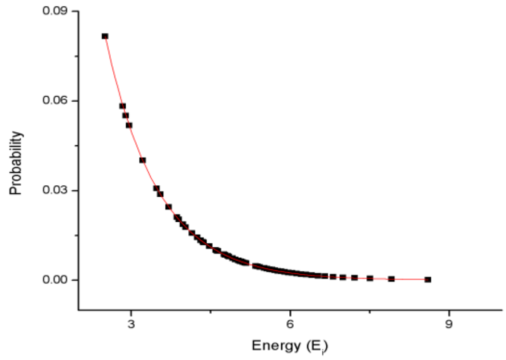


Fig. 6.2. Distribution of Probability (of occurrences of note-duration combinations) vs. 'Energy'. This constitutes the p vs E plot for the working corpus made up of every note-duration combination computed in all three ragas: *Marwa*, *Puriya* and *Sohini*.

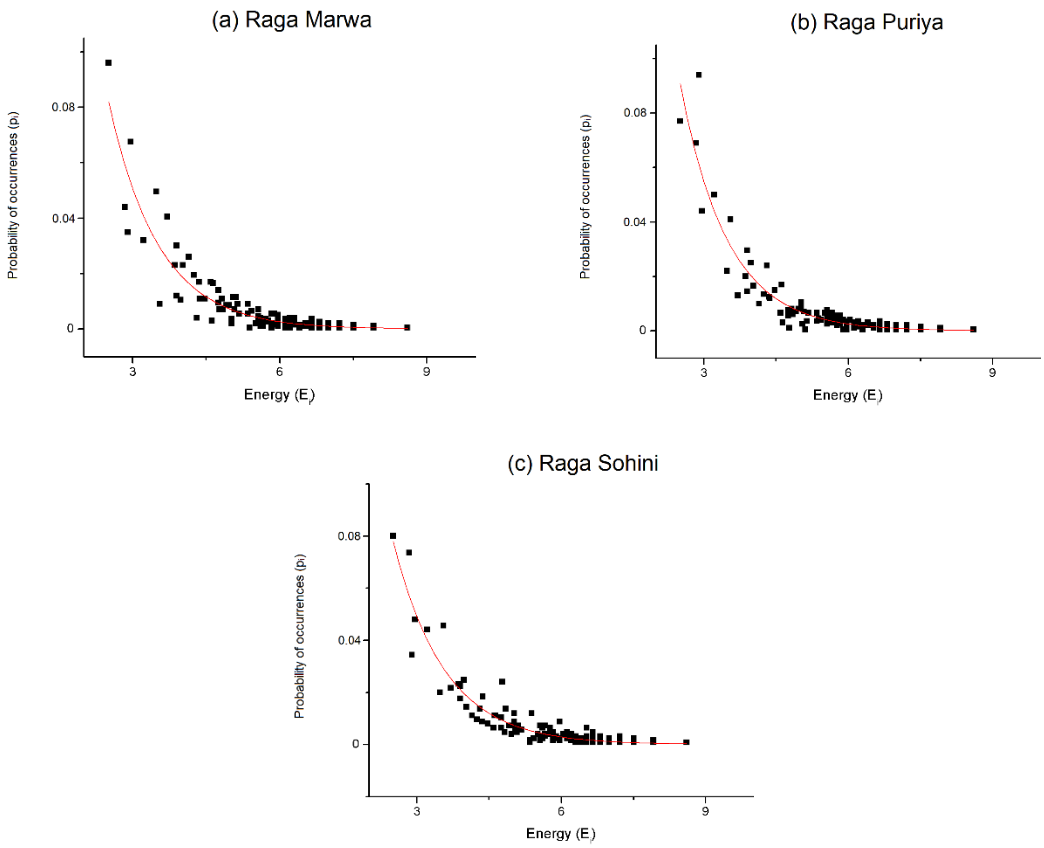


Fig. 6.3 (a-c). Occurrence probability vs Energy plots for a) Raga Marwa, b) Raga Puriya and c) Raga Sohini. Occurrence probability of each note-duration combination is computed with respect to all such combinations present in the corpus and the energies are calculated via eq. (2).

The red line corresponds to the best fit of the data. They are then fitted to eq. (5) which will provide Maxwell-Boltzmann temperature T_{MB}. The results of the curve fitting are given in Table 6.4. Rego et al (2014) converted these sets of plots into ln (p_i) vs E_i by taking logarithms on both sides of eq. (2) and compared the resulting slopes. According to them, such ‘comparative thermodynamic analysis’ reveals the comparative changes in data even when the corpus or corpora shows similar Zipf distribution characteristics. Following the same treatment here, the respective ln (p_i) vs E_i plots for each Raga along with the corpus is given in Fig. 6.4 (a-d). Here, we can see that the working corpus has a slope value of -1, as it should since it has a perfect fit (Fig. 6.2). Values of the slopes corresponding to Raga Marwa, puriya and Sohini, respectively, are: -0.94621, -0.95547 and -0.92181. The semi-log plot corresponding Raga Puriya has the highest slope whereas it is lowest in case of Raga Sohini. The change in slope generally stipulates change in temperature. As it was discussed in the previous chapter, the change in slope values usually indicate that the ranks of constituent note-duration combinations have altered from one raga to the next, hence altering the energy and the temperature associated with it. It definitively demonstrates that although the three ragas use the same set of notes, the importance of those notes changes course from one raga to the other. This, of course, sounds all too obvious intuitively, because these are distinct ragas in reality. But what this information entails is how subtle the change of context should be when performing or listening to these sets of ragas. Seeing the proximity of their note distributions with a corpus of their constituent notes, no wonder it is uncomfortable even for the most discerning listeners to differentiate between the ragas.

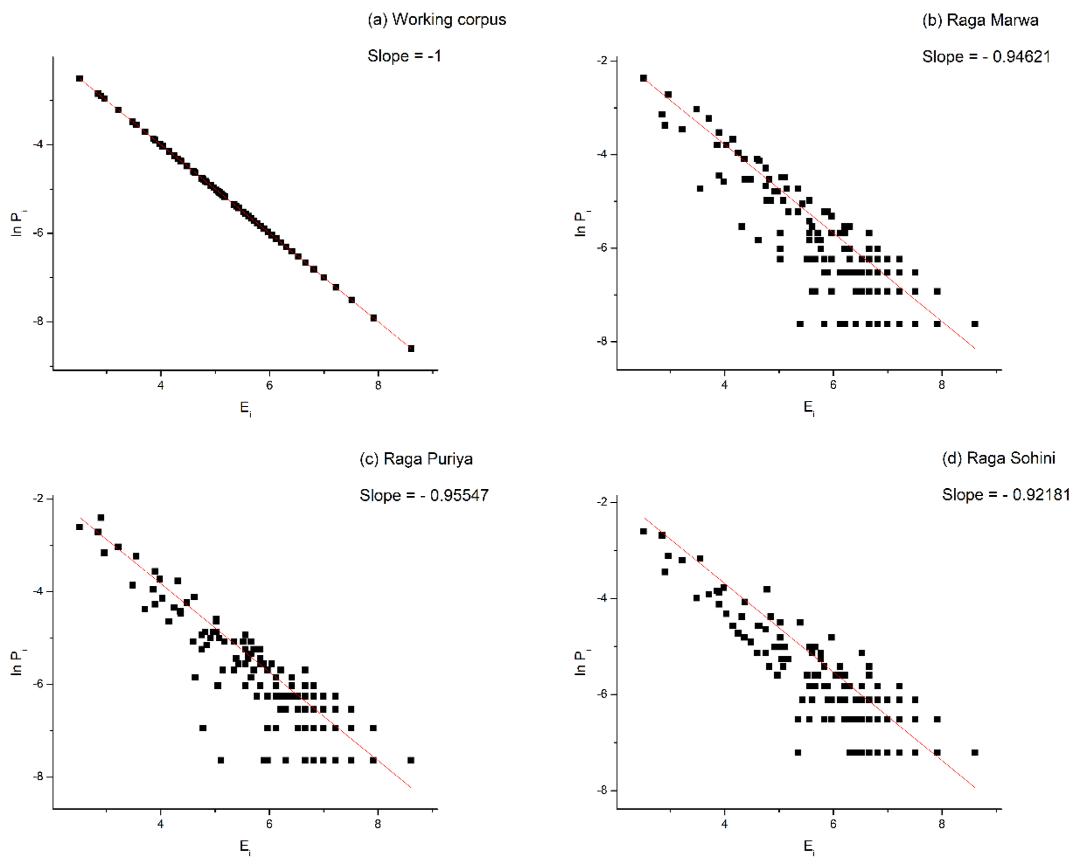


Fig. 6.4 (a-d). ln (p_i) vs E_i plots for a) Working corpus, b) Raga Marwa, c) Raga Puriya and d) Raga Sohini along with respective slope values. It is computed by taking logarithm of both sides of eq. (2). Different slopes indicate different temperatures (Rego et al., 2014).

6.10.2. Plots regarding BE Distribution

The rank-frequency distribution plots for the three ragas are given in fig. 6.5 (a-c). Horizontal plateaus in the figure having high ranks and low frequencies correspond to a large number of components having the same frequency. The longest plateau corresponds to frequency 1, i.e., elements occurring only once in the sample, or '*hapaxes*' (Rovenchak & Buk, 2011).

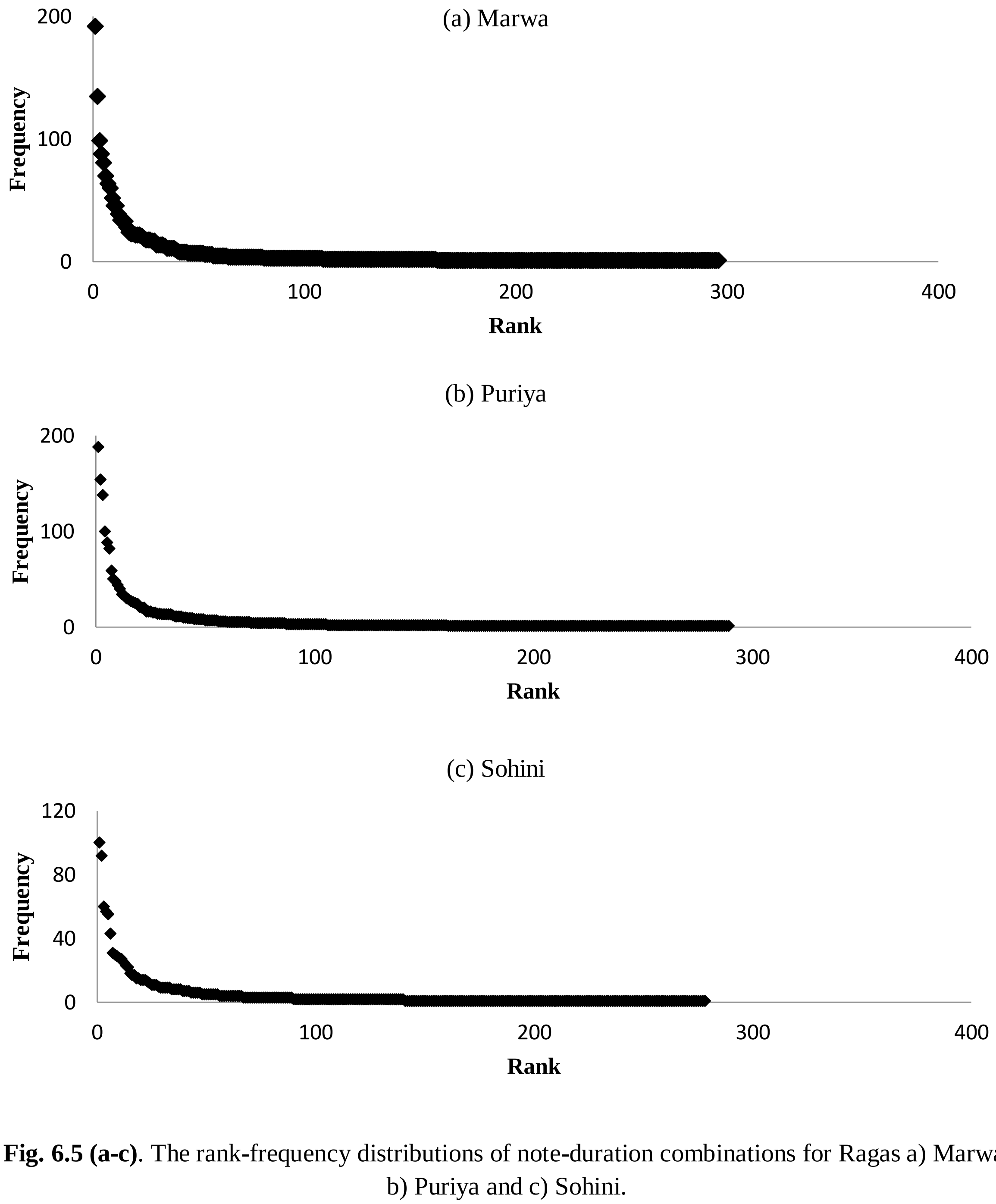


Fig. 6.5 (a-c). The rank-frequency distributions of note-duration combinations for Ragas a) Marwa, b) Puriya and c) Sohini.

Clearly, they follow the Zipf's distribution (Lestrade, 2017). Now we invert the rank-frequency data to energy level j and their occupant number N_j. Corresponding plots are given in fig. 6.6 (a-c). They resemble the exponential decay pattern of BE distribution.

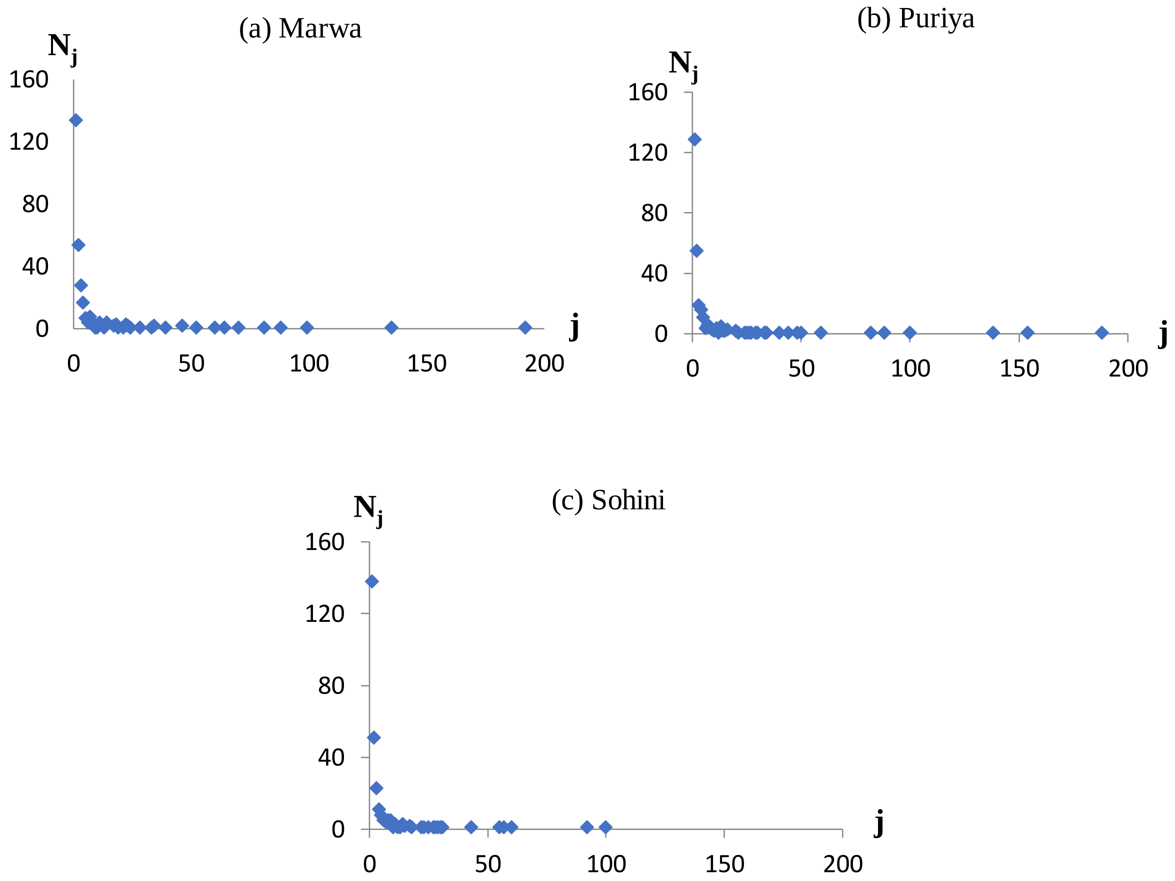


Fig. 6.6 (a-c). j vs N_j plots for a) Raga Marwa, b) Raga Puriya and c) Raga Sohini follow the decay pattern of BE distribution

Evidently, figures 6.3 and 6.5 take after respective distribution patterns. Now, we will move on to the curve fitting data and compute the desired parameters.

6.10.3. Results from Curve Fitting

In this step, the collected data are fitted to equations (5)-(8). The parameters that emerge from it are: N (total number of note-duration combinations), N_1 (total number of once occurring elements), Maxwell-Boltzmann temperature T_{MB}, z (fugacity) and Bose-Einstein temperature T_{BE}. The results of the fitting are elaborated in **Table 6.4**.

This table represents the crux of the experiment. It bears the parameters we were set out to seek and also their interesting implications. Evidently, among the three ragas, the highest number of note-duration combinations (N) were present in case of raga Puriya, which was closely followed by raga Marwa. Sohini, on the other hand, uses the least number of such combinations. But one thing that remains noteworthy is that despite greater difference in their N values, Sohini has the highest N_1 value. That is, the number of elements that occurred only once, is higher in raga Sohini than the other two ragas. We shall come back to explore this soon.

Raga	N (total number of note-duration combination)	N_1 (total occupancy in lowermost state)	T_{MB} (K)	z	T_{BE}	$\tau = \ln T_{BE}/\ln N$
Marwa	2038	134	1.01 ± 0.03	0.9926	132.6 ± 18.8	0.642
Puriya	2075	129	0.98 ± 0.02	0.9923	135.38 ± 18.7	0.643
Sohini	1350	138	1.07 ± 0.02	0.9928	75.77 ± 9.4	0.601

Table 6.4. Using equations (5)-(8), parameters N, N_1, T_{MB}, z and T_{BE} are obtained. An additional parameter $\tau = \ln T_{BE}/\ln N$ is also calculated for categorization purposes.

The next parameter, T_{MB} or the Maxwell-Boltzmann temperature, is found by fitting the probability and corresponding energy data to eq. (5). Table 6.4 shows that the value of T_{MB} is highest in case of raga Sohini (1.07 K). The raga that displayed the lowest T_{MB} is Puriya. Recorded at 0.98 K, the value is even lower than the corpus temperature of 1K. Marwa, marginally higher than Puriya in terms of T_{MB}, is the closest to the corpus, with 1.01 K. From the discussions in the previous chapter, it can be said that proximity to the corpus temperature indicates how closely or frequently a raga uses its characteristic notes or phrases in its movements. Not unlike the efficiency in using the vocabulary of a language. The further it is from the corpus the less efficient it is in using its characteristic notes. In Rego et al. (2014), authors argue that the temperature parameter is associated with the vocabulary complexity of a language. It says that the language having lower temperature has greater complexity, therefor requiring greater effort from the reader whereas languages with high T has lower complexity (due to presence of synonyms to express similar ideas) and requires lesser efforts to be read. Since the analytical parallel of words, in our case, is the note-duration combinations, it can be said that Puriya has the highest complexity in terms of note usage, whereas Sohini has the lowest. Puriya has a *Vadi-Samvadi* pair of Ni2 and Ga2 in exact 5th harmonic relationship, it uses ascending and descending notes in a balanced capacity, moves slowly and known to be of a serene nature. Sohini, on the other hand, has a faster tempo than others, and evokes joyfulness among the listeners. Marwa, sitting between the two, has a Vadi-Samvadi pair of Dha2-Re1 which are not in fifth harmonic relationship. Also, it employs equivalent weightage on Sa and Re1 which renders the raga some tension over its otherwise calm nature. To sum up, it seems that the temperature parameter T_{MB} is generally suggestive of the raga's kinetic nature and the its efficiency of using the 'musical vocabulary'.

Another noteworthy observation that could be made from the above data is that the ratio of N_1 and N (N_1:N) follows the pattern expressed in T_{MB}, i.e., Sohini > Marwa > Puriya. Similar trend was observed in Chapter 5 dealing with Marwa renditions of three artists. A possible explanation could be made by going back to the origin of the analogy – gas molecules in a container. It is well known in thermodynamics that the Maxwell-Boltzmann distribution, when represented in the population vs velocity graph, broadens and shifts toward higher velocity part with increasing temperature (illustrated in Fig. 6.7). This posits that the probability of finding greater number of molecules having high velocity increases, i.e., if one molecule is taken at random, it is more likely that it will have most probable velocity. In our experimental scenario, ratio (N_1:N) represents the probability of finding a random note-duration combination having highest occurrence state 1 (as Fig. 6.5 evidently points out, 1 is the highest occurring frequency among the available elements). The linear relation between T_{MB} and (N_1:N) suggests that more combinations are likely to be found in the most probable state N = 1 as T_{MB} climbs higher. Here, Sohini has the highest T_{MB}. And with its high N_1 value, the probability of finding a random note-duration combination occurring once is justifiably higher in case of Sohini than the other two. This observation further advocates for the analogy of music with statistical ensemble of particles.

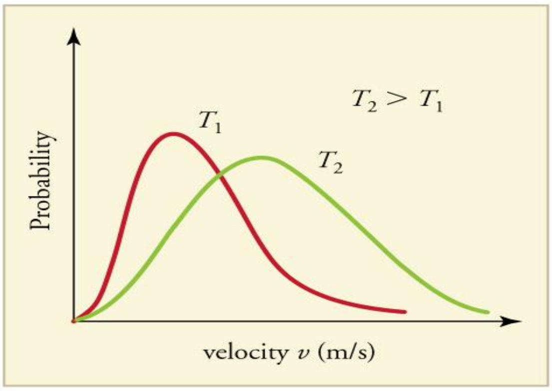


Fig. 6.7. The Maxwell-Boltzmann distribution is shifted towards higher velocities and broadened at higher temperatures (Urone & Hinrichs, 1998).

One final observation that we can find from the Maxwell-Boltzmann calculations, is that, the slope of the ln (p_i)-E_i plots also shares the T_{MB} trend (refer to figure 6.4). We could find the same in the previous chapter as well. This could mean that the slopes of comparative thermodynamic analysis carry the same implications as the temperature. It needs to be investigated further.

The other important parameter that the Table 6.4 provides us, is T_{BE} or the Bose-Einstein temperature. It is found by fitting the N_j-j data into eq. (6). Raga Puriya shows the highest value of T_{BE}, 135.38 (± 18.7), closely followed by raga Marwa at 132.6 (±18.8). Raga Sohini, at 75.77 (±9.4) registers the lowest T_{BE}. Again, referring to the discussions of Chapter 5, it can be seen that this parameter T_{BE} is a testament of 'musical analyticity'. Reduction in diversity of note usage and *alankar*-presence marks the analytic level of the rendition. Lower the T_{BE}, higher its analytic levels, i.e., the resistance to note-diversity and ornamentation; with the lowest level being $T_{BE} = 0$, where all the elements have occurred only once each (analogous to the elusive BEC). Now, in our case, the rendition of Sohini displays the least keenness on inflective additions and note diversity, as it can be confirmed from its note distribution pattern in fig. 6.8. Other than the dominant Sa in the middle octave, rest of the notes are short spanned. Which, possibly, explains why despite having the lowest number of note-duration combinations (N = 1350), raga Sohini has the highest N_1, elements occurring only once (N_1 = 138).

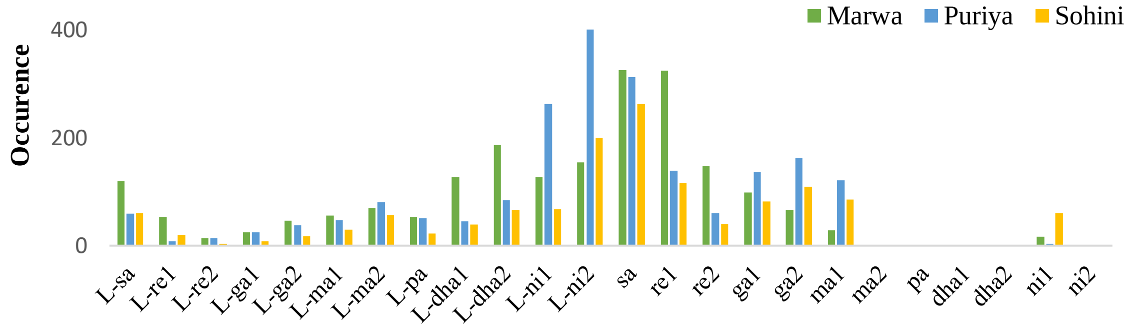


Fig. 6.8. Distribution of the occurrence frequency of notes for three ragas Marwa, Puriya and Sohini. 'L-' indicates the lower octave. '1' and '2' after the notes indicate *komal* and *shuddha* versions of the note, respectively (except for '*ma*', where 1 and 2 denotes *shuddha* and *tivra*). It is seen that Sohini uses more short-spanned notes across the spectrum.

The parameter z corresponds to fugacity. Its value is supposed to be close to 1, the ideal case being z = 1. In each of the Raga renditions, the z value is found to be sufficiently closer to the value of 1.

Table 6.4 also provide an additional parameter, τ, which is the ratio of the logarithms of T_{BE} and N. It remains fairly unchanged with the growth of the sample size (Rovenchak & Buk, 2011). This property makes it a useful parameter for comparative purposes. We will use the N-τ plane here to demonstrate the categorization potential of this parameter. The N-τ plane is given in Fig. 6.9.

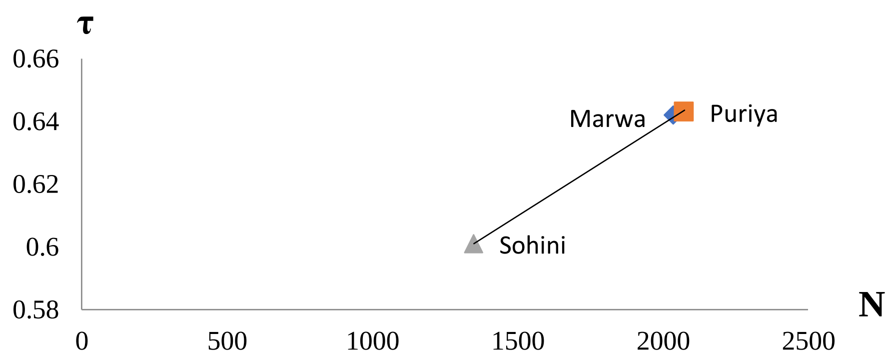


Fig. 6.9. The position of the three Ragas on the N-τ plane.

In linguistic analysis, lower τ value corresponds to higher 'analyticity', i.e., the tendency for words to be single and isolated (contrary to 'syntheticity' – prone to more syntactic relations and more inflections). Our study also corresponds to that notion in lower τ regions. Analyticity also seems to have a relationship with the N_1:N ratio. It is found that the higher the ratio of N_1:N, the more the analyticity (evident from the note distribution structure for the Ragas given in Fig. 6.8). This correlation also asks for further explorations.

6.10.4. Measurement of Goodness-of-fit:

Goodness-of-fit was tested using the determination co-efficient R^2 which seems most suitable (Rovenchak, 2015; Mačutek & Wimmer, 2013) for the data. R^2 is given as eq. (9):

$$R^2 = 1 - \frac{\sum_i (f_i - NP_i)^2}{\sum_i (f_i - \bar{f})^2} \tag{9}$$

Here, f_i = the observed frequency of the value i, P_i = the theoretical probability of the value i, N = the sample size (the total number of the observations) and $\bar{f}$ is the mean of the observed frequencies ($\bar{f} = \sum_i^n f_i/n$). In case of BE data, the summation runs from j = 2 (since j = 1 is fixed by z), for the rest it runs from j = 1.

The R^2 value are given in Table 6.5.

Generally, the R^2 value ≥ 0.9 is considered satisfactory, although ≥ 0.8 is also acceptable (Mačutek & Wimmer, 2013). For MB distribution, the data fits in a satisfactory manner. Whereas, for BE, it lacks the precision. Presumably, this scattering happens due to the relatively low number of data points obtained from the rank-frequency distribution. Increased sample size could increase the efficiency of the fit significantly.

Raga	**MB distributions**	**BE distributions**
Sohini	0.91	0.66
Marwa	0.87	0.68
Puriya	0.90	0.73

Table 6.5. R^2 values for the Ragas Marwa, Puriya and Sohini. The MB distribution data fits acceptably but in case of BE distribution the fit is imprecise.

6.11. CONCLUSIONS

Raga is the backbone of Indian Classical Music (ICM). It is a complex structure made up of several nuances which are almost entirely learned by intuition and experience, even after the presence of melodic grammar. A Raga performer uses a definite set of notes/note combinations in a certain boundary of scale and transits from one note to the next in diverse ways specific to that Raga to colour the listener's mind with a certain emotion. Before the scientific advancements, much of the above process remained qualitative. Acoustic researches in recent years have changed that scenario. Abundance of studies are now being conducted to identify the features and cues that could replicate the intuitive ideas into quantitative parameters with various degrees of success. Fields like artist identification, instrument identification, genre classification, raga classification, music emotion recognition have blossomed which are constantly pushing the horizon of knowledge further and further. Similarly, in this work, we attempted a novel endeavour of analysing music as a physical system using the point of view of statistical mechanics. We set out to quantify acoustic features found in the structure of three ragas, namely, Marwa, Puriya and Sohini, who are so collegial that it takes very accomplished performers of classical music to dare to perform them in succession. Our study didn't wish to focus on their musicological or phraseological differences in distinguishing them, for that, it would not be a task different from a trained musician's. Our objective was to use the informational content found in their structure and develop parameters which would help specify properties that remain elusive to a naïve listener (some of them are too intuitive and abstract to explain even by the skilled professionals). For this, first we used a mid-level feature of musical information, pitch, to spot the notes or *swara*s that provide the necessary skeleton of the raga structure. Then, using that note distribution and their respective durations, we described the musical piece as a gas contained in a vessel with the note-duration combinations as its constituent units, or molecules, so to speak. This treatment has enabled us to ascribe real physical properties on the abstractness of music.

The main conclusions that could be summarised from this study are:

1. The fitting of the probability vs. energy graphs in the Maxwell-Boltzmann distribution (**Fig. 6.3-6.4**) appears as satisfactory. This points to the fact that even with a small corpus, significant results can be obtained.

2. The temperature parameter T_{MB} demonstrates how efficiently a raga's *chalan* or exposition adheres to its corpus characteristics, that is, how much or how little the raga rendition deviates from a certain standard reference that has been set. This helps differentiate between ragas that have similar note patterns with very subtle differences, even for a non-trained listener. Additionally, T_{MB} serves two more purpose. Firstly, its potential in hinting at the vocabulary complexity. By vocabulary complexity we imply the complexity in the note usage patterns for that specific sample. Higher the complexity, greater efforts it will require for the performer and audience alike to discern its signatures and conveying the mood. Secondly, T_{MB} indicates the kinetic nature of the Raga (quite analogous to thermodynamics). Overall, it has the potential to be used as a Classification parameter for musical information research.

3. The consistency in the z-value indicates that the BE distribution analogy with the low frequency data is significant.

4. The parameter $\mathbf{T_{BE}}$ can be used to indicate **diversity in note variations**, that is, the level of musical ornamentation and inflective movements are present in the sample. This clearly has a potential to be used as an improvisational parameter that can, besides distinguishing features between different artists, also tell how an individual's creativity differs in each rendition of the same song he performs.

5. The consistency of parameter τ could be useful in larger categorization problems of different Ragas based on the note distribution pattern. Planes like N-τ can be used for similar tasks.

6. Based on the τ value, ragas as well as individual performances can be categorized in 'musically analytic' to 'musically synthetic' spectrum. This puts emphasis on the improvisational and ornamental features which are not sufficiently explored from musical information perspective.

7. Few interesting observations which need further explorations: the consistency of the slopes of $\ln(p_i)$-E_i plots with the Maxwell-Boltzmann temperature T_{MB}. It is seen that further the slope from the corpus, the more distant the temperature is from the reference temp. This suggests that the slope of the semi-log plot encompasses similar potentials as that of T_{MB}. This could potentially be used in genre classification, if found to be held universally.

8. Another such interesting find was the consistency of the ratio (N_1:N) with the analyticity level of the experimental sample. We didn't seem to find such correlations in the work of the previous chapter; hence it remains a doubt whether this pattern is universal or a random one-off.

To conclude, this was a novel attempt in exploring Raga characteristics with the help of robust methods of statistical mechanics - methods that are generally eponymous with the microscopic reality. Although the initial findings are interesting and encouraging, it lacks a few important components that we must mention before conclusion. Firstly, the corpus problem. Lack of datasets in the field of Indian classical music is a concern that researchers have to face quite frequently. Although the musical tradition of the country is one of the oldest in the world, documentation and digitization of the same is yet to come at par with that of the west. Task-specific datasets are even rarer, discussed at length by Murthy & Koolagudi (2018). They raise the concern that more often the existing literatures stand limited due to their datasets having several constraints. It is, therefore, remain a serious issue which this study also suffers from. Secondly, a pitfall of parametrization is mapping higher dimensional systems or spaces into, say, a two-dimensional plane. To combat this dimensional reduction, further analysis like multivariate discriminant analysis needs to be utilized. This prevails as a future direction at this point.

All in all, we barely scratched the surface in terms of exploring acoustic features of Indian Classical Music. The attempts will continue in the future as well. A broad horizon of work could be sought after using this idea, for example: on the cognitive aspects of the dynamic characteristics of music, acoustic cues responsible for the emergence of associated mood and so on. Algorithmic approach mixed with psychological experiments could help highlight the significant aspects of those features, since these are not isolated properties, rather a concomitance of seemingly endless components working in tandem.

REFERENCES

Aerts, D., & Beltran, L. (2020). Quantum structure in cognition: Human language as a Boson gas of entangled words. *Foundations of Science*, *25*(3), 755-802.

Balkwill, L. L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music perception*, *17*(1), 43-64.

Belle, S., Joshi, R., & Rao, P. (2009). Raga identification by using swara intonation. *Journal of ITC Sangeet Research Academy*, *23*(3).

Bello, J. P. (2007, September). Audio-Based Cover Song Retrieval Using Approximate Chord Sequences: Testing Shifts, Gaps, Swaps and Beats. In *ISMIR* (Vol. 7, pp. 239-244).

Beltran, L. (2021). Quantum Bose–Einstein Statistics for Indistinguishable Concepts in Human Language. *Foundations of Science*, 1-13.

Bhattacharjee, A., & Srinivasan, N. (2011). Hindustani raga representation and identification: a transition probability based approach. *IJMBC*, *2*(1-2), 66-91.

Burrows, J. (2002). 'Delta': a measure of stylistic difference and a guide to likely authorship. *Literary and linguistic computing*, *17*(3), 267-287.

Chang, M. C., Yang, A. C. C., Stanley, H. E., & Peng, C. K. (2017). Measuring information-based energy and temperature of literary texts. *Physica A: Statistical Mechanics and its Applications*, *468,* 783-789.

Chordia, P., & Rae, A. (2007, August). Understanding emotion in raag: an empirical study of listener responses. In *International symposium on computer music modeling and retrieval* (pp. 110-124). Springer, Berlin, Heidelberg.

Datta A.K., Sengupta, R., Dey, N., Nag, D. and Mukerjee A (9-10 January, 2006), *A methodology of note extraction from the song signals.* Proc. Frontiers of Research on Speech and Music (FRSM-2006), U P Technical University, Lucknow.

Dighe, P., Karnick, H., & Raj, B. (2013, November). Swara Histogram Based Structural Analysis And Identification Of Indian Classical Ragas. In *ISMIR* (pp. 35-40).

Fu, Z., Lu, G., Ting, K. M., & Zhang, D. (2010). A survey of audio-based music classification and annotation. *IEEE transactions on multimedia*, *13*(2), 303-319.

Gabaix, X. (1999). Zipf's Law and the Growth of Cities. *American Economic Review*, *89*(2), 129-132.

Ghosh, M. (Ed.). (2002). *Natyashastra: A treaties on ancient Indian dramaturgy and histrionics ascribed to Bharata Muni*. 1–2 vols. Varanasi: Chowkhamba Sanskrit Series Office.

Haro, M., Serrà, J., Herrera, P., & Corral, A. (2012). Zipf's law in short-time timbral codings of speech, music, and environmental sound signals. *Plos One*, *7*(3), e33993.

Holmes, D. I. (1992). A stylometric analysis of Mormon scripture and related texts. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 155*(1), 91-120.

Holroyde, P. (1972). *The Music of India*. Praeger.

Johnston, D. (1999). Cool edit 2000. *Syntrillium Software Corporation, Phoenix, Arizona, EEUU*.

Katte, T. (2013). Multiple techniques for raga identification in Indian classical music. *International Journal of Electronics and Computer Engineering, 4*(6), 82-7.

Katz, J., & Pesetsky, D. (2011). The identity thesis for language and music. *Manuscript, January*.

Koduri, G. K., Gulati, S., Rao, P., & Serra, X. (2012). Rāga recognition based on pitch distribution methods. *Journal of New Music Research*, *41*(4), 337-350.

Kosmidis, K., Kalampokis, A., & Argyrakis, P. (2006). Statistical mechanical approach to human language. *Physica A: Statistical Mechanics and its Applications*, *366*, 495-502.

Krumhansl, C. L. (2001). *Cognitive foundations of musical pitch*. Oxford University Press.

Kumar, V., Pandya, H., & Jawahar, C. V. (2014, August). Identifying ragas in indian music. In *2014 22nd International Conference on Pattern Recognition* (pp. 767-772). IEEE.

Lestrade, S. (2017). Unzipping Zipf's law. *PloS one*, *12*(8), e0181987.

Mačutek, J., & Wimmer, G. (2013). Evaluating goodness-of-fit of discrete distribution models in quantitative linguistics. *Journal of Quantitative Linguistics*, *20*(3), 227-240.

Manaris, B., Romero, J., Machado, P., Krehbiel, D., Hirzel, T., Pharr, W., & Davis, R. B. (2005). Zipf's law, music classification, and aesthetics. *Computer Music Journal*, *29*(1), 55-69.

Mathur, A., Vijayakumar, S. H., Chakrabarti, B., & Singh, N. C. (2015). Emotional responses to Hindustani raga music: the role of musical structure. *Frontiers in psychology*, *6*, 513.

McNeil, A. (2007). Improvisation as Conversation: A Cross-Cultural Context Perspective. *Indian Musicological Society. Journal of the Indian Musicological Society*, *38*, 4.

Miyazima, S., & Yamamoto, K. (2008). Measuring the temperature of texts. *Fractals*, *16*(01), 25-32.

Montemurro, M. A., & Zanette, D. H. (2002). Entropic analysis of the role of words in literary texts. *Advances in complex systems*, *5*(01), 7-17.

Mosteller, F., & Wallace, D. L. (1963). Inference in an authorship problem: A comparative study of discrimination methods applied to the authorship of the disputed Federalist Papers. *Journal of the American Statistical Association*, *58*(302), 275-309.

Mukerjee, R. (1965). " Rasas" as Springs of Art in Indian Aesthetics. *The Journal of Aesthetics and Art Criticism*, *24*(1), 91-96.

Murthy, Y. S., & Koolagudi, S. G. (2018). Content-based music information retrieval (cb-mir) and its applications toward the music industry: A review. *ACM Computing Surveys (CSUR)*, *51*(3), 1-46.

Naldi, M. (2003). Concentration indices and Zipf's law. *Economics Letters*, *78*(3), 329-334.

Neuman, D. M. (1990). *The life of music in North India: The organization of an artistic tradition*. University of Chicago Press.

Perotti, J. I., & Billoni, O. V. (2020). On the emergence of Zipf's law in music. *Physica A: Statistical Mechanics and its Applications*, *549*, 124309.

Reˆgo, H. H. A., Braunstein, L. A., D′ Agostino, G., Stanley, H. E., & Miyazima, S. (2014). When a text is translated does the complexity of its vocabulary change? Translations and target readerships. *PloS one*, *9*(10), e110213.

Reed, J., & Lee, C. H. (2009, April). On the importance of modeling temporal information in music tag annotation. In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 1873-1876). IEEE.

Rosso, O. A., Craig, H., & Moscato, P. (2009). Shakespeare and other English Renaissance authors as characterized by Information Theory complexity quantifiers. *Physica A: Statistical Mechanics and its Applications*, *388*(6), 916-926.

Rovenchak, A. (2015, May). Models of frequency spectrum in texts based on quantum distributions in fractional space dimensions. In *2015 20th International Conference on Control Systems and Computer Science* (pp. 645-649). IEEE.

Rovenchak, A., & Buk, S. (2011). Application of a quantum ensemble model to linguistic analysis. *Physica A: Statistical Mechanics and its Applications*, *390*(7), 1326-1331.

Shetty, S., & Achary, K. K. (2009). Raga mining of Indian music by extracting arohana-avarohana pattern. *International Journal of Recent Trends in Engineering*, *1*(1), 362.

Sjölander, K., & Beskow, J. (2000). Wavesurfer-an open source speech tool. In *Sixth International Conference on Spoken Language Processing*.

Suma, S. M., & Koolagudi, S. G. (2015). Raga classification for carnatic music. In *Information Systems Design and Intelligent Applications* (pp. 865-875). Springer, New Delhi.

Tweedie, F. J., & Baayen, R. H. (1998). How variable may a constant be? Measures of lexical richness in perspective. *Computers and the Humanities*, *32*(5), 323-352.

Urone, P., & Hinrichs, R. (1998). College physics.

Waghmare, K. C., & Sonkamble, B. A. (2017). Raga Identification Techniques for Classifying Indian Classical Music: A Survey. *International Journal of Signal Processing Systems*, *5*(4), 1-6.

Wieczorkowska, A. A., Datta, A. K., Sengupta, R., Dey, N., & Mukherjee, B. (2010). On search for emotion in Hindusthani vocal music. In *Advances in music information retrieval* (pp. 285-304). Springer, Berlin, Heidelberg.

Yang, Y. H., Lin, Y. C., Su, Y. F., & Chen, H. H. (2008). A regression approach to music emotion recognition. *IEEE Transactions on audio, speech, and language processing*, *16*(2), 448-457.

Zanette, D. H. (2006). Zipf's law and the creation of musical context. *Musicae Scientiae*, *10*(1), 3-18.

Zhang, Q., & Sornette, D. (2011). Empirical test of the origin of Zipf's law in growing social networks. *Physica A: Statistical Mechanics and its Applications*, *390*(23-24), 4124-4130.

Zipf, G. K. (1949). Human behaviour and the principle of least-effort. Cambridge MA edn. *Reading: Addison-Wesley*, *24*.

CHAPTER 7

AN IN-DEPTH QUANTITATIVE EXPLORATION OF ABSTRACTNESS OF IMPROVISATION INHERENT IN INDIAN CLASSICAL MUSIC USING STATISTICAL PHYSICS APPROACH

"The ability to perceive or think differently is more important than the knowledge gained."

David Bohm

ABSTRACT

Indian classical music (ICM) is considered to be one of the oldest and most creative art forms existing in this world and *Raga*, in spirit, is the structural unit that binds together the vast expansion of this music genre. A musician expresses the *Raga* according to his present psychological state. Thus, the performances, even with the same Raga, have some subtle differences with each other. These differences are generally called “improvisation”. In this study, we intend to quantify such improvisations using measurable parameters inspired by statistical tools used in Physics. To study them quantifiably, we introduce methods based on well-known concepts of Statistical Physics (especially thermodynamics), namely Maxwell-Boltzmann (MB) statistics and Bose-Einstein (BE) distribution. In this present study, these distributions have been applied to find new parameters (equivalent to ‘temperature’ in physical systems) to identify different features of improvisation in different Hindustani classical music performances of the same Raga by the same artist. Music clips chosen were 6 different renditions of the same *Vilambit bandish* of the same Raga sung by legendary classical music maestro Kumar Gandharva on 6 different occasions. The resulting analysis gives a number of parameters (they come from the analogy between the rank-frequency distribution and the respective statistical distribution) that help in identification and categorization of the improvisational changes in the chosen 6 renditions and thus concepts such as individual improvisation pattern which were previously considered as abstract and not easily explainable to naive listeners (individuals without musical training in ICM) can now be analysed from a quantitative approach. The methods studied here are novel in the music research field and can prove to be useful in the fields of music and speech as quantifying parameters for artist style, Raga identification and genre classification.

Keywords: Maxwell-Boltzmann distribution, Bose-Einstein distribution, Indian classical music, Raga, improvisation, temperature

7.1. INTRODUCTION: A BRIEF NOTE ON IMPROVISATION

7.1.1. Musical improvisation: comparison between Indian and Western Classical Music Scenario

Musical improvisation, rarely explored with as much scientific rigor as other musical features, is an essential part of music evolution and practice among cultures across the globe. Now, the term 'improvisation' is approached rather cautiously by musicologists because of its abstract nature and it remains mostly debated. However, improvisation usually refers to those features or details which are extemporaneous, added by the performer within the heart of the composition but without changing the composition's identity. The nuances and subtle variations generated by the artist that were not pre-planned or pre-composed, can be recognized as an 'improvisation' (Nettl, 1974).

According to Ravi Shankar (2007), to perform a raga in Indian classical music, other than the ascending and descending notes (*Aarohan* and *Avarohan*), characteristic phrase (*Chalan*) and the structural composition (*Bandish*) - everything else is a product of performer's own imagination and creativity, which he acquires via vigorous practice. During a performance of a raga, the pre-composed phrases either gets stretched or constricted. The existing melodic motives, too, may get added on by additional motives, either in front or followed by. Phrases can be cleaved or collapsed. Motives or phrases may be sequenced through different registers as well (Neuman, 1990). In this way, a performer plays with the grammatical strictness, redefining it while staying within its boundary. This is where Indian classical music (ICM) finds its distinctive character, as the raga and its grammar are only means and not the ends here. Therefore, the performer's contribution in ICM appears to be much greater than only reproducing the pre-existing scores and memorized segments cover a minority of the performance in general. Such person-specific interpretation of the musical grammar is the essence of improvisation for ICM. McNeil's (2007) characterization of raga in this regard seems rather appropriate: Raga is unpredictable and ever-blooming, flourishing out to newer forms in each performance and this is the crux of improvisation.

Unlike the Western counterpart, ICM doesn't deal with standardized notation system. Here, the singer/player is himself the composer, for most of the duration of the performance. Indian classical performers have excelled in improvisation as music traditionally has a living oral existence, whereas Western classical music follows a long tradition of reproducing scores that were written previously. A raga is essentially the extension of the performer's psychological and existential state rather than a rigid vessel of rules. Thus, one rendition of the same raga differs from another in more than one subtle way. Even if an artist sings/plays the same *Bandish* of the same *Raga* multiple times, almost never they remain exactly similar to each other. These differences in the rendition of the same raga on different performances are generally called improvisation. Musical improvisation is created in the moment, combining musical grammar with emotions and instrumental/vocal techniques (Alperson, 1984).

Multiple ethnomusicological studies credit the longstanding musical tradition and the interactions to be crucial factors in developing the social hierarchy of North Indian Classical music (Clayton 2005; Clayton et al. 2019). For the majority of Western musicology, improvisation is considered inferior as it is a symbol of opposition to 'composition', which is regarded as the proverbial elite of the musical hierarchy (Sadie & Tyrell, 2001). A glaring contrast in observed in Indian classical music, where "improvisation" takes the central stage in raga performances. Hence, to summarize, Improvisation is pivotal part of Indian classical Music whose expression relies upon the imaginative insight of a particular artist and it can be best identified by analyzing the subtle variations introduced by the same artist in different renditions of the same raga.

7.1.2. Musicological roots of improvisation

Historically, improvisation is majorly ignored by ancient Indian texts on music apart from fleeting mentions about how certain parts of the performance belong to the artist rather the composure. One such aspect was the *Alap*, the lyric-less and rhythmless preface that introduces a raga. In later years, many aspects of music have been attached to this notion. Most prominent of them being vocal genre of '*Khyal*', literally translates to 'imagination', which allows the performer more freedom to improvise than an older stricter component called '*Dhrupad*'. Similarly, in Carnatic classical music, the composition and improvisation parts are differentiated as '*Kalpita*' and '*Kalpana*', respectively (Nooshin & Widdess, 2006). The role of improvisation in North Indian musical tradition followed a rather bittersweet journey. Usually the tradition of '*Guru-shishya parampara*' dominated the scene, especially in instrumental music. This led to the overemphasis on learning memorized materials over improvised ones. Only after years of practice students are allowed to go beyond them. Although comparatively liberal, vocal music also shared similar scenario to some extent (Widdess & Sanyal, 2004).

According to Nooshin & Widdess (2006), signatures of improvisation in ICM is displayed in a handful of components:

- *Vistar* or melodic expansion: the gradual expansion of melody to assimilate successive higher or lower pitches.
- Rhythmic intensification: the gradual escalation of tempo to include different techniques with each level.
- Permutation: re-ordering or permutation of pitches/melodic motif within the boundaries of raga structure.
- Development of individual pitches: focusing on a single pitch and playing around with it by emphasising, prolonging or repeating.
- *Alankar* or sequential transposition: a melodic motif being repeated several times starting on successively higher or lower degrees of the scale.

Thus, the performers in Indian classical music use compositional materials and use techniques in variety of contexts to change their impressions in an adroit manner. A skilled performer can deploy many such strategic changes in various combinations and appearances, a lot of which are novel as well as spontaneous, thereby, providing improvisational attributes to the composition itself.

7.1.3. Previous works on improvisation

There seem to be a dearth in experimental approaches to study improvisation in Indian classical music scenario. Most of the reports are based on Western music. Jazz and folk music have received a lot of attention in this regard (Johnson-Laird, 1991; Berliner, 2009; Sertan & Chordia, 2011). In recent years, improvisational music has been studied in context to its therapeutic capacity. It is reported to have specific benefits for patients of neurological damage in improving mental health conditions and reducing anxiety and stress levels. Also, improvement in communication skills and attentional traits are observed in children with autistic spectrum disorders (Lee, 2000; Kim et al. 2009; MacDonald & Wilson, 2014). The neuroscientific signature of improvisation is explored in a number of studies, with the jazz musicians in the focus (Berkowitz & Ansari, 2010; Donnay et.al., 2014; Beaty, 2015). These studies reported to have identified brain regions involved in the generation of improvised pieces and spontaneous performance of them, simultaneously controlling for the influence of memory (Bengtsson et al., 2007; Limb and Braun, 2008). Walton et al. (2015) studied the role of cooperation with the co-performer in the process of producing musical improvisation in jazz musicians. Contrasting to such diverse range of research, studying improvisation in Indian classical music has mostly been relevant only as a context to another component, like performative gestures. Gestures are generally considered an important part of performance and means of (Gritten & King, 2011). In Indian music, gestures are associated with melodic patterns and movements, but it is not learned, rather improvised. Rahaim (2008) studied the significance of gestures in Indian music context and learned that they help visualize the melodic movements in spatial manner, in three-dimensional space. Other than this and few ethnomusicological takes on improvisation, perhaps the only study that attempts to tackle this issue by a rigorous scientific approach is done by Sanyal et al. (2016). Here, the authors applied chaos based non-linear methods like MFDFA and MFDXA on multiple performances of the same raga by the same artist. Their results show promise in parametrizing possible improvisational cues. To our knowledge, no other work has met this research question with similar scientific tools in such a headfast outlook.

7.2. APPROACH USED IN THE PRESENT WORK

It is not difficult to see why studies on improvisation in the Indian music is fairly limited. The subtleties of such a feature could be hard to characterize, especially when they are so ingrained into the fabric of musical treatise. To analyze the footprints of improvisation in the ICM scenario, in the present work we intend to use tools that have been an integral part of physics and have a wide usage, mostly in the atomic and molecular level. Origin of this approach begins from an empirical law, frequent in the field of linguistic analysis, called Zipf’s law. In 1949, Zipf argued that the speaker/writer prefers to minimize word usage, that is, few words for most meanings whereas the hearer/reader needs things maximized, that is, every meaning has a different word. The higher the degree of satisfaction of the requirements (of one of the two), the less its effort. Zipf's law emanates from this exchange of requirement between the speaker and the audience (Zipf, 2016 reprint). Using a word’s frequency and its rank in a certain text/corpus, eq. (1) expresses this law in mathematical form. If we assign the rank j = 1 to the most frequent word of a language, j = 2 to the second one, etc., then the frequency of occurrence f(j) of a given word varies with its rank j as:

$$f(j) \sim 1.\, j^{-\alpha} \tag{1}$$

Here, α is an exponent which is to be determined from the rank vs. frequency distribution.

Zipf’s law is observed in case of diverse natural phenomena – linguistics, city population distribution, internet traffic, ranking of company sizes, bibliometric data, earthquake probability and more (Li, 2002). The possibility of using it in music related research has already been explored in western classical music. The studies cover music aesthetics (Manaris et al., 2005), creation of musical context (Zanette, 2006), statistical distributions of short-time timbral codings (Haro et al. 2012) to name a few. In a recent study, authors argued that the emergence of Zipf's law is a mere consequence of music being more efficient and complex form of language (Perotti & Billoni, 2020). Zipf's famous law has found a new dimension in linguistic analysis in the last decade or so. A growing body of work has developed which is based on the assumption that human language can be analogous to a physical system (e.g., gas particles in container) within the boundaries of equilibrium statistical mechanics (Miyazima & Yamamoto, 2008; Chang et al. 2017). Very recently, in a novel attempt in music domain, we have applied the idea on Raga categorization in Indian classical music with a good degree of success (Roy et al., 2021). The theoretical fundamentals of the idea have been discussed in Chapter 2; therefore, we will restrict ourselves within a brief overview of it. The central idea can be summarised as: different words with different occurrence frequency has different energies ('word energies') and their energy distribution follows Maxwell-Boltzmann (MB) distribution. It is connected to the Zipf's law via direct correlation of the Hamiltonian with the rank of the word (usefulness) (Kosmidis et al. 2006).

Now, according to statistical mechanics, for a system of particles at constant temperature T, the probability p_i that it is found at a given state i with energy E_i is:

$$p_i \sim 1.\exp(-\beta E_i) \tag{2}$$

Here, $\beta = 1.(kT)^{-1}$; k is the Boltzmann constant ($= 1.38 \times 10^{-23}$ J/K) and T is absolute temperature, which is regarded as the 'measure' of the interaction between system and environment. Taking k as 1, computing p_i values and fitting them into eq. (2) gives the value of T. This 'text temperature' has been used in diverse measures such as authorship disputes, changes in complexity of vocabulary etc.

Music and language share similarities in both cognitive and structural fronts (Katz & Pesetsky, 2011). The processing of both involves creating expectations of what is to come. Fine-scale voluntary vocal production and the ability to imitate it are involved in the learning and producing of both. And most importantly, they share the characteristic of individual's ability to create and improvise, in order to be emotionally meaningful and also, distinct. This prompts us to suspect that the statistical models used in linguistic studies could also be of use in the categorization problems regarding music. In ICM, the *swara*s or notes are the fundamental structural units of the Raga framework. Hence, if a musical rendition was to be assumed a physical system, then the notes are the natural 'Zipfian' units. But unlike words, music has a temporal component which must be taken into account. The notes themselves are only as important as what duration they are performed or sung for. A note sung for 100 ms. is entirely different from it being sung for 200 ms, in context to the raga performance, its structure and its movements. The Zipfian units, in the case of music should have a temporal representation which will take into account the duration of the notes as well. Considering this factor, it was decided that the notes and their occurrence duration together would form the organisational units. The occurrence probability of such a combination would ascertain a corresponding 'energy' whose distribution should, theoretically, give us a characteristic temperature regarding that music sample.

Other than MB distribution which has been associated with high-frequency words, Rovenchak and Buk (2011) suggested a new set of parameters by comparing low-frequency words with a bosonic system within the grand canonical formulation (The rationale is also explained in

Chapter 2). The basis of the idea is reflected in the fact that the rank-frequency distribution of the words in text show similarity with Bose-Einstein (BE) distribution pattern. The result of this analysis includes a temperature like parameter which correlates with the analyticity level of a language. Recently, in unrelated studies, authors have explored the possibility of describing human language as a Boson gas and argued in favour of using ‘words’ as fundamental units carrying different energies. Interestingly, they too related the energy levels with Zipf rankings (Aerts & Beltran, 2020; Beltran, 2021). Low frequency data is of particular interest in context to ICM and improvisation. Since improvisations are spontaneous and deviates from pre-determined compositions, there occurrence in the sample should be fewer compared to the other segments and note distributions. More so, when different renditions of the same raga of the same artist is being compared to each other, because the stylistic base always remain the same. The improvisations, however subtle they are, should reside on the low frequency end of occurrence. Hence, the dynamics regarding the lower frequent notes are of prime importance in this context.

To sum up, our approach renders the music piece as a physical system with computable statistical parameters. Since the improvisation is rendered more accessible when a piece of performance is compared to another rendition of the same piece performed by the same artist, we have decided to analyze multiple performances of the same raga performed by the same artist. Here, we explore the dynamics of used notes (and the duration performer decides to spend on them) and attempt to characterize the improvisational cues the performer might have used in his renditions. To assess our findings on improvisation, we have also consulted a musician who has extensive training in Indian classical music. He opines about the samples used without having any prior knowledge about the experiment or access to the results of the analysis, thereby, eradicating confirmation bias. Our previous endeavours in using these methods in ICM related works have provided some interesting and novel outcomes which has made us hopeful about this study as well.

7.3. MAXWELL-BOLTZMANN (MB) AND BOSE-EINSTEIN (BE) DISTRIBUTIONS IN BRIEF

7.3.1. Maxwell-Boltzmann distribution

The Maxwell-Boltzmann (MB) statistics is generally used for distribution of an amount of energy between identical but distinguishable particles. MB statistics predicts that the probability of finding a particle with a specific energy decreases exponentially with increasing energy, considering the system consists of a huge number of non-interacting distinguishable particles.

The distribution function has the following form:

$$f(E_i) = 1/Ae^{Ei/kT} \qquad (3)$$

where $f(E_i)$ is the probability of a particle having energy E_i, A is the normalisation constant, E_i is the energy of the i-th state, k is the Boltzmann constant, and T is absolute temperature.

7.3.2. Bose-Einstein distribution

Bose-Einstein (BE) statistics describes the dynamics of an ensemble of identical and indistinguishable particles occupying discrete energy states. The distribution function indicating the energy distribution looks like:

$$f(E_i) = 1/[Ae^{E_i/kT} - 1] \quad (4)$$

where $f(E_i)$ is the probability of a particle having energy E_i, 1/A denotes the degeneracy, i.e., how many particles are having particular energy state E_i, E_i is the energy of the i-th state, k is the Boltzmann constant, and T is absolute temperature. The -1 factor in the denominator recognises the fact that the particles are indistinguishable, unlike MB statistics. BE distribution applies to a very particular kind of particles who have integer spin values, known as Bosons. They do not obey the Pauli's exclusion principle and hence unlimited number of particles can occupy the same energy state (The particles that obey the Pauli's exclusion principle are called Fermions). This unusual property of the BE distribution helps in applying this concept in various systems beyond the sub-atomic world.

Fig. 7.1 shows the MB and BE distributions on a population vs. energy graph.

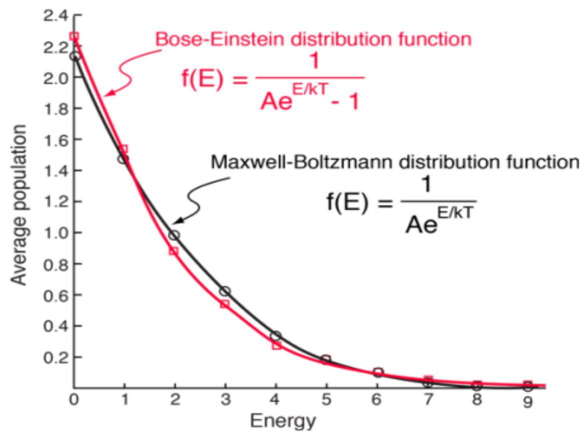


Fig. 7.3. Maxwell-Boltzmann and Bose-Einstein distributions, population vs energy plot. Mathematically, they are quite similar, except for the -1 factor in BE to factor in the indistinguishability of the particles. [*Image courtesy*: http://hyperphysics.phy-astr.gsu.edu; *URL*: http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/imgqua/disbemb.png]

7.4. OVERVIEW OF THE WORK

Improvisation and related characteristics are integral parts of Indian classical music (ICM) and its emotional fabric. It has been a distinctive feature of the musical tradition of this country for centuries. Yet researches, using rigorous scientific paradigm to study improvisation, still remain scarce. Musicological studies are mostly qualitative. In view of this, we intend to quantify such abstractness using tangible parameters. To study musical information, we introduce methods based on well-known concepts used in Statistical Physics, namely Maxwell-Boltzmann (MB) and Bose-Einstein (BE) distribution. The approach we follow in this study is based on the analogy between the rank-frequency distributions (using Zipf's law) of a note-duration combination of a music sample and the statistical distributions (both the MB distribution and the BE distribution in grand canonical formulation). To apply MB statistics to music, it is assumed that musical combination units (of notes and their durations) with different occurrence frequencies are at different energy levels, the distribution of which follows the MB distribution pattern. In case of BE statistics, a rank-frequency distribution of the time durations of various notes of different ragas is studied. The resulting analysis gives rise to a number of parameters that help to characterize the musical information. One of which is the equivalent of Temperature in case of physical systems. This 'temperature' parameter is quite familiar in linguistic research and it has been used to specify the underlying dynamics of various languages, authorship disputes, changes in complexity of vocabulary and many more. When dealing with Indian classical music, such parameters help in highlighting different features of improvisational aspects. The chosen music clips belonged to the *Vilambit Bandish* part of raga *Sur Malhar* performed by Pandit Kumar Gandharva, a maestro of Indian classical music. We have used six renditions of the same raga sung by the great artist in six different performances. With the resulting parameters, we found features that separate these versions on the basis of improvisation. Also, we could categorise them accordingly. The methods described here are novel in the music research field and can prove to be useful in the study of Indian classical music as quantifiers for improvisational characteristics.

7.5. EXPERIMENTAL DETAILS

7.5.1: Choice of music clips

The music clips we chose to study belonged to raga '*Sur Malhar'*, performed by Pandit Kumar Gandharva, one of the most renowned vocalists of ICM. Sur Malhar (or *Surdasi Malhar*) belongs to *Thaat* Kafi and the characteristic notes that accompany this raga are: Sa/high Sa, Shuddha Re (Re2), Shuddha Ma (Ma1), Pa, Shuddha Dha (Dha2), Komal Ni (Ni1), Shuddha Ni (Ni2). The symbol of the notations was discussed in Chapter 6. The *Vadi-Samvadi* pairs are: Ma1 and Sa. We used six different performances of raga Sur Malhar sung by the artist in six different occasions to examine their improvisational differences. The part of the composition under study was from the *Vilambit Bandish* segment. *Bandish* is said to be the literature equivalent of the raga, as in, it holds the compositional structure and characteristic notes specific to the raga. The reason for choosing *Bandish* part was the fact that the note usage in all the renditions would remain the same, thereby, the consequent changes in the structural patterns would be because of the lyrical or melodic improvisations included by the artist each time he performed the raga. This treatment would help us segregate the improvisational features. The chosen clips were of 4 minutes duration.

7.5.2: Processing of Music Signals

All the audio signals were digitized at a sampling rate of 44.1 kHz 16-bit mono channel format using Cool Edit software of Syntrellium Corporation (Johnston, 1999) and amplitudes of all the clips were normalized to 0 dB. These sound signals constitute our database for this experiment. Pitch extraction technique is defined in next section. For this, Wavesurfer software of KTH, Stockholm was used (Sjölander & Beskow, 2000).

7.5.3: Pitch Extraction and Note Profiles

The sound signals which were analyzed in this study have all been normalized to 0 dB, and hence intensity or loudness is not being considered here. Normalization helps removing any DC offset (which causes distortion) from the track by centering the waveform on the 0 dB amplitude level (maximum amplitude level). This is done by averaging all the sample values in the selection, and then subtracting the average value from the exact value of all the samples.

The extraction of notes from the obtained pitch profile of each of the acoustic signal was done with the help of an experienced musician who has profound knowledge of the tonic. A skilled musician was asked to listen to the signal files one after another to detect the position of tonic 'Sa' in the signal file. The notes were extracted from the signal file following the methodology of Datta et.al (2006). The ratio-intervals were evaluated by first dividing the smoothed pitch values for each song by the pitch value of the 'Sa' tonic of that *raga* rendition. This gives the frequency ratios for each pitch data. From the ratio data, steady state sequences were created with all consecutive pitches in a sequence, which is terminated when $| x_{i+1} – M | > M/30$, where $M = (1/i) \Sigma x_i$ where x_{i+1} is the $(i+1)^{th}$ pitch and x_i is the i^{th} pitch. If the duration of any sequence were found to be less than that of a certain minimum value then the sequence is rejected. Elements extracted from these sequences were taken as suitable candidate data for this analysis. Whenever the ratio was found to be less than 1, it is multiplied by a factor of 2 and when it is greater than 2 it was divided by 2. This exercise effectively folds all pitch data into the middle octave. All the extracted pitch values were distributed in 1200 bins of one-cent width each. The peaks of these distributions for each musical signals are expected to be indicative of the note positions for that *raga* rendition. After detecting the value of 'Sa', we have calculated the pitch value of other notes (Datta et.al, 2006), thus each note was found out along with their time duration.

7.6. GENERAL METHODOLOGY

The general methodological approach, described in Chapter 2 and Chapter 5, remains the same. Here, we will only refer to the relevant equations and parameters that we wish to seek in this work.

7.6.1: Maxwell-Boltzmann (MB) Distribution

Our 'working corpus', in this study, consists of all the available note-duration combinations from all the six renditions of the raga that we are analysing. Firstly, the 'energies' of each note-duration combination is evaluated using eq. (2). Then, they are compiled and plotted against the probability of occurrence in the following seven cases: the corpus and all six renditions each, separately, with the help of model equation Eq. (5).

$$p(E) = y0 + A1.\ \exp(-E/t1) \tag{5}$$

Here, t1 denotes the desired Maxwell-Boltzmann temperature (T_{MB}) in Kelvins. The results from the curve fitting will also be used in comparison purposes.

7.6.2: Bose-Einstein (BE) Distribution

Here, our focus is on the low-frequency end of the data. The rank-frequency distribution plots of note-duration combinations are plotted. Followed by this, we invert the rank frequency data to find the corresponding N_j vs j plots. The relevant equation here is given by eq. (6):

$$N_j = \frac{1}{z^{-1}.\exp(\varepsilon_j/T_{BE}) - 1} \tag{6}$$

Here, N_j is occupancy number of j^{th} energy level and z is fugacity. The energy of the j^{th} level is ε_j and T_{BE} is the Bose-Einstein temperature. The spectrum of ε_j is given by eq. (7).

$$\varepsilon_j = (j - 1)^{\alpha} \quad (7)$$

Unity is subtracted to make sure that the lowermost energy state, j = 1, has zero energy. According to Rovenchak and Buk (2011), this spectrum is effective only for low-frequency data because in high frequencies, the spectrum of ε_j might have a different configuration.

Value of z is found from eq. (6) using j = 1,

$$N_1 = \frac{z}{1-z} \quad (8)$$

The other parameters are determined from eq. (6) by fitting the acquired data to the equation.

7.7. RESULT AND DISCUSSIONS

7.7.1: Plots regarding MB Distribution

The probability of occurrence of the note-duration combinations is plotted against the respective energies, for seven cases: the working corpus, and the six raga renditions. The plot for probability vs energy for the working corpus is given in Fig. 7.2.

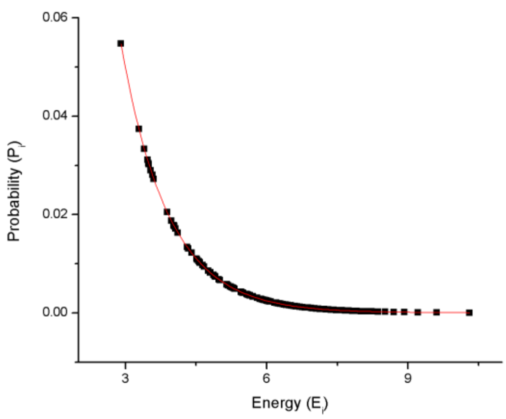


Fig. 7.2. Distribution of Probability (of occurrences of note-duration combinations) vs. 'Energy' plot. This constitutes the p_i vs E_i plot for the working corpus made up of every note-duration combination computed in all six renditions of raga *Sur Malhar.*

As we can expect, the plot is perfectly exponential and perfectly fits the data since the energies were computed by taking k = T = 1 in eq. (2). Of course, it follows the MB distribution. We assumed the corpus temperature to be 1K for comparison purposes, as this will serve as the reference temperature. Now, the individual p_i-E_i plots for six performances are given in Fig. 7.3 (a-f).

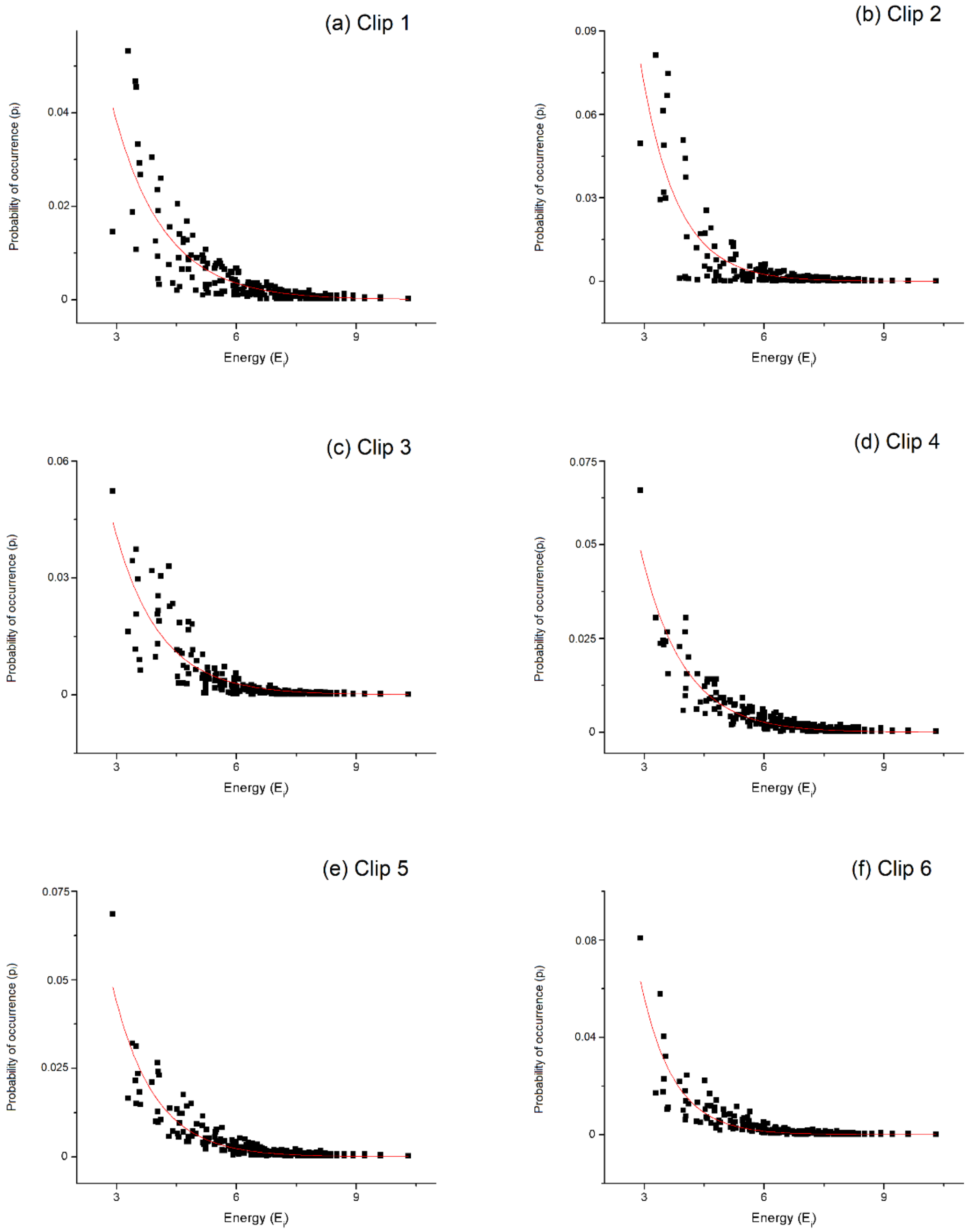
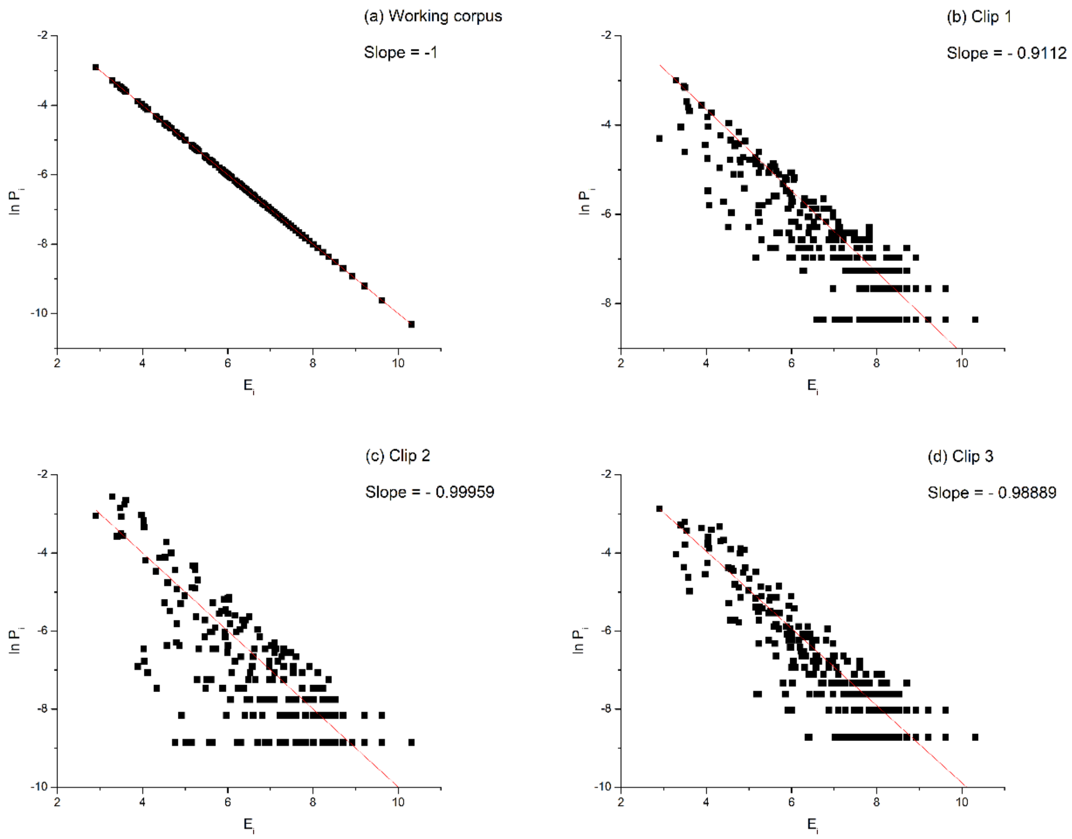


Fig. 7.3 (a-f). Occurrence probability vs Energy plots for Clips 1-6 of six raga renditions. Occurrence probability of each note-duration combination is computed with respect to all such combinations present in the corpus and the energies are calculated via eq. (2).

The red line corresponds to the best fit of the data. It is clearly seen that the plots do follow MB distribution pattern. Next, the respective data are fitted to eq. (5) which will provide Maxwell-Boltzmann temperature T_{MB}. The results of the curve fitting are given later in Table 7.2. The semi-log plots of the above samples are (taking logarithm on both sides) used by Rego et al. (2014) in so-called 'comparative thermo-linguistic analysis'. It is seen that comparing the slopes of the samples can be of significance, just as in the language example, the slope value differed even when the corpus of the samples showed similar Zipf distribution characteristics. In our music samples, plots for the comparative thermodynamic analysis are given in Fig. 7.4 (a-g). Values of corresponding slopes are given in Table 7.1.

Sample	Working corpus	Clip 1	Clip 2	Clip 3	Clip 4	Clip 5	Clip 6
Value of slope	-1	- 0.9112	- 0.99959	- 0.98889	- 0.92954	- 0.9175	- 0.93526

Table 7.1. Slope values obtained from corresponding E_i vs ln (p_i) plots given in Fig. 7.4 (a-g). Clip 2 and 3 are closest to the reference corpus, Clip 1 and 5 are furthest.



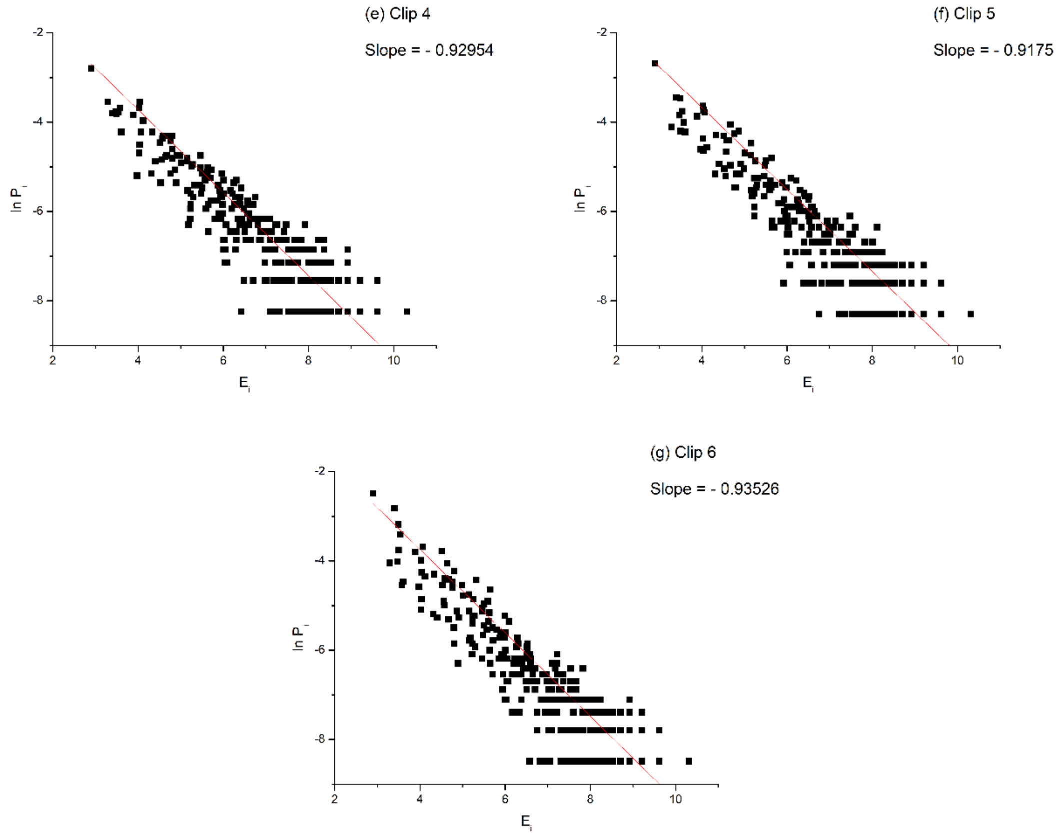
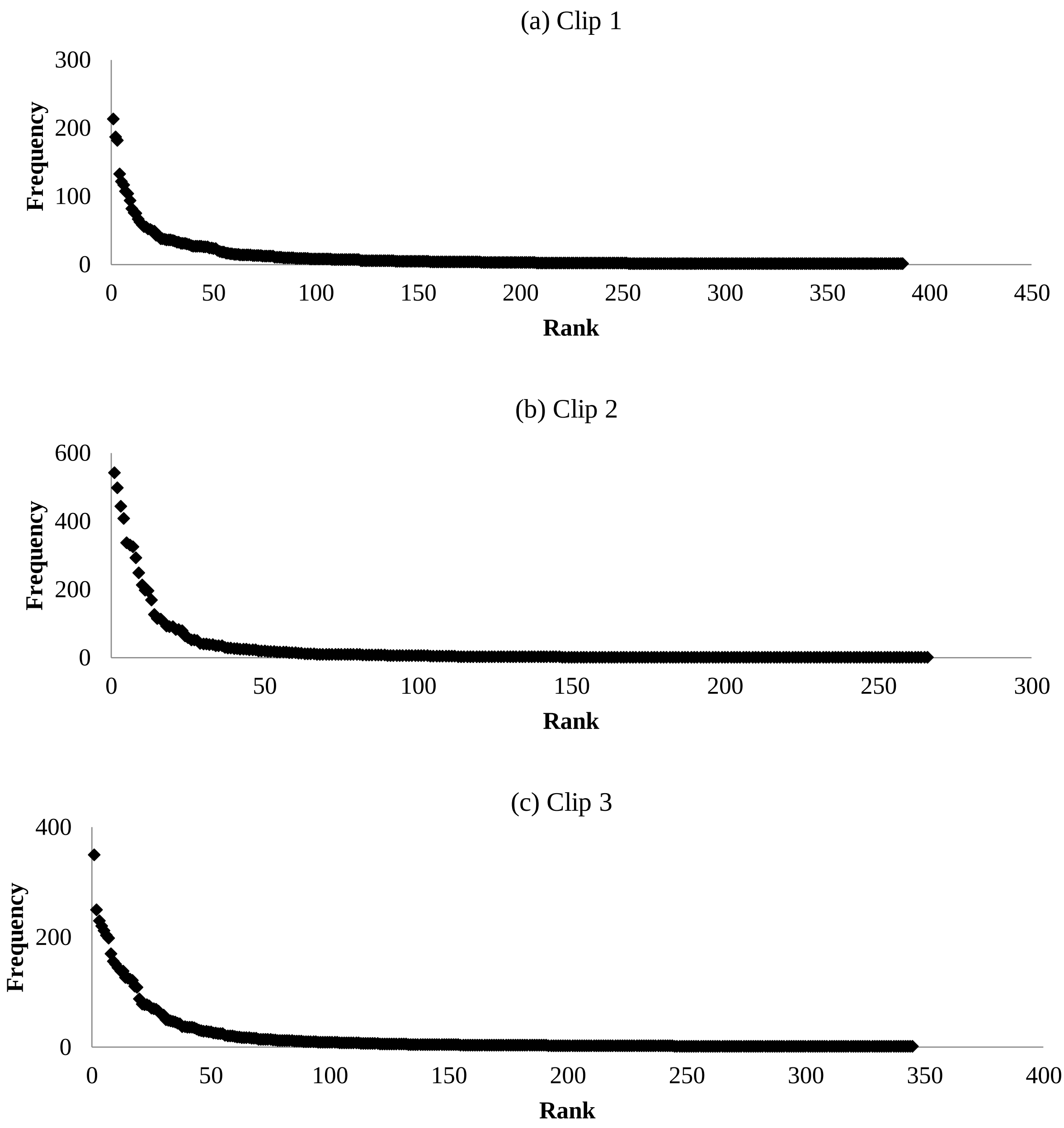


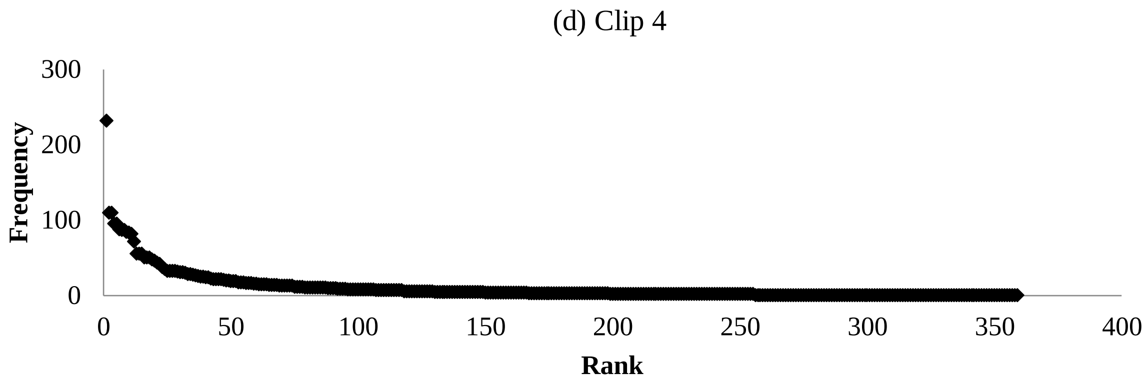
Fig. 7.4 (a-g). ln (p_i) vs E_i plots for Working corpus (a) and six raga renditions (b-g) along with respective slope values. It is computed by taking logarithm of both sides of eq. (2). Different slopes indicate different temperatures (Rego et al., 2014).

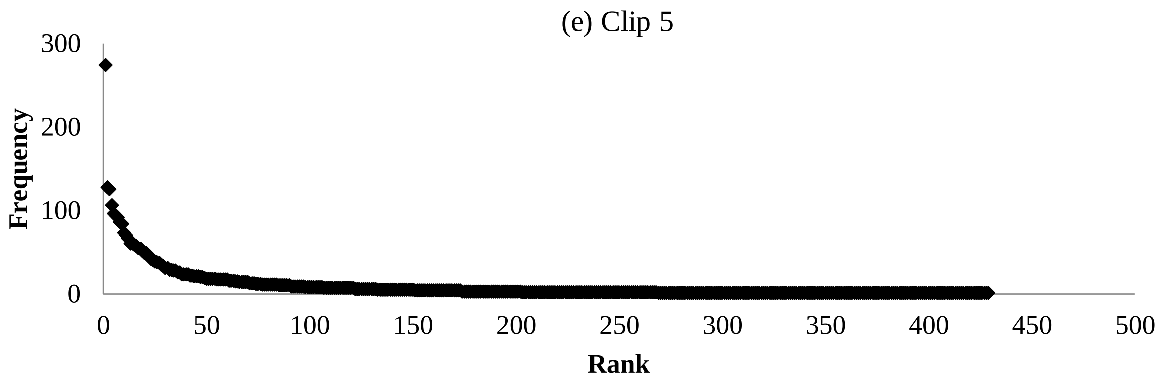
From Fig. 7.4 we can see, expectedly, the slope of the semi-log plot for the working corpus is -1. This was the reference for our temperature evaluation. Now, from Table 7.1, the sample whose slope was closest to that of the corpus is Clip 2 (- 0.99959), closely followed by Clip 3 (- 0.98889). Clip 1 and 5 were the furthest with slopes – 0.9112 and – 0.9175, respectively. The change in slope is a general indicator for the change in temperature, since the ranks of the constituent elements also change (Rego et al., 2014). Also, it shows that even though it is the same raga was performed by the same artist in all six occasions, the performances themselves are not the same. Note distribution and their durations have gone through subtle changes. Hence, it can definitively be said that there were elements of improvisation involved that brings the difference in these performances. Also, in the previous chapters, we had seen that the slope values have a possible correlation with Maxwell-Boltzmann temperature T_{MB}, as they share a similar trend: the further the slope of a sample is from reference, the more distant its T_{MB} is from corpus temperature. We will check from the fitting data whether this trend holds here as well.

7.7.2: Plots regarding BE Distribution

The rank-frequency distribution plots for the six samples are given in Fig. 7.5 (a-f). Horizontal plateaus in the figure having high ranks and low frequencies correspond to a large number of components having the same frequency. The longest plateau corresponds to frequency 1, i.e., elements occurring only once in the sample, or '*hapaxes*' (Rovenchak & Buk, 2011).







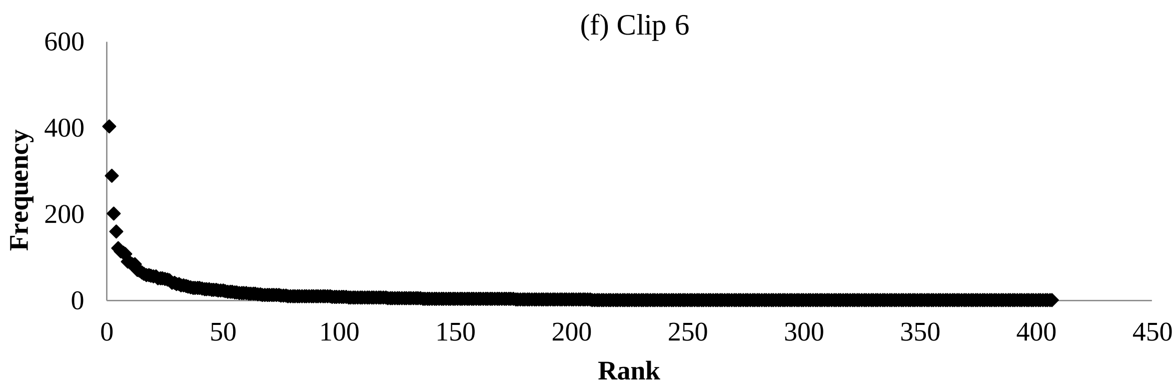


Fig. 7.5 (a-f). The rank-frequency distributions of note-duration combinations for six performance clips. The plateaus correspond to elements having same frequency, the longest plateau being 1.

Following the example given by Lestrade (2017), it is clear that the rank-frequency distributions of the six samples follow Zipf's distribution pattern. Now we invert the rank-frequency data to energy level j and their occupant number N_j. Corresponding plots are given in fig. 7.6 (a-f). They resemble the exponential decay pattern of BE distribution.

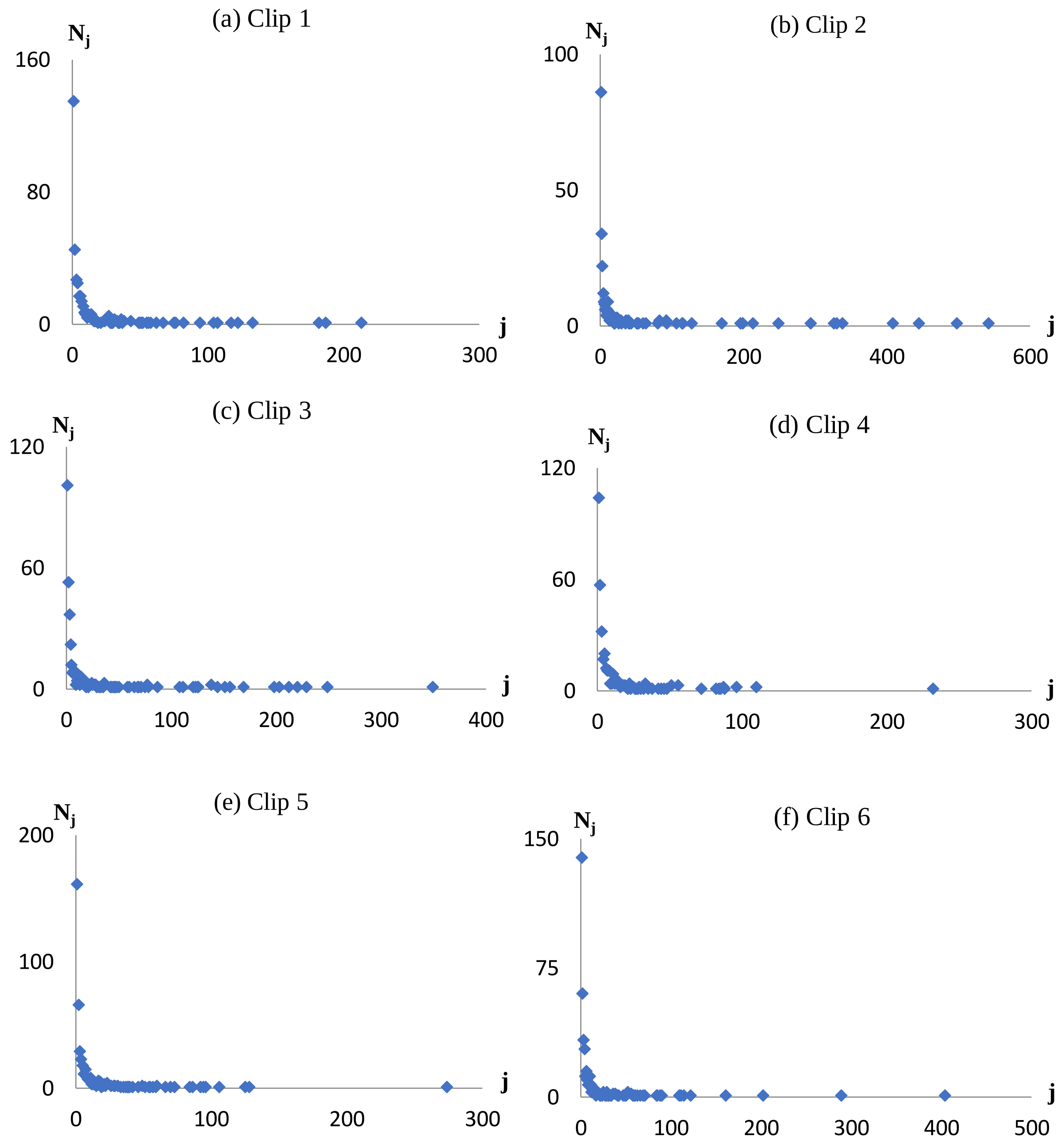


Fig. 7.6 (a-f). N_j vs j plots for six performance clips follow the decay pattern of BE distribution

Evidently, figures 7.3 and 7.5 take after respective distribution patterns. Their goodness-of-fit are given in section 7.7.4. Now, we will move on to the curve fitting data and compute the desired parameters.

7.7.3: Results from Curve Fitting

In this step, the collected data are fitted to equations (5)-(8). The parameters that emerge from it are: N (total number of note-duration combinations), N_1 (total number of once occurring elements), Maxwell-Boltzmann temperature T_{MB}, z (fugacity), α and Bose-Einstein temperature T_{BE}. The results of the fitting are elaborated in **Table 7.2**.

Table 7.2 represents the essence of the experiment. It holds all the parameters we were looking for and also their interesting implications. Among the six samples, it is seen that the number of note-duration combinations have varied. N is lowest for Clip 4 and highest for Clip 2. However, the occupancy number of lowermost occurrent state or N_1 is highest for Clip 5 and lowest for Clip 2. It doesn't have a linear relationship (as in, high N doesn't ensure high N_1). If a strict grammatic avenue was followed thoroughly in each performance, this linearity should be maintained since every sample consists the same *Bandish* part sung by the same artist. Evidently the improvisations impacted the performances.

Clip No.	**N (total number of note+duration combination)**	**N_1 (total occupancy in lowermost state)**	**T_{MB} (K)**	**z**	**α**	**T_{BE}**	**τ = (ln T_{BE}/ln N)**
1	4275	135	1.26 ± 0.05	0.993	1.30	121.49 ± 6.56	0.575
2	7004	86	0.91 ± 0.04	0.988	0.84	29.90 ± 1.41	0.394
3	6115	101	1.13 ± 0.04	0.990	1.73	154.22 ± 19.08	0.578
4	3798	104	1.07 ± 0.02	0.991	1.89	285.87 ± 30.22	0.686
5	4015	161	1.03 ± 0.03	0.994	1.18	91.60 ± 7.68	0.544
6	4862	139	0.82 ± 0.03	0.993	0.79	46.07 ± 2.42	0.451

Table 7.2. Using equations (5)-(8), parameters N, N_1, T_{MB}, z, α and T_{BE} are obtained. An additional parameter $\tau = \ln T_{BE}/\ln N$ is also calculated for categorization purposes.

The next parameter we will consider is the Maxwell-Boltzmann temperature or T_{MB}. Clips 1 and 3 has the highest T_{MB} with 1.26 K and 1.13 K respectively. Two of the clips, 2 and 6, recorded T_{MB} values lower than the corpus temperature, i.e., 0.91 K and 0.82 K. Clips 4 and 5, with 1.07 K and 1.03 K respectively, had the closest T_{MB} values to the corpus. Now, we had seen in previous chapters that the closer the value of T_{MB} to the corpus temperature, the more efficient that sample is in using its constituent note-duration combinations. From that, we can say in the performances of clips 4 and 5, the note distributions were used most efficiently than the other four renditions. It was established that T_{MB} can be used in two more ways in characterizing the experimental sample. One of which, the measure of complexity in 'musical vocabulary' of a raga, is not applicable here as the samples all belong to *Sur Malhar*. Yet, it can be suggested that the renditions having lower T_{MB} (clips 2 and 6) corresponds to greater complexity in the musical exposition or *Chalan*. The other one is the measure of the kinetic nature. We consulted a trained classical musician about his opinion on the samples. His observations are compared with the T_{MB} observations in Table 7.3. It was seen that the renditions closest to corpus temperature has the most balanced distribution of fast moving *Taan*s (the combination of notes rendered in a faster speed) and slow moving *Meend*s (a particular technique of Indian music for smooth gliding from one note to other). Proper presence of *Vadi-samvadi* pairs standing notes (where the melodic pattern 'rests') was also observed in their cases. The renditions that have higher T_{MB}, here the improvisation patterns were present due to both *Taan*s and *Meend*s, but very little lyrical variations were present. Among lower T_{MB} renditions, Clip 6 was observed to have more pauses and limited *Meend*s but no lyrical variations. Clip 2, however, had fast-moving jumping notes with almost no *Meend*s or standing notes.

Therefore, it can be seen that barring the exception of clip 2, the kinetic nature of the other renditions follows their T_{MB} trends as predicted.

Clip No.	T_{MB} (K)	Observations made by expert
1 3	1.26 1.13	Improvisation patterns due to presence of *Taan*s and *Meend*, almost no Lyrical variations
2 6	0.91 0.82	Fast note transitions in Clip 2, rhythmic transitions, Clip 6 dominated by *Meend*
4 5	1.07 1.03	Better balance of *Taan*s and *Meend*s. Proper presence of standing notes (*Vadi-samvadi* pairs)

Table 7.3. Comparison of T_{MB} observations with opinions of Classical music expert.

Another noteworthy observation is that the slope of the semi-log plot has not followed the temperature T_{MB}. One of the renditions, clip 5, which had the second highest slope deviation, is seen to have the closest temperature to the corpus. Similarly, Clip 2, whose slope value was very close to the reference, has a bigger temperature deviation.

Moving on to the next important parameter, T_{BE} or the Bose-Einstein temperature, it was observed that clip 4 has the highest T_{BE} among the six, 285.87 (± 30.22) whereas clips 2 and 6 are among the lowest T_{BE} obtainers, with 29.90 (± 1.41) and 46.07 (± 2.42), respectively. T_{BE}, as we know, is the testament of 'musical analyticity', indicating measures of note diversity and ornamentation level in the sample. Lower the T_{BE}, higher its analytic levels, i.e., the resistance to note-diversity and ornamentation; with the lowest level being $T_{BE} = 0$, where all the elements have occurred only once each (analogous to the elusive BEC). This makes it a great parameter for improvisation related affairs. Here, as mentioned by the expert, clip 2 has no *Meend*s and faster transitions, making the diversity and ornamentation spectrum limited. Clip 6 has *Meend* presence, but interspersed with more pauses than the other samples. These possibly are the reasons for which these clips register lowest T_{BE} values. Similarly, clip 4 displays greater note variations with *Taan*s and *Meend*s present in balanced shares. However, T_{BE} measures fail to hold up in the case of clip 5. It is seen that even after the presence of comparable diversity, it has a significantly lower T_{BE} than clip 4.

The parameter z corresponds to fugacity. Its value is supposed to be close to 1, the ideal case being z = 1. In each of the six samples, the z value is found to be sufficiently closer to the value of 1.

The parameter α is the exponent of a power excitation spectrum and for lower j's, $1 < \alpha < 2$ is expected (Rovenchak, 2014) which implies good fitting of the data to the energy spectrum. In this study, α values remained within limits in four cases and exhibits slightly lower values for clips 2 and 6. This indicates that the data fitting for BE distributions were mostly adequate.

One more parameter was included in Table 7.2. That is τ, which is the ratio of the logarithms of T_{BE} and N. It is useful for comparison and general classification purposes, when samples are plotted in planes like α-τ and N-τ. Here, we will use the α-τ plane, illustrated in Fig. 7.7. Low values of α and τ are associated with higher analyticity, which means samples residing in low α-low τ area includes fewer range of inflexions, i.e., musical ornamentations. This is resonated with the plot in Fig. 7.7 as well.

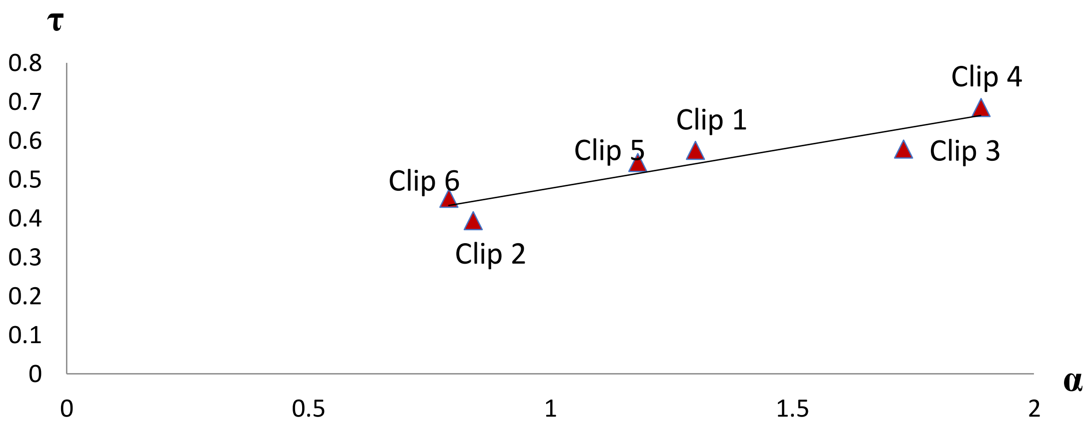


Fig. 7.7. The position of the raga renditions on the α-τ plane. It is seen that low α-low τ region hosts high 'musical analyticity'.

It was seen in the previous chapter that the $(N_1{:}N)$ ratio also corresponds to analyticity levels. High $(N_1{:}N)$ indicates higher analytic rendition. In this example, although not in exact 1:1 correspondence, similar trends can be observed except for clip 2 whose $(N_1{:}N)$ ratio and analyticity level both were found to be lowest. Also, clip 5 is in the middle zone in terms of analyticity despite having the highest $(N_1{:}N)$ ratio. Hence, this pattern shows promise but doesn't seem to be conclusive as of yet.

7.7.4: Measurement of Goodness-of-fit:

Goodness-of-fit was tested using the determination co-efficient R^2 which seems most suitable (Rovenchak, 2015; Mačutek & Wimmer, 2013) for the data. R^2 is given as eq. (9):

$$R^2 = 1 - \frac{\sum_i (f_i - NP_i)^2}{\sum_i (f_i - \bar{f})^2} \quad (9)$$

Here, f_i = the observed frequency of the value i, P_i = the theoretical probability of the value i, N = the sample size (the total number of the observations) and $\bar{f}$ is the mean of the observed frequencies ($\bar{f} = \sum_i^n f_i/n$). In case of BE data, the summation runs from j = 2 (since j = 1 is fixed by z), for the rest it runs from j = 1.

Clip	MB distributions	BE distributions
1	0.84	0.64
2	0.76	0.57
3	0.83	0.61
4	0.87	0.70
5	0.83	0.64
6	0.79	0.59

Table 7.4. R^2 values for the six raga renditions. The MB distribution data fits acceptably but in case of BE distribution the fit is imprecise.

The R^2 value are given below in Table 7.4. Generally, the R^2 value ≥ 0.9 is considered satisfactory, although ≥ 0.8 is also acceptable (Mačutek & Wimmer, 2013). Similar to the previous occasions, the fitting for Maxwell-Boltzmann distribution is satisfactory but for Bose-Einstein, it lacks the precision. Presumably, such scattering happens due to the relatively low number of data points obtained from the rank-frequency distribution, when converted to j and N_j. Increased sample size could increase the efficiency of the fit significantly.

7.8. CONCLUSIONS

Improvisation is a concept as revered as it is debated. Some have criticised music's alleged affiliation to it (Hanslick, 1891). On opposite end of such repudiation, lie historical accounts of ancient Greeks whose musical traditions were replete with improvisations (Burkholder et al. 2019). The western music of Baroque era, jazz and some folk also produce improvisations. So does our very own Indian Classical Music. In fact, here, it is one of the distinctive features of music practice, specially in vocal renditions. Hence, improvisations have always dominated the musical discourse. But more often than not it is really hard to decipher what is meant by it, specifically. Improvisation usually refers to spontaneously generated elements or patterns in the structure of the musical composition. Elements that were not pre-determined or pre-planned. Skilled performers introduce these subtleties that amazes the audience and conveys the mood of the song. In Indian classical music, a singer practically acts as a composer because he improvises while he sings memorized melodic patterns. In the body of a raga, there are multiple such segments where a performer leaves his signature with the improvisation that he adds. It is the innovativeness of the performer that fuels this creative addition. Without stepping out of a fixed framework of a *raga,* a musician can make variations in the lyrics, rhythm, or melodic structure. This constitutes as the spirit of Indian classical music.

It was on this basis, we started our exploration to find improvisational characteristics in the body of ICM. Our objective was to try to quantify the abstractness of improvisation, using novel analytical tools based on statistical physics. The experimental samples belonged to multiple renditions of the same raga sung by the same artist. This treatment could help us to quantitatively analyse how the improvisation occurs, if any at all. Using the combination of notes and their durations as units, we treated the musical piece as a gas in a container and came up with parameters that might be useful in studying the improvisational attributes.

The major conclusions that could be summarised from this study are:

1. The fitting of the probability vs. energy graphs in the Maxwell-Boltzmann distribution appears as satisfactory. Therefore, it points to the fact that even with a small corpus, significant results can be obtained.

2. From the calculated slopes of the energy vs probability semi-log plots indicate clear temperature differences between six samples, which in turn, shows that even though the same artist performed the same raga, each performance is distinct from the other. This leads us to a two-fold conclusion. Firstly, it definitively proves that however subtle it is, improvisations exist and that is the reason for the change in temperature values as with each improvisation, the ranks of the note-duration elements change within the rendition due to the addition/ removal of note variations. Secondly, the 'comparative thermodynamic analysis' can be used as an indicator of improvisational segregation. In chapter 5, we discussed about the existence of such a possibility. This confirms our findings regarding its potential as a parameter that can classify improvisational features.

3. The Maxwell-Boltzmann temperature T_{MB} has again demonstrated that it can successfully determine how efficient a music sample's usage of its corpus characteristics are. Renditions with temperature closest to the corpus were found to have the most balanced structural traits. It was also seen that T_{MB} can predict the kinetic nature of the music sample to a good degree of success. Apart from clip 2, kinetic properties of the other samples were corroborated with respective T_{MB} values.

4. It can also be observed from the data that the slope obtained from the 'comparative thermodynamic analysis' and T_{MB} does not follow the same pattern, something which was consistent in the previous chapters. One possible reason could be the increase in N values. That is, the correlation could only be existing for lower value of N. However, the reasoning is not dismissive enough to disregard it completely. It should be investigated in future endeavours too.

5. The consistency in the z-value indicates that the BE distribution analogy with the low frequency data holds true.

6. Barring two occasions, consistency in α values indicate that the energy spectrum fits well for low frequency data. For higher frequencies, spectrum of ε_j needs to be investigated. It also has an inverse correlation with the analyticity level of the rendition.

7. The total number of elements and the total number of hapaxes do not necessarily follow a linear growth. High N does not guarantee high N_1. This could also be an indicator of improvisation in this particular example. Since the baseline was set equal for all six renditions (same raga, same artist), only the presence of improvisational changes could explain such an irregularity.

8. The Bose-Einstein temperature T_{BE} is a measure of diversity in note variations and indicates the presence of musical ornamentation and inflective movements, or 'musical analyticity'. It clearly differentiates six samples based on their improvisational shift in note usage and related variations. However, it fails to explain clip 5 and its structural changes.

9. Comparing the T_{BE} with T_{MB} results, one can notice that clips that have faster note transitions and limited *Meend* portion, also display lesser improvisational variety and clips that are dominated by *Meend* have higher diversity of note usage (barring clip 6). It suggests that there might be a possible correlation between presence of *Meend* and the improvisation patterns present of the performance. Similar ideas were proposed by Ghosh et al. (2018). Our analysis also hints at the possibility. Future works specifically dedicated to exploring this relation could help advancing our knowledge about improvisation further.

10. The ratio (N_1:N) and the analyticity level of the samples also could have a possible relationship. For now, it doesn't have a high enough correlation and can be disregarded in relation to our present work.

11. α-τ plane can be used to categorize and compare between different genres, different ragas of the same melodic origin, different artists of the same genre or different renditions of the same performance as they vary in information content and contextual liabilities. τ remains fairly unchanged under the growth of the sample size which makes it an ideal parameter in comparative studies. It is seen that the low α- low τ region includes performances with higher analytic quality and the more inflective ones reside in high α-high τ part of the plane.

12. Using τ values, we can segregate the six samples into three clear categories (Fig. 7.7). Clips 4 and 3 belong to 'musically synthetic' group, exhibiting more ornamentation and melodic diversity. 'Musically analytic' group can be constituted with clips 2 and 6 which have higher resistance to inflective movements. Rest of the clips fall in the middle who show traits of both analytic and synthetic. Therefore, once T_{BE} is computed and analyticity levels are exposed, then τ can categorise samples accordingly into clusters.

To conclude, here we have attempted to present a novel idea of investigating the musical information of Indian classical music using fundamental statistical tools (M-B and B-E distributions), extensively used in the domains of physical world. The parameters it yielded can categorize different performances and their improvisational characteristics on the basis of note occurrence and presence of note-duration variation. Usage of such statistical methodologies as a classificatory algorithm in the music domain is unique. With larger data and rendition diversities, further correlation between the parameters and finer categorization of musical information could be possible, we believe. Number of parameters can also be extended to other thermodynamic variables. The early results are promising and indicative that this method could be used in the fields of speech and music for style identification, classification and improvisational feature extraction.

REFERENCES

Aerts, D., & Beltran, L. (2020). Quantum structure in cognition: Human language as a Boson gas of entangled words. *Foundations of Science*, *25*(3), 755-802.

Alperson, P. (1984). On musical improvisation. *The Journal of Aesthetics and Art Criticism*, *43*(1), 17-29.

Beaty, R. E. (2015). The neuroscience of musical improvisation. *Neuroscience & Biobehavioral Reviews*, *51*, 108-117.

Beltran, L. (2021). Quantum Bose–Einstein Statistics for Indistinguishable Concepts in Human Language. *Foundations of Science*, 1-13.

Bengtsson, S. L., Csíkszentmihályi, M., & Ullén, F. (2007). Cortical regions involved in the generation of musical structures during improvisation in pianists. *Journal of cognitive neuroscience*, *19*(5), 830-842.

Berkowitz, A. L., & Ansari, D. (2010). Expertise-related deactivation of the right temporoparietal junction during musical improvisation. *Neuroimage*, *49*(1), 712-719.

Berliner, P. F. (2009). *Thinking in jazz: The infinite art of improvisation*. University of Chicago Press.

Burkholder, J. P., Grout, D. J., & Palisca, C. V. (2019). *A History of Western Music: Tenth International Student Edition*. WW Norton & Company.

Chang, M. C., Yang, A. C. C., Stanley, H. E., & Peng, C. K. (2017). Measuring information-based energy and temperature of literary texts. *Physica A: Statistical Mechanics and its Applications*, *468*, 783-789.

Clayton, M. (2005). Communication in Indian raga performance. *Musical communication*, 361-381.

Clayton, M., Jakubowski, K., & Eerola, T. (2019). Interpersonal entrainment in Indian instrumental music performance: Synchronization and movement coordination relate to tempo, dynamics, metrical and cadential structure. *Musicae Scientiae*, *23*(3), 304-331.

Datta A.K., Sengupta, R., Dey, N., Nag, D. and Mukerjee A (9-10 January, 2006), *A methodology of note extraction from the song signals.* Proc. Frontiers of Research on Speech and Music (FRSM-2006), U P Technical University, Lucknow.

Donnay, G. F., Rankin, S. K., Lopez-Gonzalez, M., Jiradejvong, P., & Limb, C. J. (2014). Neural substrates of interactive musical improvisation: an FMRI study of 'trading fours' in jazz. *PLoS one*, *9*(2), e88665.

Ghosh, D., Sengupta, R., Sanyal, S., & Banerjee, A. (2018). Improvisation—A New Approach of Characterization. In *Musicality of Human Brain through Fractal Analytics* (pp. 185-212). Springer, Singapore.

Gritten, A., & King, E. (Eds.). (2011). *New perspectives on music and gesture*. Ashgate Publishing, Ltd.

Hanslick, E. (1891). *The beautiful in music: A contribution to the revisal of musical aesthetics*. Novello.

Haro, M., Serrà, J., Herrera, P., & Corral, A. (2012). Zipf's law in short-time timbral codings of speech, music, and environmental sound signals. *Plos One*, *7*(3), e33993.

Johnson-Laird, P. N. (1991). Jazz improvization: a theory at the computational level. *Representing musical structure*.

Johnston, D. (1999). Cool edit 2000. *Syntrillium Software Corporation, Phoenix, Arizona, EEUU*.

Katz, J., & Pesetsky, D. (2011). The identity thesis for language and music. *Manuscript, January*.

Kim, J., Wigram, T., & Gold, C. (2009). Emotional, motivational and interpersonal responsiveness of children with autism in improvisational music therapy. *Autism*, *13*(4), 389-409.

Kosmidis, K., Kalampokis, A., & Argyrakis, P. (2006). Statistical mechanical approach to human language. *Physica A: Statistical Mechanics and its Applications*, *366*, 495-502.

Lee, C. (2000). A method of analyzing improvisations in music therapy. *Journal of Music Therapy*, *37*(2), 147-167.

Lestrade, S. (2017). Unzipping Zipf's law. *PloS one*, *12*(8), e0181987.

Li, W. (2002). Zipf's Law everywhere. *Glottometrics*, *5*(2002), 14-21.

Limb, C. J., & Braun, A. R. (2008). Neural substrates of spontaneous musical performance: An fMRI study of jazz improvisation. *PLoS one*, *3*(2), e1679.

MacDonald, R. A., & Wilson, G. B. (2014). Musical improvisation and health: a review. *Psychology of Well-being*, *4*(1), 1-18.

Mačutek, J., & Wimmer, G. (2013). Evaluating goodness-of-fit of discrete distribution models in quantitative linguistics. *Journal of Quantitative Linguistics*, *20*(3), 227-240.

Manaris, B., Romero, J., Machado, P., Krehbiel, D., Hirzel, T., Pharr, W., & Davis, R. B. (2005). Zipf's law, music classification, and aesthetics. *Computer Music Journal*, *29*(1), 55-69.

McNeil, A. (2007). Improvisation as Conversation: A Cross-Cultural Context Perspective. *Indian Musicological Society. Journal of the Indian Musicological Society*, *38*, 4.

Miyazima, S., & Yamamoto, K. (2008). Measuring the temperature of texts. *Fractals*, *16*(01), 25-32.

Nettl, B. (1974). Thoughts on improvisation: A comparative approach. *The musical quarterly*, *60*(1), 1-19.

Neuman, D. M. (1990). *The life of music in North India: The organization of an artistic tradition*. University of Chicago Press.

Nooshin, L., & Widdess, R. (2006). Improvisation in Iranian and Indian music. *Journal of the Indian Musicological Society*, *36*, 104-119.

Perotti, J. I., & Billoni, O. V. (2020). On the emergence of Zipf's law in music. *Physica A: Statistical Mechanics and its Applications*, *549*, 124309.

Rahaim, M. (2008). Gesture and melody in Indian vocal music. *Gesture*, *8*(3), 325-347.

Reˆgo, H. H. A., Braunstein, L. A., D′ Agostino, G., Stanley, H. E., & Miyazima, S. (2014). When a text is translated does the complexity of its vocabulary change? Translations and target readerships. *PloS one*, *9*(10), e110213.

Rovenchak, A. (2014). Trends in language evolution found from the frequency structure of texts mapped against the Bose-distribution. *Journal of Quantitative Linguistics*, *21*(3), 281-294.

Rovenchak, A. (2015, May). Models of frequency spectrum in texts based on quantum distributions in fractional space dimensions. In *2015 20th International Conference on Control Systems and Computer Science* (pp. 645-649). IEEE.

Rovenchak, A., & Buk, S. (2011). Application of a quantum ensemble model to linguistic analysis. *Physica A: Statistical Mechanics and its Applications*, *390*(7), 1326-1331.

Roy, S., Banerjee, A., Sanyal, S., Ghosh, D., & Sengupta, R. (2021, April). A study on Raga characterization in Indian classical music in the light of MB and BE distribution. In *Journal of Physics: Conference Series* (Vol. 1896, No. 1, p. 012007). IOP Publishing.

Sadie, S., & Tyrrell, J. (2001). *Dictionary of music and musicians*. New York: Oxford University Press. Yónatan Sánchez.

Sanyal, S., Banerjee, A., Patranabis, A., Banerjee, K., Sengupta, R., & Ghosh, D. (2016). A study on improvisation in a musical performance using multifractal detrended cross correlation analysis. *Physica A: Statistical Mechanics and its Applications*, *462*, 67-83.

Sertan, S., & Chordia, P. (2011). Modeling melodic improvisation in Turkish folk music using variable-length Markov models. In *12th International Society for Music Information Retrieval Conference* (pp. 269-274).

Shankar, R. (2007). *My music, my life*. Mandala Publishing Group.

Sjölander, K., & Beskow, J. (2000). Wavesurfer-an open source speech tool. In *Sixth International Conference on Spoken Language Processing*.

Walton, A. E., Richardson, M. J., Langland-Hassan, P., & Chemero, A. (2015). Improvisation and the self-organization of multiple musical bodies. *Frontiers in psychology*, *6*, 313.

Widdess, R., & Sanyal, R. (2004). *Dhrupad. Tradition and Performance in Indian Music*. Ashgate.

Zanette, D. H. (2006). Zipf's law and the creation of musical context. *Musicae Scientiae*, *10*(1), 3-18.

Zipf, G. K. (2016). *Human behavior and the principle of least effort: An introduction to human ecology*. Ravenio Books.

CHAPTER 8

NONCLASSICALITY IN THE COGNITION OF AUDITORY STIMULUS WITH AMBIGUOUS MUSICAL PROPERTIES

"If a man will begin with certainties, he shall end in doubts; but if he will be content to begin with doubts, he shall end in certainties."

Sir Francis Bacon

ABSTRACT

Sensory organs in human body receive external experiences in various forms of stimuli (visual, auditory, olfactory, tactile) which get transported to the brain via electrical impulses. But the exact nature of its 'cognition' is still largely unknown. Another equally important concern is whether our present physical theories are sufficient to describe such complex phenomenon. We start with the hypothesis that cognitive phenomena regarding ambiguous auditory stimuli can't be described using classical physical or mathematical theories. In classical physics, the well-known formula of total probability is represented by the sum of discrete probability values of discrete events including the respective conditional probabilities of A and B, if they are two dichotomous random variables. Here we seek to examine whether an induced dichotomy in auditory signal makes the cognitive properties adhere to this classical notion of total probability. For this study, as auditory stimuli we took two pairs of contrast emotion (happy-sad) music clips (each clip of 30 seconds duration and absolute tempo or BPM fixed for each pair) and for the perceptual experiment, we took total 100 participants, divided them into two equal groups (say Group 1 and Group 2), and posed two different experimental conditions. 100 participants were divided into two equal groups. Group 1 was exposed to a pair of music clips (having similar tempo but contrasting musical structure, separated by five seconds) and were asked whether the two clips had the same tempo. Group 2 was made to listen to a different such pair first, then followed by the previous pair after 15 seconds and was asked the same question for each pair. To eradicate musician/non-musician bias, participants were asked not to use any means of tempo measurements, including finger/foot-tapping. The answer 'yes' to the posed question on Pair A gave p(A+), and 'no', gave p(A-). Responses from Group 2 gave us probabilities p(B+), p(B-), p(A+/B+), p(A-/B+), p(A-/B+), p(A-/B-). If classical theories hold true, the collected data should match the law of total probability. But our analysis clearly revealed that for Group 2, the calculated values of p(A+) and p(A-) using the law of total probability, i.e., p(A±) = p(B+).p(A±/B+) + p(B-).p(A±/B-), featured significant difference from the p(A+) and p(A-) values calculated for Group 1. Thus, we have found an intriguing result of violation of classical probability laws in case of ambiguous auditory signals. This data agrees with previous reports made on visual stimuli and hence, can be of serious significance in further explorations on the nature of the non-classicality in cognitive properties.

Keywords: Non-classicality, Probability, Ambiguity, Auditory stimulus, Music, cognition

8.1. INTRODUCTION: CLASSICAL MODELS UNDER SCRUTINY

The brain is a complex system of biological materials exhibiting self-organized characteristics (Rosu 1994). Despite years of efforts, we are yet to decipher the intricacies of those characteristics in their entirety, including how it interacts with external environment (Hopfield 1984). The well-known idea is that sensory organs in human body receive external experiences in various forms of stimuli (visual, auditory, olfactory, tactile) which is transported to different parts of the brain dedicated to receive different set of stimuli through neuronal activation. Various complex processes of information delivery involving electrical impulses have also long been studied. But the number of researches on the possible causes and locations of properties such as perception, cognition and most importantly, consciousness are few and far in between. Where, when and how in the path of travelling nerve impulses these properties emerge still remain as open questions. A number of theories throughout the years were proposed about the working principle of this complex biological organ. For some, it is a complex neural network, obeying classical theories (Mézard et al. 1987, Amit & Amit 1992, Harvey 1994). On the other hand, a growing community of authors have forwarded several non-classical arguments including entanglement, superposition, superconductivity, Bose-Einstein condensate, superflourescence to be responsible for the neuropsychological or electrophysiological activities in the brain (Penrose 1989; Penrose et al. 2000; Stapp 2004; Walker 1970; Marshall 1989; Zohar & Marshall 1990; Ricciardi & Umezawa 1967; Vitiello 1995; Hameroff & Watt 1995). Although the field of studies vary from one another, some shared notions do exist across them. It is agreed that any conscious system has to be self-aware. The sensory information and abstract actions like feelings, intentions and judgements regarding that information are constantly feed off each other to 'create' the sensory world and environment around us. And as human cognitive experience is an ever-continuous process of unionizing varied aspects like information congregation, its evaluation and its awareness, similar complementarity is also expected from the knowledges acquired via neurobiological and psychological reports, regardless of the explanation behind such phenomena. While the advancement of neuroscience and neuro-imaging has made it convenient to study neuronal actions and their evolution in spatio-temporal domains resulting in identification and observation of cognitive properties in functional aspect, but there exists different phenomenology that are more easily and accurately described via psychological terms. However, the association of neurobiological findings with the data described in psychological studies still remain elusive. In an ideal scenario, the quantifiable parameters of neuroscience should be adequate enough to explain the abstract psychological concepts - something which rarely occurs in reality. Some have argued that the reason behind such aperture stems from the absence of appropriate physical theories which could be applied to explain psychological properties (Schwartz et al. 2005). The critic against the ruling theoretical approach of determinism and reductionism in neuroscience is echoed in the light that modern physics has themselves abandoned it: "*[Physicists] invented the deterministic-reductionistic philosophy and taught it to the biologists, only to walk from it themselves*" (Loewenstein 2013).

One such methodological shortcoming in cognitive modelling is the usage of classical logic and probabilistic framework. There have been reports in cognitive psychology in the last few decades that urge for alternative theories in explaining the psychological space. The main derivations were observed in two major sections of this field – 'concept theory' and 'decision theory' (Sozzo 2017). Concept theory, in classical terms, alludes to the 'graded typicality' of a concept, i.e., a concept typically belongs more to type 1 than type 2 (for example, the concept of *Robin* is more typical of *Bird* than *Stork*) (Rosch 1973). Although classical fuzzy set logics appear adequate to explain it, but researches in concept conjunction have indicated otherwise. The famous 'Pet-Fish' (Osherson & Smith 1981) and 'Fruit-Vegetable' problems (Hampton 1988) demonstrate that typicality of a concept can be associated more with conjunction of the types ('*Pet-Fish'*, instead of '*Pet'* or '*Fish'*), as well as disjunction ('*Fruit or Vegetable*', instead of '*Fruit*' or '*Vegetable*' separately). Classical models of fuzzy set logic have failed to explain such conceptual grading. In the second example involving decision theory, it is seen that the probability of estimating the conjunction of two events is higher than the probability of each event independently (Tversky & Kahneman 1983). Hence, it is in a direct disagreement with classical probability laws. One more such effect regarding decision theory is the 'disjunction effect' where it is seen that human choice between two actions 1 and 2 might change depending on whether prior knowledge about an event E is available to them or not. If they know for sure that the event is occurring (or not occurring), they prefer 1 over 2. But in case of uncertainty over E, their preference changes towards 2 (Tversky & Shafir 1992). This fallacy, too, violates the law of total probability and in extension, the classical modelling framework.

The nature of such arguments against the usage of traditional theories has compelled researchers to look for different models involving advanced physical concepts.

8.2. THEORETICAL BACKGROUND

One alternative approach in describing the cognitive space is known as 'quantum cognition model'. Although originally quantum theory was devised to deal with subatomic particles in dimensions unimaginable to everyday human experience, but due to its rigorous theoretical framework and astounding mathematical accuracy it has transcended the microscopic world to enter meso and macro-worlds in theory and, specially, in application. Constant experimental verification helped extend this physical theory to have many practical applications which hold a lion's share of today's economic and technological developments. Quantum cognition, contrasting as it may sound, employs the conceptual and mathematical structure of quantum physics in the cognitive modelling problems. This approach is somewhat distinct from the quantum neurobiology approaches which discuss the possibility of non-trivial quantum effects present in the structure and inner workings of human brain (Jedlicka 2017). Instead, it focuses on developing 'quantum-like' models to explain cognitive properties (like decision, memory, learning, question order effects), mainly using quantum probabilities.

One of the most revered scientific developments, Quantum theory, came into existence a century back in the early 1900s. Even in the early days of its inception, the prospect of using it in cognitive domain have intrigued its founding fathers. In 1929, Alfred Whitehead, on the question of the nature of reality, viewed the 'reality' as a complex collection of quanta of 'events' (Whitehead & Sherburne 1957). These 'events' are repeatedly occurring in space-time and constitutes the 'subjective experience'. Consciousness is an emerging coherent by-product of such temporal stream of discrete events (Hameroff, 2003). Later, Shimony found the compatibility of Whitehead's ideas with the process of quantum state reduction described in the widely accepted Copenhagen interpretation of quantum theory. State reduction represents the Whiteheadian '*events*'. Hence, according to Shimony, quantum theory represents 'reality' in microscopic scale (Shimony 1993). Similar reductionist approaches to the connection of mental space and quantum physical properties have been forwarded by different physicists and philosophers at different points of time. As the Copenhagen interpretation advocates the subjectivity of reality as a direct act of observation (or measurement), authors have argued in favour of a one-to-one correspondence between quantum mechanical events and basic mechanisms of cognitive properties. Some of the pioneers of modern quantum theory has participated in such explorations including Heisenberg (Heisenberg & Northrop 1958), Wigner (1963) and Mermin (1985). The discourse of material-mental connection since has been bifurcated into two broad directions. First one involves the idea that studying the material structure of brain is sufficient enough to understand the mental properties, i.e., anything non-material (consciousness, aware of self, cognition) is an epiphenomenal outcome of the complex biochemical interaction. This point of view is strongly reductionist and have been severely critiqued as well (Kim 1998). An offshoot of this view is a far less reductionist 'emergent phenomena' approach where it is suggested that mental properties are emergent phenomena of the material brain and knowledge about the latter is not sufficient to include the former. In this picture there remains some residual if the non-material is reduced to material. Evidently, a question of the effect of causality arises between the two. This is where quantum theory, Copenhagen interpretation to be precise, has been used to argue that the collapse of the wave function is associated with an act of 'thinking' (Orlov 1982). Materially put, a wave function usually describes a quantum system and a collapse happens to the function when it is under measurement or observation. In the mental domain, authors argue, this collapse represents realization/ actualization of one event from the superposition of infinite number of possible alternatives. And such state reduction provides the basis of cognitive processes (Conte et al. 2003). Scientists like Shimony, Wigner, London and Bauer all has advocated for such connections between the human consciousness and the quantum measurement process (London & Bauer 1939). Stapp, further radicalising the view, has proposed that conscious intentions and decisions can affect the brain processes - as in, the supposed randomness of a wave function collapse in not entirely random, rather is influenced by the mental properties of the observers, provided there exist some sort of correlation between them (Stapp 2015). Concisely put, the 'intention' that guides the 'action', will affect its outcome. Needless to mention that this idea is radical and, of course, much debated. The general counterargument against these reductionist approaches criticises the lack of conclusive evidence to make such mental-material connections, rendering them as merely unproven hypotheses. Additionally, it is argued that the dimensions needed for quantum properties to deal in are too microscopic to influence (or testable, regardless) the activities inside a human brain full of noise and unfavourable atmosphere (Conte et al. 2003; Tegmark 2000).

An alternative to the reductionist/emergent view has existed parallelly where the mind and the material is separate from each other but connected via a third category, a 'background reality'. This entity is neutral from psychological and physical aspects but the sense of existence is experienced when the two correlate via this underlying third mode. Whitehead's idea of 'process philosophy' falls into this genre of thought that sees existence as a product of processes rather than substances. Different theories have appeared over the years on how this third domain relates to both mind and material (for details one should see (Atmaspacher 2014)). Most of them are beyond the scope of discussion barring two as they employ principles of quantum theory extensively. Both of these theories consider the third domain to be holistic, i.e., a collective unconscious shared by individuals not unlike entangled ensembles between system and environment described in quantum theory. Proponent of the first one is Wolfgang Pauli, well-known for the ‘exclusion principle’ explaining atomic stability. Through his correspondence with psychologist Carl Jung between 1932-1958, Pauli developed a general framework to describe the mental-physical-neutral connection. The Pauli-Jung conjecture entails that the third psychophysically neutral domain (PPN) can decompose itself in mental (M) or physical (P) domains with acausal correlations among themselves. In quantum theoretical terms, the simplified idea is given in **Fig 8.1**, where Φ represents a state in the respective domain. Φ_{PPN} not being a product state aligns with the quantum theoretical idea of non-commutativity and complementarity. Φ_M and Φ_P shares an acausal correlation ($\Phi_M \sim \Phi_P$) as they are created from the same entity Φ_{PPN} (Atmanspacher 2020). But unlike quantum mechanical correlations, which are quantitative and reproducible, Pauli-Jung conjecture prescribes the mental-physical correlations to be qualitative, subjective (‘meaningful’) and hence, irreproducible. These are termed as ‘synchronistic events’, which represents the forgotten remains of intrinsic holism existing in the neutral entity.

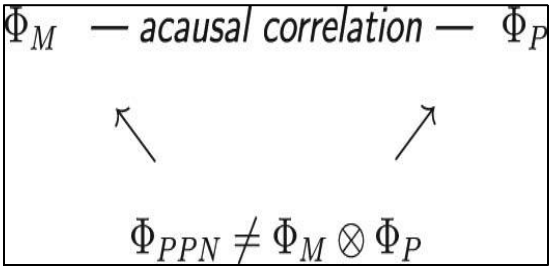


Fig.8.1: Schematic idea of Pauli-Jung conjecture (Atmanspacher 2020)

The second of the holistic and decompositional ideas came from another prominent quantum physicist David Bohm. Theory of ‘implicate and explicate orders’, later advanced by Hiley and Pylkkänen, pans toward a similar direction of Pauli and Jung but with important distinctions (Bohm 1990; Hiley 2001; Hiley & Pylkkanen 2005). Bohm’s interpretation of the undivided holistic third domain is of dynamic ‘implicate orders’ or *holomovements* from which mental and physical properties emerge as decomposed ‘explicate orders’. The implicate orders are multi-layered and capable of ordering the explicate orders (mental and physical) using ‘active information’ and pilot waves, mirroring the idea of *archetype*s in the Pauli-Jung conjecture. However, Bohm’s idea differs from the former in two crucial aspects. Here, the implicate orders/underlying reality is multi-layered, meaning a relativity component comes at play – an implicate order could turn out to be explicate with respect to a more implicate order. Such kind of relative nature of the psychophysically neutral domain was not discussed by Pauli and Jung. Also, in Pauli’s conjecture, the third domain is neither mental nor physical. But in Bohm’s interpretation, implicate orders are both mental and physical - meaning they have a fundamental knowledge of the difference of M and P even before being unfolded. In mathematical terms, this represents $\Phi_{PPN} = \Phi_M \otimes \Phi_P$, with Φ_{PPN} being a product state (Atmanspacher 2020). Bohm’s idea has found a theoretical stronghold as Hiley subsequently established it in strong mathematical framework.

Instead on decomposing the third domain into mental and physical, a school of thought championed the idea of composing the mental/physical aspects using the third domain as building blocks. The configuration of the composition decides whether it constitutes material or non-material elements. This is the ‘compositional’ view of holistic third category (also known as ‘Neutral Monism’). Although this idea is not related to or explained using quantum theory, but one of the participants, psychologist William James had an indirect but indelible contribution in the history of the microphysical theory and its influence in the mind-brain problem. Complementarity, a term introduced by James (1890), was borrowed by Niels Bohr and was used as a cornerstone in building quantum theory. The basic idea behind complementarity is that in case of certain psychological characteristics involving more than one step of judgement or decisions, context of the previous step would influence following ones. In quantum theoretical terms it translates as the results of carrying out two operations A and B would depend on their sequential operation. So, $AB \neq BA$, or in a more familiar term, non-commutativity of A and B (Atmanspacher 2012). Complementarity also entails that in a Boolean (two-valued) framework, two incompatible ideas can't be used in describing the system wholly.

Bohr's intention of extending the idea of complementarity in mind-matter discussions has paved the way for another facet of quantum mechanical applications in this field. This application doesn't intend to discuss how neurobiologically involved the quantum mechanical properties are, or whether the mental properties are an emergent offshoot of a quantum field. Instead, it restricts itself on addressing purely psychological constructs that have been discussed and studied for long: bistable perceptions, decision making (prisoner's dilemma), question order effects and more. This field of research came to be known as 'quantum cognition'. The theoretical idea is based on the assumption that the cognitive properties are describable using a 'mental state space'. Quantum cognition (although the nomenclature has varied from 'non-classical cognition' to 'non-quantum cognition', where Hilbert space structure from classical optics and field theory has been used to describe non-boolean cognitive phenomena (Ghose 2012)) has not only been used to develop a well-defined framework of addressing psychological concepts, it enjoyed surprisingly accurate empirical success too. Moreover, it has prompted several groups of researchers to work in this field all across the globe. Some of the issues that were covered under the quantum cognition canon are:

1. Bistable perception: It is a cognitive phenomenon where the viewer feels a dynamic change between two different interpretations of the same stimulus (Sterzer et al. 2009). Examples include Necker cube, Rubin's vase, Maltese cross etc. Atmanspacher and Filk gave a novel framework to explain bistable perception based on the quantum zeno effect. Their model used the idea that the 'operation' on mental states display non-commutativity. Predicted times scales were experimentally confirmed as well (Atmanspacher & Filk 2013).

2. Question order effects: This is a well-documented phenomenon routinely observed in surveys and polls that the order of asked questions tend to influence the responses (Schuman et al. 1981). It is seen that quantum probability models based on contextual features can account for this intuitively irrational cognitive effect. The model explains order effects by transforming a state vector with different sequences of operators for different orderings of information (Trueblood & Busemeyer 2011; Wang et al. 2014).

3. Decision making: The decision theory discusses the reasoning behind a rational choice. Quantum cognition models replace the classical probability with complex valued quantum probability amplitudes. This introduces interference effects which can't be accounted classically (Busemeyer et al. 2006). Furthermore, quantum probabilities helped in addressing the conjunction and disjunction fallacies as well (Pothos & Busemeyer 2009).

8.3. QUANTUM PROBABILITIES IN A NUTSHELL

Classical physics employs classical probabilities which are constrained by determinism. They assume deterministic descriptions as a universal law lying in the basis of reality. Quantum probabilities, unlike the classical counterpart, champions the factor of indeterminism as its core. According to quantum theory such probabilities are an integral part of our reality, although not always manifested perceivably in macroscopic phenomenon. The key difference that separates quantum probability with classical probabilities is a particular process – 'potentiality-actualization' (Conte et al. 2007). Usually, a *measurement* is made during a physical process to quantify any physical property. To explain such a phenomenon using classical probability methods restricts the properties into not exhibiting any potentiality. Instead, it is assumed that their values are determined well before the measurement and the outcomes are merely automatic expression of those values. Hence – 'potentiality does not exist before measurements and no effective actualization occurs during measurements'. Whereas in case of quantum probabilities, the measurement process involves potentiality-actualization, i.e., probabilities are related to the process of actualization of one potentiality among many disparate ones. In this case, context becomes an important factor to be taken into account. Using the following examples in Conte et al. (2007), the differences between classical and quantum probabilities can be demonstrated in detail.

Let us consider a variable A which is dichotomous in nature, i.e., A has two possible values '+' or positive and '–' or negative (A = +,-). A participant is asked a question that whether or not he has read a certain author, Dante Alighieri in this case. The answer to this question is either A = + (yes/positive) or A = - (no/negative); a situation analogous to that of the measurement process mentioned above. Here, the aim of the measurement is to quantify the value of variable A (A = yes or no). This is an example of a classical framework that deals with classical probabilities. The possible responses and opinions here are predetermined and the property has a definite value even before the measurement begins. In such a deterministic frame, there is no existence of potentiality or context. Hence, outcomes are also predefined and the measurement becomes merely a formality to reach automated outcomes. Now, a second example is presented which has an entirely different premise than the above. In this case, the participant is shown a photo of a woman, previously unseen by him, followed by the question: do you like this woman? Contrary to the previous example, here the participant has no established opinion and the response will be 'actualized' among two possible and potential options that will be formed at cognitive level only moments after the question will be posed. So, it was the act of measurement (posing the question) that prompted the cognitive process leading to the outcome which wasn't determined before. Unlike the prior example, measurement isn't a formality but a necessity. In this case, the situation reaches beyond classical statistical framework since the response (A = + or A = -) will depend on factors like the state of the participant at the moment and the context in which the question is posed. As the authors put it: "the answers A=+ or A = - do not pre-exist but will be formed (actualized) only at the moment of the interaction between the participant and the question starting soon the participant with the two quantum-like superposition of potentialities A = + and A = - established at his cognitive level." Such inherent indeterminism, due to the existence of potentialities which then collapses onto an actualization prompted by an act of measurement is what makes quantum probabilities unique and highly context-dependent. In some related works, authors have argued in favour of such indeterminism and contextuality lying at the core of various biological processes (Zbilut et al. 1996; Zak 1997; Conte et al. 2004). Many of the processes, instead of being perfectly periodic, are often influenced by noise fluctuations (deemed as 'context'). These contextual intervals are majorly responsible in determining the future trajectory of the system, that is, out of different possible potentialities which one would be actualized. Such noise containing system are considered to be similar to 'acting' and 'thinking' systems as they exhibit both transition from one equilibrium point to another ('acting') and decision-making under the influence of noise ('thinking'), further cementing the role of context in non-deterministic scenario (Zak et al. 2008). The processes of potentiality, actualization and contextuality is intrinsic to the cognitive domain as well. The interconnected dynamics of these acts correspond to the emergence of conceptual entities like perception and memory. The conscious experience of the environment is emerged via the contest between filtration through innate categories and actualization of learned categories, which shift fluidly depending on contextual relevance associated with it. This context dependency of the concepts such as 'democracy', 'truth', 'belief' or 'falsehood' assures that they continue to be evaluated through cognitive processes as perception or other forms of human interactions (Aerts et al. 2011; Khrennikov 2015).

The core of classical probability theory is set up on Boolean logic which is the cornerstone of 20th century scientific achievements like information theory and digital electronics. Classical logic presents us with a basic and remarkable idea called conditional probability, i.e., in case of two independent events A and B, the probability of event B occurring given the condition that event A has already occurred is given by Eq. (1):

$$P(B|A) = \frac{P(B\cap A)}{P(A)} \tag{1}$$

In its heart, eq. (1) carries intrinsic Boolean logic as it features the axiomatic Boolean operation ∩ (or intersection). This leads to one of the fundamental laws of classic probability theory, known as the Bayes' formula of total probability (FTP), given as Eq. (2):

$$P(A) = P(B_1).P(A|B_1) + \cdots + P(B_n).P(A|B_n)$$

$$= \sum_{i=1}^{n} P(B_i).P(A|B_i) \tag{2}$$

Eq. (2) describes the probability of an event A where B_1, B_2,……, B_n are mutually exclusive and exhaustive events in a sample space (also known as σ-space). This plays a major role in classical statistical theory and decision theory. On the other hand, quantum mechanical theory, and by extension the quantum probabilities are governed by Born's rule (Khrennikov 2015):

$$p(x) = |\psi(x)|^2 \tag{3}$$

Here, p(x) is the probability to find a particle at point x and ψ(x) is the complex wave function of the particle.

Quantum theory acknowledges wave nature of microscopic particles and hence, bears some unique attributes in the probability structure which are absent in classical framework, like - interference of probability functions, analogous to constructive/destructive interferences in the double slit pattern. Von Neumann, most notably, compared the classical and quantum probability structures and noted that it was impossible to reproduce the quantum properties on a classical probability space (via hidden variables) (Von Neumann 2013). Quantum probabilistic framework, hence, introduces new concepts of interference, complementarity and entanglement which violates various rules of classical Boolean logic such as - Commutative law, Distributive law, FTP - all these classic fundamental laws are either violated or relaxed at the advent of quantum logic. Many rudimentary principles of quantum theory were deeply imbued in the psychological and cognitive intuitions since the days of its inception. Therefore, it is natural to try to use the mathematical foundations of quantum theory (mainly, quantum probabilities) to tackle problems of psychology, cognitive science, social science, economics and many other fields which tend to overlap with probabilistic nature of human judgement and decision making (Khrennikov 2014, 2015; Aerts & Aerts 1995; Busemeyer & Bruza 2012). Out of such attempts, the studies exploring non-classical effects in regards to visual ambiguity are of particular interest for the present work. Using the idea that FTP violation during cognitive task can be an indicator of non-classicality (Khrennikov 2014), experiments have been made in recent past to test context-dependency in recognizing ambiguous figures. Curiously enough, evidences of FTP violation were found in the statistical data. Moreover, the presence of an 'interference term' of complex nature suggests the involvement of quantum-like behaviour in visual perception, highly sensitive to contextual influence (Conte et al. 2007, 2008, 2009).

8.4. MATHEMATICAL APPROACH TO THE EXPERIMENT

Following the arguments from the previous section, the violation of total probability can be deemed as a marker of experimentally verifying the existence of non-classicality in cognitive domain. The method to translate the theoretical ideas into empirical evidence is given in following manner: First, two dichotomous observables (questions, in this case) A and B are prepared which can assume two potentialities or answers, i.e., + ('yes') and – ('no'). Then, the questions are posed to a large ensemble of cognitive systems (here, human subjects), denoted with U. Now, as observable A is operated on the ensemble U, the subsequent probabilities emerge – P(A+) and P(A-). The former is probability of the outcome being A = + or yes and the latter gives the same for A = - or no. It is evident that

$$P(A+) + P(A-) = 1 \tag{4}$$

Similar results can be obtained using observable B on U.

$$P(B+) + P(B-) = 1 \tag{5}$$

Next, the larger ensemble U is divided equally into two sub-ensembles U_1 and U_2. The first ensemble, U_1, is asked to respond to question A only and the corresponding probabilities P(A+) and P(A-) are generated. Ensemble U_2, however, is asked to respond to question B first, immediately followed by question A. This produces two pairs of conditional probabilities P(A+ | B±) and P(A- | B±), which indicate the probability of obtaining A+ or A- having obtained B+ or B- already. Now, using eq. (2) from previous section, the formula of total probability is computed. Theoretically, if the classical logic is followed in human cognitive sphere, FTP should not be violated from the acquired statistical data. But, in presence of non-classicality (and quantum probabilities), an interference-like term is spawned into existence to account for the missing data from the unbalanced equation of total probability. The equation is modified as (Conte et al. 2007):

$$P(A\pm) = P(B+).P(A\pm|B+) + P(B-).P(A\pm|B-) + 2\sqrt{P(B+).P(A\pm|B+)P(B-).P(A\pm|B-)}\cos\theta(\pm) \tag{6}$$

Here, θ is phase angle and cos $\theta(\pm)$ represents the interference co-efficient which indicates the incompatibility of different contexts (Conte et al. 2009). The systems that follow eq. (6) are said to exhibit contextuality and non-classical (quantum-like) behaviour.

8.5. EXPERIMENTAL DETAILS

8.5.1. Auditory Stimulus Used

Almost entirety of the works regarding classicality violation in cognitive domain deals with visual stimulus. The basics of it lies with a principle of psychology originated in 1920s called the Gestalt psychology that identifies the human perceptual process as holistic – that is to say, any complex stimulus is perceived as a whole (or groups of whole, 'Gestalten') which includes both the primary object and its relations to the surrounding - instead of classical reductionist approach that stimulus is a sum of elementary features that are singularly necessary and jointly sufficient enough to be perceivable (Ellis 2013). When one sees a visual stimulus, not only they see the object or its shape and form, but its relation to the surrounding space as well. Sometimes the stimulus could feel kinetic even though it is fixed in a two-dimensional space. One of the major contributions of Gestalt psychology is to underline such latent contextuality and related ambiguity in perception. The rationale for using this interpretation of psychology in the topic at hand is that the manifested outcome is probabilistic due to the subjective experience of the stimuli. Now, unlike vision, which is spatial and two/three-dimensional, auditory stimulus is only temporal. Despite, it provides a complex scene involving multiple sources of human speech, vocal and instrumental music, animal sounds and other nature noises, occasionally all occurring at the same time, each with its own sub-phrasing and structure. Therefore, even though Gestalt principles like proximity, similarity, closure are predominantly used in visual domain, auditory Gestaltens can also be created (Bregman 1994; Kubovy & Van Valkenburg 2001). For example, silence or background noise, interrupted by a loud sound, followed again by silence or noise, is an auditory analogue of a figure on a ground. Likewise, a regular series of identical short clicks interspersed by equal/unequal time interval or pairs of identical sounds separated by equal intervals but coming from different directions (analogue to visual proximity principle), equal-interval grouping of pairs of auditory samples which are similar in auditory attributes like amplitude, pitch or timbre (analogue to visual similarity principle) are some of the examples of auditory analogues to visual Gestaltens (Todorovic 2008).

In the present work an attempt was made to test whether the non-classical nature of cognition holds true for musical cognition as well. Music is a complex acoustic stimulus with numerous sub-components like pitch, rhythm, amplitude, timbre etc. Out of these, role of rhythm in the acoustic processing is undeniable (Kotz et al. 2018). It is also well known that musical rhythm (and subsequently, the pace of it, i.e., tempo) is innate and important in the evolution of music from both personal and societal context (Winkler et al. 2009; Leongómez et al. 2022). In view of this, the component tempo was chosen to be the component of exploration for our study. As the work aims to deal with musical tempo, it was necessary to create an acoustic sample that will offer an auditory scene involving context as well as sufficient ambiguity. The sample used consisted of two music clips of 30 seconds stitched together with a silent rest period of 5 seconds in between. These two stitched clips have the same value of tempo, i.e., beats per minute or BPM but all the other musical elements like melody or timbe were kept dissimilar. Additionally, the musical arrangement of the two clips conveyed entirely different emotions (confirmed via listening tests beforehand). This provides a holistic auditory scene where the object of perception is amalgamated with its surrounding. The resting period in between provides the necessary backdrop/ground on which the stimulus could be presented on and at the same time, not long enough that the perceptual residue to the first clip can fade away.

8.5.2. Methodologies

100 participants from various age groups, ranging from 18 to 67 (38 females, mean = 23.76, Standard Deviation = ±9.47) took part in the study. Among them, 4 participants' data had to be discarded as they faced technical issues while listening to the clips. Rest of them were divided into two equal groups (Group 1 and 2). Group 1 was exposed to a music clip as auditory stimuli (clip A) for 65 seconds, which included two 30 sec. music pieces with an interval of 5 sec of silence. Each of the pieces had the same BPM (=72). After that, participants were asked whether the two clips they heard had the same tempo or not. The answer 'yes' to the posed question on clip A gave P(A+), and 'no', gave P(A-). The other group, Group 2, were made to listen to another similarly prepared clip (clip B, where both pieces had 130 beats per minute) first, followed by the clip A (the resting period in between was 60 seconds). Responses of this group gave the remaining sets of probabilities, i.e., P(B+), P(B-), P(A+|B+), P(A-|B+), P(A-|B+), P(A-|B-). The participants belonged to various age groups and of various music tastes. All had normal hearing capabilities. To eliminate musician/non-musician bias, they were asked not to use any means of tempo measurements, including finger/foot-tapping. Finally, in all the cases, to avoid the risk to influence the participants, the question to be asked by tests was posed in the most neutral form.

8.6. RESULTS

Table 8.1 illustrates the results of the computed probabilities. The difference of P(A+) between the one calculated only from test A and the one from test B followed by A is evident and it its value happens to be $|\Delta P(A\pm)| = 0.0626$, as given in **Fig. 8.2**. This difference is completely unaccounted for in classical logic (Eq. 2) and it clearly indicates that non-classicality is indeed present in the cognition of musical tempo. Now, statistical test needs to be performed in order to confirm the resulting difference. For a collection of samples, if the measured difference in their averages is large, null hypothesis would get rejected as such differences could occur due to chance. Student's T-Test could be applied in such cases. In this case, the obtained two-tailed P value equals 0.0352 ($t = 5.1855$, $df = 2$, Standard error = 0.044). By conventional criteria, this difference is considered to be statistically significant, that is, if the null hypothesis were true, then 97% of experiments would lead to a difference smaller than the one observed, and 3% of experiments would lead to a difference as large or larger. Finally, following eq. (6), cos $\theta(\pm)$ are computed to indicate the trigonometric measure of transformation of probabilities that are generated by transitioning from one context to another (from test A to test A|B) (Conte et al. 2007). The higher the value of θ (the phase angle), the stronger the role of context in probability transforming. From Table 8.1, it is seen that $\theta(+) = 1.518$ and $\theta(-) = 1.609$, both of which are satisfactory enough to suggest the influence of context as well as the non-classical quantum-like nature of cognition in the present case.

Probabilities	Values obtained
P(A+)	0.4167
P(A-)	0.5833
P(B+)	0.2917
P(B-)	0.7083
P(A+ | B+)	0.3571
P(A+ | B-)	0.3529
P(A- | B+)	0.6429
P(A- | B-)	0.6471
P(A+) = P(B+).P(A+ | B+) + P(B-).P(A+ | B-)	0.3541
P(A-) = P(B+).P(A- | B+) + P(B-).P(A- | B-)	0.6458
ΔP(A+)	0.0626
ΔP(A-)	-0.0626
cos θ(+)	0.0526
θ(+)	1.518
cos θ(-)	-0.0389
θ(-)	1.609

Table 8.1: Obtained values of the probabilities, computed from participants' choices (+ or -) to the posed question. Here, ΔP(A±) = [P(A±) from only test A - P(A±) from test B, followed by test A], cosθ(±) gives a trigonometric measure of probability transformation where θ is the phase angle, denoting context-dependency

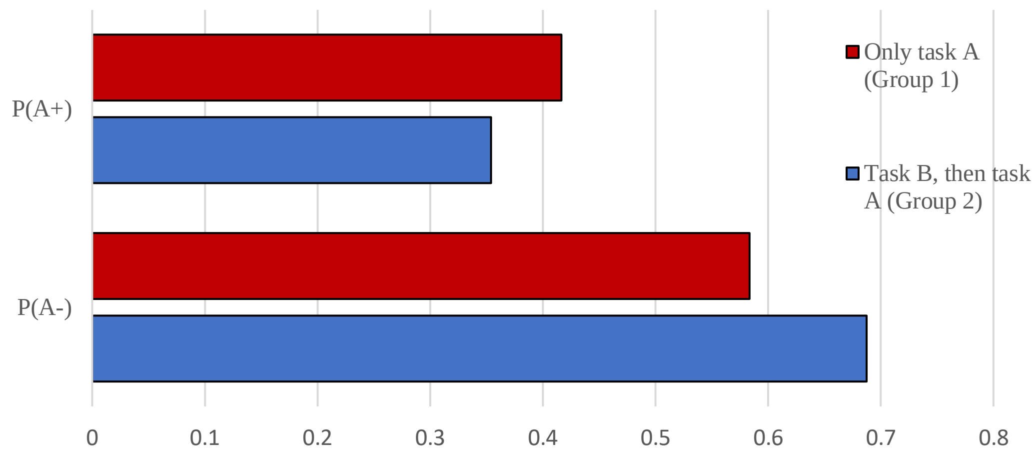


Fig. 8.2: Difference in values of obtained probabilities P(A±) in two experimental scenarios. The probabilities of A± computed from two different tasks clearly indicate towards the contextual influence of task B in the second set-up, thus violating the rules of total probability and subsequently, classical logic.

8.7. DISCUSSION

The present study yields result which indicate towards the violation of classical probability laws in case of human auditory experience. The experiment sheds light on the fact that cognitive systems can behave in the quantum-like way producing nonzero coefficients of interference. The observations are in agreement with previous reports using visual stimuli. The intriguing results found in this work strongly suggest the presence of non-classicality in the perception of musical tempo. Moreover, with the emergence of quantum-like components it can be argued that quantum mechanics do have a role to play in the dynamics of stimulus perception.

However, we consciously shy away from proposing that this non-boolean event, or any such, is a precursor of a quantum brain idea. This is because a quantum-like framework in explaining a cognitive phenomenon doesn't necessarily have to establish any link to quantum structures in neurobiology. The basis of the former revolves around the state space description. But, the physical interpretation of such a state remains as elusive and debated as it was decades ago. Ideally, the higher brain functions or high-level brain features should be describable from the lower-level descriptions, with some stability conditions applied preferably that can explain the sustainability from transient lower to the stable higher. Yet we see that the latter is necessary but not sufficient to address the former. The lack of continuity and robustness in constructing the state space description can often lead to incompatibility. Hence, it is difficult to argue, at this point, that quantum-like properties are a definite consequence of a quantum brain. Additionally, there seems to be argument against it on the basis of existing Hilbert space notion in classical physics. In classical optics, the idea of nonquantum entanglement is not alien. The principle of superposition is responsible for the optical interference events where the phase of the wave makes the corresponding wave function complex-valued. Phenomena such as violation of Bell inequalities, usually associated with quantum physics, have also been reported in optical polarization (Spreeuw 2001; Simon B. N. et al. 2010). Considering these arguments, we only restrict ourselves to describe the study as a unique demonstration of non-classical effects in auditory cognition. The underlying reasons for this, however, need more investigations to be ascertained.

Historically, cognitive and quantum entities share a lot of commonalities and the long-standing dream of several researchers of finding a bridge between the two might not be far-fetched. The resulting inconsistency in the probability data may lead to the agreement on a modified nondeterministic approach over the classical probability model as it fails to recognize concepts like contextuality or interference and employs a highly reductionist approach. In addition, context envelopes used in this experiment produced the coefficients of interference that, in turn, provided a quantification of the incompatibility between those contexts. It certainly won't be out of line to cautiously venture into a prediction that more and more aspects of cognitive properties could exhibit such context-dependency and eventual non-classical behaviour. Of course, more such cases of theoretical and experimental work are needed to establish this connection in a conclusive manner. However, this preliminary work hopes to start the proverbial ball rolling in the field of auditory (especially musical) perception.

REFERENCES

Aerts, D. (2014). Quantum theory and human perception of the macro-world. *Frontiers in Psychology*, *5*, 554.

Aerts, D., & Aerts, S. (1995). Applications of quantum statistics in psychological studies of decision processes. *Foundations of Science*, *1*(1), 85-97.

Aerts, D., Broekaert, J., & Gabora, L. (2011). A case for applying an abstracted quantum formalism to cognition. *New Ideas in Psychology*, *29*(2), 136-146.

Amit, D. J., & Amit, D. J. (1992). *Modeling brain function: The world of attractor neural networks*. Cambridge university press.

Atmanspacher, H. (2012). Dual-aspect monism à la Pauli and Jung. *Journal of Consciousness Studies*, *19*(9-10), 96-120.

Atmanspacher, H. (2020). The Pauli–Jung Conjecture and Its Relatives: A Formally Augmented Outline. *Open Philosophy*, *3*(1), 527-549.

Atmanspacher, H., & Filk, T. (2013). The Necker–Zeno model for bistable perception. *Topics in cognitive science, 5*(4), 800-817.

Atmaspacher, H. (2014). 20th century variants of dual-aspect thinking. *Mind and Matter*, *12*(2), 245-288.

Bohm, D. (1990). A new theory of the relationship of mind and matter. *Philosophical psychology*, *3*(2-3), 271-286.

Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound*. MIT press.

Busemeyer, J. R., & Bruza, P. D. (2012). *Quantum models of cognition and decision*. Cambridge University Press.

Busemeyer, J. R., Wang, Z., & Townsend, J. T. (2006). Quantum dynamics of human decision-making. *Journal of Mathematical Psychology*, *50*(3), 220-241.

Conte, E., Federici, A., & Zbilut, J. P. (2004). On a simple case of possible non-deterministic chaotic behavior in compartment theory of biological observables. *Chaos, Solitons & Fractals*, *22*(2), 277-284.

Conte, E., Khrennikov, A. Y., Todarello, O., Federici, A., Mendolicchio, L., & Zbilut, J. P. (2009). Mental states follow quantum mechanics during perception and cognition of ambiguous figures. *Open Systems & Information Dynamics*, *16*(01), 85-100.

Conte, E., Khrennikov, A., Todarello, O., Federici, A., & Zbilut, J. P. (2008). A preliminary experimental verification on the possibility of Bell inequality violation in mental states. *NeuroQuantology*, *6*(3).

Conte, E., Todarello, O., Federici, A., Vitiello, F., Lopane, M., & Khrennikov, A. (2003). A preliminar evidence of quantum like behavior in measurements of mental states. *arXiv preprint quant-ph/0307201*.

Conte, E., Todarello, O., Federici, A., Vitiello, F., Lopane, M., Khrennikov, A., & Zbilut, J. P. (2007). Some remarks on an experiment suggesting quantum-like behavior of cognitive entities and formulation of an abstract quantum mechanical formalism to describe cognitive entity and its dynamics. *Chaos, Solitons & Fractals*, *31*(5), 1076-1088.

Ellis, W. D. (2013). *A source book of Gestalt psychology*. Routledge.

Ghose, P. (2012). Nonquantum Cognition. *arXiv preprint arXiv:1210.4473*.

Hameroff, S. (2003). Consciousness, Whitehead and quantum computation in the brain: Panprotopsychism meets the physics of fundamental spacetime geometry. *Searching for New Contrasts. Whiteheadian Contributions to Contemporary Challenges in Neurophysiology, psychology, Psychotherapy and the Philosophy of Mind New York: Lang*, 61-86.

Hameroff, S. R., & Watt, R. C. (1982). Information processing in microtubules. *Journal of Theoretical Biology*, *98*(4), 549-561.

Hampton, J. A. (1988). Disjunction of natural concepts. *Memory & cognition*, *16*(6), 579-591.

Harvey, R. L. (1994). *Neural network principles*. Prentice-Hall, Inc.

Heisenberg, W., & Northrop, F. S. C. (1958). *Physics and philosophy: The revolution in modern science* (p. 256). New York: Harper.

Hiley, B. J. (2001, June). Non-commutative geometry, the Bohm interpretation and the mind-matter relationship. In *AIP Conference Proceedings* (Vol. 573, No. 1, pp. 77-88). American Institute of Physics.

Hiley, B. J., & Pylkkanen, P. (2005). Can mind affect matter via active information?. *Mind and Matter*, *3*(2), 8-27.

Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the national academy of sciences*, *81*(10), 3088-3092.

James, W. (1890). The principles of psychology (vol. 1) New York: Henry Holt & Co. *Inc. http://dx. doi. org/10.1037/11059-000*.

Jedlicka, P. (2017). Revisiting the quantum brain hypothesis: toward quantum (neuro) biology?. *Frontiers in molecular neuroscience*, *10*, 366. Whitehead, A. N., & Sherburne, D. W. (1957). Process and reality. New York, NY: Macmillan.

Khrennikov, A. (2015). Quantum-like modeling of cognition. *Frontiers in Physics*, *3*, 77.

Khrennikov, A. Y. (2014). *Ubiquitous quantum structure*. Berlin: Springer.

Kim, J. (1998). *Mind in a physical world: An essay on the mind-body problem and mental causation*. MIT press.

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The evolution of rhythm processing. *Trends in cognitive sciences*, *22*(10), 896-910.

Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition*, *80*(1-2), 97-126.

Leongómez, J. D., Havlíček, J., & Roberts, S. C. (2022). Musicality in human vocal communication: an evolutionary perspective. *Philosophical Transactions of the Royal Society B*, *377*(1841), 20200391.

Loewenstein, W. (2013). *Physics in mind: a quantum view of the brain*. Basic Books.

London, F., & Bauer, E. (1939). *La théorie de l'observation en mécanique quantique* (Vol. 775). Hermann & Cie.

Marshall, I. N. (1989). Consciousness and Bose-Einstein condensates. *New ideas in Psychology*, *7*(1), 73-83.

Mermin, N. D. (1985). Is the moon there when nobody looks? Reality and the quantum theory. *Physics today*, *38*(4), 38-47.

Mézard, M., Parisi, G., & Virasoro, M. A. (1987). *Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications* (Vol. 9). World Scientific Publishing Company.

Orlov, Y. F. (1982). The wave logic of consciousness: A hypothesis. *International journal of theoretical Physics*, *21*(1), 37-53.

Osherson, D. N., & Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition*, *9*(1), 35-58.

Penrose, R. (1989). The emperor’s new mind Oxford University Press. *New York*.

Penrose, R., Shimony, A., Cartwright, N., & Hawking, S. (2000). *The large, the small and the human mind*. Cambridge University Press.

Pothos, E. M., & Busemeyer, J. R. (2009). A quantum probability explanation for violations of ‘rational’decision theory. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1665), 2171-2178.

Ricciardi, L. M., & Umezawa, H. (1967). Brain and physics of many-body problems. *Kybernetik*, *4*(2), 44-48.

Rosch, E. H. (1973). Natural categories. *Cognitive psychology*, *4*(3), 328-350.

Rosu, H. C. (1994). Essay on mesoscopic and quantum brain. *arXiv preprint gr-qc/9409007*.

Schuman, H., Presser, S., & Ludwig, J. (1981). Context effects on survey responses to questions about abortion. *Public Opinion Quarterly*, *45*(2), 216-223.

Schwartz, J. M., Stapp, H. P., & Beauregard, M. (2005). Quantum physics in neuroscience and psychology: a neurophysical model of mind–brain interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1458), 1309-1327.

Shimony, A. (1993). *The Search for a Naturalistic World View: Volume 1* (Vol. 1). Cambridge University Press.

Simon, B. N., Simon, S., Gori, F., Santarsiero, M., Borghi, R., Mukunda, N., & Simon, R. (2010). Nonquantum entanglement resolves a basic issue in polarization optics. *Physical review letters*, *104*(2), 023901.

Sozzo, S. (2017). Effectiveness of the quantum-mechanical formalism in cognitive modeling. *Soft Computing*, *21*(6), 1455-1465.

Spreeuw, R. J. (2001). Classical wave-optics analogy of quantum-information processing. *Physical Review A*, *63*(6), 062302.

Stapp, H. P. (2004). Mind, matter, and quantum mechanics. In *Mind, matter and quantum mechanics* (pp. 81-118). Springer, Berlin, Heidelberg.

Stapp, H. P. (2015). A quantum-mechanical theory of the mind-brain connection. *Beyond Physicalism*, 157-193.

Sterzer, P., Kleinschmidt, A., & Rees, G. (2009). The neural bases of multistable perception. *Trends in cognitive sciences*, *13*(7), 310-318.

Tegmark, M. (2000). Importance of quantum decoherence in brain processes. *Physical review E*, *61*(4), 4194.

Todorovic, D. (2008). Gestalt principles. *Scholarpedia*, *3*(12), 5345.

Trueblood, J. S., & Busemeyer, J. R. (2011). A quantum probability account of order effects in inference. *Cognitive science*, *35*(8), 1518-1552.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological review*, *90*(4), 293.

Tversky, A., & Shafir, E. (1992). The disjunction effect in choice under uncertainty. *Psychological science*, *3*(5), 305-310.

Vitiello, G. (1995). Dissipation and memory capacity in the quantum brain model. *International Journal of Modern Physics B*, *9*(08), 973-989.

Von Neumann, J. (2013). *Mathematische grundlagen der quantenmechanik* (Vol. 38). Springer-Verlag.

Walker, E. H. (1970). The nature of consciousness. *Mathematical Biosciences*, *7*(1-2), 131-178.

Wang, Z., Solloway, T., Shiffrin, R. M., & Busemeyer, J. R. (2014). Context effects produced by question orders reveal quantum nature of human judgments. *Proceedings of the National Academy of Sciences*, *111*(26), 9431-9436.

Whitehead, A. N., & Sherburne, D. W. (1957). *Process and reality*. New York, NY: Macmillan.

Wigner, E. P. (1963). The problem of measurement. *American Journal of Physics*, *31*(1), 6-15.

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences*, *106*(7), 2468-2471.

Zak, M. (1997). Dynamical simulations of probabilities. *Chaos, Solitons & Fractals*, *8*(5), 793-804.

Zak, M., Zbilut, J. P., & Meyers, R. E. (2008). *From instability to intelligence: complexity and predictability in nonlinear dynamics* (Vol. 49). Springer Science & Business Media.

Zbilut, J. P., Hübler, A., & Webber, C. L. (1996). Physiological singularities modeled by nondeterministic equations of motion and the effect of noise. In *Fluctuations and order* (pp. 397-417). Springer, New York, NY.

Zohar, D., & Marshall, I. N. (1990). *The quantum self: Human nature and consciousness defined by the new physics*. William Morrow & Co.

CHAPTER 9

AMBIGUITY INVOKES CREATIVITY: LOOKING THROUGH QUANTUM PHYSICS

"The future is uncertain... but this uncertainty is at the very heart of human creativity."

Ilya Prigogine

ABSTRACT

Creativity, defined as 'the tendency to generate or recognize new ideas or alternatives and to make connections between seemingly unrelated phenomena', is too vast a horizon to be summed up in such a simple sentence. The extreme abstractness of creativity makes it harder to quantify in its entirety. Yet, a lot of efforts have been made both by psychologists and neurobiologists to identify its signature. A general conformity is expressed in the 'Free association theory', i.e. the more freely a person's conceptual 'node's are connected, the more divergent thinker (also, creative) he or she is. Also, tolerance of ambiguity is found to be related to divergent thinking. In this study, we approach the problem of creativity from a theoretical physics standpoint. Theoretically, for the initial conceptual state, the next 'jump' to any other node is equally probable and non-deterministic. And to study such a probabilistic system, Quantum theory has been proven the most successful, time and again. We suggest that this collection of nodes form a system which is likely to be governed by quantum physics and specify the transformations which could help explain the conceptual jump between states. Our argument, from the point of view of physics is that the initial evolution of the 'creative process' is identical, person or field independent. To answer the next obvious question about individual creativity, we hypothesize that the quantum system, under continuous measurements (in the form of external stimuli) evolves with chaotic dynamics, hence separating a painter from a musician.

Keywords: Creativity, divergent thinking, quantum physics, ambiguity, Wigner function

9.1. INTRODUCTION

"Others have seen what is and asked why. I have seen what could be and asked why not." ― Pablo Picasso.

From the very dawn of civilization, Creativity has been inspiring and reshaping human existence continuously. It was creativity that gave rise to the likes of Picasso, da Vinci, and Einstein - who, with their endless wonders on and off the paper, changed the course of human history and civilization time and again. We,in turn, have wondered why these stalwarts possess such unique dispositions and how it separates them from a 'not-so-creative' individual. Lexically put, Creativity is the development of new ideas and original products in a novel and appropriate way (Sternberg & Lubert 1991, 1995; Lubert 1994; Ochse & Ochse 1990). And theories and ideas about understanding the creative process stem from far back in history since it is a particularly human characteristic (Ryhammer & Brolin 1999).

Though it started as far back as late 1800s, the systemic search of creativity blossomed in the twentieth century, where its roots have been searched in the lights of a plethora of diversified disciplines (Craft 2001):

- **Psychoanalytic approach**: This school of thought, put forward by Sigmund Freud, argues that the act of creativity is a socially acceptable defense mechanism against the morally unaccepted sexual desire of the unconscious depth of human mind (known as 'id'). The defense mechanism, termed as Sublimation, manifests the unconscious fantasy into conscious cultural diversities via the creative process (Freud 1958, 1959). Later, Donald Winnicott, in his seminal works on psychoanalysis, claims Creativity to be an intrinsic part of human nature (Winnicott 1990, 1991). He defines a 'positional space' in the early developmental years: which embeds a unique paradox of experience in the infant. This experiential paradox is a part of his nature afterwards and this is where the creative process blooms throughout the human life.

- **Cognitive approach:** Originated from Galton's work on hereditary genius (Galton 1870). In the latter half of the nineteenth century, study of creativity started with the study on the origins of '*genius*'. Sir Francis Galton, himself a cousin of Sir Charles Darwin, systematically studied family histories of eminent high-ranking people of society and deduced the idea that 'genius' is hereditary, i.e., most of the successful professionals have high percentage of other successful professionals as their immediate relatives. Although this claim was unsupported by studies later, it did inspire a lot of research into the origins of creativity. Other cognitive approaches include Mednick's associative process idea (1962) where he states: "..the forming of associative elements into new combinations which either meet specific requirements or are in some way useful. The more mutually remote the elements of the new combination, the more creative the process or solution." In a similar line of thought, J.P. Guilford explored the Creativity problem with Divergent Thinking with emphasis on Fluency, Flexibility, Originality and Elaboration, essential components for the divergent production of thoughts (Guilford 1967a, 1967b).

- **Behaviourist approach:** J.B Watson (one of the early developers of the behaviourist psychology to refute the psychoanalytic abstractness and focus only on the observable behaviourial manifestation) believed that the environment conditions one's behaviour patterns and this conditioning stays latent in his subconscious (Watson 1913). B.F. Skinner later studied the effect of applying/withdrawing rewards on the behavioural patterns, famously known as Operant Conditioning. He has claimed that the Operant Conditioning, coupled with the subconscious memories is the driving force behind the creative act ('behaviourial mutations') (Skinner 1965, 1990).

- **Neurological/Biological approach:** These kinds of studies include using modern instruments like EEG or fMRI to pinpoint the brain areas activated during creativity, both in spatial and/or temporal scale. Blossomed with the advent of highly accurate scientific bio-sensors in the recent past, this approach has allured psychologists and neurobiologists alike. Although the pursuit found many takers, as of yet not much conformity has been found among them (Dietrich & Kanso 2010; Abraham 2012).

The latter half of the 20^{th} century of creativity research is dominated by Psychometrics. Psychometric approaches to creativity were begun by psychologist J.P. Guilford, who developed a tool for measuring the extent of divergent thinking, which he later developed into the concept of 'divergent production' (Guilford 1966, 1967b). Divergent production (or thinking), also loosely called 'lateral thinking', is a method used to generate multiple related ideas for a given topic or a problem. Despite criticism, the idea of divergent thinking has become important in the scientific study of creativity because many widely used tests for creativity are measures of individual differences in divergent thinking ability like the Torrance tests of creative thinking (Torrance 1972, 1988).

9.2. FREE ASSOCIATION THEORY OF CREATIVITY

Divergent thinking tasks have been widely used because traditionally creativity has been understood in terms of the accessibility of concepts in our long-term memory systems. Concepts are connected in our brains in 'semantic networks'. Fig. 9.1 presents a sample schematic of a semantic network, with each concept 'node' of the network accessible from the concept 'street' via other nodes.

Psychologists have proposed that individual differences in creativity are due to differences in whether these kinds of associative networks were 'steep' or 'flat' – those with 'flat' networks have numerous and loose conceptual connections, enabling them to be more creative. Those with 'steep' networks tend to have more logical, linear associations between nodes (Mednick 1962; Freedman 1965).

The idea of free association is said to be the core of the divergent thinking process (Benedek et al. 2012). But in the recent past, association tests have been claimed to be less influential in determining creative abilities. The main argument against it is that they require convergent thinking instead of divergent, i.e., it follows a very systematic path of deductive reasoning to arrive at a single conclusion, instead of traversing a diverging array for the possible solutions (Brophy 2001, Lee & Therriault 2013). Fig. 9.2 illustrates the basic difference between divergent and convergent thinking.

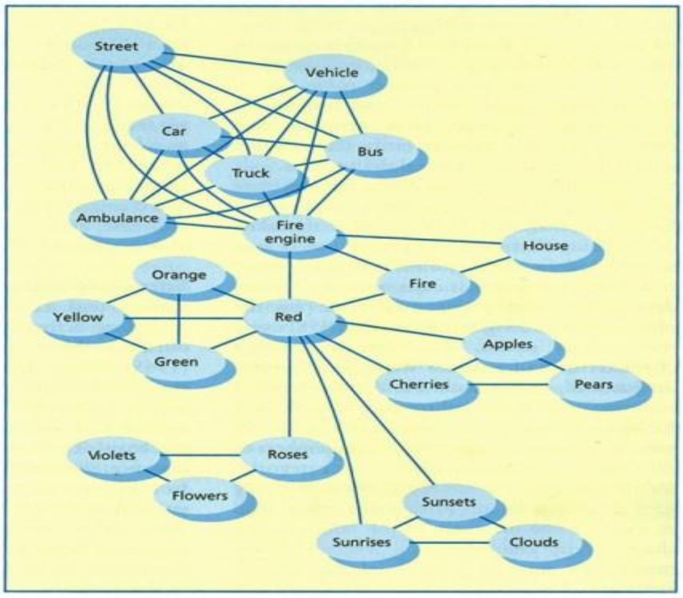


Fig. 9.1. Schematic diagram of a semantic network. Each concept node is connected to each other via other nodes. (Collins & Loftus 1975)

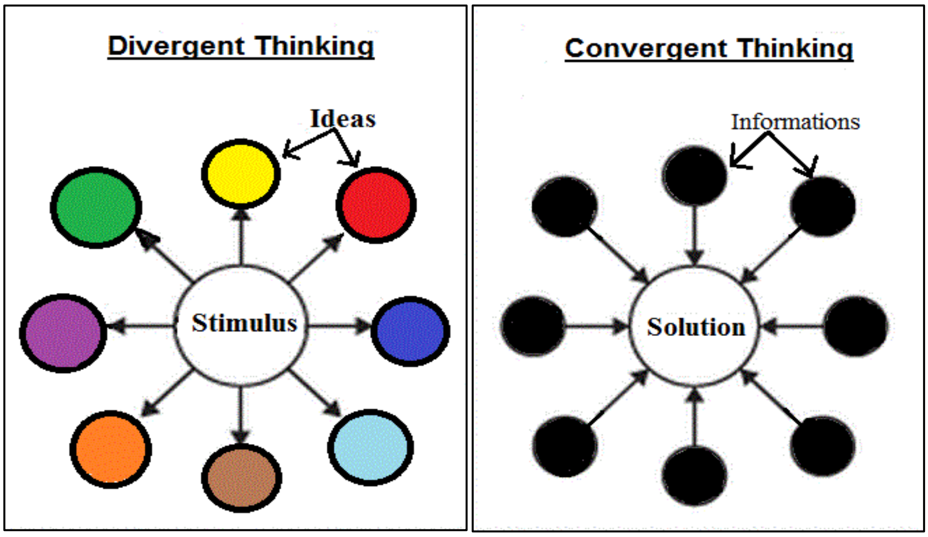


Fig. 9.2. Schematic diagram of Divergent & Convergent Thinking patterns. Unlike divergent thinking, convergent process follows a systematic deductive reasoning path.

But, comparing associative tests with other convergent as well as divergent tests show that divergent measures are significantly higher (Taft & Rossiter 1966). Smith et al. (2013) has shown that association-based test initiates from a divergent thinking process before evolving into convergent ones. Considering merits and demerits of both the perspectives, we can see that associative tests or ideas could be responsible for both the processes, hence crediting it for one while discrediting it for the other would definitely be an error in judgment (Marron & Faust, 2018).

In this paper, however, we investigate the idea of divergent thinking from a theoretical point of view of Physics, more specifically, Quantum physics.

9.3. WHY USE PHYSICS? AND WHY QUANTUM PHYSICS?

Answer to the first question is quite obvious. Human brains, similar to any other physical system, are made up of atoms- joined together to form molecules, bounded into specialized cells, i.e., neurons, which can communicate with each other. All the other parts and functions of brain- including neurotransmitters, synaptic conductance of electrical impulse- are governed by physical laws describable using Physics. Hence, like everything else in this world, neurons and brains must obey the laws of physics.

Now, to address the second part of the question: human cognition, and other higher brain functions like creativity, is intuitively macroscopic in nature. This lies within the realm of day-to-day human behavior to such an extent that drawing a connection between it and quantum mechanics, a highly successful theory devised mainly to explain microscopic phenomena, may sound wildly far-fetched. Yet, there are good scientific reasons to do so (Busemeyer & Bruza, 2012). Quantum theory is the best experimentally confirmed scientific theory in the history of physics. It is essential to every natural physical process and their practical applications, starting from microscopic explanations like stability of the atomic structure, to macroscopic applications such as Laser. It has further paved the way for new groundbreaking ideas like superconductivity and superfluidity. Unlike classical physics, which is limited by determinateness and objectivity, Quantum theory possesses unique properties that make it enable to tackle the problems regarding cognitive functions which appear to exceed the known nature of classical laws. And the fundamental ideas that weave these specialties into the theoretical base of Quantum theory have been inspired by psychological concepts themselves, hence relating the two since the inception. Subsequently, it has the potential to provide coherent and mathematically principled explanations for the puzzles and challenges in human cognitive research. A very brief overview of the unusual nature of the theory is presented as a precursor to the idea we want to convey.

- **Discreteness of nature:** Quantum physics emphasizes that our world is built on discrete particles that are bound in finite systems of discontinuous energies (Rechenberg 1982). Unlike classical physics, where energies have continuous distributions.

  Also in the neurological framework, to describe the dynamics of neuron firings - evidently a discrete and discontinuous process - quantum theory can be used for more precision.

- **Wave-particle dual nature:** Particles, like electron or photon, can exhibit both- particle and wave- characteristics, an event that is entirely non-classical in nature (Eisberg & Resnick, 1985).

  Proposal of the manifestation of this property in mind-matter relationship is not an alien notion (Bohm 1990).

- **Quantum tunneling:** Quantum wave effects allow tunneling through an energy barrier which would classically be insurmountable (Razavy 2003). It is suggested that macroscopic spread of quantum effects in human brain may involve the tunneling effect (Hameroff 1998).

- **Quantum superposition:** Before a measurement, a particle can be in a state which is a superposition of all the possible energy configurations available for the particle (Monroe et al. 2002). Dealing with a mental space which consists of several possible states, the role of quantum superposition and quantum probability amplitudes is surely undeniable.

- **Non-deterministic nature:** Quantum process is non-deterministic (i.e., the process of measurement introduces non-determinism). Unlike classical counterpart, only knowing the initial state of a system is not enough to predict its future course of evolution (Braginsky et al. 1995). Neurological investigation hasn't been able to determine the exact state of the psychological functions in any sort of experiments. Limitations of the deterministic ideas definitely point towards a quantum intervention.
- **Quantum entanglement:** Entanglement is the inseparable quantum correlation of two or more particles or degrees of freedom which determines the states of these two spatially separated systems simultaneously as soon as one of them is acted or measured upon (Schrödinger 1935). Entanglement effects in human brain have been of a topic of several researches, including memory (Thiagarajan et al. 2010) or cognitive processes (Fisher 2015).

These are some of the unique phenomena that marks the stark differences between quantum and classical physics, making the former more suitable to approach the issues of explaining a number of brain and cognitive functioning, including Creativity (Arndt et al. 2009).

9.4. QUANTUM LEAP INTERPRETATION

Before we explain the hypothesis, let us steer the reader towards the path we hope to take. From the viewpoint of cognitive neurobiology, we now well understand the nature of nerve cell activity: the creation of action potentials, ion exchange, the use of energy, axonal transport, the vesicle cycle, and the production, cycle and breakdown of neurotransmitters. Yet, how these unconscious materials produce the stream of consciousness or perform the individual higher brain functions so smoothly, classical theories are still mum on that. Hence, we take help from the emerging field of quantum cognition that explores these issues with an alternate approach. According to quantum cognition, although the jury is still out on whether quantum physics is involved in neurobiological processes (consciousness, memory, internal experiences), the processes of choice and decision making, which are the products of the warm-wet-noisy brain, may be easily explainable using the operations of quantum physics (Tarlaci 2010). To make our intention crystal clear, we would like to state that we never claim our interpretation to be the only valid one by which the creative process can be explained. Our argument is that, this point of view can help paving the way for a foundation to further understanding this complex cognitive process.

Keeping the above preface in mind, and also the Free association theory of creativity, we separate the whole 'creative process' in three distinct parts:

1. **Pick up stage**: Brain receives the external stimulus which constitutes the first 'node' in the creative process (from here the mental state, represented by a vector in a Hilbert space, will start its divergent evolution). This is equivalent to the situation where a creative person 'picks up' his inspiration.

2. **Leap stage**: The mental state, therefore, starts evolving in a 'steep' or 'loose' path to subsequent nodes by taking 'Leap's between them. This is where the divergent thinking helps the 'node' to make distant nodal connections, expanding the creative process spatially and/or temporally.
3. **Chaotic Evolution stage**: With continuous measurement of the evolution function (measurement here signifies the continuous bombardment of contextual inputs from memory and/or the external stimulus), the evolution of the vector is encountered by a 'white noise', under whose influence, it exhibits nonlinearity and a chaotic nature. This is the phase, which separates two creative individuals (say a musician and a painter) depending on the nature of the memory inputs and stimulus received.

9.4.1. DETAILS OF THE STAGES:

9.4.1.1. Pick up stage

We start by using the standard method to explain any dynamical system and its evolution. i.e., constructing a phase space. Phase space is a space which contains all possible states of a physical system. In this case, the physical system we are interested in can be represented by a 'Mental State function', a wave function on the 'Mental state space'. Using Dirac notation, we denote this state function as $|\Psi>$. Unlike the classical approach where state of a system is denoted by a specific point in phase space, in the quantum mechanical approach, this is a complex valued wave function whose position and momentum cannot be determined simultaneously (Heisenberg's uncertainty principle) (Robertson 1929). The state function can evolve in time and can be changed by 'interactions' with external stimulus or memory (or experiences). Also, our constructed space is made up of all the probable states that this state function can achieve after these 'interactions'. So, to generalize, we can express $|\Psi>$ as a linear combination of all these probable states, which serve as orthogonal basis of the Hilbert space of all mental states, using Eq. (1):

$$|\Psi> = \sum_{k=1}^{n} c_k |k> \tag{1}$$

where c_k's are complex coefficients and $|k>$ are basis vectors which span the Hilbert space of all probable mental states.

Now, when an external stimulus interacts with the wave function, $|\Psi>$ changes to a different state from its initial state. Manousakis (2009) showed that this interaction happens via a concept known as operators. In quantum mechanical point of view, operators are the equivalent of observables in classical physics. So, when the state function interacts with environment, a particular operator, say $\hat{Q}$ (it has a matrix representation), operates on $|\Psi>$ and changes it to a new state $|\Psi'>$. This is given in Eq. (2).

$$|\Psi'> = \hat{Q}|\Psi> \tag{2}$$

This state transition depends on the structure and properties of $\hat{Q}$. Manousakis used the example of binocular rivalry and showed how the two possible state transitions are achieved. In our interpretation, the initial stimulus (which 'provokes' the creative process) operates in the same manner, culminating in the new state $|\Psi'>$. This $|\Psi'>$ works as the first 'node' of the creative network. Hence, change in the Mental state function is what starts the creative process, i.e., the ‘Picking up’ stage.

9.4.1.2. Leap stage

The transition state $|\Psi'>$ transits (or ‘jump’s) between conceptual nodes in this phase. According to the Free association theory, the extent of creativity lies in the distant connections between conceptual nodes. Flatter the connection, the more divergent and novel it is. So, what makes these connections, which are spatially and/or temporally separated, happen?

To help visualize the reader, let’s take the analogy of the structure of an atom (Martin 2004). Electrons move in circular or elliptical paths centering the nucleus. Each path is separated from the other via energy barriers. When the atomic structure is perturbed with an external agent, say a stream of photons (basically, energy is given to the electrons in different shells), electrons in lower energy states move to a higher energy shell, provided they receive sufficient energy to cross the barrier. Depending on the quanta of energy introduced, ground state electrons could reach any of the higher shells (sometimes even out of the atom breaking the binding energy).

Something similar, we predict, could be seen in case of divergent thinking as well. The state function $|\Psi'>$ which constitutes the initial node, is previously acted upon by the external stimulus in form of an operator. The stimulus, we believe, is the necessary perturbation that pushes the transition state $|\Psi'>$ to leap to the next node (the probability of leaping to the next node is equally distributed amongst all the available unless there is an introduction of contextuality by memory or interaction with environment). But what is that ‘energy’, similar to the atomic analogy, which is essential in this leap from primary to a secondary node?

Dependence on Ambiguity: Here, ambiguity plays a crucial role. The ambiguous nature of the external stimulus provides $|\Psi'>$ the ‘energy’ to leap to a secondary node. Similar phenomenon is quite common in the fields of nonlinear and quantum optics where transition or absorption rate of a particle (electron, photon) is proportional to the intensity of the perturbing light. Likewise, in this case, the nodal transition rate is dependent on the ambiguous nature of the stimulus received.

Let us use an example to clarify the point. When a person hears a word, say ‘tiger’, which is absolutely unambiguous in nature (i.e., the person has a well-constructed knowledge or idea about the concept of ‘tiger’), it is unlikely that the person would produce a unique or novel way of associating ‘tiger’ with, say ‘suspension bridge’ (provided the person has a well-developed idea about it too). But a person having an ambiguous (or vague) knowledge about both concepts has a better chance of associating them in a novel way.

The tolerance of ambiguity has been studied before as a factor of a person’s creative aspect (Zenasni et al. 2008; Furnham & Ribchester 1995). According to our interpretation, it is reasonable to inspect this idea as ambiguity has an important role in the divergent production.

9.4.1.3. Chaotic Evolution stage

The Third and final phase of the creative process is Chaotic evolution. Before starting this part, a brief idea of chaos theory is needed. It designates a specific class of dynamical behaviour. According to SH Kellert (1994), it is “the qualitative study of unstable aperiodic behaviour in deterministic dynamical systems”. ‘The aperiodic’ reinforces the point that the same state is never repeated twice. Chaos theory has three essential properties: firstly, they are very sensitive to initial conditions. Secondly, they can display a highly disordered behavior; and third, they are deterministic, that is they obey some laws that completely describe their motion (Faure & Korn, 2001). So, where is the relevance of such a theory to our quantum leap interpretation? The phases we dealt with till now has a finite time limit set on them, namely, tens or thousands of a fraction of a second. But what happens when we try to push the limit of time gradually higher? Also, what if, instead of a single external perturbation, we had a continuous flow of stimulus to deal with? Each of these perturbations is analogous to performing a measurement on a quantum system. And (unlike classical systems) each of the interactions with environment causes some irreducible effect on the system. Quite a few numbers of literature in the recent past have suggested that continuous measurements on a quantum system that is evolving with time is equivalent to averaging over all the possible trajectories that the particle might have taken. Also, this kind of measurements introduces some interesting conditions on the Wigner function of the quantum system (Bhattacharya et al. 2003; Habib et al. 2005). The Wigner function, introduced by Wigner (1997, originally published in 1932), is a probability distribution (more specifically, a quasidistribution) which helps to transform the trajectory of a quantum system or operator in phase space (in terms of its position and momentum variables, i.e., x and p respectively) from Hilbert space. Unlike a classical system, it is not possible to measure x and p of a quantum system simultaneously, thanks to Uncertainty principle. Hence, we resort to the Wigner function, described as eq. (3).

$$W(x,p) = \frac{1}{h}\int e^{-\frac{ipy}{\hbar}}\Psi\left(x+\frac{y}{2}\right)\Psi^*\left(x-\frac{y}{2}\right)dy \tag{3}$$

Where, ψ is the wavefunction, Ψ^* is its complex conjugate and x, p are position and momentum variables and h is Planck's constant. The centroid of the Wigner function is the phase space point defined by the mean values of x and p, i.e., (<x>, <p>).

For this quantum system to act as a classical one, the Wigner function needs to be 'localized', that is, its distribution needs to be sharply peaked about the phase space variables so that its evolution can be described classically in terms of these variables. The time evolution of the centroid of a Wigner function resembles Ehrenfest equations, given in eq. (4).

$$\begin{aligned} \langle\dot{x}\rangle &= \langle p\rangle/m \\ \langle\dot{p}\rangle &= \langle F(x,t)\rangle = -\left\langle \frac{\partial V}{\partial x}\right\rangle, \end{aligned} \tag{4}$$

Where <□> is the expectation value, F(x, t) is the force and V is the potential.

Now, for a highly localized Wigner function, the expectation value < F(x, t) > can be expanded about <x> with a Taylor expansion. Doing so, and neglecting the higher terms, readily makes the equation (4) perfectly Newtonian:

$$\begin{aligned} \langle\dot{x}\rangle &= \langle p\rangle/m \\ \langle\dot{p}\rangle &= F(\langle x\rangle, t) \end{aligned} \tag{5}$$

Thereby, it falls into the classical realm.

Since the Wigner function of an unobserved quantum system rarely remains localized, measuring the system continuously with a high rate of information extraction introduces Gaussian white noise into the evolution equation:

$$<\dot{x}> = <p>/m + (8k)^{1/2}\sigma_x^2\xi(t)$$
$$<\dot{p}> = <F(x, t)> + (8k)^{1/2}C_{xp}\xi(t) \; ; \qquad (6)$$

Where, σ_x^2= variance of x, C_{xp} = co variance of x and p, and ξ(t)= Gaussian white noise. This makes the localized Wigner function close to a Gaussian distribution in nature (Bhattacharya et al., 2004).

The stronger and frequent the measurement is, the more noise is introduced to the system. Careful simulation conditions reveal that the trajectory of such a noise-induced system (provided that the distribution is >> ħ) is classical and chaotic (Bhattacharya et al., 2000, 2004).

The mental state function |Ψ>, an essentially quantum system, also can be described using a Wigner distribution which undergoes similar evolutions. Continuous exposure to the external perturbations introduces the necessary noise which, in principle, gives rise to classical and chaotic behaviors (unless of course the action of the system is small compared to Planck's constant ħ. This is unlikely, since the action of the state function is much larger as one can see from neurobiological signatures like fMRI or EEG, both in spatial and temporal domain). We hypothesize that the above discussed chaotic trajectory is the reason why individual creativity is person specific. Evolution of the mental state, in this environment, depends heavily upon the initial conditions and exhibits disorderness. That's why even after going through the same two initial phases, creative process is found to be different for everyone. The nature of the stimulus, again, decides who will be a painter and who shall excel in music.

9.5. OVERVIEW OF THE HYPOTHESIS AND ITS ADVANTAGES

Here, we have proposed a hypothesis (Quantum Leap interpretation) that tries to explain the creative process from its inception to its evolution. This hypothesis uses quantum physics and its unusual but highly effective approach to divide the whole process into three distinct stages. In the primary or 'Pick up' stage the mental state function (operating in a space we call 'Mental state space') interacts with an external stimulus (denoted by an operator) and begets a new state function. This change marks the start of the creative process. In the next stage, which we denote as 'Leap' stage, this state function (primary node) leaps between the conceptual nodes further, analogous to electrons jumping from lower to higher shells absorbing external energy inside an atom. Ambiguity, similar to energy, is the active agent that makes the state function leap the conceptual nodes. During the evolution, the state function continuously interacts with external environment in the form of contexts, memories or stimulus. These interactions, keep injecting noise in the evolving Wigner distribution of the state function, ultimately localizing it and making it classical and furthermore, chaotic (subjected to strict conditions). This constitutes the third and final stage, called the 'Chaotic evolution' stage. We propose that this final stage indicates individual creativity. That is, since the state function evolves in a disorderly chaotic manner, being highly sensitive to initial conditions and also since this stage is dominated mostly by external stimulus, hence, the nature of stimulus can affect the chaotic evolution heavily. This variation in interaction with stimulus is, we believe, the reason why every creative individual has a unique nature of creativity even after going through the first two common stages.

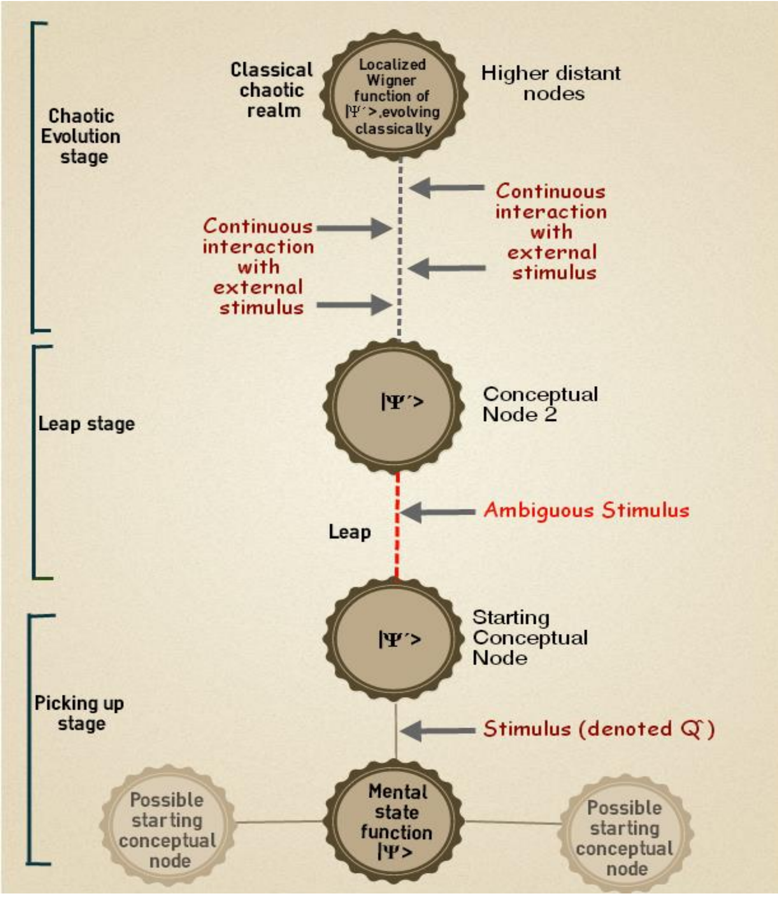


Fig. 9.3: Schematic diagram of the Quantum Leap hypothesis. The three stages (Pick-up stage, Leap stage and Chaotic evolution stage) are marked with the episodic changes in state function ψ

A schematic diagram of the hypothesis is given in Fig. 9.3. The main features and advantages of this interpretation can be summarized as:

1. This novel hypothesis takes into account the fact that explaining creativity using classical realm can't describe the process in its entirety and hence, a quantum physical approach needs to be introduced.

2. Unlike existing ideas which are necessarily top-down (spotting the creativity, followed by the investigation of an explanation), ours is bottom-up. We try to explain the starting point of the process and discuss how it can evolve into the creative effects that one observes as a final picture.

3. The effect of stimulus and its nature (degree of ambiguity, continuity) is given the utmost importance in this interpretation.

9.6. DISCUSSION AND COMPARISON WITH EXISTING IDEAS

In this section, we shall present an overview on how our theory stands amidst some of the existing ideas on creativity and whether it can complement or consummate them in explaining the creative process. Some of the long-standing ideas on creativity include (Sternberg 2018):

- **Psychoanalytic Model:** According to Freud (1920, 1958, 1959), creativity stems from the depths of subconscious. In Quantum Leap Hypothesis (QLH), the initiation and evolution of the mental state function describes the path by which this process begins and further takes course.
- **Four C Model:** The four C's (little-C, Big-C, Pro-C and mini-C) introduced by Kaufman and Beghetto (2009), can be identified using QLH. The everyday creativity or the mini-C is the leaps of the state function to different degrees, further evolution of which falls under the domain of little-C. The early stages of chaotic evolution represent the little-C or 'personal creativity'. The rest of the two (Pro-C and Big-C) are dependent on higher time evolution and exposure to external stimulus. Hence, depending upon the evolution time and nature of external factors, the C categorizations can change.
- **The Componential Model:** Suggested by Amabile (1985), the four components of creativity are: Task Motivation (a creative task must be motivated by internal rewards rather than materialistic ones), Domain-relevant Skills (knowledge of the specific domain), Creativity-relevant Processes (processes which inspire novelty on task performances) and Social Environment (a space where creativity thrives). Two of these four components have direct representation in QLH, namely Creativity-relevant Processes and Social Environment: the nodal jumps of the state function are the foundation of creative process, which, again, needs the external stimulus from the environment to evolve further in the chaotic evolution phase. Also, Domain-relevant Skills constitute the very nodes that the state function is transitioning to and from; the elements that can essentially combine for the creative process. That leaves us with the interesting case of Task Motivation. According to Amabile, creativity is expressed most when it is motivated primarily by interest, enjoyment, satisfaction– the intrinsic motivators-and not by extrinsic motivators like rewards, competition etc. This is identified as the first step of creativity. In QLH interpretation, the earliest step that changes the function $|\Psi\rangle$ to $|\Psi'\rangle$ is a stimulus in the form of an operator $\hat{Q}$. We believe it represents the Task Motivation component from Amabile's interpretation, since this $|\Psi'\rangle$ forms the first node, i.e., 'picking up' or the inception of the creative process. The only point QLH doesn't address is the nature of this operator. From QLH's point of view, the motivation is what propels the process. Hence, the theory focuses on its existence and it is yet to factor in the intrinsic/extrinsic nature, quantitatively.
- **The Theory of Creativity in the Domain and Field:** Creativity, is a product of vigorous screening between Domain (set of some specific rules or procedures) and Field (set of people at the helm of the Domains), says Csikszentmihalyi (1997). This knowledge of rules and procedures constructs the conceptual nodes in the leap stage, for QLH. Also, the external environment of the chaotic evolution stage, similar to the Field component, screens the evolution of the state function. This indicates that the two theories don't conflict, in principle. Although it must be mentioned that Csikszentmihalyi's explanation of Field deals more with the durability of novel ideas (or '*how* novel they are'), whereas QLH attempts to describe the basic principles of creative processes, irrespective of their future survival rates.
- **Geneplore Model:** This model, forwarded by Finke et al. (1992), consists of two main phases: the Generative phase and the Exploratory phase. In the first phase, one constructs a mental representation called the 'preinventive structures', which are explored in the

second phase to generate creative ideas. Also, product constraints could be enforced in any of the two phases. This model differs from QLH in a sense that both the phases of Geneplore model are represented in the Chaotic Evolution phase, where the Wigner distribution is localized due to continuous exposure to external perturbation, i.e., the imposing product constraints. In QLH, the construction and evolution of the first two phases are not adjustable once the wave function evolves into the third phase.

- **Triangular Theory/Investment theory:** Sternberg (2018) has suggested that there are three facets of creativity: defying self, defying crowd and defying Zeigeist (i.e., prevailing and widespread ideas). These three factors combine to form different levels of creativity. This theory measures creativity as a product of the ability to defy three hurdles and evidently reliant on time. Also, it segregates creative process in individualistic form as the nature of creativity Sternberg describes could change forms in one individual over a long period of time. Hence, Sternberg's approach is equivalent to QLH's chaotic evolution, since it also depends on time and external perturbation (self-entrenchments, Zeitgeist-defying, domain-specific elements) for evolving. On this note, it might be an important point to discuss whether the inception or initial evolution of the creative process described by QLH is dependent on this nature of perturbation or is it essentially the same for an individual irrespective of time, turn or context. In QLH, the initial two phases for a creative process doesn't change for any perturbations that are occurring further ahead of time in later phases. But, should a creative process develop again via the same path, the external perturbations and applied conditions/constraints experienced on previous cases has the ability to change certain elements of the 'Pick up' and 'Leap' stages: the operator (representing Task motivation (Amabile 1996)) in the former and the nodal connections in latter is hugely influenced by external stimulus, hence viable to change. Sternberg's idea of variations in creative combinations for the same individual over a large period of time might be the result of such influences on the inception phases of QLH during the evolution of future creative processes.

These are some of the theories on creativity that has been proposed over the years. QLH, the youngest of them all, provides a bridge that on one hand attempts to explain the creative process from a cognitive standpoint and on the other hand, is compatible with the theories of the field. It is different from the prevailing ideas because it is based on the sound physical theories and at the same time incorporates factors whose importance have been argued by many in creativity research such as: environmental factors, tolerance of ambiguity and divergent association followed by convergent selection process.

9.7. SUGGESTED EXPERIMENTAL PROTOCOL

The experimental procedure has to test following factors:

a) Existence of the operator or the Task Motivation in 'Pick up' stage

b) Tolerance of ambiguity in the 'Leap' stage

c) Assessment of the Creative work

d) Assessing the change in creative output when domain-specific constraints are applied

We suggest a protocol that will include a control and an experimental group. Both the groups will perform a creative work, whose creativity is to be rated by experts on that specific field. The experimental group, along with the creative task, will perform tests designed to incorporate task motivation (Amabile 1985) and measure of ambiguity tolerance or MAT (Norton 1975). We predict that the participants of the experimental group scoring high on the Task motivation test and MAT shall do better on the creative assessment task as well. The assessment of creative change due to constraints is to be done by dividing the original experimental group into two and to ask them to perform the creative task again with applying domain-specific constraints on one of them. In this case, we predict that the quality of novelty and quantity of creative output shall be reduced in the conditioned group.

9.8. SUGGESTED EXPANSION OF THE HYPOTHESIS

Further extension of this idea may be envisaged as following:

1. Development of knowledge about the creative process by detailing each of the steps more extensively.
2. Finding the causal relationship between Ambiguity and creativity which in turn can help promote creative aspects in young individuals.
3. In long term, encouraging 'democratic creativity': Instead of narrowing the term 'Creativity' by associating it with genius individuals, spreading it to otherwise mundane society and system will generate productivity across every social platform.
4. Improvements in the development of the skills of creativity, critical thinking and producing novel ideas are essential for developing the next generation of researchers.

9.9. DISCUSSION

This paper presents a novel idea on the link between two well researched and talked about concept – Ambiguity and Creativity. The approach is based on the usage of Quantum physics, the most successful theory in the realm of the behavior of subatomic particles. The rationale behind the approach is: creativity, which, besides being influenced by external environment like society and culture, also remains a deeply cognitive and psychological attribute. This prompts one to find its roots in the depths of cognitive domains. Also, the external factors are non-deniable too. The role of classical laws in the structural and functional aspects of the human brain remains hotly debated. An alternative is provided by the emerging field of quantum cognition. It doesn't involve itself in the dispute of how or whether mind is emerged from quantum structures or fields in macroscopic and thermodynamically open system like human brain. Instead, it uses the mathematical frameworks of quantum physics in explaining cognitive phenomena like decision errors, bistable perception, semantic networks etc. The non-commutative nature of cognition is naturally drawn to the quantum framework as it applies probabilistic amplitudes to circumnavigate the Boolean logic. Based on this idea, we apply quantum mechanical constructs to explain a cognitive process like creativity from the bottom-up perspective. We present a novel idea, Quantum Leap Hypothesis, that describe creativity as a three-step process. Here, we hypothesize a state function can be ascribed to psychological states in a mental state space. The function is a linear superposition of basis states that span the whole mental space. These eigenstates contribute to the state function via probability amplitudes. Once the external stimulus interacts, the state function is 'operated' on by an operator to transition into a new form. This completes the first step. In the next, the event of jumping or 'leaping' of the state function to a distant conceptual node is actualized. This 'leap' is prompted by the embedded ambiguity the concepts under focus carry within. In the third and final stage, the Wigner function corresponding to the quantum system is localized by continuous intervention of external stimulus, thereby turning classical and evolving chaotically. This hypothesis, though in a very nascent stage, gives ambiguity and external stimulus important credits in this entire process to bring about. In the recent times, one other approach to creativity has been developed that utilizes similar proponents in a different manner. It is the so called 'Honing Theory' of creativity put forward by Liane Gabora (2017; Scotney et al. 2019). It discusses the emergence of creativity via self-honing or self-restructuring of one's 'worldview'. HT has three main facets:

(1) Creative individuals are swarmed with ideas that are underdefined, residing in a state of potentiality. It only becomes well-structured in the process of considering them in different contexts. This changes the probability amplitudes associated with the ideas and pushes the state function towards, according to Gabora, the direction of lower 'psychological entropy'.

(2) Creative outputs are the final outcome of a process where new contexts are evaluated or honed and it self-organizes a person's internal worldview into a more stable and low entropy state. The worldview, imagined to be a complex system, self-organizes into a criticality and small 'perturbations' trigger conceptual changes and restructuring of the worldview and internalized knowledge.

(3) More than the output, HT emphasizes on the ways creativity influences the cultural evolution by exchanging concepts and contexts and contributing to the holistic restructuring of stable worldviews among the society.

There are some common aspects that we can find between QLH and HT. Not unlike QLH, HT also discusses the potentiality actualization via probability amplitudes. It also discusses how the creative process starts from a 'gap' in the worldview and it restructures itself accordingly. This, we reckon, is a parallel to the role ambiguity plays in the QLH scenario. The distance between the conceptual nodes is how a novel nodal transition is accelerated. Moreover, HT's emphasis on the external influence on the creative process is also shared by Quantum Leap Hypothesis. Honing Theory differs from QLH in the aspect that it imagines the creative process as a mean for the internal system to reach its lowest possible energy configuration, so to speak. Whereas, QLH maintains an agnostic stand on this topic. It is not possible for this theory, in its current form, to venture an assumption regarding the restructure of 'worldview' which is not a well-defined concept to start with. HT, unlike QLH, doesn't attempt to consider the subsequent evolution of the function, instead it dwells on the non-separability of concepts and effects of contexts on them. Another point on which QLH disagrees with HT is that the process is, at least as of now, is how an individual creative process evolves. The holistic approach of Honing Theory (how creativity influences cultural as a whole), therefore, falls outside the realm of QLH.

To sum up, creativity is too vast a horizon to be consummated by one singular approach. It is embedded in human cognition which in itself is full of controversial research questions and their explanations. Yet, in this work, we tried to scratch the surface using the emergent field of quantum cognition. Although quantum theory mostly deals with the microscopic reality, physical and mathematical ideas regarding it also approve this kind of approach. Lack of empirical evidence do hold against it, but we can hope that the propositions suggested here can be regarded as a trigger that will enrich the knowledge of this very critical and complex domain of Creativity involving the human brain and its mechanisms.

REFERENCES

Abraham, A. (2012). The neuroscience of creativity: a promising or perilous enterprise?. *Creativity and cognitive neuroscience*, 15-24.

Amabile, T. M. (1985). Motivation and creativity: Effects of motivational orientation on creative writers. *Journal of personality and social psychology*, *48*(2), 393.

Amabile, T. M. (1996). Creativity in Context. Boulder, Colorado.

Arndt, M., Juffmann, T., & Vedral, V. (2009). Quantum physics meets biology. *HFSP journal*, *3*(6), 386-400.

Benedek, M., Könen, T., & Neubauer, A. C. (2012). Associative abilities underlying creativity. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(3), 273.

Bhattacharya, T., Habib, S., & Jacobs, K. (2000). Continuous quantum measurement and the emergence of classical chaos. *Physical review letters*, *85*(23), 4852.

Bhattacharya, T., Habib, S., & Jacobs, K. (2003). Continuous quantum measurement and the quantum to classical transition. *Physical Review A*, *67*(4), 042103.

Bhattacharya, T., Habib, S., & Jacobs, K. (2004). The Emergence of Classical Dynamics in a Quantum World. *arXiv preprint quant-ph/0407096*.

Bohm, D. (1990). A new theory of the relationship of mind and matter. *Philosophical psychology*, *3*(2-3), 271-286.

Braginsky, V. B., Braginskiĭ, V. B., & Khalili, F. Y. (1995). *Quantum measurement*. Cambridge University Press.

Brophy, D. R. (2001). Comparing the attributes, activities, and performance of divergent, convergent, and combination thinkers. *Creativity research journal*, *13*(3-4), 439-455.

Busemeyer, J. R., & Bruza, P. D. (2012). *Quantum models of cognition and decision*. Cambridge University Press.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review*, *82*(6), 407.

Craft, A. (2001). An analysis of research and literature on creativity in education. *Qualifications and Curriculum Authority*, *51*(2), 1-37.

Csikszentmihalyi, M. (1997). Flow and the psychology of discovery and invention. *HarperPerennial, New York*, *39*.

Dietrich, A., & Kanso, R. (2010). A review of EEG, ERP, and neuroimaging studies of creativity and insight. *Psychological bulletin*, *136*(5), 822.

Eisberg, R., & Resnick, R. (1985). *Quantum physics of atoms, molecules, solids, nuclei, and particles* (p. 864).

Faure, P., & Korn, H. (2001). Is there chaos in the brain? I. Concepts of nonlinear dynamics and methods of investigation. *Comptes Rendus de l'Académie des Sciences-Series III-Sciences de la Vie*, *324*(9), 773-793.

Finke, R. A., Ward, T. B., & Smith, S. M. (1996). *Creative cognition: Theory, research, and applications*. MIT press.

Fisher, M. P. (2015). Quantum cognition: The possibility of processing with nuclear spins in the brain. *Annals of Physics*, *362*, 593-602.

Freedman, J. L. (1965). Increasing creativity by free-association training. *Journal of experimental psychology*, *69*(1), 89.

Freud, S. (1958). On Creativity and the Unconscious: Papers on the Psychology of Art, Literature, Love. *Religion*, 15.

Freud, S. (1959). Creative writers and day-dreaming. Standard Edition, 9: 141-153. *London: Hogarth Press*, *78*, 59-137.

Furnham, A., & Ribchester, T. (1995). Tolerance of ambiguity: A review of the concept, its measurement and applications. *Current psychology*, *14*(3), 179-199.

Gabora, L. (2017). Honing theory: A complex systems framework for creativity. *Nonlinear Dynamics, Psychology, and Life Sciences*.

Galton, F. (1870). *Hereditary genius: An inquiry into its laws and consequences*. D. Appleton.

Guilford, J. P. (1966). Intelligence: 1965 model. *American psychologist*, *21*(1), 20.

Guilford, J. P. (1967a). Creativity: Yesterday, today and tomorrow. *The Journal of Creative Behavior*, *1*(1), 3-14.

Guilford, J.P. (1967b). *The nature of human intelligence*. McGraw Hill, New York, NY.

Habib, S., Bhattacharya, T., Greenbaum, B., Jacobs, K., Shizume, K., & Sundaram, B. (2005). Chaos and quantum mechanics. *Annals of the New York Academy of Sciences*, *1045*(1), 308-332.

John, B. W. (1913). Psychology as the behaviorist views it. *Psychological Review*, *20*(2), 158-177.

Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four c model of creativity. *Review of general psychology*, *13*(1), 1-12.

Kellert, S. H. (1993). *In the wake of chaos: Unpredictable order in dynamical systems*. University of Chicago press.

Lee, C. S., & Therriault, D. J. (2013). The cognitive underpinnings of creative thought: A latent variable analysis exploring the roles of intelligence and working memory in three creative thinking processes. *Intelligence*, *41*(5), 306-320.

Lubart, T. I. (1994). *Product-centered self-evaluation and the creative process* (Doctoral dissertation, Yale University).

Manousakis, E. (2009). Quantum formalism to describe binocular rivalry. *Biosystems*, *98*(2), 57-66.

Marron, T. R., & Faust, M. (2018). 15 Free Association, Divergent Thinking, and Creativity: Cognitive and Neural Perspectives. *The Cambridge Handbook of the Neuroscience of Creativity*, 261.

Martin, R. M. (2004). *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press.

Mednick, S. (1962). The associative basis of the creative process. *Psychological review*, *69*(3), 220.

Monroe, C., Meekhof, D. M., King, B. E., & Wineland, D. J. (1996). A "Schrödinger cat" superposition state of an atom. *Science*, *272*(5265), 1131-1136.

Norton, R. W. (1975). Measurement of ambiguity tolerance. *Journal of personality assessment*, *39*(6), 607-619.

Ochse, R. A., & Ochse, R. (1990). *Before the gates of excellence: The determinants of creative genius*. CUP Archive.

Razavy, M. (2003). Quantum Theory of Tunneling. Singapore, SGP.

Rechenberg, H. (1982). *The Historical Development of Quantum Theory* (Vol. 1). Springer Science & Business Media.

Robertson, H. P. (1929). The uncertainty principle. *Physical Review*, *34*(1), 163.

Ryhammar, L., & Brolin, C. (1999). Creativity research: Historical considerations and main lines of development. *Scandinavian journal of educational research*, *43*(3), 259-273.

Schrödinger, E. (1935). Mathematical proceedings of the cambridge philosophical society. In *Mathematical Proceedings of the Cambridge Philosophical Society* (Vol. 31, pp. 555-563).

Scotney, V. S., Weissmeyer, S., Carbert, N., & Gabora, L. (2019). The ubiquity of cross-domain thinking in the early phase of the creative process. *Frontiers in psychology*, *10*, 1426.

Sigmund, F. (1920). A general introduction to psychoanalysis. *The Journal of Nervous and Mental Disease*, *52*(6), 548-549.

Skinner, B. F. (1965). *Science and human behavior* (No. 92904). Simon and Schuster.

Skinner, B. F. (1990). *The behavior of organisms: An experimental analysis*. BF Skinner Foundation.

Smith, K. A., Huber, D. E., & Vul, E. (2013). Multiply-constrained semantic search in the Remote Associates Test. *Cognition*, *128*(1), 64-75.

Sternberg, R. J. (2018). A triangular theory of creativity. *Psychology of Aesthetics, Creativity, and the Arts*, *12*(1), 50.

Sternberg, R. J., & Lubart, T. I. (1991). An investment theory of creativity and its development. *Human development*, *34*(1), 1-31.

Sternberg, R. J., & Lubart, T. I. (1995). *Defying the crowd: Cultivating creativity in a culture of conformity*. Free press.

Stuart, H. (1998). Quantum computation in brain microtubules? The Penrose–Hameroff 'Orch OR 'model of consciousness. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *356*(1743), 1869-1896.

Taft, R., & Rossiter, J. R. (1966). The remote associates test: divergent or convergent thinking?. *Psychological reports*, *19*(3_suppl), 1313-1314.

Tarlaci, S. (2010). Why we need quantum physics for cognitive neuroscience. *NeuroQuantology*, *8*(1).

Thiagarajan, T. C., Lebedev, M. A., Nicolelis, M. A., & Plenz, D. (2010). Coherence potentials: loss-less, all-or-none network events in the cortex. *PLoS biology*, *8*(1), e1000278.

Torrance, E. P. (1972). Predictive validity of the torrance tests of creative thinking. *The Journal of creative behavior*.

Torrance, E. P. (1988). The nature of creativity as manifest in its testing. *The nature of creativity*, 43-75.

Wigner, E. P. (1997). On the quantum correction for thermodynamic equilibrium. In *Part I: Physical Chemistry. Part II: Solid State Physics* (pp. 110-120). Springer, Berlin, Heidelberg.

Winnicott, D. W. (1990). Creativity and its origins. *Essential papers on the psychology of women*, 132-145.

Winnicott, D. W. (1991). *Playing and reality*. Psychology Press.

Zenasni, F., Besançon, M., & Lubart, T. (2008). Creativity and tolerance of ambiguity: An empirical study. *The Journal of Creative Behavior*, *42*(1), 61-73.

CONCLUDING NOTE: NOW AND BEYOND

In this thesis, few interesting concepts were put in focus which will be summarized in this section briefly. It started with exploring the neuro-cognitive effects of external stimulus, both visual and auditory. Changes in the neural mechanism, due to a stimulus, is directly manifested by bio-signals like EEG or fMRI. And the changes in the cognitive domain is indirectly manifested via expression, actions and choices. Both of these means were used in two different aspects. It was previously reported how an auditory stimulus like music can affect biological signatures and their complexities (Ghosh et al., 2018). The stimulus, in fact, itself is highly self-similar and exhibit long-range correlations (Hsü & Hsü, 1991). Naturally, the question that arises next is, whether this stimulus-induced biological changes have cognitive ramifications as well. Also, do these processing networks cross paths with other perceptory networks like visual perception and cognition of visual stimulus? In view of this, cognitive effects of auditory stimulus were studied in relation to its visual association. It was found that the stimulus complexity does affect the color choices of the participants. To study the effects of stimulus on a neurological level, next, the focus was on the direct effects of visual stimuli via the EEG data. Using Grey as a baseline, participants were exposed to seven colors of VIBGYOR and the corresponding bio-signals were recorded. The complexity analysis of the same showed an overall increase during the color Blue, followed by Red and Green. It was also revealed that the processing of the perceptual information is not domain-specific even at temporal level, as in, multiple brain areas were involved (Frontal, Parietal and Occipital, in this case) in the process at the same time. Hemispherical asymmetry is also observed, as right hemisphere of the brain is found to be more aroused than the left during visual information processing. These studies shed new lights on stimulus processing and its neural and cognitive correlates. But what makes a stimulus an agent of arousal? Every literature is a subset of a universal set of words used in that language. Yet, the dictionary doesn't evoke affective response whereas a play by Shakespeare does. Is it the grammar that invokes the emotive arousal, then? Semantics? Or maybe the context? Could it be because of reader's fruition/failure of expectation of an event (a word, a sentence, a literary expression, a narrative pattern)? Similarly, a music is created from the same number of notes that another piece of music is also made up of. What makes them different? Is it only the order or sequence of the notes? More importantly, how do they invoke very different kinds of emotion, even when their structural roots are the same? Inspired by these questions, it was the structure of music that was investigated to find ques which could help understand their effects on the listeners. Hardcore statistical methods were used to study Indian Classical Music and its various renditions in a novel manner. These methods have revealed parameters which can identify hitherto abstract or qualitative properties based on present musical notes and their usage patterns. It can quantify the kinetic nature of the rendition, the 'closeness' of the rendition to the traditional techniques used, the complexity of the notation vocabulary, improvisational characteristics, affinity/hinderance to ornamentation and inflective movements ('musical analyticity'). These properties are some of the essential reasons why a piece of music is aesthetically arousing. Additionally, the parameters could categorize and classify genres, artists and *Raga*s as well. Finally, the question - whether classical theories and logics are sufficient in describing the realms of cognition and perception – was explored. It was shown that the Boolean logic breaks down while accounting for auditory perception with an inherent ambiguity. This argues for a framework that will improve upon the classical limitations, something that is provided in quantum cognition models. This approach is different from the quantum neurobiological models which tries to explain the phenomenological roots of cognition and consciousness using quantum theory. Quantum cognition uses the mathematical framework of quantum theory (probability amplitudes, actualization of potentials, quantum operators) to describe the cognitive domain and resolve non-Boolean conundrums. This approach is later used to describe a model of creative cognition which prioritises ambiguity (or tolerance of ambiguity) as a driving factor. The studies and approaches described in this thesis have one underlying thread that binds them all – that is emotion. A factor that provides evolutionary advantage in the perception of stimuli and remains one of the fundamental components that plays a major role in their cognitive binding. Unfortunately, science is yet to objectively quantify this human construct which brings us to the question that whether such universality can exist at all. This thesis opens up the opportunity to amalgamate different field like physics, music, color perception, psychology, cognition and neuroscience and the author believes such interactions could answer several of these hard questions in the future.

Physics takes pride in describing the widest range of reality using the fewest concepts and principles. From the eyes of physics, the domain of cognition, perception and emotion is too peripheral and vague. However, one of the modern physics' founding fathers, Sir Issac Newton himself took on such a task, calling it 'spiritual substance' (Dempsey, 2006). His attempts remained unsuccessful and since then such discussion were reserved for philosophers and psychologists. In recent times, advances in neuroscience have put these back into the realms of hardcore science. With technical developments in scanning (PET, fMRI, MEG, EEG) providing increasingly sophisticated representations of what goes on in an individual's brain as they actually have an experience, taking the field several steps nearer to finding the neural correlates of subjective reality. According to Grossberg & Levine (1987), the mechanism of emotions are neural signals connecting instinctual and conceptual brain regions. They communicate instinctual needs to conceptual recognition-understanding mechanisms. Their function is to motivate behavioral and conceptual representation-models, which correspond to objects or events that can potentially satisfy instinctual needs, so that these models receive preferential attention and processing resources within the brain. Thus, emotions evaluate concepts for the purpose of instinct satisfaction. Our higher cognitive abilities involve many emotions, which include processes of learning, emotions in the voice prosody, emotions of cognitive dissonances, as well as musical emotions. Cognitive variants like perception of spatial relations, color space, color constancy, the pitch of sounds, tastes, numbers, objects and actions rely on neural transformations that support optimal generalization and categorization. Coming back to cognition, Linguistics and music, also, have greater roles to play in aspects of cognitive research. Chomsky tried to separate language from cognition but language is so important for thinking that it is difficult to comprehend what cognition would be without language (Chomsky, 1957). Do we think with words, or only use words as labels when a chunk of a thinking process is complete? There is virtually infinite number of possible associations between words and objects, so how is it possible that every child learns correct associations? What changes exactly does language bring in neural mechanisms of a child? The science needs to understand the mechanisms of language and cognition interactions; why they are so interdependent, and yet so separate? Emotionality of language is another frontier. Language and its main way of functioning, speech, can only function if sounds of words are perceived emotionally. If a word sound produces no emotions and no motivations, the word has no effective meaning. Emotional prosody of human voice, even if unnoticed, affects the entire psyche and even culture. This remains an area of future research. Similarly, music cognition deserves attention as well. The origin of music cognition still remains a mystery (Honing et al., 2015). The origin and the effects of emotionality associated with it is a huge area of research. Why music evokes emotion or how its emotional signatures appeal universally, answers to these questions can lead to understanding higher cognitive abilities. Traditionally a lot of these problems belonged to the subjective reality, usually away from scientific objectivity, it was the rise of cognitive psychology in the 1960s which discussed possibility and necessity of writing subjective experience back into the scientific study, establishing experimental methods and theoretical models by which to understand the hidden mental apparatus of human perception, cognition, and action (e.g. Neisser, 1976; Miller et al.,1960). To quote the great E. P. Wigner (1995): "Both Physics and Psychology claim to be all-embracing discipline; the first because it endeavours to describe all nature; the second because it deals with all mental phenomena, and nature exists for us only because we have cognizance of it. Both disciplines may yet be united into a common discipline without overtaxing our mendacity for abstraction." For physics, the very first test of a scientific theory is its elegance and beauty, as in, its ability to describe a vast area of knowledge from few basic principles, and to make experimentally testable predictions. Future researches in cognitive domain, involving all the big guns like physics, neuroscience and psychology should create a symbiotic environment where the ideas and their testability thrives beyond intellectual gatekeeping. This thesis hopes to set the proverbial ball rolling in that direction and hopes to be an initiation point of such a massive undertaking.

REFERENCES

Chomsky, N. (1957). Syntactic Structures (The Hague: Mouton, 1957). *Review of Verbal Behavior by BF Skinner, Language*, *35*, 26-58.

Dempsey, L. (2006). Written in the flesh: Isaac Newton on the mind–body relation. *Studies in History and Philosophy of Science Part A*, *37*(3), 420-441.

Ghosh, D., Sengupta, R., Sanyal, S., & Banerjee, A. (2018). *Musicality of human brain through fractal analytics*. Springer Singapore.

Grossberg, S., & Levine, D. S. (1987). Neural dynamics of attentionally modulated Pavlovian conditioning: blocking, interstimulus interval, and secondary reinforcement. *Applied optics*, *26*(23), 5015-5030.

Honing, H., ten Cate, C., Peretz, I., & Trehub, S. E. (2015). Without it no music: cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664), 20140088.

Hsü, K. J., & Hsü, A. (1991). Self-similarity of the" 1/f noise" called music. *Proceedings of the National Academy of Sciences*, *88*(8), 3507-3509.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). Plans and the structure of behavior. Henry Holt and Co.

Neisser, U. (1976). Principles and implications of cognitive psychology. *Cognition and reality*.

Wigner, E. P. (1995). The limits of science. In *Philosophical Reflections and Syntheses* (pp. 523-533). Springer, Berlin, Heidelberg.

LIST OF MY PUBLICATIONS:

<u>Book Chapters:</u>

1. **Ambiguity Creates: A Quantum Leap Interpretation**
   Souparno Roy, Archi Banerjee, Ranjan Sengupta and Dipak Ghosh
   Creativity and Cognition in Art and Design, Bloomsbury Publishers **ISBN**: 978-93-86349-88-0.

2. **Musical Influence on Visual Aesthetics: An Exploration on Intermediality using Audience Response, Feature and Fractal Analysis**
   Archi Banerjee**,** Pinaki Gayen, Shankha Sanyal, Sayan Nag, Junmoni Borgohain, Souparno Roy, Priyadarshi Patnaik and Dipak Ghosh
   Accepted for publication in SpeechMusic2022 to be published by Springer.

3. **A Fractal Approach to Characterize Emotions in Audio and Visual Domain: A Study on Cross-Modal Interaction**
   Shankha Sanyal, Sayan Nag, Archi Banerjee, Souparno Roy, Ranjan Sengupta and Dipak Ghosh
   Accepted for publication in SpeechMusic2022 to be published by Springer.

<u>Journal Publications:</u>

1. **Chaos based non-linear cognitive study of different stimulus in the cross-modal perspective**
   Souparno Roy, Chandrima Roy, Sayan Nag, Archi Banerjee**,** Ranjan Sengupta, Dipak Ghosh
   Physica A: Statistical Mechanics and its Applications (2020) 546, 122482
   https://doi.org/10.1016/j.physa.2019.122842

2. **Brain response to color stimuli: An EEG study with nonlinear approach**
   Souparno Roy, Archi Banerjee, Chandrima Roy, Sayan Nag, Shankha Sanyal**,** Ranjan Sengupta, Dipak Ghosh
   Cognitive Neurodynamics (2021), Springer Nature 15 (6), 1023-1053
   https://doi.org/10.1007/s11571-021-09692-z

3. **A novel study on perception–cognition scenario in music using deterministic and non-deterministic approach**
   Archi Banerjee, Shankha Sanyal, Souparno Roy, Sayan Nag, Ranjan Sengupta, Dipak Ghosh
   Physica A: Statistical Mechanics and its Applications (2021), 567, 125682
   https://doi.org/10.1016/j.physa.2020.125682

4. **Tagore and Neuroscience: A Non-Linear Multifractal Study to encapsulate the evolution of Tagore Songs over A Century**
   Shankha Sanyal, Archi Banerjee**,** Sayan Nag**,** Uddalok Sarkar, Souparno Roy, Ranjan Sengupta, Dipak Ghosh
   Entertainment Computing (2020), 100367;
   https://doi.org/10.1016/j.entcom.2020.100367

5. **Ragas in Bollywood music — A microscopic view through multifractal cross-correlation method**
   Shankha Sanyal, Archi Banerjee, Souparno Roy, Sayan Nag, Ranjan Sengupta, Dipak Ghosh
   The Journal of Acoustical Society of India (2021) (ISSN: 0973-3302) 48 (1-2), 91-97

6. **Categorization of Indian Classical Music Using MB-BE Distributions**
   Souparno Roy, Archi Banerjee, Shankha Sanyal, Dipak Ghosh, Ranjan Sengupta
   Journal of Software Engineering Tools & Technology Trends 6(3), 2019, 9-15

7. **From Speech to Songs via Recital–A Case of Phase Transition? A Fractal Study of Tagore's Work**
   Archi Banerjee, Shankha Sanyal, Souparno Roy, Ranjan Sengupta, Dipak Ghosh
   Trends in Electrical Engineering 3 (9), 2020, 30-36

8. **Improvisation in Indian Classical Music: Probing with MB and BE Distributions.**
   Souparno Roy, Archi Banerjee and Shankha Sanyal
   Jadavpur Journal of Languages and Linguistics 4, no. Sp. I (2020): 130-143.

9. **An Acoustical and Neuro-cognitive Study on the Effects of Lyrics in Song from Non-linear Perspective**
   Archi Banerjee, Shankha Sanyal, Souparno Roy, Priyadarshi Patnaik, Dipak Ghosh
   Jadavpur Journal of Languages and Linguistics 4, no. Sp. I (2020): 40-53.

10. **A Fractal Approach to Characterize Emotions in Audio and Visual Domain: A Study on Cross-Modal Interaction**
    Sayan Nag, Uddalok Sarkar, Shankha Sanyal, Archi Banerjee, Souparno Roy, Samir Karmakar, Ranjan Sengupta, Dipak Ghosh
    Journal of Image Processing & Pattern Recognition Progress, 6(3), 2019 1-7

National/International Conference Publications:

1. **A study on Raga characterization in Indian classical music in the light of MB and BE distribution**
   Journal of Physics: Conference Series, vol. 1896, no. 1, p. 012007. IOP Publishing, 2021.

2. **An acoustical and psychological study on contribution of lyrics in raga-based happy and sad Indian music**
   Journal of Physics: Conference Series, vol. 1896, no. 1, p. 012017. IOP Publishing, 2021.

3. **Lyrics on the melody or melody of the lyrics?**
   The Journal of the Acoustical Society of America (2020), *148*(4), 2699-2699
   https://doi.org/10.1121/1.5147477

4. **In Search of Rhythm: A Correlational Study between Different Emotive Poetry pieces and Corresponding manifestation in the EEG signals using Fractal Analytics - a pilot study**
   Proc. of International Conference cum Workshop on Rhythm in Speech and Music (ICRSM-2020), Jadavpur University, Kolkata, January 16-18, 2020

5. **An acoustical and neuro-cognitive study on the appraisal of reading, recitation and song in nonlinear perspective**
   Proc. of the National Symposium on Acoustics-2019 (NSA-2019), 17-19 October, 2019, Ravenshaw University, Cuttack, India

6. **Analyzing Music with Bose-Einstein Distribution: A Case Study with Indian Classical Music**
   Proc. of Western Asian Pacific Acoustic Conference, New Delhi November 11-15, 2018; ISSN-0973-3302

7. **Neuroaesthetic study on the association of 'rasa's and colors in human brain in perspective of Indian classical dance: Do emotions need a new governing paradigm?**
   Proc. of International Symposium FRSM-2017, December 14-15, 2017, NIT, Rourkela, India

8. **A Simultaneous EEG and EMG study to quantify emotions from Hindustani Classical Music**
   Proc. of National Symposium on Acoustics (NSA - 2017), October 28-30, 2017, Aligarh Muslim University, Aligarh, India

9. **Improvisation In Hindustani Vocal Music**
   Proc. of International Symposium FRSM-2016, November 11-12, 2016, North Orissa University, Baripada, India; Ird India, ISBN: 978-93-81693-07-03

10. **Music, color and emotion: A novel study from psycho-neuro-cognitive perspective**
    Proc. of International Symposium FRSM-2016, November 11-12, 2016, North Orissa University, Baripada, India; Ird India, ISBN: 978-93-81693-07-03

11. **Chaos Based Study on Association of Color with Music in the Perspective of Cross-Modal Bias of the Brain**
    Vol. 14 CIC 2016 Special Issue, International Journal of Computer Science and Information Security (IJCSIS),36-42; Proc. of International Conference on Advances in Computational Intelligence and Communication (CIC 2016), Pondicherry Engineering College, Puducherry, India October 19 & 20, 2016; ISSN 1947-5500

12. **Does ambiguity indicate non-classicality in mental states?**
    Proc. of the West Bengal State Science and Technology Congress, February 28-29, 2016, Presidency University, Kolkata

13. **Cue for the categorization of Rabindrasangeet**
    Proceedings of the International Symposium FRSM-2015, November 23-24, 2015, Indian Institute of Technology (IIT), Kharagpur, India

14. **Categorization & quantifying effect of ragas and colors on human brain: A neurocognitive nonlinear study**
    Proceedings of the International Symposium FRSM-2015, November 23-24, 2015, Indian Institute of Technology (IIT), Kharagpur, India

15. **Non-classicality in mental states: An experimental study with ambiguous audio (music) stimuli**
    Proceedings of the International Symposium FRSM-2015, November 23-24, 2015, Indian Institute of Technology (IIT), Kharagpur, India

16. **Study of objective cues for Improvisation in Hindustani music**
    Proceedings of the International Symposium FRSM-2015, November 23-24, 2015, Indian Institute of Technology (IIT), Kharagpur, India

17. **Detrended Fractal Analysis of EEG signals with Music Stimuli of Contrast Emotions**
    Proceedings of the National Symposium on Acoustics-2014 (NSA-2014), 12-14 November, 2014, All India Institute of Speech and Hearing, Mysore, India

arxiv preprint:

1. **Non-classicality in mental states: an experimental study with ambiguous audio (music) stimuli**

   arXiv preprint arXiv:1603.06401 (2016)

2. **Neural (EEG) Response during Creation and Appreciation: A Novel Study with Hindustani Raga Music**

   arXiv preprint arXiv:1704.05687 (2017)

3. **Gestalt Phenomenon in Music? A Neurocognitive Physics Study with EEG**

   arXiv preprint arXiv:1703.06491 (2017)

4. **Ambiguity invokes Creativity: looking through Quantum physics**

   arXiv preprint arXiv:1702.07734 (2017)

Physica A 546 (2020) 122842

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Chaos based non-linear cognitive study of different stimulus in the cross-modal perspective

Souparno Roy [a,b,*], Chandrima Roy [b,c], Sayan Nag [d,e], Archi Banerjee [b,f], Ranjan Sengupta [b], Dipak Ghosh [b]

[a] Department of Physics, Jadavpur University, Kolkata 700032, India
[b] Sir C. V. Raman Centre for Physics and Music, Jadavpur University, Kolkata 700032, India
[c] Department of Electronics & Communication Engineering, Heritage Institute of Technology, India
[d] Department of Electrical Engineering, Jadavpur University, Kolkata 700032, India
[e] Department of Medical Biophysics, University of Toronto, Canada
[f] Rekhi Centre of Excellence for the Science of Happiness, IIT Kharagpur, India

ARTICLE INFO

Article history:
Received 20 August 2018
Received in revised form 10 June 2019
Available online 28 September 2019

Keywords:
Music
Color
MFDFA
RGB triplet
Hexadecimal representation
Cross-modal correspondence

ABSTRACT

The relationship between color and music as part of the complex system consisting of visual and auditory domain has been investigated in this study. As both the stimulus forms are processed in the same part of the human body, i.e., the brain, it will be really interesting to examine whether they share a similarity in perception. Needless to say that color and music both have strong impact on emotion and feelings & also a few studies have been reported in literature to explore causal relationship between color and emotion. This work reports a neuro-cognitive study on response of brain to two different stimulus and their cross-modal associations. In this study the correlation between emotional arousal and the effect of audio and visual stimuli has been studied from a new perspective. 93 participants were asked to hear 6 different music pieces (each of 30 s duration). The type of emotion elicited by different music pieces were identified by the participants from a given collection of possible emotional responses. Then they are asked to assign a color associating the emotion from a given color wheel (structured according to Munsell color system/RGB color space). Each color, associated with a particular music piece, is a mixture of specific Red, Green and Blue values (RGB triplet) and has a specific HEX number (hexadecimal representation), which is recorded for each response. Then, the musical pieces used were further zoomed with the help of fractal technique to identify different emotions related to music in a quantitative approach. Here, to analyze the complexity of the sound signal (which are non-stationary and scale varying in nature), we have used Multifractal detrended fluctuation analysis (MFDFA), which is capable of determining multifractal scaling behavior of non-stationary time series. From the experimental data, it is seen that the visual and emotional response to the auditory stimulus follows a specific trend which is directly related to the stimulus complexity.

* Corresponding author at: Department of Physics, Jadavpur University, Kolkata 700032, India.
E-mail address: thesouparnoroy@gmail.com (S. Roy).

https://doi.org/10.1016/j.physa.2019.122842
0378-4371/

Cognitive Neurodynamics (2021) 15:1023–1053
https://doi.org/10.1007/s11571-021-09692-z

RESEARCH ARTICLE

Brain response to color stimuli: an EEG study with nonlinear approach

Souparno Roy[1,2] · Archi Banerjee[2,3] · Chandrima Roy[2,4] · Sayan Nag[5] · Shankha Sanyal[2,6] · Ranjan Sengupta[2] · Dipak Ghosh[2]

Received: 12 August 2020 / Revised: 22 May 2021 / Accepted: 11 June 2021 / Published online: 6 July 2021

Abstract
Color perception is a major guiding factor in the evolutionary process of human civilization, but most of the neurological background of the same are yet unknown. This work attempts to address this area with an EEG based neuro-cognitive study on response of brain to different color stimuli. With respect to a Grey baseline seven colors of the VIBGYOR were shown to 16 participants with normal color vision and corresponding EEG signals from different lobes (Frontal, Occipital & Parietal) were recorded. In an attempt to quantify the brain response while watching these colors, the corresponding EEG signals were analysed using two of the latest state of the art non-linear techniques (MFDFA and MFDXA) of dealing complex time series. MFDFA revealed that for all the participants the spectral width, and hence the complexity of the EEG signals, reaches a maximum while viewing color Blue, followed by colors Red and Green in all the brain lobes. MFDXA, on the other hand, suggests a lower degree of inter and intra lobe correlation while watching the VIBGYOR colors compared to baseline Grey, hinting towards a post processing of visual information. We hope that along with the novelty of methodologies, the unique outcomes of this study may leave a long term impact in the domain of color perception research.

Keywords Color perception · EEG · Nonlinear study · MFDFA · MFDXA · VIBGYOR

Introduction

From the advent of human civilization, color and perception of color has been intimately involved with it. For survival or evolutionary purposes such as choosing safe foods, finding safe routes to navigate or perception of time during the day, for aesthetic purposes such as variations in artistic expressions of different era, for changing range of emotional experiences to various stimuli, even in the modern world for corporate branding—color reshapes the richness of complex visual information (Hanson 2012). And that is precisely what both helps and hinders research on the effect of color on humans: the sheer volume of research done on visual than any other sensory modality is due to the fact that our interaction with the world has historically depended more on the vision and processing visual information (Hutmacher 2019; Pike et al. 2012). And the hindrance stems from the fact that the experience of color is very subjective and to some extent, context dependent (Lotto and Purves 2002; Elliot and Maier 2012). Nevertheless, the study of color perception and its effects in human brain is fascinating as well as important because it entertains both practical and theoretical concerns.

The existing literature on this component of visual perception highlights two main aspects: psychological and physiological.

Color and psychological functioning

The theory that colors can cause psychological arousal dates back to Nineteenth century when Goethe (Goethe

✉ Souparno Roy
thesouparnoroy@gmail.com

1 Department of Physics, Jadavpur University, Kolkata, India
2 Sir C.V. Raman Centre for Physics and Music, Jadavpur University, Kolkata, India
3 Rekhi Centre of Excellence for the Science of Happiness, IIT Kharagpur, Kharagpur, India
4 Department of Electronics & Communication Engineering, Heritage Institute of Technology, Kolkata, India
5 Department of Medical Biophysics, University of Toronto, Toronto, Canada
6 School of Languages and Linguistics, Jadavpur University, Kolkata, India

Physica A 567 (2021) 125682

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

A novel study on perception–cognition scenario in music using deterministic and non-deterministic approach

Archi Banerjee [a,b,c,*], Shankha Sanyal [a,f], Souparno Roy [a,d], Sayan Nag [e], Ranjan Sengupta [a], Dipak Ghosh [a]

[a] *Sir C.V. Raman Centre for Physics and Music, Jadavpur University, India*
[b] *Rekhi Centre of Excellence for the Science of Happiness, IIT Kharagpur, India*
[c] *Shrutinandan School of Music, Kolkata, India*
[d] *Department of Physics, Jadavpur University, India*
[e] *Department of Medical Biophysics, University of Toronto, Canada*
[f] *School of Languages and Linguistics, Jadavpur University, India*

ARTICLE INFO

Article history:
Received 21 May 2019
Received in revised form 13 December 2020
Available online 19 December 2020

Keywords:
Pitch
Loudness
Timbre
Perception–cognition
EEG
MFDFA

ABSTRACT

In the last few decades, nonlinear science and chaos theory has provided several robust non-deterministic tools by means of which the complexity of a nonlinear audio waveform can be measured precisely. On the other hand, sound signal analysis in linear deterministic approach has reached a new dimension where a number of well equipped software have been developed which can minutely measure and control the basic parameters of sound like pitch, intensity, tempo etc. The main objective of the present work is to quantitatively study the changes in acoustic signal complexity (measured using chaos based fractal technique) with individual variation in pitch, loudness and timbre of a sound signal. EEG (Electroencephalography) was also performed on 10 participants to see how the neuro-cognitive attributes of a sound change, i.e. when these basic components — pitch, loudness and timbre of the sound vary, one at a time. Single strokes of a piano were recorded where pitch and loudness of the sound signals were varied one at a time keeping the other parameters fixed. Then the sounds of 14 different musical instruments playing the same pitch at same loudness were recorded, which effectively served the purpose of timbre variation. EEG experiment was conducted with these audio signals as stimuli for the participants. The multifractal spectral widths were calculated for all the music signals as well as the corresponding EEG signals using Multifractal Detrended Fluctuation Analysis (MFDFA) and compared with each other. The results point towards the direction of a correlation between the conventional linear parameters and the latest nonlinear features in the acoustic domain, while the changes in the multifractal values of the different EEG waves reveal new information about the cognition of the basic features of sound in human brain. This study is a novel attempt to provide new data in engulfing apparent objective (acoustics) - subjective (EEG) connection, which is highly needed for building any model for perception–cognition connectivity.

* Corresponding author at: Rekhi Centre of Excellence for the Science of Happiness, IIT Kharagpur, India.
E-mail address: archibanerjee7@gmail.com (A. Banerjee).

https://doi.org/10.1016/j.physa.2020.125682
0378-4371/

Entertainment Computing xxx (xxxx) xxxx

Contents lists available at ScienceDirect

Entertainment Computing

journal homepage: www.elsevier.com/locate/entcom

Tagore and neuroscience: A non-linear multifractal study to encapsulate the evolution of Tagore songs over a century☆

Shankha Sanyal[a,b,*], Archi Banerjee[b,d], Sayan Nag[c], Uddalok Sarkar[b], Souparno Roy[b,e], Ranjan Sengupta[b], Dipak Ghosh[b]

[a] *School of Languages and Linguistics, Jadavpur University, India*
[b] *Sir C.V. Raman Centre for Physics and Music, Jadavpur University, India*
[c] *Department of Medical Biophysics, University of Toronto, Canada*
[d] *Rekhi Centre of Excellence for the Science of Happiness, IIT Kharagpur, India*
[e] *Department of Physics, Jadavpur University, India*

ARTICLE INFO

Keywords:
Rabindrasangeet
Evolution
EEG
Hurst exponent
MFDFA
Multifractal spectral width

ABSTRACT

The verses of Rabindranath Tagore have been sung by various artistes over generations spanning over almost 100 years. There are few songs which were popular in the early years and have been able to retain their popularity over the years while some others have faded away in the course of time. In this study we tried to find cues in the singing style of these songs, sung by different singers spanning over almost five generations, which have kept them alive for all these years. For this, we took 3 min clips of four Tagore songs which are being sung by atleast five generations of artistes over 100 years and analyzed the acoustic signals with the help of latest nonlinear technique Multifractal Detrended Fluctuation Analysis (MFDFA). Next EEG data was collected from 5 persons who listened to 30 sec clips of two Tagore songs sung over five generations of artistes in chronological order. The EEG response from the participants were analyzed with the help of the same MFDFA technique and the multifractal spectral width was considered as the parameter which can help in the identification of cognitive evolution of the Tagore songs. The multifractal spectral width is a manifestation of the inherent complexity of the signal and in future, may prove to be an important parameter to identify the singing style of a particular generation of singers and how this style varies over different generations. The EEG responses from the participants reflect how the perception and cognition of the same Tagore songs evolve over generations. The results and implications are discussed in detail.

1. Introduction

1.1. Rabindranath Tagore: A historical and scientific throwback

The great visionary from Bengal, Rabindranath Tagore once said "*Whatever fate may be in store in the judgment of the future for my poems, my stories and my plays, I know for certain that the Bengali race must accept my songs, they must all sing my songs in every Bengali home, in the fields and by the rivers… I feel as if music wells up from within some unconscious depth of my mind, that is why it has certain completeness*" (Tagore, Bangla 1407). More than 75 years after his demise, we know for sure that his songs have sustained the various changes that have come in our society and been modified accordingly to keep itself relevant. Tagore (1861–1941) was a Bengali poet, philosopher, artist, playwright, composer and novelist. India's first Nobel laureate, Tagore won the 1913 Nobel Prize in Literature for *Gitanjali or Song Offerings*. He composed the text of both India's and Bangladesh's respective national anthems. Tagore travelled widely and was friends with many notable 20th century figures such as William Butler Yeats, H.G. Wells, Ezra Pound, and Albert Einstein. Accordingly, a number of his songs had been influenced by amalgamation from a number of different cultures. While caring for both the traditions, classical and folk, he respected the inviolable sanctity of neither and freely took from each what suited his purpose. He was not even averse to borrowing from western melodies, although he did very little of that and made his own whatever he took from other sources. Tagore's music cannot be separated from his literature, most of which went on to become lyrics for his songs. Primarily influenced by the *thumri* style of Hindustani classical music, Tagore's

☆ This paper has been recommended for acceptance by Matthias Rauterberg. (Where JONG-IL PARK is the name of the Handling Editor).

* Corresponding author at: School of Languages and Linguistics, Jadavpur University, India.
E-mail address: ssanyal.sanyal2@gmail.com (S. Sanyal).

https://doi.org/10.1016/j.entcom.2020.100367
Received 27 May 2019; Received in revised form 3 February 2020; Accepted 8 April 2020

A study on Raga characterization in Indian classical music in the light of MB and BE distribution

S Roy[1,2], A Banerjee[1,3*], S Sanyal[1,4], D Ghosh[1] and R Sengupta[1]

[1]Sir C.V.Raman Centre for Physics and Music, Jadavpur University, Kolkata, India
[2]Department of Physics, Jadavpur University, Kolkata, India
[3]Rekhi Centre of Excellence for the Science of Happiness, IIT Kharagpur, India
[4]School of Languages and Linguistics, Jadavpur University, Kolkata, India

*archibanerjee7@gmail.com

Abstract. Raga characterization in Indian classical music is an important aspect of music learning in this country. But the methods usually followed are mostly qualitative. In this study, we intend to quantify such abstractness using measurable parameters. To study musical information congregation quantifiably, we introduce methods based on well-known concepts used in Statistical Physics, namely Maxwell-Boltzmann (MB) and Bose-Einstein (BE) distribution. In this present study, these distributions have been applied on the chosen acoustic signals to find new parameters (equivalent to 'temperature' in physical systems) which can distinguish between different features of different ragas (containing the same notes) in Indian classical music. Music clips chosen were the 'Alap' part of these three different ragas (Marwa, Puriya, Sohini) sung by a legendary classical music maestro. All of the chosen three ragas are based on the following same note structure: Sa, komal Re, shuddh Ga, tivra Ma, shuddh Dha, shuddh Ni. To apply MB statistics to music, it is assumed that different notes with different occurrence frequencies are at different energy levels, the distribution of which follows the MB distribution pattern. In case of BE statistics, a rank-frequency distribution of the time durations of various notes of different ragas is studied. The resulting analysis gives rise to a number of parameters that help to categorize the individual characteristics of ragas. The methods studied here are novel in the music research field and can prove to be useful in the fields of music and speech as quantifying parameters for style identification.

Keywords: Maxwell-Boltzmann distribution, Bose-Einstein distribution, Indian classical music, Raga characterization, temperature

1. Introduction and background

Raga, in spirit, is the structural unit that holds the huge body of Indian classical music. It is (but not limited to) a combination of musical notes- comprising an expression of one or more emotions-following certain rules. *Raga*s have a well-defined structure consisting of a series of four/five (or more) musical notes upon which its melody is constructed. However, more than the notes themselves, the way they are approached and used in musical phrases is more fundamental in defining a specific *Raga* rendition. This allows ample scope of improvising beyond the structured framework, a feature that provides Hindustani Classical Music a distinguished character among world music scenario. The

An acoustical and psychological study on contribution of lyrics in *raga*-based happy and sad Indian music

A Banerjee[1,2]*, S Sanyal[2,3], P Gayen[4], S Roy[2,5], P Patnaik[1,4] and D Ghosh[2]

[1] Rekhi Centre of Excellence for the Science of Happiness, IIT Kharagpur, India
[2] Sir C.V. Raman Centre for Physics and Music, Jadavpur University, Kolkata, India
[3] School of Languages and Linguistics, Jadavpur University, Kolkata, India
[4] Department of Humanities and Social Sciences, IIT Kharagpur, India
[5] Department of Physics, Jadavpur University, Kolkata, India

*Corresponding author, e-mail address: archibanerjee7@gmail.com

Abstract. A perfect complementary relationship between the lyric and the melody can give birth to a beautiful song. The melodic expression of a song is universal but the lyrical expression is not – lyric is culture specific because of its language dependence. Can melody itself communicate the core emotions of a song? Or does the addition of a lyrical sense significantly change its emotional experience? This study looks for the answers focusing on a unique subgenre of Indian Classical Music – *Ragashroyi* compositions, where the melodic movements sincerely follow the *Raga* pathways but the lyrics explore a much deeper and wider variety of emotions compared to *Raga bandishes*. Recordings were collected from two eminent vocalists (1 male, 1 female), each of whom was asked to sing (with proper lyrics) and hum (without meaningful lyrics) any two Bengali *Ragashroyi* compositions of two opposite emotions – happiness and sadness. Hurst Exponents, obtained from robust non-linear Detrended Fluctuation Analysis (DFA) of the recorded acoustic waveforms, were compared for each song-humming pair having same melodic structure to understand the acoustic contribution of the lyrics in a song quantitatively. A comparative audience response study was also conducted where several humming and song clips were played randomly and two groups (one which understands Bengali, the other which does not), each having 30 participants, were asked to mark the emotions and the characteristic features corresponding to each clip on 5 point Likert scale, and their responses were compared for each song-humming pair. This pilot study on *Ragashroyi* Indian music explores in depth the contribution of lyrics in vocal music from both the perspectives of computational acoustics and audience psychology.

Keywords: Singing & Humming, Indian Classical Music, Human response, Non-linear analysis, DFA

1. Introduction

From mimicking the sounds of nature to expressing the most profound realizations of this universe in songs – vocal music has evolved through millions of years and human voice was probably the first musical instrument. Songs are generally made up of two intrinsically connected parts: poetry (which is in the form of lyrics) and melody. A very good complementary relationship between the poetry/ lyric and the melody can give birth to a beautiful song, capable of communicating the complete emotion of the composition to the audience. The proper fit between these parts seems to be made by acoustic features that encompass the relationship between them, representing two fields of sonic communication: musical and verbal communication [1]. While lyrics convey semantic meaning, melody enhances its emotional intention, filling informational gaps and enhancing the significance of poetic meaning of lyrics that otherwise would be incomplete. Within a song, the melodic expression is universal but the lyrical expression is not. Lyrics grow with help of language and as we move across the globe languages vary. Without proper understanding of a language it is impossible to understand

The Journal of the Acoustical Society of America

Published Online: 02 December 2020

Lyrics on the melody or melody of the lyrics?

The Journal of the Acoustical Society of America **148**, 2699 (2020); https://doi.org/10.1121/1.5147477

Archi Banerjee

- IIT Kharagpur, Kharagpur 721302, India, archibanerjee7@gmail.com

Souparno Roy

- Jadavpur Univ., Kolkata, India

Priyadarshi Patnaik

- IIT Kharagpur, Kharagpur, India

Vol. 14 CIC 2016 Special Issue International Journal of Computer Science and Information Security (IJCSIS)
https://sites.google.com/site/ijcsis/
ISSN 1947-5500

Chaos Based Study on Association of Color with Music in the Perspective of Cross-Modal Bias of the Brain

Chandrima Roy
Department of Electronics &Communication Engineering
Heritage Institute of Technology
Kolkata, India
chandrimaa.roy@gmail.com

Souparno Roy[1], Dipak Ghosh[2]
[1]Researcher, [2]Professor Emeritus
Sir C.V. Raman Centre for Physics & Music
Kolkata, India
deegee111@gmail.com

***Abstract*— The relationship between color and music as part of the complex system consisting of visual and auditory domain has not yet been systematically investigated. As both color and music had derived and evolved their forms from the nature that we, humans, perceive through our senses and as both the forms are processed in the same part of the human body, i.e., the brain, it will not be a wildly invalid assumption that they share a similarity in perception. Needless to say that color and music both have strong impact on emotion and feelings & also a few studies have been reported in literature to explore causal relationship between color and emotion. This work reports a neuro-cognitive study on response of brain to different color stimuli. Red, Green, Blue: three primary colors utilizing the signals of electro-encephalograms and multi-fractal methodology to access the degree of complexity with the help of quantitative parameter. In this study the correlation between emotional arousal and the effect of audio and visual stimuli has been studied. This investigation explores the problem from a new perspective. 15 participants were asked to hear 6 different music pieces (each of 30 second duration). The type of emotion elicited by different music pieces were identified by the participants from a given collection of possible emotional responses. Then they are asked to assign a color associating the emotion from a given color wheel (structured according to Munsell color system). Each color, associated with a particular music piece, is a mixture of specific Red, Green and Blue values (RGB triplet) and has a specific HEX number (hexadecimal representation), which is recorded for each response. Then, the musical pieces used were further zoomed with the help of fractal technique to identify different emotions related to music in a quantitative measure. Here, to analyze the complexity of the sound signal (which are non-stationary and scale varying in nature), we have used Multifractal detrended fluctuation analysis (MFDFA), which is capable of determining multifractal scaling behavior of non-stationary time series. Hence, with the data collected, we can correlate color, emotion and music quantitatively.**

Keywords-music, color, MFDFA, RGB triplet, Hexadecimal representation

I. INTRODUCTION

The correlation between color and music with effect causing emotional arousal has always been an integral part of interest for researchers. It was this thought that made none other than Sir Issac Newton curious enough to propose such a correspondence in his book '*Opticks*' back in eighteenth century. Since then, researchers have attempted to identify systematic links between music and color. Perhaps the most direct connection comes from the fascinating phenomenon of music–color synesthesia. Studies also show that non-synesthetic people also have music-to-color associations. So why and how this music-color association works in human? It is strongly suspected that emotion plays a key role in mediating these two stimuli in the brain. That is to say, both color and music has similar emotional qualities that inspire arousal in a similar manner. Why so? The reason is: music and color have been found to instigate emotional arousal time and again. The strong relation between music and emotion has been repeatedly reported in various studies. Music has been shown to affect the emotional state across age, culture and language boundaries. The mood a song induces is so reliable that music is often used as a mood-inducer in psychological studies. Similarly, association of color with emotions is also reported in previous literatures. Not only this, the complexity in bio-signals using color as a visual stimulus has been found to be higher than that of music. In light of these, this study is conducted to find whether a quantitative correlation can be found between music, color and emotion since both the stimuli is connected to emotion in a consistent manner.

This study provides new data in regard to association of color to emotion. Main outcomes may be summarized as follows: the weighted average of emotion 'Joy' is higher in clip 1 whereas, the emotion 'anxiety' is prevalent in clip 3 and 4. Also, participants reported 'romantic' emotion in clip 6. The respective color choices made by participants show a particular trend in these music pieces. The average values of Green and Red are found to be higher in case of emotions- Joy and Anxiety respectively. Again, the clip 6 (romantic) has an average higher Red value than Green or Blue. But, comparing it with the clip corresponding anxiety, we found that even though the Red value is high in both, anxiety is far more pronounced and easily associated with high Red values. Same

Creativity and Cognition in Art and Design

Editors

DR. ANEESHA SHARMA
DR. JAMUNA RAJESWARAN

राष्ट्रीय डिज़ाइन संस्थान
NATIONAL INSTITUTE OF DESIGN

NATIONAL INSTITUTE OF MENTAL HEALTH AND NEUROSCIENCES

BLOOMSBURY
NEW DELHI • LONDON • OXFORD • NEW YORK • SYDNEY

First published, 2017

ICCCAD gratefully acknowledges the support received from Tata Trusts in the publication of this book

BLOOMSBURY PUBLISHING INDIA PVT. LTD.
New Delhi London Oxford New York Sydney

ISBN: 978-93-86349-88-0

10 9 8 7 6 5 4 3 2 1

Published by Bloomsbury Publishing India Pvt. Ltd.
DDA Complex LSC, Building No. 4, 2nd Floor
Pocket 6 & 7, Sector C
Vasant Kunj, New Delhi 110070

Cover design by Aneesha Sharma
Logo design by Aneesha Sharma, Vinayak Hegde

Laser Typeset by FORTUNE GRAPHICS, Naraina, New Delhi

Printed at ANVI COMPOSERS, Paschim Vihar, New Delhi

Ambiguity Creates: A Quantum Leap Interpretation

Souparno Roy, Archi Banerjee, Ranjan Sengupta and Dipak Ghosh
Sir C.V. Raman Centre for Physics and Music, Jadavpur University, Kolkata
E-mail: thesouparnoroy@gmail.com

ABSTRACT

Creativity, defined as 'the tendency to generate or recognize new ideas or alternatives and to make connections between seemingly unrelated phenomena', is too vast a horizon to be summed up in such a simple sentence. The extreme abstractness of creativity makes it harder to quantify in its entirety. Yet, a lot of efforts have been made both by psychologists and neurobiologists to identify its signature. A general conformity is expressed in the 'Free association theory', i.e. the more freely a person's conceptual 'node's are connected, the more divergent thinker (also, creative) he or she is. Also, tolerance of ambiguity is found to be related to divergent thinking. In this study, we approach the problem of creativity from a theoretical physics standpoint. Theoretically, for the initial conceptual state, the next 'jump' to any other node is equally probable and non-deterministic. And to study such a non-deterministic system, Quantum theory has been proven the most successful, time and again. We suggest that this collection of nodes form a system which is likely to be governed by quantum physics and specify the transformations which could help explain the conceptual jump between states. Our argument, from the point of view of physics is that the initial evolution of the 'creative process' is identical, person or field independent. To answer the next obvious question about individual creativity, we hypothesize that the quantum system, under continuous measurements (in the form of external stimuli) evolves with chaotic dynamics, hence separating a painter from a musician.

Keywords: *Creativity, Divergent Thinking, Quantum, Ambiguity*

INTRODUCTION

"Others have seen what is and asked why. I have seen what could be and asked why not." - Pablo Picasso

From the very dawn of civilization, Creativity has been inspiring and reshaping human existence continuously. It was creativity that gave rise to the likes of Picasso, da Vinci, and Einstein - who, with their endless wonders on and off the paper, changed the course of human history and civilization time and again. We, in turn, have strived to understand the experiences of them and have questioned what, if anything, we ourselves have in common with these amazing individuals. Creativity is the development of new ideas and original products in a novel and appropriate way [1][2]. And theories and ideas about understanding the creative process stem from far back in history since it is a particularly human characteristic [3].

Though it started as far back as late 1800s, the systemic search of creativity blossomed in the twentieth century, where its roots have been searched in the lights of a plethora of diversified disciplines [4]: