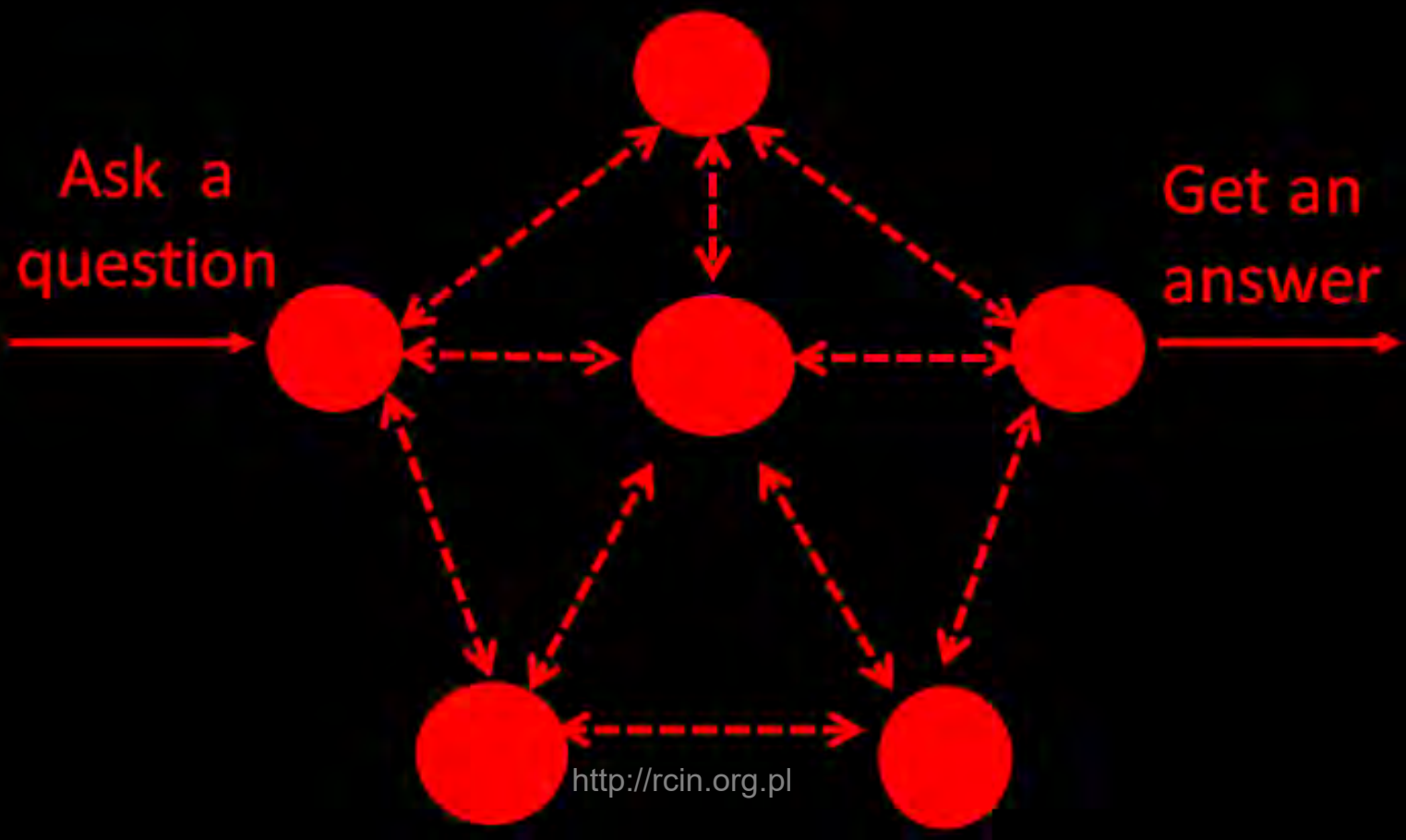
Evolutionary algorithms as a tool for designing chemical computers

ASHMITA BOSE



IChF

FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

Evolutionary algorithms as a tool for designing chemical computers

A thesis

submitted in fulfillment of the requirements
for the award of the degree of

Doctor of Philosophy

in the field

Chemical Sciences

submitted by

Ashmita Bose

Under the Supervision of

Prof. dr. hab Jerzy Gorecki

&

Prof. dr. hab Peter Dittrich

Institute of Physical Chemistry, Polish Academy of Sciences
ul. Kasprzaka 44/52, 01-224 Warszawa

April, 2022

Declaration

I hereby declare that the research included within this thesis was carried out by myself or with support by others included in acknowledgments.

I state that I have exercised care to ensure that the work is original and contains no previously published material or written by another person, except where citations have been made in the text. To the best of my knowledge, the content provided here does not violate any copyrights.

I accept that the Polish Academy of Sciences has the right to use plagiarism detection software to ensure the thesis's legitimacy.

I certify that no part of my thesis has been or will be submitted for obtaining a degree or diploma by the Institute of Physical Chemistry, Polish Academy of Science, or any other educational institution.

This thesis's copyright rests with the author, and no information derived from it may be published without the author's consent.

ASHMITA BOSE

WARSAW,POLAND

Dedicated to my grandmother

ACKNOWLEDGMENTS

The journey towards Ph.D. was a unique one. The journey is filled with surprises, sometimes pleasant and sometimes challenging. The "sometimes pleasant" moments gifted good memories, and the "challenging" moments shaped my life for good. It has been a journey where the easy times have been enjoyable and challenging times have given me opportunities to learn new things.

First and foremost, I would like to thank Professor Jerzy Gorecki. This journey from being a student to being a researcher would not have been possible without his guidance. I want to thank him for being a fantastic guide and for being so student-friendly. He took the effort to train me and make me learn new skills. Not only studies, but being a foreign national, I got his help almost in every matter (be it studies, accommodation, administration, and literally everything) . I am lucky enough to have met someone as a guide who is always ready to protect his students from every problem.

I want to express my gratitude to Professor Peter Dittrich. The 6-month secondment in his lab was one of the most memorable moments in this Ph.D. journey. His warm nature helped me quickly get integrated into his lab. I am thankful to him for those long hours of discussions that we had about our project. I would also like to express my gratitude to my labmates (both from ICHF and Jena University). I would like to mention Stephan Peter (Jena University), who has helped me immensely during my stay in Germany. I am very much thankful to Richard Loffler for always being a source of positivity. And to have always guided me whenever I was in any dilemma. I would like to especially mention Jeel Raval, who has been my colleague and my apartment mate. My sheer luck was when an acquaintance turned to be like a family member and a best friend for life.

I would like to thank Prof. Wojciech Gozdz for letting me use his lab computers. During this Covid-times, when I mostly started working from home, it was beneficial for me to use his computers for my work remotely. I would also like to thank Masakazu Kuze from Hiroshima University, who helped me with my experimental work. I am thankful to Prof. Robert Kolos for arranging seminars where we presented our talks and for his valuable suggestions on improving presentation skills. I would also like to express my gratitude towards the administration department for having helped in every step. I want to thank Ms. Aleksandra Bernatek to have helped me with many administrative works related to my stay in Poland. A special mention to Ms. Edyta Slojewska and Ms. Patrycja Niton for their warm and welcoming nature.

It is almost unforgettable how my friends have been a support system to me. Abhiri, Atreyee, Nishat, being themselves doing Ph.D. in their respective fields, we practically have lived this journey experience together.

Last but definitely not least, words fall short of expressing how lucky I feel to have such a wonderful family. My parents and grandparents have always been a blessing to me. My father has been a source of optimism while my mother has been a source of strength. Having an ultimate best friend in my mother and talking to her over the phone every day has always kept me cheerful and motivated throughout this journey. A big thanks to my family for just being always there for me!

FUNDINGS

This research is part of a project that has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 711859.

Scientific work funded from the financial resources for science in the years 2017-2022 awarded by the Polish Ministry of Science and Higher Education for the implementation of an international co-financed project.

SUMMARY

Conventional computing technology based on silicon chips still holds the dominant position in modern information processing. The chips achieve high-speed functionality by shrinking component sizes and packing a large number of transistors onto a single integrated circuit. However, it is anticipated that Moore's law describing the progress of semiconductor technology will break down, and alternative information processing technologies are necessary to keep up with the progress. The motivation of my research was to move towards unconventional computing, consider information processing with a chemical medium, and design a "chemical computer".

Results on information processing with various types of chemical media have been reported by many groups. Successes in forming chemical logic gates with Belousov-Zhabotinsky (B-Z) reaction made this reaction a suitable candidate to work with. However, the bottom-up approach from gates to complex devices is not appropriate when we don't know exactly how to combine simple components in order to obtain the anticipated results. Hence, in this thesis, I applied the top-down strategy to design information processing with networks of chemical oscillators. My study is based on numerical simulations. A realistic two-variable Oregonator model was used to simulate the chemical oscillations. Specific reactions between activators of different oscillators were introduced to describe interactions between them. The photoinhibition of B-Z reaction was used for the external control over individual oscillators and for data input. The number of activator maxima observed on a selected oscillator in the network was considered as the output. Oscillator networks were optimized using an evolutionary algorithm to perform the required information processing function with the maximum mutual information or accuracy on the training dataset of cases.

The thesis is concerned with the applications of networks of interacting chemical oscillators to solve two geometrically oriented and two medically inspired problems. Regarding geometrically oriented problems, we demonstrated that a network of three interacting chemical oscillators could be optimized to predict with 95% accuracy the color of a randomly selected point of a Japan flag on the basis of point coordinates. Another reported result is the design of a 6-oscillator network to distinguish two branches intertwined spirals apart. I demonstrated that the network with optimized parameters could separate points belonging to different branches with an accuracy greater than 80%. I also investigated the influence of the geometry of interactions among oscillators on network accuracy. For the spiral separation problem, the networks with simple interactions gave higher accuracy. Concerning medically oriented

problems I demonstrated that a simple network of chemical oscillators could be trained and optimized to diagnose schizophrenic patients with an accuracy of 82% on the test dataset of 45 ill and 39 healthy patients. For this problem, I investigated networks of different numbers of nodes. It was found that the accuracy did not significantly dependent on network size. I also presented the strategy of voting within a group of different networks to distinguish if the patient had schizophrenia or not. This strategy increased the accuracy. An idea of a three-stage schizophrenia diagnosis, based on separation of the original EEG signal into shorter time domains, was also presented, leading to an accuracy of 90%. In another considered medically oriented problem, networks of three interacting chemical oscillators were trained to predict the relationship between the gene expression data and the bortezomib drug efficiency for cancer patients with multiple myeloma. The networks were optimized on a training set containing data from 239 therapies. Networks using a single gene expression value showed 71% accuracy in differentiating between non-responsive and responsive patients. When the results of single-gene networks were combined into the concilium with the majority voting strategy, the accuracy of predicting the patient's response to bortezomib drug therapy increased to ~ 85 %.

In the final part of the thesis I presented experimental results that can be important for building an operational chemical computer in the future.

Streszczenie

Konwencjonalna technologia przetwarzania informacji oparta na krzemowych układach scalonych ma obecnie dominujaca pozycje. Wysoka funkcjonalność mikroprocesorów uzyskano dzieki zmniejszeniu rozmiarów komponentów i upakowaniu wielkiej liczby tranzystorów w pojedynczym układzie scalonym. Przewiduje sie jednak, że prawo Moore'a opisujace postep technologii półprzewodnikowej w końcu przestanie obowiazywać i, że alternatywne technologie przetwarzania informacji beda niezbedne dla podtrzymania tempa rozwoju technik informatycznych. Motywacja dla moich badań był rozwój niekonwencjonalnych metod przetwarzania informacji, a w szczególności wykorzystanie w tym celu reakcji chemicznych i zaprojektowanie komputera chemicznego".

Badania nad przetwarzaniem informacji przy wykorzystaniu różnych mediów chemicznych sa prowadzone w wielu ośrodkach. Sukcesy z eksperymentalna realizacja chemicznych bramek logicznych w oparciu o reakcje Biełousowa-Żabotyńskiego (B-Z) skłoniły mnie do wykorzystania tej reakcji w badaniach, których rezultaty opisuje w rozprawie. Jednak strategia oddolnego (bottom-up) budowania układów chemicznych nietrywialnie przetwarzajacych informacje jest trudna w realizacji, gdyż nie wiemy dokładnie, jak połaczyć proste elementy, aby uzyskać oczekiwane wyniki. Dlatego w tej pracy zastosowałem strategie odgórna (top-down) do projektowania sieci oscylatorów chemicznych przetwarzajacych informacje. Moje badania opieraja sie na symulacjach numerycznych. Do symulacji oscylacji chemicznych wykorzystano realistyczny model Oregonatora z dwiema zmiennymi. Wprowadzono specyficzne reakcje miedzy aktywatorami różnych oscylatorów, aby opisać oddziaływanie miedzy nimi w sieci. Idea inhibitowanej światłem reakcji B-Z została wykorzystana do zewnetrznej kontroli poszczególnych oscylatorów oraz do wprowadzania danych wejściowych. Założyłam, że odpowiedzia sieci na postawiony problem jest liczba maksimów aktywatora zaobserwowanych na wybranym oscylatorze w sieci. Prezentowane sieci oscylatorów zostały zoptymalizowane za pomoca algorytmu ewolucyjnego, tak aby uzyskać maksymalna informacje wzajemna albo dokładność dla wybranego zbioru przykładowych rozwiazań.

Wyniki składajace sie na rozprawe dotycza zastosowania sieci oddziałujacych oscylatorów chemicznych do rozwiazania dwóch problemów geometrycznych oraz dwóch zagadnień inspirowanych diagnostyka medyczna. Jeśli chodzi o problemy zorientowane geometrycznie, wykazaliśmy, że zoptymalizowana sieć trzech oddziałujacych oscylatorów chemicznych potrafi przewidzieć kolor losowo wybranego punktu na fladze

Japonii na podstawie jego współrzednych z 95% dokładnościa. Innym opisanym w rozprawie wynikiem jest opisanie sieci złożonej z 6 oscylatorów, która potrafi rozróżnić punkty zlokalizowane na spirali utworzonej przez dwie splecione ze soba gałezie. Wykazałam, że sieć o zoptymalizowanych parametrach potrafi przypisać gałaź do losowo wybranego punktu należacego do spirali z dokładnościa wieksza niż 80%. Zbadałam również wpływ geometrii oddziaływań miedzy oscylatorami na dokładność sieci oscylatorów rozwiazujacej problem separacji spirali. Wyższa dokładność otrzymano dla sieci z prostsza struktura oddziaływań. W rozprawie rozważono dwa problemy inspirowane diagnostyka medyczna. Pokazano, że prosta sieć złożona z 6 oscylatorów chemicznych można nauczyć i zoptymalizować tak, aby diagnozowała pacjentów, którzy moga chorować na schizofrenie z dokładnościa ok 82% na próbce 45 chorych i 39 zdrowych osób. W przypadku problemu diagnozowania schizofrenii testowałam sieci o różnej liczbie wezłów. Stwierdziłam, że dokładność diagnozy nie zależała istotnie od wielkości sieci. Aby podnieść dokładność testu czy pacjent choruje na schizofrenie, czy nie przedstawiłam strategie "głosowania" z jednoczesnym wykorzystaniem kilku sieci oscylatorów. Ta strategia poprawiła dokładność diagnozy. Przedstawiłam też idee trzystopniowej diagnostyki schizofrenii oparta o podział sygnału EEG na podprzedziały czasowe. Pozwoliło to podnieść dokładność diagnozy do 90%. Drugim rozważanym problemem medycznym było opracowanie sieci złożonej z trzech oddziałujacych oscylatorów chemicznych, tak, aby przewidywała zwiazek miedzy danymi dotyczacymi ekspresji genów, a skutecznościa działania lek bortezomib u pacjentów chorych na raka (szpiczaka mnogiego). Sieci zostały zoptymalizowane na podstawie danych z 239 terapii. Sieci wykorzystujace wartość ekspresji pojedynczego genu jako input wykazały 71% dokładność w różnicowaniu pacjentów niereagujacych i reagujacych na bortezomib. Kiedy połaczono przewidywania sieci zoptymalizowanych na znalezienie korelacji miedzy wynikiem terapii a ekspresja pojedynczego genu w konsylium ze strategia głosowania wiekszościowego, wtedy trafność przewidywania wyniku terapii pacjenta na leczenie bortezomibem wzrosła do $\sim 85\%$.

W końcowej cześci rozprawy omówiłam wyniki prac eksperymentalnych, które moga mieć znaczenie dla budowy w przyszłości działajacego komputera chemicznego opartego o sieci oddziałujacych oscylatorów.

Contents

List of figures

List of frequently used symbols

u	Concentration of activator in the two-variable Oregonator model
v	Concentration of inhibitor in the two-variable Oregonator model
ε	Scaling parameter in time evolution equations of two-variable Oregonator model.
f	Stoichiometric coefficient used in two-variable Oregonator model
$\phi()$	Illumination in two-variable Oregonator model
α	The rate of decay of activator u
β	The coupling factor between two oscillators
$s_{j,k}$	Represents the interaction between j^{th} and k^{th} oscillator
$I()$	Mutual information
Z_k	Excitation class; stores the number of oscillations of k^{th} oscillator
Z_O	Output class; contains the labels related to input cases
$H()$	Shannon entropy
$p_{i,n}$	i^{th} predictor value of n^{th} case
t^j_{illum}	The time of illumination of j^{th} oscillator
t_{max}	The total time of simulation
t_{start}, t_{end}	Parameters used to transform input value to illumination time of an input oscillator

List of tables

Chapter 1

Introduction

The chapter gives an insight into the world of artificial intelligence(AI) and touches the topic of conventional and unconventional computing as one of the most important products of AI. Further, it describes the motivation and the research objective of my work.

Although conventional computers are still dominating the world, they have their own limitations related to the design complexity or their disadvantage in working under certain extreme conditions. Keeping in mind the limitations, alternative computation methods have been introduced. I postulate that information processing can be performed by a network of chemical oscillators coupled with reactions. I demonstrated the usefulness of such an approach to solve different geometry and medically oriented problems successfully.

1.1 Computing and artificial intelligence

Artificial Intelligence (AI) refers to the field of science aimed at providing *objects* with the capacity of performing functions such as logic reasoning, planning, learning and perception [1]. This *object* can be anything: a machine or any system that can be taught to be intelligent. Thus AI is focused on training an *object* to replicate human intelligence. A schematic representation of how an *object* is trained to perform AI has been shown in Figure 1.1. The basic steps behind it are: first, perceive the things happening in the environment, then interpret this information by reasoning the knowledge collected and finally decide the best action to be taken to achieve a set goal [2]. In Figure 1.1, the sensor is shown as an eye. However, it can be anything like an input device or a sensor with physical quantities like temperature or pressure. In general, a sensor should be something that can perceive the data present in the environment which are relevant to achieve the goal given to an *object* to perform AI. Then based on the collected data or information by the sensor, the system will reason on this acquired knowledge and will understand what should be the best action that can be taken to achieve the desired goal. Once the action has been decided, the system is ready to perform it by actuators. In Figure 1.1 the actuators are depicted as arms and legs. But it necessarily do not have to be some physical body. An AI system can also simply output a signal that dictates the action that has been decided [3].

AI is a vast field with many subsets. However, here I would like to discuss about the two most commonly used subsets. They are machine learning [4] and deep learning [5]. Machine learning is simply based on the idea to "train" a machine with respect to the information that it gathers from a certain dataset. They gather these informations by locating some kind of patterns in a dataset [6]. And they remember the patterns to later use this as an acquired knowledge, for processing of future datasets. Machine learning technique uses statistical analysis to find the patterns. Some example of commonly used machine learning techniques are : regression trees [7],logistic regression analysis [8], linear regression analysis [9] to name a few.

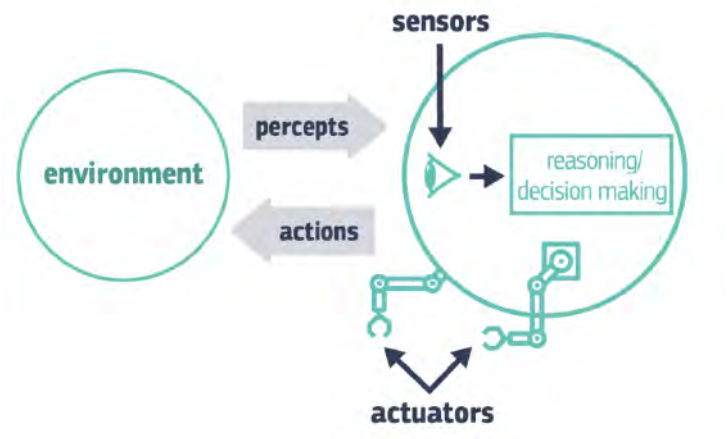


Figure 1.1: A schematic representation of an AI system is shown [3]

Deep learning is a subset of machine learning. Like machine learning, it also finds patterns among the past obtained information in the form of data. The method of deep learning describes an algorithm that analyzes data with a logic structure similar to how a human would draw conclusions [10](first perceive, then analyze and finally act). The idea of designing a deep neural network is inspired by the living creatures who have a control unit due to which they can perform many important tasks (vision, hearing, sensory control, mobility control etc.). The method is called "deep" as a number of layers are added in the creation of a network (Figure 1.4). Layers in gen-

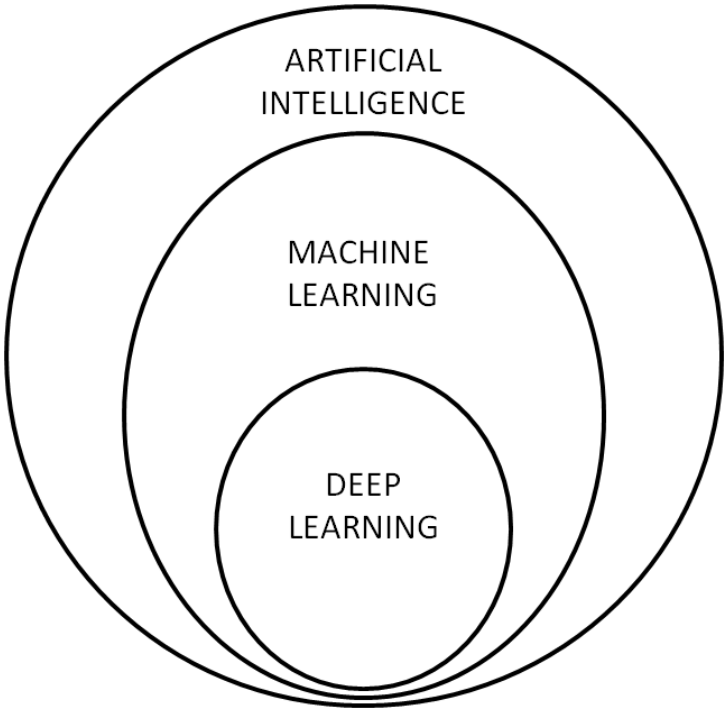


Figure 1.2: A schematic representation of the relationship between the important fields of AI.

eral refers to the collection of so called "neurons" found in living organisms, connected to each other to perform certain functions together. The more the number of layers, deeper the network is formed. The control unit or brain in a living being contains a dense number of neurons which form complex networks by interconnecting with each other at points called synapses. The neuron is the fundamental building block of these neural networks. The axon, cell body, and dendrite are the three primary components of a neuron. The neuron's center, the cell body, contains genetic information and also serves as a source of energy. Dendrite [11] is a fibrous root that branches out from the cell body and has the primary purpose of gathering information from other cells in the body and transporting it to the cell body. An axon [12] is the neuron's long tail-like structure that carries information away from the cell body to synapses of the neurons (Figure 1.3) [13]. The basic function of a neural network in a living organism is as follows: The neurons form a mesh that keeps them connected. The dendrites of a neuron accept input from sensory organs or inputs from the surrounding environment. Electrical impulses are created as a result of the inputs [14]. The signal is then carried along the axon's length. There are certain branches like structures at the end of the axon that pass on the signals to the dendrites of other associated neurons [15]. As a result, information in a neuron flows from the dendrite to the cell body, then to the axon, and lastly to the synapses' end tips. This high efficiency of neural networks in living organisms to encode, process and learn information [16] has encouraged researchers to design deep neural networks. The basic idea of designing this is to provide some information to a certain layer, also can be called as the input layer (with the analogy of neural network in a living organism, the input neurons are those which receive information directly from the cell of the body) and extract output from the output layer. The hidden layers or the in between layers perform operations over the inputs and pass it forward to the output layer. In our work, we have created a model with the inspiration of deep neural network. However our network is made up of chemical oscillators and interact with each other by internal coupling and inflow/outflow of input/output are designed keeping in mind the chemical properties.

Computers also perform artificial intelligence tasks and thus we use AI almost every day. Though the human brain is better in qualitative analysis than computers, still in quantitative analysis computers almost always beat the human brain by their speed [17]. Computers were created to try and replace the human intellectual effort [18]. Conventional computers are based on the classic principle that a "switch" can be in ON or OFF state. Large amount of information can be coded in sequences of binary states. The switches are the transistors and the logic 1 is represented by

a high electrical potential and the logic 0 is represented by a low one. All of the letters, characters and numbers are represented by long strings of 0's and 1's [19]. Conventional computers are based on silicon chip technology. They achieve high-speed functionality by shrinking component sizes and packing a large number of transistors onto a single IC (integrated circuit) chip [20]. According to Moore's law the number of transistors placed on IC chips has doubled approximately every two years [21]. This gives rise to design complexity. If the trend keeps on continuing it will create complex design challenges. Moreover the physical properties of IC chips also impose certain restrictions in which they can operate, eg. biological ones (it is difficult to introduce IC chips on a cellular level as it can cause bio-compatible issues and has the potential to influence biological processes) or under extreme conditions eg. high

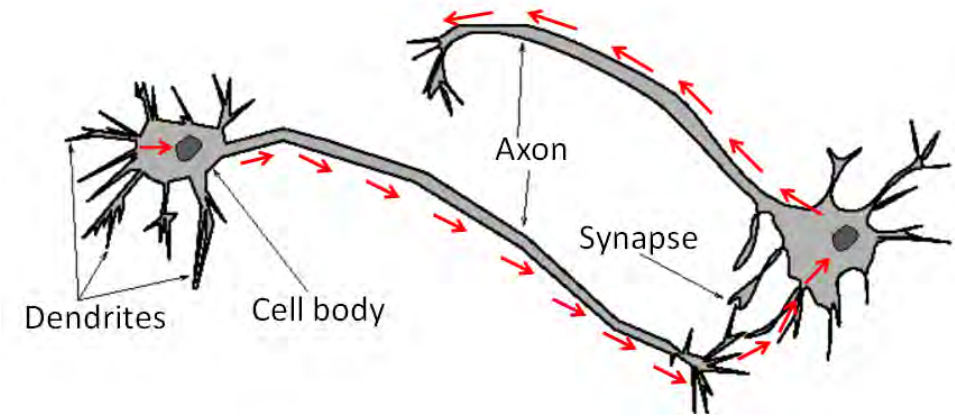


Figure 1.3: Schematic representation of biological neuron [15]. The red marked arrows show the flow of information from one neuron to another.

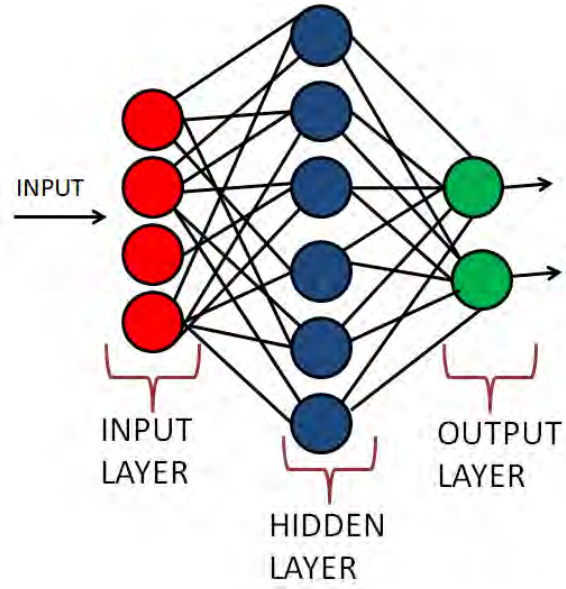


Figure 1.4: A schematic representation of a simple deep neural network comprising of input, hidden and output layers.

temperature [22]. All these factors led to a new research arena to search for another way to perform computing. Now, computing is a term that can sometimes get easily misunderstood. Mostly people associate this term with computers only. Computing simply means processing certain inputs to get desired results as output. For example, suppose the object is a leaf. A leaf can modify it's orientation to optimize sunlight [23]. After harvesting sunlight as it's input, the leaf processes this energy by passing it from one chromophore (photosynthetic molecule) to another unless it reaches the "reaction center" where the photosynthesis takes place [24]. Finally the solar energy is transferred into a chemical one. Thus objects around us can perform some kind of computing.

Unconventional computing

As the name suggest unconventional computing is unlike conventional computers. Now as we know that any object can perform some kind of computing. This drives us to this new topic where we don't use "conventional" IC chips for computational purposes but try and perform computation in "unconventional" ways. The conventional computing has an already existing programming model providing instructions to perform desired tasks, but for unconventional computation we really have to search for the perfect programming that can give us an acceptable solution [25]. We discuss the concept of "embodiment" for unconventional computation. Embodied computation can be defined as a computation in which physical characteristics of a system is an important factor to perform information processing with or the computation is having it's physical realization in the environment [26]. Like the example of a leaf performing photosynthesis: the leaf is itself embodied with it's physical characteristics to track sunlight and process it (by passing the energy from one chromophore to another) all by itself. Or a neural network which is a physical realization of the neural network in living beings.

The foundation of many algorithms for unconventional computing are inspired by physics: for example, "simulated annealing search algorithm", where the physical process of metal being cooled down slowly to achieve it's ground state of minimum energy has been used. Here the minimum energy state was used as an analogy to the desired solution and the temperature acted as a control parameter for the movement around the search space to achieve a particular solution. With higher temperature the search space is widened and giving the algorithm variety of options to choose the optimal one. Later as the search proceeds and the algorithm is nearer to finding it's desired solution, the temperature is decreased so that the algorithm reduces the extent of it's search to converges to a minimum [27] [28]. Another example well known

is quantum computation [29].

There is also population based computations. This kind of algorithm is inspired by biological systems. The basic idea behind this kind of algorithm is to consider a huge population, where each individual is characterized by a fitness value corresponding to a particular task we want to achieve in a search space. We get rid of those individuals with lower fitness values and move towards a population of fit individuals which leads us to higher fitness value in the search space. Some examples of already established algorithms are evolutionary algorithm [30] and swarm optimization [31]. As our work uses an evolutionary algorithm to design a chemical computer, a further detailed explanation of it is given in Chapter 2.

Another type of unconventional computing is network based computation. The realization of it can be drawn straight from the biological point of view of how our nerves are connected to each other forming a network and interacting with each other to process information (as explained in the concept of deep neural network). Mostly networks-based algorithms are the learning algorithms. The network gets trained over a dataset to find a pattern. The algorithm is designed in a way that the network parameters are adjusted so that it moves in the direction of our desired result.

At this point if we compare the conventional computer and neural network based computation, we can say that the network based computation can do many things together. It is designed from the onset to work parallel, whereas conventional computer processing is a sequential one and tasks follow one another. Thus the computation speed of neural network is high. The traditional computer learns to perform a certain task based on certain well defined rules in the form of a definite algorithm whereas network based algorithms are created in a way that it self learns from past experiences. Thus while the conventional computer can only learn based on the rigid steps mentioned in the algorithm, neural network learning is more flexible. It can upgrade it's learning lessons by gaining new experiences in the form of different inputs and acquiring information from past events [32].

1.2 Research objective

My main objective is to design an unconventional computer that can perform non-trivial information processing. I design a chemical computer by considering a certain chemical medium as its physical representation. My thesis is concerned with network based computation. The network is formed by interacting chemical oscillators. Previously published papers [33] [22] considered oscillators described by the oversimplified event-based model. I apply well established Oregonator model and use chemical re-

actions to describe interactions. The population search algorithm is used to find the best parameters for the networks to perform different information processing tasks and have successfully shown progress towards designing a chemical computer.

1.3 Motivation for the present research work

The creation of chemical reactions based artificial intelligence is especially interesting as it is used by living organisms to perform information processing in everyday life. I present an idea of designing a chemical computer by using the Belousov-Zhabotinsky (B-Z) reaction as a computing medium. At specific conditions, a spatially distributed medium of this reaction can be locally excited and the excitation can propagate in space. This type of behaviour resembles the propagation of nerve impulse in living organisms and hence many researchers designed neuron-like chemical computers using this medium. [34] [35]. We think that networks of interacting chemical oscillators represent more realistic models of biological neural computing than typical artificial neural networks with arbitrarily selected activity rules.

Most of all the works on chemical computing that are done with B-Z reaction have been concerned with designing based on the bottom-up approach. However, as the bottom-up approach deals with linking individual components to form a complete system, we should know first how to combine together simple components to obtain the anticipated result. There are many problems for which we can define the input variables and specify the set of anticipated outputs, but we do not know the algorithm that links the input with the output. For example, problems inspired by the needs of medical diagnostic belong to such class. The input information is collected from a number of medical tests and on this basis we are expected to conclude if the patient is healthy or ill. But in many medically oriented problems, we do not know the algorithm that produces the answer on the basis of the input data. For such types of problems the top-down design strategy seems to be more appropriate than the bottom-up one. Unlike the bottom- up approach, in top-down strategy an overview of the system is formulated first (without going into details of any part of it). After that we refine each part unless the model is validated to perform certain task [36]. Here I presented the top down approach to perform information processing with chemical oscillators and have shown how it can solve geometrically oriented and medical problems successfully.

I have used the two-variable Oregonator model [37] to simulate the B-Z reaction and designed networks of chemical oscillators. The networks were trained by using an evolutionary algorithm to solve complex information processing tasks.

1.4 Organization of the thesis

The results presented in the thesis are organized and structured in the form of seven chapters, which are briefly described as follows:

i) **Chapter 1** describes the concept of artificial intelligence and unconventional computing briefly. It tells us about the motivation behind my work.

ii) **Chapter 2** gives information about the chemical medium which has been considered to design our chemical computers. It also gives an insight into information theory and evolutionary algorithm which has been used to train our systems.

iii) **Chapter 3** provides comprehensive details about the model I have used. It tells us about how the model has been modified to design an oscillatory network. The chapter describes how we have introduced inputs to the system and how to extract output from it. The chapter also deals with the information of how we have trained and optimized our system.

iv) **Chapter 4** presents the result on the application of oscillator network on geometrically oriented problems.

v) **Chapter 5** discusses the results we obtained using our system as a classifier. It describes the application of oscillator networks for certain medically oriented problems, such as: diagnosing schizophrenia and predicting the drug response in cancer treatment.

vi) **Chapter 6** summarizes my experience in experiments with a chemical medium that can be used for oscillator networks.

vii) **Chapter 7** concludes the thesis with the overall findings of the present research work.

Chapter 2

The basic concepts of chemical media used for information processing, elements of information theory and evolutionary optimization

This chapter introduces the basic concepts of the topics relevant to my work in the thesis. It provides explanation about Belousov-Zhabotinsky (B-Z) reaction, both quantitatively and qualitatively. Further, the basic concepts of information theory has been explained. It gives an insight into Shannon entropy and mutual information. In the end, the concept of the evolutionary algorithm, which was used to design the chemical computer is explained.

2.1 A chemical medium with potential information processing applications:

As our main aim is to design a chemical computer, this section sheds light on the qualitative and quantitative explanation of the chemical medium I focused on. I also discuss some noteworthy information processing work already reported as performed by this chemical medium.

2.1.1 An introduction to Belousov-Zhabotinsky (B-Z) reaction

In the year 1950, B.P. Belousov was working to model catalysis in the Kerbs cycle by using metal ion cerium instead of protein-bound metal ion which is common in enzymes of living cells [38]. It was expected to see how the reaction proceeds towards equilibrium, using the visible sign of progress as a change in colour of the solution from being colourless (cerium in the reduced state) to being pale yellow (cerium in the oxidized state). However to his surprise, the reaction turned colourless to yellow and back to yellow and again colourless, oscillating many times, for as long as an hour, between the oxidized and reduced state of the cerium ions [39]. But the investigation was never published. The reviewers thought that this chemical oscillator violates the 2^{nd} law of thermodynamics. Moving away from the equilibrium position will require the Gibbs free energy to increase, BUT it must always decrease [38]. However that was not the case. A better model for understanding chemical oscillations is the grandfather clock. We can notice that the hands of the clock pass twice daily through the same position, but the energy stored in the elevation of weights decreases as the clock runs. Similarly, in the oscillating chemical reaction, the reagent concentration passes repeatedly through the same value, but the energy-dissipating reaction proceeds towards equilibrium through concentration oscillations [40].

Chemical mechanism of B-Z reaction

The B-Z reaction is rather a complex reaction to understand. The model that can explain the complex kinetics of this reaction is the FKN model [41]. This model depicts the basic steps contributing to the oscillating mechanism of the reaction. For the reaction to happen we need : an oxidizing agent - $NaBrO_3$, an organic substrate-usually the malonic acid (Adamčíková and Ševčík [42] demonstrated oscillations in a system where NaH_2PO_2 was used instead of malonic acid), an inorganic acid-

H_2SO_4 that is a source of H^+ ions and a catalyst - cerium or ferroin. Oxidation of Ce^{3+} ion happens by $BrO^{\cdot}$ radicals in the presence of H^+ ions. The concentration of Ce^{4+} increases with the accumulation of bromomalonic acid. The reduction of Ce^{4+} occurs due to the production of Br^- from the bromoderivative of malonic acid [43]. The detailed mechanism can be explained by dividing it into three different processes. The processes are as follows:

Process 1 : In the initial mixture of reagents, consumption of Br^- ions take place by BrO_3^- and results in the formation of $HBrO_2$. This $HBrO_2$ further consumes Br^- ions and forms $2HBrO$. $HBrO$ then reacts with Br^- and gives Br_2 as a product. For all these above mentioned steps to occur, an ample amount of Br^- ion is required. Hence Process 1 can only occur when the concentration of Br^- ion is high.

$$BrO_3^- + Br^- + 2H^+ \longleftrightarrow HBrO_2 + HOBr$$

$$HBrO_2 + Br^- + H^+ \longleftrightarrow 2HOBr$$

$$HOBr + Br^- + H^+ \longleftrightarrow Br_2 + H_2O$$

Process 2 : At low concentration of Br^-, $HBrO_2$ changes to $BrO_2^{\cdot}$ radicals. The $BrO_2^{\cdot}$ radical then gets reduced by Ce^{3+} and changes to two $HBrO_2$, which in turn changes Ce^{3+} to Ce^{4+} ion. Thus two $HBrO_2$ molecules are produced by one molecule of $HBrO_2$ through self catalytic positive feedback reaction. This self catalytic reaction gets suppressed when all Ce^{3+} gets consumed.

$$2HBrO_2 \longleftrightarrow HOBr + BrO_3^- + H^+$$

$$BrO_3^- + HBrO_2 + H^+ \longleftrightarrow 2BrO_2^{\cdot} + H_2O$$

$$BrO_2^{\cdot} + Ce^{3+} + H^+ \longleftrightarrow Ce^{4+} + HBrO_2$$

Process 3 : Ce^{4+} gets reduced to Ce^{3+} and Br^- is generated through subset process involving organic reaction [44].

$$Br_2 + CH_2(COOH)_2 \rightarrow BrCH(COOH)_2 + Br^- + H^+$$

$$6Ce^{4+} + CH_2(COOH)_2 + 2H_2O \rightarrow 6Ce^{3+} + 2CO_2 + HCOOH + 6H^+$$

$$4Ce^{4+} + BrCH(COOH)_2 + 2H_2O \rightarrow 4Ce^{3+} + 2CO_2 + 5H^+ + Br^- + HCOOH$$

$$Br_2 + HCOOH \rightarrow 2Br^- + CO_2 + 2H^+$$

A schematic representation of different processes happening in B-Z reaction has been illustrated in Figure 2.1

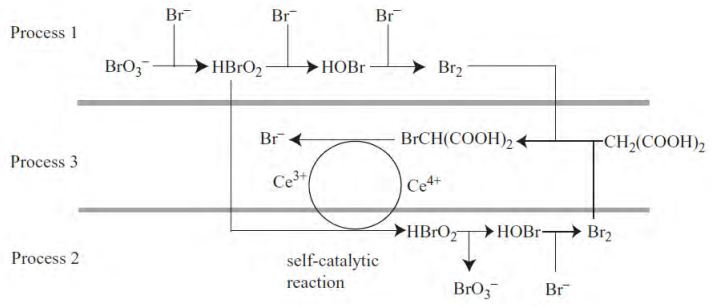


Figure 2.1: Schematic representation of different processes among the reagents in a B-Z reaction [45]

The switching mechanism (Figure 2.1) between the processes can be understood as : Suppose the Br^- concentration is high. Thus process 1 is dominant. Due to the consumption of Br^- ion in the process, its concentration gradually decreases. Thus eventually, Process 2 becomes dominant and Ce^{3+} gets oxidized to Ce^{4+}. Again by Process 3 reduction of Ce^{4+} to Ce^{3+} takes place and also Br^- ion gets formed. With the increase in Br^- concentration , again Process 1 gets dominant.

Chemical oscillations in B-Z reaction

The above mentioned mechanism can clearly explain why it is an oscillating reaction. Let us consider in the start that the concentration of Ce^{4+} ion is high, thus bromination of the organic compound takes place and Br^- ions are formed rapidly. So autocatalytic oxidation of Ce^{3+} to Ce^{4+} gets inhibited. Hence no new Ce^{4+} is formed. Thus slowly the concentration of Ce^{4+} decreases. With the decrease in the concentration of Ce^{4+}, the formation of Br^- ions decreases. When the concentration of Ce^{4+} reaches the lower threshold, a drop in production of Br^- ion occurs. With this decrease in Br^- ion, the autocatalytic oxidation to change Ce^{3+} to Ce^{4+} is again possible. Again when the formation of Ce^{4+} reaches a higher threshold, Br^- ion increases, and the same cycle repeat and keeps on going back and forth [46]. This change can be noted down by the change in colour of the solution from yellow to colourless and vice versa. To have more of stark contrast in the change of colours, Zhabotinsky suggested using ferroin as a catalyst in place of cerium, and the same mechanism of change of Fe^{2+} to Fe^{3+} back and forth, results in oscillation with visible colour change from red to blue and blue to red. That is, the concentration of the reagents spontaneously changes with a time period and we can witness the temporal oscillations of the reagent concentrations.

It may come to mind by taking the first glance of the mechanism that the reaction will always oscillate, however that is not the case. One can imagine that if Process 2 completely and permanently suppresses Process 1 or vice-versa or if the pathways come to steady state, then no oscillations will be observed. Hence to observe the oscillations, we must find the conditions under which the mechanism can exhibit sustain oscillations [39].

In 1970, Zaikin and Zhabotinsky described that a thin unstirred layer of B-Z reagents could be observed in the propagation of waves of oxidation [39]. The identification of factors leading to the formation of spatial phenomena has been apprehended by two different interpretations. The first interpretation of the formation of a spatial wave is the heterogeneity in the system: some foreign particle might be present where the bromine ion gets adsorbed on the surface of the foreign object, due to which at that particular point, the concentration of bromine ions in smaller and chemical wave gets triggered [47]. Another theory is considering the existence of internal local fluctuations of the concentrations of reagents that leads to spatial patterns. The formation of these spatial patterns can be understood as follow: When nucleation of the reaction occurs at a certain point, bromide from the surrounding travel to this point via diffusion. Now, as in this point Br^- ion is dominant, it resets the mechanism and Process 1 starts at this point. In the surrounding, due to deficiency of Br^- ion, Process 2 becomes dominant, which supports the formation of $HBrO_2$. Thus progression of a circular wave outward from the nucleation point gets witnessed. If there exists a favourable condition for the oscillation being continuous, a second circular wave will be formed from the nucleation site, resulting in the development of numerous concentric rings in a "target pattern". Else if the condition does not satisfy for continuous oscillation, then only a single wavefront will expand until it gets annihilated at the edge of container [38]. One such example of spatial pattern formation is : Let us say we created perturbation by touching a silver wire to the B-Z solution with ferroin used as a catalyst. At the contact point the concentration of Br^- ion is low and hence the production of $HBrO_2$ is dominant (observed by a blue dot at the contact surface). After some time the diffusion of $HBrO_2$ takes place expanding the wavefront in the direction of diffusion. With the diffusion, the place from where it started now consists of a higher concentration of Br^- ion (visible via a change from blue to red colour). This gives rise to many targeted spatial wave patterns.

The speed of the wave propagation depends on the frequency of oscillation. With an increase in frequency, the wave speed decreases because with high frequency, the waves follow closely upon each other, due to which the oxidation waves are forced to propagate through a medium that is still recovering from the reduced state, and

hence the oxidation wave travels more slowly in incomplete recovered medium [39].

Our work mainly concerns to the temporal oscillation caused in the B-Z reaction as these oscillations were used to code information to perform artificial intelligence.

Illumination controlled B-Z reaction

In order to design a chemical computational medium with a B-Z reaction, it is a very important thing to control this reaction externally, so that we can modify it according to the computational task we want to perform. As in this thesis we have used the characteristics of temporal oscillations of the reaction as a method to encode different information, we wanted an external control factor that can affect and alter these oscillations.

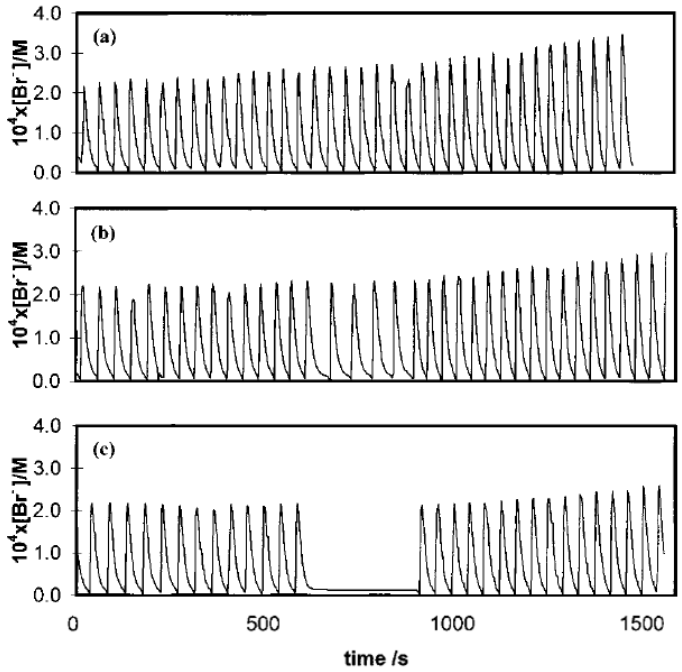


Figure 2.2: Effect of illumination on a B-Z system, for different light intensities. The blue light was switched on at 600s and off at 900 s. Light intensity PMT voltages:(a)1.0 V (B)2.5 V (C) 2.30 V [48]

One of the easiest ways to control oscillation in B-Z reaction is to use ruthenium catalyst and illumination as the control factor. It has been seen that the intensity of blue light can manipulate the oscillation frequency and above the threshold value it can stop oscillations in B-Z reactions. The light intensity controls the formation of bromide ions in the solution which results in affecting the oscillations. It was found that the oscillation period depends on the intensity of light. When the light intensity is low, no effect is observed in the oscillations. With the increase in light intensity, the oscillation period increases (Figure 2.2 (b)) and when the intensity of

light increases above the threshold value, the oscillations seize to exist [48]. The inhibition of oscillations by illumination is explained as follows: When the blue light is flashed over the reaction, the $Ru(II)$ catalyst gets excited to $Ru(II)^*$ and the following reaction takes place: [49]

$$4Ru(II)^* + HBrO_2 + 3H^+ \rightarrow Br^- + 2H_2O + 4Ru(III)$$

In the above reaction $Ru(II)^*$ reacts with $HBrO_2$ and produces Br^-. When $HBrO_2$ gets fully exhausted, a mechanism which was proposed by Kuhnert [50]

$$6Ru(II)^* + BrO_3^- + 6H^* \rightarrow 6Ru(III) + 3H_2O + Br^-$$

contributes to additional Br^- production. With increased concentration of Br^- the system remains in the reduced state and hence no oscillation is witnessed.

2.1.2 Model representing B-Z reaction

The two-variable Oregonator model

FKN model contains many different chemical species, which in turn can be difficult to solve numerically. Hence Field and Noyes [37] came up with a new model consisting of just three variable concentration and still being able to describe the essence of B-Z mechanism. The model was named as "Oregonator model". The model consists of only 5 chemical reactions taken from FKN mechanism.

$$A + W \rightarrow U + P \tag{2.1.1}$$

$$U + W \rightarrow 2P \tag{2.1.2}$$

$$A + U \rightarrow 2U + 2V \tag{2.1.3}$$

$$2U \rightarrow A + P \tag{2.1.4}$$

$$B + V \rightarrow hW \tag{2.1.5}$$

These 5 equation has been written in a simplified way where $A = [BrO_3^-]$, $B = [BrCH(COOH)_2]$, $P = [HOBr]$, $U = [HBrO_2]$, $V = [Ce^{4+}]$ and $W = [Br^-]$. The

concentration of A is assumed to be decreasing slowly and hence it is treated as a constant parameter. The stoichiometric factor h tells us about the number of $[Br^-]$ ions produced through an organic oxidation reaction. The variables of the model are U, V and W [45]. The rate equations can be written as:

$$\frac{dU}{dt} = k_1AW - k_2UW + k_3AU - k_4U^2$$

$$\frac{dV}{dt} = 2k_3AU - k_5BV$$

$$\frac{dW}{dt} = -k_1AW - k_2UW + k_5hBV$$

where k_i(i=1,2,3,4,5) are the rate constants of equation listed in Eqn. 2.1.1 - 2.1.5 . The above equations can be expressed in non-dimensional form as:

$$\varepsilon\frac{du}{dt} = qw - uw + u - u^2$$

$$\frac{dv}{dt} = u - v$$

$$\varepsilon^{'}\frac{dw}{dt} = -qw - uw + fv$$

where $u = [2k_4/k_3A]U$, $v = [k_4k_5B/(k_3A)^2]V$, $w = [k_2/k_3A)]W$, $t = k_5Bt$, $\varepsilon = k_5B/k_3A$, $\varepsilon^{'} = 2k_4k_5B/(k_2k_3A)$, $q = 2k_1k_4/(k_2k_3)$, $f = 2h$.
As $\varepsilon >> \varepsilon^{'}$, so w can be considered as a fast variable. Hence we can set, $dw/dt = 0$ (steady state approximation). Then the Oregonator model can be expressed as follows:

$$g_1(u,v) = \frac{du}{dt} = \frac{1}{\varepsilon}(u - u^2 - fv\frac{(u-q)}{u+q}) \tag{2.1.6}$$

$$g_2(u,v) = \frac{dv}{dt} = u - v \tag{2.1.7}$$

Now, we know that if we consider a close system, then the system should reach the equilibrium state. And, even if the system is slightly perturbed (for example, suppose we add a small amount of reactant), it will return back to its equilibrium state. However, on considering an open system in a continuously stirred flow reactor(CSTR), we might witness one or more steady states [51]. And we do not have an idea if these steady states are stable. That is, suppose we apply perturbation. But here, we do not know if the system will return back to its previously acquired steady state or not. Hence in this case, linear stability analysis will help us in predicting the time

evolution of the system after the steady state is perturbed.

To understand how linear stability analysis is performed, let us consider a simple example of a dynamical system described by equations 2.1.8 and 2.1.9 [52]. This simplification is sufficient for oscillating systems considered in my thesis.

$$\frac{dx}{dt} = f_1(x, y) \tag{2.1.8}$$

$$\frac{dy}{dt} = f_2(x, y) \tag{2.1.9}$$

The steady state (x_s, y_s) of this system is achieved when $f_1(x_s, y_s) = 0$ and $f_2(x_s, y_s) = 0$. Let us now see how the stability analysis is performed to check the "stability" of the steady point (x_s, y_s) :
We search for the solution of Eqn. 2.1.8 and 2.1.9 assuming the following form:

$$x = x_s + \delta x(t) \tag{2.1.10}$$

$$y = y_s + \delta y(t) \tag{2.1.11}$$

where, contribution of δx and δy are small. If we neglect the higher order terms then the time evolution of $\delta x(t)$ and $\delta y(t)$ is described by:

$$\begin{bmatrix} \frac{d}{dt}(\delta x(t)) \\ \frac{d}{dt}(\delta y(t)) \end{bmatrix} = \begin{bmatrix} \frac{df_1}{dx}|_{x_s,y_s} & \frac{df_1}{dy}|_{x_s,y_s} \\ \frac{df_2}{dx}|_{x_s,y_s} & \frac{df_2}{dy}|_{x_s,y_s} \end{bmatrix} \begin{bmatrix} \delta x(t) \\ \delta y(t) \end{bmatrix} \tag{2.1.12}$$

Let us represent Jacobian matrix as $J = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. We can search the solution of the linear equation 2.1.12 in the form :

$$\begin{bmatrix} \delta x(t) \\ \delta y(t) \end{bmatrix} = Ce^{\lambda t} \tag{2.1.13}$$

Considering both eqn 2.1.12 and 2.1.13 we get,

$$JC = \lambda C \tag{2.1.14}$$

In the above equation, C represents the eigenvector and the constant λ is the eigen-

value. On rearranging the above equation, we get:

$$(J - \lambda I)\,C = 0 \tag{2.1.15}$$

where I is known as the identity matrix. Now, the above equation can have nontrivial solution if

$$(J - \lambda I) = 0 \tag{2.1.16}$$

upon using the terminology of the Jacobian matrix we can rewrite the above equation as,

$$(a_{11} - \lambda)(a_{22} - \lambda) - a_{21}a_{12} = 0 \tag{2.1.17}$$

On multiplying it out, we get:

$$\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0 \tag{2.1.18}$$

Now $a_{11} + a_{22}$ is actually the trace of matrix J ($Tr(J)$) and the last term of the equation is the determinant ($det(J)$). Hence we can rewrite the equation as:

$$\lambda^2 - Tr(J) + det(J) = 0 \tag{2.1.19}$$

and it's solutions are

$$\lambda_{1,2} = -\frac{1}{2}Tr(J) \pm \sqrt{Tr(J)^2 - 4det(J)} \tag{2.1.20}$$

There are the following relationship between solution:

$$\lambda_1 + \lambda_2 = Tr(J) \tag{2.1.21}$$

$$\lambda_1\lambda_2 = det(J) \tag{2.1.22}$$

Following equation 2.1.20 , we consider a few cases [53].
1. If $Tr(J) > 0$ and $det(J) > 0$ then it means the sum as well as the product of eigenvalues are positive. In terms of a differential equation, this represents an exponential growth in the direction of vectors c_1 and c_2. Thus the trajectories will move away from the fixed point. Hence we call this situation as an unstable node.(Figure 2.3(a))

2. If $Tr(J) > 0$ and $det(J) < 0$: This signifies that the product of eigenvalue is negative. Thus one eigenvalue is positive and another one is negative. So we can witness the exponential growth in the direction represented by the positive eigenvalue and we will witness an exponential decay in the direction represented by the negative eigenvalue. This is represented as "saddle", points move towards the fixed point in one direction but move away from it in another direction.(Figure 2.3(b))
3. If $Tr(J) < 0$ and $det(J) > 0$: It means that both the eigenvalues are negative. Thus we will witness exponential decay in both c_1 and c_2 direction. So it implies a stable node. (Figure 2.3(c))
4. If $Tr(J) < 0$ and $det(J) < 0$: which again means that any one of the eigenvalue is negative and another is positive and hence the situation corresponds to the saddle. (Figure 2.3(b))

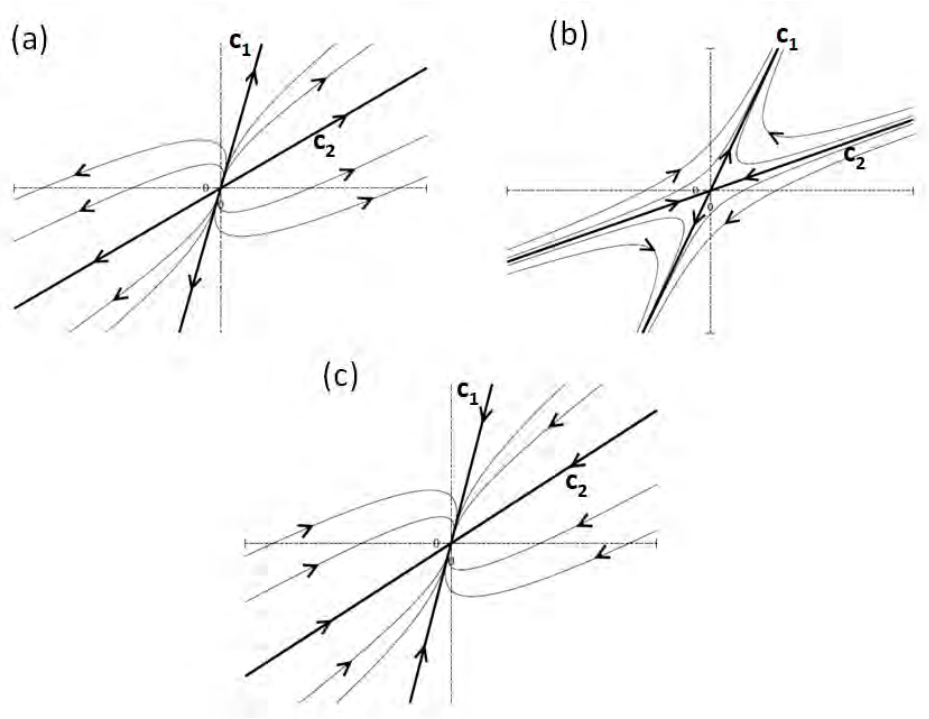


Figure 2.3: The steady state point is at (0,0). Illustration of: (a) an unstable node with $\lambda_1 > \lambda_2 > 0$. The trajectory leave the equilibrium point. (b) a saddle, with $\lambda_1 > 0 > \lambda_2$. The trajectory approach the equilibrium in direction c_2, but leave in direction of c_1. (c) a stable node with $0 > \lambda_1 > \lambda_2$. [53]

Now let us consider the case where we have complex eigenvalues. (i.e when $tr(J)^2 < 4\ det(J)$ in Eqn. 2.1.20) The eigenvalues are then written as : $\lambda_1 = \rho + i\omega$ and $\lambda_2 = \rho - i\omega$. Now, in complex eigenvalue case , $e^{\lambda t}$ is in the form of $e^{(\rho \pm i\omega)t}$. The real part of the eigenvalue describes growth and decay. Let us consider three cases:
1. If $Tr(J) > 0$: Because the eigenvalues have a positive real component, exponential

growth will occur. As a result, points drift away from the equilibrium point while oscillating. This is referred to as an unstable spiral (Figure 2.4(b)).
2. If $Tr(J) < 0$:We see exponential decline because the eigenvalues have a negative real component. We see oscillations while they approach closer to equilibrium. This is referred to as a stable spiral (Figure 2.4(a)).
3. If $Tr(J) = 0$: When two eigenvalues are fully imaginary with real components zero, the consequence is a motion with points moving around on an ellipse: no net motion is observable away from or towards the steady state (Figure 2.4(c)).

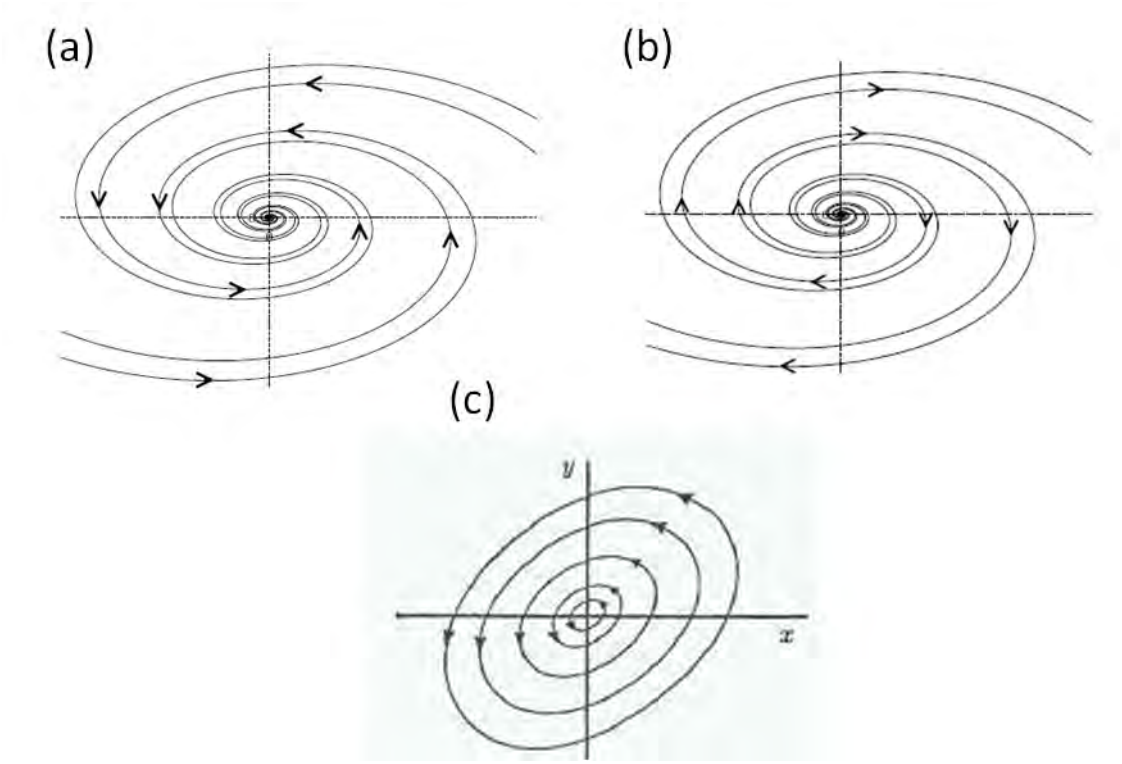


Figure 2.4: (a) Stable spiral (b) unstable spiral (c) center : no net motion is seen away or towards the steady state [53].

Now let us apply stability analysis to two-variable Oregonator model [54]:
The Oregonator model has a steady state at $du/dt = 0$ and $dv/dt = 0$. The steady state we get is : $u_0 = v_0 = \frac{1}{2}[1 - q - f + \sqrt{(1-q-f)^2 + 4q(1+f)}]$

We can construct a Jacobian matrix in the form as Eq. 2.1.12 with $f_1 \to g_1$ and $f_2 \to g_2$
On solving for the Jacobian matrix, the elements we get are:

$$a_{11} = \frac{\varepsilon_H}{\epsilon}$$

$$a_{12} = -\frac{f(u_0 - q)}{\epsilon(u_0 + q)}$$

$$a_{21} = 1$$

$$a_{22} = -1$$

here, $\varepsilon_H = 1 - 2u_0 - \frac{2qfu_0}{(u_0+q)^2}$. Thus the Jacobian has a sign structure of $\begin{bmatrix} + & - \\ + & - \end{bmatrix}$ So to achieve the condition of stability, the trace of this matrix should be negative and the determinant should be positive. To achieve stability , the condition which is to be followed is:

$$1 - 2u_0 - \frac{2qfu_0}{(u_0+q)^2} < f\frac{u_0 - q}{u_0 + q} \tag{2.1.23}$$

The above condition only holds for typical values of f and q. Hence for particular values of these parameters , the fixed point will be stable, otherwise it will be called as unstable point and we can only witness oscillations for certain parameter values (Figure 2.5)

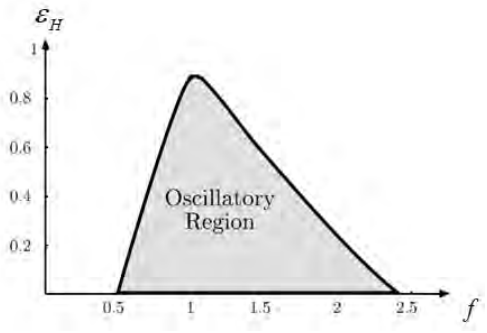


Figure 2.5: Plot of ε_H vs. f is shown with $q = 10^{-3}$. The suitable condition of oscillations appear below the curve [55].

To give a rather proper visual explanation, we take into consideration the effect of parameter f over the stability of the fixed point and explain this concept with the help of a phase portrait diagram. Figure 2.6 shows the graph of u vs. v. The nullclines ($du/dt = 0$, $dv/dt = 0$) are plotted. The curves intersect at a single point which is the steady state of our system. Now the position of the intersection point can vary with different values of f. If the value of f is low, the nullclines intersect to the right side of the maximum of the u-nullcline. If the value of f is high, then the intersection shift to the left [56]. In Figure 2.6(b) the graph has been plotted with f= 1.0 , q= 0.015, $\varepsilon = 0.03$. Now, let us consider that the initial conditions are represented by a point located below the u-nullcline (Figure 2.6 (b) point A). So at point A both du/dt and dv/dt are positive, and the direction of trajectory in u ,

v phase space move upwards. On moving upward, dv/dt is still positive but du/dt is negative, causing the trajectory to move towards the left branch of the v-nullcline. Now at this position dv/dt also becomes negative, which makes the trajectory to go downward and approach the minimum. At this point the value of $dv/dt < 0$, so the trajectory leaves the u-nullcline and again proceed towards the right branch to repeat the whole cycle. Hence there is no possibility for the system to reach the stable point but instead converges to a "limit cycle" trajectory. This signifies its oscillatory behaviour. Now let us assume another case (Figure 2.6(c)) with $f = 10$. Now the intersection point got shifted to the left branch and is located very close to the minimum of the u-nullcline. Suppose we start with the initial condition (Figure 2.6 (c) Point A). The corresponding trajectory will move upward, followed by moving

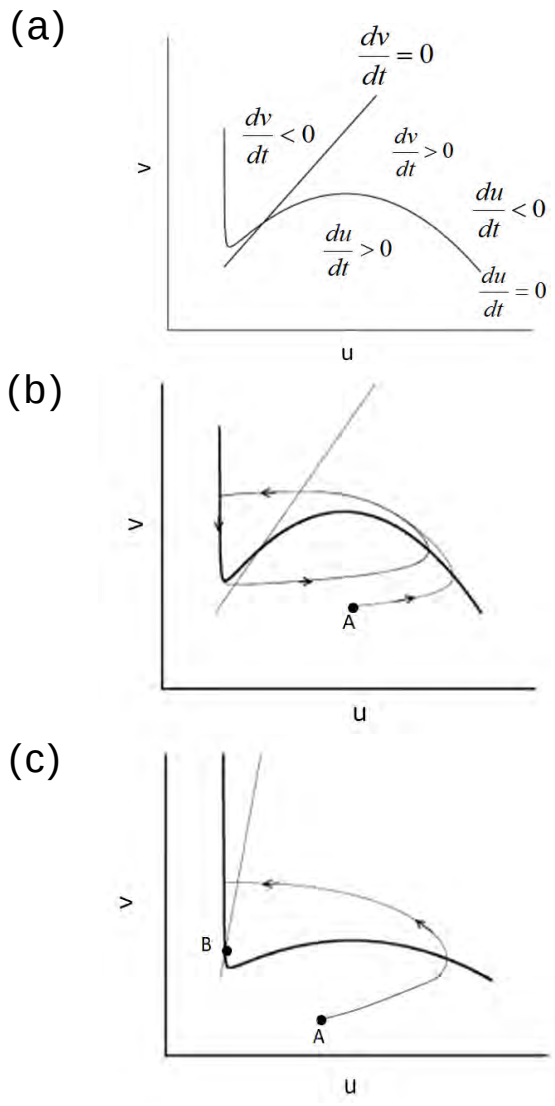


Figure 2.6: Phase space representation of the two-variable Oregonator model with $\varepsilon = 0.03$.(a) f=1.0 (b) f=1.0, oscillatory regime (c) f=10, excitable regime [56]

towards the left branch and finally to the right branch of the u-nullcline. And it begins to descend along the left branch and finally reaches a steady state. So in this case the intersection point is a stable one. This situation is also analogous to the condition of excitability of the system where it gets perturbed and then the perturbation rapidly increases and finally slowly returns back to its stable state.

Thus we can conclude that the system approached the stable stationary state. As it is stable, it's small perturbation decays with respect to time and the system returns back to it's steady state. However with large perturbation, the system responds to it by giving excitation peaks. If an excitable medium gets spatially distributed, we can witness propagation of pulse [57]. Using of excitable media for information processing comes from the analogy of signaling happening in neural system due to excitable biochemical reactions [58]. The most basic method of information coding that can be used is an excitable media where a single pulse represents a bit of information. The information coded in these pulses can be processed via pulse to pulse interactions. Some noteworthy works on information processing by chemical media of B-Z reaction are mentioned in the next section.

Illumination controlled Oregonator model

Krug et al. [59] modified the already existing Oregonator model by adding another step to the kinetic process, which shows that the production of bromide ion increases with the increase in light intensity through the term ϕ. After the scaling of equations suggested by Tyson [60], the equation becomes:

$$\frac{\partial u}{\partial t} = \frac{1}{\varepsilon}(u - u^2 - (fv + \phi)\frac{u - q}{u + q}) \tag{2.1.24}$$

$$\frac{\partial v}{\partial t} = u - v \tag{2.1.25}$$

This new form of Oregonator model can be understood by replacing f with f_{eff}, where $f_{eff} = f + \phi/v$. When f is fixed and $\phi = 0$, we can obtain the original Oregonator model. With the increase in the value of ϕ, f_{eff} will also increase and the model undergoes a "photoinduced" transition from oscillatory behaviour to a steady state [61]. This model can explain the effect of light over B-Z reaction effectively. Hence the oscillations in the B-Z reaction can be easily controlled externally by means of illumination.

We have modified these already existing Oregonator equations to establish interactions among the oscillators. The B-Z network designed by Oregonator model is

explained in detail in Chapter 3.

2.2 Chemical computing with B-Z reaction

Chemical information processing can be witnessed in living organisms. The important functions of living organisms : be it cognitive tasks, vision, plannings or be it the immune system response all happens due to some chemical information processing going on within our body. This natural phenomenon inspired the researchers to design an artificial chemical processor.

In the majority of chemical reactions, the intermediate compounds are formed during the initiation reaction and slowly die down when the reaction is over. But in oscillatory reaction, the intermediates oscillate periodically. This type of reaction can be witnessed where a set of parallel chemical processes exist with at-least one being auto-catalytic and one being inhibitory [62]for eg B-Z reaction. The reaction became popular and a famous choice for chemical information processing when the formation of logic gates came to reality. In 1993, Toth and Showalter [63] performed experiments to create logic gates by B-Z media in capillary tubes. Thin layers of B-Z media were connected to the capillary tube of different diameters. It was seen that the wave initiated by the reaction enters the capillary tube and forms a hemispherical shape at the exit point. And when the tube size is below the critical size of wave nucleation, the hemisphere of the excitation wave collapses and no travelling wave is formed. Thus this phenomenon can create a chemical analog of on/off switch. With

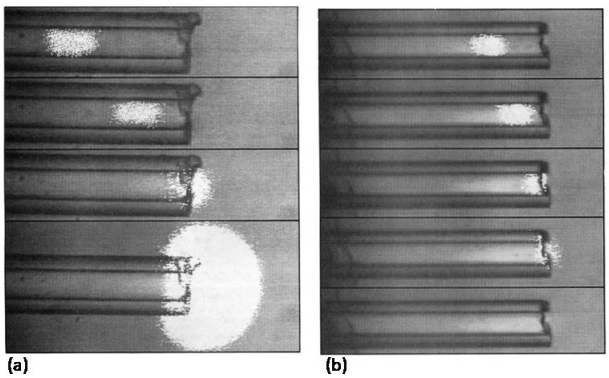


Figure 2.7: Images showing chemical wave at 5 s intervals travelling through the capillary tube of 160 μm.(a) The wave grows in a hemisphere at the tube opening which then generates an expanding circular wave. Concentration of BrO_3^-= 0.15M (b)Concentration of BrO_3^-= 0.10M. At this concentration, the tube size is below the critical size for wave nucleation and thus the hemisphere of excitation collapses and no travelling wave develops [63].

an excitation wave as a single input, one can get the result of output as binary state

0 or ON (when no output wave is observed, in Figure 2.7(b)) or output as binary state 1 or OFF (when output wave is observed, in Figure 2.7(a)). Further after this observation, the authors constructed logic gates with respect to the chemical waves by creating an array of tubes. The gates were created by using two coaxial capillary tubes forming two input chambers on either side, with the output chamber in the center. The reading based on the chemical wave was done as follows: If input 1 is 0 (that is no wave generated in input tube 1) and input 2 is 1 (wave generated in input tube 2), then the output will be either 0 (no wave in the output chamber) or 1 (generation of wave occurred in the output chamber). Configuration of AND gate from OR gate is possible with a slightly lesser wave velocity resulting in a larger critical radius of the tube [64]. Further they also constructed other gates by constructing them with the combinations of the basic ones. This formation of logic gates where geometrically controlled by critical radius, but after that many other logic gates were formed with B-Z reaction using different methods such as chemical diode logic gates [65] or light controlled formation of logic gates [66].

Another mention worthy work done with B-Z catalyst as a media in the arena of information processing was image processing. Kuhnert et al. [50] have shown that using the photosensitive property of B-Z reaction and its oscillatory nature can perform image processing operations such as alteration of a positive and negative image, contrast enhancement, contour detection, smoothing discontinuous images. In this experimental investigation, they projected a half tone image onto a thin layer of reagent and the light intensity was varied as a function of spatial coordinate and optical density. A simple mechanism of image processing can be explained by Figure 2.8 as follows:

Figure 2.8: Light sensitive B-Z reaction with a negative image projected onto it. [50]

Frame 1 corresponds to the phase of negative image projection onto it. In frames 2 and 3 the portions which are illuminated have a larger period of oscillations. Hence oscillation first starts from non illuminated parts followed by the illuminated ones. So the initial negative image changes to a positive one (i.e, the region of orange blue oscillations changes to blue orange respectively). Now due to the reaction's own oscillatory nature the blue will change to orange and the orange region to blue, hence the image will be negative again [50] . And this process repeats periodically. It is also to be noted that the chemical medium has some kind of short term memory, where it

was able to store the image and perform image processing.

Further, Motoike et al. [67] demonstrated the memory on an excitable media by constructing an experiment by Nafion membrane assembly (consisting of a circular track with a small gap of 0.168mm and an input channel) soaked in B-Z catalyst ferroin and immersed in B-Z solution. When the input channel is introduced above the gap, waves were observed in anticlockwise direction. If the input channel is placed below the gap, the waves move clockwise. And when the input channel was detached the clockwise waves were still present. Thus they concluded this observation as the waves will always be anticlockwise if the inputted channel is above, hence retaining the memory of moving in an anticlockwise direction. The memory gets erased when the input channel is placed below the gap and so the wave starts moving in clockwise direction. The authors compare this finding to a real time memory which can be erased and rewritten (Figure 2.9).

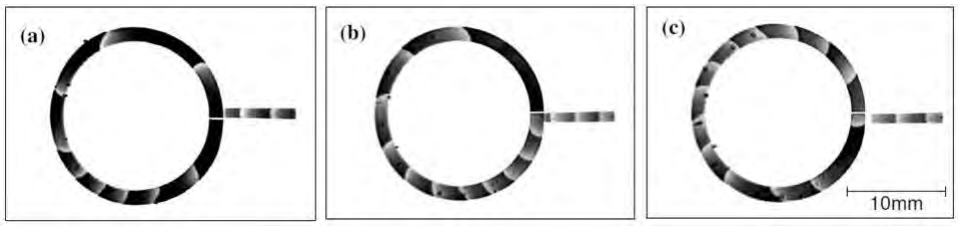


Figure 2.9: Real time memory on an excitable media. (a) Input signals introduced on the ring moves in anticlockwise direction. (b) After the input channel is shifted to lower position, the wave rotates in clockwise direction. (c) The input channel is removed and the wave continues to move in clockwise direction. [67] [35]

In recent works M.S. Tsompanasa et al. [68] have shown how simple B-Z probed liquid marbles can be used to control the movement of a robot. Also Adamatzky et al [69] explored the geometry of London's streets using a computational mode of an excitable chemical system, Belousov-Zhabotinsky (B-Z) medium. The streets were virtually filled B-Z medium and they studied the propagation of excitation waves for a range of excitability parameters, gradual transition from excitable to sub-excitable to non-excitable with the aim of understanding the dynamics of pedestrian and vehicular traffic in the cities. We can conclude from all these works (computational and experimental) that nonlinear reactions as a media seem to be an interesting candidate in designing chemical computers as the medium evolution can be controlled externally. Forming an interacting network of chemical oscillators can give rise to more complex behaviour.

All of the above mentioned works were based on bottom up approach. We performed information processing with B-Z reaction by using a top-down strategy. We used an evolutionary algorithm as a tool to train the system and used the concept

of information theory as a technique to teach our system to perform information processing. The basic concepts of these are described in the next section.

2.3 An insight to information theory

Information can be treated as an abstract concept where it receives it's physical meaning only when it involves two correspondents: one generating it (the information source) and another receiving it (the destination). The transmitted signal acts as a physical medium to transport information from the receiving to the final end [70]. Shannon tried to quantify this abstract notion by presenting the relation between entropy and information. He introduced this concept in communication theory and formed the foundation of information theory. Information theory is still primarily concerned with communication theory, but it does have a wide range of applications in statistics, data processing, and computers. [71].

Entropy and information

Let us consider random phenomena. Random phenomena are mostly unpredictable and if randomness contains a certain amount of *structure* or pattern, it's predictability increases. So to find a certain definite pattern among all the randomness, we try to describe it in terms of probability. To do so, we can define a set of outcomes, which is called the sample space. A probability distribution on a sample space $A = \{E_1, E_2, ..., E_N\}$ can be defined as a function p that assigns a probability to each outcome in a sample space. p is a map from A to the unit interval, $p : A \rightarrow [0, 1]$,which must satisfy $\sum_{i=1}^{N} p(E_i) = 1$.

Considering our everyday daily routine *structured* life, if we want to add some unpredictability to our existing monotonous life, some random surprise events can be introduced. Usually more unlikely and unpredictable the event is, the more surprising it is. We can quantify this idea by using probability distribution as follow:

If E is an event in a sample space A, we define the surprise of E to be $s(E) = -log(p(E))$. Events for which $p(E) = 1$ has zero surprises, i.e are certain to occur and if $p(E) = 0$, then the event is unlikely to occur and hence has an infinite surprise. Defining the surprise as the negative logarithm of the probability gives us the required limiting values tending to 0 or 1. It also makes surprise additive, meaning that if several independent events happen in succession, the overall surprise generated is equal to the sum of their individual surprises. This expectation value of the *surprise* can be linked to be the entropy of the probability distribution and can be formulated

as

$$H(p) = -\sum_{i=1}^{N} p(E_i)log(p(E_i)) = \sum_{i=1}^{N} p(E_i)s(E_i)$$

Entropy is considered to be a measure of the extent to which a system is disorderd [72]. And if we want to understand the *amount of surprise* stored in an event, it is to calculate the amount of information present. The less is the surprise the less is the information the event can provide. Hence the amount of surprise is proportional to the amount of information present. Shannon quantified the amount of surprise or information present in events with entropy and connected the concept of entropy with that of predictability. It was then realized that entropy is a property of any stochastic system [73]. Let us now understand how entropy is related to predictability. A plot of $-plogp$ as a function of p is shown in Figure 2.10. It is to note that $0 < p < 1$:

$$\lim_{p\to 0} plog_2 p = 0$$

and

$$\lim_{p\to 1} plog_2 p = 0$$

This indicates that events that are more likely to happen and those that are less likely to happen both contribute little to entropy. Events whose probability are somewhere in between make a comparatively large contribution to the entropy. Now, coming to

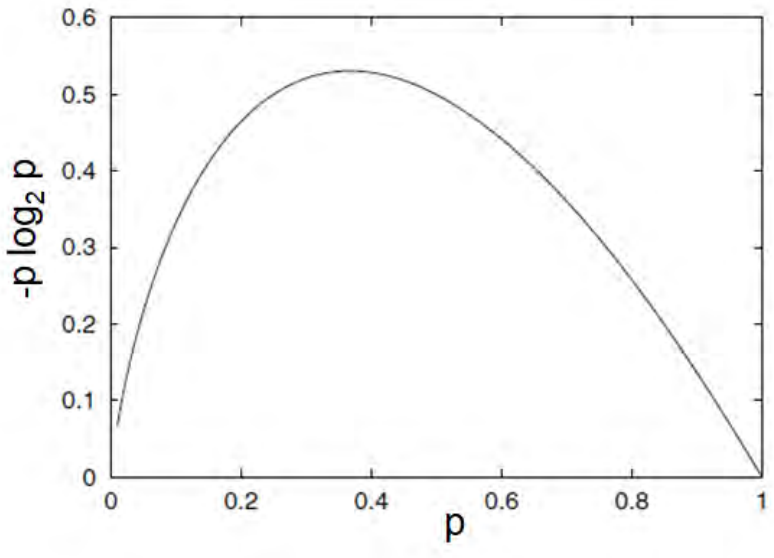


Figure 2.10: Plot of $-plog_2(p)$ as a function of p. [71]

the entropy, let us understand the relation of entropy with predictability and information with certain examples:

Example 1: Let us consider $A = \{s_1, s_2\}$ with $p(s_1) = 0.5 = p(s_2)$. The entropy (using

log_2) is

$$H(p) = -(0.5)(-1) - (0.5)(-1) = 1$$

Thus with entropy equal to one, it means both the cases s_1 and s_2 are equally likely to occur and hence the situation is as unpredictable as possible.
Example 2:
$A = \{s_1, s_2\}$ with $p(s_1) = 1.0$, $p(s_2) = 0.0$. Using the convention that $0log(0) = 0$, the entropy is 0, i.e the situation is entirely predictable and it is clear that s_1 will always occur. So no surprise element is left and hence hardly this case can provide us with large amount of information.

The unit in which entropy is measured depends upon the base of logarithm. It can be of base 2 or e. If it is base 2, the entropy is measured in *bits* and if its of base e, the entropy is measured in *natural unit*. The above examples are calculated with log_2 base. Now as we know the relation between information and entropy, let us now understand how we can quantify information by using the equation of Shannon entropy.

Let us take example 1. So, as H(p)=1 *bit*, this implies that 1 bit of memory is necessary to store the information [73] about this particular case. In the case of example 1, the situation was as unpredictable as possible and hence contained a good amount of surprise and information. It is to be noted that if an event is easily predictable, then the amount of surprise content in that event is less, which in turn means that the amount of information content is also less. And, thus the requirement of the number of bits to store the information for that particular case will also be less.

Joint and conditional entropy

Joint entropy is a measure of entropy [74] that provide us with the information about the amount of uncertainty present if two variables are entertained together. Suppose two variables X and Y, take x and y values, respectively. Let p(x,y) be the joint probability of x and y. The joint entropy of X and Y is defined as [75]

$$H(X, Y) = -\sum_{x,y} p(x, y) log\, p(x, y)$$

Conditional entropy can be understood as, if we want to find the entropy of random variable y, provided we already know the information of random variable x.

The conditional entropy $H(Y|X)$ is defined as: [76]

$$\begin{aligned} H(Y|X) &= \sum_x p(x) H(Y|X=x) \\ &= -\sum_x p(x) \sum_y p(y|x) log\, p(y|x) \\ &= -\sum_x \sum_y p(x,y) log\, p(y|x) \end{aligned} \tag{2.3.1}$$

We can establish a relation between conditional and joint entropy as follows:

$$\begin{aligned} H(X,Y) &= -\sum_x \sum_y p(x,y) log\, p(x,y) \\ &= -\sum_x \sum_y p(x,y) log\, p(x)p(y|x) \\ &= -\sum_x \sum_y p(x,y) log\, p(x) - \sum_x \sum_y p(x,y) log\, p(y|x) \\ &= -\sum_x p(x) log\, p(x) - \sum_x \sum_y p(x,y) log\, p(y|x) \\ &= H(X) + H(Y|X) \end{aligned} \tag{2.3.2}$$

This is known as the chain rule.

Mutual Information

Mutual information finds the degree of dependence between two variables. Or rather, we can express it as the measure of the amount of information two random variable shares *mutually*. Thus, it tells us the amount of information that can be extracted about variable x by observing y or vice versa. Mutual information can be expressed as :

$$I(X;Y) = \sum_x \sum_y p(x,y) log \frac{p(x,y)}{p(x)p(y)}$$

We can also define the mutual information, in another way as shown below [76]:

$$\begin{aligned} I(X;Y) &= \sum_{x,y} p(x,y) log \frac{p(x,y)}{p(x)p(y)} \\ &= \sum_{x,y} p(x,y) log \frac{p(x|y)}{p(x)} \\ &= -\sum_{x,y} p(x,y) log\, p(x) + \sum_{x,y} p(x,y) log\, p(x|y) \\ &= -\sum_{x} p(x) log\, p(x) - (-\sum_{x,y} p(x,y) log\, p(x|y)) \\ &= H(X) - H(X|Y) \\ &= H(X) + H(Y) - H(X,Y) \end{aligned} \tag{2.3.3}$$

2.4 Evolutionary algorithm

Charles Darwin was the first one to explain about evolution: explaining how every life form adopts to the surrounding environment for the matter of survival. The concept of evolution and survival of the fittest has long escaped biology and made its way to mathematics, physics and computer science as well [77] and came to be known as evolutionary algorithm (EA). EA mimicked the evolution and selection process of the life form to exist and became a famous numerical optimization method. It can be categorized as a stochastic search algorithm [78].

Evolutionary algorithm starts with a randomly initialized population and undergoes selection, mutation, recombination in order to give us the optimal solution [79]. The idea of EA is shown in Figure 2.11. The termination condition of the shown flowchart can be selected based on two criteria:

- resource criteria - We can beforehand decide the maximum number of generations or the maximum amount of cpu time we want to run the algorithm and introduce the conditions as termination criteria.

- convergence criteria - We can also define a condition to terminate the program by checking that a certain desired fitness value is achieved and it remains the same for say, in next few generations.

The three main operations performed by EA are : mutation, recombination and selection. Selection exploits the fitness function in order to guide the search into a promising search space region, whereas mutation explores the search space [79]. Now, while mutation performs search step based on only one information of the parent,

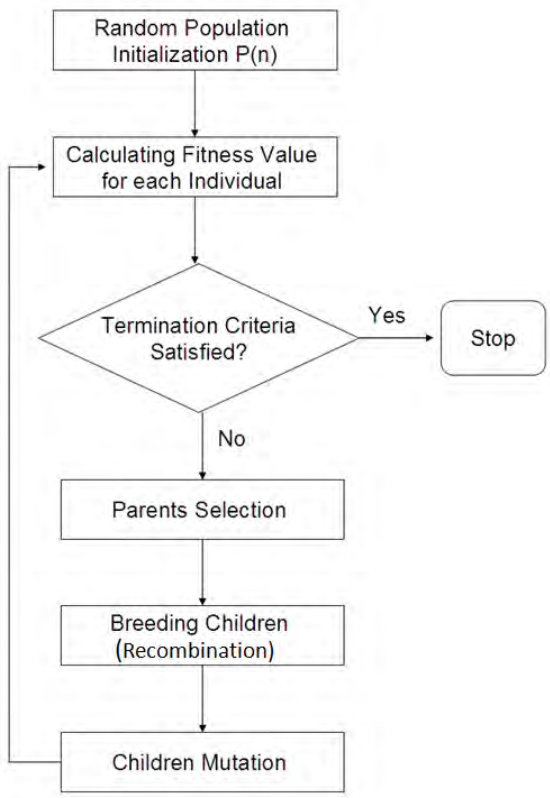


Figure 2.11: Schematic representation of evolutionary algorithm [80]

recombination shares the information from ρ parent individual. The search space is a space with all possible solutions. Every point lying in the search space represents a feasible solution and each point's feasibility is represented by its fitness value. EA helps us to find that one point from the search space with a higher fitness value. That point will be our optimum solution.

The properties of the three main operators of EA are described below:

Mutation

The mutation operator helps to bring a kind of variation in the set of the existing populations. The type of mutation we want to perform, depends upon the type of problem we are dealing with. Some rules which have been proposed [81] are:

- Reachability : Given a parental state (or the already existing state we are in) we should be able to reach any another state with a finite number of mutation steps or generations.
- Unbiasedness: The mutation operator should not be biased as its selection should not depend on the fitness function.
- Scalability: It says that the variation rate should be introduced in a tunable

manner with the sole purpose of increasing the overall fitness value under each generation.

Recombination

The recombination operator helps in mimicking the process of reproduction by creating a new set of offsprings from the existing population (the parents). There are two ways in which recombination can be performed, they are : dominant or discrete recombination and intermediate recombination.

In the dominant or discrete recombination, the parents are chosen randomly to undergo recombination. Only a few discrete features from each parent are chosen to produce offspring. These parents are randomly chosen from a slot or group of parents who are dominant over others. Whereas in intermediate recombination, one takes all parents into the account (and not only choosing from the dominant slot) to perform recombination. Here, the features to pass on to new offspring can be selected by considering the average of all the values of the particular characteristic present in the whole population.

Selection

The selection operator can guide the search into the promising regions of the optimal solution space. That is, it guides the search towards a high fitness value. There are two versions of selection techniques: plus selection$(\mu/\rho + \lambda)$ and comma selection $(\mu/\rho, \lambda)$, where μ is the population size, ρ is the number of parents taking part in recombination or in the production of offsprings and λ is the number of offspring generations. In $(\mu/\rho, \lambda)$ only the λ newly generated offspring defines the selection pool and the parents from the previous generation are forgotten, irrespective if they have higher fitness value compared to the offspring. Whereas in $(\mu/\rho + \lambda)$ strategy, both the parent and the offsprings are copied to the selection pool of size $(\mu/\rho + \lambda)$.

Due to its simplicity, flexibility and robustness, EA has become a popular solving technique for optimization and learning and can solve a wide range of problems [82].

2.4.1 Working of evolutionary algorithm in function approximation

Evolutionary algorithm is a great tool to be used when we want to find the optimal values of parameters that give the best solution for a certain problem. Let us take a simple example to understand the working: In this case the usefulness of EA has been investigated in function approximation. I define a function $f(x)$ by it's values at selected points of x.

$$x \in \{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}$$

And the corresponding values of $f(x)$ were:

$$\{49, 36, 25, 16, 9, 4, 1, 0, 1, 4, 9\}$$

I decided to approximate this unknown function by a polynomial equation

$$f_{appr}(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$$

which will be subjected to the EA in order to find out the best fitted coefficient values. As I have said before that the individual with more fitness survives in the population, it becomes very important to define the fitness criteria perfectly. In the case of function approximation, the main desire was to obtain a curve that matches the curve of $f(x)$ with high accuracy. So, the fitness function I selected was the error determination in between the actual and the approximate point values. Hence, in this case, the fitness function was the minimization. That is, the lower the value of fitness function, the more is the survival rate of that individual. Mean squared error(MSE) was used as a fitness function, which was defined as:

$$fitness = \frac{1}{n} \sum (f_{appr}(x) - f(x))^2$$

In this particular case, I took the population size to be 1000 (later on I have varied the population size). The element of population were the coefficient values (a_0, a_1, a_2, a_3) of $f_{appr}(x)$. The initial set of the population was formed by generating random values for each element of the population. The total number of random values generated was equal to the population size. I kept the range of coefficients in between -10 to +10. I copied 5% of the best fitted individual (based on fitness value) directly to the next generation. The top 50% of the population was chosen as potential parents to produce new offsprings by performing recombination, hence

creating a new generation consisting 5% best fitted old population and 95% new offsprings. A schematic representation of how the recombination took place in this case has been shown below:

$$parent1 : a_0 + a_1x + a_2x^2 + a_3x^3$$

$$parent2 : b_0 + b_1x + b_2x^2 + b_3x^3$$

$$parent3 : c_0 + c_1x + c_2x^2 + c_3x^3$$

$$offspring : a_0 + b_1x + c_2x^2 + a_3x^3$$

The values of parameter a_0, a_3, b_1 and c_2 were copied from parent1, parent2 and parent3 respectively to create a new offspring. It was a very important step to let the population evolve with a number of generations. I introduced mutation to the coefficient a_3 by giving it altogether a new value from the range of -0.1 to +0.1. I allowed every single individual to undergo mutation.

$$parent1 : a_0 + a_1x + a_2x^2 + a_3x^3$$

$$parent2 : b_0 + b_1x + b_2x^2 + b_3x^3$$

$$parent3 : c_0 + c_1x + c_2x^2 + c_3x^3$$

$$offspring : a_0 + b_1x + c_2x^2 + a_3x^3$$

$$offspring(with\, mutation) : a_0 + b_1x + c_2x^2 + d_3x^3$$

where d_3 is a random value ranging in between -0.1 to +0.1. I got the minimized fitness value to be 0.105 ran up-till 500 generations. It is to be noted that in this case fitness value was the error value, so an error of 0.105 was considered as a low error value, hence carrying a capability to give the best approximation. The coefficient values, I got with this approximation are given in Table 2.1.

Below, we can see the fitness vs. generation graph (Figure 2.12). It was seen that with an increase in generation, the fitness kept on getting better, that is the fitness was minimized until converged and the fittest value of the coefficients were obtained. In Figure 2.13 the main function was plotted and was fitted with the best optimized polynomial approximation equation. It was seen that it fitted perfectly with an average minimal error (or in our case we can call it the minimized fitness) of 0.105.

Table 2.1: Optimized coefficient values for $f_{appr}(x)$

population size	1000
best fitness value	0.105
a_0	3.598075
a_1	-4.079368
a_2	1.034429
a_3	0.004143

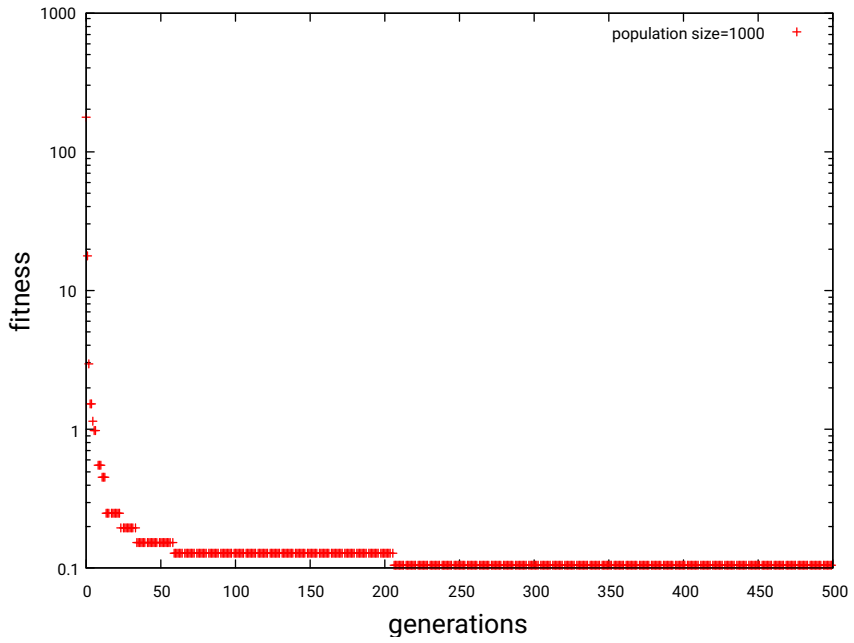


Figure 2.12: fitness vs. number of generation (Population size=1000). The y axis was plotted in log scale, which gives a better picture of how the fitness function evolved with respect to the generation.

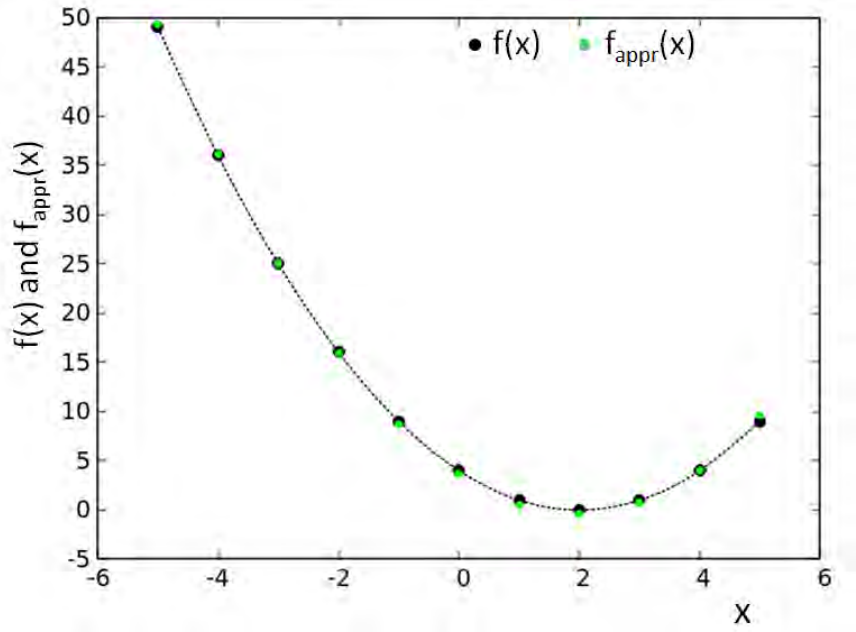


Figure 2.13: The function values plotted along with it's approximation obtained by evolutionary algorithm.

To know the characteristic features of evolutionary algorithm, I decided to investigate how different parameters of evolutionary algorithm results in affecting the fitness value. The investigation goes as follow:

Effect of Population size

I changed the population size, ranging in between 100 to 1900, and noted down the minimized fitness value obtained after 500 generations in each case. The mutation and recombination rules (as mentioned before) were kept the same during evolution with each population size. Below, are the results showing the variation of fitness value with that of the population size. The fitness value seemed to decay exponentially with respect to the size of the population (Figure 2.14). Further, it can be said from the obtained graph that, a smaller population size led to a higher fitness value, which in turn meant the error was high and the fitness value minimization was not of refined quality. Moreover, we can also argue that in the case of a small population, the variation in "genes" is lower. In Figure 2.15, we can see that for a smaller population, the fitness value converged very easily, which means with an increase in generation, it was not able to evolve due to less number of variations among the individuals. Whereas in case of a higher population, the evolution of the fitness value was witnessed for a long time after which it reached its convergence value.

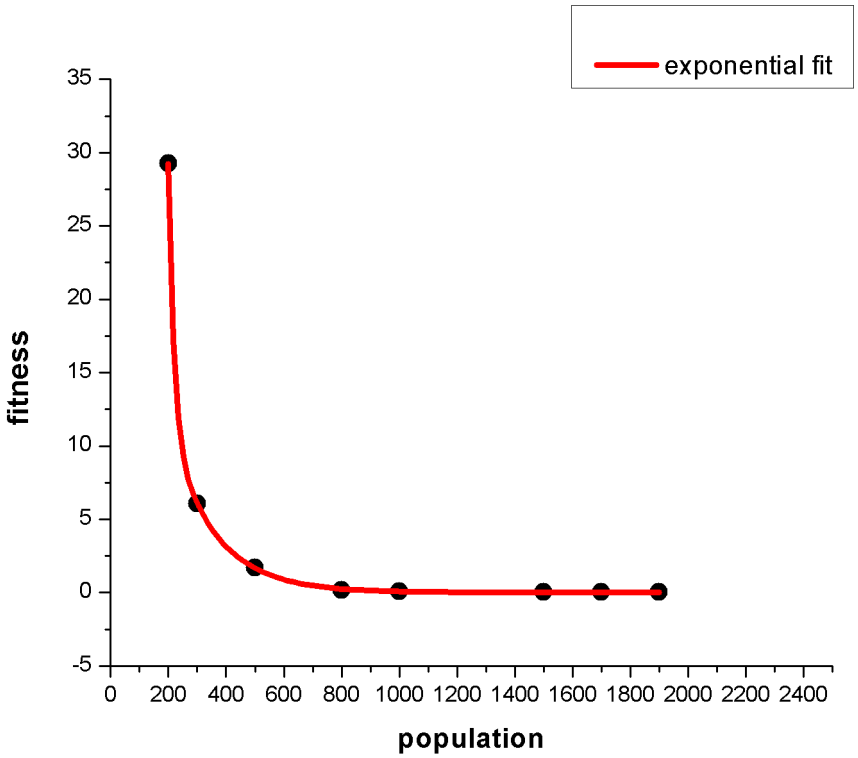


Figure 2.14: Minimized fitness (in log scale) vs. population curve. The graph was fitted with exponential decay function

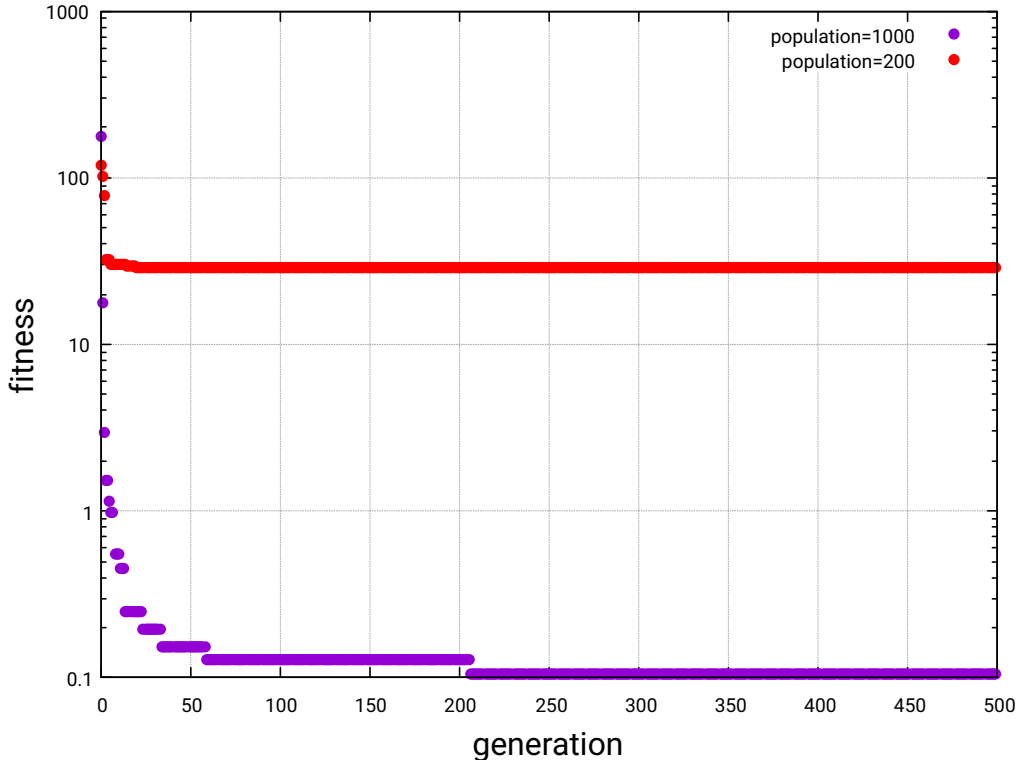


Figure 2.15: Fitness (in log scale) vs. generation graph for population size 200 (red) and 1000 (violet) respectively. It shows the evolution of fitness with respect to number of generations

Mutation Probability

I introduced a mutation rate or mutation probability in my algorithm. That is from a technical point of view, I generated random probabilities for every individual and only allowed individuals with certain probabilities to undergo mutation. Hence introducing this mutation rate, I put a constraint to allow the only selective individuals to undergo mutation. I noted down the change of minimized fitness value with respect to the mutation rate for population of 500, 800, 1500 and 1900 individuals respectively (Figure 2.16). For every size of the population, minimized fitness value varied in an oscillatory way with that of mutation rate. Moreover, the larger variation in fitness value with respect to the mutation rate was observed in cases with a low population (500 and 800) and the variation gradually decreased with an increase in population size. With a population of 1900 individuals, the variation in the fitness value with respect to mutation rate was seen to be less. When the population size was large, the change in mutation rate did not show a significant change in the value of fitness. However, the fitness value was dependent on the choice of mutation rate in the case of a smaller size population. In small size population (500 and 800) the fitness value was high when the mutation rate was chosen in the range of 0.8 to 0.9.

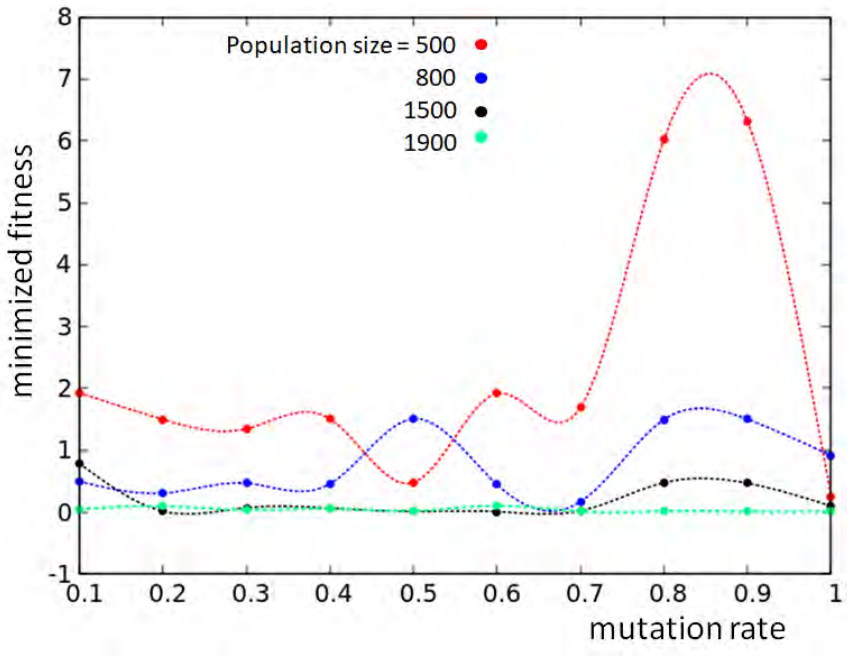


Figure 2.16: The minimum fitness vs. mutation rate plotted with varying population size.

Chapter 3

Computing with interacting chemical oscillators- network designing and teaching strategy

This chapter gives an insight into the working of chemical network based computation. The chapter starts with an overview of how an artificial neural network works. The next section describes a single node chemical oscillator and explains how to perform computation with it. The formation and working of networks of chemical oscillators are also explained. The chapter wraps up by explaining the teaching strategy used to perform information processing with the chemical network.

3.1 Basic working of network based computers

I have already introduced the basic concept of network based computation in chapter 1. As my work also deals with network based computation, I would like to further discuss how a neural network performs computation [83] [84]. To understand it's working, let us consider the simplest form of neural network with two nodes. A schematic representation of it is given in Figure 3.1.

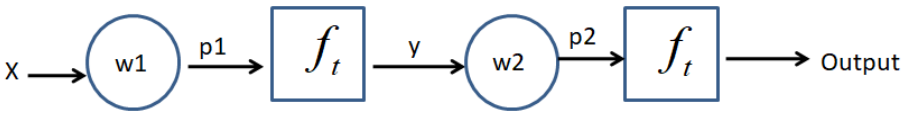


Figure 3.1: Schematic representation of artificial neural network with two nodes. Here x is the input given to the network. w1 and w2 represents the weight multiplied to the input at each stage, giving p1 and p2 as product values respectively. Function f_t is a threshold function.

A single input is given to the network. This input can be considered analogous to the information received by dendrite branches in a neuron (See Chapter 1, Figure 1.3). However for simplification purposes, I considered only single input (as in, an input received from a single branch of a dendrite). The input is then multiplied with a weight value. This weight provides certain kind of importance to the inputs. It is analogous to synaptic connections in between the neurons. If there is a connection, the input should proceed forward and hence the weight value will go up, else the value is less. Then the product value (of input and the weight assigned) passes through the threshold function f_t. The function checks if the obtained product value passes a set threshold value. If the product crosses the threshold value, it is passed on to the next "neuron" where again a new weight is multiplied and the process of passing the answer again to the threshold function repeats. In the final stage, if the processed input value crosses the threshold value, the output is 1 else the output we get is 0. This can be considered analogous to the situation where the decision is taken if the neuron will be activated or not. In supervised learning we consider a training dataset, so we already have our inputs and we beforehand know the output. We want to train a network in a way that it can give our desired output with respect to the input. In case of a neural network, the weight values get tuned for the training purpose.

In my work, I have also created network based computation, but with chemical oscillators. We think that biological neural computing can be represented more realistically by the network of interacting chemical oscillators compared to the typical artificial neural networks with arbitrarily selected activity rules [85].

3.2 Computing with a single chemical oscillator

Let us start with the simplest case by considering modeling of a single chemical oscillator. To simulate a chemical oscillator we used an illumination controlled two-variable Oregonator model.

$$\frac{\partial u}{\partial t} = \frac{1}{\varepsilon}(u - u^2 - (fv + \phi(t))\frac{u - q}{u + q}) \tag{3.2.1}$$

$$\frac{\partial v}{\partial t} = u - v \tag{3.2.2}$$

here, u represents the concentration of the activator and v represents the concentration of the inhibitor. All other parameters have the same meanings as said in Chapter 2. The oscillations obtained are shown in Figure 3.2. I simulated this equation by using the Cash-Karp RK45 method [86]. The time step used here was 10^{-4}. The parameter values of the Oregonator model that was used are: $q = 0.0002$, $f = 1.1$ and $\varepsilon = 0.2$. The initial concentrations of u and v were 0.001 and 0.003 respectively. In Figure 3.2 we illustrate oscillations of activator and inhibitor observed when $\phi = 0.0001$ in time interval [0,50]. With the increase in inhibitor concentration, the activator concentration falls and rises only after the inhibitor decreases. Thus we observe the oscillations.

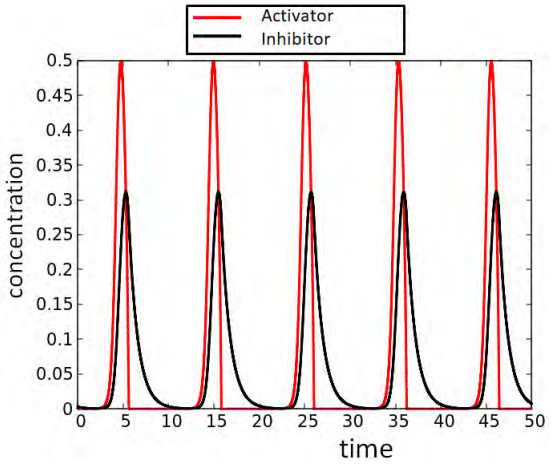


Figure 3.2: Plot of concentration of activator and inhibitor vs. time. The red colour signifies activator concentration and the black one represents inhibitor concentration. Parameter values used are: $q = 0.0002$, $f = 1.1$, $\epsilon = 0.2$ and $\phi = 0.0001$.

In my thesis, I have done my work of information processing by coding information with the activator maxima only. However, we can also code information with the inhibitory oscillations and use the same concepts of inflow and outflow of informations and teaching strategy, as explained in this chapter. We can think of oscillation as: when the activator concentration shoots above a threshold value, it represents the

binary 1 state, and when it hits the low, it represents the binary 0 states. This is analogous to the case of neural networks when the input value needs to pass a certain threshold to get output as "1" state. In my case, the threshold value of activator concentration should have gone above 0.05, for it to be considered as an activator maximum. Also we have used a non-binary information coding process to design a chemical computer.

Let us define ϕ as a function of simulation time:

$$\phi(t) = 0.1 \cdot (1.001 + \tanh(-10(t - t_{illum})) \tag{3.2.3}$$

where t_{illum} stands for the time of illumination. It describes the time before which the system does not oscillate and remain in a steady state. An example of it is shown in Figure 3.3. All the parameters (except the constant ϕ value) used for simulation are the same as mentioned above. The value of ϕ remains high till $t_{illum} - \delta$ ($\delta = 0.1$). The value of δ describes the speed of transition between the steady state and oscillations. It can be reduced by increasing the multiplier under tanh () function. The increased value of ϕ causes the system to go to a steady state and hence we do not witness any oscillations.

This illumination control factor has been used to give input to an oscillator. The process to give input in the form of illumination time is shown below.

Inflow and outflow of information

We give input to a chemical oscillator by converting it into illumination time (t_{illum}) that appears in Eq.3.2.3. The inputs are given in the form as:

$$Input^r_{\#} = t_{start} + (t_{end} - t_{start}) * In^r_{\#} \tag{3.2.4}$$

where t_{start}, $t_{end} \in [0, t_{max}]$ (t_{start}, t_{end} are random parameters belong in the range of zero to total time of simulation). $In_{\#}$ is the input which we will be "feeding" the oscillator with. r denotes the input number in a dataset. The meaning of Eq. (3.2.4) can be decoded as : the input receiving oscillator remains inactive between time $[0, Input^r_{\#} - 0.1]$. The value of $Input^r_{\#}$ will corroborate with the value of t_{illum} in Eq. (3.2.3). So, two steps are followed to feed input to a oscillator:

$Step1: \ Conversion\ of\ data\ inputs\ to\ illumination\ time \quad t_{start} + (t_{end} - t_{start}) * In^r_{\#}$

$$Step2: \ Define\ the\ range\ of\ inactivity\ [0, Input^r_{\#} - 0.1]$$

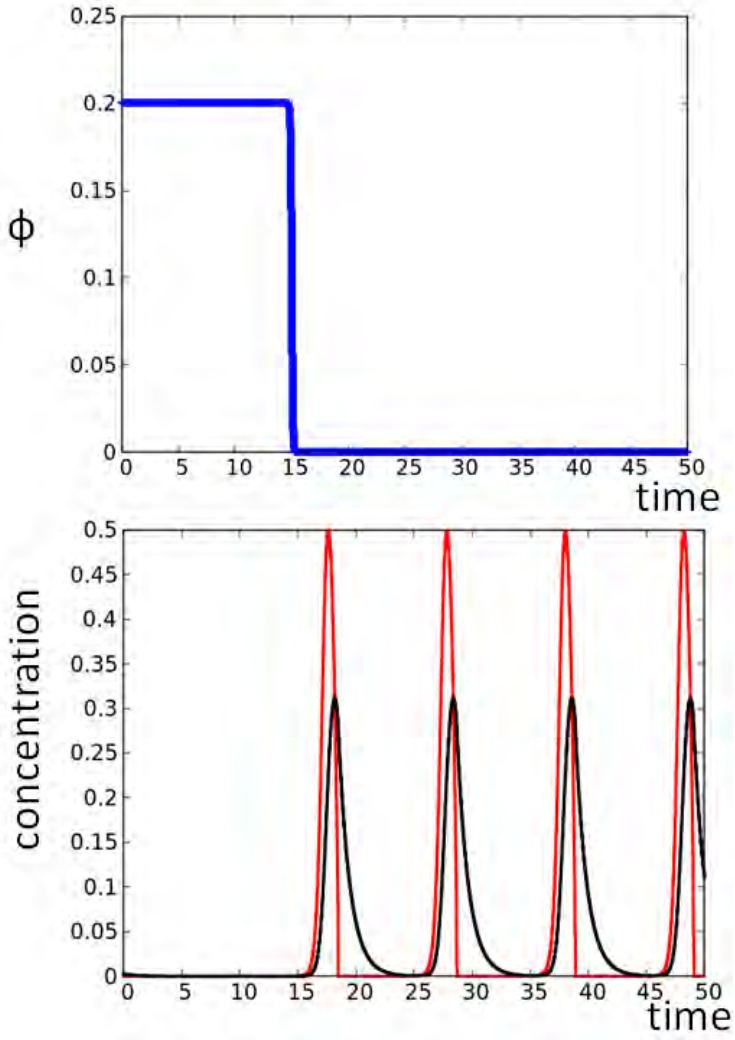


Figure 3.3: In top figure: Plot of ϕ as a function of t ($t_{illum} = 15$). The value of ϕ is high till time $t_{illum} - 0.1$ and after that it gradually decreases. In the bottom figure the oscillations of activator and inhibitor concentrations are shown for $\phi(t)$ illustrated above. Parameter values used for simulation are: $q = 0.0002$, $f = 1.1$ and $\epsilon = 0.2$.

Each input value can produce a different time of illuminations. Let us say that $In_{\#} \in \{0.2, 0.4, 0.6\}$ and assume that $t_{start} = 10$, $t_{end} = 40$ and $t_{max} = 50$. So our 1^{st} step is to always convert the inputs to the illumination time. Following Eq (3.2.4), the relationship between input $Input^r_{\#}$ and the range of inactivity of the input oscillator following the step 2 is:

r	$In^r_{\#}$	$Input^r_{\#}$	Range of inactivity
1	0.2	16	0-15.9
2	0.4	22	0-21.9
3	0.6	28	0-27.9

Now on simulating a single oscillator (with same parameters in Figure 3.3, initial concentrations of u and v : 0.001 and 0.003) by considering these different ranges of

inactivity, we can see the number of oscillations observed for each different case of input is different (Figure 3.4). In the analysis below, we treat the activator maxima as our output. Our main thought behind it is with different counts of maxima of each case, we can differentiate between the inputs and it's unique characteristics. For example, suppose the oscillator is asked to predict if the value, input x is greater than 0.2 and less than 0.6. So, the first step is to change the input into illumination time and feed it to the oscillator. Now, I have already trained the oscillator with different inputs and observed the oscillation distribution graph (Figure 3.4). For input x, if the oscillator gives 3 activator maxima as output, then considering Figure 3.4 data, there are high chances that the input given is $0.2 < x < 0.6$. Thus this chemical oscillator has predictive power achieved by top-down approach.

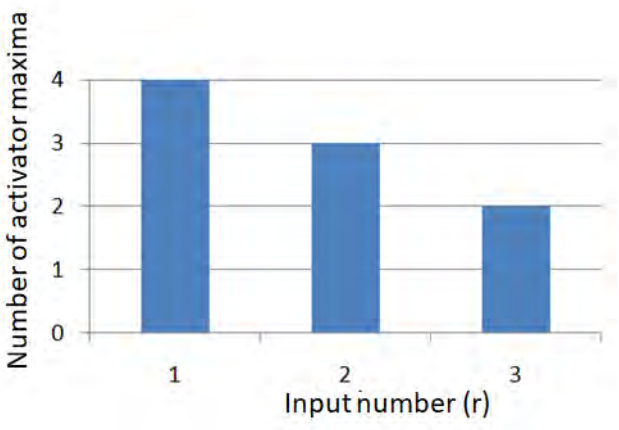


Figure 3.4: Oscillation distribution graph is plotted for each input.

The situation of changing the input to illumination time can be considered as analogous to the concept of changing an input by multiplying it with a weight in the artificial neural network. Just like the weight in the neural network can be finely controlled, similarly the illumination time associated with the input can be controlled by adjusting the parameter values of t_{start}, t_{end}.

The value of parameters: t_{start}, t_{end} are the important ones to perform information processing. Their values should be chosen in a way so that it can distribute the "input illumination time" to cover a good amount of range in $[0, t_{max}]$, which is an important criterion for the differentiation of inputs based on the number of oscillations. Considering the above example, we can see that with respective values of t_{start}, t_{end} , the input illumination time got a good distribution between the range 10 to 40. Now, if the parameter values are chosen in a wrong manner, then the input illumination time will be in a smaller range which in turn might result in giving a similar number of oscillations in every case (say, we might get the same number of oscillations irrespective of the input $0.2 < x < 0.6$ or $x < 0.2$) . Thus the oscillator loses it's predictive

power. In this particular case, we can easily choose some values of t_{start}, t_{end} that satisfy the criteria for getting different activator maxima for cases $0.2 < x < 0.6$ and $x < 0.2$. But in complex problems where there are more inputs, one oscillator will not work. In fact we will need a network of chemical oscillators interacting with each other. The increase in the number of oscillators, the introduction of more than one input and the interactions among the oscillator will make it difficult to choose these parameters by ourselves. Hence we use evolutionary algorithm to do this work. The method of finding the parameter values is discussed in Section 3.4.

3.3 Designing of a network of chemical oscillators described by Oregonator model

In my thesis, I have performed information processing with a network of chemical oscillators. We assumed that interactions between the oscillators #k and #j appeared as the results of reactions involving the activators U_k and U_j of these oscillators:

$$U_j + B_j \rightarrow U_k + C_k \tag{3.3.1}$$

$$U_k + B_k \rightarrow U_j + C_j \tag{3.3.2}$$

with the same reaction rate constant k_B.

We also assumed that the activator of each reaction could spontaneously decay with the reaction rate constant k_D:

$$U_j + D_j \rightarrow products \tag{3.3.3}$$

Here symbols B, C and D denote other molecules involved in these reactions.

Therefore, the changes in concentrations of U_j and U_k as the result of reactions (3.3.1) and (3.3.2) are:

$$\frac{\partial u_j}{\partial t} = -k_B b_j u_j \tag{3.3.4}$$

$$\frac{\partial u_k}{\partial t} = -k_B b_k u_k \tag{3.3.5}$$

and the changes in concentration of U_j as the result of reaction (3.3.3) is:

$$\frac{\partial u_j}{\partial t} = -k_D d_j u_j \tag{3.3.6}$$

In Eq.(3.3.4-3.3.6) b_j, b_k and d_j denote concentrations of B_j, B_k and D_j respectively. We assume that these concentrations were high with respect to concentrations of activators involved and the same for all oscillators. Therefore, the concentrations of B, C, and D were regarded as constants during the network evolution, and there is no need to include them in the model of network evolution. Let us introduce the parameters $\alpha = k_D d_j$ and $\beta = k_B b_j$. Keeping in mind that α and β can be modified by concentrations of B_j and D_j, we can treat them as free parameters that can be easily adjusted. Including reactions listed above, the time evolution of the network is described by the following set of kinetic equations:

$$\frac{\partial u_j}{\partial t} = \frac{1}{\varepsilon}(u_j - u_j^2 - (fv_j + \phi_j(t))\frac{u_j - q}{u_j + q}) \\ - (\alpha + \beta \sum_{k=1,m} s_{j,k})u_j + \beta(\sum_{k=1,m} s_{j,k}u_k) \tag{3.3.7}$$

$$\frac{\partial v_j}{\partial t} = u_j - v_j \tag{3.3.8}$$

where,

$$\phi_j(t) = 0.1 \cdot (1.001 + \tanh(-10(t - t^j_{illum})) \tag{3.3.9}$$

The last two terms in Eq. 3.3.7 represent the coupling in between k^{th} and j^{th} oscillators and the activator decay. The symbol $s_{j,k}$ is defined as:
$s_{j,k} = 0$ if j = k or if j≠k and oscillators # j and #k do not interact,
$s_{j,k} = 1$ if j ≠k and oscillators # j and #k do interact. Here, m is the total number of oscillators considered in a network. Mathematically the terms resulting from the processes 3.3.1 and 3.3.2 have a similar form to those describing the coupling between CSTRs resulting from the exchange of equal volumes of reagents (and assuming that there is no transport of inhibitor).

The role of α is a rate decay constant. To understand its behaviour towards oscillation, I considered simulating a single oscillator (initial concentration $u_j = 0.001$, $v_j = 0.003$) with Cash-Carp RK45 method (time-step: 10^{-4}, $t_{illum} = 0$ and all other parameters were similar to Figure 3.3) and observed the oscillatory nature with respect to different values of α (Figure 3.5). We saw with increase in the value of α the amplitude of concentration decreased and after a certain threshold value the decay rate of concentration of activator was too high that the oscillations were not witnessed. The threshold value of α lies somewhere in [1.5,2].

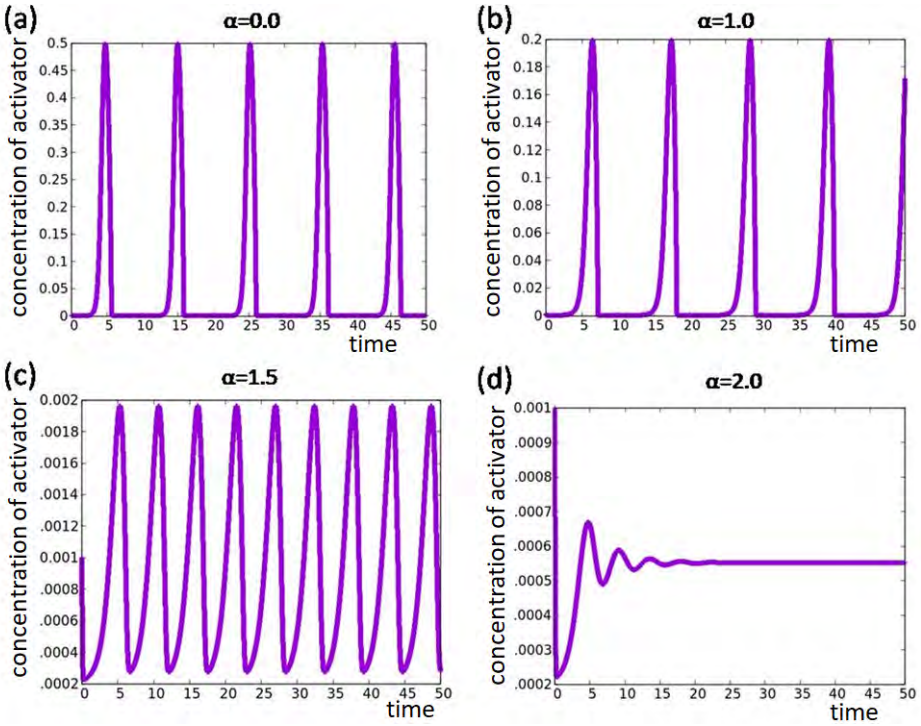


Figure 3.5: The oscillations of the concentration of the activator of a single oscillator has been plotted with respect to time for (a) $\alpha = 0.0$ (time period of oscillations = 10.97) (b) $\alpha = 1.0$ (time period of oscillations = 10.21) (c) $\alpha = 1.5$ (time period of oscillations = 5.39) (d) $\alpha = 2.0$

To understand the effect of parameters α and β on oscillation synchronization between coupled oscillators, I took two oscillators (i, j) in consideration (considered initial concentration : $u_j = 0.000202$, $v_j = 0.0315$, $u_i = 0.03$, $v_i = 0.0001115$, $t^j_{illum} = t^i_{illum} = 0$ and all other parameters were same as in Figure 3.3) and observed the time difference (t_d) between activator maxima of these two oscillators. With $\alpha = 0$ and $\beta = 0$, t_d was 5.96.

At first I considered $\alpha = 0$ and varied β to observe the change in t_d (Figure 3.6). We can note that with just an increase in $\beta = 0$ to $\beta = 0.1$, the value of t_d decreases almost instantaneously from 6 to 1.8 causing a strong synchronization. With a further increase in the value, the strength of synchronization increases.

When $\beta = 0$ and α was varied, a reverse phenomenon was noted (Figure 3.7). The rate of decay of activators in both the oscillators was high and t_d between the oscillations gradually increased with an increase in value of α.

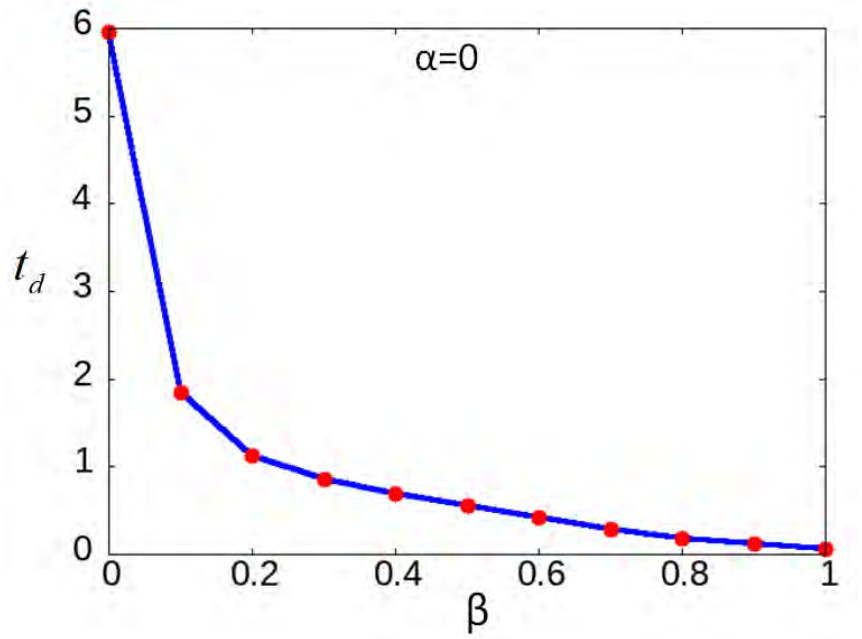


Figure 3.6: Time difference(t_d) between activator maxima of two oscillators has been plotted against coupling factor β with $\alpha = 0$. With increase in value of β the coupling strength between the oscillators increases.

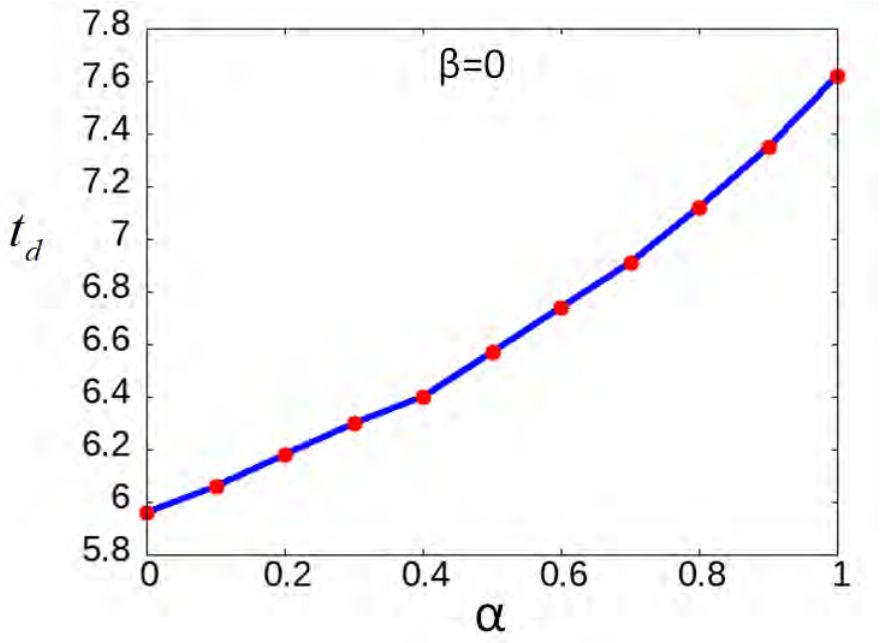


Figure 3.7: Time difference(t_d) between activator maxima of two oscillators has been plotted with respect to the value of α with $\beta = 0$.

I also plotted the values of t_d with a change in β with α as a parameter (Figure 3.8). It has been seen that, previously when α was set to zero (Figure 3.6), a very strong instantaneous synchronization was witnessed among the oscillators. However, in this Figure 3.8 we can see, that with an increase in value of α, the coupling between

the oscillators was not so instantaneous. When $\alpha > 0.6$ and $\beta = 0.1$, t_d was very high. t_d gradually decreased and the coupling became strong with an increase in value of β. When $\alpha < 0.6$, even at a very small value of $\beta = 0.1$, we were able to witness coupling. With the increase in the value of β, (irrespective of what the value of α is) strong coupling was witnessed. The coupling was the strongest when α had a smaller value (in this figure $\alpha < 0.8$) and β had a larger value. Thus we can say that for the considered parameters of the Oregonator model (as well as a few others we tried) the interactions between oscillators described by reactions 3.3.1 and 3.3.2 were difficult to control without the process of the rate of decay (Eq 3.3.3). The model, depending only on the value of β gave too weak or too strong coupling between oscillators. The introduction of reaction (Eq 3.3.3) allowed us to control the value of activator concentration around its maximum and to moderate these interactions without the need to optimize all parameters of the Oregonator model.

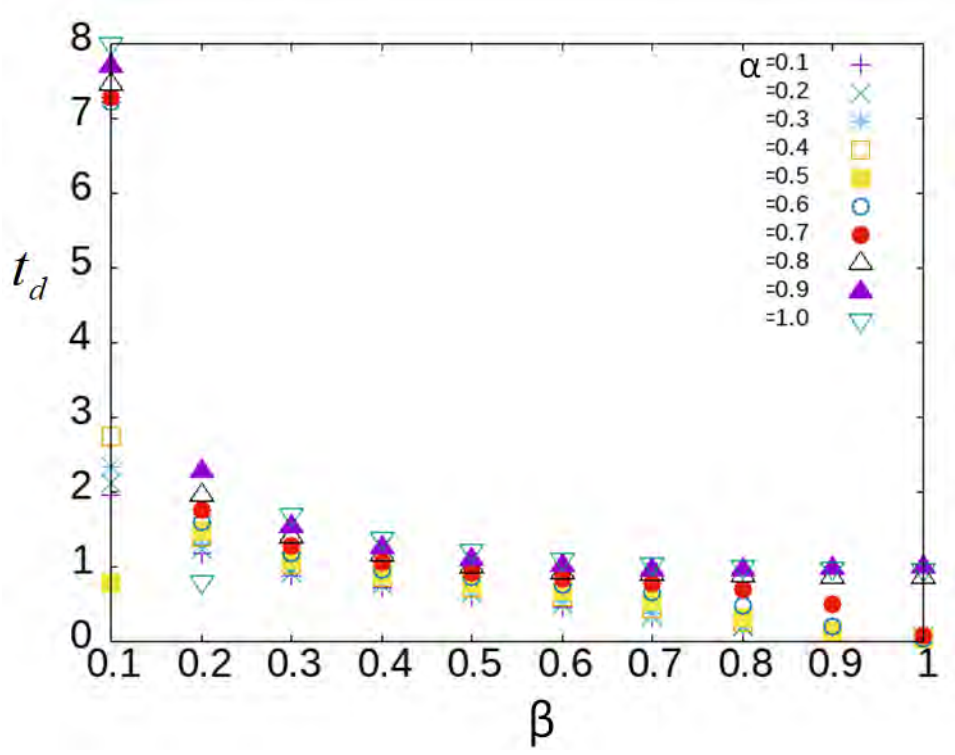


Figure 3.8: The change in t_d has been shown with respect to β, with different values of α as a parameter

An example of 6-interacting chemical oscillators

To study the dynamics of a network of chemical oscillators, I designed a model consisting of six interacting oscillators (m = 6) each represented by a two-variable Oregonator model. The time evolution of the network is described by the following

set of kinetic equations:

$$\frac{\partial u_j}{\partial t} = \frac{1}{\varepsilon}(u_j - u_j^2 - (fv_j + \phi_j(t))\frac{u_j - q}{u_j + q}) - (\alpha + 2\beta)u_j + \beta(\sum_{i=1,m} s_{j,i}u_i) \tag{3.3.10}$$

$$\frac{\partial v_j}{\partial t} = u_j - v_j \tag{3.3.11}$$

The last two terms in Eq. 3.3.10 represent the coupling in between i^{th} and j^{th} oscillators. Here $\sum_{i=1,m} s_{j,i} = 2$. In the considered network, the interactions are fixed and illustrated in Figure 3.9.

The initial concentrations of each oscillator were chosen randomly and each oscillator had different starting concentration than the other (only based on concentration value and if we consider $\alpha = \beta = 0$ and $t^j_{illum} = 0$ for $j \in \{1, 2, 3, 4, 5, 6\}$ the order of oscillations were oscillator 1,2,4,5,6 and 3). The chosen value of α and β were 1.2 and 0.7 (with these values and $t^j_{illum} = 0$ for $j \in \{1, 2, 3, 4, 5, 6\}$, the order of flow of excitations were $4 \to 5 \to 4 \to 6 \to 3 \to 1 \to 2$). The total time of simulation (t_{max}) was 50. All other parameter values were the same as mentioned in Figure 3.3. Cash Carp RK45 method was used for simulation with time steps 10^{-4}.

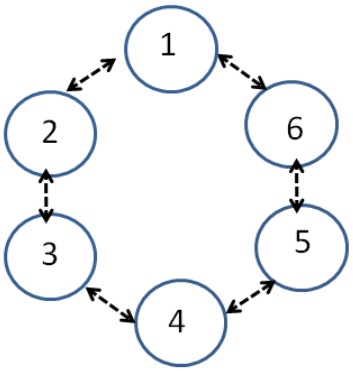


Figure 3.9: The network is made up of six oscillators that interact with one another. All of the oscillators are numbered. Interactions between the oscillators are represented by the dotted lines between them. The parameter values that are used for simulation are same as in Figure 3.3

Let us externally control the system by illumination time and decide arbitrarily the illumination time of each oscillator. The illumination time chosen were: $t^1_{illum} = 5$, $t^j_{illum} = 10$ with $j \in \{2, 3\}$ and $t^j_{illum} = 15$ with $j \in \{4, 5, 6\}$. The corresponding function $\phi(t)$ for every oscillator are shown in Figure 3.10.

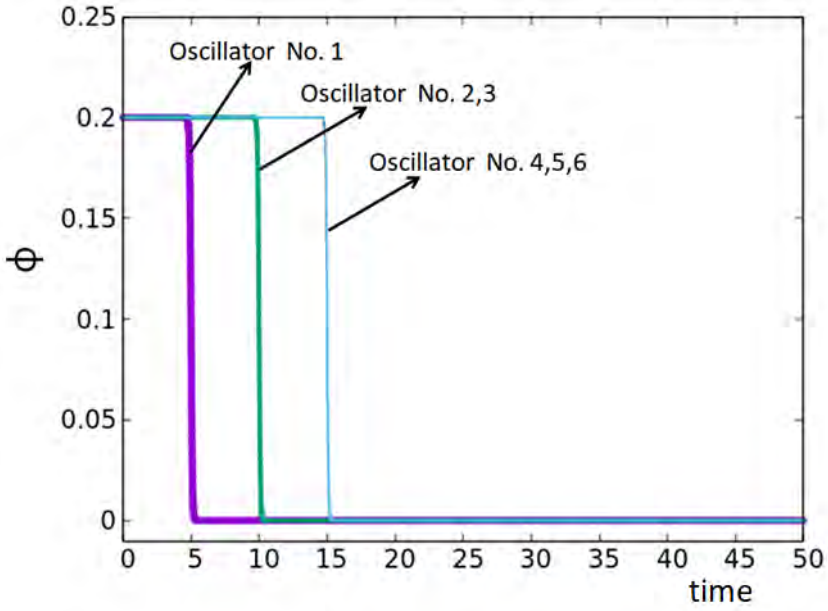


Figure 3.10: Functions $\phi_j(t)$ of all oscillators present in the network in Figure 3.9.

In Figure 3.11 we have noted the time in which all the oscillator oscillates. In Figure 3.11(a) the order of oscillation of the oscillators is kind of controlled externally by the assigned illumination time. The order of oscillation changed from $4 \rightarrow 5 \rightarrow 4 \rightarrow 6 \rightarrow 3 \rightarrow 1 \rightarrow 2$ (without illumination control) to $2 \rightarrow 1 \rightarrow 3 \rightarrow 6 \rightarrow 4 \rightarrow 5$. This new order of flow of excitation was due to the effect of forced illumination. In the time range of 30.80 to 30.89, the same kind of order of oscillation was maintained. After this time, although every oscillators were almost coupled but the excitation passed from fully coupled oscillator 1, 2 to oscillator 5 followed by oscillations witnessed in oscillator 6, 4 at a same time and then to oscillator 3. If we observed (a) and (b), the order of passing on the excitation from one oscillator to another, which was forcefully induced by us with different illumination times, was kind of retained till the time of 30.89. This can be said or referred as a memory of the system. After time 30.89, the complexity of interactions came into play and total synchronization was witnessed in between oscillators.

Now till this point, we have chosen many values of parameters randomly. We do not know what exactly should be the values of α, β, total time of simulation (t_{max}). Nor we know what the values of t_{start}, t_{end} should be. In the next section, I have presented a teaching method based on which we can find the optimum values for these parameters.

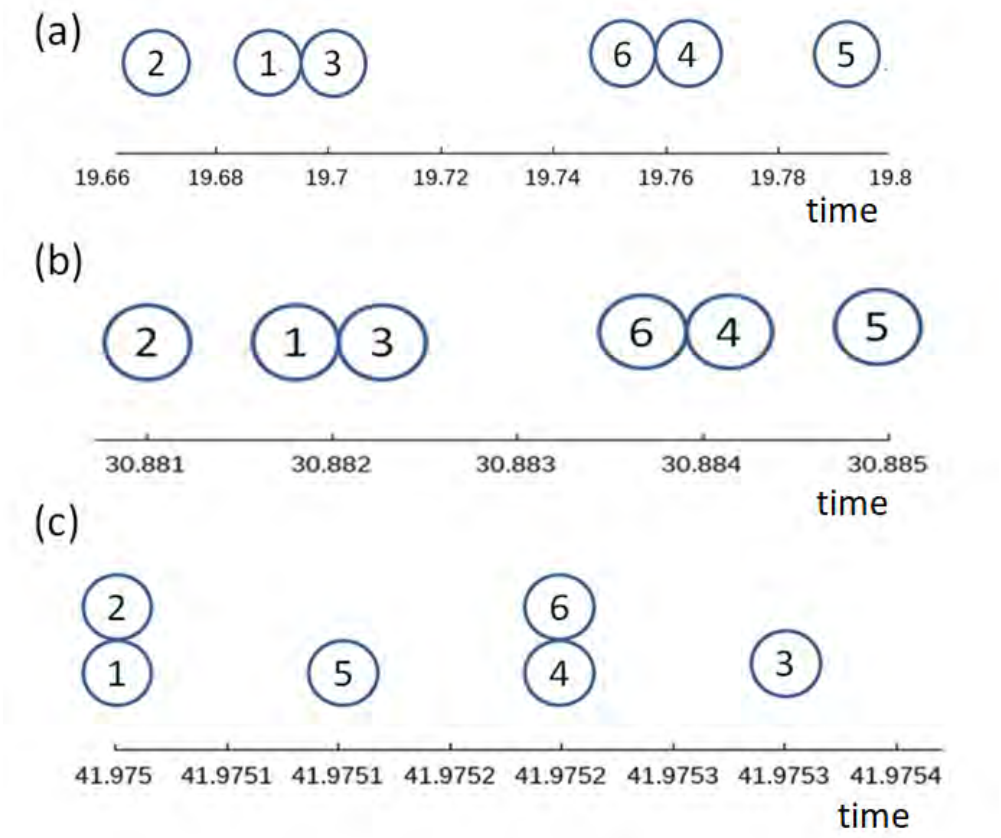


Figure 3.11: Figure depicts the time in which the oscillators oscillate. Each circle symbolizes $\#i^{th}$ oscillator. If the oscillators are stacked on top of one another, it means that they oscillated in the same time.

3.4 The teaching strategy

We used supervised learning to teach our system. It uses something called the classification technique to train the network of chemical oscillators. The basic requirement of using this technique is the dataset used to train should be "labelled". Let us take a simple example : in the shown Figure 3.12, suppose we want to train our system in a way so that it can say if the point is blue or red. So to train it we give the point's coordinate (x,y) as input and also we give information to the system of how the points are labelled. If points are labelled by zero, then the point is red and if labelled one then the point is blue. So this technique helps the system to be trained in a way that it can create an imaginary line by finding certain kinds of patterns in the existing dataset. This line can separate two labelled groups. In this simple case the system can easily find the pattern in the dataset which is: when $x < 0.5$ the point is always red and when $x > 0.5$ the point is blue. So following this pattern, the system draws an imaginary line AB. After training we can say, if the (x,y) coordinate lies on left of the line AB then it most probably is red in colour. This is the simplest case

where even our naked eyes can find certain patterns in the dataset. But in complex datasets, it is difficult to find patterns on our own. So we apply an evolutionary

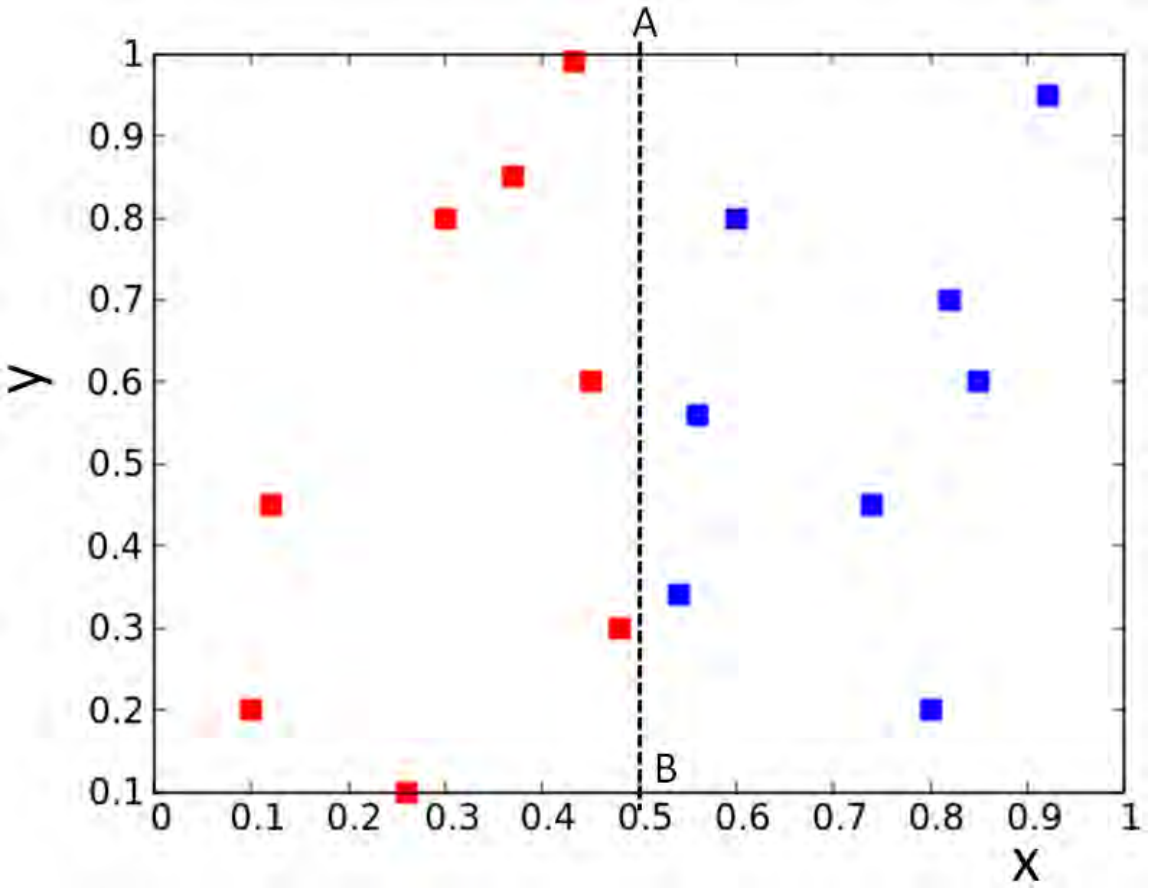


Figure 3.12: The graph shows the distribution of random points(red or blue) within the range of x, y coordinate ($x, y \in [0, 1]$). Line AB is an imaginary boundary which can actually distinguish between point red and point blue. If points lie on left side of the line it is surely red and if the point lie on right side of the line it is surely blue.

algorithm to find the optimum values for the parameters of our chemical system, so that based on these parameter values, we get such outputs (number of oscillations) which will be different for different labelled groups. The distinguished output result will give information about the group to which the considered input belongs. The parameters (the set of the population considered) to be optimized are:

1. α, β - the rate of decay constant and the coupling factor,
2. t_{start}, t_{end} - the relationship between inputs and illumination time,
3. t_{max} - the total time of simulation,
4. the oscillator type in considered network geometry,
5. the illumination time of normal oscillators.

The oscillator type means that the selection of the oscillator, whether it will be an input one where the data will be fed or it will be a normal oscillator whose function will be just to integrate the signals in the network, are decided by the evolutionary algorithm (EA). Also if it is a normal oscillator, whether it will be allowed to oscillate or it will be illuminated the whole interval $[0, t_{max}]$ or in a part of it is decided by the optimized parameter t^j_{illum}.

As seen in the example in Chapter 2, the optimized parameters a_0, a_1, a_2, a_3 were found based on the fitness function described. In that example the desired value of the parameters is those which minimize the fitness function. This fitness function value gives an idea to the EA in which direction of the search space it should move to find the optimum parameter values. Keeping that in mind, in this case we defined the fitness function in the form of mutual information which goes as follow:

$$I(Z_k; Z_O) = H(Z_k) + H(Z_O) - H(Z_k Z_O) \tag{3.4.1}$$

where H is the Shannon entropy, Z_O refers to the output class; the class containing the labels of the input cases, Z_k is the excitation class; it stores the number of oscillations (of k^{th} oscillator) received for each case of inputs and $I(Z_k; Z_O)$ tells us the mutual information (MI) shared between this two class in bits. So, if the number of excitations of a particular oscillator is always the same, independent to the class level, then there is no information that the oscillation number can provide us about the labels or groups, thus $H(Z_k) = 0$ and $H(Z_k Z_O) = H(Z_O)$, which makes $I(Z_k; Z_O) = 0$. Now if the excitation class is perfectly correlated with the output class then, $H(Z_k) = H(Z_O)$ and $H(Z_k Z_O) = H(Z_O)$, so $I(Z_k; Z_O) = H(Z_O)$, which tells us that all the information about the dataset can be extracted from the number of excitations of the output oscillator [22].

The main purpose is to find a higher mutual information which will signify that the number of oscillations received is correlated with the class labels. In simpler words, EA finds the optimized parameter for which the value of MI will be higher. The oscillator with the maximum mutual information will be our output oscillator. The output oscillator can either be a normal one or an input oscillator.

In most of our information processing works in chapter 4 and 5, the following rules are used (unless mentioned otherwise) : The size of the population we have always used were 100 or 200. Top 2% of the population based on high mutual information was reserved for the next generation as it is, i.e neither did it undergo recombination nor mutation. Rest 98% of the population results from recombination within top 40% of the population as potential parent candidates. The population size always remains

constant in every generation. Mutation occurred with a certain mutation rate. In order to obtain these new offspring, the operations performed are as follows: [87]

- Another network is created by mixing parameters of two networks selected from top 40% of the population. They undergo parameter recombination and a new offspring was created with new values of t_{start} and t_{end} , forming a new set of parameters: $t_{max}, t_{start}^{offspring}, t_{end}^{offspring}, \alpha, \beta$. The oscillator type, whether it will be an input oscillator or a normal one, also underwent recombination. The parents were again chosen randomly the oscillator type can be altered or recombined. There were no constraints given among selecting the type of oscillator. There could be many or no input oscillators for a given input. There may be no normal oscillators if it optimizes the network. And an input oscillator was allowed to be the output one. If oscillators are selected as normal ones, then the modification on the illumination times (t_{illum}) of normal oscillators was also performed by recombination by randomly selecting a parent from 40% of the best fitted population and copying its values of t_{illum} to the next generation.

- Mutation on the coupling factors and on rate of decay of activator: The coupling factor(β) and the rate of decay of activators(α) were mutated with a mutation rate of 0.5. The mutated values of α and β were the sums of a fraction of their old value and a random number.

The parameters of the Oregonator model (f, ϵ and q) were fixed and did not undergo optimization.

3.5 Computing with network of chemical oscillators

The basis of computation with a network of chemical oscillators is similar to what I have explained in case of a single oscillator (section 3.2). However as there is more number of oscillators now, many inputs can be given to the system at the same time. Each oscillator can be fed with one input. It was seen in the case of neural networks, that the value of weight (which is to be multiplied to the input) could be less , for the input to be not considered. Similarly if our system finds an input to be not important, it could choose the "weight" of t_{start} and t_{end} in a way that t_{illum}^{j} of the input receiving j^{th} oscillator will be close to t_{max}. Thus the oscillator will be inactive for most of the simulation time and hence that particular input wont have any major role in influencing the dynamics of the system.

I have presented Figure 3.13 which summarizes the computation method. The figure gives the step by step measures to be taken to train a network of interacting chemical oscillators.

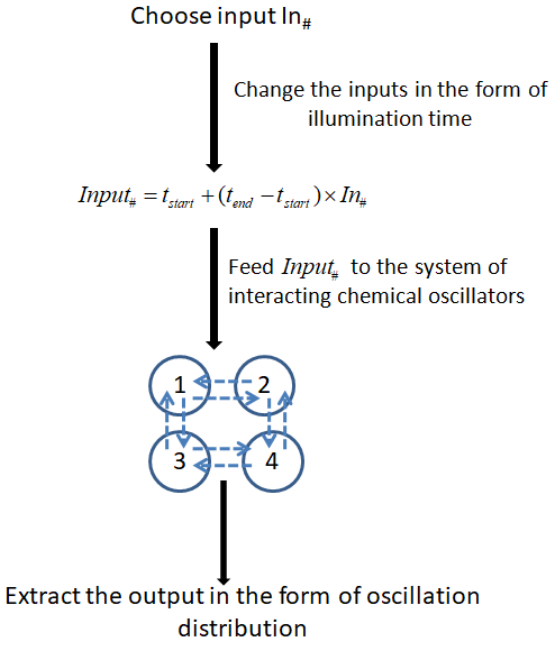


Figure 3.13: A schematic representation of how a 4 interacting network of chemical oscillators can be taught to perform information processing has been shown. The oscillators have been marked by numbers. The dotted lines between them show the interactions.

The steps are as follows:

- Decide on what input (and how many inputs) we want to feed to the system, to perform a certain information processing task. Prepare a dataset, say , $D = \{(x_n, y_n, g_n), n = 1, N\}$ where x_n, y_n are the inputs of n^{th} case and g_n is the attached label (0 or 1) to the inputs.
- Change input x, y in the form of illumination time of the input oscillators.
- Use the evolutionary algorithm to the system to get an optimized network for the maximum value of mutual information.
- Extract the output in the form of the number of oscillations from the output oscillator.

Chapter 4

Networks of chemical oscillators as classifiers of geometrically inspired problems

The chapter shows the results we obtained by training a network of chemical oscillators to solve two geometrically oriented problems. One of which is to train the network to find the colour of a randomly selected point on the Japan flag. The other problem is to differentiate between two intertwined spirals. In this problem I also give an insight into the dependency of interaction complexity in the system vs. the accuracy in the determination of spirals.

To understand the intelligence power of a network of chemical oscillators as a classifier, we considered two geometrically oriented problems and optimized the networks to check their solving capacity.

4.1 Classification of point colour on the Japan flag by a network of interacting chemical oscillators

In this section, we concentrate on a geometrically oriented problem illustrated in Figure 4.1. The results related to this topic are published in [88] We considered a red disk (sun) located in the mid of a white square (here represented by the Cartesian product $[0,1] \times [0,1]$). We postulate that a chemical computer can answer if a randomly selected point $(x, y) \in [0,1] \times [0,1]$ is located in the red or in the white region. To make the problem difficult, the disk radius $r = (\sqrt{2\pi})^{-1}$ is selected such that the areas of the sun and the white region are equal. Therefore, a device that gives a random answer or a device that gives the same answer (red, white) to all inputs solve the problem with 50% accuracy (or with 50% chance to obtain the wrong answer). We show below that a chemical medium can solve the problem with much higher accuracy.

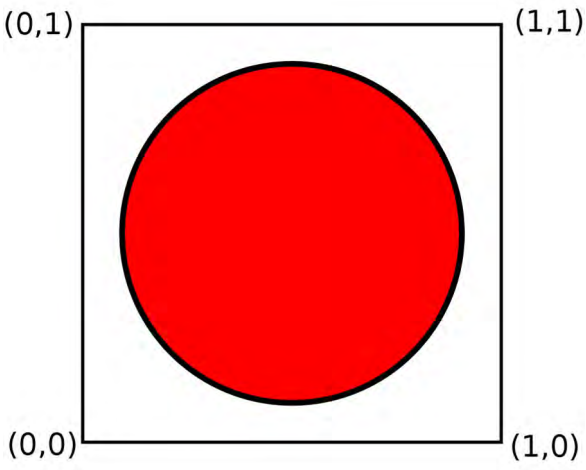


Figure 4.1: The graphical illustration of the problem that is solved by the optimized network. The disk is located at the square center. Its radius was selected such that the areas of the red sun and the white surrounding are equal. The network is supposed to answer if a randomly selected point $(x, y) \in [0,1] \times [0,1]$ belongs to the red or to the white regions.

Model

We postulate that the problem of attributing colour to a point on the Japan flag

defined by its coordinates can be approximately solved by a network of interacting chemical oscillators (cf. Figure 4.2). The time evolution of each individual oscillator is described by two-variable Oregonator model combined with the additional reactions (3.3.1 - 3.3.3), as mentioned in chapter 3.

Within our model, the time evolution of the network is described by the following set of kinetic equations:

$$\frac{\partial u_j}{\partial t} = \frac{1}{\varepsilon}(u_j - u_j^2 - (fv_j + \phi_j(t))\frac{u_j - q}{u_j + q}) - (\alpha + 3\beta)u_j + \beta(u_1 + u_2 + u_3) \tag{4.1.1}$$

$$\frac{\partial v_j}{\partial t} = u_j - v_j \tag{4.1.2}$$

The parameter values were same as mentioned in Figure 3.3 (Chapter 3). The initial concentrations (at $t = 0$) were $u_j = 0.0002$ and $v_j = 0.0002$. The input is introduced as the illumination time of the input oscillator.
If the j^{th} oscillator functions as the input one for the coordinate $x \in [0, 1]$ then:

$$t^j_{illum} = t_{start} + (t_{end} - t_{start}) \cdot x \tag{4.1.3}$$

Having in mind the symmetry of the problem the input oscillator #k for the y value is inhibited for time:

$$t^k_{illum} = t_{start} + (t_{end} - t_{start}) \cdot y \tag{4.1.4}$$

and the values of parameters t_{end} and t_{start} in Eq.(4.1.3) and Eq.(4.1.4) are identical. We here already assumed oscillator #1 to be our input oscillator for input x and oscillator #2 for input y. The type of oscillator #3 is determined by the optimization procedure. The function of oscillator #3 is undecided, so it can be an input oscillator

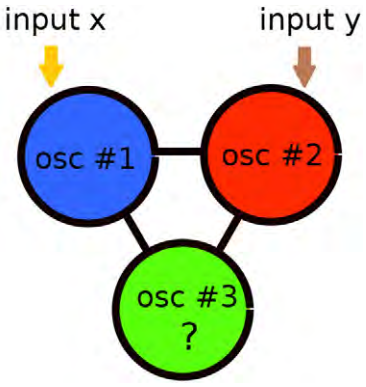


Figure 4.2: The structure of the considered chemical oscillator network. The black line indicates interactions between oscillators. The type of oscillator #3 was optimized to achieve the maximum mutual information between locations of points and the number of activator maxima.

or a normal one.

Network Optimization

In order to apply the evolutionary algorithm we need a training dataset $D_S = \{(x_n, y_n, g_n), n = 1, N\}$ of records in the form (x, y, g) where $x, y \in [0, 1]$ and $g \in \{0, 1\}$. The numbers x, y denote the point coordinates and g is the record type; $g = 1$ for points located in the red disk and $g = 0$ for points outside it. Here we used D_S of $N = 800$ records with randomly generated points inside $[0, 1] \times [0, 1]$ located as illustrated in Figure 4.4(A). Let us also introduce the discrete random variable of record types Z_O, defined as: $Z_O = \{g_n, n = 1, N\}$. Also, assume that $z_1(n)$, $z_2(n)$ and $z_3(n)$ are the numbers of activator maxima observed on oscillators #1, #2, and #3, respectively. Now, let us combine the results obtained for all inputs from the training dataset together and introduce the random variables $Z_k = \{z_k(n), n = 1, N\}$ for k^{th} oscillator, $k \in \{1, 2, 3\}$. We postulate that information about the point colour can be extracted from the number of activator maxima recorded on a selected oscillator of the network during the time interval $[0, t_{max}]$. The quality of an oscillator network for solving a specific problem can be estimated in the following way. Let us consider a record $(x_n, y_n, g_n) \in D_S$ and study the network evolution for the input (x_n, y_n). The training dataset (Figure 4.4(A)) contained 377 points in the sun area (red) and 423 points in the surrounding region (blue). The adjustable parameters defining each network (type of oscillator #3, t^3_{illum}, t_{max}, t_{start}, t_{end}, α and β) were randomly generated. The population size considered was 200. And we used an evolutionary algorithm to find the optimized set of parameters. The fitness (Chapter 3 Eq. 3.4.1) of each network was calculated using the whole training dataset. The next generation comprised of 20% of most fit networks of the previous population and of 80% of offspring generated by recombination and mutation operations applied to oscillators from top 50% networks of the previous population. We randomly selected two parents from 50% of the fittest networks and next recombined their parameters randomly to obtain an offspring. After recombination, we applied mutations to randomly selected parameters. The probability of this operation was selected, such that on average, a single parameter of the network was mutated. The maximum change in the chosen parameter value was restricted to 10% of the original one. Next, the fitnesses of networks belonging to the new generation were calculated and the procedure was repeated. The optimization procedure was continued for 1000 generations.

Results

The optimization procedure returned the network illustrated in Figure 4.3(A) in which the oscillator #3 is of the normal type and $t^3_{illum} = 6.37$. The other parameters of the network are: $t_{max} = 20.23$, $t_{start} = 3.78$, $t_{end} = 12.10$, $\alpha = 0.849$, $\beta = 0.251$. The oscillator #3 is also the output one. For each input, we observed one or two activator maxima at the output oscillator.

Figure 4.4(C) illustrates the relationship between the point location and the number of activator maxima. The statistic of the network outputs for all inputs from the training dataset is illustrated in Figure 4.4(D). For the majority of cases located in the sun area, we observed a single maximum (369 cases); two maxima appeared for only 8 points (marked by dark blue dots in Figure 4.4(C)). On the other hand, for most of the background points, two activator maxima were observed (393 cases), whereas only

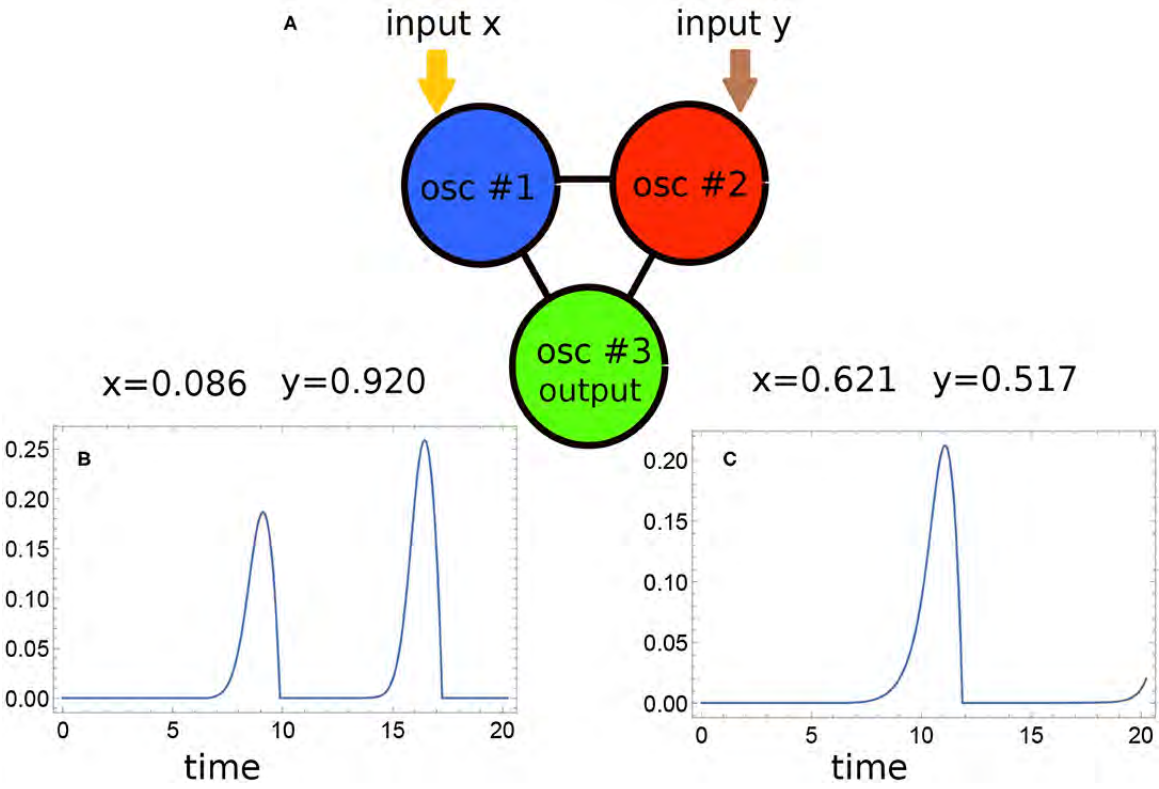


Figure 4.3: (A) The idea of information processing with a network of oscillators illustrated using the parameters of the optimized network. Oscillators #1 and #2 are inhibited by times related to the values of x and y respectively. The output information is coded in the number of activator maxima observed on the oscillator #3. The time evolution of activator observed on this oscillator for points located in the white and red regions are shown in subfigures (B) and (C) respectively. The evolution was calculated using the parameters of the optimized network: the oscillator #3 is the normal one with $t^3_{illum} = 6.37$, $t_{max} = 20.23$, $t_{start} = 3.78$, $t_{end} = 12.10$, $\alpha = 0.849$, $\beta = 0.251$.

30 cases produced a single maximum (dark red points in Figure 4.4(C)). We can use the majority rule and declare that all points for which the network outputs a single maximum correspond to the sun, and all points for which two maxima are observed correspond to the surrounding white area. For the training dataset D_S such majority rule leads to $(369 + 393)/800 \sim 0.95$ accuracy. It is interesting that the distribution of incorrectly attributed cases is not rotationally symmetric.

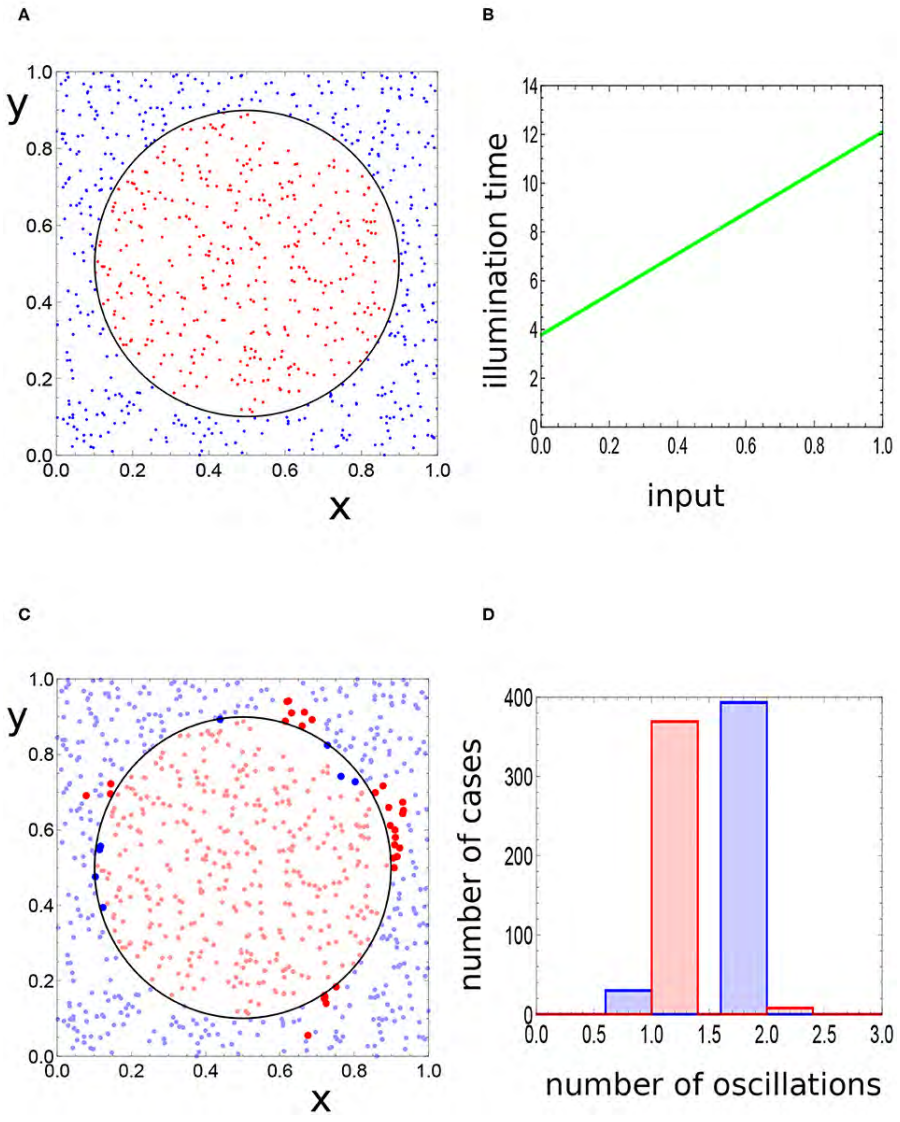


Figure 4.4: (A) Locations of 800 points representing the records of the training dataset D_S. The red points are in the sun area, the blue ones outside it. (B) The relationship between t_{illum} and the input value, obtained for the optimized network solving the Japan flag problem (Eqn.4.1.3 and 4.1.4). (C) The response of the optimized network to inputs from the training dataset. The dark red points are located outside the sun area and generate a single maximum of activator at the output oscillator. The dark blue points are in the sun area and generate two maxima of activator. (D) The red and blue bars correspond to the red and blue points in (A). The majority of red points produce a single maximum of the activator on oscillator #3, whereas most of the records corresponding to blue points generate two maxima.

For a more objective evaluation of the accuracy of the optimized network, we considered a large test dataset D_T of 100,000 random, uniformly distributed points in the square $[0,1] \times [0,1]$. Figure 4.5 shows the comparison between the location of a point from D_T and the number of activator maxima observed on the output oscillator within the time interval $[0, t_{max}]$. The red light points are located inside the sun disk and produced a single maximum (48028 cases). The light blue points are located in the surrounding area, and they forced two maxima of the output oscillator (47117 cases). Using the majority rule introduced for the training dataset D_S we can say that the total number of correctly located points was 95145 thus the classifier accuracy is $\sim 95\%$. The dark colours mark points that are incorrectly attributed. The dark red points were located outside the sun, but they force a single activator maximum (2967 cases). The dark blue points produced two activator maxima, but they were located in the sun area (1888 cases). The red points in Figure 4.5 (both light and dark) can be

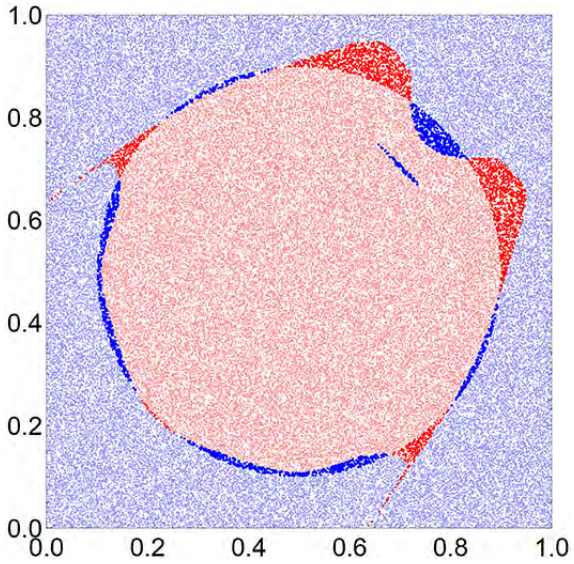


Figure 4.5: The figure shows how the optimized network sees the Japan flag. Each dot corresponds to a record from the testing dataset that contains 100000 records. The red points produced a single activator maximum thus, they are considered as the sun. For the records represented by the blue points, two maxima were observed, so they are classified as the surrounding white area. The light red and light blue points are those that are classified correctly. The dark red points belong to the white region in Figure 4.1, but the network thinks they belong to the sun. The dark blue points are the points belonging to the sun, but the network incorrectly classifies them as the points of the white surrounding area.

regarded as the image of the sun area seen by the optimized network. Instead of the disk, the network senses a complex, two horned shapes. In order to describe it more precisely we introduce new coordinates: $p = x - y$ and $q = x + y$. In these coordinates, the points producing a single activator maximum on the output oscillator are located,

as shown in Figure 4.6(A). The upper fitting curve (blue) is described as:

$$F_U(p) = 1.49928 + 5.56776p^2 - 81.5484p^4 + 503.745p^6 - 2275.64p^8 + 7512.65p^{10} - 14690.6p^{12} + 11764.8p^{14} \tag{4.1.5}$$

and the lower fitting curve (green) is:

$$F_D(p) = 0.432601 + 1.65454p^2 - 2.0677p^4 - 107.24p^6 + 1120.59p^8 - 3887.3p^{10} + 5377.69p^{12} - 2552.69p^{14} \tag{4.1.6}$$

Now we can forget about the original problem and consider the question if a chemical computer can correctly distinguish the points of the unit square that are located between the curves $F_D(x-y)$ and $F_U(x-y)$. Such a problem looks rather difficult, but the answer is simple: a good candidate for such a chemical computer is the optimized network described above. Its accuracy is around 99 %. The incorrectly attributed points are marked in Figure 4.6(B). Most of them are located at the boundary, and the error may be connected with oversimplified fitting. There are also some points for which the attribution error is hard to explain (the blue horizontal line at $q \sim 1.4$ on Figure 4.6(B)).

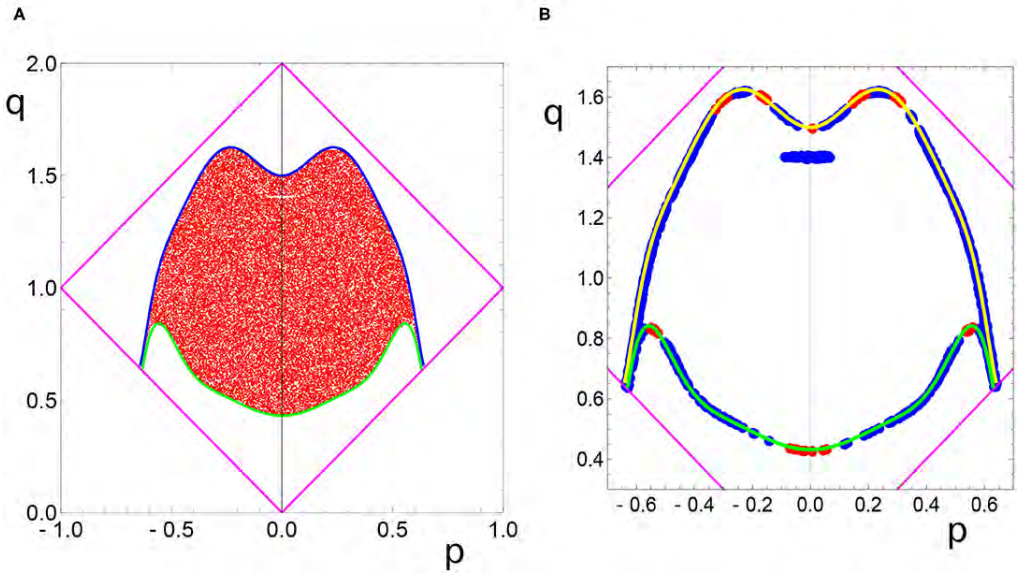


Figure 4.6: (A) The shape of the sun area, as seen by the optimized network in the p, q coordinates. The blue and green lines are the fits of the sun boundaries (cf. Eq. (4.1.5, 4.1.6). (B) The positions of incorrectly classified points that are located in the region between the fitting lines. The blue points are classified as belonging to the region outside, whereas they should be in. The red points are classified as belonging to the red region, but they fail outside the fitting lines. In (B) for better visibility, the upper fitting line is represented by a yellow curve.

4.2 Classifying two intertwined spirals apart by a network of interacting chemical oscillators

Throughout the past years "Telling two intertwined spirals apart" [89] [90] [91] has been a challenging problem. It has always been seen as a benchmark to test the performance of an AI system. It is specifically a difficult task because no single line drawn can differentiate between these two spirals. Moreover this problem resembles an XOR gate problem which has always been considered challenging for the AI to solve.

Our main aim is to teach the network of interacting chemical oscillators to tell if a point (x,y) belongs to spiral 1 or spiral 2. We considered two spirals (Figure 4.7) with coordinates belonging to $[0,1] \times [0,1]$. The coordinate values x, y are used as network inputs. The training dataset includes 449 points depicting spiral 1 and spiral 2 consists of 337 points. Thus the training dataset $M = \{(x_n, y_n, g_n), n = 1, N\}$, where N= 786 , $x, y \in [0,1]$ and $g_n = 0$ if point (x,y) represents spiral 1 or $g_n = 1$ if (x,y) lies in spiral 2. The total entropy of the training dataset (449+337) is around 0.985 bits.

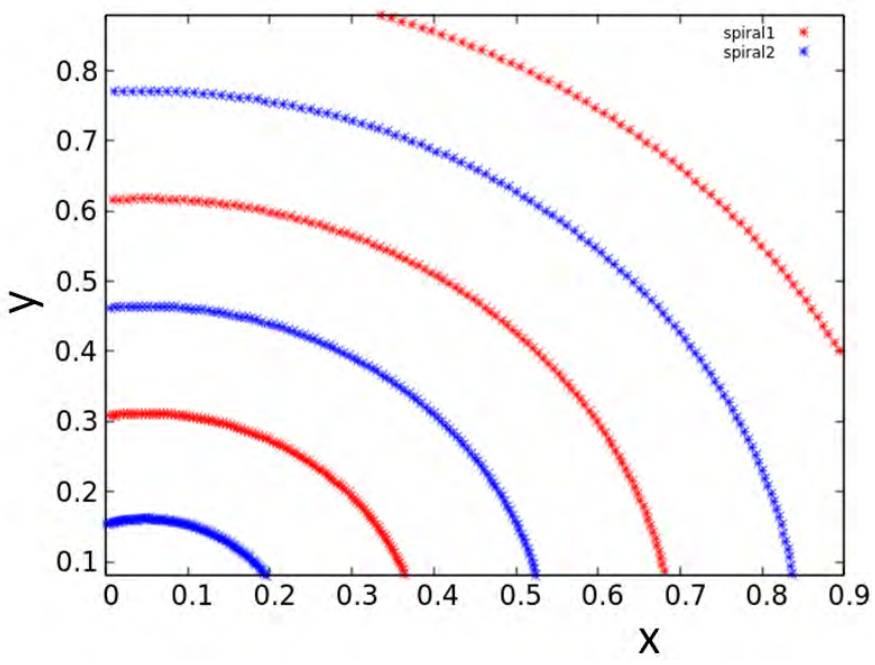


Figure 4.7: Two intertwined spirals with points lying in $[0,1] \times [0,1]$ considered as a training dataset in the classification problem.

Networks used

We have considered spiral classification with three different types of networks. The network size was always fixed: each network included 6 chemical oscillators. However the interactions between the oscillators were different. Figure 4.8 shows the three networks we have considered for this particular task. The degree of complexity of interactions was the smallest in the first network (Figure 4.8(a)), medium in the second network (Figure 4.8(b)) and was maximum in the third one (Figure 4.8(c)). The inputs (In_1, In_2) considered were the x, y coordinates, representing a point in a particular spiral. They were introduced to the system in the form of illumination time (same way as in Eq. 4.1.3 and 4.1.4). All oscillators were simulated by a two-variable Oregonator model(Eq. 3.3.7 and 3.3.8) as explained in Chapter 3. The parameter model values used for all oscillators in the networks were : $\varepsilon = 0.2$, $q = 0.0002$ and $f = 1.1$.

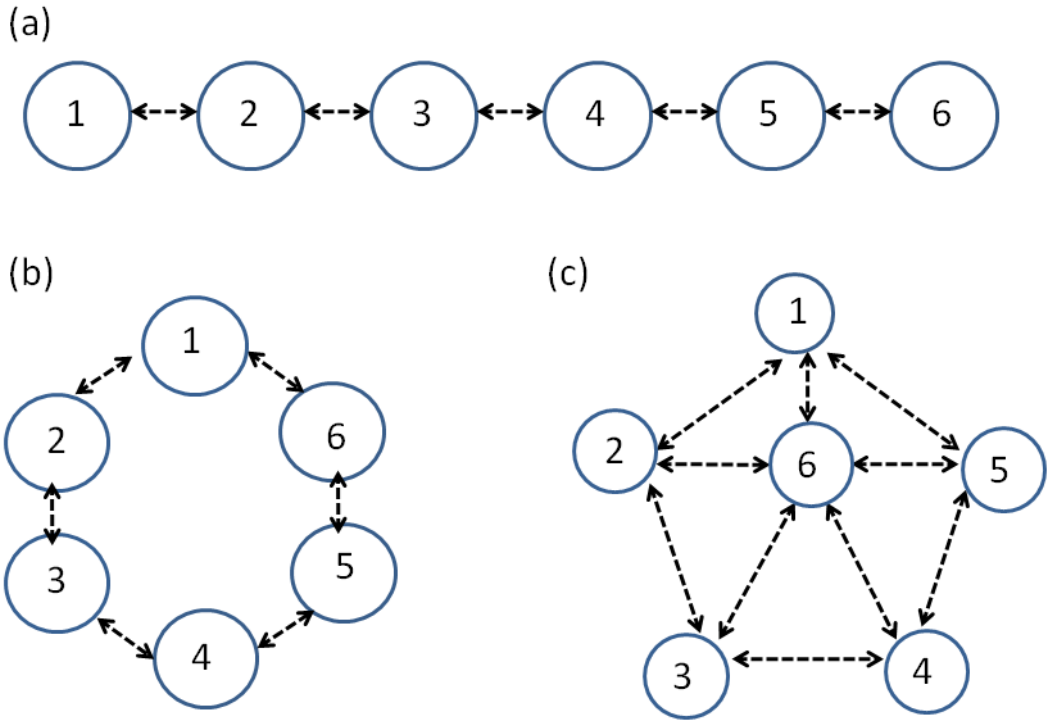


Figure 4.8: Three different oscillator networks considered for classification of intertwined spirals. All networks include six oscillators but have different structure of interactions. The oscillators are all numbered. The dotted lines between the oscillators represent interactions among them. (a) System 1: The simplest linear geometry among the oscillators. (b) System 2, the oscillator ring: The complexity of interactions is greater than that of System 1. (c) System 3, the pentagon geometry of oscillators: the complexity of interaction among the oscillators is the highest in the lot.

Results

We used evolutionary algorithms to find the optimized parameters of the systems. The details of the parameters optimized (for systems 1, 2 and 3) and the rules of recombination and mutations followed were discussed in Chapter 3, section 3.4. The population size considered for all three systems was 100. The evolution of all the structures was carried out for 500 generations. The optimization progress of system 1 is shown in Figure 4.9(a). The maximum mutual information (Chapter 3 Eqn. 3.4.1) obtained by this system at 500^{th} generation was 0.536 bits for oscillator #6. The values of optimized parameters are shown in Table 4.1.

Table 4.1: Optimized parameters of 6-oscillator linear classifier (Figure 4.8(a))

parameter	value
t_{max}	90.45
t_{start}	4.71
t_{end}	86.71
α	0.32
β	1.23
t^1_{illum}	4.71
t^3_{illum}	24.41
t^4_{illum}	58.91

The optimized structure is shown in Figure 4.9(b). There were three normal oscillators and three input oscillators. In this particular case, one of the input oscillators was also the output oscillator. In Figure 4.9(c) I plotted the oscillation distribution with respect to the cases obtained from the output oscillator. On this basis, I postulated the following classification rule: if the number of oscillations observed at the output oscillator are 0,1,4,6 or 9, then the input point inserted to the system belongs to spiral 1 else the point belongs to spiral 2. By following this rule, the classifier was able to distinguish two spirals apart with 85% of accuracy for points of the training dataset. 359 (out of 449) and 309 (out of 337) points of spiral 1 and spiral 2 were correctly identified.

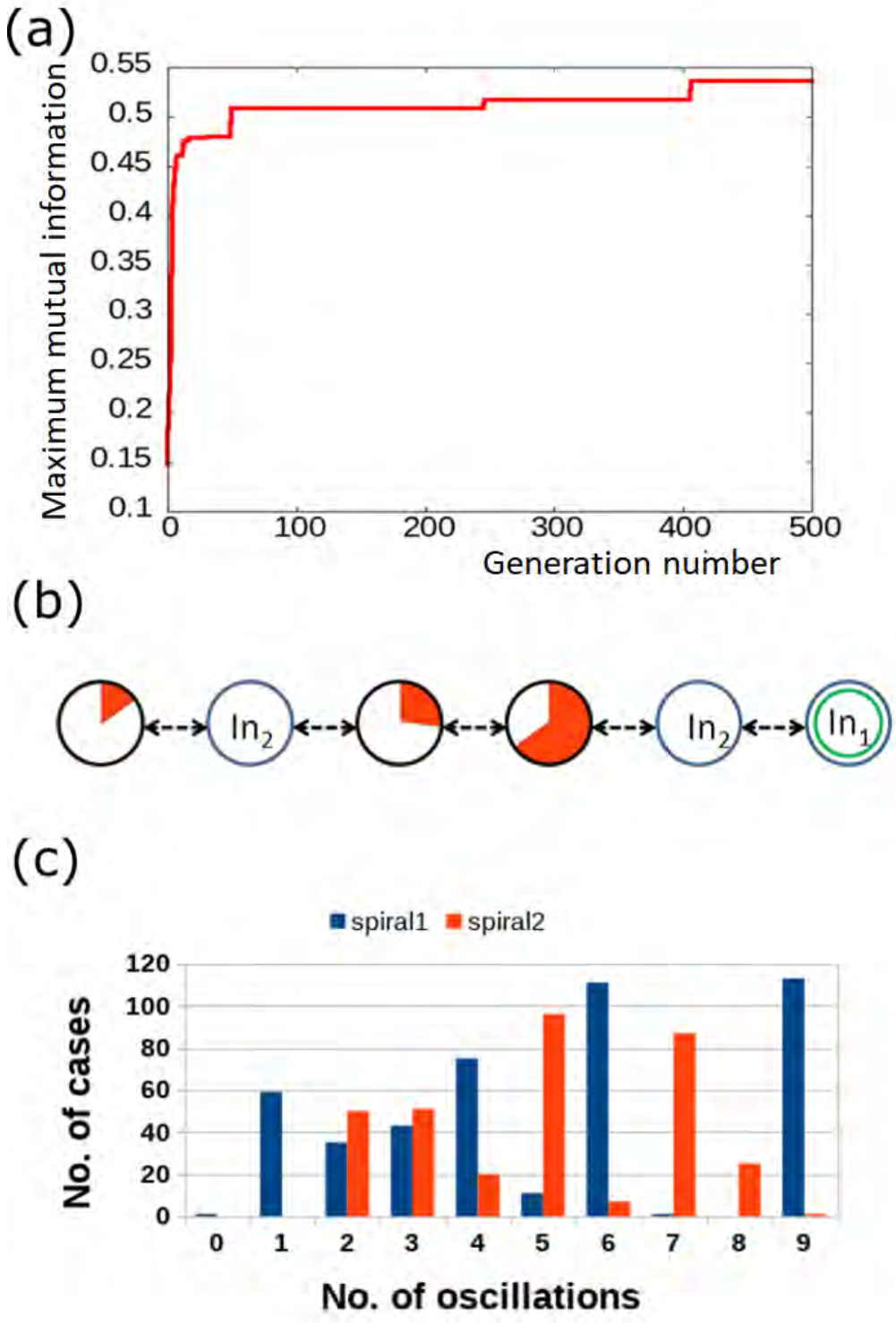


Figure 4.9: Results for System 1 : (a) The progress of optimization; the *Fitness* as a function of generation number. (b) The structure of the optimized network. The disks within a black circle are the normal oscillators. The ratio between the surface of the red shaded part and the disk surface represents the ratio between t^j_{illum} and t_{max}. In_1 and In_2 mark inputs for x and y coordinate respectively. The disk with the green circle inside is the output oscillator. (c) The distribution of the number of cases for which a given number of activator maxima was observed on oscillator #6. Colours indicate records representing spiral 1 and spiral 2.

To observe the dependency of accuracy over the complexity of interactions, I optimized two other systems (System 2 (Figure 4.8(b)) and System 3 (Figure 4.8(c)). The maximum mutual information obtained (by oscillator 5) of system 2 at 500^{th} generation was 0.54 bits. The optimized structure (Figure 4.10(b)) was almost similar to that obtained in the case of System 1. There were three normal oscillators and three input oscillators and like System 1, in this case also one of the input oscillators was the output one. The values of optimized parameters are shown in Table 4.2. After plotting the oscillation distribution (Figure 4.10(c)) obtained from output oscillator number 5, I derived a rule that if the number of oscillations is 1,3,4,6 or 8, then the input point belongs to the spiral 1. With that I got an accuracy of 84%. 404 (out of 449) and 257 (out of 337) points of spiral 1 and spiral 2 were correctly detected.

Table 4.2: Optimized parameters of 6-oscillator ring classifier (Figure 4.8(b))

parameter	value
t_{max}	90.4
t_{start}	4.71
t_{end}	86.71
α	0.606
β	1.104
t^1_{illum}	67.36
t^2_{illum}	7.36
t^3_{illum}	32.38

Table 4.3: Optimized parameters of 6-oscillator pentagon classifier (Figure 4.8 (c))

parameter	value
t_{max}	85.83
t_{start}	12.54
t_{end}	79.34
α	0.36
β	1.23
t^2_{illum}	1.38
t^3_{illum}	16.08
t^4_{illum}	94.69
t^5_{illum}	1.38

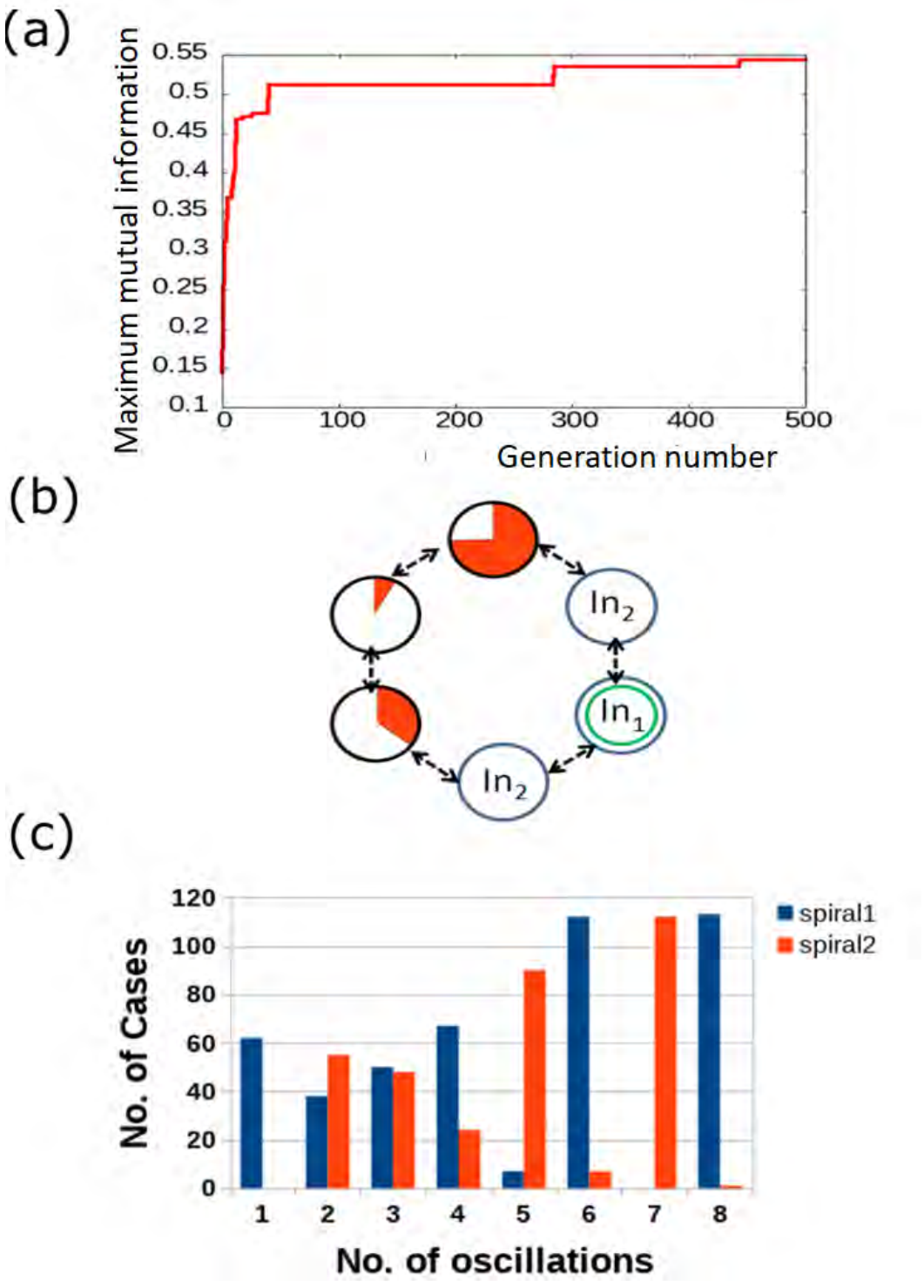


Figure 4.10: Results for System 2 : (a) The progress of optimization; the *Fitness* as a function of generation number. (b) The structure of the optimized network. The disks within a black circle are the normal oscillators. The ratio between the surface of the red shaded part and the disk surface represents the ratio between t^j_{illum} and t_{max}. In_1 and In_2 mark inputs for x and y coordinate respectively. The disk with the green circle inside is the output oscillator. (c) The distribution of the number of cases for which a given number of activator maxima was observed on oscillator #5. Colours indicate records representing spiral 1 and spiral 2.

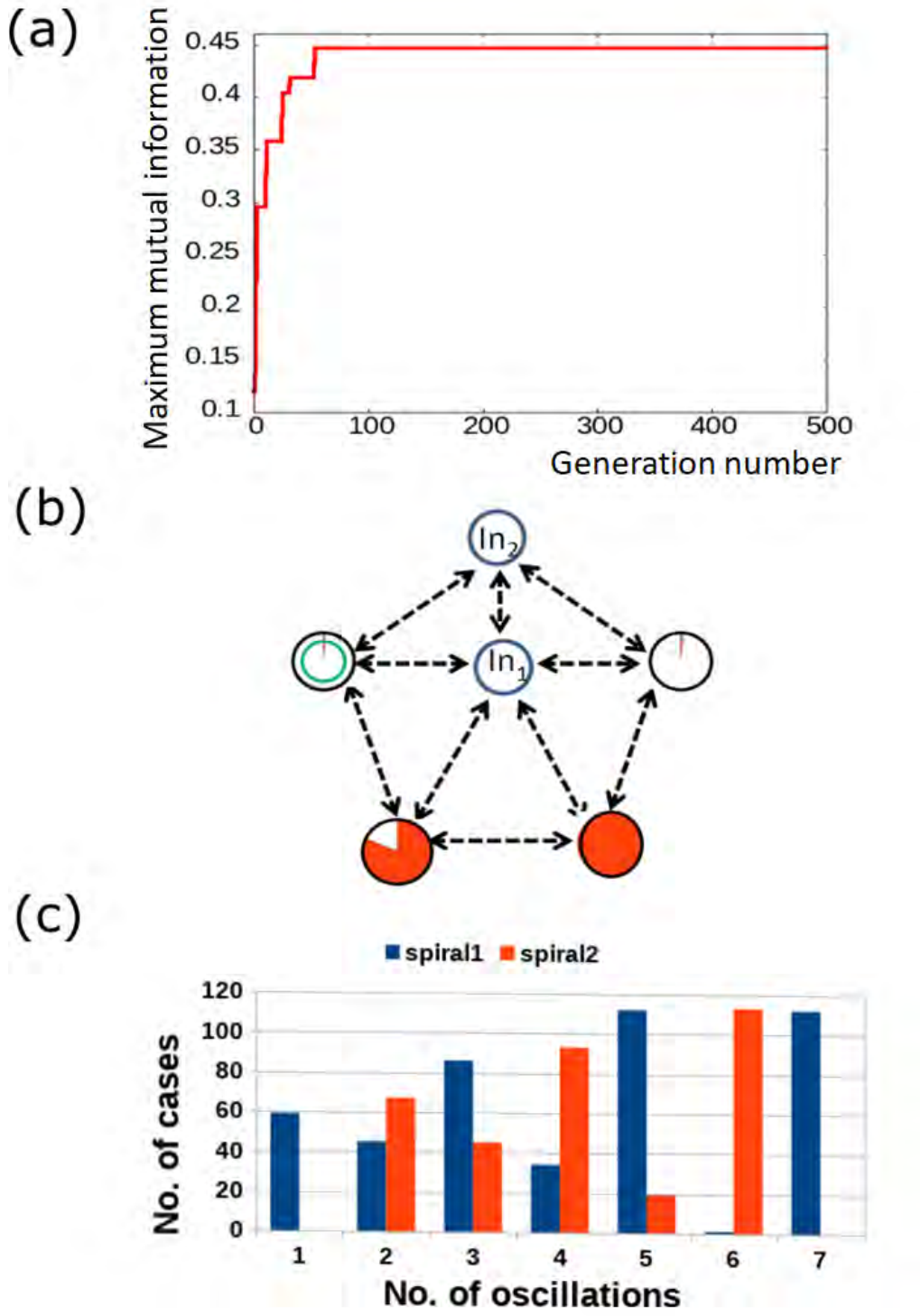


Figure 4.11: Results for System 3 : (a) The progress of optimization; the *Fitness* as a function of generation number. (b) The structure of the optimized network. The disks within a black circle are the normal oscillators. The ratio between the surface of the red shaded part and the disk surface represents the ratio between t^j_{illum} and t_{max}. In_1 and In_2 mark inputs for x and y coordinate respectively. The disk with the green circle inside is the output oscillator. (c) The distribution of the number of cases for which a given number of activator maxima was observed on oscillator #2. Colours indicate records representing spiral 1 and spiral 2.

The optimized parameters of System 3 are provided in Table 4.3. The optimization progress is shown in Figure 4.11(a). Unlike the other two systems, the optimized structure of this system was different. The structure consisted of four normal oscillators (Figure 4.11(b)) and only two input oscillators, each taking different inputs. The 4^{th} oscillator was illuminated throughout the time of simulation and it can be treated as an empty slot in the network. It was never oscillating. The normal oscillator number 2 was the output oscillator contrary to the other two cases where the input oscillators were the output one. The maximum mutual information observed was 0.44 bits. On plotting the oscillation distribution from the output oscillator (Figure 4.11(c)), I derived a rule that if the number of oscillations witnessed are 1, 3, 5 or 7, then the input point belong to the spiral 1, else it belongs to spiral 2. With this, we got an accuracy of 81.6%. 369 (out of 449) and 273 (out of 337) spiral 1 and spiral 2 points were correctly detected.

To check the qualities of the optimized networks, I prepared a test dataset F with 22401 points from which 12800 points (x,y) represent spiral 1 and the rest of the points belong to spiral 2. These new points were fed as inputs to all three optimized systems. The number of activator maxima was noted down from the output oscillators of each system. The same classification rules obtained with respect to the training dataset were used to distinguish among the two spirals. In this test dataset, I received an accuracy of 85.2% (system 1), 84.06% (system 2) and 81.8% (system 3) in determining the two spirals apart.

4.3 Conclusions

In this chapter, we solved two geometrically oriented problems successfully.

We demonstrated that a simple network of just 3 coupled chemical oscillators could predict the colour of a randomly selected point on the Japan flag with 95% accuracy. Another interesting result is that the network delivers a fast answer. The output information if a point belongs to the white or the red region on the flag appears just within two oscillation periods. It confirms that even small networks of interacting oscillators can perform complex computations. It is worth noticing that the points classified as belonging to the sun area group into an interesting, horned shape illustrated in Figure 4.6(A). The boundaries of this shape can be approximated by complex polynomials of a high order. The problem of finding if a point is located inside a horn-shaped area is computationally more complex than the determination of point location with respect to a disk located at the center of the unit square. A high accuracy algorithm ($\sim$ 99% accuracy) if a point belongs to the horned region is

given by the network of 3 oscillators we optimized to see the Japan flag. We doubt if any multilayered neural network can produce an equally simple algorithm of finding a point in the horn-shaped area.

The network of six interacting chemical oscillators was able to tell these two intertwined spirals apart (Figure 4.7) with a maximum of 85% accuracy. I have not only showed that the system could accurately distinguish between the two spirals but have also explored the possibility of the influence of different geometry of interactions among the chemical oscillators, on the accuracy of differentiating between spirals. It was seen that the accuracy was maximum in the case of System 1 and the smallest in the case of System 3. For this particular problem, the smallest interaction complexity improved the accuracy. However, irrespective of the interaction complexity, I was able to train every system in a way that the accuracy was always greater than 80%. To check the optimized network quality, I introduced a test dataset with 22401 points. All three of the classifiers were able to distinguish if the points belong to spiral 1 or spiral 2 with greater than 80% of accuracy.

Chapter 5

Networks of chemical oscillators as classifiers for medically oriented problems

The chapter shows how information processing with oscillator networks can help to solve medically oriented problems. I report results on diagnosing schizophrenia and on predicting the drug response in the treatment of patients with multiple myeloma.

5.1 Detection of schizophrenia by a network of chemical oscillators

Introduction

Schizophrenia is the most common form of psychotic behaviour. The statistics of people getting diagnosed every year is one in 4000, which means about 1.5 million people will be diagnosed approximately every year worldwide. The prevalence of schizophrenia compared to other well-known diseases is shown in Figure 5.1. There are many different kinds of symptoms a schizophrenic patient can face. The common form of those symptoms observed in schizophrenia is listed in Table 5.1. Its characteristic syndrome is : patients suffering from reality distortion often get delusional and also deal with hallucinations, which makes it difficult for them to distinguish between real things and their fantasies. Patients dealing with the effect of disorganisation, disrupt the connection between the thought process and the way to express it. The actions to express their thoughts get disrupted and it results in irrelevant speech and disorganized behaviour. Patients suffering from the symptoms of psychomotor poverty : exhibit a decrease in the amount of thought process, voluntary motor activity and emotions. In affective syndrome: depression is something common the

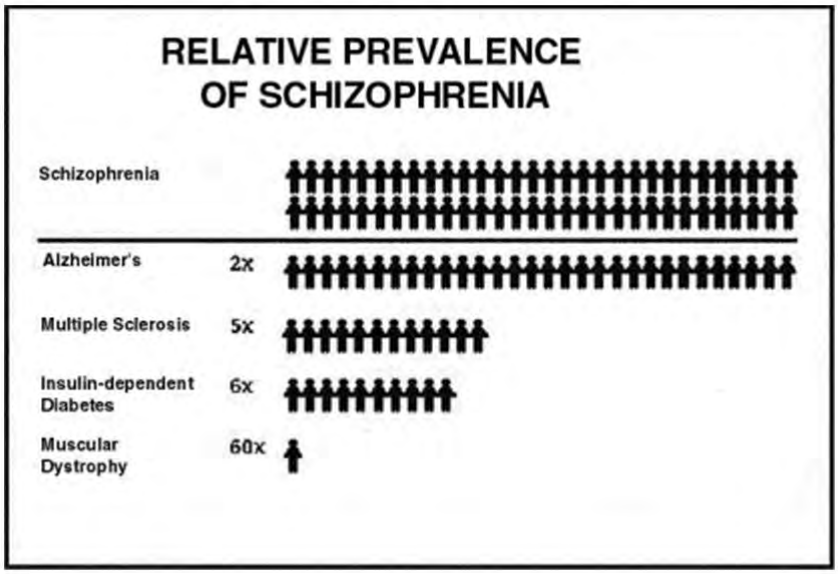


Figure 5.1: A schematic representation of relative prevalence of schizophrenia in the United States with respect to other disease [92] [93]

Table 5.1: Common form of symptoms observed in schizophrenic patients [94] [95].

Characteristic syndromes	Affective syndromes
Reality distortion	*Depression*
-Delusions	-Low mood
-Hallucinations	-Low self-esteem
Disorganization	-Hopelessness
-Formal thought disorder	-Suicidality
-Inappropriate affect	*Psychomotor Excitation*
-Disorganized or bizarre behaviour	-Irritability
Psychomotor Poverty	- Pressure of speech
-Flat affect	-Motor agitation
-Poverty of speech	Non-specific symptoms
-Decreased voluntary motor activity	- Anxiety

patient can go through, or some motor agitation (repetitive involuntary movements) can also be witnessed [94] [95]. The symptoms can vary from person to person and hence schizophrenia is difficult to detect and easily go undetected. It is believed that information about schizophrenia can be extracted from the EEG signals recording brain activity [96]. Such signals are recorded from electrodes placed in different parts of the scalp. Moreover, it was seen that the signals from the frontal lobes could be used to distinguish schizophrenic patients from the healthy ones. I present a method to diagnose the schizophrenic patients using a network of interacting chemical oscillators with averaged electric potentials from the frontal lobe of the brain as the input. The works related to this topic are published in [87] [97].

Preparation of input dataset

Signals (the time dependent potential values) were recorded from electrodes located in different parts of the scalp (see Figure 5.2). For the network optimization purpose, we used signals received from F7 and F8 channels marked blue and red in Figure 5.2 respectively . The dataset available on the web [98] containing signals recorded on $N = 84$ patients, out of which $N_h = 39$ were healthy and the other had symptoms of schizophrenia ($N_s = 45$). The EEG signals were recorded with the sampling rate 128 Hz for 1 minute. Therefore, for each patient, we have 16 data files corresponding to different electrodes and each data file contains $W = 7680$ values of recorded potential (in μV). The time between consecutive potential values is $\Delta t = 7.8125ms$. In order to reduce the amount of input data we characterized each data file by a single number. Let $V^l(n, w)$ denote the potential recorded for n^{th} patient, on the l^{th} electrode and at the time $t_w = w \cdot \Delta t$. For each patient we introduced

16 time averaged potentials defined as:

$$x_n^l = \sum_{w=0}^{W} V^l(n, w) \tag{5.1.1}$$

To proceed with the analysis we normalized the time averaged potentials over the set of patients. To do it we introduced:

$$\mu_l = \frac{1}{N} \sum_{n=1}^{N} x_n^l$$

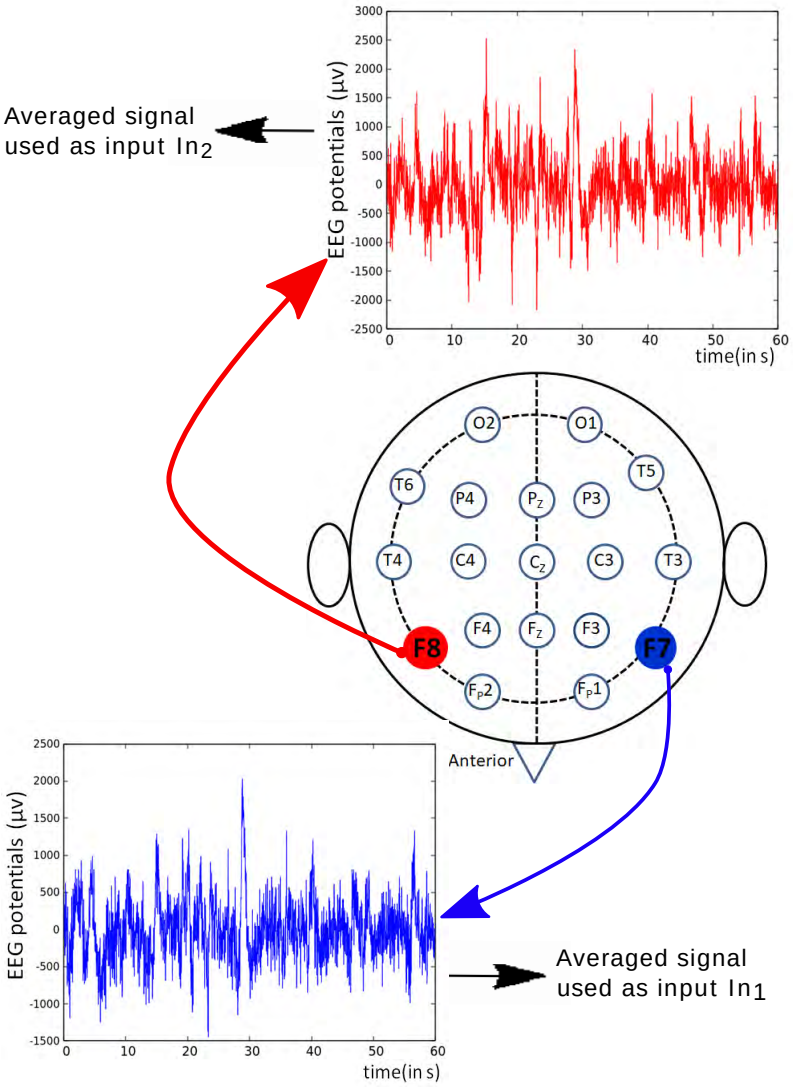


Figure 5.2: Schematic representation of positions of different electrodes used for recording potential values. The potentials derived from the red and blue marked channels are the one which were used as inputs for the classifier described in this section.

and

$$\sigma_l = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_n^l - \mu_l)^2}.$$

Previous studies have shown that the signals obtained from the frontal lobe of the brain (F7, F8) clearly show the difference in the brain activity of a schizophrenic patient from that of a healthy subject [96]. Therefore, our problem of schizophrenia diagnosis is reduced to the classification of the dataset: $D_S = \{(p_{1,n}, p_{2,n}, g_n), n = 1, N\}$ where the values of predictors $p_{1,n}$ and $p_{2,n}$ for the patient n were defined as:

$$p_{1,n} = \frac{x_n^{F7} - \mu^{F7}}{\sigma^{F7}}, p_{2,n} = \frac{x_n^{F8} - \mu^{F8}}{\sigma^{F8}} \tag{5.1.2}$$

and the record type $g_n = 0$ for a schizophrenic patient and $g_n = 1$ for a healthy subject. On the basis of available medical data we calculated the values $\mu_{F7} = 7.724\mu V$, $\mu_{F8} = 2.46\mu V$, $\sigma_{F7} = 20.3\mu V$ and $\sigma_{F8} = 15.10\mu V$.

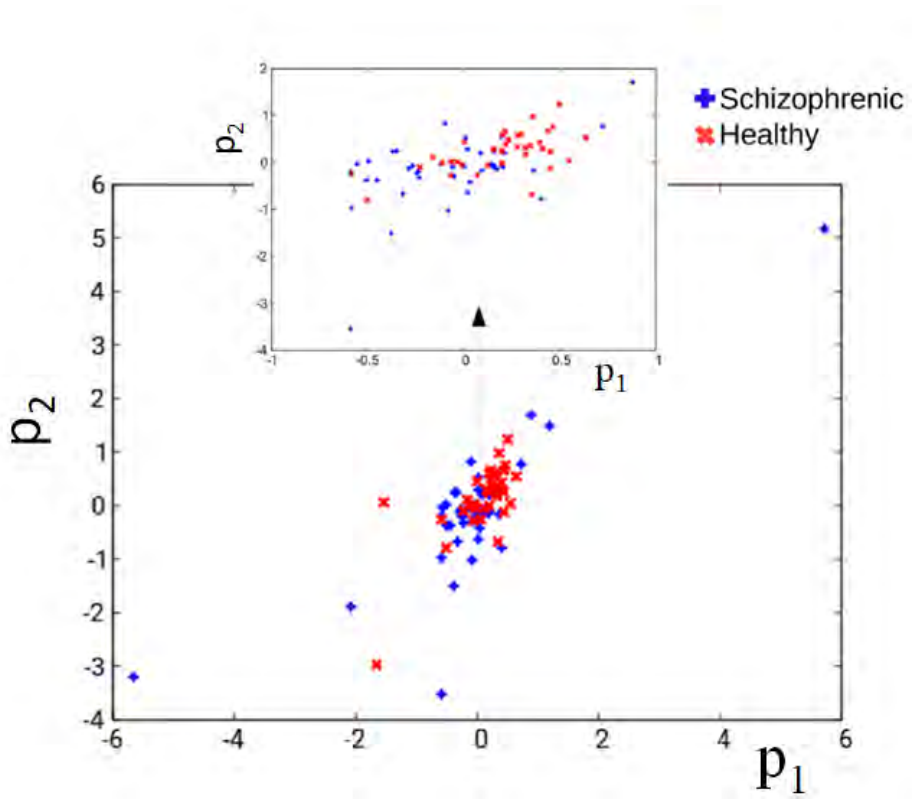


Figure 5.3: The distribution of records in the D_S database in the (p_1, p_2) coordinates. Blue and red crosses correspond to schizophrenic and healthy cases, respectively.

Network geometry

At the initial stage of our study, we proposed that a network formed by six interacting oscillators (Figure 5.4) can distinguish between schizophrenic and healthy patients. The arrows (in Figure 5.4) illustrate interactions between the oscillators. The time evolution of reactions proceeding in each oscillator was described by two-variable Oregonator model (Eq. 3.3.7 and 3.3.8). In our simulations we used the following values of model parameters for all oscillators ($1 \leq j \leq 6$): $\varepsilon = 0.2, q = 0.0002$ and $f = 1.1$. The parameters of the Oregonator model were fixed and did not undergo optimization. The method of providing input to the system was same as mentioned in Chapter 3 (Eq. 3.2.4). Here In_1 and In_2 corroborates to the predictor value of $p_{1,n}$ and $p_{2,n}$ of the considered dataset.

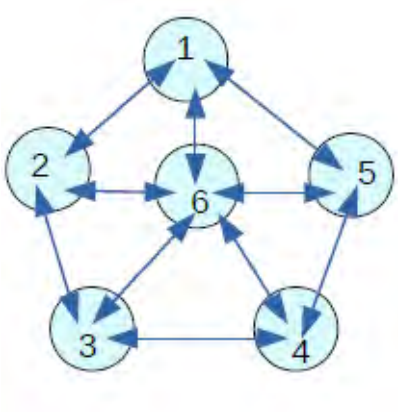


Figure 5.4: Geometry of 6-oscillator network considered for the schizophrenia diagnosis. The disks represent individual oscillators that can act as input oscillators or normal ones. Arrows show interactions among the oscillators. The numbers are used to mark individual oscillators.

Results

The classifier made of 6 oscillators arranged in the geometry illustrated in Figure 5.4 was optimized for 1000 generations. The initial population consisted of 200 classifiers. The optimization process and rules of recombination and mutation were similar to those reported in Chapter 3. The *Fitness* observed after 1000 optimization steps was 0.422 bit. The evolution of mutual information values with respect to the generations is shown in Figure 5.5(a). The parameters of the best classifier are given in Table 5.2.

Table 5.2: Optimized parameters of 6-oscillator pentagon classifier (Figure 5.4)

parameter	value
t_{max}	79.5
t_{start}	72.1
t_{end}	4.9
α	0.25
β	0.77
t^2_{illum}	22.5
t^3_{illum}	0.57

Figure 5.5(b) illustrates the structure of the optimized classifier. There are 2 normal oscillators, 2 oscillators that act as inputs of the predictor $p_{1,n}$ and 2 oscillators representing inputs of $p_{2,n}$ for N patients. The input of $p_{2,n}$ was also the classifier output. Moreover, in the central oscillator was the input of $p_{1,n}$. Figure 5.5(d) shows the distribution of numbers of activator maxima observed on the oscillator #1. On its basis, we can deduct the following classification rule: a patient is healthy if the number of activator maxima is 1, 3, 4, or 5. The observation of any other number of activator maxima indicates that the patient is ill. Application of this rule gives 15 errors for 84 cases included in D_S (82 % accuracy). It is worth noticing that for the majority of schizophrenic cases, the six node optimized classifier did not produce a single activator maximum at the output oscillator. Figure 5.6 illustrates the positions of correctly and incorrectly classified cases for 6-oscillator network in the phase space (p_1, p_2).

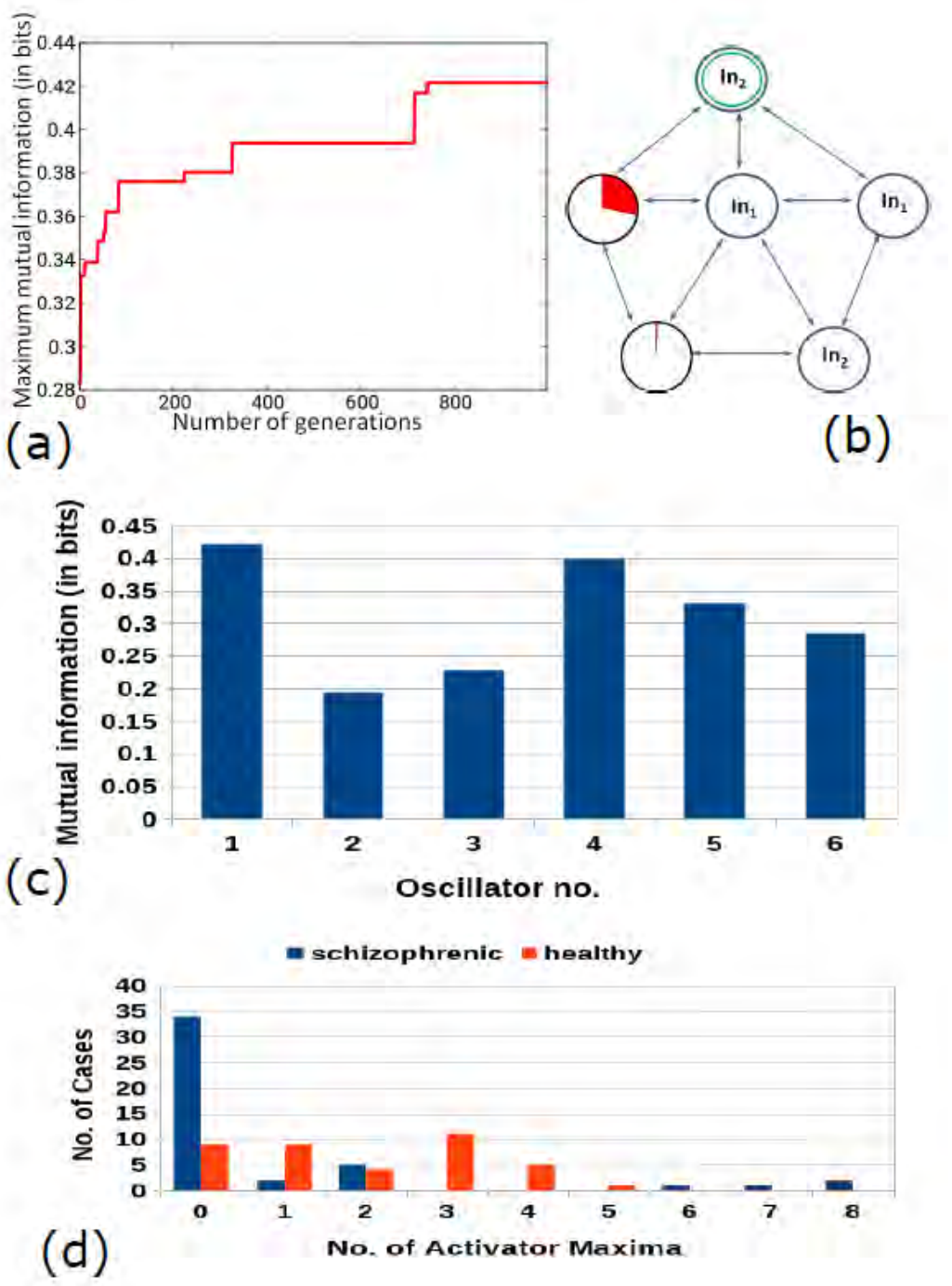


Figure 5.5: Results for 6-oscillator network (cf. Figure 5.4): (a) The progress of optimization; the *Fitness* as a function of generation number. (b) The structure of the optimized network. The circles with black rings are the normal oscillators. The red shaded parts of the piecharts represent the ratio between t^j_{illum} and t_{max}. The circles with blue rings mark input oscillators (In_1 represents the input from F7 channel, In_2 represents the input from F8 channel). The circles with a double ring shows the output oscillator. (c) The mutual information $I(Z_k; Z_O)$(chapter 3, Eqn 3.4.1) for k^{th} oscillator, $k \in \{1, 2, 3, 4, 5, 6\}$. The function $k \to I(Z_k; Z_O)$ has the maximum at $k = 1$. (d) The distribution of the numbers of cases for which a given number of activator maxima was observed on oscillator #1 for records representing schizophrenic and healthy patients.

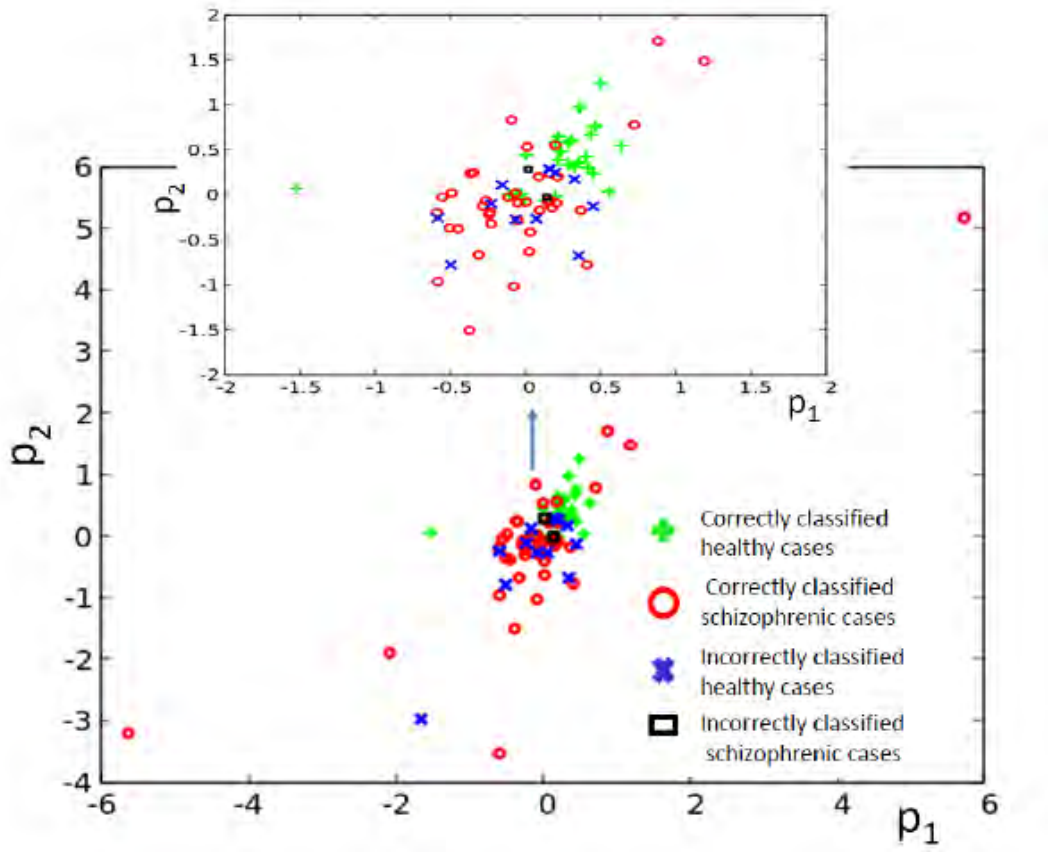


Figure 5.6: The distribution of correctly and incorrectly classified cases for 6-oscillator network (cf. Figure 5.4) in the phase space (p_1, p_2).

6-node classifier operating on partial signals

The classifier described above has shown 82% accuracy in schizophrenia diagnosis. But can it be improved? The obvious approach is to use a larger network that would allow considering a dataset with a larger number of predictors. However, such an approach can be strongly biased by a limited number of cases in the training dataset. We would like to present another approach. Each 60 second signal $V^l(n,w)$ for n^{th} patient, on the l^{th} electrode at time $t_w = w \cdot \Delta t$ is composed of $W = 7680$ values of recorded potential. We divide this signal into signals:

$$V1^l(n,w) = V^l(n,w) \;\; for \;\; w = 1, 2560$$

$$V2^l(n,w-2560) = V^l(n,w) \;\; for \;\; w = 2561, 5120$$

$$V3^l(n,w-5120) = V^l(n,w) \;\; for \;\; w = 5121, 7680$$

It means that instead of one 60 s long signal, we considered 3 signals $V1, V2$ and $V3$ recorded during the time intervals $[0, 20s]$, $[20s, 40s]$ and $[40s, 60s]$ respectively.

The signals $V1, V2$ and $V3$ were averaged and normalized using the same algorithm as applied to the signal V (Eq. 5.1.1, 5.1.2). As a result we obtained three datasets D_{S1}, D_{S2} and D_{S3}, where the averaged, normalized potentials observed in subintervals of time were matched with the parameter describing the patient health. Next, we performed the optimization procedure to obtain oscillator networks that classify the datasets D_{S1}, D_{S2}, D_{S3}. For each case we considered population of 100 classifiers and optimized it for 500 generations. The progress of optimization is illustrated in Figure 5.7. Subfigures (a), (b) and (c) correspond to datasets D_{S1}, D_{S2} and D_{S3} respectively. We can see that in all cases the fitness of the best classifier (0.33 bit, 0.323 bit and 0.371 bit, respectively) was lower than that achieved within 500 generations for the network classifying the dataset D_S.

The structures of the most fit classifiers are shown in Figure 5.8 (subfigures (a), (b) and (c) represent D_{S1}, D_{S2} and D_{S3} respectively). Like in Figure 5.5(b) the circles with black rings are the normal oscillators. The red shaded parts of the piecharts represent the ratio between t^j_{illum} and t_{max}. The circles with blue rings mark input oscillators (In_1 represents the input from F7 channel, In_2 represents the input from F8 channel). The circle with a double ring shows the output oscillator. The values of parameters describing the classifiers are:

1. For the classifier of D_{S1}:
   $t_{max} = 85.8$, $t_{start} = 12.5$, $t_{end} = 79.3$, $\alpha = 0.71$, $\beta = 0.52$, $t^3_{illum} = t_{max}$.

2. For the classifier of D_{S2}:
   $t_{max} = 89.6$, $t_{start} = 94.6$, $t_{end} = 10.1$, $\alpha = 0.38$, $\beta = 0.11$, $t^2_{illum} = 24.4$, $t^6_{illum} = 56.8$.

3. For the classifier of D_{S3}:
   $t_{max} = 77.9$, $t_{start} = 28.1$, $t_{end} = 78.6$, $\alpha = 0.76$, $\beta = 0.17$, $t^1_{illum} = 31.2$, $t^3_{illum} = t^5_{illum} = 9.37$, $t^6_{illum} = t_{max}$.

Therefore, these structures and parameters describing classifiers of signals coming from different subintervals are very different. This may suggest insufficient optimization (but still optimization with similar size of population and number of generations led to reasonable results for other problems [33]), or, more likely too small and too divergent training dataset. Despite the differences, it is interesting to note that in all cases the output oscillator is also the input of a predictor. It is $p_{1,n}$ for the classifier of D_{S2} and $p_{2,n}$ for classifiers of D_{S1} and D_{S3}. Not surprising the relationships between the number of activator maxima recorded on the output oscillator and the patient health are different for D_{S1}, D_{S2} and D_{S3}. They can be extracted from results illustrated in Figure 5.9. For the classifier of D_{S1} a patient is schizophrenic if the

number of activator maxima recorded on the output oscillator #4 $\in \{0, 3, 5, 6, 8\}$. For the classifier of D_{S2} a patient is schizophrenic if the number of oscillations of output oscillator #3 $\in \{0, 5, 9\}$ and for D_{S3} a patient is schizophrenic if the oscillations of oscillator #2 $\in \{0, 2, 6, 7, 9\}$. The accuracy of such classifiers were 77%, 75% and 79.7% for D_{S1}, D_{S2} and D_{S3} respectively. Now we can combine the expertise of all three classifiers together and postulate that a patient is ill if at least two of three classifiers predict schizophrenia. The accuracy of such classification increases to 90% and there were 7 healthy patients and only a single ill patient that was incorrectly diagnosed.

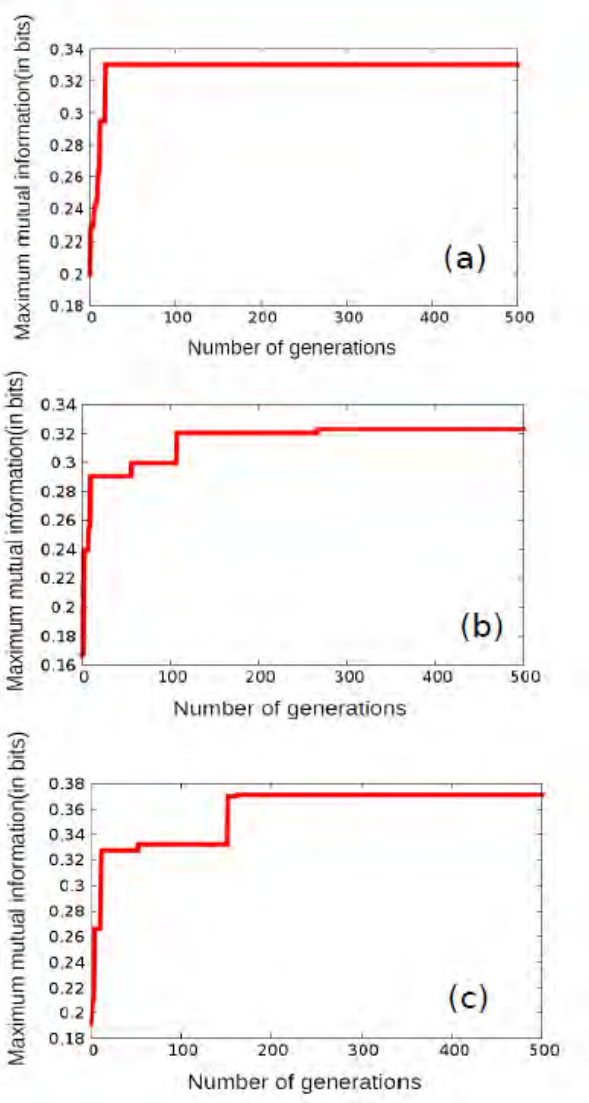


Figure 5.7: The progress of optimization of the schizophrenia classifiers processing the averaged 20 s long signals. Subfigures (a), (b) and (c) correspond to datasets D_{S1}, D_{S2} and D_{S3} respectively. The maximum values of fitness obtained after 500 generations were 0.33 bit, 0.323 bit and 0.371 bit.

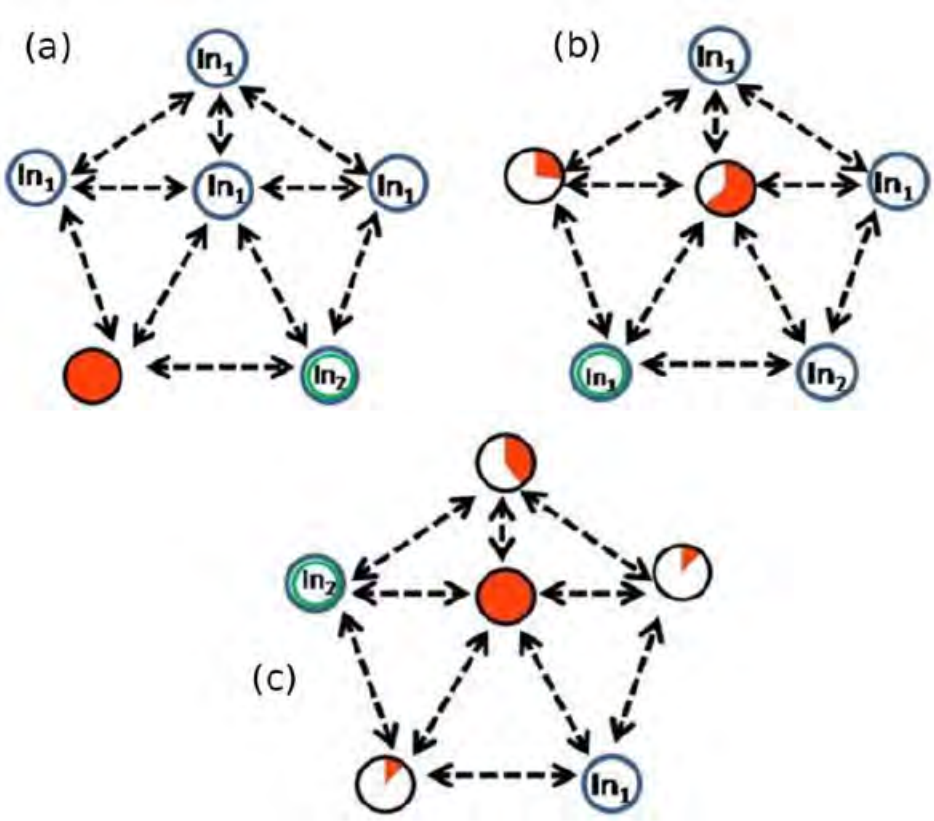


Figure 5.8: The structures of the schizophrenia classifiers processing the averaged 20 s long signals. Subfigures (a), (b) and (c) correspond to datasets D_{S1}, D_{S2} and D_{S3} respectively. The circles with black rings are the normal oscillators. The red shaded parts of the piecharts represent the ratio between t^j_{illum} and t_{max}. The circles with blue rings mark input oscillators (In_1 represents the input from F7 channel, In_2 represents the input from F8 channel). The circle with a double ring shows the output oscillator.

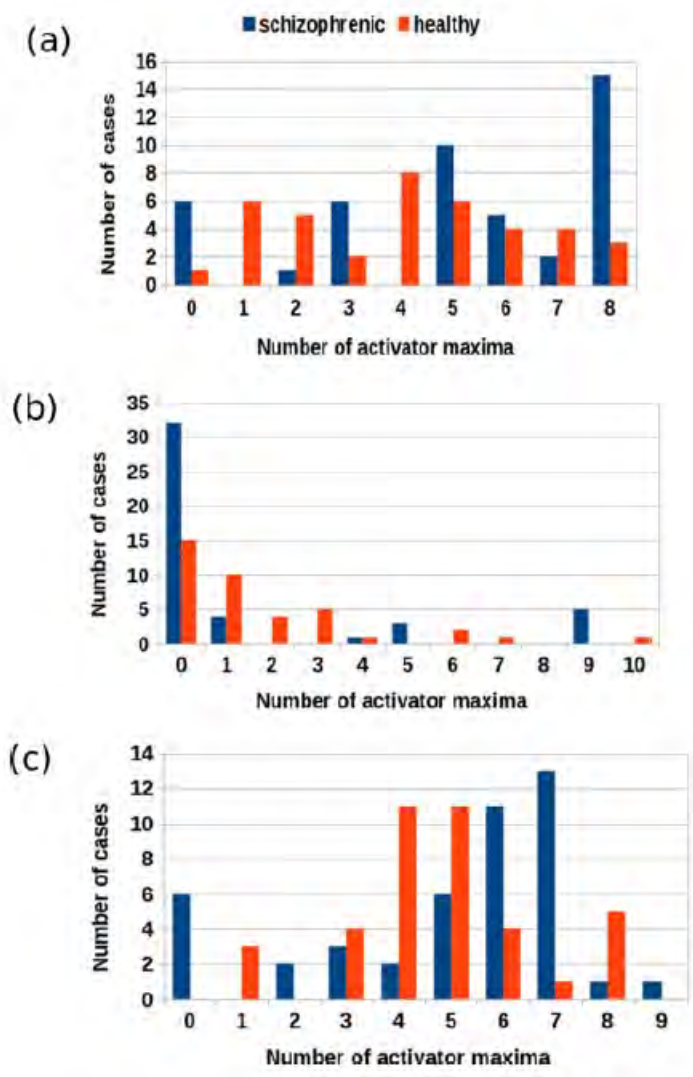


Figure 5.9: The relationship between the number of activator maxima recorded on the output oscillator and the patient health for of the schizophrenia classifiers processing the averaged 20 s long signals. The blue bars correspond to schizophrenic cases and the orange ones to the healthy ones. Subfigures (a), (b) and (c) represent D_{S1}, D_{S2} and D_{S3} respectively.

Accuracy measurement in diagnosing schizophrenia with different network size

Now, let us discuss the accuracy of optimized classifiers for networks with geometries illustrated in Figure 5.10. We decided to study the accuracy with respect to the number of oscillators used to process information. We studied the time evolution of the networks by numerical solution of Eq. (3.3.7 and 3.3.8) using Cash-Karp RK45 method [86] with 10^{-3} time steps. The number of activator maxima was calculated as the number of activator maxima, with a concentration(u) larger than 0.05, observed when the predictors $p_{1,n}$, $p_{2,n}$ (of dataset D_S) were used as the input. The optimization process was done with a population size of 200 and for 1000 generation steps. Figure

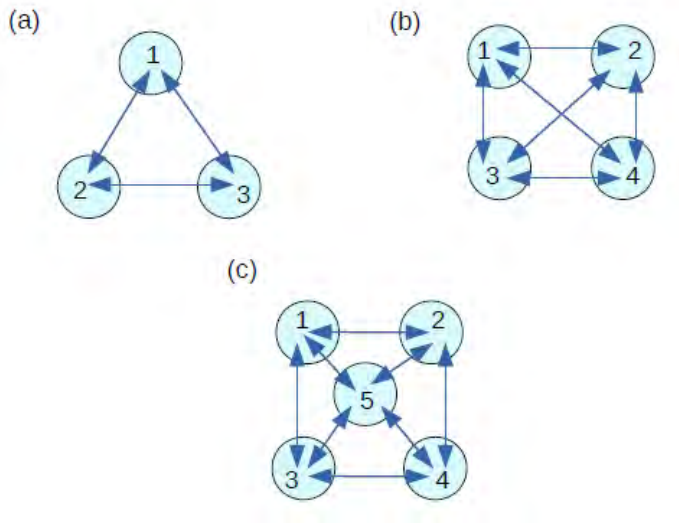


Figure 5.10: Geometries of oscillator networks considered for the schizophrenia diagnosis. The disks represent individual oscillators that can act as input oscillators or normal ones. Arrows show interactions among the oscillators. The numbers are used to mark individual oscillators in the following description of results. (a) 3-oscillator network, (b) 4-oscillator network, (c) 5-oscillator network.

5.11(a) shows the progress of optimization as the function of a number of generations for the network composed of 3 oscillators. The increase in *Fitness* is fast for the first few generations. Next, it changes into randomly distributed jumps with decreasing amplitude and frequency. Such dependence of the *Fitness* is typical for evolutionary optimization of classifiers [99]. The *Fitness* observed after 1000 optimization steps were 0.417 bit. The parameters describing the best classifier are given in Table 5.3.

Figure 5.11(b) illustrates the structure of the optimized classifier. It is interesting to notice that the normal oscillator remained inactive for the majority of the time when the network evolution was observed. This feature is reflected by the values of $I(Z_k; Z_O)$ (Eq. 3.4.1 chapter 3, section 3.4 - Teaching strategy) for k^{th} oscillator, is shown in Figure 5.11(c). The value of $I(Z_k; Z_O)$ is very small, which means that the activity of the oscillator #1 gives little information about the patient health. The value of $I(Z_3; Z_O) = 0.417$ bit is the maximum one; thus, the oscillator #3 was selected as the output one. Figure 5.11(d) shows the distribution of numbers of activator maxima observed on the oscillator #3 for schizophrenic and healthy patients. This result suggests the following classification rule: a patient is healthy if the number of activator maxima is 6 or 7. The observation of any other number of maxima diagnoses schizophrenia. The application of this rule gives 15 errors for 84 cases included in D_S (82 % accuracy). Only 3 schizophrenic patients (of 45) are diagnosed as the healthy ones. It gives over 93 % accuracy in detecting the illness. On the other hand, 12 healthy people (of 39) are diagnosed as schizophrenic ones (30 % error). If these results are confirmed using a large dataset of cases, then the 3-oscillator classifier can detect healthy patients with high accuracy. On the other hand, the verdict that a patient is ill requires further investigation. Positions of correctly and incorrectly classified cases for 3-oscillator network in the phase space (p_1, p_2) are shown in Figure 5.12.

Table 5.3: Parameters of the optimized 3-oscillator network.

parameter	value
t_{max}	99.8
t_{start}	42.5
t_{end}	6.3
α	0.89
β	0.36
t^1_{illum}	92.5

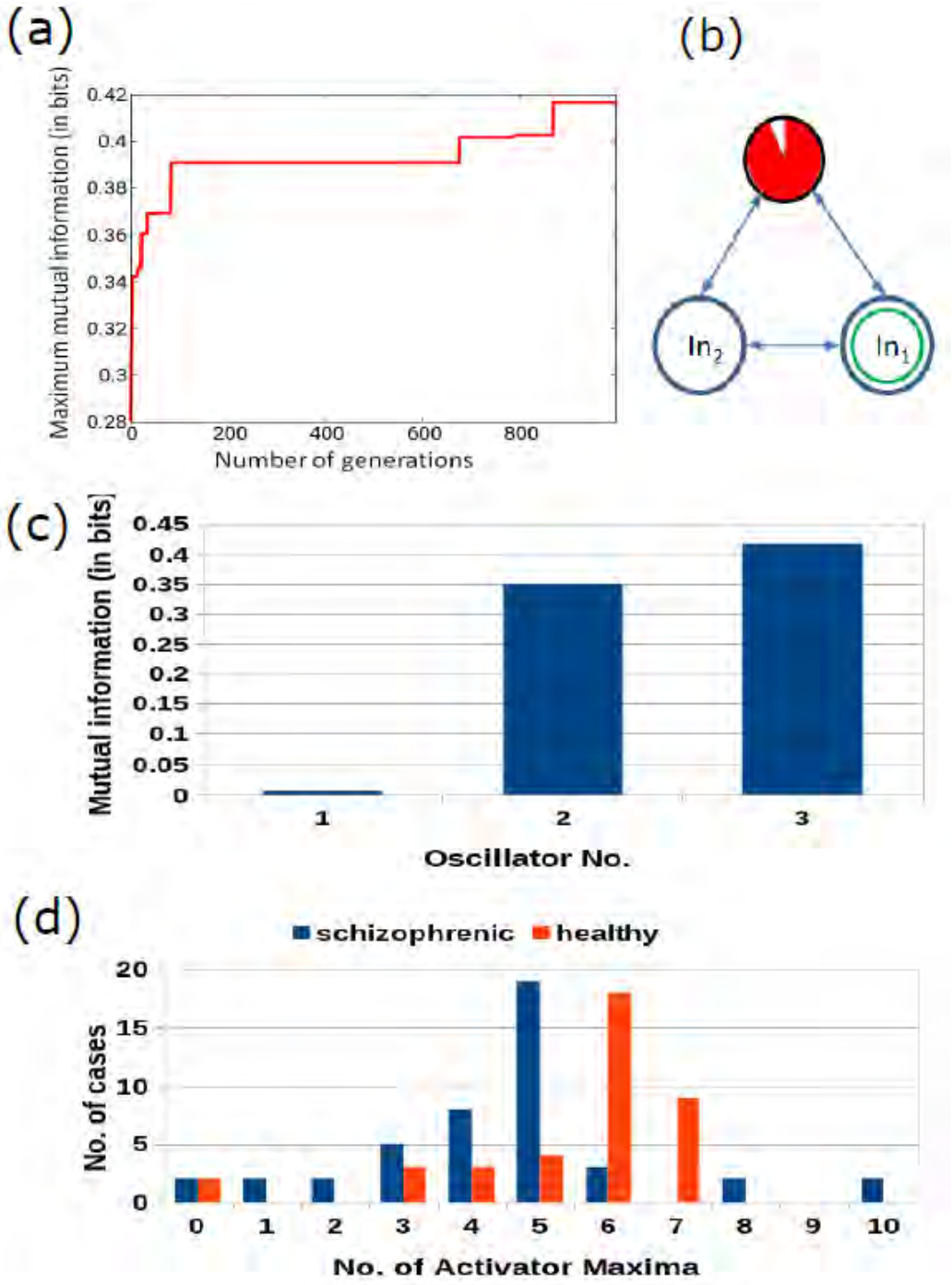


Figure 5.11: Results for 3-oscillator network (cf. Figure 5.10a): (a) The progress of optimization; the *Fitness* as a function of generation number. (b) The structure of the optimized network. The disk within a black circle is the normal oscillator. The ratio between the surface of the red shaded part and the disk surface represents the ratio between t^j_{illum} and t_{max}. In_1 and In_2 mark inputs for $p_{1,n}$ and $p_{2,n}$. The disk with the green circle inside is the output oscillator. (c) The mutual information $I(Z_k; Z_O)$ for $k \in \{1, 2, 3\}$. The mutual information has the maximum at the oscillator #3. (d) The distribution of the numbers of cases for which a given number of activator maxima was observed on oscillator #3. Colours indicate records representing schizophrenic and healthy patients.

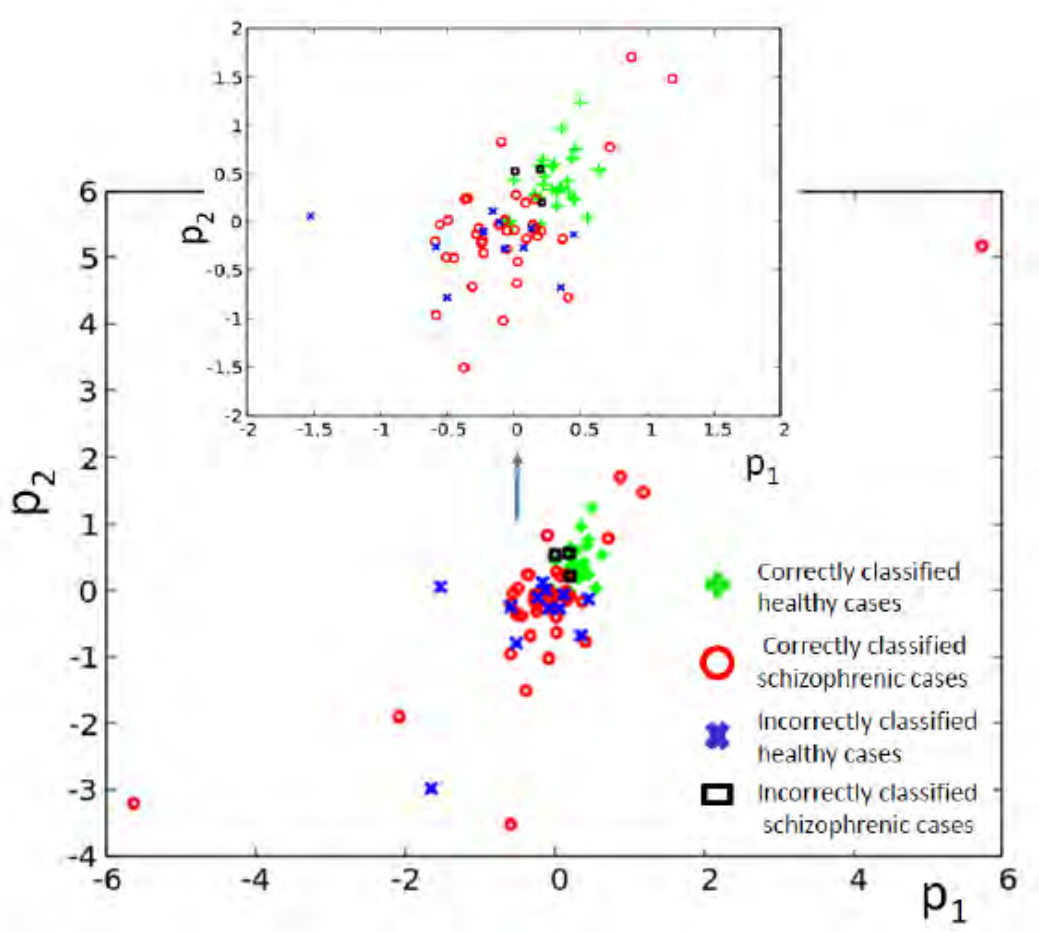


Figure 5.12: The distribution of correctly and incorrectly classified cases for 3-oscillator network (cf. Figure 5.10(a)) in the phase space (p_1, p_2).

Similar results for optimization of the 4-oscillator classifier are illustrated in Figure 5.13(a). The *Fitness* observed after 1000 optimization steps was 0.409 bit. The parameters describing the best classifier are given in Table 5.4.

Table 5.4: Parameters of the optimized 4-oscillator network.

parameter	value
t_{max}	94.1
t_{start}	57.7
t_{end}	11.5
α	0.61
β	0.36
t^1_{illum}	59.6

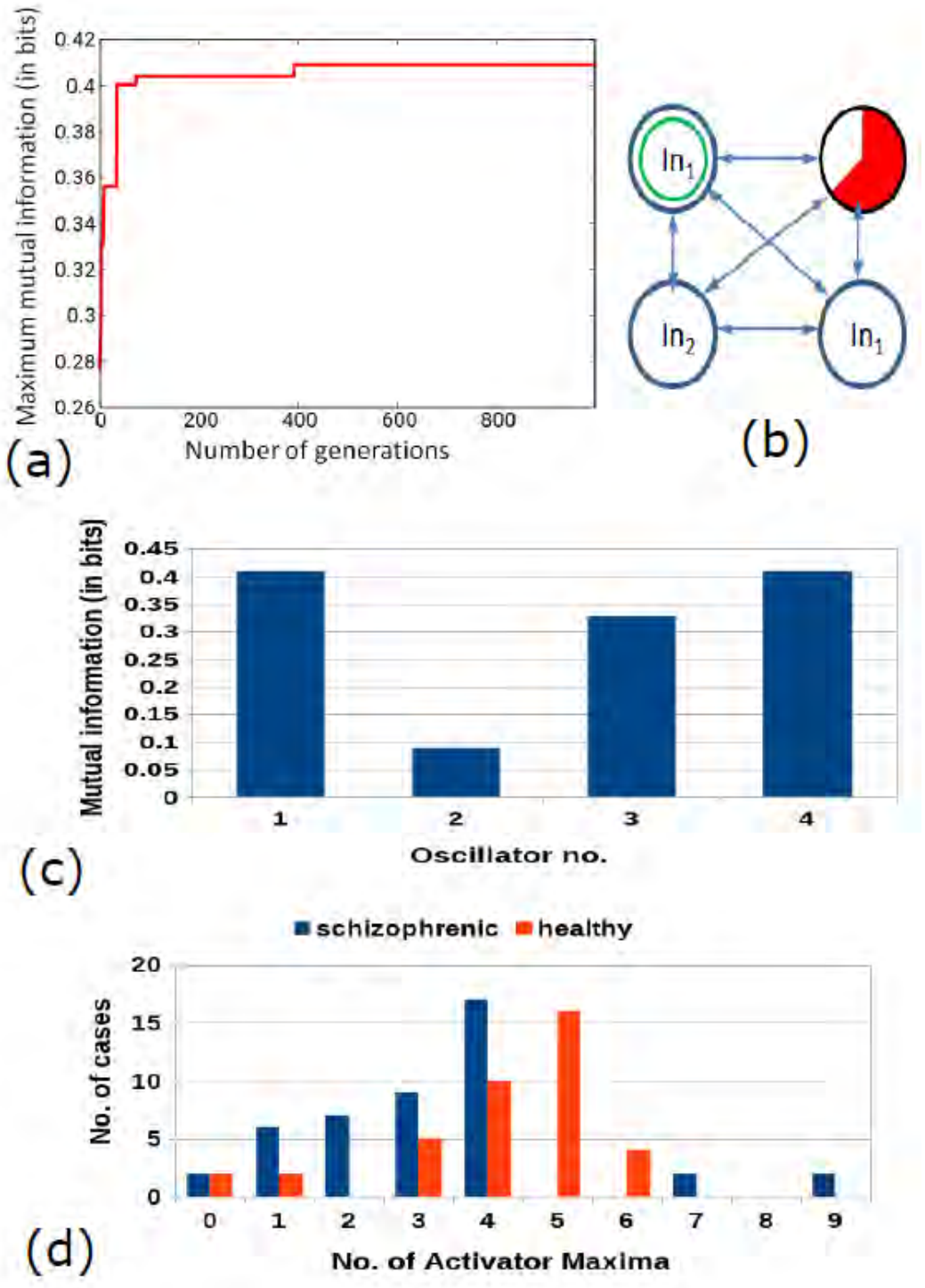


Figure 5.13: Results for 4-oscillator network (cf. Figure 5.10b): (a) The progress of optimization; the *Fitness* as a function of generation number. (b) The structure of the optimized network. Notation as in Figure 5.11(b). (c) The mutual information $I(Z_k; Z_O)$ for $k \in \{1, 2, 3, 4\}$. The function $k \to I(Z_k; Z_O)$ has the maximum at $k = 1$ and $k = 4$. $k = 1$ is chosen as an output oscillator. (d) The distribution of the numbers of cases for which a given number of activator maxima was observed on oscillator #1 for records representing schizophrenic and healthy patients.

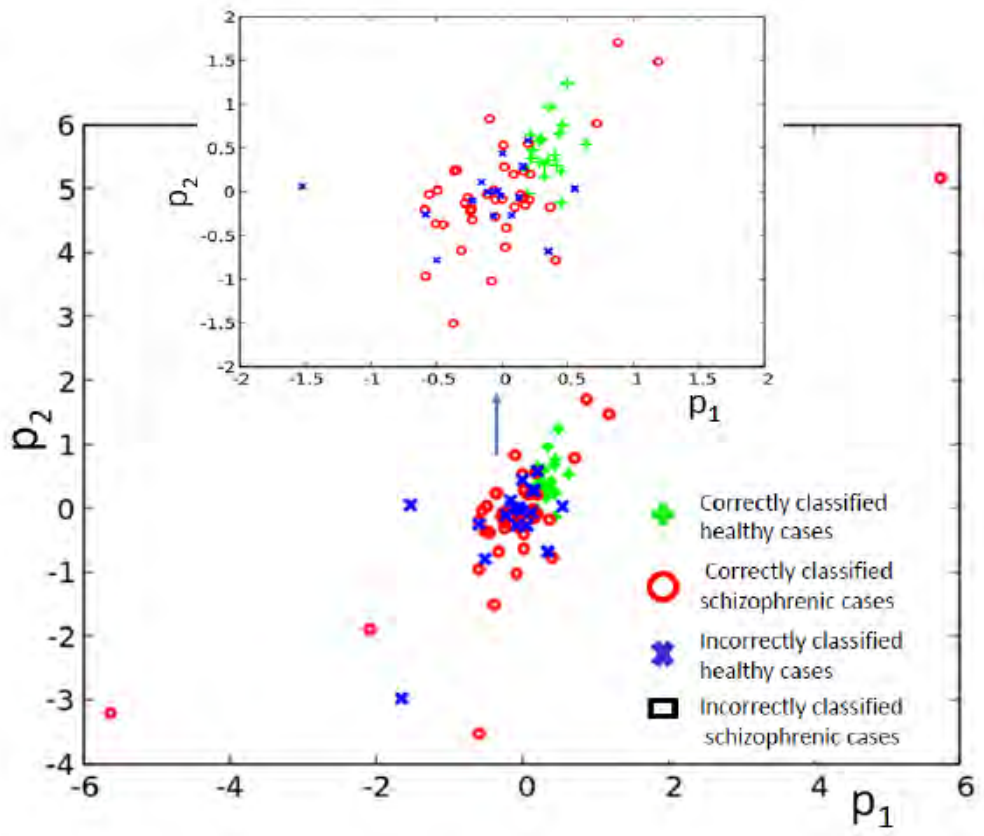


Figure 5.14: The distribution of correctly and incorrectly classified cases for 4-oscillator network (cf.Figure 5.10(b)) in the phase space (p_1, p_2).

Figure 5.13(b) illustrates the structure of the optimized classifier with four nodes. There are 2 oscillators that act as inputs of the predictor $p_{1,n}$ and a single input for predictor $p_{2,n}$. Due to the network symmetry $I(Z_1; Z_O)$ and $I(Z_4; Z_O)$ (see Figure 5.10(b)). These values (0.409 bit) are the maximum ones; thus, both oscillators #1 and #4 can be selected as the output one. In Figure 5.13(b), we marked the first of them. Figure 5.13(d) shows the distribution of numbers of activator maxima observed on the oscillator #1 for schizophrenic and healthy patients. The classification rule based on the majority of cases for a given number of activator maxima is: a patient is healthy if the number of activator maxima is 5 or 6. The observation of any other number of maxima diagnoses schizophrenia. The application of this rule gives 19 errors for 84 cases included in D_S (77 % accuracy). All incorrectly diagnosed patients are the healthy ones who are diagnosed as being schizophrenic. On the other hand, ALL schizophrenic patients were correctly diagnosed. Figure 5.14 presents locations of correctly and incorrectly classified cases for 4-oscillator network in the phase space (p_1, p_2).

The results for optimization of the 5-oscillator classifier are illustrated in Figure 5.15(a). The *Fitness* observed after 1000 optimization steps was 0.407 bit. The parameters describing the best classifier are given in Table 5.5. Figure 5.15(b) illustrates the structure of the optimized classifier. It is highly asymmetric and includes 3 normal oscillators. There are 2 oscillators that act as inputs of the predictor $p_{1,n}$ and a single input for predictor $p_{2,n}$. The highest value of $I(Z_k; Z_O)$ was observed for the oscillator $k = 3$ (see Figure. 5.15(c)) that has no direct contact with the input of predictor $p_{2,n}$. Figure 5.15(d) shows the distribution of numbers of activator maxima observed on the oscillator #3. As for the 3- and 4- oscillator cases the output oscillator does not generate small nor large numbers of activator maxima for healthy patients. The 5-oscillator network diagnoses a patient as a healthy one if the number of activator maxima is 2, 4, 5, or 6. The observation of any other number of maxima diagnoses schizophrenia. The application of this rule gives 15 errors for 84 cases included in D_S (82 % accuracy). The schizophrenic patients are diagnosed with very similar accuracy as the healthy ones (82.2 % vs. 82.1 %). Correctly and incorrectly classified cases for 5-oscillator network are located in the phase space (p_1, p_2), as shown in Figure 5.16.

Using our optimization method, we could not increase the classification accuracy above 82 % for any considered geometry (Figure 5.4 and 5.10) within 1000 optimization steps. However, the accuracy can be increased if one combines answers of different classifiers using the voting strategy. We considered three classifiers that showed the highest accuracy. They were based on 3, 5, and 6 oscillators. The same record was processed by all classifiers, and the majority verdict was taken as the answer. Such a method gave only one mistakenly diagnosed case for 45 schizophrenic records from D_S. The classification accuracy for healthy patients (39 records in D_S) was lower, and 10 such cases were misdiagnosed. Therefore, the overall accuracy of classification increased to 86.9%.

Table 5.5: Parameters of the optimized 5-oscillator network.

parameter	value
t_{max}	99.3
t_{start}	72.5
t_{end}	13.8
α	1.11
β	0.23
t^1_{illum}	28.4
t^3_{illum}	4.64
t^4_{illum}	4.64

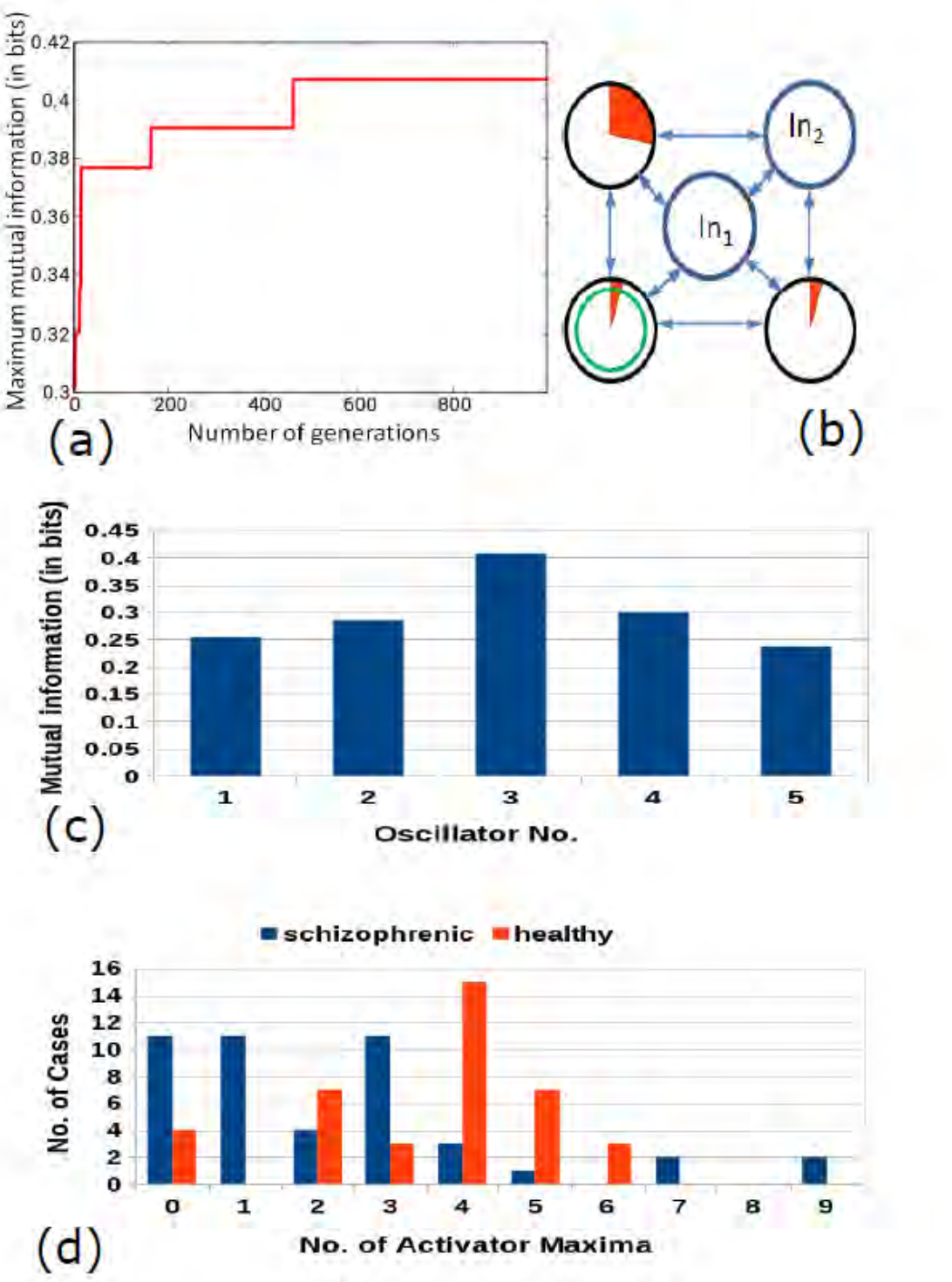


Figure 5.15: Results for 5-oscillator network (cf. Figure 5.10(c)): (a) The progress of optimization; the *Fitness* as a function of generation number. (b) The structure of the optimized network. Notation as in Figure 5.11(b). (c) The mutual information $I(Z_k; Z_O)$ for $k \in \{1, 2, 3, 4, 5\}$. The function $k \to I(Z_k; Z_O)$ has the maximum at $k = 3$. (d) The distribution of the numbers of cases for which a given number of activator maxima was observed on oscillator #3 for records representing schizophrenic and healthy patients

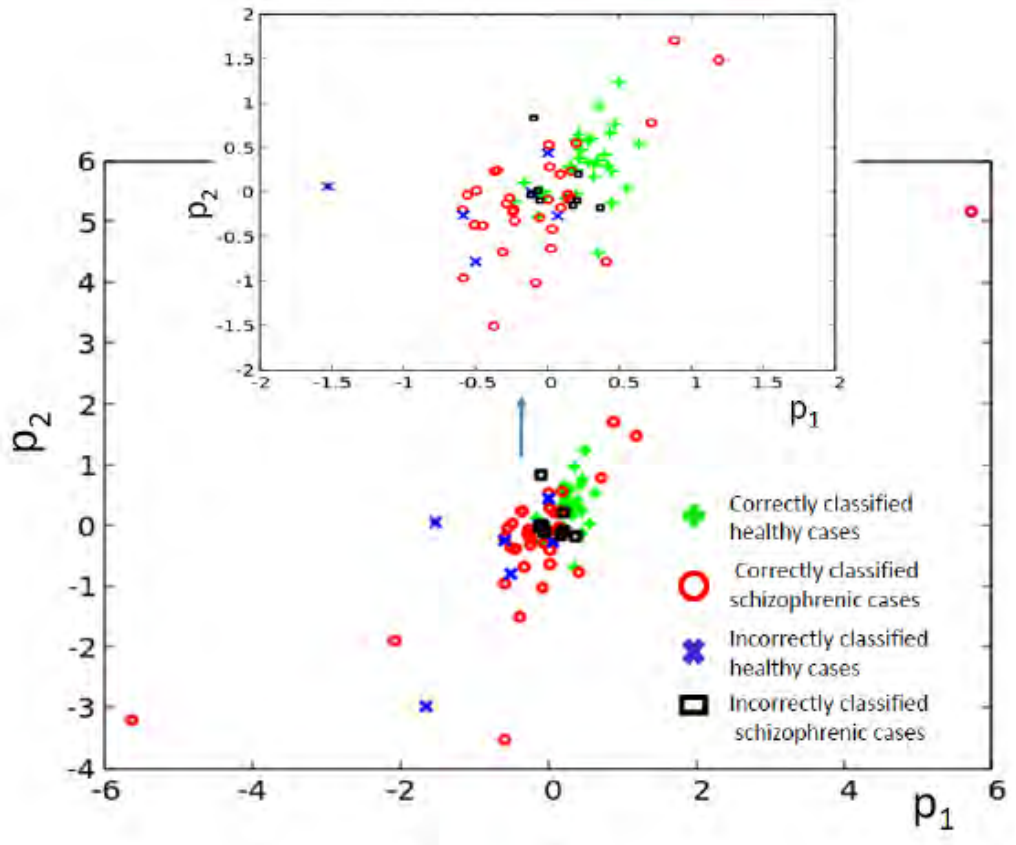


Figure 5.16: The distribution of correctly and incorrectly classified cases for 5-oscillator network (cf 5.10(c)) in the phase space (p_1, p_2)

5.2 Prediction of drug-response on patients with multiple myeloma by interacting chemical oscillators

The medical therapy of the future will be based on drugs individually selected for patient needs. Individual treatment is necessary because it has been observed that one standard treatment does not work for all patients with the same type of cancer [100]. One example is the therapy for patients with multiple myeloma. The bortezomib drug has been found to have a good anti-myeloma effect, and it has been approved for treatment for patients with multiple myeloma [101]. But literature reports suggest that myeloma, unlike other cancer types, consists of many variants with different molecular pathologies [100] [102]. For certain subtypes, the bortezomib drug can be effective, and for others, it may not. One can hope that using gene expression profiling, it is possible to determine if the drug will be effective on a particular patient or not [103]. Considering the challenges one can face with genomic analysis for each

patient [104], we present a simple method that can be used to predict the effectiveness of the bortezomib drug.

Our approach is based on processing the gene expression data with a network of chemical oscillators optimized for correlations between a single gene expression and the result of therapy. The result of therapy is described by a single bit; its value is 1 for effective therapy and 0 if the bortezomib drug does not help the patient. The information processing network takes the gene expressions as the input, and its activity is optimized for maximum correlations between the result of therapy and the number of oscillations observed on the output oscillator, which is regarded as the network answer. The ideal network should process information coming from many genes simultaneously. However, it has been observed that in the case of many inputs, the evolutionary optimization of an information processing chemical network is long and ineffective because there are many networks that are the local attractors of the algorithm. In order to simplify the numerical simulations, we restricted our attention to a simple network made of 3 oscillators that process data for a single gene expression. Networks were optimized using a set of records, including information on 239 therapies. In this set, there were 113 drug-responsive cases and 126 nonresponsive. The typical optimized network returns ~ 68 % accuracy of predicting if for a patient with a given gene expression level, the therapy using bortezomib is an effective treatment for multiple myeloma or not. We combined answers of single gene networks and made a concilium of networks that used the majority voting strategy. Such strategy predicted the patient response to bortezomib drug treatments with the accuracy of 84.5% measured on the training dataset. The results are discussed in details in the following sections.

Bortezomib therapy data and their classification

In the available dataset $R = \{r_n, n = 1, 239\}$ we have 239 patient records including both responsive (113 cases) and nonresponsive (126 cases) therapy. Each record r_n contains information on the expression of 15 genes $(e_{i,n}, i = 1, 15)$ listed in the Table 5.6. The outcome of the therapy $g_n \in \{0, 1\}$, where 0 stands for nonresponsive therapy and 1 denotes responsive one, can be regarded as the record type. Therefore, a record r_n has the form of 16-tuple: $r_n = (e_{1,n}, ..., e_{15,n}, g_n)$. The problem of deciding if the therapy with bortezomib can be effective for a given patient reduces to the problem of finding an algorithm that gives the correct record type g provided that the predictor values $e_i, i = 1, 15$ are known. It can be noticed that on this test group of patients, two trivial algorithms:
(1) always use bortezomib

and
(2) do not use bortezomib because it will not help
lead the correct therapy in 47 % and 53 % cases.

Table 5.6: Information on the range of gene expression recorded in R.

Gene No. i	Gene Name	Gene expression value range in the set R: $Min_n(e_{i,n}) - Max_n(e_{i,n})$	Histogram Accuracy
1	SERP1	315.8 – 1879.8	56.9%
2	NPM1	1556.8 – 9849.1	56.9%
3	PIK3R1	59.8 – 437.1	57.7%
4	APEX1	200.7 – 1741.2	58.9%
5	DAPP1	69.9 – 564.5	59.4%
6	NRAS	38.8 – 679.2	55.6%
7	RRAGC	148.9 – 679.2	61.08%
8	CFLAR	135.7 – 2000.7	56.9%
9	CXCL5	1.09 – 58.8	57.3%
10	IL15	13.3 – 562.05	58.9%
11	NFKβ2	54.7 – 2848.01	56.9%
12	COX7C	559.5 – 5476.9	61.5%
13	RPS7	1142.48 – 12167.3	62.3%
14	RPS13	2079.7 – 23208.7	60.6%
15	UQCRH	370.82 – 5554.15	60.2%

The values of a single gene expression weakly correlate with the efficiency of the bortezomib therapy. Figure 5.17 illustrates the correlations between the expression of the RPS7 gene and the responsive/nonresponsive results of bortezomib therapy. The whole range of gene expression numbers is divided into 10 subintervals (bins) of the same length with the ranges given in Table 5.7. It can be noticed that if the gene expression value is within bins no. 1, 2, 3 or 10, then the probability of successful therapy is higher than its failure. On the other hand, if the expression of the RPS7 gene is above 4445 but below 11065 then it is more likely that multiple myeloma will not respond to the drug. Using this rule we can plan the therapy with the accuracy of 62.3 %, which is much higher than that of the trivial algorithms mentioned above. Of 239 cases included in the dataset R we obtain 70 correctly determined nonresponsive cases and 79 correctly determined responsive ones. We also observe 56 wrongly determined nonresponsive cases and 34 wrongly determined responsive ones. The accuracy of rules following from histograms for the other gene expression values are given in the Table 5.6. In the following, we show how this

accuracy can be improved with simple classification networks based on interacting chemical oscillators.

Table 5.7: The relationship between the bin number and the range of gene expression values.

bin number	range of gene expression values in the bin
1	1142.4 - 2244.9
2	2244.9 - 3347.4
3	3347.4 - 4449.9
4	4449.9 - 5552.4
5	5552.4 - 6654.8
6	6654.8 - 7757.3
7	7757.3 - 8859.8
8	8859.8 - 9962.3
9	9962.3 - 11064.8
10	11064.8 - 12167.3

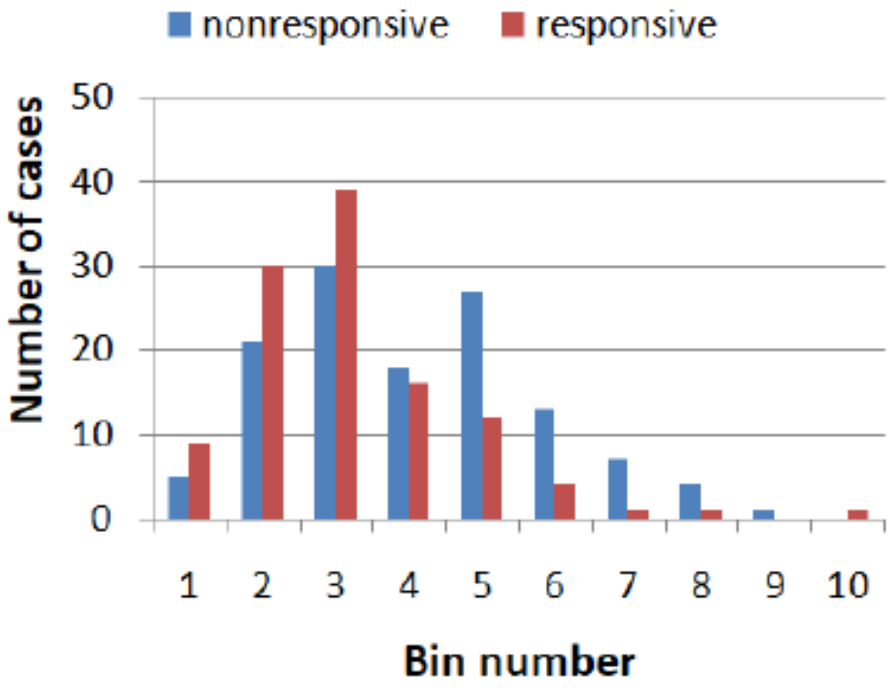


Figure 5.17: The histogram showing correlation between the expression of the RPS7 gene and responsive (red)/nonresponsive (blue) results of bortezomib therapy. The range of gene expression values corresponding to each bin are given in Table 5.7

It can be seen in Table 5.6 that the range of results significantly differs between genes. In order to unify the data we normalized the set of gene expression values for each gene i for all 239 patients using the formula:

$$p_{i,n} = \frac{e_{i,n} - Min_n(e_{i,n})}{Max_n(e_{i,n}) - Min_n(e_{i,n})} \tag{5.2.1}$$

As an example the data $\{p_{13,n}, n = 1, 239\}$ are illustrated in Figure 5.18. For better visibility of data we show points $\{(p_{13,n}, y_n), n = 1, 239\}$ where y_n is a random number in the interval $[0, 1]$. The random y-coordinates $Y = \{y_n, n = 1, 239\}$ were added to make the distribution easier to see, but of course, only the x-coordinate has a medical meaning. The same set of random y-coordinates Y is used in Figures 5.19, 5.22 and 5.24. The red and blue points in Figure 5.18 mark responsive and nonresponsive results of bortezomib therapy, respectively. It can be seen that points corresponding to responsive and nonresponsive cases are not differentiated by the x-coordinate, related to the expression of gene number 13.

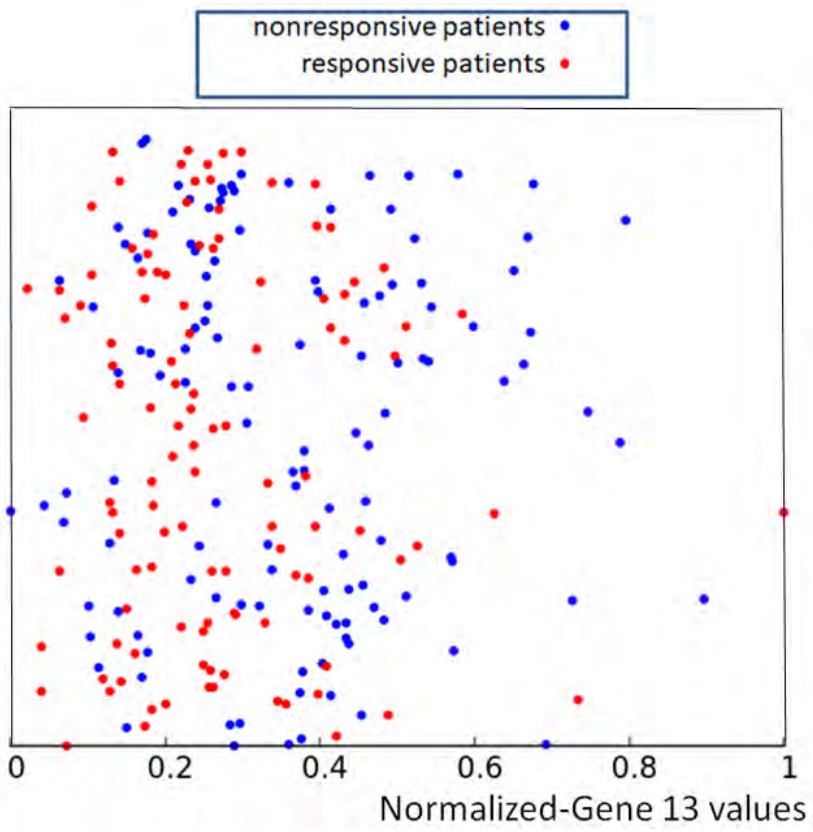


Figure 5.18: The distribution of normalized values of the RPS7 gene expression corresponding to responsive (red) and nonresponsive (blue) results of bortezomib therapy respectively. The values of $p_{13,n}$ are represented by the x-coordinate of marked points. The y-coordinate was randomly generated to differentiate points.

If we apply the rule following from the histogram shown in Figure 5.17 to the dataset of normalized data $Q = \{m_n, n = 1, 239\}$ where $m_n = (p_{1,n}, ..., p_{15,n}, g_n)$ we obtain the distribution of correctly and wrongly determined results of bortezomib therapy illustrated in Figure 5.19

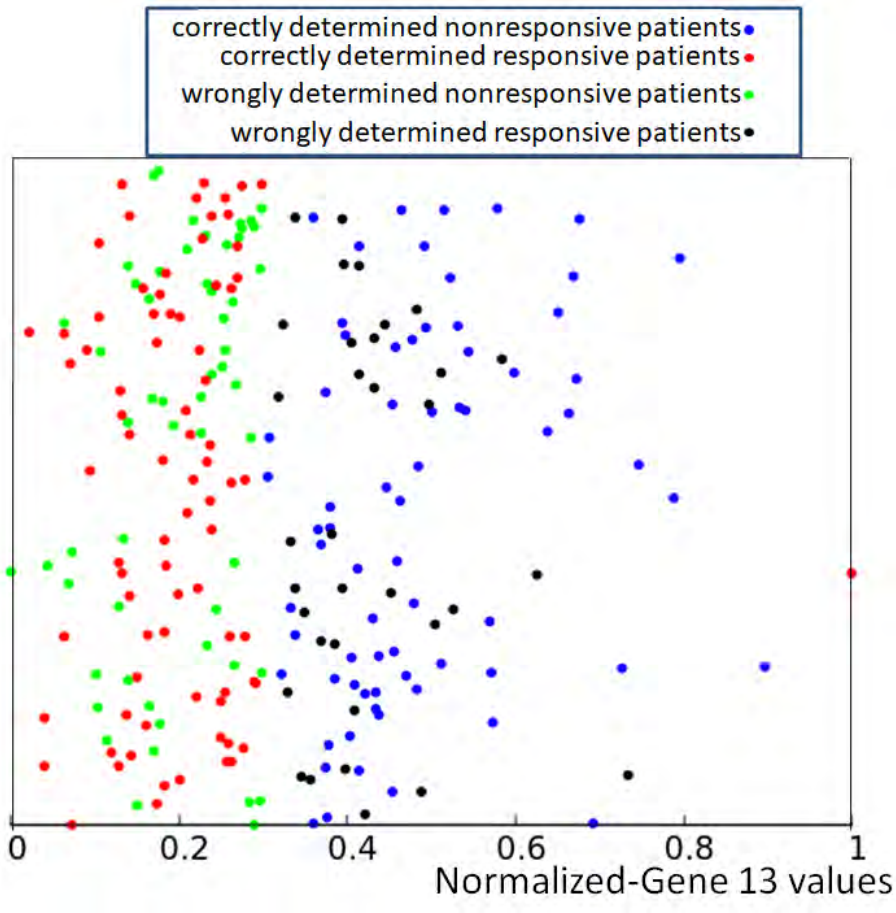


Figure 5.19: The distribution of normalized values of correctly and wrongly determined therapies using the accuracy rule for the histogram of Figure 5.17. Red and blue points mark correctly determined responsive and nonresponsive records respectively. Black and green points mark wrongly determined responsive and nonresponsive records. The values of $p_{13,n}$ are represented by the x-coordinate of marked points. The y-coordinate was randomly generated to differentiate points and it is the same as in Figure 5.18.

We chose a network of chemical oscillators with three nodes to solve this problem. The oscillators were simulated with two-variable Oregonator model (Eq 3.3.7 and 3.3.8, with m=3). 5^{th} order Cash-Carp algorithm [105] was used to solve the differentials with 10^{-3} time steps. The simulation was carried out till time t_{max}.

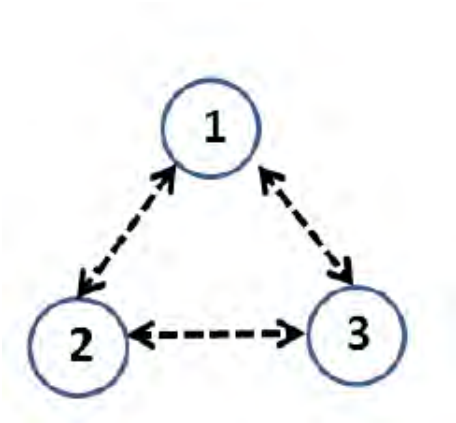


Figure 5.20: The network comprises of three oscillators. The oscillators are marked with numbers. The dotted lines signifies interactions among them.

The oscillators on the network can be either an input one or a normal oscillator. If on optimization, the j^{th} oscillator is considered as an input of i^{th} predictor $p_{i,n}$ for patient n, then the value of t^j_{illum} is functionally related to $p_{i,n}$. The relation is shown as follow:

$$t^j_{illum} = t_{start} + (t_{end} - t_{start}) * p_{i,n} \tag{5.2.2}$$

If the j^{th} oscillator is the normal one, then its illumination time t^j_{illum} will be decided by the evolutionary algorithm and it wont be having any dependency on the predictor value. For fixed parameters describing the network, we select the output oscillator as the one that produces the highest accuracy of predicting record types for the dataset used for network training. In order to determine the network accuracy we applied the following method. The network time evolution is simulated for all records of Q. For each oscillator and for each number of activator maxima we formulate the relationship between the number of activator maxima and the outcome of therapy based on majority of cases. The oscillator for which the number of errors is minimized is considered as the output one. Keeping in mind the symmetry of the considered network, we can assume that the oscillator #1 is the input oscillator and the oscillator #3 is a normal one. Of course, the role of the oscillator #2 is the subject of optimization. So, the parameters to be optimized are : t_{max}, α, β, t_{start}, t_{end}, t^3_{illum}, the role of oscillator #2 and if it is a normal one t^2_{illum}.

We optimize the network by the evolutionary algorithm. The optimization started with 100 networks with randomly generated parameters. The next generation of networks included 5 top fit networks of the previous generation and 95 networks formed by recombination of parameters of two networks selected from 40 best networks of the previous generation. Each network obtained by recombination was then allowed to mutate. Mutations included the values of all parameters and the type of oscillator #2.

The optimized network was obtained after 500 evolutionary steps. The fitness function of evolutionary optimization was the maximum accuracy between the network output coded in the number of activator maxima observed on one of the oscillators and the record type g.

Results

The evolutionary optimization described in the previous section was used to design networks with a high correlation between the number of activator maxima observed on one of the oscillators and the result of bortezomib therapy. As an example, we show such network for linking the RPS7 normalized gene expression value with the success of therapy. For this input the evolutionary optimization produced the network illustrated in Figure 5.21(a). It is composed of two input oscillators, marked with In_{13}, that accept the value of $p_{13,n}$. The rightmost oscillator, marked with the double circle, is the output one. The circle marking this oscillator is also a base for a pie chart representing the ratio t^3_{illum}/t_{max} by the surface of red slice. Figure 5.21(b) shows the distribution of a number of activator maxima observed on the output oscillator for responsive and nonresponsive multiple myeloma treatment with the drug bortezomib. On this basis, we can define the rule that ensures the highest accuracy on the training dataset Q. For this network, the rule is:
- if 1,3,4,5,9 or 10 activator maxima are observed on the output oscillator then the patient belongs to the responsive group.
- if another number of activator maxima is observed, then unsuccessful treatment with the drug bortezomib is expected.
If applied to the training dataset Q this rule gives 71.1% of correct answers, which is by 5% higher than the trivial rule based on the distribution of gene expression values (cf. Figure 5.17). Of 239 cases included in the dataset Q we obtain 91 correctly determined nonresponsive cases and 79 correctly determined responsive ones. We also observe 35 wrongly determined nonresponsive cases and 34 wrongly determined responsive ones. The distribution of correctly and incorrectly classified points from the training dataset is illustrated in Figure 5.22. Here again, the y-coordinate is random, and it is the same as introduced to differentiate the values of $p_{13,n}$ in Figure 5.18.

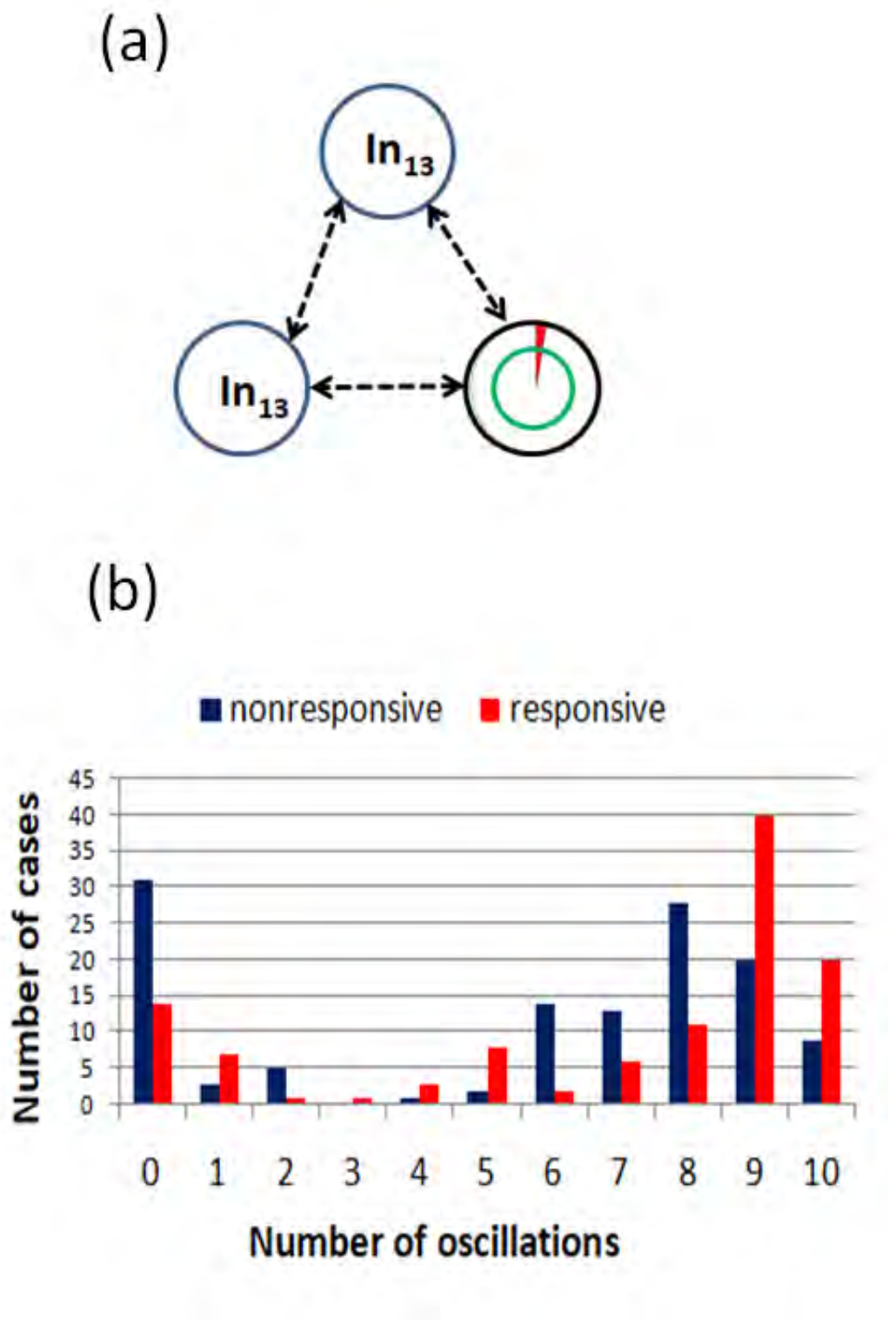


Figure 5.21: (a) The structure of a network that produces a high correlation between the RPS7 gene expression and the result of bortezomib therapy. The symbol In_{13} marks input oscillators, taking values of $p_{13,n}$. The rightmost oscillator is a normal one, and it is also the output oscillator what is indicated with a double circle. (b) Correlations between the success of therapy and the number of activator maxima observed on the output oscillator.

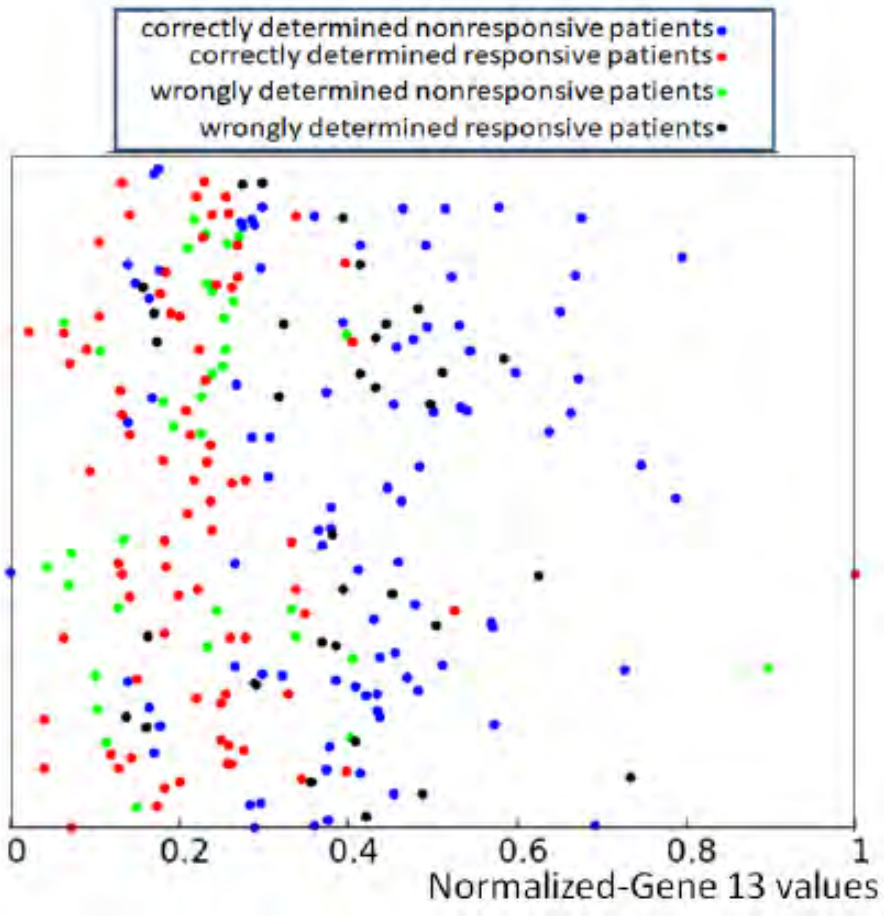


Figure 5.22: Location of correct and incorrect predictions on the result of bortezomib therapy by a network of oscillators optimized for the RPS7 gene. The values of $p_{13,n}$ are represented by the x-coordinate of marked points. The y-coordinate is randomly generated to differentiate points and is the same as in Figure 5.18

We optimized 3-oscillator networks for all genes listed in Table 5.6. The parameters of all optimized networks are listed in Table 5.8. The gene numbers in Table 5.8, Table 5.9, Table 5.11 and Figure 5.23 correspond to those in Table 5.6. The rules that translate the number of activator maxima on the output oscillator and the therapy result are given in Table 5.9. The accuracy of optimized information processing networks is shown in Figure 5.23, and in Table 5.9. It is in the range between 66.5 % (genes SERP1, CXCL5 and IL15) to 71.1% (gene RPS7).

Table 5.8: The parameters of networks that give the best correlations between the number of activator maxima on the output oscillator and the therapy result.

Gene No.	t_{max}	t_{start}	t_{end}	α	β	input oscillators	output oscillator	normal oscillators	t^j_{illum}
1	77.35	28.11	2.31	0.69	0.08	1	2	2, 3	t^2_{illum}= 17.63 t^3_{illum}=7.24
2	77.55	2.03	33.77	0.7	0.07	1	1	2, 3	t^2_{illum}= 8.73 t^3_{illum}=4.33
3	80	63.81	2.26	0.7	0.08	1, 2	1	3	t^3_{illum}=8.58
4	77.45	6.02	35.37	0.68	0.074	1	1	2,3	t^2_{illum}=8.26 t^3_{illum}=2.84
5	80	2.45	68.46	0.62	0.06	1, 2	3	3	t^3_{illum}=4.34
6	80	2.1	42.37	0.7	0.09	1	1	2, 3	t^2_{illum}==2.58 t^3_{illum}=12.11
7	80	1.87	69.8	0.63	0.06	1	2	2, 3	t^2_{illum}=3.51 t^3_{illum}=5.28
8	70.3	1.82	32.68	0.70	0.081	1	1	2, 3	t^2_{illum}=2.64 t^3_{illum}=10.68
9	74.63	2.49	28.12	0.70	0.08	1, 2	3	3	t^3_{illum}=10.56
10	80	2.56	29.14	0.7	0.06	1, 2	3	3	t^3_{illum}=4.33
11	80	2.74	29.05	0.69	0.07	1	2	2, 3	t^2_{illum}=3.16 t^3_{illum}=7.25
12	70.79	70.65	1.57	0.59	0.09	1, 2	1	3	t^3_{illum}=13.42
13	80	2.48	51.09	0.7	0.07	1, 2	3	3	t^3_{illum}=2.0
14	80	36.28	2.79	0.7	0.08	1, 2	3	3	t^3_{illum}=4.39
15	75.71	62.57	3.52	0.66	0.08	1, 2	3	3	t^3_{illum}=0.78

Table 5.9: The rules that translate the number of activator maxima on the output oscillator and the effective therapy using the bortezomib drug.

Input gene no.	number of activator maxima for responsive therapy	number of activator maxima for nonresponsive therapy	accuracy %
1	0, 1, 3, 5, 7, 10	2, 4, 6, 8, 9, 11	66.5
2	3, 6, 7, 9, 10	1, 2, 4, 5, 8, 11	69.03
3	2, 6, 7, 8, 9, 11	1, 3, 4, 5, 10	68.2
4	1, 2, 3, 4, 7, 9	0, 5, 6, 8	67.7
5	1, 5, 7, 9	3, 4, 6, 8, 10, 11	68.2
6	2, 5, 6	0, 1, 3, 4, 7, 8, 9	67.7
7	2, 6 ,9	1, 3, 4, 5, 7, 8, 10, 11	69.03
8	0, 2, 4, 6, 7, 8, 9	1, 3, 5, 10	69.4
9	0, 2, 3, 5, 7, 8	1, 4, 6, 9, 10	66.5
10	1, 3, 6, 7, 10	2, 4, 5, 8, 9, 11	66.5
11	1, 2, 6, 7, 11	3, 4, 5, 8, 9, 10	69.8
12	2, 3	1, 3, 4, 5, 6, 7, 8	69.8
13	1, 3, 4, 5, 9, 10	0, 2, 6, 7, 8	71.1
14	0, 1, 2, 4, 7	3, 5, 6, 8, 9, 10, 11	68.2
15	3, 4	1, 2, 5, 6, 7, 8, 10	69.4

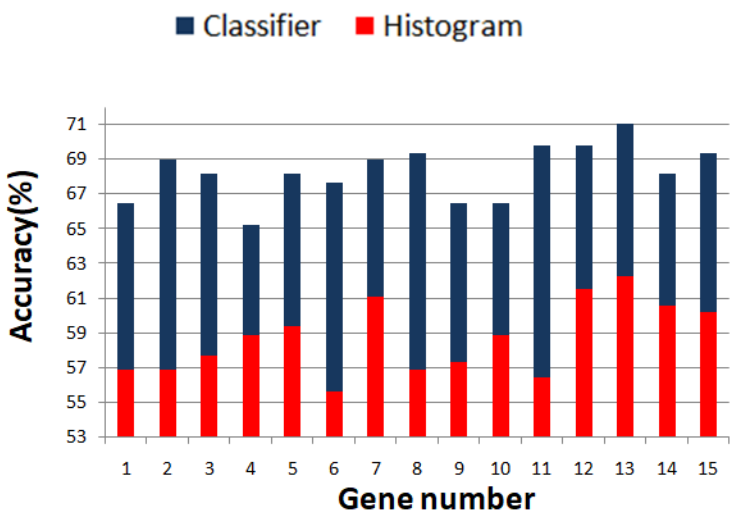


Figure 5.23: Comparison between the accuracy of predictions of the result of bortezomib therapy based on the histogram of gene expression values (red bars) and the optimized network of oscillators (blue bars). The gene numbers correspond to those in Table 5.6.

To increase the accuracy in determining the success or failure of bortezomib therapy, we called a concilium of optimized networks. Each member of concilium is a network specialized in finding correlations between the expression value of one gene and the success of bortezomib therapy and has one vote. The final decision is taken on the basis of majority voting. The accuracy of determining if the therapy using the bortezomib drug is efficient or not for different majority rules applied to the concilium is shown in Table 5.10. We can see that the decision based on the opinions of more than half of concilium members is accurate for almost 85 % of cases included in the training dataset. Of 239 cases included in the dataset Q we obtained 117 correctly determined nonresponsive cases and 86 correctly determined responsive ones. We also observe 9 wrongly determined nonresponsive cases and 27 wrongly determined responsive ones. Their distribution is illustrated in Figure 5.24.

Table 5.10: The accuracy of determination if the therapy using the bortezomib drug is efficient for different majority rules. The results for the optimized network.

the majority rule	the concilium accuracy
14 or more votes for	7.1%
13 or more votes for	17.1%
12 or more votes for	30.9%
11 or more votes for	43.0%
10 or more votes for	61.5%
9 or more votes for	77.4%
8 or more votes for	84.9%

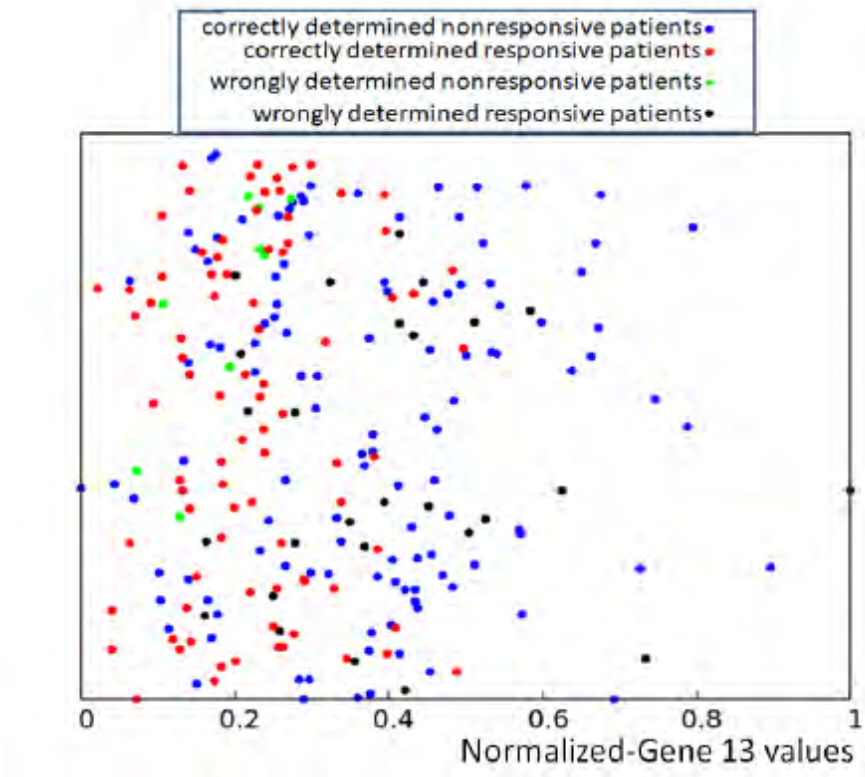


Figure 5.24: Location of correct and incrorrect predictions on the result of bortezomib therapy predicted by a concilium of networks of oscillators optimized for all genes. The values of $p_{13,n}$ are represented by the x-coordinate of marked points. The y-coordinate is the same as in Figure 5.18

But do we need oscillator networks? Correlations between the values of single gene expression corresponding to responsive and nonresponsive cases from the training dataset and the results of bortezomib therapy can be extracted from the histograms of single gene expression values, as it was done for the RPS7 gene (cf. Figure 5.17). For this gene the accuracy of such method is 62.3 % which is not much lower than that of optimized classifier (71.1 %). So why not try to make the concilium based on histograms for all genes? We have tested such approach dividing the whole range of gene expression values into 10 sub intervals and introducing the rule based on the majority of cases in subintervals for each gene. Next we applied the majority voting strategy for all records of Q. The accuracy of such concilium was 66.9%. Of 239 cases included in the dataset Q we obtained 100 correctly determined nonresponsive cases and 60 correctly determined responsive ones. We also observe 26 wrongly determined nonresponsive cases and 53 wrongly determined responsive ones. Their distribution is illustrated in Figure 5.25. Therefore the accuracy of such concilium is much smaller than the concilium based on networks of oscillators optimized for correlations between the single gene expression value and the outcome of therapy.

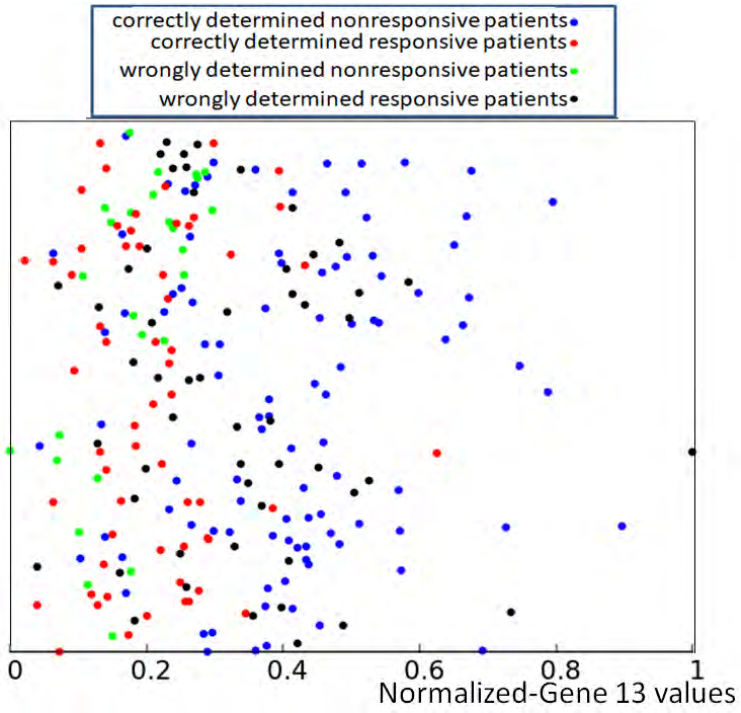


Figure 5.25: The distribution of normalized values of correctly and wrongly determined therapies using the majority rule for the histogram . Red and blue points mark correctly determined responsive and nonresponsive records respectively. Black and green points mark wrongly determined responsive and nonresponsive records. The values of $p_{13,n}$ are represented by the x-coordinate of marked points. The y-coordinate was randomly generated to differentiate points and it is the same as in Figure 5.18.

In order to check how sensitive to fluctuations of parameters are the results produced by the concilium made optimized networks, we considered random modifications in parameter values. For each optimized network we selected one parameter in random and decreased or increased its value by ±1%. The details on applied modifications and their influence on the accuracy of each network are given in Table 5.11. In all cases the accuracy decreased by 2-3% (comparing accuracy in Table 5.11 and Table 5.9). A similar decrease of accuracy is observed for the decision of concilium (82.8 % for the concillium of modified networks). Still such accuracy is high enough to claim that concilium strategy for determination of bortezomib effectiveness is robust to random changes of parameters and fluctuations in the medium.

Table 5.11: The accuracy of modified oscillator networks for correlations between the gene expression value and the therapy result. The table also defines modification introduced to the optimized network.

Gene no.	New parameter=old parameter±1% of old parameter	output oscillator	Accuracy
1	$t_{max} + 1\%$	2	64.8%
2	$\beta + 1\%$	1	66.9%
3	$\beta - 1\%$	1	66.9%
4	$t_{start} + 1\%$	1	62.7%
5	$t^3_{illum} - 1\%$	2	65.6%
6	$t_{end} + 1\%$	1	64.01%
7	$\beta + 1\%$	2	67.3%
8	$t^2_{illum} - 1\%$	1	66.9%
9	$\alpha - 1\%$	2	62.3%
10	$t^3_{illum} + 1\%$	3	63.1%
11	$t^2_{illum} + 1\%$	2	62.3%
12	$t_{end} + 1\%$	3	69.4%
13	$\beta - 1\%$	3	69.8%
14	$t_{max} + 1\%$	2	66.1%
15	$t^3_{illum} - 1\%$	3	69.03%

Table 5.12: The accuracy of determination if the therapy using the bortezomib drug is efficient for different majority rules. The results for the network with modified parameters (cf. Table 5.11) .

the majority rule	the accuracy of concilium with modified networks
14 or more votes for	8.3%
13 or more votes for	17.1%
12 or more votes for	28.03%
11 or more votes for	40.5%
10 or more votes for	56.9%
9 or more votes for	69.03%
8 or more votes for	82.8%

5.3 Conclusions

In the presented study, we assumed that schizophrenia could be detected by a chemical oscillator network that analyses EEG signals recorded from electrodes located on a patient scalp. We used two-variable Oregonator model to simulate the oscillators and to establish interactions among them. A network of 6 coupled oscillators gave 82% of correct diagnosis for cases in the considered dataset. Its modification based on three stage diagnosis on parts of the recorded signal produced even higher, 90% accuracy of diagnosis. Then we considered three different network sizes to study the effect of determination on accuracy. We expected that the schizophrenia diagnosis accuracy would increase with the network size. However, the networks formed of 3, 5, and 6 oscillators gave 82% of correct answers for cases included in the training dataset. The accuracy increased to 86.9% if 3 networks process a case and the majority rule is used to select the final answer. The fact that larger networks did not produce better results than the small ones could be related to inefficient optimization for a large number of parameters that were taken into account. The problem can be overcome by a larger population of classifiers and a larger number of optimization

steps. However, both methods increase the numerical complexity of the optimization.

In section 5.2 I discussed the application of information processing network formed by chemical oscillators for determination of the outcome of a therapy with bortezomib drug. The network input information comes form the gene expression values. Information was processed by simple networks, each made of 3 oscillators. Each oscillator was optimized to find correlations between the expression value of a particular gene and the outcome of the therapy. Individual classifiers gave the accuracy in the range between 66.5% to 71.1% (cf. Figure 5.23). To improve the determination of the therapy outcome we considered the concilium of 15 classifiers and accepted the majority decision. Such strategy increased the accuracy to almost 85%, which seems to be a promising result for the further development of the method.

Chapter 6

Towards experimental realization of a chemical computer using Belousov-Zhabotinsky reaction

This chapter explores the possibility of designing chemical computers experimentally. It gives an insight into the experimental works done previously with B-Z (Belousov-Zhabotinsky) reaction. Experimental designing for coupling between B-Z droplets and controlling excitation by blue light has been presented. The idea of producing confined oscillations in "beads" has also been discussed. I also presented some preliminary findings of my experimental results.

In this chapter, I discuss the possibilities to show chemical computation by B-Z oscillators experimentally. In our computational approach, the main criteria of designing a chemical computer were the introduction of a network of chemical oscillators, establishing interactions between them and using illumination control to train the system. Hence, here I have given a brief overview of a few noteworthy works done in the past years, providing information about the possible formation of a network of B-Z oscillators, studying the characteristics of coupling between these oscillators and methods of controlling oscillations in each individual oscillator.

6.1 Illumination controlled B-Z droplets

Gyzinski et al. [106] set up such experiments, where they showed how each single B-Z droplet (i.e a droplet with solution of B-Z reaction reagents surrounded by an organic solvent) could be controlled by blue light illumination. They presented an experimental setup (Figure 6.1) showing how each droplet was separately illuminated with blue light which was transmitted through the optical fibers attached to the diodes. These light emitted fiber tips were placed just below the corresponding droplets. The illumination time was controlled using a personal computer. By performing experiments with non interacting oscillating droplets and illuminating a few of those oscillators in particular time intervals, they compared the time period of non-illuminated droplets with that of the time period of the illuminated droplets observed after the light was switched off. They found that the period of the droplets was the same before and post illumination operation. Hence concluded that the sequence of illuminations does not change the inherent period of oscillations in droplets.

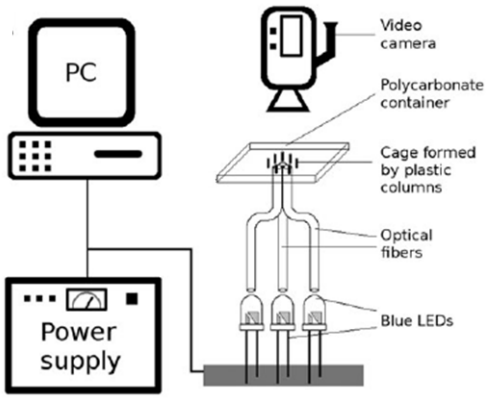


Figure 6.1: Schematic representation of the experimental setup. Each droplet has been separately illuminated with blue light transmitted by the optical fibers with it's tip attached to the bottom of the container [106].

In the next section, an experimental realization of designing a chemical "memory" is presented using illumination as its control factor.

6.2 Realization of existing chemical memory in interacting B-Z droplets

The potential of storing information in oscillating B-Z droplets, controlled by illumination was investigated experimentally by Gizynski et. al. [106]. They prepared three droplets (arranged in a triangular manner) containing B-Z solution by pipetting B-Z reagent into a polycarbonate container filled with asolectin in decane. The plastic columns were attached at the bottom of the container to fix each droplet at a particular location (schematic representation of the arrangement shown in Figure 6.1). At the bottom of the container, the droplets flattened. These flattened droplets, each caged under stabilizing columns, expands toward each other by forming a stable lipid bilayer which can be penetrated by the activators of the B-Z reaction establishing interactions among the droplets. Chemical oscillation was controlled by blue light. Light emitting fiber tips were used for each droplet, so that the droplets could be controlled individually.

They observed different oscillation modes in three droplets interacting with each other, as shown in Figure 6.2. The applied illumination strategy was that all three droplets were allowed to oscillate in dark conditions for a few periods. And then each one of them was illuminated with the same light intensity. The time of illumination considered was equal to the inherent oscillation period (60 s). In Figure 6.2 the in-phase mode implies the situation when all droplets get excited at the same time. The anti-phase implies that the excitation of one droplet is shifted in time with respect to others by T/2. In the forcing modes: 2a, 2b and 2c the first two oscillators were in phase with each other and together activated the third one. In modes 3a, 3b and 3c excitation of one of the droplets force synchronize the other two. Another type of forcing mode was observed when one droplet excites the other and that droplet in turn excites the third droplet and this rotation of excitation can be either clockwise or anticlockwise. The stable modes where the clockwise or anticlockwise movement was sustained for a long time were studied. They observed mode 6 for 145 s and then reverted the illumination, which resulted in a stable clockwise mode. To go back to the anticlockwise mode was achieved after choosing the right illumination sequence. Therefore, the fact that illumination can cause switching between clockwise and anticlockwise modes confirms that if information can be coded in states of interacting

droplets, it can be a good candidate for a chemical memory.

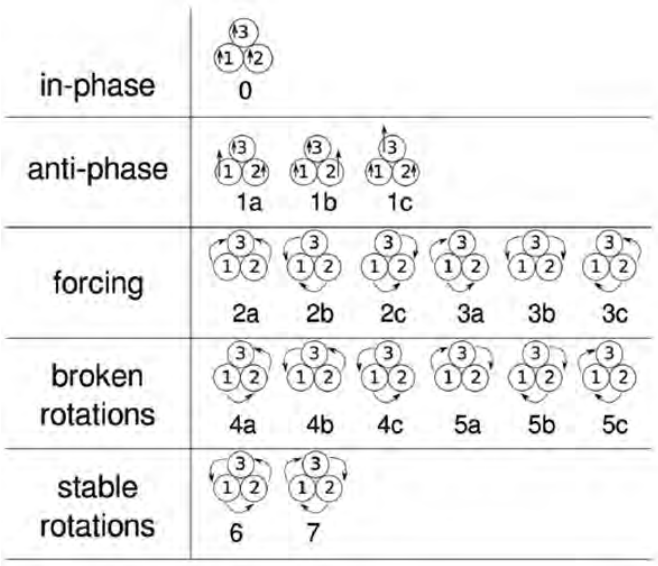


Figure 6.2: The model of three coupled B-Z droplet arranged in triangular form. Vertical arrow depicts the movement of the pulse from the droplet center towards the surface. The arrows pointing from one droplet to the another illustrates the propagation of chemical excitation [106].

6.3 B-Z oscillations in DOWEX beads

Many experimental systems were studied for wave-propagations in B-Z reaction. Some systems were made with gels [107], droplets [108] and cation-exchange resin beads. As using the ferroin loaded cation exchange resin (DOWEX) beads as our oscillator can be useful for spatially distributing it, we chose to work with these. Moreover, the raisin beads can offer a bubble free system for many hours. The formation of bubbles can specially create a problem when we are trying to stabilize interactions among two oscillating beads.

Coupling between B-Z oscillators

As our main aim is to create a network of B-Z oscillators to perform information processing, I focused on the results of coupling between two oscillators. Understanding the interactions in between oscillating beads has been studied for many years. Nobuaki Nishiyama and Kaori Eto [109] investigated the 3 coupled oscillator interactions. The cation-exchanged beads (DOWEX, 50W-X2) loaded with ferroin (bead size 110 $\mu m, 90\mu m, 90\mu m$) were arranged in a triangular form in a petri dish containing B-Z solution. They reported the observation of two types of out of phase oscillations. One, the phase wave moves from the direction of 1 to 3 to 2 and the

other corresponds to the phase wave moving in the direction 2 to 3 to 1 (Figure 6.3). Finally, the author proposed the idea that the present coupled oscillator system can be easily extended to that with much more oscillators and may serve as a realistic model for neural network systems.

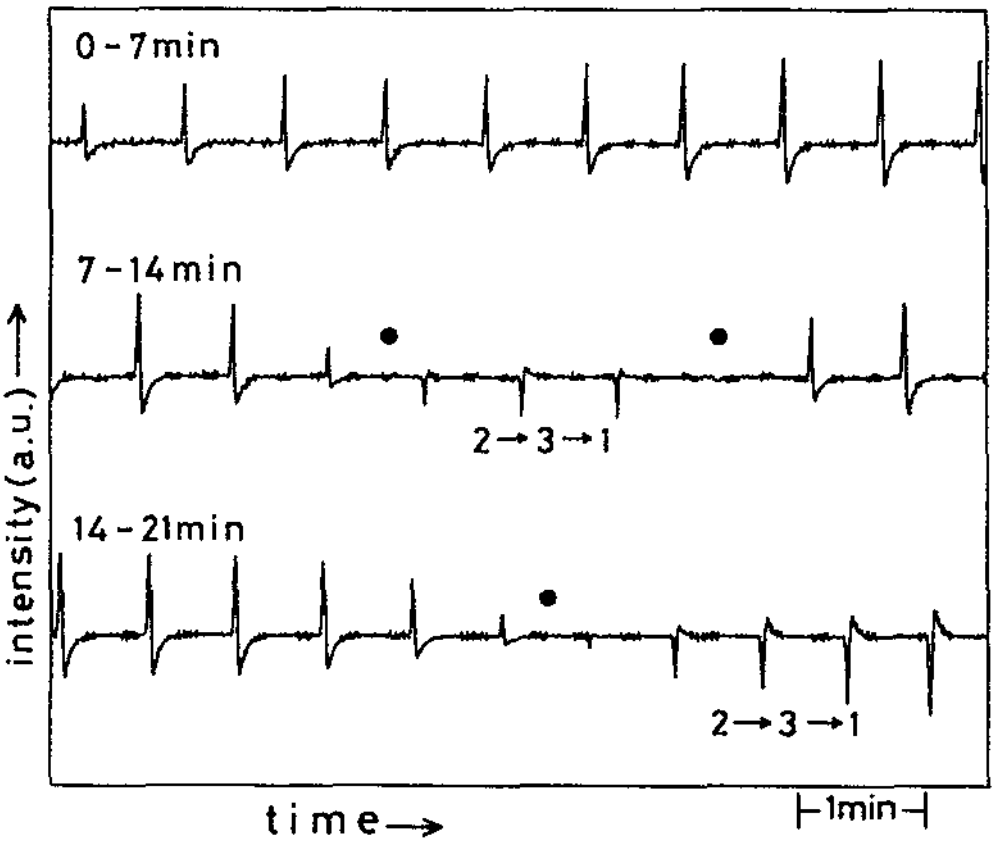


Figure 6.3: Intensity difference of the lights transmitted by bead 1 and bead 2. The two out-of-phase modes, corresponds to the phase waves running in opposite directions ($2 \rightarrow 3 \rightarrow 1$ and $1 \rightarrow 3 \rightarrow 2$) [109]. The dark circles represent the switching time between these modes. Distance between the beads are : $170\mu m$ (bead 1-2), $150\mu m$ (bead 2-3) and $140\mu m$ (bead 1-3). Diameters of the three beads are 110 μm (bead 1), 90 μm (bead 2), and 90 μm (bead 3). Periods of three beads before coupling are 55s (bead 1), 62 s (bead 2) and 48 s (bead 3).

The recent work Kuze et al. [110] also demonstrated the coupling between ferroin loaded oscillating resin beads. The experimental setup was the following: Two ferroin loaded beads were placed in two separated silicone rubber plates, immersed in a rectangular vessel filled with B-Z solutions. One of B-Z beads was placed in a silicone rubber plate connected to a stepper motor to alter the distance between two beads (l). They showed that with the decrease in distance between the two beads, the oscillation period between them drastically decreased (Figure 6.4).

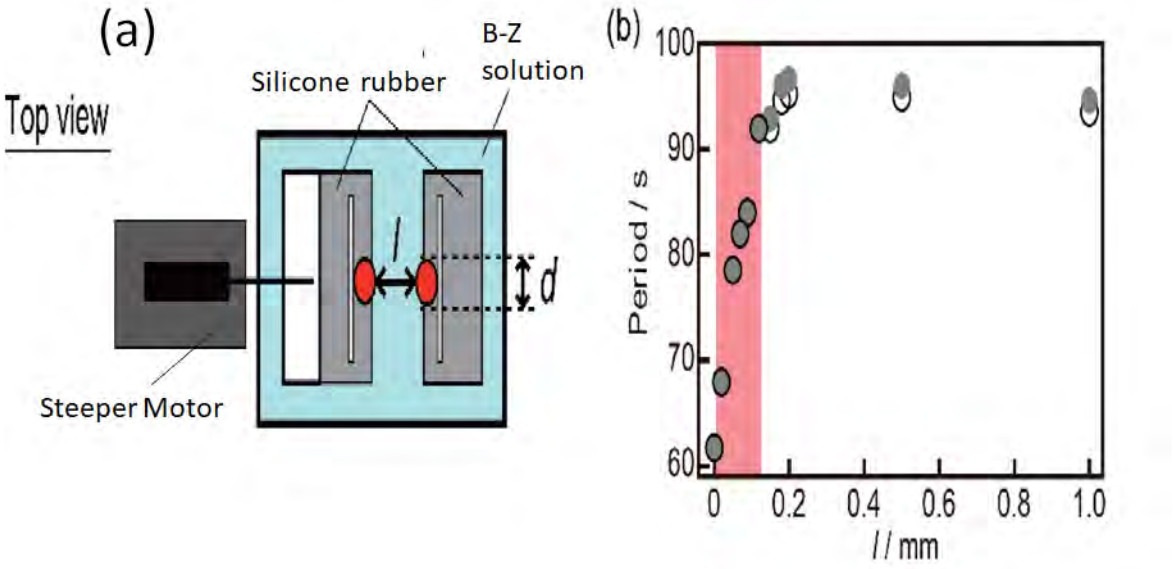


Figure 6.4: (a) Schematic representation of the experimental system (top view). (b) Lower (empty circle) and higher (gray circle) periods of the beads vs. the distance (l) between them [110]. The diameter of the beads were 0.73 mm

Electrode potential controlled B-Z beads

Kuze et al. [111] set up an experiment where they have shown that the oscillation of the B-Z reaction can be controlled and manipulated by the electric potential. Figure 6.5 (figure on the left) shows the experimental setup. They used ferroin loaded Dowex 50W-X4 beads [112] and immersed them in a B-Z solution without a catalyst. A single bead was placed on the platinum plate electrode. With electric potential(E) < -0.75 V, they observed travelling waves. At $0.50V <= E <= 0.75V$ only global oscillations were observed and with $E >= 1V$, no oscillations were observed [111]. With electric potential being negative, Br^- ion gets repelled from the surface of the B-Z beads which are in contact with the electrode and hence travelling wave is observed rising from the contact point. The positive electric potential attracted the Br^- ion to the surface of the beads and compared to the surface, the center remains deprived of it. Thus global oscillations start from the center of the beads.

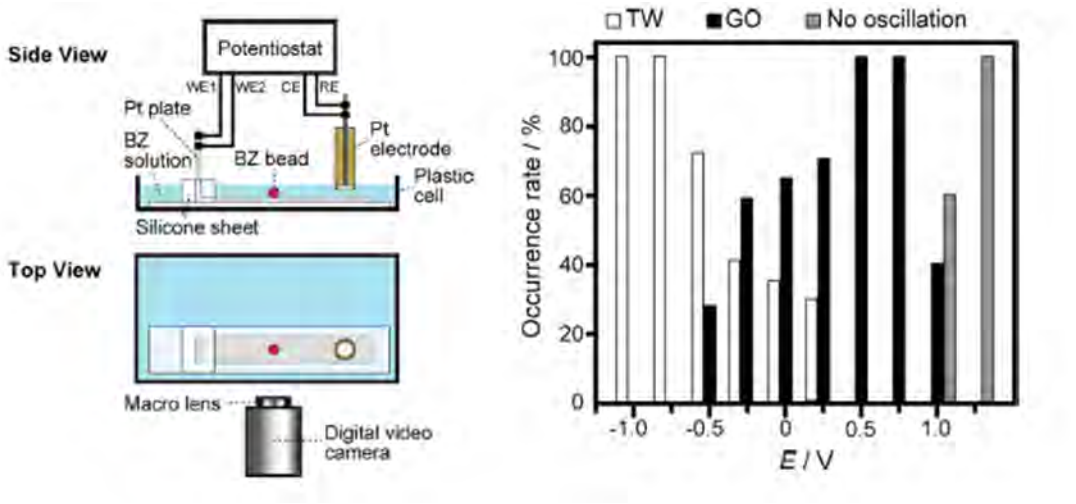


Figure 6.5: Schematic representation of the experimental setup (left). In the right, the occurrence rate of different types of oscillation with respect to electric potential (E) has been shown. Travelling waves (TW), global oscillations (GO) and no oscillation condition are depicted by white, black and grey bars respectively. [111]

6.4 My experimental results

Designing of experiment with B-Z oscillators

I used cation-exchange resin beads to design B-Z system. I used the recipe described by Kuze et al. [112]. I loaded 0.2 g of cation-exchange resin (DOWEX 50W-X4) beads with 1.1×10^{-3} M aqueous ferroin solution ($78\mu l$ of ferroin catalyst + $1.92ml$ of water) under gentle stirring at room temperature. After one hour, it was seen that the solution was colourless and the ferroin was transferred to each bead. After a few hours, the ferroin was distributed to the surface of the beads. Within the next 48 hours, the beads were fully loaded. The fully loaded beads were put in a catalyst free BZ solution. I prepared 2 ml of B-Z solution using the following recipe:

$$[H_2SO_4] = 0.6M$$

$$[NaBrO_3] = 0.35M$$

$$[CH_2(COOH)_2] = 0.35M$$

$$[NaBr] = 0.06M$$

Confined oscillations were observed within the beads. I mainly investigated if the size of the DOWEX beads can have some effect in oscillations. I also noted the results on coupling two different sizes of beads and investigated how the distance between them

plays a role.

I considered a larger bead(D_L) and another bead which was relatively smaller(D_S). Both consisted of ferroin catalyst and were immersed in B-Z solution at the same time. It was observed that D_S almost instantaneously started to show global oscillations, whereas D_L was exhibiting noisy oscillations in the time interval from 0 to 2600 s and then finally started showing stable global oscillations(Figure 6.6). These oscillations started from the center of the bead and spread towards the edge. This reason can be understood as the beads consist of sulfonate ion and it repels the production of Br^- ions which cause the oscillations to start from the center. As D_S has a smaller surface area, the excitation wave was quick to reach the surface edge and global oscillations were almost instantaneously witnessed.

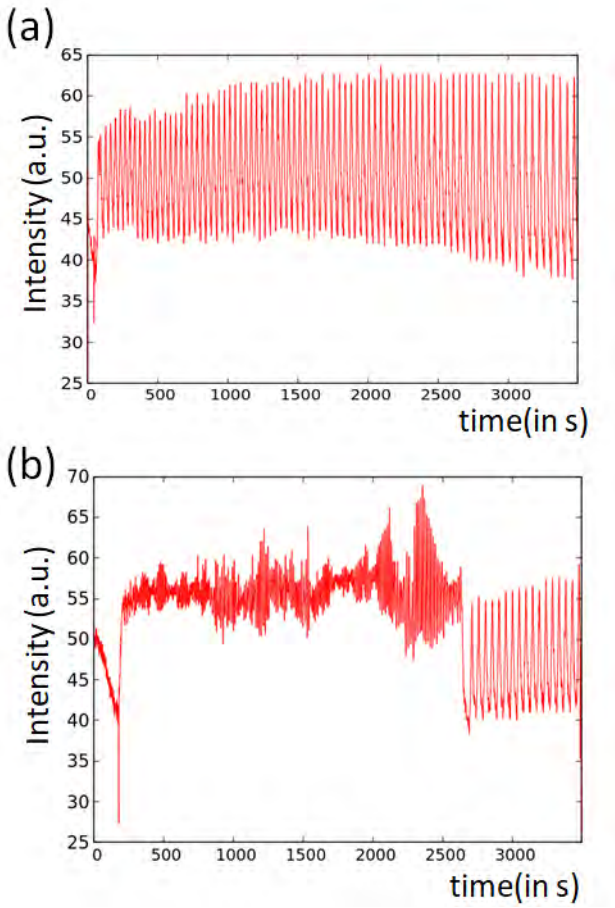


Figure 6.6: The intensity (a.u.) of blue light (in the recorded movie of the experiment) vs. time graph is plotted for oscillators (a) D_S and (b) D_L

Experimental study of coupling between different sizes of B-Z oscillators

All the above mentioned works were done by considering coupling between two oscillators of almost similar sizes. I investigated the dependence of distance in coupling among these two beads(D_L and D_S) which were significantly of different sizes. The main aim behind it was to see if the coupling between different sizes of oscillators influences their interactions or not. In Figure 6.7 the oscillations of both D_L and D_S were noted for 1^{st} 1 min, after the oscillators were coupled with distance $d_{\#i}$. It was observed that irrespective of the two different sizes of the beads used here for coupling, the results I got were similar to what we would expect if similar sizes of oscillators were coupled. The average time difference between the oscillations of D_L and D_S decreased from 3 s to 1 s with the change in distance from d_1 to d_2. The coupling strength between the oscillators increases with a decrease in the distance($d_{\#i}$) among them, resulting in strong synchronization. This phenomenon can be understood as follows: with oxidation of ferroin, $HBrO_2$ molecules are produced which diffuses from the ferroin loaded beads to the bulk phase. When distance $d_{\#i}$ between two beads are short, the local concentration of $HBrO_2$ around the beads is high. This triggers the oxidation of ferroin on the other bead . This whole phenomena may induce in-phase synchronization [110]. The time taken for the process for $HBrO_2$ to diffuse from the edge of one bead to the other can be explained by Eq.6.4.1

$$t_c = \frac{(d_{\#i})^2}{c} \tag{6.4.1}$$

where, t_c is the time required for diffusion, c is the diffusion coefficient. Thus with decrease in distance between the beads, t_c will be less and quick synchronization will be witnessed [113].

The studies related to coupling create a future prospect in creating a neural network with B-Z oscillators. Moreover, in order to teach the neural networks, we can use illumination control factor to have our domination over the oscillations within the beads.

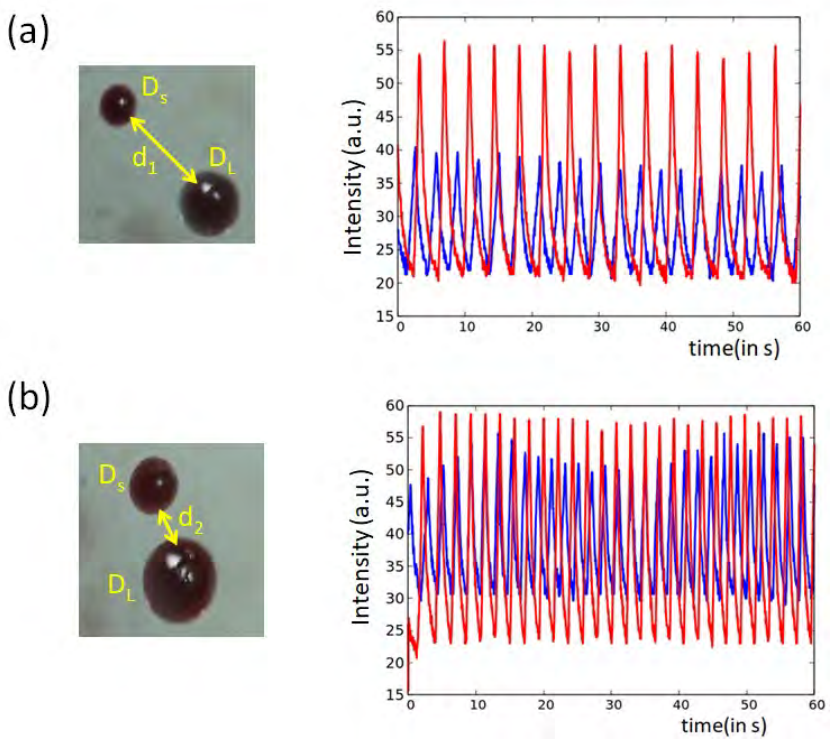


Figure 6.7: Oscillations of coupled oscillators D_S and D_L at distance (a) d_1 (b) d_2 ($d_1 > d_2$) has been shown. The blue and red colour corresponds to oscillator D_L and D_S respectively.

6.5 Conclusions

I have given an overview of how the experiments are conducted to study B-Z oscillators. I have further mentioned some noteworthy work performed in studying coupling between oscillators and the methods to control oscillations by illumination. I performed preliminary experiments in studying the effect of coupling with two different sizes of DOWEX beads. Coupling strength between the oscillators increases with a decrease in the distance among them.

All the experiments done by DOWEX beads produce promising results in designing a chemical computer. The future work of this experimental designing of chemical computers is being continued by F. Muzika and J. Gorecki [114].

Chapter 7

Conclusions and discussions

The final chapter of this thesis concludes the overall findings of the work done.

The research work embodied in this thesis has presented the idea of designing a chemical computer by networks of interacting chemical oscillators with B-Z reaction as its chemical medium. My work focuses on a top-down strategy for training the system and shows how this strategy has been helpful to solve complex information processing.

I used a realistic two-variable Oregonator model to design the networks of chemical oscillators. We think that networks of interacting chemical oscillators represent more realistic models of biological neural computing than typical artificial neural networks with arbitrarily selected activity rules [85].

First and foremost, an interacting network of chemical oscillators was designed by modifying the already existing two-variable Oregonator model. The coupling factor β was able to establish interactions among the oscillators by the change in concentration of the activators. However, instantaneous synchronization among the oscillators was observed. The introduction of rate decay constant α successfully forbids the system from having instantaneous synchronization. In short, the factor α proved to be a useful one to control the regime of coupling (weak, moderate or strong). After successfully designing the model, I presented the idea of teaching the system and explained how we have used evolutionary algorithms to fulfill this purpose.

We trained the system to solve two geometrically oriented problems. The first one we considered, was an image recognition problem. We trained our system to visualize a Japan flag. A classical image visualization problem is based on multilayered neural networks in which the output of an artificial neuron is a single number [85]. In our approach, the time evolution of a neuron (oscillator) is more complex. The output of a separated artificial neuron remains constant, whereas, in our medium, the activator concentration in a non-interacting oscillator periodically changes in time. The output also depends on the time t_{max} within which it was observed. It seems that to determine the colour of a point on the Japan flag with a high accuracy, using a standard neural network, one needs more nodes than the number of oscillators used by our medium [115]. However, just three interacting chemical oscillators were able to visualize the Japan flag with 95% accuracy. This confirms the fact that even a very simple network of chemical oscillators can solve complex information processing tasks. The system visualized the sun area of the Japan flag as some "horn" shaped. The problem of finding if a point is located inside a horn-shaped area is more complex than the determination of point location in the sun area of a Japan flag. However, the 3-interacting chemical oscillators could solve this with 99% of accuracy.

The next geometrically oriented problem we considered, was a problem to train the system to tell two spirals apart. System 1 solved this complex information processing

task with 82% of accuracy in the training dataset. I further investigated the effect of different interactions between the oscillators on the accuracy of solving the above mentioned problem. I found that establishing less complex interactions among the oscillators can be a better candidate to train for this particular case. I also introduced a test dataset to check the quality of the optimized network. I found that all three of the classifiers could distinguish if the test points belong to spiral 1 or spiral 2 with greater than 80% of accuracy.

One of the most significant challenges of civilization is how to use Artificial Intelligence (AI) for various life-inspired problems [85]. AI techniques can be beneficial for medical applications where the knowledge is accumulated as information on previously cured cases. To diagnose a new patient, one should search for similarities with the previous ones. Here I presented applications of AI methods for designing a chemical system that can help to (a) diagnose schizophrenia and (b) predict the effectiveness of bortezomib drug on patients with multiple myeloma.

In diagnosing schizophrenia : we assumed that schizophrenia could be detected by a chemical oscillator network that analyses EEG signals recorded from electrodes located on a patient scalp. The chemical system was able to diagnose schizophrenia with 82% accuracy. We successfully increased the accuracy of detection to 90% by using three different optimized systems for analyzing the EEG signal in parts and then taking the decision by voting. Although the presented results are encouraging, the access to data for the larger number of patients seems important for further studies. We think our results can be affected by the small size of available data. It can be seen when we divided the recorded potentials into three time sub-intervals and optimized the classifiers separately for each sub-interval. Seeing typical randomness of recorded signal, it is hard to expect that results measured for the first 20 s sub-interval are qualitatively different than those for the next one. Nevertheless, we obtained significantly different networks for the diagnosis in different subintervals. It can happen that there are many local maxima of optimization, but we believe that the difference in our optimization comes from a small sample of test cases. We believe, that for unbiased set of data the results of each 20 s interval belonging to healthy and ill patients should be similar, because the signals should be independent of the time when they are measured. Therefore, the difference between classifiers optimized for different time intervals illustrates the randomness in the considered dataset. We further considered three additional information processing networks characterized by different numbers of nodes. Even smaller node classifiers could produce accuracy of determination of schizophrenic patients greater than 70%. The fact that just a few oscillators can perform a complex information processing function, confirmed by the

results for schizophrenia diagnostics reports, opens the door for experimental realization of chemical instant machines with systems of interacting oscillators reported in the literature [116] [117] [118] [119] [106]. To further increase the diagnostic accuracy, we proposed the idea to choose best three networks of the lot based on accuracy and then do voting among the three to determine if the patient is schizophrenic or not. On taking a conjoint decision of the networks, we saw the detection accuracy increased to 87%, which is a 5% increase of accuracy if we consider only the decision of the single best network. It is anticipated that the accuracy of diagnosis should improve if the information on signals recorded on more than two electrodes is included in the input. The presented optimization algorithm can be easily modified to do this if one includes input oscillators of any important signal into the network. If additional signals do not increase the *Fitness*, then networks with the inputs of irrelevant signals will vanish from the population. After a successful classifier optimization is completed, its application does not require significant computing power. There are just two steps of the algorithm: 1) normalization of patient data with parameters (μ, σ) obtained for the training dataset and 2) numerical solution of differential Eq 3.3.7 and 3.3.8 and activator maxima counting. A modern laptop needs a few seconds to execute these tasks. The whole procedure can be incorporated into EEG equipment software or distributed as a laptop or smartphone application.

To predict the drug response on patients with multiple myeloma, we used a network with 3 interacting oscillators. We gave a gene expression value as input to the system. We obtained 15 optimized networks for 15 gene expression values. Individual classifiers gave the accuracy in the range between 66.5% to 71.1%. To improve the determination of the therapy outcome, we considered the concilium of 15 classifiers and accepted the majority decision. Such a strategy increased the accuracy to almost 85%, which seems to be a promising result for the further development of the method. We can suggest two ways in which the accuracy in the determination of therapy effectiveness can be increased: The first one is to consider the voting strategy with more complex networks formed by a larger number of oscillators that are used to determine correlations between a single gene expression value and the therapy outcome. One can expect that "wiser" members of concilium can produce more accurate answers. However, the strategy of employing top specialists does not guarantee top results. Our simulations have shown that the synergy between concilium members is also important and should be taken into account. We continued optimization for some networks processing gene expression values and obtained higher accuracy than what we got before. However, if we replace the optimized network for the gene SERP1 that gives the accuracy of 66.5% by a wiser member (accuracy 67.7%,

$t_{max} = 80$, $t_{start} = 3.52$, $t_{end} = 32.23$, $\alpha = 0.71$, $\beta = 0.09$, normal oscillators 2 and 3, $t^2_{illum} = 3.38$, $t^3_{illum} = 4.24$) then the accuracy of concilium decreases to 83.6%. Alternatively, one can consider networks that are processing expression values for more than a single gene. However, trying a few pairs of gene expression values as inputs (with 1000 generations, 200 population size and a network with 6 oscillators) could not produce high accuracy. Therefore, it can be anticipated that a larger oscillator network is necessary for the data classification and its optimization will be numerically complex. In our opinion, it will be worthwhile to proceed with the approaches mentioned above when access to a larger training dataset is provided.

The presented classification methods (in Chapter 3, 4 and 5) are based on many assumptions. All of them can be lifted in search of the best network for a certain information processing task. The optimization of interactions in the medium can be directly included in the optimization program. The presented results were obtained assuming that interactions between oscillators were fixed. The information about interactions was included in equations describing the time evolution of oscillators as the binary parameters $s_{k,j}$ in Eq. (3.3.7). The values of these parameters can be included in classifier optimization. The network model includes the activatory coupling between oscillators. It means that an excited oscillator can speed up the excitations of the other oscillators that interact with it. Such coupling is observed, for example, in droplets containing reagents of B-Z reaction. Alternatively, one can consider a medium with inhibitory coupling where excitation of one oscillator slows down the activity of those oscillators that interact with it [117–119]. Allowing for different types of coupling within a single network can help to identify the best medium for a given computing task.

Results presented in the Chapter 4 and 5 of the thesis were obtained based on computer simulations, but the Oregonator model can quantitatively describe B-Z reaction, so it brings information for potential experiments on the chemical computation. Systems of interacting oscillators have been studied experimentally using a few techniques [120], [121], [122], [34]. Interacting droplets containing reagents of B-Z reaction can be stabilized by a solution of lipids in the organic phase [123]. If the photosensitive variant of B-Z reaction is used, then oscillations in droplets can be individually controlled. Experiments on the control of 3 coupled droplets mechanically stabilized inside a plastic cage were reported in [106]. However, the experiments have demonstrated [106] that it is rather difficult to stabilize even 3 droplets if a standard variant of B-Z reaction with the malonic acid is used. The bubbles of gas can appear between droplets, deform them, and change interactions between oscillators. Therefore, we believe that solid objects loaded with the catalyst, like DOWEX beads or

silica gel beads, seem to be more suitable for experiments with information processing using a network of oscillators. I performed preliminary experiments to notice coupling between two different sizes of DOWEX beads. It was found that the coupling strength between the oscillators increases with a decrease in distance. The finding was similar as to what we see if two similar sizes of oscillators are coupled.

To sum up the conclusions, it can be said that a network of chemical oscillators can be trained to perform complex information processing tasks (like geometrically and medically oriented problems) easily. The results are promising towards designing an unconventional chemical computer.

List of publications

[1] Ashmita Bose and Jerzy Gorecki (2020), **Determination of Psychotic Behaviour Using a Network of Chemical Oscillators.** pp 67-71 in Sergey Y. Yurish, Editor *Advances in Signal Processing and Artificial Intelligence* ASPAI' 2020 Conference Proceedings Publisher: International Frequency Sensor Association (IFSA) Publishing, S. L. ISBN: 978-84-09-21931-5 e-ISBN: 978-84-09-21930-8 BN-20200705-XX https://www.sensorsportal.com/

[2] Jerzy Gorecki and Ashmita Bose (2020), **How Does a Simple Network of Chemical Oscillators See the Japanese Flag?** *Frontiers in Chemistry* Vol 8. doi: 10.3389/fchem. 2020. 580703

[3] Ashmita Bose and Jerzy Gorecki (2021), **Can A Network of Chemical Oscillators Help to Diagnose Schizophrenia?** *International Journal of Unconventional Computing*,Vol 16 , pp 1-18.

[4] Ashmita Bose and Jerzy Gorecki (2022), **Computing With Networks of Chemical Oscillators and its Application for Schizophrenia Diagnosis.** *Frontiers in Chemistry*. Vol 10. doi: 10.3389/fchem.2022.848685.

[5] Ashmita Bose, Peter Dittrich and Jerzy Gorecki. **The concilium of information processing chemical networks for determining Bortezomib drug response in patients with multiple myeloma.** (submitted)

Bibliography

[1] J. Andreu-Perez, F. Deligianni, D. Ravi, and G.-Z. Yang, "Artificial Intelligence and Robotics," *arXiv preprint arXiv:1803.10813*, vol. 147, 2018.

[2] C. Vocke, C. Constantinescu, and D. Popescu, "Application potentials of artificial intelligence for the design of innovation processes," *Procedia CIRP*, vol. 84, pp. 810–813, 2019.

[3] H. AI, "High-level expert group on artificial intelligence," *Ethics guidelines for trustworthy AI*, 2019.

[4] K. Das and R. N. Behera, "A Survey on Machine Learning: Concept, Algorithms and Applications," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 5, pp. 1301–1309, 2017.

[5] E. Sánchez-DelaCruz and D. Lara Alabazares, "Deep learning: concepts and implementation tools," in *LANMR*, pp. 142–149, 2019.

[6] `https://www.javatpoint.com/subsets-of-ai`.

[7] R. J. Lewis, "An introduction to classification and regression tree (CART) analysis," in *Annual meeting of the society for academic emergency medicine in San Francisco, California*, vol. 14, Citeseer, 2000.

[8] C.-Y. J. Peng, K. L. Lee, and G. M. Ingersoll, "An Introduction to Logistic Regression Analysis and Reporting," *The Journal of Educational Research*, vol. 96, pp. 3–14, 2002.

[9] R. Sarmento and V. Costa, *Introduction to Linear Regression*. IGI Global, 2017.

[10] `https://levity.ai/blog/difference-machine-learning-deep-learning`.

[11] M. Shah, "Dendrites," in *Encyclopedia of the Neurological Sciences (Second Edition)* (M. J. Aminoff and R. B. Daroff, eds.), p. 970, Oxford: Academic Press, second edition ed., 2014.

[12] D. Debanne, E. Campanac, A. Bialowas, E. Carlier, and G. Alcaraz, "Axon Physiology," *Physiological Reviews*, vol. 91, pp. 555–602, 2011.

[13] `https://www.healthline.com/health/neurons#parts`.

[14] H. S. Das and P. Roy, "Chapter 5 - A Deep Dive Into Deep Learning Techniques for Solving Spoken Language Identification Problems," in *Intelligent Speech Signal Processing* (N. Dey, ed.), pp. 81–100, Elsevier, 2019.

[15] `https://www.teco.edu/~albrecht/neuro/html/node7.html`.

[16] J. J. Torres and P. Varona, *Modeling Biological Neural Networks*, pp. 533–564. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

[17] M. Hassan and D. Q. Rizvi, "Computer vs human brain: An analytical approach and overview," *Computer*, pp. 580–583, 2019.

[18] M. Oltean, "Unconventional computing: A short introduction," *Studia Universitatis Babes-Bolyai : Series Informatica*, vol. 54, 07 2009.

[19] Y. Faroukh, "Quantum Computers Vs Conventional Computers: A Study on the Larger Scale," 2018. doi:10.13140/RG.2.2.29502.18243.

[20] S. Hassan, M. Asghar, *et al.*, "Limitation of silicon based computation and future prospects," in *2010 Second International Conference on Communication Software and Networks*, pp. 559–561, IEEE, 2010.

[21] G. Moore, "Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp.114 ff," *Solid-State Circuits Newsletter, IEEE*, vol. 11, pp. 33 – 35, 2006.

[22] K. Gizynski and J. Gorecki, "Cancer classification with a network of chemical oscillators," *Physical Chemistry Chemical Physics*, vol. 19, pp. 28808–28819, 2017.

[23] H. Smith, "Plants that track the Sun." `https://www.nature.com/articles/308774a0.pdf?origin=ppub`, 1984.

[24] "`https://www.scientificamerican.com/article/when-it-comes-to-photosynthesis-plants-perform-quantum-computation/`."

[25] H. Hamann and H. Wörn, "Embodied Computation," *Parallel Processing Letters*, vol. 17, pp. 287–298, 2007.

[26] B. J. MacLennan, "Mapping the Territory of Computation Including Embodied Computation," in *Handbook of Unconventional Computing: VOLUME 1: Theory*, pp. 1–30, World Scientific, 2022.

[27] "`https://www.mathworks.com/help/gads/what-is-simulated-annealing.html`."

[28] Y. Eren, İbrahim B. Küçükdemiral, and İlker Üstoğlu, "Chapter 2 - Introduction to Optimization," in *Optimization in Renewable Energy Systems* (O. Erdinç, ed.), pp. 27–74, Boston: Butterworth-Heinemann, 2017.

[29] S. Jain, "Quantum computer architectures: A survey," in *2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom)*, pp. 2165–2169, 2015.

[30] C. A. Coello Coello, "An introduction to evolutionary algorithms and their applications," in *International Symposium and School on Advancex Distributed Systems*, pp. 425–442, Springer, 2005.

[31] M. Amira, A. Abd El-Aziz, A. Hesham, and A. Ahmed, "Swarm Optimization Techniques: A survey," *Egyptian Computer Science Journal*, vol. 42, 2018.

[32] `https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/Comparison/comparison.html`.

[33] K. Gizynski and J. Gorecki, "A Chemical System that Recognizes the Shape of a Sphere," *Computational Methods in Science and Technology*, vol. 22, pp. 167–177, 2016.

[34] P. L. Gentili, M. S. Giubila, R. Germani, A. Romani, A. Nicoziani, A. Spalletti, and B. M. Heron, " Optical Communication among Oscillatory Reactions and Photo-Excitable Systems: UV and Visible Radiation Can Synchronize Artificial Neuron Models.," *Models. Angewandte Chemie (International ed. in English)*, vol. 56, pp. 7535–7540, 2017.

[35] A. Adamatzky, B. D. L. Costello, and T. Asai, *Reaction-Diffusion Computers*. USA: Elsevier Science Inc., 2005.

[36] "https://www.geeksforgeeks.org/difference-between-bottom-up-model-and-top-down-model/."

[37] R. J. Field and R. M. Noyes, "Oscillations in chemical systems. IV. Limit cycle behavior in a model of a real chemical reaction," *Journal of Chemical Physics*, vol. 60, pp. 1877–1884, 1974.

[38] D. Hill and T. Morgan, "Pattern Formation and Wave Propagation in the Belousov-Zhabotinskii Reaction," *University of California, San Diego–Physics Department. Neurophysics, ucsd. edu/courses/physics_173_273/bz_paper. PDF. Accessed*, vol. 10, 2020.

[39] S. A. Levin, E. Perlin, and C. Levin, "Frontiers in mathematical biology," pp. 569–584, Springer, 1994.

[40] B. Z. Shakhashiri, *Chemical Demonstrations: A Handbook for Teachers of Chemistry*, vol. 2. Univ of Wisconsin Press, 1985.

[41] C. Egami, "Mechanism for the color transition of the Belousov-Zhabotinsky reaction catalyzed by cerium ions and ferroin," *Discrete Continuous Dynamical Systems - B*, vol. 23, pp. 2527–2544, 2018.

[42] L. Adamčíková and P. Ševčíak, "A completely inorganic oscillating system of the Belousov–Zhabotinskii type," *International Journal of Chemical Kinetics*, vol. 14, pp. 735–738, 1982.

[43] A. Cassani, A. Monteverde Videla, and M. Piumetti, "Belousov-Zhabotinsky type reactions: the non-linear behavior of chemical systems," *Journal of Mathematical Chemistry*, vol. 59, pp. 792–826, 2021.

[44] I. Barzykina, "Chemistry and Mathematics of the Belousov–Zhabotinsky Reaction in a School Laboratory," *Journal of Chemical Education*, vol. 97, pp. 1895–1902, 2020.

[45] J. Miyazaki, "2 - Belousov–Zhabotinsky Reaction," in *Pattern Formations and Oscillatory Phenomena* (S. Kinoshita, ed.), pp. 61–83, Boston: Elsevier, 2013.

[46] A. M. Zhabotinsky, "Belousov-Zhabotinsky reaction," *Scholarpedia*, vol. 2, p. 1435, 2007. revision #91050.

[47] "`http://rcin.org.pl/Content/51549/PDF/WA333_70884_F-B468-15_Gizynski.pdf`."

[48] S. Kádár, T. Amemiya, and K. Showalter, "Reaction Mechanism for Light Sensitivity of the $Ru(bpy)_3^{2+}$-Catalyzed BelousovZhabotinsky Reaction," *The Journal of Physical Chemistry A*, vol. 101, pp. 8200–8206, 1997.

[49] P. K. Srivastava, Y. Mori, and I. Hanazaki, "Photo-inhibition of chemical oscillation in the $Ru(bpy)_3^{2+}$-catalyzed Belousov-Zhabotinskii reaction," *Chemical Physics Letters*, vol. 190, pp. 279–284, 1992.

[50] L. Kuhnert, K. Agladze, and V. Krinsky, " Image processing using light-sensitive chemical waves," *Nature*, vol. 337, pp. 244–247, 1989.

[51] R. J. Field, "An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos (Epstein, I. R.; Pojman, J. A.)," *Journal of Chemical Education*, vol. 77, p. 450, 2000.

[52] A. Panvilov, K. Tusscher, and R. J. de Boer, "Matrices, Linearization, and the Jacobi matrix," *Theoretical Biology, Utrecht University*, 2021.

[53] "`http://alun.math.ncsu.edu/wp-content/uploads/sites/2/2017/01/linearization.pdf`."

[54] E. Zemskov, K. Kassner, M. Hauser, and W. Horsthemke, "Turing space in reaction-diffusion systems with density-dependent cross diffusion," *Physical Review E*, vol. 87, p. 032906, 2013.

[55] C. Gray, "An analysis of the Belousov-Zhabotinskii reaction," *Rose-Hulman Undergraduate Mathematics Journal*, vol. 3, 2002.

[56] F. Sagués and I. Epstein, "Nonlinear Chemical Dynamics," *Dalton Transactions*, vol. 34, pp. 1201–1217, 2003.

[57] J. Gorecki, J. Gorecka, and Y. Igarashi, "Information processing with structured excitable medium," *Natural Computing*, vol. 8, pp. 473–492, 2009.

[58] J. Murray, *Mathematical biology*. Springer, Berlin, Germany, 1989.

[59] Krug, Hans Juergen. and Pohlmann, Ludwig. and Kuhnert, Lothar., "Analysis of the modified complete Oregonator accounting for oxygen sensitivity and photosensitivity of Belousov-Zhabotinskii systems," *The Journal of Physical Chemistry*, vol. 94, pp. 4862–4866, 1990.

[60] J. J. Tyson, "Scaling and reducing the Field-Koros-Noyes mechanism of the Belousov-Zhabotinskii reaction," *The Journal of Physical Chemistry*, vol. 86, pp. 3006–3012, 1982.

[61] R. Toth and A. Taylor, "The Tris(2,2'-Bipyridyl)Ruthenium-Catalysed Belousov–Zhabotinsky Reaction," *Progress in Reaction Kinetics and Mechanism*, vol. 31, pp. 59–115, 2006.

[62] R. M. Noyes, R. Field, and E. Koros, "Oscillations in chemical systems. I. Detailed mechanism in a system showing temporal oscillations," *Journal of the American Chemical Society*, vol. 94, pp. 1394–1395, 1972.

[63] A. Toth, V. Gaspar, and K. Showalter, "Signal transmission in chemical systems: propagation of chemical waves through capillary tubes," *The Journal of Physical Chemistry*, vol. 98, pp. 522–531, 1994.

[64] Tóth and K. Showalter, "Logic gates in excitable media," *The Journal of Chemical Physics*, vol. 103, pp. 2058–2066, 1995.

[65] K. Agladze, R. Aliev, and T. Yamaguchi, "Chemical Diode," *The Journal of Physical Chemistry*, vol. 100, pp. 13895–13897, 1996.

[66] K. Agladze, A. Toth, T. Ichino, and K. Yoshikawa, "Propagation of Chemical Waves at the Boundary of Excitable and Inhibitory Fields," *Journal of Physical Chemistry A*, vol. 104, pp. 6677–6680, 2000.

[67] I. N. Motoike, K. Yoshikawa, Y. Iguchi, and S. Nakata, "Real-time memory on an excitable field," *Physical Review E*, vol. 63, p. 036220, Feb 2001.

[68] M.-A. Tsompanas, C. Fullarton, and A. Adamatzky, "Belousov-Zhabotinsky liquid marbles in robot control," *Sensors and Actuators B: Chemical*, vol. 295, pp. 194–203, 2019.

[69] A. Adamatzky, N. Phillips, R. Weerasekera, M.-A. Tsompanas, and G. C. Sirakoulis, "Excitable London: Street map analysis with Oregonator model," *arXiv preprint arXiv:1803.01632*, 2018.

[70] M. Borda, *Fundamentals in Information Theory and Coding*. Springer Berlin Heidelberg, 2011.

[71] R. Togneri and C. J. S. DeSilva, *Fundamentals of Information Theory and Coding Design*. USA: CRC Press, Inc., 2003.

[72] M. Popovic, "Research in entropy wonterland: A review of the entropy concept," *Thermal Science*, vol. 22, pp. 1163–1178, 2018.

[73] S. Vajapeyam, "Understanding Shannon's entropy metric for information," *arXiv preprint arXiv:1405.2061*, 2014.

[74] E. G. Learned-Miller, "Entropy and Mutual Information," *Department of Computer Science, University of Massachusetts, Amherst*, 2009.

[75] G. Zeng, "A Unified Definition of Mutual Information with Applications in Machine Learning," *Mathematical Problems in Engineering*, vol. 2015, 2015.

[76] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley, 2012.

[77] R. Dilao, "From Charles Darwin to Evolutionary Genetic Algorithms," *Memóricas Da Academia das Ciências de Lisboa*, pp. 1–11, 2009.

[78] F. Neumann and C. Witt, *Stochastic Search Algorithms*, pp. 21–32. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010.

[79] T. Bäck, G. Rudolph, and H.-P. Schwefel, "Evolutionary programming and evolution strategies: Similarities and differences," in *In Proceedings of the Second Annual Conference on Evolutionary Programming*, pp. 11–22, Citeseer, 1993.

[80] S. Rahnamayan, H. R. Tizhoosh, and M. M. A. Salama, "Opposition-Based Differential Evolution Algorithms," *2006 IEEE International Conference on Evolutionary Computation*, pp. 2010–2017, 2006.

[81] H.-G. Beyer and H.-P. Schwefel, "Evolution strategies - A comprehensive introduction," *Natural Computing*, vol. 1, pp. 3–52, 2002.

[82] P. Vikhar, "Evolutionary algorithms: A critical review and its future prospects," in *International Conference on Global Trends in Signal Processing, Information Computing and Communication*, pp. 261–265, 2016.

[83] "`https://www.analyticsvidhya.com/blog/2021/03/basics-of-neural-network/`."

[84] "`https://www.youtube.com/watch?v=uXt8qF2Zzfo`."

[85] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.

[86] W. T. V. William H. Press, Saul A. Teukolsky and B. P.Flannery, *Numerical recipes in C ,the art of scientific computing, 2nd edn.* USA: Cambridge University Press, 1992.

[87] A. Bose and J. Gorecki, "Can a Network of Chemical Oscillators Help to Diagnose Schizophrenia?," *International Journal of Unconventional Computing*, vol. 16, pp. 1–18, 2021.

[88] J. Gorecki and A. Bose, "How Does a Simple Network of Chemical Oscillators See the Japanese Flag?," *Frontiers in Chemistry*, vol. 8, 2020.

[89] K. Lang, "Learning to tell two spiral apart," in *Proceedings of the 1988 connectionist models summer school*, pp. 52–59, 1989.

[90] S. K. Chalup and L. Wiklendt, "Variations of the Two-Spiral Task," *Connection Science*, vol. 19, pp. 183–199, 2007.

[91] J. R. Álvarez-Sánchez, "Injecting Knowledge into the Solution of the Two-Spiral Problem," *Neural Computing & Applications*, vol. 8, pp. 265–272, 1999.

[92] D. Corring, "Exploring schizophrenia within the framework of the International Classification of Functioning, Disability and Health (ICF)," *Research Insights of the Regional Mental Health Care, London/St. Thomas.*, vol. 3, 2005.

[93] "`http://www.chovil.com/`."

[94] P. C. Tiwari, "Review on Schizophrenia entitled"Schizophrenia- A New Insight"," *International Journal of Research*, 2012.

[95] P. F. Liddle, *Introduction to schizophrenia*, pp. 3–17. Basel: Birkhäuser Basel, 2000.

[96] D. B'en, A. Kaplan, S. Timashev, G. Vstovskiĭ, and B. Pak, "Variability of the EEG autocorrelation structure in adolescents with schizophrenia spectrum disorders," *Fiziologiia cheloveka*, vol. 33, pp. 138–40, 2007.

[97] A. Bose and J. Gorecki, "Computing With Networks of Chemical Oscillators and its Application for Schizophrenia Diagnosis," *Frontiers in Chemistry*, vol. 10, 2022.

[98] "`http://brain.bio.msu.ru/eegschizophrenia.htm`."

[99] G. Gruenert, K. Gizynski, G. Escuela, B. Ibrahim, and P. Gorecki, J. Dittrich, "Understanding Networks of Computing Chemical Droplet Neurons Based on Information Flow," *International Journal of Neural Systems*, vol. 25, p. 1450032, 2015.

[100] G. Mulligan, C. Mitsiades, B. Bryant, F. Zhan, W. J. Chng, S. Roels, E. Koenig, A. Fergus, Y. Huang, P. Richardson, W. L. Trepicchio, A. Broyl, P. Sonneveld, J. Shaughnessy, John D., P. Leif Bergsagel, D. Schenkein, D.-L. Esseltine, A. Boral, and K. C. Anderson, "Gene expression profiling and correlation with outcome in clinical trials of the proteasome inhibitor bortezomib," *Blood*, vol. 109, pp. 3177–3188, 2006.

[101] A. Field-Smith, G. J. Morgan, and F. E. Davies, "Bortezomib (Velcade™) in the treatment of multiple myeloma," *Therapeutics and Clinical Risk Management*, vol. 2, p. 271, 2006.

[102] T. Hideshima, P. L. Bergsagel, W. M. Kuehl, and K. C. Anderson, "Advances in biology of multiple myeloma: clinical applications," *Blood*, vol. 104, pp. 607–618, 2004.

[103] L. J. Lesko and J. Woodcock, "Translation of pharmacogenomics and pharmacogenetics: a regulatory perspective," *Nature Reviews Drug Discovery*, vol. 3, pp. 763–769, 2004.

[104] F. Zhan, Y. Huang, S. Colla, J. P. Stewart, I. Hanamura, S. Gupta, J. Epstein, S. Yaccoby, J. Sawyer, B. Burington, *et al.*, "The molecular classification of multiple myeloma," *Blood*, vol. 108, pp. 2020–2028, 2006.

[105] J. R. Cash and A. H. Karp, "A variable order Runge-Kutta method for initial value problems with rapidly varying right-hand sides," *ACM Transactions on Mathematical Software*, vol. 16, pp. 201–222, 1990.

[106] K. Gizynski and J. Gorecki, "Chemical memory with states coded in light controlled oscillations of interacting Belousov–Zhabotinsky droplets," *Physical Chemistry Chemical Physics*, vol. 19, pp. 6519–6531, 2017.

[107] T. Yamaguchi, L. Kuhnert, Z. Nagy-Ungvarai, S. C. Mueller, and B. Hess, "Gel systems for the Belousov-Zhabotinskii reaction," *The Journal of Physical Chemistry*, vol. 95, pp. 5831–5837, 1991.

[108] K. M. Chang, M. de Planque, and K.-P. Zauner, "Towards Functional Droplet Architectures: a Belousov-Zhabotinsky Medium for Networks," *Scientific Reports*, vol. 8, pp. 1–12, 2018.

[109] N. Nishiyama and K. Eto, "Experimental study on three chemical oscillators coupled with time delay," *The Journal of Chemical Physics*, vol. 100, pp. 6977–6978, 1994.

[110] M. Kuze, Y. Hiranishi, Y. Okamoto, A. Shioi, and S. Nakata, "Coupling of Two Microbeads Exhibiting Different Features of Oscillations in the Belousov-Zhabotinsky Reaction," *Chemistry Letters*, vol. 48, pp. 847–850, 2019.

[111] M. Kuze, M. Horisaka, N. J. Suematsu, T. Amemiya, O. Steinbock, and S. Nakata, "Chemical Wave Propagation in the Belousov–Zhabotinsky Reaction

Controlled by Electrical Potential," *Journal of Physical Chemistry A*, vol. 123, pp. 4853–4857, 2019.

[112] M. Kuze, H. Kitahata, O. Steinbock, and S. Nakata, "Distinguishing the Dynamic Fingerprints of Two- and Three-Dimensional Chemical Waves in Microbeads," *Journal of Physical Chemistry A*, vol. 122, pp. 1967–1971, 2018.

[113] K. Yoshikawa, R. Aihara, and K. Agladze, "Size-Dependent Belousov Zhabotinsky Oscillation in Small Beads," *Journal of Physical Chemistry A*, vol. 102, pp. 7649–7652, 1998.

[114] F. Muzika and J. Górecki, "Identification of the best medium for experiments on chemical computation with Belousov–Zhabotinsky reaction and ferroin-loaded Dowex beads," *Reaction Kinetics, Mechanisms and Catalysis*, pp. 1–23, 2022.

[115] L. Zammataro, "Solving the hole in the square problem with a neural network," *Contribution from the Wolfram Demonstration Project " solving the Hole in the Square Problem with A Neural Network"*, 2010.

[116] J. M. Parrilla-Gutierrez, A. Sharma, S. Tsuda, G. J. Cooper, A.-C. Gerardo, K. Donkers, and L. Cronin, "A programmable chemical computer with memory and pattern recognition.," *Nature Communications*, vol. 11, p. 1442, 2020.

[117] V. K. Vanag and V. O. Yasuk, "Dynamic modes in a network of five oscillators with inhibitory all-to-all pulse coupling," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, p. 033105, 2018.

[118] V. K. Vanag, ""Cognitive" modes in small networks of almost identical chemical oscillators with pulsatile inhibitory coupling," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, p. 033106, 2019.

[119] I. S. Proskurkin and V. K. Smelov, P. S.and Vanag, "Experimental Investigation of the Dynamical Modes of Four Pulse-Coupled Chemical Micro-Oscillators. ," *Chemphyschem : a European Journal of Chemical Physics and Physical Chemistry*, vol. 20, pp. 2162–2165, 2019.

[120] V. Vanag and I. Epstein, "Stationary and Oscillatory Localized Patterns, and Subcritical Bifurcations," *Physical review letters*, vol. 92, p. 128301, 2004.

[121] A. Kaminaga, V. Vanag, and I. Epstein, ""Black spots" in a surfactant-rich Belousov-Zhabotinsky reaction dispersed in a water-in-oil microemulsion system," *The Journal of Chemical Physics*, vol. 122, p. 174706, 2005.

[122] A. Kaminaga, V. Vanag, and I. Epstein, "A Reaction–Diffusion Memory Device," *Angewandte Chemie (International ed. in English)*, vol. 45, pp. 3087–9, 2006.

[123] J. Szymanski, J. Gorecka, Y. Igarashi, K. Gizynski, J. Gorecki, K.-P. Zauner, and M. de Planque, "Droplets with Information Processing Ability," *International Journal of Unconventional Computing*, vol. 7, pp. 185–200, 2011.