UNIVERSITY OF BELGRADE

FACULTY OF PHYSICS

Miroslav Anđelković

Algebraic topology of complex networks and topological aspects of nonlinear dynamical systems

Doctoral Dissertation

Belgrade, 2019

UNIVERZITET U BEOGRADU

FIZIČKI FAKULTET

Miroslav Anđelković

Algebarska topologija kompleksnih mreža i topološki aspekti nelinearnih dinamičkih sistema

Doktorska disertacija

Beograd, 2019

Podaci o mentoru i članovima komisije:

Mentor:

dr Milan Rajković
naučni savetnik, Institut za nuklearne nauke "Vinča"

Članovi komisije:

prof. dr Milan Knežević
redovni profesor, Fizički Fakultet, Univerzitet u Beogradu

prof. dr Sunčica Elezović-Hadžić
redovni profesor, Fizički Fakultet, Univerzitet u Beogradu

prof. dr Rade T. Živaljević
naučni savetnik, Matematički Institut SANU, Univerzitet u Beogradu

dr Slobodan Maletić
naučni saradnik, Institut za nuklearne nauke "Vinča"

Datum odbrane: 8. mart 2019.

Acknowledgments

I wish to express gratitude to my mentor and supervisor dr Milan Rajković for helping me with guidance and encouragement through whole process of dissertation, starting from teaching me concepts and the various skills of academic research. Also I want to thank him for going through and correcting my publications drafts, my presentations and finally my dissertation.

For big support and guidance to my colleagues dr Slobodan Maletić and prof. dr Bosiljka Tadić. Their effort had big impact on my professional development from the begging.

I wish to thank the members of my dissertation committee: prof. dr Milan Knežević, prof. dr Sunčica Elezović-Hadžić, prof. dr Rade Živaljević, and dr Slobodan Maletić, for their support and defense of the ideas presented in this dissertation.

The support of the colleagues from the Laboratory of Thermal Engineering and Energy at the Institute of nuclear sciences Vinča was not lacking, and for that and their understanding, I am grateful to them.

I am so grateful to my parents Dragan and Mirjana for believing in me my whole life and dedicating their lives to me. Last, but the most important, I dedicate this dissertation to my biggest support, my girlfriend Jovana for being strong and trustful pillar of my life, for all tolerance and patience she had all the time.

Title:

Algebraic topology of complex networks and topological aspects of nonlinear dynamical systems

Abstract

To understand and eventually predict the behavior of complex systems arising from diverse areas of science such as physics, economics or biology, which have a widespread impact on our lives, many powerful methods and models have been developed in the recent years. Research in the area of complex system bifurcates in different directions using various methods and recent years have witnessed significant increase in publications and results related to the understanding how complex systems function. The tools of algebraic topology, presented in this dissertation, have a large influence on further development of new methods for obtaining not so apparent properties which eluded well established approcahes.Simplicial complexes are becoming an important part of the modern theoretical physics frameworks by virtue of increased use of methods of algebraic topology in various areas of physics and other areas of science as well. So far the study of complex systems which are represented by a large number of mutually interacting agents is to a large extent based on the concepts and methods which arise in the graph theory (complex networks). Analysis of complex networks has resulted in the deeper insight into the idiosyncrasies of complex systems such as structure and self-organization based on, for example, the study of aggregation of agents into communities.

In this thesis, the initial focus is concerned on the structure and substructure of complex systems, through complex network study, and high order aggregations of elements of complex network. A detailed illustration of construction of simplicial complexes from either the complex networks, from the data embedded in metric space or from the time series is presented. Application and results of simplicial complexes constructed from various complex systems represented by complex networks, originating from physics, social and biological sciences, reveal hidden geometry and topology of higher-order that are not observed via standard methods of graph analysis and statistical mechanics. In the second part of the thesis a recurrent property of nonlinear dynamical systems is studied with the use of simplicial complexes and with the application of the concepts and methods from the algebraic topology. Results for several benchmark systems are presented in order to show the scope of the topological framework and topological properties that can be obtained from this

analysis. Different models of systems which exhibit transition to deterministic chaos were studied and topological properties of phase space were recognized as generic indicators of dynamic regime changes. The applications of topology, Q-analysis and persistent homology in particular, to nonlinear dynamical systems open up new possibilities for further studies involving the interaction of these disciplines.

Keywords:
statistical mechanics, complex systems, graph, complex networks, combinatorial algebraic topology, simplicial complexes, topological invariant, combinatorial Laplacian, entropy, nonlinear dynamical systems, chaos

Scientific field:
physics

Specific scientific field:
statistical physics

UDC number:533.9(043.3)

Naslov:

Algebarska topologija kompleksnih mreža i topološki aspekti nelinearnih dinamičkih sistema

Rezime

Da bi razumeli i eventualno predvideli ponašanje kompleksnih sistema koji se javljaju u raznim oblastima nauke, od socio-ekonomskih do sistema iz, na primer, fizike ili biologije, i koji imaju važan uticaj na razne aspekte naših života, naučnici su razvili veliki broj metoda i modela. Istraživanja kompleksnih sistema se odvijaju u različitim pravcima analize i modeliranja i u poslednjih nekoliko godina dovela su do naglog porasta broja rezultata i publikacija koje su proširile perspektive sveta oko nas. Alati koji proističu iz algebarske topologije, predstavljeni u ovoj tezi, imali su veliki uticaj kao sredstvo za istraživanje i otkrivanje, ne tako očiglednih osobina, koje ranije nisu bile poznate. Simplicijalni kompleksi postaju važan predmet studija u modernoj teorijskoj fizici zahvaljujući potencijalu primene koncepata iz algebarske topologije u raznim oblastima nauke. Dosadašnja proučavanja kompleksnih sistema (u fizici i matematici) predstavljenih preko velikog broja elemenata koji uzajamno komuniciraju putem uparenih odnosa, u velikoj meri su bazirana na konceptima i metodama koji potiču iz teorije grafova, odnosno kompleksnih mreža. Analiza kompleksnih mreža je dovela do boljeg uvida u neke od osobina kompleksnih sistema kao što su njihova struktura i samoorganizacija, proučavajući, izmedju ostalog, mehanizme formiranja zajednica od interagujućih elemenata kompleksnog sistema.

U ovoj tezi, početni fokus se odnosi na proučavanje struktura i podstruktura nekolicine kompleksnih sistema putem primene topoloških osobina simplicijlanih kompleksa konstruisanih iz kompleksnih mreža i kroz agregaciju mrežnih elemenata na višim dimenzijama kompleksne mreže. Predstavljeni su načini na koji se simplicijalni kompleksi mogu izgraditi direktno iz kompleksnih mreža od podataka ugradjenih u metrički prostor ili pak, iz vremenskih serija. Primena i rezultati ovih metoda na različitim kompleksnim sistemima koji proističu iz različitih oblasti kao što su fizika, društvene i biološke nauke, otkrivaju skrivenu geometriju i topologiju na višim dimenzijama simplicijalnih kompleksa koja se ne primećuje primenom standardnih metoda koje kombinuju teoriju grafova sa statističkom fizikom, odnosno statističkom mehanikom. U drugom delu teze uvedene su metode za proučavanje rekurentnih osobina nelinearnih dinamičkih sistema (pojava koje se ponavljaju u faznom prostoru) koristeći algebarsku topologiju i multifraktalne osobine atraktora

u faznom prostoru. Izučavano je nekoliko standardnih nelinearnih dinamičkih sistema čija dinamika prelazi u deterministički haos i prikazani su topološki obrasci za prepoznavanje dinamičkih režima takvih sistema. Na ovaj način je prikazan nov, sveobuhvatan metod za analizu dinamičkih sistema i kompleksnih sistema generalno, te je otvoren je put za nove primene algebarske topologije u fizici.

Ključne reči:
statistička mehanika, kompleksni sistemi, grafovi, kompleksne mreže, kombinatorna algebarska topologija, simplicijalni kompleksi, topološka invarijanta, kombinatorni Laplasijan, entropija, nelinearni dinamički sistemi, haos

Naučna oblast:
fizika

Uža naučna oblast:
statistička fizika

UDK broj:533.9(043.3)

Contents

Chapter 1

Introduction

The intriguing concept of complexity draws attention of scientists from various diverse fields of science. The standard approach in physics is of finding common properties in different systems in order to generalize the methods and the domain of applicability of the methods used. For example, in the study of critical phenomena universality indicates that distinct physical systems exhibit the same behavior near critical points. Hence, it is not coincidence that the concepts from statistical physics had the biggest influence in generating new approaches and in obtaining significant results about complex systems. The widely accepted frame of reference for the study of complex systems are complex networks (graphs) as this framework offers the use of efficient graph theory methods and the methods of statistical mechanics, due to their applicability to a large number of interacting elements.

Since there is no unique and widely accepted definition of a complex system [1], it is a challenge to formulate the theory of complex systems. Nevertheless, this problem does not averts researchers to adapt concepts and methods from their field of study on a specific problem, i.e. complex system, and provide results about it. The main characteristic of complex systems is that they are formed by the large number of elements which communicate, or in general interact, among themselves in a pairwise manner. One of the main features of complex systems is that each constitutive element has rather simple behavior while the aggregation of a large group of such elements displays a very complex dynamics. Interestingly, as it will be presented in this thesis, interactions does not have to be restricted to pairwise, but they can rather be of a higher-order, i.e. simultaneously involve more than 2 elements (agents).

One of the best ways to represent complex system where elements related like a pairwise interaction is mathematical graph associating elements with nodes (vertices) and their interaction as links (edges) of a graph. Hence, the usual represen-

tation of a complex system is studied from the aspect of a complex network [2]. There are many ways of representing complex systems by complex networks, and there are some accepted quantitative and common properties which characterize complex system in general. These properties are developed by a suitable mathematical frameworks. More about complex network in general will be introduced in the Section2.1.

Some specific properties of a complex system can be obtained by studying the properties of a simplicial complex, that is the object which consist of, in the simplest terms, connected polyhedra and which lives in a discrete geometrical space. The hierarchical organization inherent in substructures of complex networks may be efficiently studied by simplicial complexes by the use of combinatorial algebraic topology [3]. Ron Atkin [4, 5], following the ideas of Dowker [6] for a building simplicial complex from the relation between the elements of two sets, introduced the method of Q-analysis [7]. More recently Atkin's methodology received a further development in the work of Barcelo and Laubenbacher [8] who introduce the A-homotopy theory ("A" in honor of Atkin).

Simplicial complexes can be used for the study of topological properties of a complex system obtained from experimental data [9] and also in discretization of exterior differential forms. Its geometry and topology may be used as the basis of many physical theories such as general relativity [10, 11], electromagnetism [12], gauge theory [13], elasticity [14] etc. To illustrate the importance of simplicial complexes in theoretical physics research, let us emphasize that the geometric and topological nature of such theories is often obscured by their formulation in vectorial and tensorial forms due to unavoidable use of coordinate systems. Hence the complete topological and geometrical nature is obscured hiding for example, local and global invariants. Exterior derivative of differential forms is, on the other hand, invariant under a coordinate system change and since every differential equation can be expressed in terms of exterior derivative of differential forms [15], many physical laws may be expressed in terms of differential forms. Discretization of differential forms using finite differences and using their coordinates leads to numerical invalidation of some basic theorems (Stokes, for example) making traditional discretization methods impractical. It turns out that proper discretization of differential forms that preserves all the fundamental differential properties is possible only on simplicial complexes [15]. Further advances in the field of complex systems were made by combining ideas of statistical mechanics and algebraic topology [16]-[20].

An important part of the thesis, represented in Chapter 4, is devoted to the concept of recurrence in nonlinear dynamical systems. Predictions of dynamical

properties of a dynamical system are not based on long term calculations of mathematical equations, but on two crucial facts from our daily life:

1. similar situations often evolve in a similar way;

2. some situations occur over and over again.

The first phenomenon is about certain determinism [21] in many real-world complex systems, which are very different kind, with different space-time scales. All these systems are modeled by deterministic differential equations and they behave deterministically in all sense - we can predict the future state of such systems. Nevertheless, chaos theory taught us that some systems, even deterministic, are very sensitive to fine fluctuations and even small perturbations of initial conditions, and they can make precise predictions (on long term scale) almost impossible.

The second phenomenon is fundamental to many systems and is main reason for the establishment of memory through experience. It helps elements of such a complex system to remember similar situation and predict the ending and perhaps help to survive. Remembering the similar situation, for example, hot and humid day can eventually lead to storm, but only if a system with all features that are driving the system recur to former states. Such recurrence is a fundamental characteristics of many dynamical systems. More about recurrence of dynamical systems (complex systems) will be presented in the Chapter4

The main results of the thesis are that we found proper way to identify simplicial communities in real-world social networks (emotion propagation network) and identify "key players" for switching emotions. Further, in social networks we are able to reveal detailed structure of the phase space manifolds that contains information about the system collective behavior. Also, we showed that the study of higher-order combinatorial structures via algebraic topology tools provides a sensitive methodology to quantify the shifts in functional brain networks, under changed activity or condition. Regarding the nonlinear dynamical systems, main results pertain to new aspects of topological properties of different dynamical regimes and changes from one regime to the other. Quantification of these topological properties are obtained by the use of Q-analysis, persistent homology and multifractal formalism of statistical mechanics.

Hence, the continuation of the research initiated in this thesis would give a further insight in broad range of research areas, like sociology, neurology, economics, etc. Furthermore, study of nonlinear dynamical systems using algebraic topology tools, i.e. simplicial complexes, is at its beginning and it opens the possibility for expansion in a many directions.

Chapter 2

Algebraic topology of complex networks

2.1 Complex networks

Complex systems are usually represented by complex networks and they should comply to qualitative properties of complex systems:

- Despite great irregularity that occurs in complex systems, they display some kind of self organization and can be termed as organized complexity;
- The behavior of a complex system can not be predicted by just knowing the pairwise relations of elements, but it but it is to the large extent influenced by the structure built by those elements and their relations;
- Complex system possess hidden hierarchical organization which is reason for creation of such complexity of a system.

Complex networks are, perhaps, the most important field in complex system theory and they are focused to transform real world phenomena in unified way that researchers from different fields can observe such system and exceed research in their native fields. Mathematical objects from graph theory defined sets of elements together with their pairwise relations has been found convenient for easy reconstruction of a broad research community.

The beginnings of graph theory originates in the work of Euler [22]. An important influence on the study of complex networks via mathematical graphs was due to Paul Erdös and Alfred Rényi. One of their biggest contribution is the introduction of random graphs and the corresponding random graph theory [23, 24]. Random graphs are formed from the set of different vertices and the set of edges which connect

randomly, with pairwise edge being formed with probability p. Interestingly, such graphs display the small-world property, which means that most pairs of vertices are connected by a short path through the graph, just like it is observed in the real world experiments [25].

In contrast to randomness in graph theory are the properties of the emergence of communities and of clustering [26]. The community structure means that the groups of densely connected nodes, at least internally, appear in the network. The property of clustering means that two adjacent vertices which are both adjacent to the third vertex have an increased probability of also being adjacent of one another. Next, the clustering of network communities assumes aggregation of a set of nodes in such a way that nodes in the same cluster are more similar to each other (based on some measure), than to those in other clusters. In order to satisfy both randomness and high clustering, Watts and Strogatz designed the small-world network model [27], which displays both properties for certain specific parameters of the network.

A main problem in modeling real world networks is that we can model a network to resemble the real one based on one property but not for the other one, so they behave either as a random graph or as the small world network. Albert and Barabási proposed the scale-free network model [28], which in its essence incorporates two mechanisms characteristic for many real-world networks: the growth and the preferential attachment. The first mechanism, the growth, means that networks are growing by adding new vertices and connecting them to already present vertices in the network, whereas the likelihood to connect to the vertex already present in the network depends on the number of neighbors the vertex has, hence preferential attachment. In other words, the more connection the vertex has, the higher probability that the newcomer will connect to it. All mentioned models are just few pioneers of this field of study and the most important models.

2.1.1 Introduction to algebraic topology and overview of simplicial networks

Lets start with finite set $V = \{v_0, v_1, ..., v_{q-1}, v_q\}$ of elements which are called *vertices* and define a convex hull of $q+1$ elements $\sigma_q = \{v_{\sigma_0}, v_{\sigma_1}, ...v_{\sigma_q}\}$ of the set V called a *q-dimensional simplex* or just a *q-simplex* [29]. The geometrical representation of q-simplex is a polyhedron embedded in $\mathbb{R}^d$, where $q \leq d$ [29]. Examples of this realization are point (0-dimensional simplex), line (1-dimensional simplex), triangle (2-dimensional simplex), etc. A *face* σ_p of a q-simplex σ_q is a subset of p vertices of σ_q, where $p \leq q$, so that $\sigma_q \leq \sigma_q$. So if two simplices share $p+1$ common vertices, they share p-face.

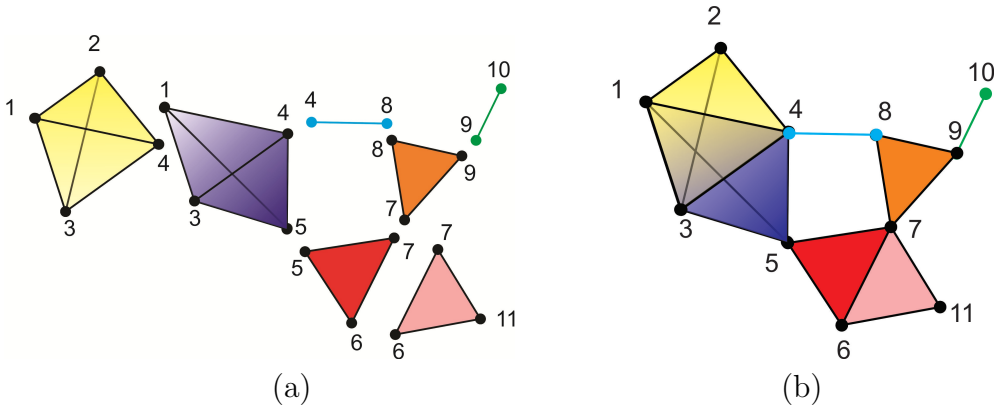


Figure 2.1: possible construction of a simplicial complex, right (b), from simplices left (a)

In Figure 2.1 (a) a geometrical representation of simplices for various dimensions of the set $V = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$ with subsets: $\{1, 2, 3, 4\}$, $\{1, 3, 4, 5\}$, $\{1, 4\}$, $\{5, 6, 7\}$, $\{7, 8, 9\}$, $\{9, 10\}$, $\{6, 7, 11\}$ is presented. In the same figure we can observe that subset $\{1, 3, 4\}$ of simplex $\{1, 2, 3, 4\}$ is also a subset of simple $\{1, 3, 4, 5\}$, indicating that these two simplices share a 2-face and by definition of a simplex this 2-face is also a simplex.

A collection of all simplices together with all their faces is called a *simplicial complex*. In more formal terms, a simplicial complex S_σ on a finite set $V = \{v_0, v_1, ..., v_{q-1}, v_q\}$ of vertices is a nonempty subset of the power set of V, such that S_σ is closed under the formation of subsets [29].The maximal dimension of a simplex in S determines the dimension of the whole simplicial complex, $D = dim(S)$. An illustration of the construction of a 3-dimensional simplicial complex from several simplices of different dimensions is presented in Figure 2.1 (b).

For practical purpose, we need some rule how to combine elements of set V into subsets which form simplices, and we must know what these simplices actually represent. Let us introduce two arbitrary sets $V = \{v_1, v_2, ..., v_n\}$ and $W = \{w_1, w_2, ..., w_m\}$ and a relation λ which together create two simplicial complexes S and S' where S' represents a conjugate complex of S [6]. The binary relation λ assigns to every element in W one or more elements in V, ie. for every $w_i \in W$ there exist $v_j \in V$ such that $w_i \lambda v_i$. The set W and the relation λ determine the subset S of the power set of V and we label each element $\{v_{\alpha 0}, v_{\alpha 1}, ..., v_{\alpha q}\} \in S$ $q \leq m$ by the element $w_i \in W$ for which $w_i\lambda, v_{\alpha 0}, w_i\lambda, v_{\alpha 1}, ..., w_i\lambda, v_{\alpha q}$. To distinguish the element w_i from the set W and its associated element from the set W due to the relation λ, the element of the set W will be labeled as σw_i. Therefore, the notation $\sigma(w_i) = \langle v_{\alpha 0}, v_{\alpha 1}, ..., v_{\alpha q} \rangle$ [32] means that an element w_i of the set W is λ-related to

q elements of set $\{v_{\alpha 0}, v_{\alpha 1}, ..., v_{\alpha q}\}$ in V. The elements of set V are called *vertices* and elements of set W are called *q-dimensional simplices* or just $q-simplices$. So, element w_i is λ-related to any subset of the set $\{v_{\alpha 0}, v_{\alpha 1}, ..., v_{\alpha q}\}$ and hence, every subset of that set is simplex, meaning that every subset is the face of a simplex, due to definition of q-face. Since each $w_i \in W$ identifies q-simplex $\sigma(a_i)$, for some q, together with all its faces, this collection is called simplicial complex S, which we denote as $S_w(V, \lambda)$ [5].

Let us introduce two sets to illustrate a construction of simplicial complex, $W = \{a, b, c, d, e, f, g\}$ and $V = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$ and suppose that elements of set W are λ-related to the elements of the set V. For example, the elements (letters) from set W may represent persons, elements of set V (numbers) represent TV shows and λ correspond whether person are watching TV shows "person a is watching TV show 1". Or, elements of set W represent doctors, element of set V represent patients and λ correspond to the property "doctor a is treating patient 1". Another example, elements of set W represent scientific article, elements of set V correspond to the coauthors and λ correspond to the property "an article a is written by coauthor 1", and so on.

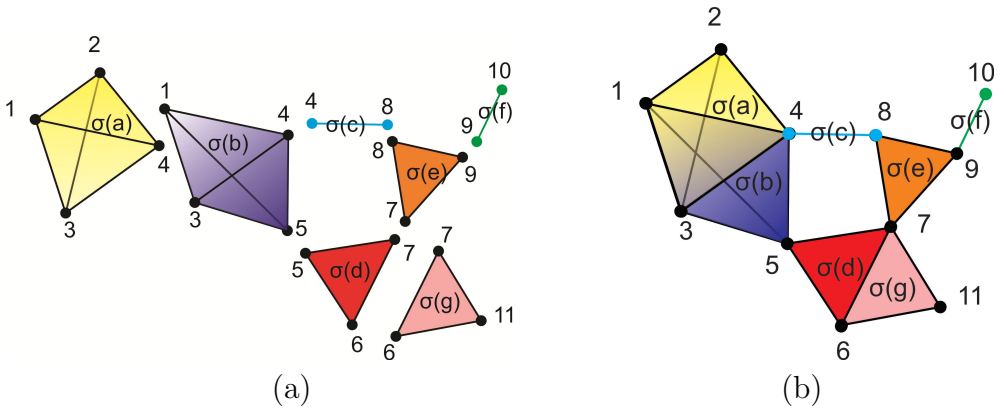


Figure 2.2: Creation of the simplicial complex from Fig 2.1 with labeled simplices

A geometrical (polyhedral) representation of simplices collected from elements from the set W, which are λ-related to the elements of the set V is presented in Fig 2.2(a). The obtained simplices are:

- $\sigma(a) = \langle 1, 2, 3, 4 \rangle$
- $\sigma(b) = \langle 1, 3, 4, 5 \rangle$
- $\sigma(c) = \langle 4, 8 \rangle$

- $\sigma(d) = \langle 5, 6, 7 \rangle$
- $\sigma(e) = \langle 7, 8, 9 \rangle$
- $\sigma(f) = \langle 9, 10 \rangle$
- $\sigma(g) = \langle 6, 7, 11 \rangle$

Fig 2.2 represents the simplicial complex formed by attaching simplices along their shared faces. Difference between Fig 2.1 and 2.2 is that simplices on Fig 2.2 have interpretation. For example, simplices are doctors and vertices are patients that are treated by different doctors. From this example we can observe that doctor a is treating patients 1,2,3,4 and doctor b is treating 1,3,4,5, and that both of them are treating patients 1,3,4. So, it is easy to catch how simplicial complex can find and display complicated relationships between properties of simplices with vertices that are constructing them, but it is crucial to assign simplices a "name" or important information can be lost.

λ relates elements of set W to elements of set V, so there must be reverse relation that connects somehow elements of set V to elements of set W. For example, if λ corresponds to the property "doctor a is treating patient 1", inverse relation λ^{-1} correspond to the property "patient 1 is visiting doctor a". So λ^{-1} relates elements of set V to elements of set W: $1\lambda^{-1}a$,$2\lambda^{-1}a$,$3\lambda^{-1}a$,$3\lambda^{-1}b$ and so on. Following the procedure, we form simplicial complex $S_v(W, \lambda^{-1})$ on the vertex set W by relation λ^{-1}, represented in figure 2.3. So elements of sets V and W switched their roles and in the complex $S_v(W, \lambda^{-1})$ vertices comes from set W and simplices from set V. A simplicial complex $S_v(W, \lambda^{-1})$ defined on sets $V = \{v_1, v_2, ..., v_n\}$ and $W = \{w_1, w_2, ..., w_m\}$ by inverse relation λ^{-1} is called the *conjugate complex* of the simplical complex $S_w(V, \lambda)$ [7, 32]. In such a way, conjugate complex is carrying important information about related system of elements from set V and W. In the previous example, each simplex in the conjugate complex stores information about all doctors that are treating the patient. In the case where for sets $W = V$, hence $S_w(W, \lambda)$ simplicial complex is the same as its conjugate complex.

Practical way to represent relation between two sets is by using an *incidence matrix* [7, 32] Λ, where rows are related to simplices and columns are related to vertices, and the matrix element $[\Lambda]_{i,j}$ is equal to 1 if simplex $\sigma(i)$ contains vertex j, otherwise is 0. The matrix representation of a conjugate complex $S_v(W, \lambda^{-1})$ of the simplicial complex $S_w(V, \lambda)$ is transpose matrix of Λ (Λ^T).

The matrix that captures the relations between simplices, in the sense of how

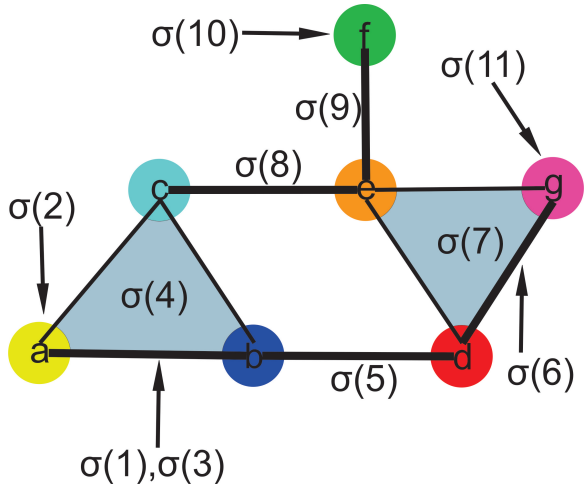
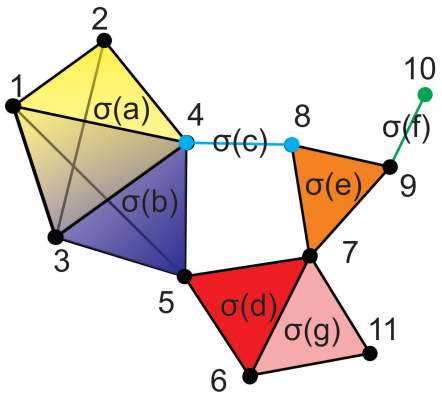


Figure 2.3: Conjugate complex of the simplicial complex from Figure2.2



$$\Lambda = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \\ g \end{array} \begin{array}{c} \begin{array}{ccccccccccc} \lambda & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Figure 2.4: Simplicial complex (left) and its incidence matrix (right)

they are "glued" to each other, is the *connectivity matrix* defined as:

$$\Pi = \Lambda \cdot \Lambda^T - \Omega, \tag{2.1}$$

where Λ is incidence matrix and Ω is matrix with all entries equal to 1. Rows and column of the matrix Π are related to simplices. Diagonal elements of connectivity matrix represent the dimension of simplices, whereas the non-diagonal element $[\Pi]_{ij}$ is equal to the dimension of the face shared by simplices i and j share or the non-diagonal element is equal to -1 if simplices i and j do not share a face. Connectivity matrix for example from Fig. 2.4:

$$\Pi = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \\ g \end{array} \begin{array}{c} \begin{array}{ccccccc} a & b & c & d & e & f & g \end{array} \\ \begin{pmatrix} 3 & 2 & 0 & -1 & -1 & -1 & -1 \\ 2 & 3 & 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 1 & -1 & 0 & -1 & -1 \\ -1 & 0 & -1 & 2 & 0 & -1 & 1 \\ -1 & -1 & 0 & 0 & 2 & 0 & 0 \\ -1 & -1 & -1 & -1 & 0 & 1 & -1 \\ -1 & -1 & -1 & 1 & 0 & -1 & 2 \end{pmatrix} \end{array}$$

Other good representation for simplicial complex, that is more intuitive, is geometrical representation, but it is not always pratical because of the size of sets of vertices and simplices.

2.1.2 Simplicial complex construction

This section contains description of construction methods of simplicial complexes from graphs, times series and from data embedded in metric space.

From complex networks (graphs)

A simplest way to represent a complex system of mutually interacting elements is by a complex network, i.e. a large graph. Mathematically speaking, graph is a set of N vertices, or nodes, connected by links (edges) which can be directed or undirected. Undirected links, bidirectional ones, are observed as links with two directions, i.e. they allow information transition in both ways.

A large number of different types of simplicial complexes may be constructed from a single graph. For example clique complex, neighborhood complex, independence complex, matching complex etc., may be constructed from one graph, depending on the focus of the study and the properties of the underlying network (graph). By building more than one simplicial complex from graph, it is possible to obtain different information about the pairwise relations of underlying graph and about whole complex system which is analyzed. The in depth analysis of these simplicial complexes reveals different properties and information about the complex network. Depending on the focus of the research and the properties of the network which are of particular interest,a simplicial complex or complexes may be constructed that best suite the purpose.

Neighborhood complex

neighborhood complex [33, 34, 35, 36] is constructed from simplices associated to each node i of the original graph. Each simplex consists of the vertices which are connected to the node i in the graph (network), so that simplices are all subsets of the vertex set of the underlying graph which have a common neighbor. Incidence matrix of the neighborhood complex is equal to the adjacency matrix of the underling graph. It is easy to deduce that the neighborhood complex and its conjugate complex are the same. In the Fig 2.5 each node is associated to its neighbors. So sets of neighbors $\{2,3\}$, $\{1,3,4\}$, $\{1,2,4,5\}$, $\{2,3,6\}$, $\{3,6\}$, $\{4,5\}$ are associated to simplices $\sigma^1(1)$, $\sigma^3(2)$, $\sigma^3(3)$, $\sigma^2(4)$, $\sigma^1(5)$, $\sigma^1(6)$ respectively. Fig 2.5(b) is geometrical representation of neighborhood complex with simplices depicted in different colors, same colors as nodes of underlying graph Fig 2.5(a). One of the applications of this type of simplicial complex is to reveal hidden relationships between elements constituting the complex network.

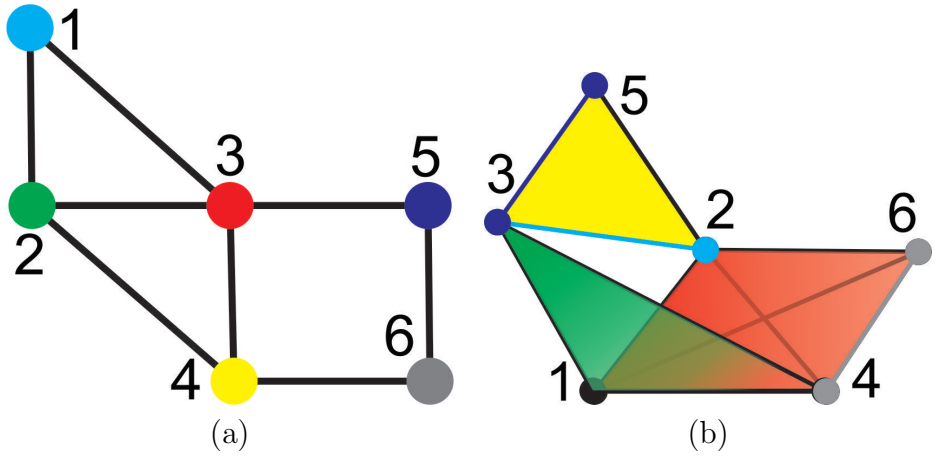


Figure 2.5: Constructing a neighborhood complex (b) from an undirected graph (a). The colors emphasize matching between nodes of the original graph and associated simplices in the neighborhood complex.

Clique complex

Cliques are originating from social sciences, i.e. social network analysis, and representing set of vertices that every two distinct vertices in the clique are adjacent, as analogy with the highly connected social group. Clique complex consists of the simplices constructed from the cliques of the underlying graph, whose faces have the maximal possible dimension. To construct the simplicial complex we need to create two sets, the first containing all maximal cliques and the second containing all

vertices of the initial graph. In a graph we try to identify all maximal cliques. For example, tetrahedron presents a maximal clique of order 3 although it is constructed from 4 triangles (cliques of order 2). So, a simplicial complex in which simplices are all maximal cliques of he underlaying graph is called *clique complex*[37, 38]. *Conjugate clique complex* is a simplicial complex in which cliques and nodes are switching roles so that the cliques assume the roles of vertices and the nodes are connected as simplices. Conjugate clique complex is some times needed when want to extract the information about the relations between nodes with respect to the cliques which they construct.

In Fig. 2.6(a) 5 maximal cliques may be identified: two triangles and three links depicted in different colors in the same figure (right). The set of vertices is $\{1, 2, 3, 4, 5, 6, \}$ is used to construct a 2-dimensional simplicial complex with the set of simplices:

- $\sigma^2(a) = \langle 1, 2, 3 \rangle$,
- $\sigma^2(b) = \langle 2, 3, 4 \rangle$,
- $\sigma^1(b) = \langle 3, 5 \rangle$,
- $\sigma^1(b) = \langle 4, 6 \rangle$,
- $\sigma^1(b) = \langle 5, 6 \rangle$,

where superscript is associated with the dimension of the simplex.

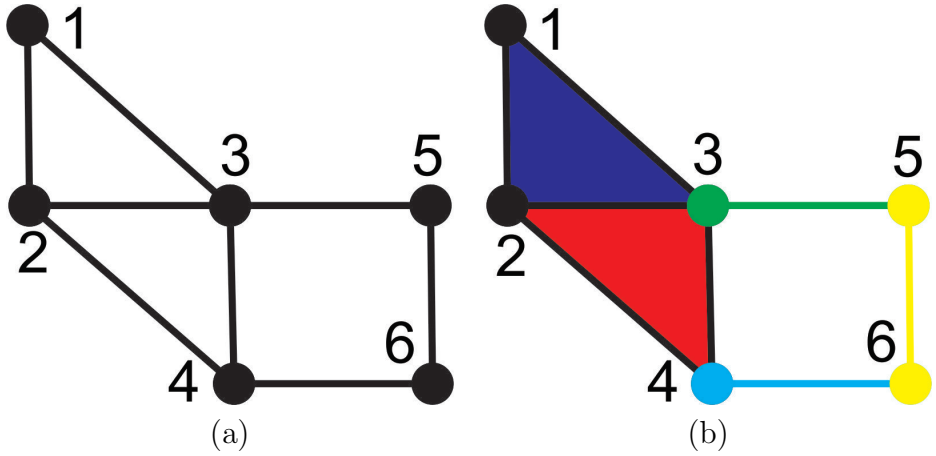
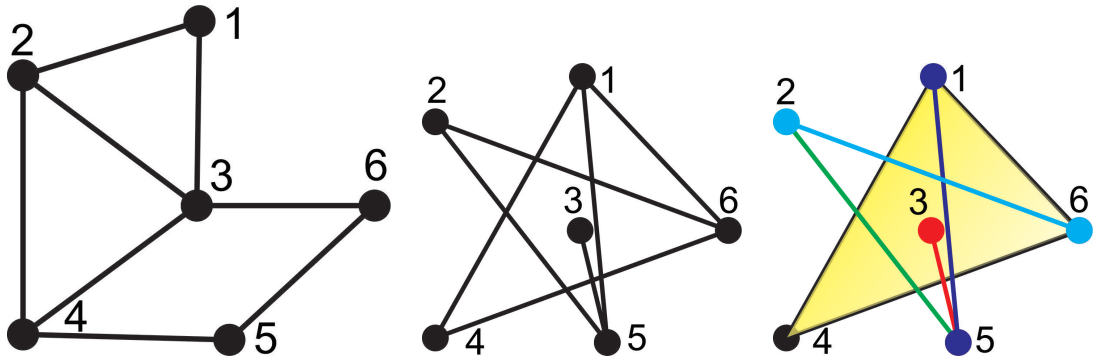


Figure 2.6: Constructing a clique complex (b) from an undirected graph (a). Different colors indicate maximal cliques, i.e. simplices

Independence complex

Relationship between nodes of a graph representing complex network is preserved through links of simplices in the simplicial complex including the information about communities and clustering. But it is also important to examine nonexisting links between nodes. In a real world complex network, it may be interesting to explore missing relationship between nodes. Therefore, it is necessary to build a graph which is complementary to the initial graph meaning that in a new graph two nodes are connected if they are not connected in the underlying initial graph, and vice versa, two nodes are not connected if they are connected in the initial graph. By finding all maximal cliques of the new graph, called anticliques or independence sets, we create *independence complex* [37]. Construction of an independence complex starting from the undirected graph is presented in Fig 2.8 . After all maximal cliques of such a graph are found an independence complex 2.8(c) is formed.



(a) (b) (c)

Figure 2.7: Building an independence complex (c) from an undirected graph (a). Different colors indicate maximal cliques, i.e. simplices, of complementary graph (b).

Matching complex

So far the focus was on the nodes and how they are connected and we switch the focus on links now. So, to examine relation between links, underlying graph can be transformed to a *line graph*. Nodes of line graph are associated with links of the initial graph, and they are connected if they share mutual node of initial graph. The independence complex of a line graph is called the *matching complex* [39]. It provide us with information of high order structures built by links that do not share mutual nodes of the underlying graph.

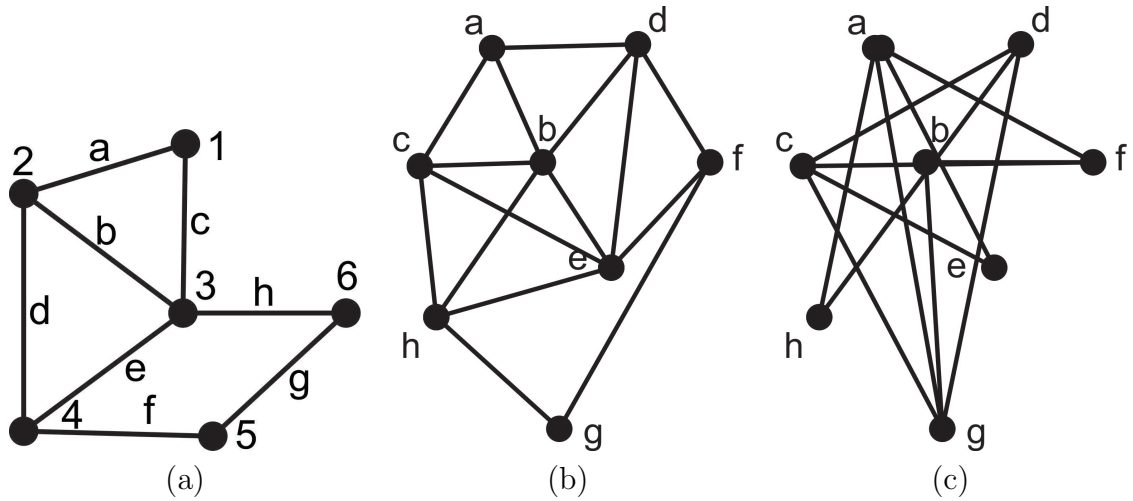


Figure 2.8: Building an matching complex (c) from line graph (b) of undelying undirected graph (a).

Construction of simplicial complex from data in phase (metric) space

When analyzing simplicial complexes constructed from graph, there is no need to observe distances between nodes, just relationship between them through links. Nevertheless, there are data sets where it is important to incorporate distances between elements of a system, like the system of antenna towers, sensors, mobile phone towers, where relationship between elements depend on overlapping of range is of great importance. Hence, for these systems, elements are associated with coordinates of an Euclidean metric space. Distances in D-dimensions are defined as

$$d(p_1, p_2) = \sqrt{\sum_{i=0}^{D} (x_{i,1} - x_{i,2})^2}.$$

It may be of great interest to reconstruct the global structure that originates from geometrical distances. For that purpose it is necessary to reconstruct the shape of data, and use Q-analysis presented in Section 2.1.3 and homology, in Section 2.2.1, to describe global and local properties. From Fig 2.9 we have a 2-dimensional Euclidean space and now we want to apply the method for simplicial complex reconstruction to extract information about the system. There are several methods of constructing the simplicial complex from the data that depends on distances between elements of which we use the Cech complex, the Vietoris-Rips complex and the witness complex, among others.

An important thing to know is that distances between elements are not the only important information, but also the open balls (as range of antenna, mobile

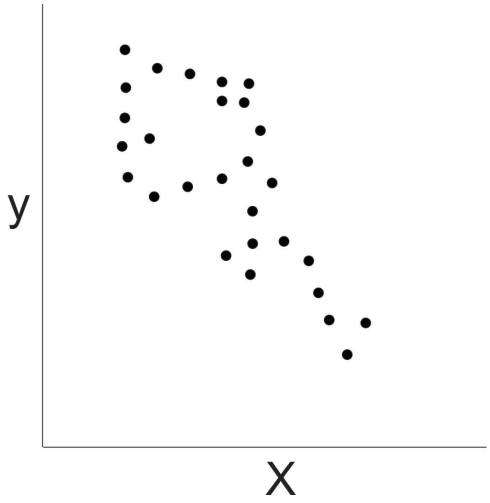


Figure 2.9: An example of data points distributed in 2-dimensional Euclidean space.

tower coverage) have a significant role because of overlapping. So, in this metric space we have a set of open balls $B = \{B_1, B_2, ..., B_N\}$, such that an aggregation of balls and their intersection is contractible and union $\cup B_i$ is the space of interest. So, making each B_i a vertex a q-dimensional simplex is formed whenever the intersection $B_{i0} \cap B_{i1} \cap ... \cap B_{iq} \neq \varnothing$ appears. With respect to so defined topological spaces, topological invariants, such as holes, i.e dimension of thre homology group (formal definition will be introduced in 2.2.1), face vector 2.2.2 are of great importance and this technique is called the nerve of the cover [40].

Cech complex

Let start with set of data points X in some metric space M and data points are centers of balls of radius $r \in \mathbb{R}$, such that $B(x_i, r)$, where for each $x_i \in X$

$$B(x_i, r) = \{m \in M \mid d(x_i, r) < r\}.$$

Links (1-simplex) are built between data points (vertices) whenever two balls have no empty intersection, triangles (2-simplex) are built where three balls intersects with each other. In a similar manner higher dimensional simplices are created and so obtained simplicial complex is called the *Cech complex* [37]. It is clear that the construction of Cech complex is related to the nerve of the cover, recovering in this way the shape of space built by data points.

In the example of a 2-dimensional data set presented in Fig 2.9, around each point a circle is drawn with radius r 2.11. So, following the definition of a Cech complex, wherever a group of circles overlap each other, a simplex is created. It is important is to make a difference between subsets of points $\{a, b, c\}$ and $\{d, e, f\}$,

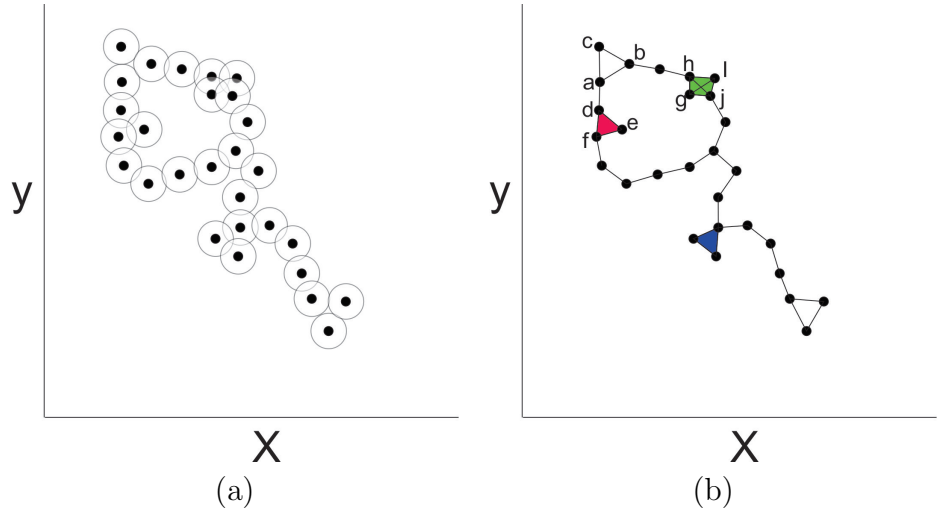


Figure 2.10: An example of building the Cech complex (b) from data points (a).

because the first subset builds 1-dimensional hole since circles overlap in pairs, not like in other subset where all three circles have common intersections and they form a triangle (2-dimensional simplex). As in second subset, same thing holds for subset $\{g, h, i, j\}$ which form a tetrahedron (3-dimensional simplex). In case of diameter increase 1-dimensional holes would be filled and only 1 large (central) hole will survive.

Vietoris-Rips complex

Vietoris-Rips complex is similar to the Cech complex and it is formed from the intersection of balls centered around the data points but with a different rule. Instead of adding a d-simplex when there is a common point of intersection of all the $(r/2)$-balls, a simplex is added when all the balls have pairwise intersections.

Unlike when forming a Cech complex in Fig 2.9 where subset of points $\{a, b, c\}$ creates a hole, the same subset creates a 2-dimensional simplex in the Vietoris-Rips complex, as is the case with other 1-dimensional holes from the Cech complex. The other higher-dimensional simplices from the Cech complex also appear in the Vietoris-Risp complex, since six 1-dimensional simplices are embedded in four 2-dimensional simplices which are faces of a 3-dimensional simplex of the Cech complex (similar procedure as when constructing a clique complex).

Witness complex

Density of data points or large data sets may be impractical and computationally very demanding. To overcome this problem, there is a need to reduce data sets and

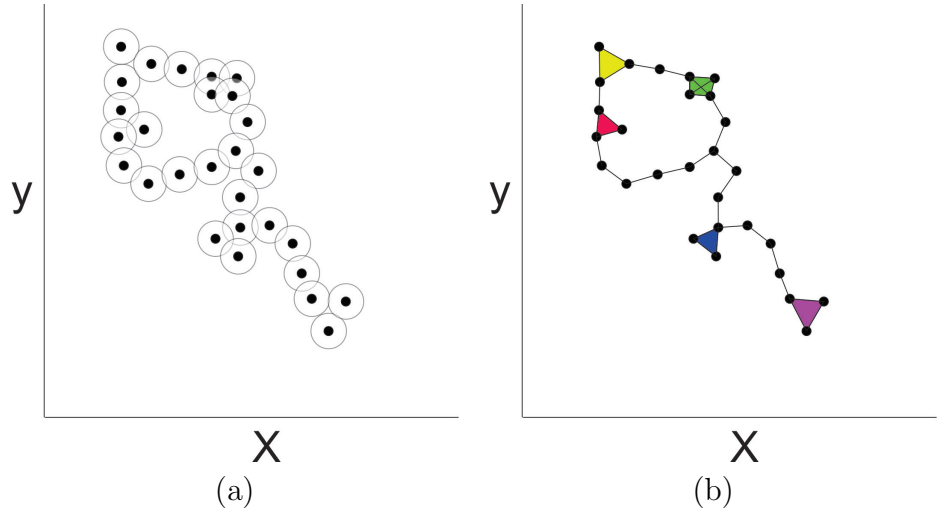


Figure 2.11: An example of building the Vietoris-Rips complex (b) from data points (a).

yet not to loose important information and properties of such system. For large data point set X in some Euclidean space, there is a choice of small set of points $L \subset X$ called *landmarks* and together with the set of points $X \setminus L$, called the *witnesses*, a *witness complex* is formed on this set of points instead of X. There the following way. A q-simplex $\sigma_q = \langle l_0, l_1, ..., l_q \rangle$ is weakly witnessed by $x \in X \setminus L$ if $d(l, x) \leq d(k, x)$ for every $l \in \{l_0, .., l_q\}$ and $k \in L \setminus \{l_0, ..., l_q\}$. Subset of points $\{l_0, ..., l_q\} \in L$ is a simplex if and only if there is a point (a witness) $x \in X \setminus L$ with every point in $\{l_0, ..., l_q\}$ closer to x than to any other point in $L \setminus \{l_0, ..., l_q\}$. The witness complex is defined as the collection of all simplices together and their faces built on the vertex set L that are weakly witnessed by a point in X [41].

Although the building of a witness complex reduces the size of the initial data point set, the good choice of landmark set preserve the topological properties of the original data set. Like in the case of previously introduced simplicial complexes from data in metric space where good choice of radius led to persistent homology calculations, same thing holds for the witness complex [42].

In Fig. 2.12 a set of points used in Fig. 2.9 we randomly select 6 landmark points which form a 1-dimensional simplicial complex and some of the simplices surround a 1-dimensional hole. This hole is preserve even if all points are connected mutually. Homological structure persists even when we reduce the number of data points from the initial sample.

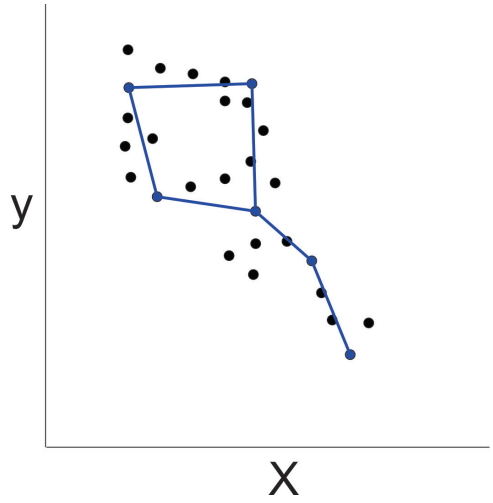


Figure 2.12: An example of building the witness complex from data points from Fig 2.9

From time series

As in the previous cases of constructing simplicial complexes from a graph and from data sets in metric space, there are several ways to create simplicial complex from a time series. All these cases require construction of both the graph and a simplicial complex. The extraction of properties that drive the dynamics of a complex system can be revealed from geometrical, combinatorial and topological features of phase space and the interest in relating them with simplcial complexes recently draw some attention.

Often the only information that are given about the dynamical complex system is time series, the time sequence of some known or unknown variable. There is a need to devise methods for determination of the number of variables, as well as methods which provide information about the relationship between them.

One of the best ways to reconstruct both a graph and a simplicial complex from data that comes from a large number of time series in order to determine relation between them is to use correlations. One of the best and widely used ways is to use the Pearson's correlation coefficient for a pair of time series A_t and B_t

$$C_{AB} = \frac{1}{N-1}\sum_{i=1}^{N}(\frac{A_i - \mu_A}{\sigma_A})(\frac{B_i - \mu_B}{\sigma_B})$$

where μ and σ are the mean and standard deviation of the corresponding time series, and N is the length of time series. Further, to separate the strong positive correlations, which are relevant in almost every context, there is way to use the filtering algorithm described in [43, 44, 45]. The algorithm enhances those

matrix elements C_{ij} that have a similar correlation pattern with the rest of the matrix elements while diminishes those with a dissimilar patterns. First we map C_{ij} to the range $[0, 1]$ using $CP_{ij} = (C_{ij} + 1)/2$. Then, each element is multiplied $CP_{ij} \rightarrow F_{ij}CP_{ij}$ by the corresponding factor F_{ij}, which is computed as Pearson's coefficient of the rearranged matrix elements from row i and column j as follows: $\{C_{ij}, C_{i1}, C_{i2}, ..., C_{iN}\}\{C_{ji}, C_{1j}, C_{j2}, ..., C_{jN}\}$. The resulting filtered correlation matrix is also transfered to a binary adjacency matrix of the graph by retaining the correlations larger than a threshold value and inserting units for the retained edges. Big challenges is the choice of threshold. Usually, by fitting distribution of obtained filtered correlation coefficient and find the value where distribution and fit are disjointed is probably good first choice. To acknowledge this choice there is a need to check some topological features whether they persist for higher value of the chosen threshold.

In order to analyze relation between values of one time series and which is particular suitable in the case of persistent fluctuations a *natural visibility algorithm* is used [46, 47]. Every time step is mapped to the node of the graph and the node is connected by undirected links with all other data points which are "visible" from that data point, where the vertical bars are considered as non-transparent. Note that by varying the mapping procedure different graphs can be obtained. The mapping procedure is illustrated in Fig. 2.13.

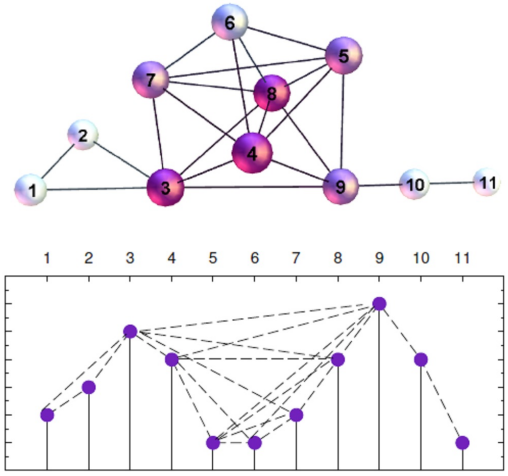


Figure 2.13: The sequence of data points in the lower pannel is mapped onto a graph in the upper pannel; each data point becomes a node of the graph, while the graphs edges are inserted according to the natural visibility between data points, indicated by broken lines.

For the purpose of analyzing the data that comes from nonlinear dynamical system where recurrence occurs, we introduced the recurrent simplicial complex. In depth explanation of implemented method will be in 4.2.

2.1.3 Chains of connectivity, global and local properties of simplicial complex

So far we have introduced the dimension of the simplex and the relationship (or adjacency) between two simplices through the shared face, which are stored in the connectivity matrix. Now we will introduce a higher aggregations of simplices induced through the shared face and, further, how they induce the intrinsic hierarchical multilevel and multidimensional organization of simplicial complex. The property that any subsimplex of a simplex is also a simplex induces various levels of adjacency between simplices, and also various levels of connectivity between collections of simplices. Two simplices are *q-near* if they share a q-dimensional face, and so on, they are also (q-1)-,(q-2),....,1 and 0-near. The collection of simplices in which any pair of simplices is connected through a sequence of simplices where a pair of connected simplices are q-near is called *q-connected component*. In other words, simplices σ and ρ are *q-connected* [4] if there is a sequence of simplices $\{\sigma, \sigma(1), ..., \sigma(n), \rho\}$ such that any two adjacent simplices share at least q-face.

Q-*vector* (*first structure vector*) [48, 32] is an integer vector with the length of maximal dimension of simplicial complex plus one. The values of Q-vector (Q_q) quantify number of q-connected component and they usually start from the number of connected components from largest dimension in descending order

$$Q = \{Q_{q_{max}}, Q_{q_{max}-1}, ..., Q_1, Q_0\}.$$

An example illustrating this vector is in Fig. 2.14 and for given example we can observe that first structure vector is $Q = \{2, 4, 5, 1\}$.

Another important quantity is called *second structure vector* N_s which value $N_s(q)$ gives the number of simplices with dimension larger or equal to q,

$$Q = \{N_s(q_{max}), N_s(q_{max} - 1), ..., N_s(1), N_s(0)\}.$$

From example from the Fig 2.14 second structure vector is $N_s = \{2, 5, 7, 7\}$.

And finally, the entries of $\bar{Q_q}$ (*third structure vectors*) [49] provides us with the degree of connectedness at all dimensions of a simplcial complex defined as

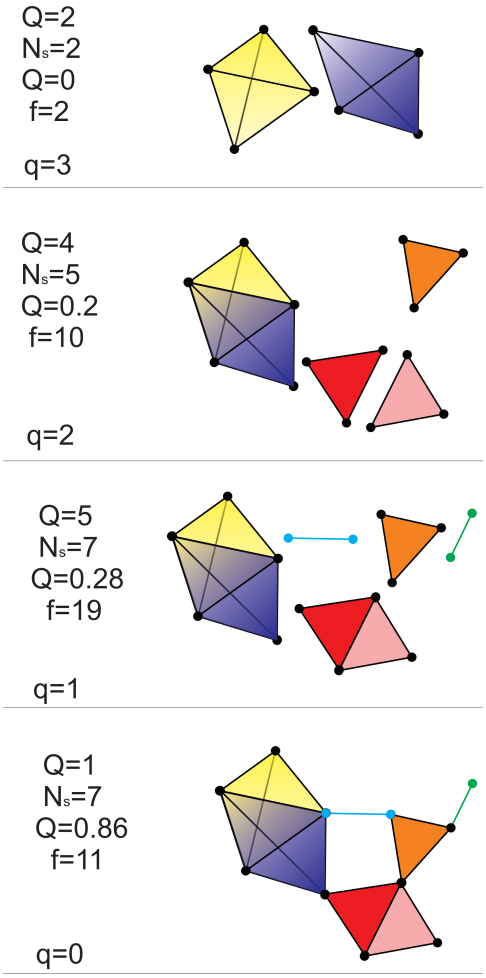


Figure 2.14: Structure vectors of simplicial complex

$$\bar{Q}_q = 1 - \frac{Q_q}{N_s(q)},$$

and it measures the number of q-connected components per number of simplices on q-level. From example in Fig 2.14 we can see that third strucure vector is $\bar{Q} = \{0, 0.2, 0.28, 0.86\}$.

One of the most important vectors that is also a topological invariant is the f-vector which represents the number of q-dimensional simplices and q-dimensional faces which are embedded in simplicial complex and from example in Fig 2.14 f-vector is $f = \{11, 19, 10, 2\}$. More about this feature may be found in Section2.2

All these vectors represent certain kind of global properties of a simplcial complex. Let us now observe some local properties of simplicial complex that describe environment of simplices in simplicial complex. We introduced *Node's Q-vector* Q^i [50] associated with node i from underlying graph which entries represents the number q-simplices in which node i participates and *topological dimension* $dimQ^i$ of the node

i as the number of all simplices in which node i participates as $dimQ^i = \sum_{q=0}^{q_{max}} Q_q^i$. The motivation for the use of the term dimension originates in view of the conjugate simplicial complex. Therefore, $dimQ_i$ corresponds to the dimension of the conjugate simplex [51].

2.2 Homology and Combinatorial Laplacian

So far, relationship between two sets have explored structural and connectivity properties of simplicial complex only through connectivity of simplices. Lets turn around to topological properties of simplicial complex and to the power set of the set over which simplicial complex is defined. Keep in mind that every subsimplex (face of a simplex) is also a simplex in simplicial complex. This premise was not explicitly emphasize, but implicitly accepted and it of great importance for upcoming concepts. So, when said q-simplices, it actually mean "all maximal q-simplices and all q-dimensional faces", and this represents the definition of f-vector.

2.2.1 Homology group, Betti numbers and reduction algorithm

To show more of the richness of algebraic topology machinery, lets start from example as before - Fig 2.15.

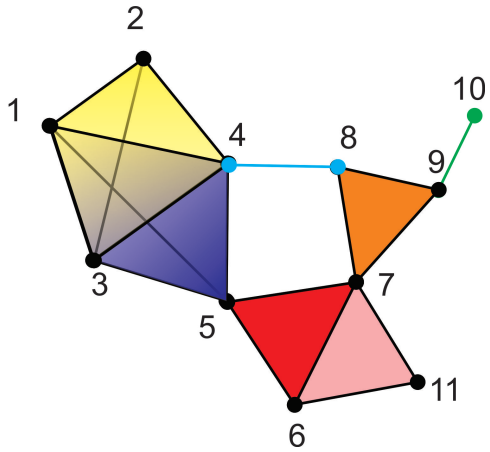


Figure 2.15: An example of simplicial complex

Let start with a finite vertex set $V = \{v_1, v_2, ..., v_n\}$. An arbitrary ordering of vertices $\{v_{\sigma 0}, v_{\sigma 1}, ..., v_{\sigma q}\}$ of a simplex defines an oriented q-simplex which will be denote as $[v_{\sigma 0}, v_{\sigma 1}, ..., v_{\sigma q}]$, and if all simplices are oriented that means that simplicial complex S_σ is oriented. Example 0-, 1-, 2- and 3-dimensional simlices is illustrated in Fig 2.16, and by convection 0-simplex does not have an orientation.

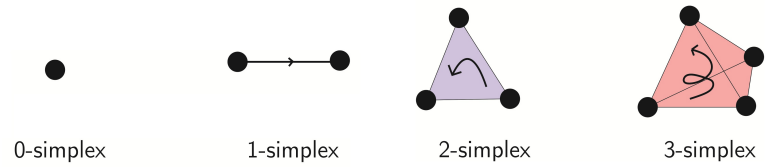


Figure 2.16: Examples of orientation of 0-, 1-, 2-, and 3-simplex

Let $C_q(S)$ be the vector space whose base is the set of all q-simplices of an oriented simplicial complex S_σ and the elements are the linear combination of bases vectors, called *chains*. THat is, a *q-chain* is a the formal sum of oriented q-simplices $c_q = \sum_i a_i \sigma_q(i)$, where coefficients $a_i \in \mathbb{Z}$ are the elements of coefficient group. $C_q(S)$ is called *chain group* [52] which dimension is equal to the q^{th} entry of already mentioned f-vector, $f = \{f_0, f_1, ..., f_D\}$ (unlike vector like measures from Q-analysis, the first entries of f-vector and other quantities in homology theory are associated with 0-dimension, second entries are associated to 1-dimension etc.). f_q is equal to the number of q-simplices (as previously mention by q-simplex we mean all maximal q-simplices and all q-faces embedded in simplicial complex) of the simplicial complex S_σ, i.e. f_0 represents the number of vertices, f_1 number of edges and from example Fig 2.15 we have $f = \{11, 19, 10, 2\}$, so this simplicial complex contains 11 vertices, 19 edges, 10 triangles and 2 tetrahedrons.

For a set of vector space $C_q(S)$ with $0 \leq q \leq dim(S)$ the linear transformation $\partial_q : C_q(S) \to C_{q-1}(S)$ is called *boundary operator* and it operates on vectors $[v_{\sigma 0}, v_{\sigma 1}, ..., v_{\sigma q}]$ as [52]

$$\partial_q[v_{\sigma 0}, v_{\sigma 1}, ..., v_{\sigma q}] = \sum_{i=1}^{q} (-1)^i [v_{\sigma 0}, v_{\sigma 1}, ...v_{\sigma(i-1)}, v_{\sigma(i+1)}, ..., v_{\sigma q}].$$

Fig 2.17 illustrates action of the boundary operator of a 3-simplex and its subsimplices.

Taking a sequence of chain groups $C_q(S)$ connected through the boundary operators ∂_q, called *chain complex*, is defined:

$$\varnothing \to C_q \overset{\partial_q}{\to} C_{q-1} \overset{\partial_{q-1}}{\to} ... \to C_1 \overset{\partial_1}{\to} C_0 \overset{\partial_0}{\to} \varnothing,$$

with $\partial_q \partial_{q+1} = \varnothing$ for all q. For example of the 3-simplex $[1, 2, 3, 4]$ from Fiq 2.17 we can illustrate how boundary operator works:

$$\partial_3[1, 2, 3, 4] = -[1, 2, 3] + [1, 2, 4] - [1, 3, 4] + [2, 3, 4],$$

and when we apply boundary operator one more time

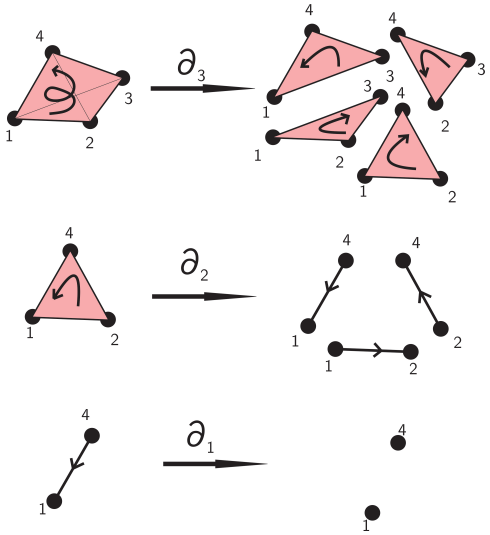


Figure 2.17: Action of ∂_3,∂_2 and ∂_1 on 3-simplex and its subsimplices

$$\partial_2\partial_3[1,2,3,4] = -[1,2]+[1,3]-[2,3]+[1,2]-[1,4]+[2,4]-[1,3]+[1,4]-[3,4]+[2,3]-[2,4]+[3,4] = 0.$$

The kernel of ∂_q is the set of q-chains with empty boundary, a q-cycle is a chain in the kernel of ∂_q (an element of the group of cycles Z_q). The image of ∂_{q+1} is the set of q-chains, which are boundaries of $q+1$-chains is denoted as B_q, known as the *group of boundaries*. B_q and Z_q are subgroups of chain group C_q. The q^{th} *homology group* [52] is defined as

$$H_q = ker\partial_q/im\partial_{q+1} = Z_q/B_q.$$

The elements of homology group H_q are equivalence classes of q-cycles which are not boundaries of $(q+1)$-chain, and can be interpreted that homology characterizes q-dimensional holes. The rank of the q^{th} homology group, i.e. q^{th} *Betti number*, $\beta_q = rank(H_q) = dim(H_q)$ is topological invariant and is equal to the number of q-dimensional holes in simplicial complex. So, β_0 is number of connected components of simplicial complex, β_1 is the number of one-dimensional or "circular" holes, and β_2 is the number of two-dimensional "voids" or "cavities". From example illustrated on Fig 2.15 we can see that $\beta_0 = 1$, there is only one connected component and $\beta_1 = 1$ (1-dimensional hole bounded by 1-dimensional simplices $[4,5]$, $[4,8]$, $[5,7]$ and $[7,8]$). Dowker [53] have proved, the homology groups of simplicial complex and its conjugate complex are isomorphic, therefore the values of Betti numbers of simplicial complex are preserved in its conjugate complex.

One of the important topological invariants, related to Betti numbers and f-

vector, is Euler characteristic [54]. For a simplicial comples S_σ with f-vector values, the Euler characteristic is defined as

$$\chi = \sum_{i=0}^{D} (-1)^i f_i,$$

and from the *Euler-Poincaré theorem*, using sum of Betti numbers is the same as Euler characteristic as

$$\chi = \sum_{i=0}^{D} (-1)^i \beta_i.$$

Computation of the Betti numbers for real-world complex networks with huge-dimensional simplicial complexes can be a big problem. So, many algorithms that are based on this problem either do with 2-dimensional simplices (low-dimension simplicial complex) or reduce (reduction algorithm) simplicial complex to just boundary simplices around holes, so many information about system is lost. Knowing this, there is a need for reduction algorithm that will preserve all topological invariants of simplicial complex and yet reduce it that computing of dimension of homology groups will be in realistic time. Algorithm is simple, just iterate as much as possible by the following rule: "Eliminate all nodes (and their links) which topological dimension is one". To be boundary of a hole, node must be part of at least 2 boundary links (which before applying reduction algorithm can be part of one simplex).

2.2.2 Combinatorial Laplacian

The content of this section relies on T. Goldberg's in-depth analysis in his manuscript *Combinatorial Laplacians of Simplicial Complexes* [55] and D.Horak's work with this subject [56]. As previously mentioned, any face of a simple is a simplex as well. So, if we have two q-simplices σ_i and σ_j, they can share a lower-dimension simplex (face), from $(q-1)$- to 0-simplex and we can say that they are neighbors through a common face. On the other hand, they can be both faces of some higher dimensional simplex, for example $(q+1)$-simplex, implying that they are neighbors as parts of same higher-dimensional simplex. So, for defining an appropriate quantity, we rely on introduced concepts of oriented simplex and boundary operator since we want to define measure that bound $q-$ and $(q-1)$-simplex for example. For formal definition, we need two q-simplices $\sigma_q(i)$ and $\sigma_q(j)$ of an oriented simplicial complex S_σ. $\sigma_q(i)$ and $\sigma_q(j)$ are *upper adjacent*, denoted $\sigma_q(i) \sim_U \sigma_q(j)$, if they are both faces of some $(q+1)$-simplex in S_σ. The *upper degree* of a q-simplex σ_q in S_σ, denoted $deg_U(\sigma_q)$ is a number of $(q+1)$-simplices in S_σ of which σ_q is a face. If oriented q-simplices

$\sigma_q(i)$ and $\sigma_q(j)$ are upper adjacent and have a common $(q+1)$-simplex τ, we say that they are *similarly oriented* if orientation of them agree with the once induced in τ. For easier understand, this definitions are illustrated at Fig 2.19. For two q-simplices $\sigma_q(i)$ and $\sigma_q(j)$ of an oriented simplicial complex S_σ we say that they are *lower adjacent*, denoted $\sigma_q(i) \sim_L \sigma_q(j)$, if they have common $(q-1)$-face. Hence, the *lower degree* $deg_L\sigma_q$ of a q-simplex is defined as the number of $(q-1)$-faces in σ_q, which is always equal to $q+1$. Simplices $\sigma_q(i)$ and $\sigma_q(j)$ from example in Fig 2.19 have twofold adjacency - they are upper adjacent through part of τ and lowe adjacent through sharing face, the simplex φ.

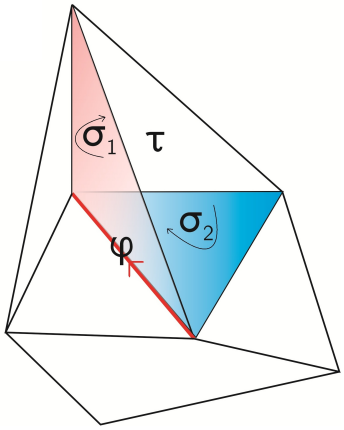


Figure 2.18: An example of upper and lower adjacency between simplices σ_i and σ_j .

For the definition of combinatorial Laplacian of simplicial complex, we have already defined the boundary operator and its adjoint. For a simplicial complex S_σ and an integer $q \geq 0$, the q^{th} *combinatorial Laplacian* is linear operator defined as $L_q : C_q \to C_q$ and given by formula [57]:

$$L_q = \partial_{q+1} \circ \partial^*_{q+1} + \partial^*_q \circ \partial_q.$$

Relating the upper/lower adjacency and definition above we can rewritten as

$$L_q = L_q^{UP} + L_q^{DN},$$

where $L_q^{UP} =$
$partial_q \circ \partial^*_q$ is referred as upper combinatorial Laplacian and $L_q^{DN} = \partial^*_q \circ \partial_q$ is referred as down combinatorial Laplacian. Corresponding matrix representation relative to some ordering of the standard bases for C_q for q^{th} *Laplacian matrix* of S_σ is

$$\mathcal{L}_q = B_{q+1}B_{q+1}^T + B_q^T B_q.$$

As in the case of the Laplacian operator we can use following notation for convenience $\mathcal{L}_q^{UP} = B_q B_q^T$ and $\mathcal{L}_q^{DN} = B_q^T B_q$.

An illustrative example is in Fig 2.2.2 where can be found the matrix representations of boundary operators and combinatorial Laplacian.

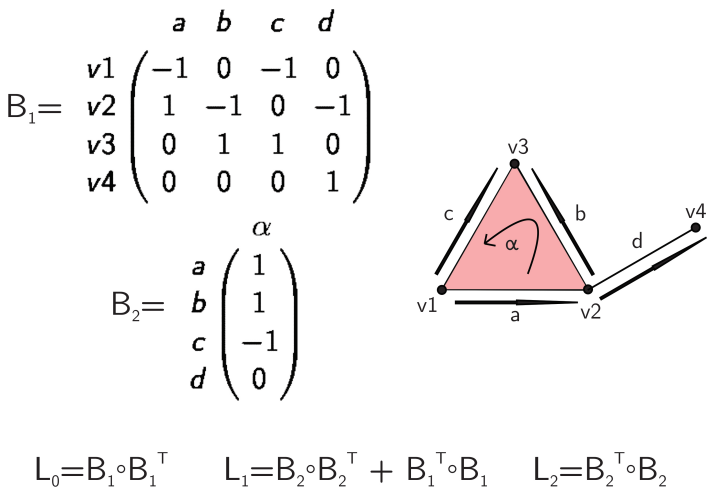


Figure 2.19: An oriented simplicial complex in which every simplex is labeled and its matrix representations of boundary operators and combinatorial Laplacian

Let assume that S_σ is an oriented simplicial complex, q is and integer with $0 \leq q \leq dim(S_\sigma)$, and let $\{\sigma_1, \sigma_2, ..., \sigma_n\}$ denote the q-simplices of simplcial complex S_σ, then from $\mathcal{L}_q = \mathcal{L}_q^{UP} + \mathcal{L}_q^{DN}$ is easy to deduce that

$$(\mathcal{L}_q)_{ij} = \begin{cases} deg_U(\sigma_i) + q + 1, & \text{if } i = j \\ 1, & \text{if } i \neq j \text{ and } \sigma_i \text{ and } \sigma_j \text{ are not upper adjacent but have a similar common lower simplex} \\ -1, & \text{if } i \neq j \text{ and } \sigma_i \text{ and } \sigma_j \text{ are not upper adjacent but have a dissimilar common lower simplex} \\ 0\ , & \text{if } i \neq j \text{ and } \sigma_i \text{ and } \sigma_j \text{ are upper adjacent or are not lower adjacent} \end{cases} \tag{2.2}$$

since $(\mathcal{L}^{UP})_{ii} = deg_U(\sigma_i)$ and $(\mathcal{L}^{DN})_{ii} = deg_L(\sigma_i)$. Detailed proof of above expressions can be found in [58].

The above definition of matrix elements is unhandy and impractical for applications of large simplicial complex, and hence we need to develop some computation-

ally convenient way to extract useful and meaningful information from combinatorial Laplacians. In that course, let us focus now on the eigenvalues and eigenvectors of q^{th} combinatorial Laplacian L_q. For an oriented simplicial complex S_σ and an integer q with $0 \leq q \leq dim(S_\sigma)$, the q^{th} *Laplacian spectrum* is denoted as $M(L_q(S_\sigma))$. It represents set of eigenvalues of $L_q(S_\sigma)$ together with their multiplicities and is independent on the choice of orientation of q-simplices in the complex S_σ. Since the q^{th} Laplacian matrix is positive semidefinite, all its eigenvalues are nonnegative. The null space of $N(L_q(S_\sigma))$ is the eigenspace of $L_q(S_\sigma)$ and corresponds to the zero eigenvalues. At this moment, as we promised, we can relate the combinatorial Laplacian and homology: the combinatorial Hodge theorem states that the q^{th} homology group $H_q(S_\sigma)$ is isomorphic to the null space of q^{th} combinatorial Laplacian [59], that is

$$H_q(S_\sigma) \cong N(L_q(S_\sigma)),$$

for each integer q with $0 \leq q \leq dim(S_\sigma)$. The multiplicity of zero eigenvalues of q^{th} combinatorial Laplacian is equal to the number of the q-dimensional holes in a simplicial combplex, i.e. Betti number. This is very useful expression providing a practical method for calculation of Betti numbers [60].

Chapter 3

Application and results 1

Application of aforementioned methods in the topological data analysis is performed on data originating from several different fields are presented in the subsequent Sections and include the social sciences [50, 61],3.13.2, traffic jamming [62]3.3, analysis of complex system based on the time series of collective charge transport [63]3.4 and neuroscience [64]3.5. All results demonstrate wide range of methods and features of data analysis of complex systems through applications of the algebraic topology and represents a significant contribution to the various fields of science.

3.1 Hierarchical sequencing of online social graphs

An online communications, patterns of conduct of individual actors and use of emotions in the process can lead to a complex social graph exhibiting multilayered structure and mesoscopic communities. Using simplicial complexes representation of graphs, we investigate in-depth topology of the online social network constructed from MySpace dialogs which exhibits original community structure. A simulation of emotion spreading in this network leads to the identification of two emotion-propagating layers. Three topological measures are introduced, referred to as the structure vectors, which quantify graphs architecture at different dimension levels. Notably, structures emerging through shared links, triangles and tetrahedral faces, frequently occur and range from tree-like to maximal 5-cliques and their respective complexes. On the other hand, the structures which spread only negative or only positive emotion messages appear to have much simpler topology consisting of links and triangles. The nodes structure vector represents the number of simplices at each topology level in which the node resides and the total number of such simplices determines what we define as the nodes topological dimension. The presented results suggest that the nodes topological dimension provides a suitable measure

of the social capital which measures the actors ability to act as a broker in compact communities, the so called Simmelian brokerage.Wealso generalize the results to a wider class of computer-generated networks. Investigating components of the nodes vector over network layers reveals that same nodes develop different socio-emotional relations and that the influential nodes build social capital by combining their connections in different layers.

3.1.1 Introduction

Structure of online social networks emerges via self-organizing processes of social dynamics, where the links are being established and used for communications between individuals. The contents (information, emotion) communicated between pairs and groups of participants affects their activity patterns and thus shape the networks evolution. Two prototypal classes of online social networks can be distinguished [65]: a hierarchically organized multi-layered structure that reflects the level of knowledge of the involved individuals in the chats-based systems, on one hand, and a wider class of networks with community structure, on the other. Recently, online social networks of both types have been studied based on high-resolution empirical data from a variety of Web portals [65]-[69].

In recent research, efforts have been made on determining the network complexity metric that permits to successfully distinguish between critical and redundant nodes [70], discover the active cores of the network as compared to the networks periphery [71] and quantify the networks multiplexity [72]-[75],[68, 76] and the role of higher-order structures in the network dynamics [77]-[79]. In this respect, several approaches have focused on introducing suitable graph-theoretic vectors that can be defined on local graphlets [80], the networks feature vector [81] or the graphs eigenvalue spectrum [82]. These approaches proved very useful in the study of biological systems, for example, in alignment of protein networks [83] and uncovering network function in cancer-related processes [84]. Similarly, characterizing topology of molecular graphs [85] as well as determining modules in socio-technological networks by eigenvectors localization [86] have been successful.

In this work, we exploit the topological concept of a simplex, structure that extends beyond the nodes and links, i.e., a polyhedron of possibly high dimension, and their aggregates or simplicial complexes. We investigate in-depth topology of online social networks and explore the role of nodes in layers and communities. The concept of simplicial complexes of graphs
citemyspace25,myspace26 allows precise definition, using topological, algebraic and combinatorial tools, of the nodes natural surroundings in the network. Consequently,

as we show in this work, it permits to measure the nodes social capital such as the Simmelian brokerage in social environments.

The considered network in this work is constructed from the original data collected from MySpace social network as described in Ref. [66]. Typically, in online social networks, such as MySpace and Facebook, certain kind of social graphs exist a priori. However, the use of connections over time as well as the dominance of positive emotions in the texts of messages [66] reveal the dynamical structure that is different from the conventional social networks.

In general, network layers appear due to different types of relationships among nodes [88]. In MySpace and Facebook social networks, where text messages of mixed information contents are communicated, the emotion contained in these words can be inferred [89, 66]. This fact offers the possibility to identify the networks layered structure in a unique manner. Specifically, one can define layers that propagate emotions with positive or negative valence (attractiveness and aversiveness). Quantitative study of emotions, based on Russells model [90], and the social dimension of emotional interactions are the subject of an intensive research in recent years [91]. In this context, the dynamics of emotion spreading on networks has been investigated by an agent-based model [92, 93]. Here, we employ this model to generate network layers propagating negative/positive emotion messages.

3.1.2 Simplicial complex analysis of the social network with communities

The components $\{Q_q^i\}$ for each node in the analyzed network can be computed from the MC matrix. Here we determine the components of each node of the network in Fig. 3.1(a). Sorting the nodes according to their topological dimensions $dimQ^i$, one can identify the influential nodes in the networks community or a layer. Fig. 3.2, shows the distribution of the nodes dimension for all nodes in the network as a function of the nodes rank. Notice that the topological dimensions exhibit a broad distribution (Zipfs law) with two rather than a single slope. Such situation often appears in the evolving complex systems [94, 95]. Recently, the origin of two slopes in the Zipfs law has been discussed [96] in connection with the scaling and innovation in the use of words in the written text of an increasing length. In the present case, the appearance of new topological forms in the dialogs-based network is related with the activity patterns of users (nodes in the network). According to the analysis in Ref. [66], three different groups of users can be distinguished considering the number of their actions in relation with the interactivity times. Consequently, very active nodes may build a larger environment resulting in a higher topological dimension.

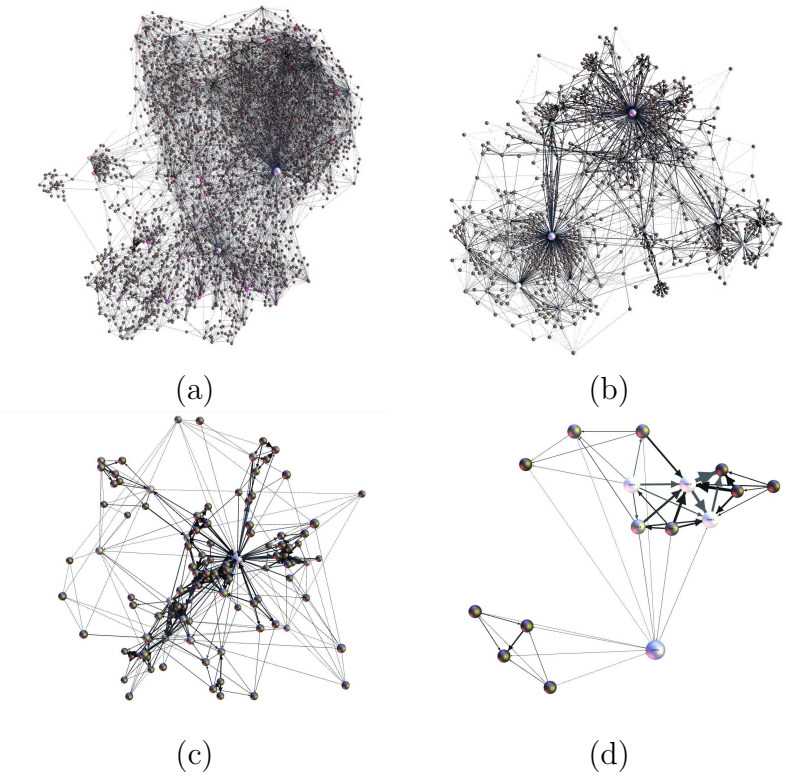
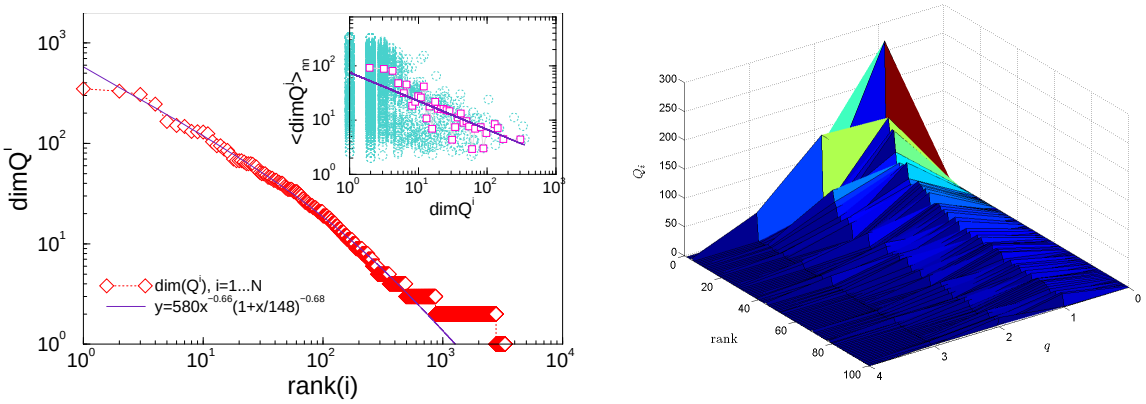


Figure 3.1: Online social network (OSN), constructed from MySpace dialogs, after removal of leaves (nodes with degree one) (a); Structure of the largest social community in that network at different topology layers corresponding to dimensions q = 2, 3 and 4 (b, c, and d, respectively).

In Fig. 3.2(left), topological dimensions of nodes obey a broad distribution with two slopes. The curve is fitted with the expression $f(x) = Ax^b(1 + x/c)^{-d}$, with the parameters $A = 580, b = -0.667 \pm 0.006, c = 148 \pm 59, d = 0.68 \pm 0.18$. In the following, we study correlations among topological dimensions of the connected pairs of nodes.

In an analogy with standard assortativity measure in social networks [97], we plot the nodes topological dimension against the average topological dimension of its neighbors. The results, shown in the inset of Fig. 3.2(left), indicate that at the level of triangles and cliques of higher dimension, the graph exhibits disassortativity. The general trend of all points can be approximated with the function $(dimQ^j)_{nn} \sim (dimQ^i)^{-0.52}$. This means that gradually fewer number of nodes with high dimension are connecting between structures of a smaller dimension. These findings complement the results of disassortativity found in Ref. [66] for the same network at the level of links (i.e., including leaves).



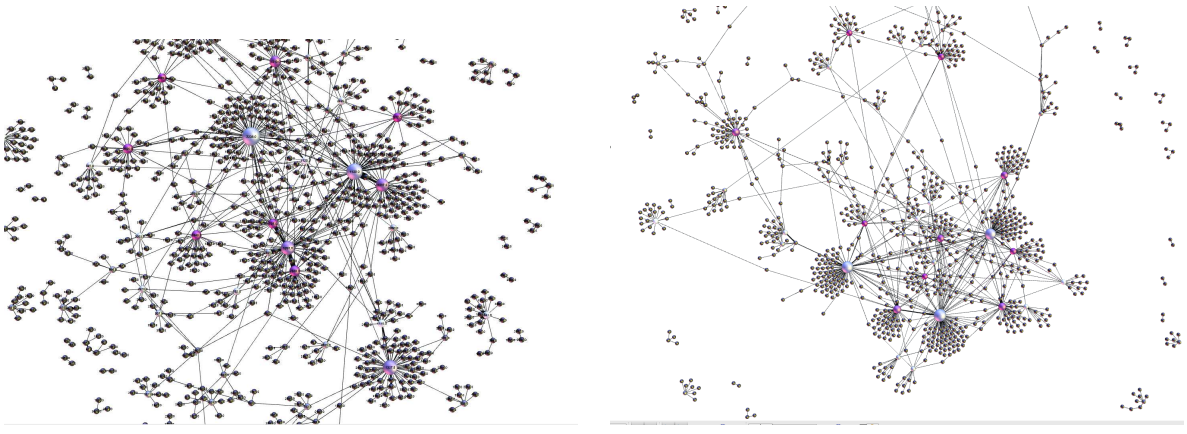
(a) (b)

Figure 3.2: (Left, main panel) Ranking plot (Zipfs law) of topological dimensions $dimQ^i$ of all nodes in the network. Inset: Average topological dimension of the nodes i nearest neighbors plotted against the nodes i dimension exhibiting disassortative behavior (slope of the fit line is $\mu = 0.527 \pm 0.026$). (Right) Components of the topology vector Q_i^q of the first 100 nodes plotted against the nodes rank and simplex dimension q. Ranking order of nodes according to the topological dimension $dimQ_i$ applies.

3.1.3 Structure vectors of emotion-propagating network layers

The emotion-propagation dynamics involves different types of contacts among individual actors in online social network. The diversity of the emotional content of communicated messages, described by two variablesemotional arousal and valence, enables to identify different network layers corresponding to a particular type of emotional content. Consequently, two network layers are recognized. The positive layer consists of the links along which messages with a positive emotion valence were communicated up to a given instance of time, and the negative layer with links carrying messages with a negative valence. In order to enhance the difference between these layers, we use the agentbased model of Ref. [92] to simulate the propagation of a specified emotion. In the model, agents are situated at nodes of the above described original network from MySpace data. Each agent is characterized by a fixed ID and two dynamical variablesarousal $a_i(t)$ and valence $v_i(t)$, that describe the agents emotional state. The agents activity profile is defined by the action delay probability $P(\Delta t)$ and the circadian cycle, which is incorporated by the driving noise $p(t)$. Both $P(\Delta t)$ and $p(t)$ are taken from the empirical data collected from MySpace. The action rules of real online social networks are implemented (see Ref. [92] for details). The agents messages are being sent along the network links; each

message contains the emotional state of the sender. Then the aggregated messages at the wall of the recipient agent affect its emotional state. In result, the variables $a_i(t)$ and $v_i(t)$ of the recipient node are updated where the influence of the input noise is also taken into account [92, 93]. Here, we numerically simulated two situations. In one, the external input noise has prevailing positive emotion astonished and, in the other, the negative emotion ashamed. As it was shown in Refs. [92, 98], in such situations the temporal correlations of message streams which mark the emotion propagation dynamics, lead to collective emotion states in networks. Here, the positive/negative valence of the input emotion eventually prevails. In this way, in each case one can distinguish a dominant layer that diffuses the winning emotion from the counter-emotion layer. These are named PP (positivepositive) and NN (negativenegative) as dominant layers in positive and negative input, respectively; counteremotion layers are NP and PN. Here NP designates a layer with links propagating positive emotion in the case when the majority of messages in the network are negative (following negative emotion input). While PN denotes a layer with the negative emotion links when the prevailing emotion is positive. Note that, by definition, the same nodes (and sometimes overlapping links between them) can belong to both layers. The focus is on in-depth-topology analysis of these layers and in quantifying the roles of relevant nodes residing in each one of them. The structure of connections in counter-emotion layers for the two cases of emotion-propagating dynamics is shown in Fig. 3.3. The results of topology analysis of all emotion propagating layers are summarized in Table 3.1. For a comparison, we also provide the standard network measures: average degree $\langle k \rangle$, clustering coefficient CC, diameter d, average path length $\langle p \rangle$, and the degree of modularity $modul$.



(a) (b)

Figure 3.3: Counter-emotion layers NP (left) and PN (right) of the online social network from Fig. 3.1(a).

Table 3.1: Standard network measures and three structure vectors of the layers in the OSN propagating negative and positive emotion messages when the majority of messages are of negative emotional content (NN and NP) and when the majority of messages in the network are of positive emotional content (PN and PP).

Layer	$\langle k \rangle$	CC	d	$\langle p \langle$	$Modul$	q	Q	N_s	$\hat{Q}$
						4	-	-	-
						3	4	4	0
$NN_\omega > 30$	4.04	0.051	13	5.11	0.62	2	198	198	0
						1	6242	6353	0.017
						0	1	6353	0.999
						4	-	-	-
						3	-	-	-
NP	0.69	0.046	21	6.24	0.84	2	12	12	0
						1	1134	1138	0.003
						0	2310	3404	0.321
						4	-	-	-
						3	-	-	-
PN	0.65	0.031	16	5.96	0.84	2	11	11	0
						1	1069	1074	0.005
						0	2360	3390	0.304
						4	-	-	-
						3	5	5	0
PP	4.05	0.051	14	5.03	0.63	2	204	205	0.005
						1	6228	6287	0.009
						0	4	6352	0.999

3.1.4 Simmelian brokerage and the nodes Q-vector

In order to estimate the social capital of nodes (users) in the network, we measure the Simmelian brokerage B_i for each node $i = 1, 2, ..., N_U$. According to Ref. [99], for a given node i Simmelian brokerage captures opportunities of brokerage between otherwise disconnected cohesive groups of contacts. Quantitatively, B_i is determined via the nodes efficiency E_i as [99]:

$$B_i = n_i - (n_i - 1)E_i,$$

where n_i is the number of neighbors of the node i inside a considered group N_i; the nodes local efficiency E_i is determined by

$$E_i = \frac{1}{n_i(n_i - 1)} \sum_{l \in N_i} \sum_{m \in N_i} \frac{1}{d_{lm}},$$

where d_{lm} is the distance between all distinct pairs of nodes $l \neq m$ in the set N_i when the node i is removed. As stated earlier in our approach, the node's neighborhood N_i is precisely defined at different topology levels by simplices and simplicial complexes in which the node i resides.

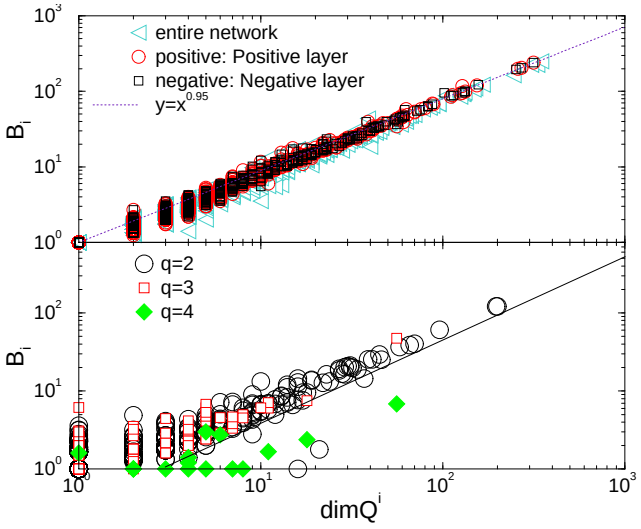


Figure 3.4: For the networks in Fig. 3.1, Simmelian brokerage B_i of nodes plotted against the nodes topological dimension $dimQ^i$. The symbols indicate social graphs at different topology levels $q = 2, 3, 4$ (bottom panel), and the entire network $q = 0$, and the two emotion-propagating layers (top panel).

Performing the computation indicated in two above equations, we determine Simmelian brokerage Bi for each node in the network shown in Fig. 3.1(a). In Fig. 3.4(a) Bi is plotted against the nodes topological dimension dim Qi, where

each point represents one node of the network. In addition, are plotted the nodes brokerage values obtained within the emotion-propagating layers. The bottom panel includes similar plots but only for the higher topological levels $q = 2, 3, 4$, matching the social graphs in Fig. 3.1(b), (c), (d), respectively. Note that the whole network corresponds to the simplicial complex at level $q = 0$. It is remarkable that in these plots the majority of nodes follow a universal pattern that can be expressed by functional dependence

$$B_i \sim (dimQ^i)^\mu,$$

where the exponent $\mu \leqslant 1$. Dispersion along vertical axis correlates with the number of higher-order cliques in the considered graph. we confirm the robustness of the functional dependence expressed by previous equation. Particularly, we present an approximate analytical expression (which is exact in some limiting geometries) as well as numerical work for a wider class of networks.

$$B_i = n_i - \frac{\sum_{l \in N_i} \sum_{m \in N_i} \frac{1}{d_{lm}}}{n_i}$$

This sum can be computed analytically in some limiting cases. For instance, consider the situation where the node i belongs to a $(q + 1)$-clique, which implies that its number of neighbors is $n_i = q$ and all distances within the clique are $d_{lm} = 1$. After removal of the node i, the remaining nodes contribute to the sum as $\sum_{l \in N_i} \sum_{m \in N_i} \frac{1}{d_{lm}} = q(q-1)$. A straightforward extension to the situation where the node i connects k such cliques leads to $\sum_{l \in N_i} \sum_{m \in N_i} \frac{1}{d_{lm}} = kq(q-1)$. In this case, removal of the node leaves the cliques separated from each other, i.e., the distance between pairs of nodes from different cliques is infinite, while the distance inside each clique is one. Note that in the topology analysis, the number of $(q+1)$ cliques related to the node i is given by the nodes q-level component, Q^i_q. Hence, the sum at q-level gives $\sum_{l \in N_i} \sum_{m \in N_i} \frac{1}{d_{lm}} = Q^i_q q(q-1)$. The situation is exact for top-level cliques, for example 5-cliques in Fig. 3.1(d). A similar reasoning can be extended to $(q-1)$-level, provided that at this level no shared faces occur between the q-level cliques. Hence, in this case we can rewrite Brokerage equation as

$$B_i = \sum_q Q^i_q q - \frac{\sum_q Q^i_q q(q-1)}{\sum_q Q^i_q q},$$

where we also note that ni can be expressed via the components of the topology vector as $n_i = \sum_q Q^i_q q$. By extending the summation over topology levels, we obtain an approximate expression that can be written as

$$B_i \simeq dimQ^i\langle q\rangle - \frac{\langle q^2\rangle}{\langle q\rangle} + 1,$$

where we used the definition about $dimQ^i$ and introduced an abbreviation $\langle q\rangle = \sum_{q=0}^{q_{max}} Q^i q / \sum_{q=0}^{q_{max}} Q^i$.

Note that in the case of tree structures, where the highest topology level corresponds to links, i.e., $q = 1$ is the maximal clique, we have that $B_i = dimQ^i$ with the exact exponent $\mu = 1$. In a more complex network cliques of higher order occur and are weakly connected at lower topology levels, as in the case of studied online social network (OSN in Table 3.3). Consequently, the contribution of the second term in last equation induces corrections eventually resulting with an exponent $\mu \leqslant 1$. The dispersion in the number of cliques in last equation and the number of their shared faces at lower q-levels can be considerably greater in the case of more compact networks. Nevertheless, a power-law dependence appears, with different values of the exponent μ. The reasons for the occurrence of such power-law dependence in a general network structure are not evident. Here, we provide a numerical proof by considering a wider class of networks.

The analysis in previous sections suggests that the relation between Simmelian brokerage of nodes and their topological dimension depends on the graph architecture. Therefore, by varying the building rules of the network, one can vary the topological dimensions of different nodes and test the robustness of brokerage equatin. We consider several types that are shown in Fig. 3.5. In these networks, the nodes neighborhood can be varied by control parameters of the growth, ranging from the tree-like to a highly clustered compact structure. These networks, consisting of approximately 1000 connected nodes, are generated by the algorithm that is initially described in Ref. [86] for growth of scale-free networks with clustering and communities. The basic idea of clustered scale-free networks by preferential attachment and preferential rewiring of Ref. [100] is implemented for the case where different communities (node groups) are allowed to grow. Thus, the attachment of new nodes is preferred within a currently growing community while rewiring can take part both within and outside of that community. Three parameters that control the structure are: $p$the probability of a new community, α and βthat appear in the preferential shift-linear rules for attachment and rewiring probabilities [100], respectively, and $M$the number of nodes added per growth step. In addition, we consider a dense single-community graph consisting of 100 nodes, shown in Fig. 3.5d. The results of the topology analysis of these networks is summarized in Table 3.3. Topological dimension $dimQ^i$ of each node in these networks is also determined. Then Simmelian brokerage is computed, according to the original formula, for each node and

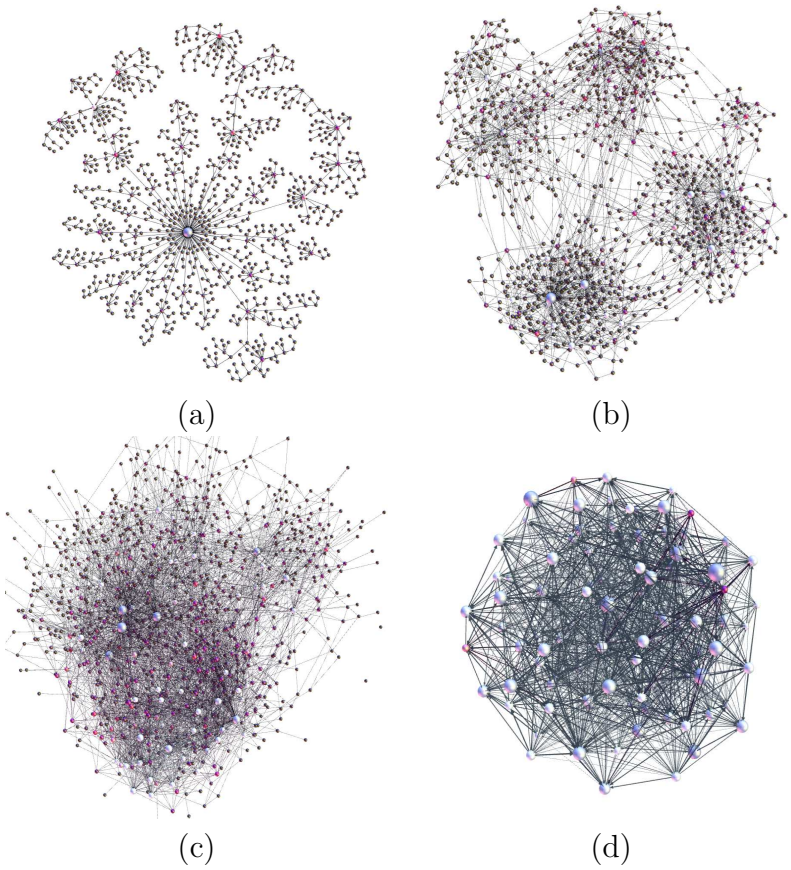


Figure 3.5: Computer-generated networks with different topology: (a) scale-free tree (SF-Tree); (b) network with clustered and weakly interlinked scale-free communities (SF-Comm); (c) same parameters as (b) but stronger connections between communities (CL-Comm); (d) a strongly clustered single community network (CL-Str).

plotted against the nodes topological dimension. The results are shown in Fig. 3.6. The powerlaw dependence holds for each network type in the corresponding range of nodes topological dimensions. Values of the corresponding exponent μ are also given in Table 3.3.

3.2 Topology of innovation spaces in the knowledge networks

The communication processes of knowledge creation represent a particular class of human dynamics where the expertise of individuals plays a substantial role, thus offering a unique possibility to study the structure of knowledge networks from online data. Here, we use the empirical evidence from questions-and-answers in mathematics to analyse the emergence of the network of knowledge contents (or tags) as the individual experts use them in the process. After removing extra edges from

Table 3.2: Standard network measures and the q-components of the three structure vectors for online social network from the MySpace data (OSN) and four computer-generated network. Error bars for the exponent are within ±0.02.

Graph	$\langle k \rangle$	*CC*	*d*	$\langle p \rangle$	*Modul*	*q*	*Q*	N_s	$\hat{Q}$	μ
SF-Tree	1	0	20	6.46	0.92	1	998	998	0	
						0	1	998	0.99	1
SF-Comm	1.94	0..201	9	4.42	0.40	3	2	2	0	
						2	328	328	0	
						1	1322	1602	0.175	
						0	10	1611	0.994	0.974
CL-Comm	3.83	0.081	7	3.36	0.40	4	16	16	0	
						3	65	80	0.188	
						2	837	900	0.07	
						1	2269	3090	0.269	
						0	12	3101	0.996	0.898
CL-Str	14.5	0.269	2	1.71	0.21	5	9	9	0	
						4	264	266	0.008	
						3	1302	1522	0.145	
						2	367	1887	0.806	
						1	1	1887	0.995	
						0	1	1887	0.995	0.442
OSN	2.30	0.183	8	4.07	0.74	4	5	5	0	
						3	90	91	0.011	
						2	1064	1103	0.035	
						1	5397	6437	0.161	
						0	1	6437	0.999	0.969

Table 3.3: Standard network measures and the q-components of the three structure vectors for online social network from the *MySpace* data (OSN) and four computer-generated networks. Error bars for the exponent are within ±0.02.

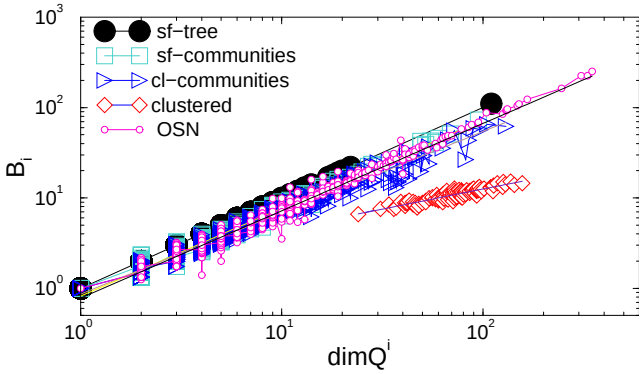


Figure 3.6: Brokerage versus topological dimension of nodes for computer-generated network types and for OSN MySpace, as indicated. Network characteristics are summarized in Table 3.3.

the network-associated graph, we apply the methods of algebraic topology of graphs to examine the structure of higher-order combinatorial spaces in networks for four consecutive time intervals. We find that the ranking distributions of the suitably scaled topological dimensions of nodes fall into a unique curve for all time intervals and filtering levels, suggesting a robust architecture of knowledge networks. Moreover, these networks preserve the logical structure of knowledge within emergent communities of nodes, labeled according to a standard mathematical classification scheme. Further, we investigate the appearance of new contents over time and their innovative combinations, which expand the knowledge network. In each network, we identify an innovation channel as a subgraph of triangles and larger simplices to which new tags attach. Our results show that the increasing topological complexity of the innovation channels contributes to network's architecture over different time periods, and is consistent with temporal correlations of the occurrence of new tags. The methodology applies to a wide class of data with the suitable temporal resolution and clearly identified knowledge-content units.

3.2.1 Introduction

The knowledge creation through online social interactions represents an emerging area of increased interest both for technological advances and the society [101] where the collective knowledge is recognised as a social value [102]-[104]. Recently studied examples include the knowledge accumulation in systems with direct questions-and-answers [105], crowdsourcing scientific knowledge production [106, 107] and scientific discovery games [108]. Similar phenomena can be observed in business/economics-associated online social networking [109]-[111]. On the other hand, the study of the collective knowledge creation opens new topics of research interests. In partic-

ular, it provides ground to examine a novel type of collective dynamics in social systems in which each actor possesses certain limited expertise. In the course of the collaborative social efforts to solve a problem, such as communications through questions-and-answers that we consider here, the tacit knowledge and the expertise of individual actors are externalised and dynamically shared with other participants who take part in the process. When a systematic tagging applies to the shared cognitive contents, the process leads to an explicit knowledge [103] as the output value (the network of knowledge contents), from which others can learn. Furthermore, the dynamics underlying knowledge creation exemplifies multi-scale phenomena related to the cognitive recognition, which may occur in a wider class of systems, social, biological and physical [112].

By the nature of the underlying stochastic processes, the knowledge networks that emerge through the collaborative social endeavours necessarily reflect the expertise and the activity patterns of the involved participants. Furthermore, these networks tend to capture the logical relationship among the used cognitive contents as it resides in the mind of each participating individual. In this regard, these networks substantially differ from the commonly studied knowledge networks, which are produced in ontological initiatives [113]-[115] such as those from the online bibliographic data and Wikipedia, or the mapping citation relationships between journal articles [116], to name a few. Also, the stochastic process of knowledge creation through questions and answers are different from the spreading dynamics of scientific memes, whose inheritance patterns are identified in citation networks [117].

In recent work [105], we have shown that the knowledge creation by questions-and-answers involve two-scale dynamics, in which the constitutive social and cognitive elements (individual experts or actors and the knowledge contents that they use) interact and influence each other on the original scale. This complex system evolves in a self-organised manner leading to the emergence of socio-technological structures where the involved actors share the accumulated knowledge. These structures are visualised as communities on the related bipartite network of actors and their artefacts [105]. Furthermore, the advance of innovation in this process, which builds on the expertise of the involved participants, leads to the expansion of the knowledge space by adding new cognitive contents. The central question for the research and applications of the collective knowledge creation is how these stochastic processes work and potentially can be controlled to converge towards the desired outcome. Furthermore, what is the structure of the emergent knowledge that can be used by others? A part of the answer relies on the structure of the networks, co-evolving with the knowledge- sharing processes among the actors possessing the required

expertise. In [105] the empirical data from the Stack Exchange site Mathematics (http://math.stackexchange.com/) were downloaded and analysed, as a prototypal example. The sequence of events in the process of questions- and-answers ($Q\&A$) suitably maps onto a growing bipartite network of actors, as one partition, and their questions and answers, as another partition. The emergent communities on these networks have been identified, consisting of the involved actors and the connected questions- and-answers. As a rule, in each community a dominant actor is found, representing an active user with a broad expertise. The knowledge elements of each question are specified according to the standard mathematical classification scheme by one to five tags (for instance, functional analysis, general topology, differential geometry, abstract algebra, algebraic number theory). Consequently, the expertise of the actor can be specified as a combination of tags that the actor had frequently used. Assuming that a minimal matching applies among the actors expertise and the contents of the answered question, and using theoretical modelling based on the empirical data, it was shown [105] that the emergent communities and the knowledge that they share strongly depend on the population of the involved experts and their activity patterns.

In this work, using the same empirical dataset, our focus is on the networks of cognitive elements (tags) that emerge in these processes with questions-and-answers. Different from the aforementioned bipartite networks, these emergent knowledge networks contain subelements of both partitions, namely, knowledge contents of questions as well as a measure of the users expertise. Such networks, supported by the current information and computer technology (ICT) systems, embody the collective knowledge that emerges via the cooperative social efforts and can be used by others to learn. Moreover, the relevance and speed of knowledge acquisition from these networks may be more efficient than from the networks generated through widescale ontological plans and efforts. We apply the techniques of algebraic topology of graphs [118]-[122] to investigate higher-order structures that characterise the connection complexity between knowledge elements in the emergent networks. Specifically, we aim to determine:

- the metrics to quantify the higher-order combinatorial structures which contain the logical units of knowledge as the actors use them in communication;
- the role of innovative contents brought over time by the experts in building the network architecture.

In addition to the standard graph-theoretic metrics and community detection in the emergent networks of knowledge units, we describe their hierarchical organisa-

tion using several algebraic topology measures. Further, we identify the appearance of new tags over time and investigate the subgraphs (innovation channels) where these new cognitive elements attach to the existing network. By tracking topology measures over the consecutive time periods for the innovation channel together with the topology of the entire network, we quantify the impact of the new-added contents. Our main findings indicate that the networks of cognitive elements map to a nontrivial hierarchical architecture which contains aggregates of high-order cliques. The increasing structural complexity of these networks over time, owing to the innovation expansion, is consistent with the logical structure of knowledge that they contain and temporal correlations in the appearance of new cognitive contents. In the following, the networks of tags are built from the empirical data for four successive one-year periods. At the initial stage, the networks are filtered to remove redundant links. At the next stage, network measures are obtained at the graph level, and the community structure is determined. At the final stage, the algebraic topology analysis of these networks for different periods and filtering levels is performed. The analysis is focused on the subgraphs, which are related to the appearance of new tags, representing the innovation channels of these networks.

3.2.2 Emergence of the tags networks - The $Q\&A$ process and structure of the empirical data

In this work, we have constructed knowledge networks from the empirical data, which are collected and described in Ref. [105]. In the data, the knowledge contents are mathematical tags used in the communications on $Q\&A$ system Mathematics Stack Exchange. In particular, the content of each question is specified (tagged) by one or more (maximum five) tags according to the standard mathematical classification scheme.While in Ref. [105] we investigated the role of expertise in the social process taking part on the co-evolving bipartite network of users-and- questions, here we focus on the network of tags as the elementary units of knowledge that are used by the actors in this process. With the help of the agent-directed modeling, in Ref. [105] we have demonstrated that the considered empirical process obeys the fundamental assumption of knowledge creation, i.e., that at least minimal matching between the contents of the question and the expertise of answering actor occurred in each event. Therefore, the emergent network of tags reflects the way in which these knowledge units are used in the process and, indirectly, the expertise of the social community. Moreover, the architecture of the emergent network of tags is expected to mirror the logical structure of knowledge, as it is presented by the experts involved in the knowledge-creation process.

To be consistent with the previous studies and the associated analysis of Ref. [105], we use the same dataset that was downloaded on May 5, 2014, from https://archive.org/details/stackexchange and contains all user-contributed contents on Mathematics since the establishment of the site, July 2010, until the end of April 2014. Specifically, the considered dataset contains 269818 questions, posted and answered by 77895 users, 400511 answers, and 1265445 comments. For the present analysis, from the available high-resolution data we use the information about questions, i.e., ID of each question, its content as a list of tags, and time stamp. The tags and their combinations define the knowledge landscape whose size is not constant but increases with time and the number of posted questions. In this way, the innovation increases as the key feature of the collective knowledge creation [105]. By investigating the network of tags, here we examine how the knowledge creation can be expressed by the topological complexity of the expanding knowledge landscape.

Mapping data to networks of tags is performed within four consecutive periods; a period is one-year long. First, the questions that are posted within the considered year period are selected, and a unique set of tags that are involved in these questions is formed. Each tag represents a node of the tags network. Two tags (i, j) are linked by multiple connections w_{ij}, where the link multiplicity $w_{ij} = 0, 1, 2, ...$ represents the number of common questions in which the considered pair of tags appeared in the selected dataset. The resulting networks are termed tagNetY-k, where $k = 1, 2, 3, 4$ indicates the considered year period.

3.2.3 Graph measures of tags networks without redundant connections

The raw networks of tags contain a large number of redundant connections leading to a large density graph, cf. an example in Fig 3.7. To move forward, we first apply an advanced procedure to eliminate the potentially redundant links.

Filtering redundant connections in a network of tags is motivated by the following facts. In the data, the number of tags is between 500 and 1000 while the number of posted questions per year are between 15 and 120 thousand, which results in a quite dense network of tags. On the other hand, a broad distribution of the tags frequencies [105] suggests that a relatively small number of tags occurs quite frequently. Among the most frequent tags are homework, proofwriting, reference-request, and terminology, which are not related to any particular field of Mathematics but rather determine the type of question asked. For this reason, these tags can occur in many different combinations of tags, thus increasing the networks density. Here, we apply an algorithm to decrease the networks density by identifying the edges that do not

incur as a result of a random process. For this purpose, the weighted network is considered as a multigraph where the weight w_{ij} represents a multiplicity of links between the pair of nodes (i, j). We apply the filtering technique described in Ref. [123]; it utilizes a random configurational model for weighted graphs that preserves the total weight of the realised links,$W = \sum_k s_k$, as well as the nodes strength $s_k = \sum_j w_{ij}$ on average. To avoid the influence of the filtering on higher structures, we apply the algorithm to each link independently.

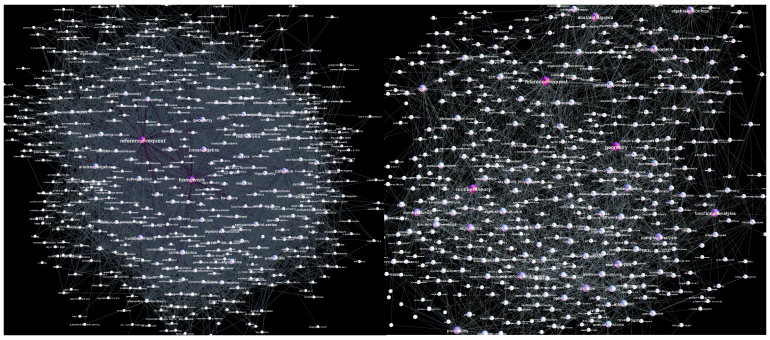


Figure 3.7: The network tagNetY-1: a close-up of unfiltered network near some large nodes (left) and the whole network filtered at confidence level $p = 0.1$ (right).

A pair of nodes (i, j) is selected proportionally to their strengths s_i and s_j. In the considered network, the selected pair is connected by the weighted link of the multiplicity w_{ij}. In the random configurational model, the occurrence of a link with multiplicity m between the selected pair of nodes is given by the conditional probability

$$P_{ij}(m|s_i, s_j, W) = \binom{W}{m}(\frac{s_i s_j}{2W^2})^m(1 - \frac{s_i s_j}{2W^2})^{W-m}$$

Then the probability that the realised weight w_{ij} of the link (i, j) occurred by chance (p-value) according to the marginal distribution given by last equation is computed as [123]

$$P_r(w_{ij}) = \sum_{m \geq w_{ij}} P_{ij}(m|s_i, s_j, W)$$

The links for which the probability $P_r(w_{ij})$ appears to be larger than a preset confidence level p are removed. The remaining edges, which satisfy the condition $P_r(w_{ij}) \leq p$, represent the filtered network with the specified confidence level. Here we examine the structure of the filtered networks obtained for several values of the parameter, $p \in \{0.1, 0.05, 0.01\}$. As an example, the right panel in Fig 3.7 shows the first year network after the filtering procedure with the confidence level $p = 0.1$.

The networks of tags for different periods and filtered at various confidence levels are analyzed by algebraic topology techniques, as presented in the following Sections. In this regard, we turn the weighted networks into binary graphs, which retain all important topological features of the weighted graphs while making the computation less demanding. Here, we first show that the filtering process leads to a reduced-density graph but preserves the relevant (nonrandom) connections. Specifically, the thematically connected groups of nodes (cf. labels of nodes in Figs 3.7 and 3.8) appear to form distinct communities on the network. In these networks, mostly non-overlapping communities occur. Consequently, they are suitably identified by methods based on the optimisation of the modularity [124]-[126]. A module is recognised as a densely connected group of nodes that are sparsely connected to nodes in other groups [127]. For a better comparison of different networks, the communities are systematically determined at the same resolution parameter (standard resolution 1.0 in *Gephi*, the open graph visualization platform http://gephi.org). This large-scale clustering of the knowledge networks appears systematically during the network growth. See also the structure of innovation channels studied in the following Section.

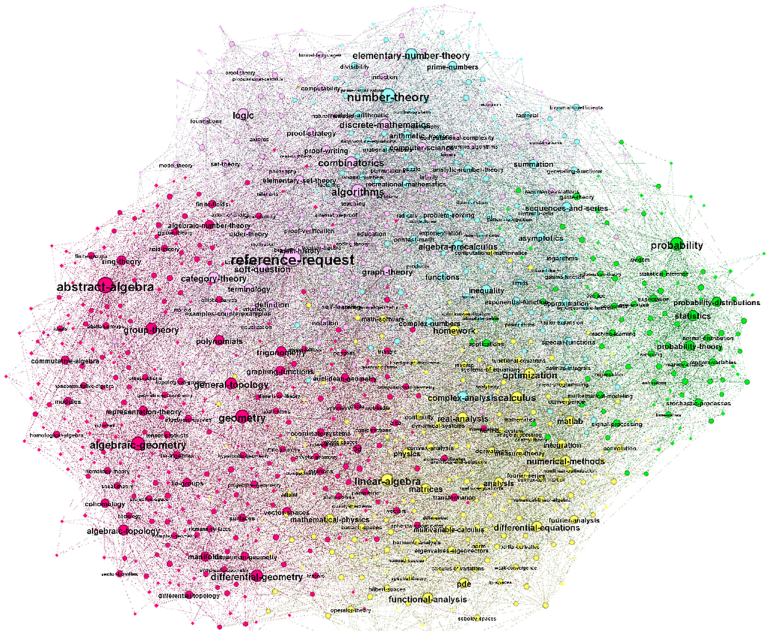


Figure 3.8: The community structure of the network of tags for the fourth period, which is filtered at $p = 0.1$. In each community, the mutually connected cognitive contents (mathematical tags) are indicated by the nodes labels.

For comparison, in Table 3.9 we summarise the standard graph-theoretic measures [127] of the networks of tags for four consecutive periods and the confidence level $p = 0.1$. Note that the network of tags grows over years by the appearance

of new tags, but also shrinks by the number of tags that appeared in the previous period and were not used in the current period.

Net	N	$\langle k \rangle$	$\langle \ell \rangle$	d	Cc	ρ	M
TagNetY-1	582	10.07	3.02	6	0.365	0.018	0.439
TagNetY-2	702	14.45	2.86	5	0.365	0.021	0.441
TagNetY-3	856	20.09	2.72	5	0.351	0.023	0.436
TagNetY-4	1033	22.52	2.68	5	0.338	0.022	0.422

Figure 3.9: The number of nodes N, average degree $\langle k \rangle$, average path length $\langle l \rangle$, diameter d, clustering coefficient Cc, graph density $\rho = \frac{L}{N(N-1)}$, and modularity M, where the summation runs over different communities.

3.2.4 Topology of the tags networks

In addition to the standard graph-theoretic analysis, cf. Table 3.9, we apply techniques of algebraic topology to determine simplices and simplicial complexes, which describe higher order structures of these networks. Definitions and detailed explanation of topological quantities used in this presentation may be found in Ref. [119] and references within. The simplices are identified as maximal cliques of all orders, i.e. dimensions. Then the topological complexity of the simplicial complex constructed from the complex network is quantified by the number of cliques at each topological level (dimension) q, starting at $q = 0$ up to the $q_{max} - 1$. A clique at level q = 0 is an isolated node while $q = 1$ is a link, $q = 2$ is a triangle and so on up to the level $q_{max} - 1$ representing the largest clique found in the network.

To demonstrate the relevance of nodes, we compute the topological dimension of each node in the original and filtered network of tags for the first-year interval, which are shown in Fig 3.7. The components at each q level of the top 40 nodes (tags), ordered according to their topological dimension, are displayed by three-dimensional plots in Fig 3.10. As this figure shows, the applied elimination of the links reduces not only the nodes topological dimension but also changes the structures at q-levels where the considered node is present. Consequently, the ranking order of a particular node can be changed (see the corresponding lists of nodes in Table 3.12), which is compatible with the altered importance of that node in the filtered network.

We further compare the role of individual nodes in the networks evolving over time, which are filtered at different confidence levels, i.e., $p = 0.1$, $p = 0.05$ and $p = 0.01$. We determine the topological dimensions of all nodes in the corresponding filtered networks for the four successive year-periods. The ranking distributions of the nodes topological dimensions are displayed in Fig 3.11a. This Figure shows that, first, nodes with a gradually higher topological dimension appear at later

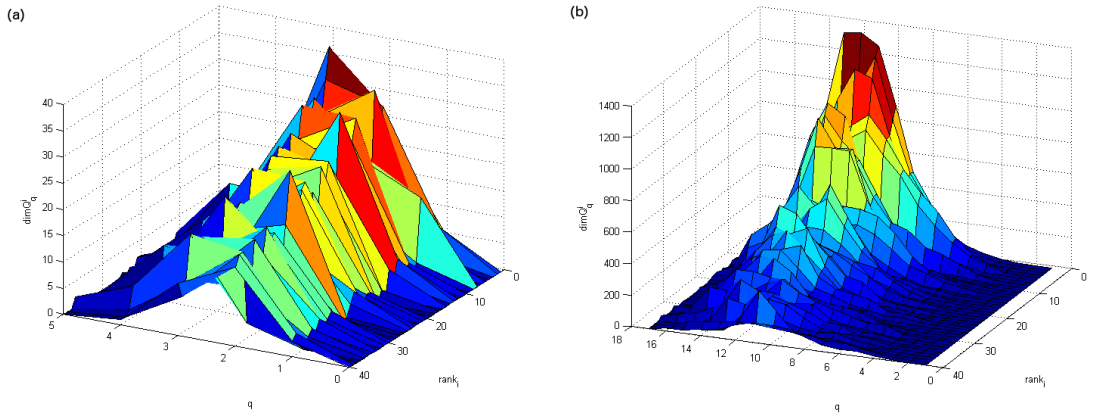


Figure 3.10: Components Q_q^i of the first 40 tags ranked by their topological $dimQ^i$ for the tagNetY-1 network filtered at $p = 0.1$ (a) and with no filtration (b).

periods, suggesting that topological complexity of tags networks increases over years. Furthermore, within each year, the reduced confidence level p results in a simpler structure of the nodes neighborhood (and possible shifts in the ranking order of nodes, as

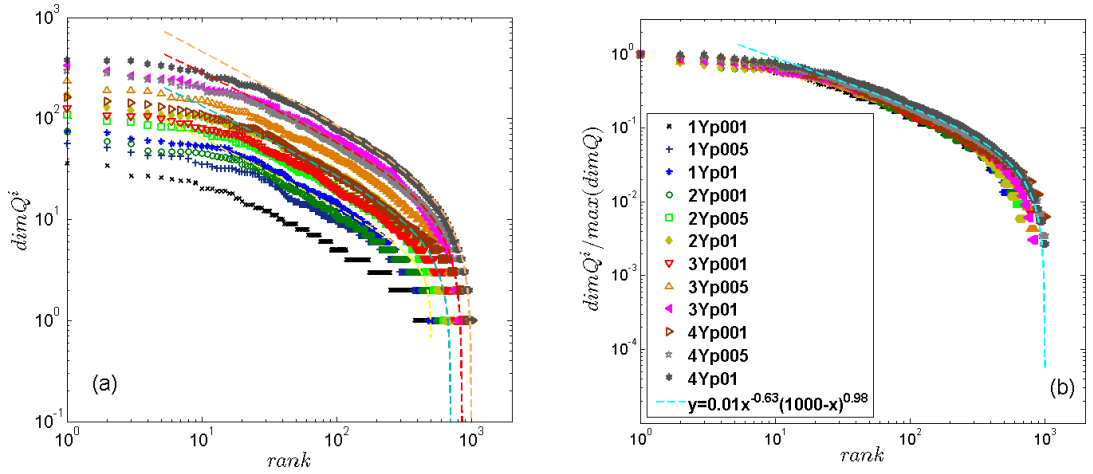


Figure 3.11: Ranking distributions of the topological dimension of tags $dimQ^i$ for all years and all p values (a) and scaled distribution $dimQ^i/max(dimQ^i)$ of all data (b). The legend abbreviations: 1Yp001 indicates the first-year network filtered at the level $p = 0.01$, and so on. Fit lines are according to the discrete generalised beta function; in panel (a) the parameter $b = 0.67 \pm 0.03$ and c varies from 0.32 for 1Yp001 and 0.71 for 2Yp001 to 0.82 for 3Yp001 and 4Yp001, with error bars ± 0.03.

mentioned above). However, all networks exhibit a broad ranking distribution of the nodes topological dimension with a power-law section. The distributions are fitted by the discrete generalised beta function

$$f(x) = ax^{-b}(N + 1 - x)^c$$

before filtering	$p = 0.1$	$p = 0.05$	$p = 0.01$
calculus	number theory	number theory	geometry
linear algebra	geometry	geometry	number theory
analysis	algebraic topology	functional analysis	calculus
homework	combinatorics	sequences and series	algorithms
reference request	abstract algebra	combinatorics	functional analysis
probability	functional analysis	algebraic topology	abstract algebra
sequences and series	algebra precalculus	abstract algebra	reference request
geometry	group theory	differential geometry	real analysis
functions	real analysis	calculus	algebra precalculus
real analysis	differential geometry	real analysis	logic
combinatorics	logic	probability	sequences and series
abstract algebra	sequences and series	algebraic geometry	probability
number theory	soft question	algebra precalculus	linear algebra
terminology	probability	algorithms	algebraic topology
complex analysis	integration	soft question	combinatorics
general topology	algorithms	logic	complex analysis
category theory	complex analysis	analysis	dierential geometry
algebraic geometry	analysis	complex analysis	soft question
discrete mathematics	differential equations	integration	discrete mathematics
logic	calculus	differential equations	analysis

doi:10.1371/journal.pone.0154655.t002

Figure 3.12: Names of the first twenty tags ordered according to their topological dimension in the network of tags before filtering and after filtering at the indicated confidence level p has been performed.

with different parameters a, b and c. The robustness of the observed scaling feature is further confirmed by the scaling collapse of all curves to a master curve, shown in Fig 3.11b. The scaleinvariant ranking, where the nodes topological dimension is scaled by the maximal dimension found in the corresponding network, suggests that the relative topological complexity of the tags networks is preserved over time and the degree of filtering.

3.2.5 Topological spaces in the filtered networks of tags

Fig 3.13 summarizes the components of two structure vectors for the tags networks emerging over different periods and varied filtering level p.

By comparing the curves for different one-year periods but fixed filtering level, say $p = 0.1$, we observe that the network topological complexity increases over time. It manifests in the increased number of connectivity classes (components of the first structure vector) at all topological levels as well as the shift of the maximum from $q = 2$ (triangles), in the first year, to $q = 3$ (tetrahedra) and $q = 4$ (5-cliques), in the fourth year. At the same time, we observe that the number of topological levels increases as well as the connectivity among the cliques at each topology level, cf. the third structure vector in the Fig 3.13b.

On the other hand, by decreasing the filtering confidence level p, a more sparse network is obtained having a smaller number of topological levels and a reduced number of simplicial complexes. However, they proportionally preserve the above-

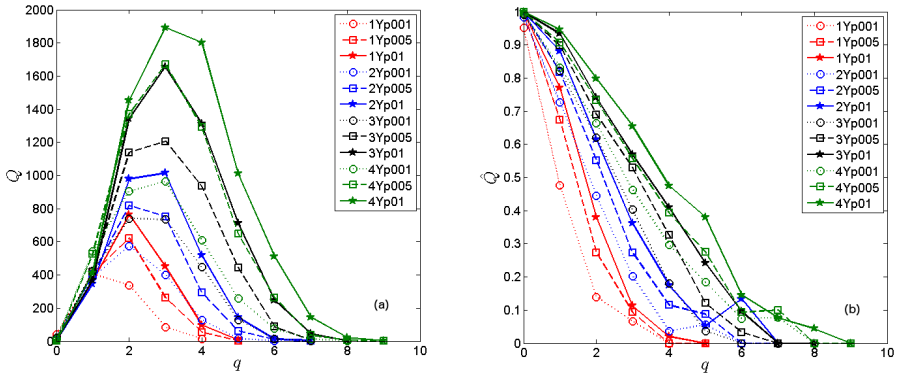


Figure 3.13: The components of (a) the first structure vector Q_q and (b) the third structure vector $\bar{Q}$ plotted against the topology level q for each year period and three filtering levels $p = 0.1, 0.05, 0.01$.

described tendency of the enhanced complexity of combinatorial spaces over time. The corresponding curves for $p = 0.05$ and $p = 0.01$ are also shown for each year-period in Fig 3.13. According to the structure vectors in Fig 5, all filtered networks exhibit a systematic shift towards richer topology in later years. Once again, these results confirm the structural stability in Fig 4 of the emergent networks of tags, which complements the logical organisation of knowledge contents in the communities in these networks, demonstrated in Fig 3.14 and in the following section.

3.2.6 Clustering of the innovative contents - Three aspects of innovation in the knowledge creation

The innovation growth [105, 128] is a crucial element of the process of knowledge creation. In the voluntary system, the innovation that comes from the expertise of the actors involved in the process was shown [105] to expand the knowledge space over time. To quantify the impact of innovation onto the architecture of the emerging knowledge networks, we consider the following three aspects of the innovation:

- the appearance of new tags due to the actors expertise;
- the occurrence of new combinations of tags expanding the knowledge space;
- the emergence of new combinatorial topological structures enriching the architecture of the knowledge network.

In the following, we discuss in detail these features of innovation.

Fig 3.14a contains time sequence of the first appearance of tags that are present in the data of each one-year period. Naturally, the sequence for Year-1 is the shortest,

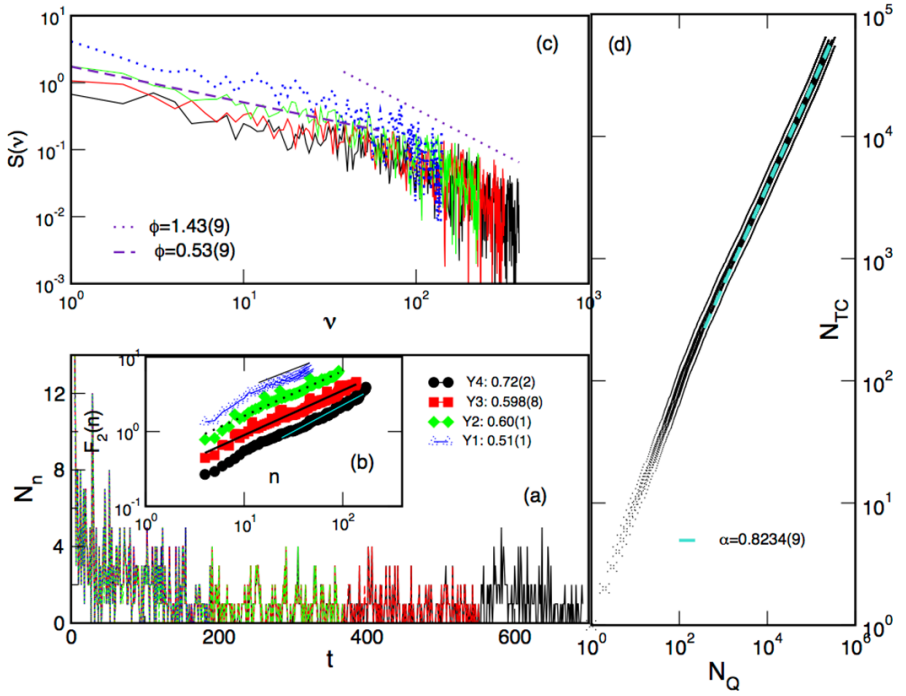


Figure 3.14: The temporal sequence of the appearance of new tags present in the networks for Year-1, Year-2, Year-3 and Year-4 periods (a). Temporal resolution is two days. The scaling of the standard fluctuation function (b) and the power spectrum (c) of these time series. Panel (d) displays increase in the number of new combinations of tags as a function of the number of questions over time.

while the sequence for Year-4 is the longest, since some tags that are present in Year-4 appeared in the earlier periods. The time series contains the number of new tags appearing in the sequence of two-day time intervals. The fractal analysis of these time series and their power spectrum, shown in Fig 3.14b and 6c suggest that the appearance of new tags is not random but exhibits long-range temporal correlations. Specifically, the plots in Fig 6b represent the fluctuation function $F_2(n)$ of the standard deviations of the integrated time series at the interval of length n. They reveal scaling regions (of different length for each time series) which permit determination of the Hurst exponent via $F_2(n) \sim n^H$. Values of the Hurst exponent H indicated in the legend suggest the fractal structure of the fluctuations. It appears that the fractality increases over time from nearly random time series with $H = 0.51 \pm 0.01$ in Year-1, to strongly persistent fluctuations with $H = 0.72 \pm 0.02$, in Year-4.

Similarly, power spectra of these time series in Fig 3.14c exhibit long-range correlations according to $S(\ni) \sim \ni^{-\phi}$ with two distinct exponents in high and low frequency regions. While the low-frequency feature is similar for all considered periods, the pronounced scaling in the highfrequency region gradually builds over years.

The number of unique combinations of tags was examined in the whole dataset and plotted against the number of posted questions in Fig 6d. The plot exhibits a power-law behaviour $N_{TC} \sim N_Q^{\alpha}$ in the range above 10^2 posted questions. It represents the Heaps law which appears to be in agreement with the ranking distribution of frequencies of the unique combinations of tags, i.e., the Zipfs law, as discussed in [5]. The occurrence of Heaps law is a manifestation of the innovation growth [105, 128] in the process of $Q\&A$. The exponent $\alpha < 1$ indicates a sublinear growth of innovation with the number of posted questions. This dependence suggests that a fraction of displayed items brings new combinations of tags while the remaining questions use the already identified combinations.

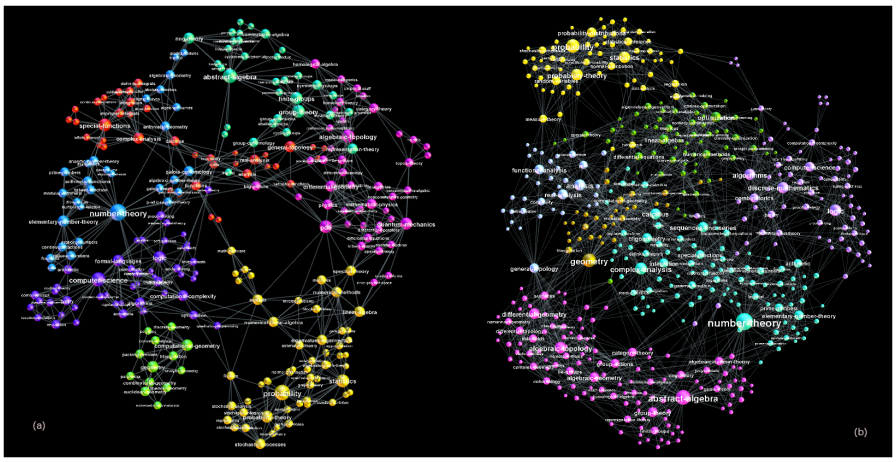


Figure 3.15: The structure of the innovation channel at the beginning of Year-2 (left) and Year-3 (right). New tags were added to the filtered tags network of the previous year, forming structures of higher dimension than a triangle. Communities of well-connected nodes show the logical grouping of mathematics subject categories, indicated by labels on nodes.

3.2.7 The structure of innovation subgraphs

The appearance of new tags in the $Q\&A$ process leads to the expansion of the knowledge network. In particular, the network grows by the addition of new nodes (cf. Table 3.9), as well as by increasing its topological complexity measured by the presence of simplicial complexes of a high order. In the remaining part of this section, we investigate how the new tags attach to the existing nodes and affect the formation of higher order structures in the knowledge network. For this purpose, we first define an innovation channel as a subgraph related with the new tags appearing at the end of a considered one-year period. Specifically, the subgraph in the network (filtered at $p = 0.1$) contains newly added tags together with the tags to which they attach

and form simplices larger than a single link (i.e., triangle or higher dimensional structure). The two plots in Fig 3.15 show the structure of the innovation channels at the beginning of Year-2 and Year-3 periods, respectively.

The innovation channels in Fig 3.15 grow over a one-year period; moreover, the innovative nodes stick with the rest of the network (previously existing nodes and links) making with them a tight structure that involves higher-order combinatorial spaces up to the largest clique. The community structure in the innovation subgraphs, which is demonstrated in Fig 3.15, reflects the thematic grouping of the entire knowledge network, as presented in Fig 3.8. For example, the newly added tag cohomology sticks to the group where we also find algebraic topology, differential geometry, abstract algebra, complex geometry and other thematically related tags, cf. the lower left community in Fig 3.15 right panel. On the other hand, the added tag computational complexity links to the community with discrete mathematics, algorithms, logics, combinatorics, computer science and others, cf. the rightmost community in the same Figure. Similarly, the node labels in all identified communities confirm their thematic closeness. Therefore, the expansion of the knowledge network by the addition of innovative contents systematically obeys the overall logical structure of (mathematical) knowledge. As mentioned earlier, the core of this feature of knowledge creation lies in the crucial role of the actors expertise in the process of meaningful cognitive-matching actions. The logical structure of individual knowledge of each actor gets externalised during the process of $Q\&A$.

According to the results in Fig 3.14, the appearance of innovative contents boosts the process of knowledge creation, leading to the observed temporal correlations, characteristic of collective dynamics. Analogously, here we show that the structure of innovation channels enriches the topological spaces of the knowledge network. In Figs 3.16 and 3.17 we summarise the topological measures of the innovation channels and compare them with the corresponding measures of the entire network. In addition to the structure vectors defined in structural vectors' equations, here we also consider the topological response function f_q to express the shifts in the topology at each level q in response to the changes in the network size. Formally, f_q is defined [120] as the number of simplices and shared faces at the level q.

Interestingly enough, the third structure vectors in Fig 3.16a and 3.16c show that the corresponding channels exhibit a better connectivity up to the level $q = 4$ of 5-clique than the background network. This feature of the innovation channels suggests the leading role of the innovative tags in the observed increase of the topological complexity of the network over years. This conclusion compares well with the number of connectivity classes at different topological levels, namely the first

structure vectors in Fig 3.16b and 3.16d. The topology of the channel determines the most ubiquitous structure in the entire network, corresponding to the peak in the first structure vector. Furthermore, the increase of the topological complexity of the knowledge graphs over consecutive periods is illustrated by the topological response function f_q, which is shown in Fig 3.17. It manifests in the increase of the number of topology levels, as well as the number of simplices and shared faces at each topology level. Also, the maximum of the function f_q shifts towards more complex structures, i.e., from triangles at Year-2 to tetrahedra in Year-4. As the plots in the lower panel of Fig 3.17 show, these topological shifts in the networks of different periods are tightly reflected in the structure of the corresponding innovation channels.

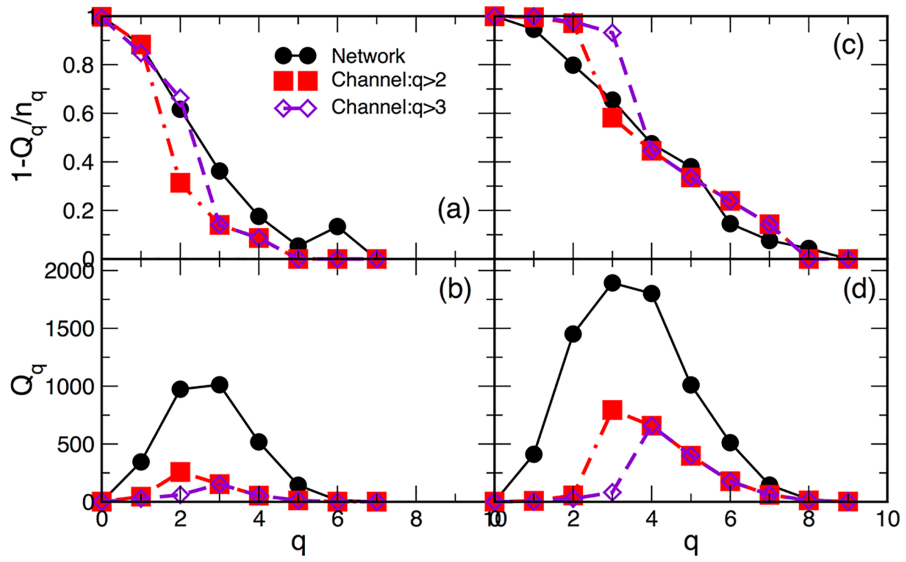


Figure 3.16: (a) and (c) The third structure vector and (b) and (d) the first structure vector for the networks of tags in Year-2 (left panels) and Year-4 (right panels) and the corresponding innovation channels above the level $q = 2$ and $q = 3$.

3.2.8 Conclusion

Information processing underlines the evolution and structure of various social networks [129][131]. The creation of knowledge through questions-and-answers requires meaningful interactions with the actors expertise adjusted to the needs of other participants; consequently, it leads to the accumulation of the sound knowledge and the expansion of knowledge space [105]]. In the studied example, we have demonstrated how the algebraic topology measures can characterise the connection complexity of the emergent knowledge networks. Using the data of questions-and-answers from the Stack Exchange system Mathematics, we have shown how the network of math-

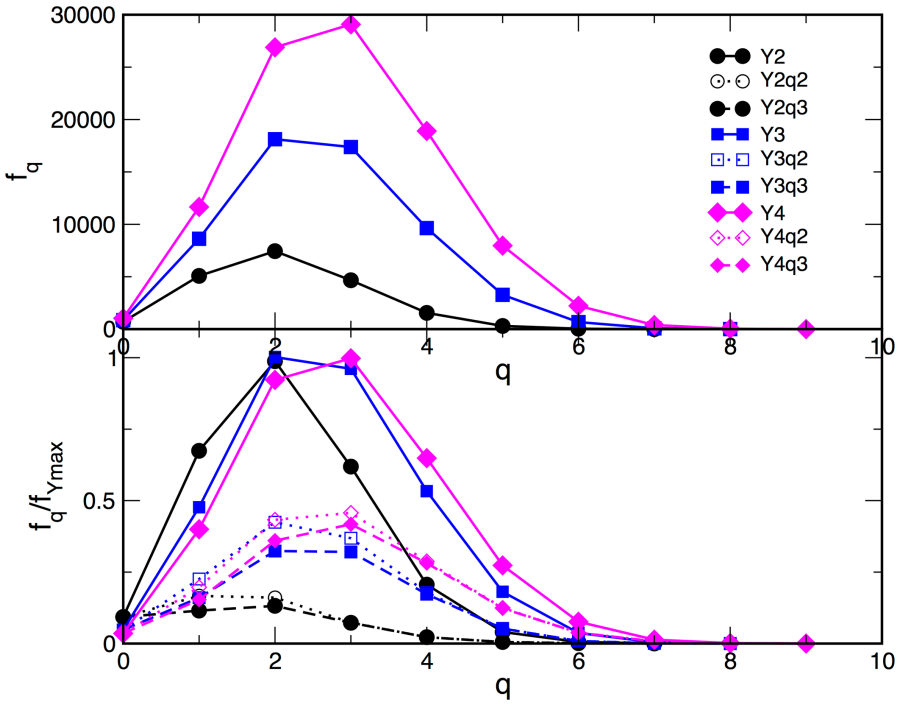


Figure 3.17: Response f_q plotted against the topology level q for networks of Year-2, Year-3, and Year-4 (top panel) and for the corresponding innovation channels scaled with the year maximum value (bottom panel).

ematical tags, as constitutive elements of knowledge, appears and evolves with the actorquestionactor-answer interactions over time.

The connections among different tags reflect their use by the actors possessing the expertise, which (at least partially) overlaps with the contents of the considered question. The networks of tags are filtered by removing the extra edges which may have appeared by chance with a specified confidence level. We have applied the filtering at the level of (uncorrelated) edges to preserve the higher-order structures, which have been the focus of this study. Our results reveal that the process preserves the genuine structure of knowledge networks consisting of thematically connected tags communities. For example, five communities in Fig 3.8 appear in the filtered network of tags in Year-4. Considering the higher-order topological spaces, the filtered networks of tags exhibit a robust structure. The hierarchy of nodes sorted out according to their suitably scaled topological dimension is represented by a unique curve, independent of the evolution time and the filtering level.

The appearance of new contents (tags) over time plays a significant role in the process of knowledge creation and the related networks. As it was shown in [105], the occurrence of new contents and new combinations of contents are chiefly related to the expertise of newly arriving users. Therefore, the introduced combinations of tags obey the logical structure as it is presented by the participating experts.

The growing number of unique combinations leads to the advance of innovation [105], as also shown in Fig 6d. Moreover, their appearance is conditioned by the cognitive-matching interactions and the users activity patterns. These features of the social dynamics are manifested in the non-random (persistent) fluctuations and long-range temporal correlations, as demonstrated in Fig 3.14a, b and c. Further, the performed algebraic topology analysis has revealed the role of these innovative contents in building the architecture of knowledge network. Specifically, we have found that:

- the newly appearing tags connect to the current network at all levels from a single link to the cliques of the highest order;
- the innovation channel is recognised as a subgraph containing simplices larger than or equal to a triangle in which at least one of the new tags occurred; its growth and the increased topological complexity over time provides the evolution pattern of the entire network;
- the growth of the innovation channel is consistent with enhanced fractal features and temporal correlations of the appearance of new contents over time; it systematically obeys the sensible connections of contents, as also demonstrated in Fig 3.15.

The presented results reveal that the creation of new combinations of knowledge contents (or innovation) is compatible with the non-random correlations in the sequence of new contents and their linking to the knowledge network. Hence the innovation expansion, as a core of each knowledge-creation process, can be additionally quantified by the fractal features of time series of new tags as well as the algebraic topology measures of the networks innovation channel. Hidden beneath these quantifiers of the emergent knowledge networks is the dynamics of human actors and their expertise, which provides the logical structure of the collective knowledge. Our approach consists of the appropriate data filtering, fractal analysis of time series, and algebraic topology techniques applied to the emergent knowledge networks and their innovative channels. The methodology can be useful to the analysis of a wide class of networks where the actors and their artefacts, as well as the cognitive elements used in the process, are clearly identified. These may include, among others, networks created by science, engineering, business and economics communities based on online collaborations. Further, such examples may also include a collection of articles (e.g. journal articles) referring to each other, where their logical units are marked. In some such situations, keywords, memes, and concepts can be identified by machine learning methods.

3.3 Hidden geometry of traffic jamming

We introduce an approach based on algebraic topological methods that allow an accurate characterization of jamming in dynamical systems with queues. As a prototype system, we analyze the traffic of information packets with navigation and queuing at nodes on a network substrate in distinct dynamical regimes. A temporal sequence of traffic density fluctuations is mapped onto a mathematical graph in which each vertex denotes one dynamical state of the system. The coupling complexity between these states is revealed by classifying agglomerates of high-dimensional cliques that are intermingled at different topological levels and quantified by a set of geometrical and entropy measures. The free-flow, jamming, and congested traffic regimes result in graphs of different structure, while the largest geometrical complexity and minimum entropy mark the edge of the jamming region.

3.3.1 Introduction

The geometrical characterization of complexity opens new horizons in the analysis of complex systems. Recent examples include introducing structural quantifiers for quantum entanglement [132], classical sandpiles [133], bioinformatics [134], and employing algebraic topology [135, 136] to characterize particulate systems. The traffic of information packets on sparse networks [137]-[139] is a paradigm of a broad class of transport processes with jamming. It exhibits complex features that are influenced by the underlying substrate, as well as by the dynamical building up of queues at nodes of high centrality, and local density fluctuations that control the global behavior [139]. At high traffic densities, the dynamical coupling between neighboring queues leads to the phenomenon of jamming. Its features resemble the congestion in granular flow [135], urban traffic [140], and other queuing problems [141]. The universality of the jamming transition on networks, however, remains a challenging problem. Recent studies provided indications for a first-order [137] as well as a continuous phase transition [138] and properties of an explosive percolation [142]. In this work, we employ methods of algebraic topology, in particular, simplicial complexes of graphs, to quantitatively characterize the jamming in systems with dynamically interacting queues. We analyze the hierarchical structure of the mathematical graphs, which suitably map the sequence of traffic density fluctuations in the sense of time-seriesgraph duality [143, 144]. In this context, we introduce geometrical and entropy measures based on the global graph structure and topological spaces around each particular vertex. These quantifiers surprisingly well describe couplings between the dynamical states in different traffic regimes, which

are rendered by interwoven topological complexes. In particular, the geometrical characterization of the jamming region is far more substantial than the conventional detrended time-series analysis. Using the model introduced in Ref. [139], we generate the traffic time series by performing simulations of the transport of information packets in dynamically distinct regimes. The transport is simulated on a substrate network with 103 nodes and a correlated cyclic scale-free structure [137]. The traffic density is controlled by a parameter R, the probability of a new packet being generated at a node per time step. The dynamics consists of guided walks from a specified origin to a destination of each packet using the next-nearest-neighbor search rule [137]. The packets at a given node make an ordered queue, whosemaximum length is fixed ($h_{max} = 1000$).At each time step, every node transmits the packet that has arrivedmost recently; that is, the topmost packet in its queue (the last-in first-out algorithm). If the queue of a neighboring node along a packets path is full, the packetwaits at its current node until the next transmission opportunity. Upon reaching its destination, a packet is removed. The diffusive traffic dynamics shows a certain regularity that simplifies the studies by transforming them into graphs [145]. With increased rates R, gradual changes in the network load occur; at the same time the strength of fluctuations differs, which permits the identification of three typical traffic regimes. They can be distinguished by various statistical measures [137, 139], and by the properties of the traffic time series. For a low generation rate $R = 0.1$, the trajectories of packets rarely intersect, resulting in the free-flow regime, which is characterized by sporadic queuing and short waiting times. The average delivery rate of the network adequately balances the creation rate R. The power-lawdistributions of the packets travel time and waiting times reflect the structure of the substrate network [139].However, for $R \geq 0.25$, finite queues can form at nodes. The jamming regime is characterized by considerable effects of queuing, which results in the increase of waiting times of packets in different queues along their paths. A large number of nodes attempt packet transmission simultaneously. A very central node can receive many packets in a single time step. Thus, a packet can get buried deeper in the queue due to increased activity of neighboring nodes. Consequently, the total number of packets that are still in the traffic, $N_p(t)$, starts rising. It has been understood [137, 138] that the onset of jamming depends on the navigation rules and network structure. A comparative study of the two substrate networks and the navigation rules that we refer to in this work can be found in Ref. [137].

3.3.2 Traffic density fluctuations and the representative traffic visibility graphs

We consider six distinct pairs of time series. Specifically, for two representative values of R in each traffic regime, we generate the time series of $N_p(t)$, the number of packets on the network, and $n_g(t)$, the number of active queues at a time t . Figure 3.18 shows the network load $N_p(t_\omega)$ for a sequence of time windows, $t_\omega = 100$ time steps; the bold lines indicate the densities where queuing affects the traffic, and the jamming can occur. Although colossal queues can form in this regime, none of the queues is full. At the edge of the jamming regime (corresponding to the maximal number of simultaneously active queues, the line $R = 0.35$ in the simulations), enormous fluctuations in packet density occur in which the system struggles to clear accumulated packets via increased activity. A natural cycle of the length ~ 400 time windows appears (cf. Fig. 3.18). When the first queue is full (i.e., at a hub) only one packet can enter that queue after one has left. Hence, the congestion spreads to the neighbor nodes, and so on. The spreading manifests in a drop of the overall activity and reduced fluctuations, which are distinguishing features of the congested regime. $N_p(t_\omega)$ exhibits a sharp increase after a certain simulation time, which is shorter for large R. The growth rate $\lambda = log_{10}N_p(t_\omega)/log10t_\omega$, which serves as an order parameter, is plotted in the top panel of Fig. 3.18 against the packet generation rate R.

The first six time series which are displayed in Fig. 3.18 span three traffic regimes where the fluctuations dominate load increase; they are transformed into six graphs, that we call traffic visibility (TV) graphs. Each TV graph contains 10^3 vertices; each vertex represents a traffic load $N_p(t_\omega)$ in the sequence of $t_\omega = 1, 2, ..., 1000$ time windows. The resulting TV graphs have different structure and should not be confused with the substrate network; see Fig. 3.19. As was pointed out in the Introduction, in this work our focus is on the analysis of the in-depth topology of these TV graphs beyond standard network measures. The closeup of a typical structure with interconnected high-order cliques studied in this work is shown in Fig. 3.19(d).

For comparison, we also consider the graphs representing a randomized version of these time series. Each data point $i = 1, 2, ..., N$ in the time series is exchanged with a point randomly selected among the other $N - 1$ data points. Note that the y-axis value of a particular data point is preserved; only its surrounding data represent some random points from the original series. Note that the resulting TV graphs are not random graphs because triangles of neighboring data points form due to the visibility methods used here. However, they have reduced complexity;

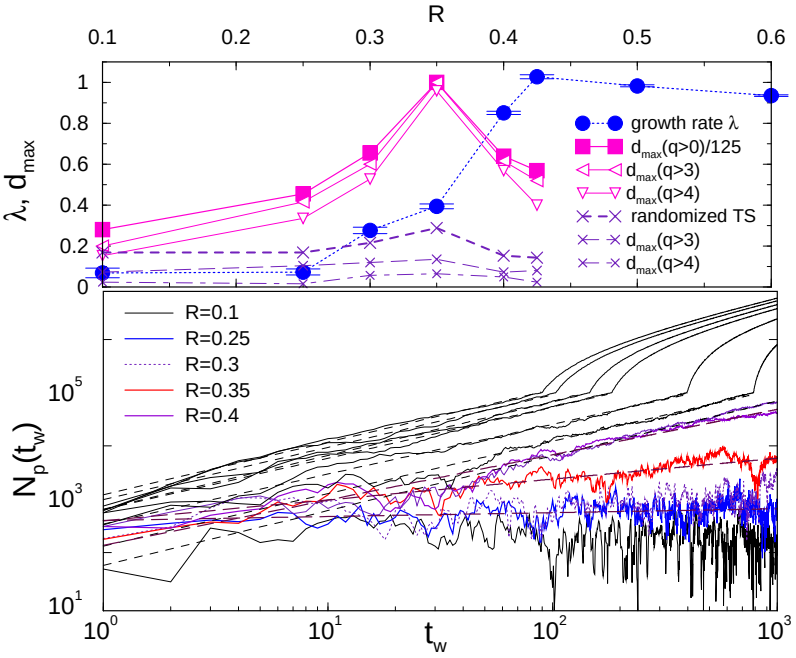


Figure 3.18: (Color online) Bottom panel: Double-logarithmic scale of traffic load fluctuations $N_p(t_\omega)$ against time window t_ω for different driving rates $R = 0.1, 0.25, 0.3, 0.35, 0.4, 0.425, 0.5, 0.6, 0.7, and 0.8$ (bottom to top). Top panel: The logarithmic growth rate λ of the time series $(\cdot)$ plotted against R; the scaled maximal topological dimension $d_{max}/125$ of a node in the corresponding TV graph computed for the range of topology levels q indicated in the legend (solid symbols) and in the graphs of the randomized time series $(\times)$ plotted against R.

see Table 3.4 and Fig. 3.20.

For completeness, we determine the ranking distribution of degree Fig. 3.19(c), and the average degree $\langle k \rangle$, the clustering coefficient CC, the average diameter d, the average path length $\langle p \rangle$, and the modularity of these TV graphs, Table 3.4. These standard network measures can yield some insights into the studied dynamical regimes. However, in the case of TV graphs, rather small variations of these network measures are observed; cf. Table 3.4. As we show in the next section, the major discrepancies between the studied traffic regimes are observable at the level of higher-order simplexes and their aggregates occurring in TV graphs.

3.3.3 Simplicial complexes of tv graphs

To reveal the connections between dynamical states $N_p(t_\omega)$ in the sequence of time, we analyze and compare higher topological features of these TV graphs using the concept of simplicial complexes of graphs [87]. In the method that we use [87, 146] simplices are identified as maximal complete subgraphs (cliques). Topological

components of the TV graphs representating the three traffic regimes are shown in Fig. 3.20. The complete data are given in Table 3.5.

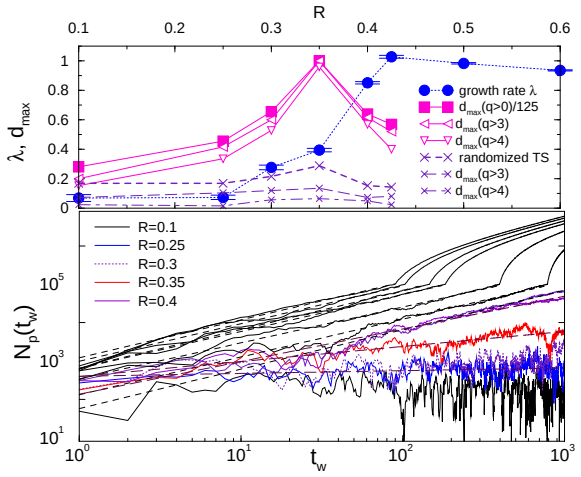


Figure 3.19: (Color online) (a) The giant cluster of the substrate network with two hubs and different scale-free distributions for incoming and outgoing links, finite clustering coefficients and disassortative mixing characteristic for Webgraph [137]. The vertex size correlates with its degree. (b) TV graph representing the traffic density $N_p(t_\omega)$ time series for $R = 0.35$. Here, vertices stand for data points while edges indicate natural visibility connections between data points. (c) Ranking distributions for the vertex degree k_i in the studied TV graphs, indicated by R values, against the vertex rank r_i . (d) Three-dimensional rendering of a closeup of the TV graph for $R = 0.35$ illustrating a typical pattern of connections analyzed in this work. Varying colors from dark to pale indicate vertices of a larger degree.

Topology and entropy measures reveal the nature of traffic jamming

In particular, in the spirit of characterizing local graph properties [134], we describe the extended surroundings of each vertex by the Q^i vector [136]

$$Q^i = \{Q^i_{q_{max}}, Q^i_{q_{max}-1}, ..., Q^i_1, Q^i_0\}$$

where Q^i identifies the number of different simplices of order q in which the vertex i participates. Then the vertex topological dimension $dimQ^i = \sum_{q=0}^{q_max} Q^i_q$ represents the number of all simplices to which the vertex i belongs. By definition, each vertex of the visibility graph represents one dynamical state of the original system. Then the vertex topology vector directly describes the structure of the manifold in the phase space to which a given dynamical state belongs. Each component Q^i further specifies the number of states that are connected to state i, i.e., by links or triangles, tetrahedra, and higher structures involving other states. Thus, the dimension $dimQ^i$

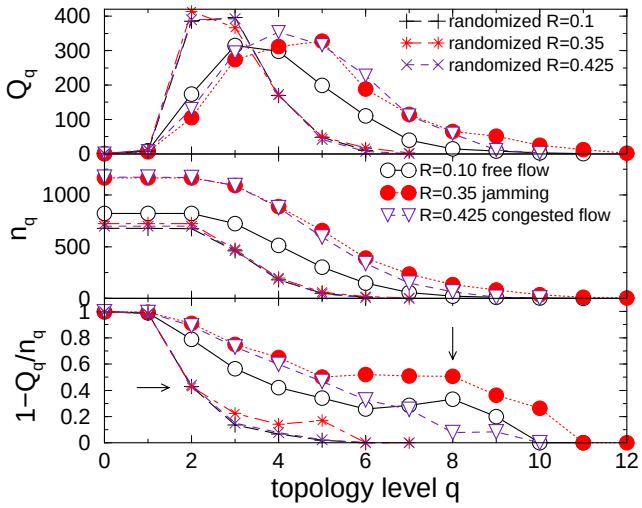


Figure 3.20: The first, second, and third structure vectors (top to bottom panel) of the visibility graphs representing the traffic density fluctuation in three different traffic regimes (solid symbols) and the corresponding randomized time series $(+, \star, \times)$. In the latter group, the connectivity at $q = 2$ drops below 40% (indicated by $\rightarrow$). On the other hand, the most complex graph has 51% connectivity at the topology level $q = 8$ (indicated by $\downarrow$).

includes not only the number of states to which the state i is connected (and can influence them), but also the geometrical forms in which these states organize.

Then the entropy of a topological level q is defined by

$$S_Q(q) = -\frac{\sum_i p_q^i log_{10} p_q^i}{log_{10} p_q^i}$$

where $p_i = \frac{Q_q^i}{\sum_i Q_q^i}$ is the occupation probability of the q level. The sum runs over all vertices, and $M_q = \sum_i (1 - \delta_{Q_q^i,0})$ is the number of vertices with a nonzero entry at the level q in the entire graph. In this context, the entropy measures the degree of

	$\langle k \rangle$	CC	d	$\langle p \rangle$	$Modularity$
$R = 0.1$	8.2	0.76	11	4.92	0.86
$R = 0.25$	8.62	0.75	10	4.56	0.85
$R = 0.3$	9.38	0.74	9	4.60	0.84
$R = 0.35$	10.37	0.72	9	4.8	0.81
$R = 0.4$	10.55	0.73	12	5.16	0.85
$R = 0.425$	10.2	0.72	10	4.8	0.80
TS-randomized $R = 0.35$	5.82	0.75	12	5.18	0.87

Table 3.4: The standard network dimensions for the traffic visibility graphs corresponding to time series at different R, and the graph from a randomized version of the time series (TS) for $R = 0.35$.

	R=0.1			R=0.25			R=0.3			R=0.35			R=0.4			R=0.425			randomized R=0.35		
q	Q	N_s	$\dot{Q}$	Q	N_s	$\dot{Q}$	Q	N_s	$\dot{Q}$	Q	N_s	$\dot{Q}$	Q	N_s	$\dot{Q}$	Q	N_s	$\dot{Q}$	Q	N_s	$\dot{Q}$
0	1	823	0.99	1	906	0.99	1	979	0.99	1	1138	0.99	1	1148	0.99	1	1174	0.99	1	724	0.99
1	9	823	0.99	8	906	0.99	4	979	0.99	6	1168	0.99	9	1148	0.99	8	1174	0.99	12	724	0.98
2	174	822	0.79	136	905	0.85	117	978	0.88	105	1168	0.91	103	1147	0.91	130	1172	0.89	414	723	0.43
3	315	723	0.56	296	823	0.64	275	908	0.7	274	1093	0.75	260	1078	0.76	295	1094	0.73	367	473	0.22
4	298	515	0.42	340	622	0.45	310	715	0.57	312	891	0.65	298	888	0.66	354	877	0.6	170	198	0.14
5	199	302	0.34	235	362	0.35	307	485	0.37	328	657	0.50	292	654	0.55	316	595	0.47	49	59	0.7
6	110	148	0.26	152	177	0.14	171	238	0.28	188	392	0.52	259	422	0.39	227	337	0.33	18	18	0
7	40	56	0.29	43	49	0.12	74	94	0.21	115	234	0.51	149	208	0.28	111	149	0.26	3	3	0
8	14	21	0.33	7	9	0.22	23	32	0.28	65	132	0.51	68	87	0.22	59	64	0.08			
9	8	10	0.2	3	3	0	11	12	0.08	51	80	0.36	31	32	0.03	11	12	0.08			
10	3	3	0				3	3	0	25	34	0.26	6	6	0	3	3	0			
11	1	1	0				1	1	0	12	12	0									
12										2	2	0									

Table 3.5: The numerical values of the components of three structure vectors, defined within the theory of simplicial complexes on graphs, for the TV graphs generated from $N_p(t_b)$ at all studied traffic rates R. For comparison, we also display the structure vectors of the graph corresponding to the randomized version of the time series for $R = 0.35$.

cooperation among vertices. The vertices that comprise an isolated clique result in a higher entropy than the vertices that share different cliques at a particular level, which causes a decrease in the entropy. Unlike the second structure vector n_q , the components f_q measure interconnection between simplexes at the level q. By varying R, the changes that occur in the traffic load and intensity of fluctuations of the time series cause the cliques to acquire new connections. The quantity $f_q(R)$ reflects such rearrangements of the structure and size of the cliques agglomerates over q levels, thus representing a sort of response function in the topological space. Considering the TV graph for a given R value, we find the vertex with the maximum topological dimension $d_{max} = max(dimQ^i)$, $i = 1, 2..., N$. In the top panel of Fig. refslikatraffic1 we plot d_{max} against the corresponding R value. Remarkably, the vertex maximum topological dimension exhibits a sharp peak at the same value $R = 0.35$ at which the order parameter λ, estimated from the time series, shows a steep increase. The smaller values of $max(dimQ^i)$ appear in all other TV graphs. The dimensions of the same vertex but computed for higher topological levels $q > 3$ and $q > 4$ are shown in Fig.3.18 (top). They demonstrate that the best-connected vertex in the $R = 0.35$ graph also belongs to a topological space of higher complexity than the vertices in other graphs, in agreement with the global topology data in Fig. 3.20. In contrast, in the graphs from the randomized time series, the vertex dimension has much lower values with increased q threshold. These findings suggest that the most complex structure is discovered at the edge of the jamming region, whereas the free-flow as well as the congested regimes map onto graphs with a relatively simpler structure.

The results for the entropy and f-vector components, shown in Fig. 3.21, support this conclusion. The concave shape of the entropy $S_Q(q)$ as a function of topology

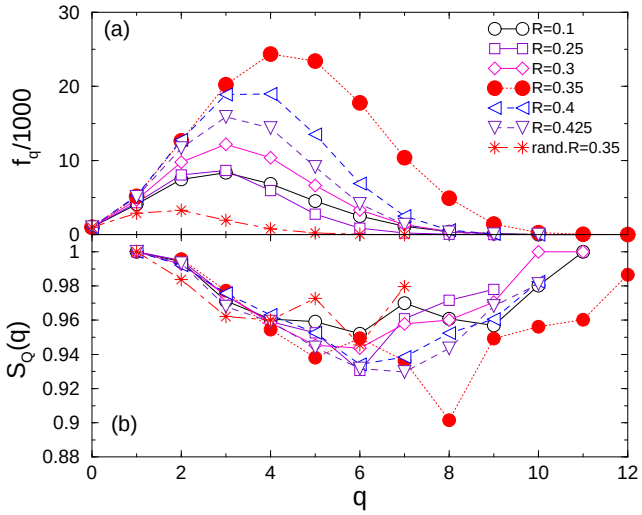


Figure 3.21: (a) The geometrical response function f_q , and (b) the entropy of occupation of topology level q, $S_Q(q)$, plotted against q for graphs representing time series for different rates R. The corresponding quantities of the graph from the randomized time series for $R = 0.35$ are shown by $\star$

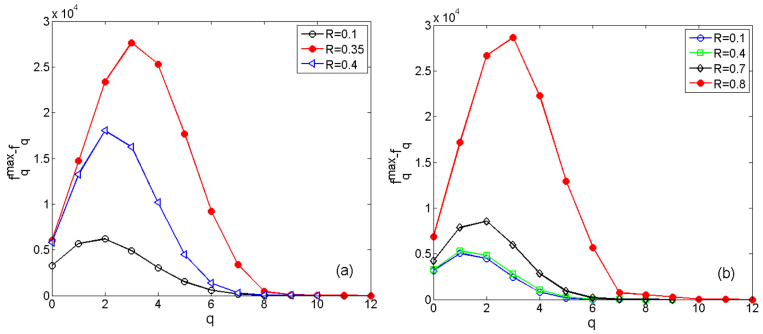


Figure 3.22: Difference between the upper bound f_q^{max} and the actual f_q value for various TV graphs related to the traffic density time series on the substrate network Webgraph (a) and a homogeneous network Statnet [137], where the jamming occurs at $R = 0.8$ (b).

level q reflects the fact that at both ends isolated cliques may occur (links at $q = 1$ and (q_{max}+1) cliques at q_{max} level), leading to the maximum entropy. Otherwise, the occurrence of shared ($q = 1$)-order faces between different simplices at a particular q level make them mutually dependent and thus reduces the entropy. The lowest entropy is observed in the most complex structure, the graph for $R = 0.35$, at $q = 8$. At this level, indicated by the vertical arrow in Fig. 3.20, 132 simplices are organized into 65 distinct connected components Table 3.5.

Thus the degree of connectivity between simplices is 51% at this topology level, compared with 33% in the free-flow case and 8% in the congested case. Similarly, in Fig. 3.21(a) the highest peak among the f_q components corresponds to the graph for

$R = 0.35$. Other TV graphs exhibit peaks with lower values, which gradually move towards smaller q levels, resembling the familiar glassy response in the frequency domain [147]. The upper bound of the f_q component, f_q^{max} , which would correspond to the situation where all simplices at the q level were independent, can be computed from the structure vectors in Fig. 3.20. Specifically,

$$f_q^{max} = \sum_{k=q}^{q_{max}} (n_k - n_{k+1}) \binom{k+1}{q+1}$$

Hence, $f_q^{max} - f_q \geq n_q - Q_q = n_q \hat{Q}_q$. The largest difference between the actual f_q and its upper bound is observed in the most complex graph at $R = 0.35$. A qualitatively similar behavior is found for traffic on a substrate network of different structure, where the jamming occurs at a larger density; see Fig. 3.22.

3.3.4 Conclusion and outlook

We have demonstrated how the complexity of connections between the states of a large dynamical system, which has profound consequences for the systems collective dynamics, can be quantified by algebraic topology techniques. Exemplified by the mapping of a suitable time series, the manifolds of connected states in the systems phase space receive within this approach a clear geometrical interpretation. We have applied this advanced graph theory analysis in the context of the time-seriesnetworks duality. Our results reveal a type of graph representing the time series of a collective dynamics. While the standard network measures of these graphs are similar (cf. Table 3.23), they possess a rich structure at higher topological levels. The observed variety of simplicial complexes closely reflects the strength of fluctuations in the underlying time series. Notably, enhanced fluctuations near the phase transition transform into visibility graphs of higher topological complexity.

The geometric quantifiers of graphs emanating from the fluctuations of time series can be useful in many general contexts. As a prominent example, we have studied traffic jamming, a collective phenomenon that is caused by dynamically interacting queues. The fractal analysis of traffic activity time series gives evidence of a self-organized dynamics among the nodes of the substrate network. The long-range temporal correlations, persistent fluctuations, and avalanches of clustered events occur with gradually changing scaling features throughout the traffic regimes. The phase space manifolds of this collective dynamics are represented by the studied TV graphs. Using the algebraic topology techniques, we have explored simplicial complexes of these TV graphs. In this concept, we have introduced appropriate topological and entropy measures, which can distinguish the edge of jamming and

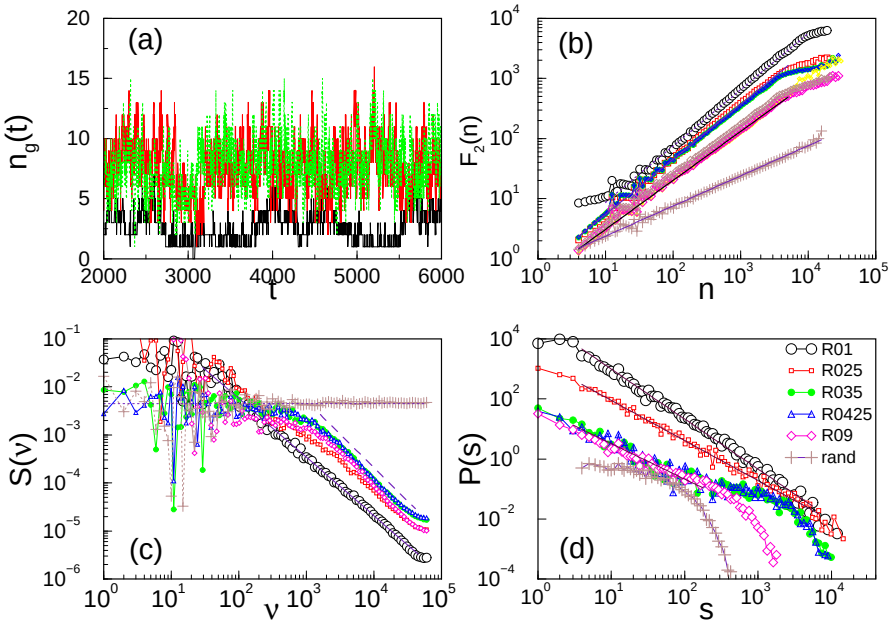


Figure 3.23: (a) Examples of time series $n_g(t)$ for $R = 0.1$ (black) and $R = 0.35$ [orange (gray)] line; (b) fluctuations around local trends of all time series; (c) power spectrum; and (d) the distributions of avalanche sizes for different R values. The legend applies to all three panels (b)(d). The corresponding quantities of a randomized time series in (b), (c), and (d) are shown by $\times$ symbols.

give insight into the nature of the jamming transition. In particular, the geometrical response function $f_q(R)$ exhibits a glassy-type behavior suggesting the underlying hierarchy of time scales across the jamming region. It should be stressed that this feature of traffic jamming has been elusive to standard methods, for example, the fractal analysis. In the present approach, it is revealed by the structure of the topological levels, which encode the correlations in the dynamical states of the primary system. Further research is necessary to uncover the relevant time scales in traffic jamming.

In conclusion, the detailed structure of the phase-space manifolds revealed by the algebraic topology technique contains vital knowledge about the systems collective behavior. The methods introduced here can yield insights into a variety of dynamical regimes occurring in complex systems, in particular, systems exhibiting a phase transition, percolation, explosive percolation, and others. The topological structure vectors are sensitive to differences in the dynamically appearing phases, thus complementing the behavior of the physical order parameter across the transition. We note that the studied problem of traffic on networks is of great interest from the applications point of view. It appears in contexts ranging from traffic on the internet, urban transport [138, 140], and traffic-related problems [148]. The analysis

of topological spaces can also reveal relevant relationships [136] in real and data-driven complex networks. The response and entropy measures introduced here for studying the jamming region can be applied to quantify the congestion in physical systems. Moreover, the presented methodology provides the elements for influencing the course of the dynamics in complex systems by the efficient methods of graph manipulations. Our work may provide an impetus for inclusion of the algebraic topology techniques as an additional standard in the analysis of complex dynamical systems and networks.

3.4 Fractal time series of collective charge transport and topology of phase space manifolds

Charge transport in the Coulomb blockade regime of two-dimensional nanoparticle arrays exhibits nonlinear I-V characteristics, where the level of nonlinearity strongly associates with the array's architecture. Here, we use different mathematical techniques to investigate the collective behavior of the charge transport and quantify its relationship to the structure of the nanoparticle assembly. First, we simulate single-electron tunneling conduction in a class of nanoparticle networks with a controlled variation of the structural characteristics (branching, extended linear segments) which influence the local communication among the conducting paths between the electrodes. Further, by applying an innovative approach based on the algebraic topology of graphs, we analyze the structure of connections in the manifolds, which map the fractal time series of charge fluctuations in the phase space. By tracking the I-V curves in different nanoparticle networks together with the indicators of collective dynamics and the topology of the phase space manifolds, we show that the increased I-V nonlinearity is fully consistent with the enhanced aggregate fluctuations and higher connection complexity among the participating states. Also, by determining shifts in the topology and cooperative transport features, we explore the impact of the size of electrodes and local charge disorder. The results are relevant for designing the nanoparticle devices with improved conduction; they also highlight the significance of topological descriptions for a broader understanding of the nature of fluctuations at the nanoscale.

3.4.1 Introduction

In the science of complex systems, understanding the emergence of distinct properties on a larger scale is one of the central problems, which requires the use of

advanced mathematical and numerical approaches [149]. Nanostructured materials are examples of complex systems exhibiting new functionality at the assembly level [150, 151]. The concepts of nanonetworks ([152] and references there) facilitate the application of graph theory methods for a quantitative study of complexity at a nanoscale. Here, we combine graph theory techniques with additional mathematical methods to investigate the connection between the cooperative electron transport through nanoparticle assemblies at applied bias within the Coulomb blockade regime and the architecture of the assembly.

In conducting nanoparticle arrangements, the Coulomb blockade conditions provide the single-electron tunneling (SET) conduction between neighbouring nanoparticles ([153, 154, 155] and references within). The prototypes of systems with the Coulomb blockade transport are self-assembled 3-dimensional arrays [156] and nanoparticle films on substrates [157, 158], consisting of small metallic nanoparticles with capacitative coupling along the tunneling junctions. Recently, similar conduction mechanisms have been described in quantum dot arrays in the reduced graphene oxide [159, 160], a new electronic and optoelectronic material [161]. In this case, quantum dots of graphene are separated by non-conducting areas through which the electrons can tunnel at the applied voltage bias. The relevance of the Coulomb blockade transport has been experimentally investigated in a variety of other nanostructures including nanowires [162], granular metals [163], and thick films [164] as well as different molecular arrays [165, 166]. At low temperatures, SET represents a main process in the assemblies of small nanoparticles arranged with a fixed pattern of tunneling junctions [167, 168, 169, 170, 171, 157, 158] and graphene quantum dots [159, 160]. Beside tunneling, another dynamical regimes, occurring in hybrid nanocomposites [172] and separate time scales due to the motion of molecular linkers [173], have been investigated.

In the nanoparticle films, at low temperatures and the applied weak bias at the electrodes encasing the array of nanoparticles the current–voltage characteristic is nonlinear in a range of voltages V above a threshold V_T,

$$I(V) \sim (V - V_T)^{\zeta} \ . \tag{3.1}$$

It has been recognized that the degree of nonlinearity, which is measured by the exponent $\zeta \in (1, 5)$, robustly correlates with the structure of nanoparticle films [157, 158, 160] and their thickness [164, 174]. The origin of this phenomenon is in the *cooperative charge transport* that involves multi-electron processes along the conduction paths, dynamically emerging between the electrodes. Precisely, the SET conduction through a single Coulomb island between the electrodes results in a

linear $I(V)$ dependence. Similarly, $\zeta \simeq 1$ is found in one-dimensional chains of nanoparticles supporting a single conduction path [167, 155]. A detailed theoretical analysis of the tunneling conduction through a chain of metallic grains with charge disorder can be found in [163]. In contrast, multiple conduction paths establish between the electrodes in the case of two-dimensional arrays and thick films. Depending on the structural characteristics of the array at local and global scale, such paths form drainage basins. Consequently, the cooperative tunneling events may occur in such basins involving several conducting paths, which increase current through the system and leads to enhanced nonlinearity in $I(V)$ curves [157, 174]. While fluctuations of the current are standardly measured at the electrode [167, 168, 170, 157, 160, 164, 175, 174], a direct observation of the tunneling events inside the sample remains a challenging problem to the experimental techniques [176]. Therefore, the genesis of the collective charge fluctuations and its connection with the structure of the array remains in the domain of numerical modeling. In this regard, the idea of nanonetworks [152] provides the framework to quantify the structure of diverse nanoparticle assemblies by graph theory methods [177]. Furthermore, the numerical implementation of SET on an array of the *arbitrary structure represented as a nanonetwork* has been introduced [178, 155].

In this work, we combine different numerical techniques to study collective features of charge transport through two-dimensional nanoparticle assemblies; the aim is to examine cooperative dynamical behavior involving different conduction paths in connection with the structural elements of the assembly and the size of electrodes as well as the effects of charge disorder. Our approach consists of three levels illustrated in Fig. 3.24, relating to a suitable mathematical modeling (see Methods for a detailed description).

In particular: (i) We construct several two-dimensional assemblies of nanoparticles connected by tunneling junctions of a given structure. Then, setting the electrodes (cf. an example in Fig. 3.24 bottom) and slowly ramping the voltage bias, we simulate SET through such assemblies. (ii) We sample the time series of the relevant observables and perform the fractal analysis of these time series, as described in Methods, to determine the quantitative indicators of aggregate fluctuations. Here, we consider temporal fluctuations of the number of tunnelings per time unit, $N_Q(t_k)$, an example is shown in Fig. 3.24 middle. (iii) Given recent developments of time-series–graphs duality [179, 180, 181], we convert these time series into graphs. The sets of data points specify a manifold in the phase space of the system's states that are involved in the course of events; the resulting graph then contains connections between these states. To explore the connection complexity

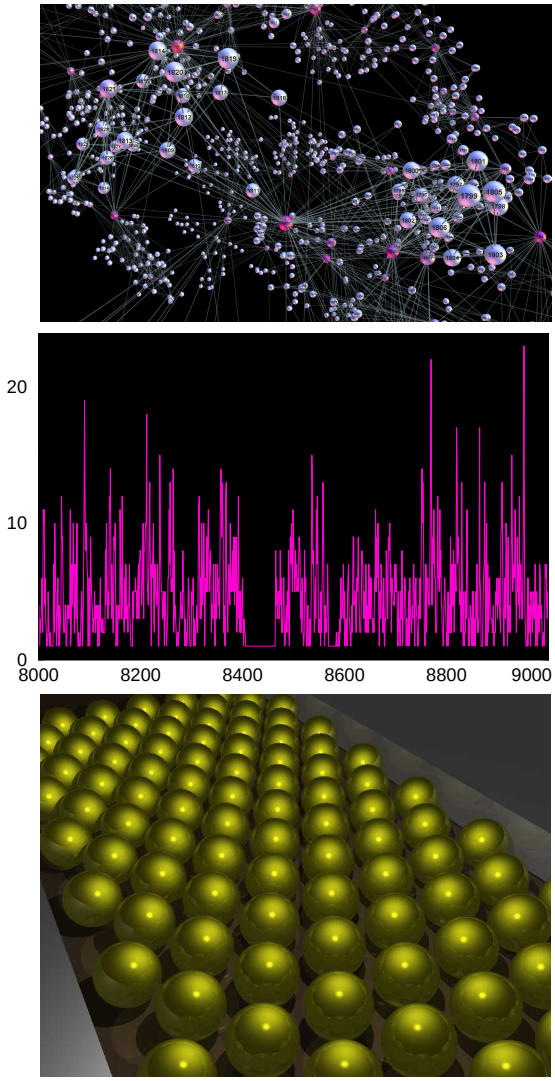


Figure 3.24: Illustration of three levels of the charge transport modeling. Bottom: Simulations of SET processes on nanoparticle array between the electrodes; Middle: Fractal analysis of the time series of the number of tunnelings per time unit in the assembly; Top: Algebraic topology analysis of the graph representing the phase-space manifold behind the sequence of events in the time series.

among the system's states, we use algebraic topology techniques (see Methods), and study higher order combinatorial spaces (simplexes) of these graphs. Our comparative analysis of different nanoparticle assemblies reveals a consistent correlation between the occurrence of collective charge transport, topological complexity of the phase space manifolds, and the $I(V)$ nonlinearity.

3.4.2 Methods

The modeling approach on three levels mentioned in the Introduction encompasses different scales and algorithms, from the single-electron tunnelings at each junction of the nanonetwork in real space to collective charge fluctuations of the whole assembly, which are studied in time and the abstract phase space manifolds. The computational hierarchy is described as follows.

I. Simulations of SET on nanonetworks. The SET processes are simulated in four different nanoparticles assemblies represented by the nanonetworks in Fig. 3.26. First, the structure of the nanonetwork is included by specified adjacency matrix $\mathbf{A}$ and the electrodes are set to some periphery nodes. Usually, we chose $P/4$ nodes in the case of extended electrodes, or a single node, in the case of point-size electrodes, at two opposite sides of the network; here, P is the number of periphery nodes of a given structure (see Fig. 3.24). Then the capacitance matrix and its inverse are computed and the vectors $\mathbf{Q}$ and $\mathbf{V}_+$ are initialized as zeros. To start the process, the time t is initialised and the bias voltage set. For each value of the bias in the range $V \in [0, V_{max}]$ the number of steps are performed as follows. At each step, a tuneling is attempted along each junction $i \to j$ and the values V_c for all nodes updated and the corresponding energy charges $\Delta E_{i\to j}(t)$ are computed. Then, each tuneling rate $\Gamma_{i\to j}(t)$ is determined; the delay time Δt_{ij} of the tuneling along the $i \to j$-junction is estimated. Following the description in [155], $\Delta t_{ij} = \frac{-log(1-x_{ij})-\sum_{k>k_0}\delta t_k \Gamma_{i\to j}(t_k)}{\Gamma_{i\to j}(t)}$, where δt_k are time intervals between successive tunnelings in the system following the time t_{k_0} of the last tunneling at the junction, and x_{ij} is a uniform random number. The tunneling along the junction with a minimal delay time is processed and the time increased accordingly $t \to t + \Delta t_{ij}$. Subsequently, the charge Q_i and V_i are updated for all nodes, and the process is repeated to find next tunneling event and so on. The simulated data are in the limit $C/C_g = 10^{-4}$, where a faster algorithm for computing the inverse of the capacitance matrix can be used [155]. We also keep $T = 0$ and $\Phi_- = \Phi_g = 0$ while $\Phi_+ = V$, the applied voltage. The simulations are performed until $V \gtrsim 10V_T$ in the corresponding nanonetwork. To sample time series of interest for this work, we set an appropriate time unit according to the average tunneling rate $\delta t = a/\langle\Gamma_{ij}\rangle$ in each nanonetwork. The parameter $a = 1$ in the one-dimensional array and takes different values [155] a =4.11 and 17.1 in CNET and NNET, respectively, in connection with the estimated number of potential paths in these irregular arrays. Then the time series represents the sequence of the number of tunnelings $N_Q(t_k)$, $k = 1, 2\cdots$ in the entire array per the identified time unit.

II. Fractal time series analysis and temporal correlations. The occurrence of collective charge fluctuations manifests in long-range temporal correlations, clustering

of events and fractal features of time series. Various indicators of the collective behavior are determined by analysis of the time series of the number of tunnelings $N_Q(t_k)$. In particular, the temporal correlations occurring in the streams of events result in the power spectrum

$$S(\nu) \sim \nu^{-\phi} \tag{3.2}$$

with a power-law decay in a range of frequencies ν. The avalanches of tunnelings accompany such temporal correlations. Using the standard numerical procedure [184, 185], the avalanches are identified in the stationary signal above V_T. The size of an avalanche comprises the area below a regular part of the signal above the zero thresholds. A broad distribution of avalanche sizes with a power-law tail suggests that, in a state with long-range correlations, the size of a triggered avalanche is not necessarily proportional to the triggering action. Moreover, the system's relaxation in the response to the external driving is characterized by the differences between the size of consecutive events (first return). The following expression, characteristic of the nonextensive statistical mechanics [186, 187, 188],

$$P_\kappa(X) = A\left[1 - (1-q_\kappa)\left(\frac{X}{X_0}\right)^\kappa\right]^{-1/1-q_\kappa} \tag{3.3}$$

satisfactorily reproduces the distributions of the avalanche sizes, i.e., q_1-exponential, for $\kappa = 1$, and the returns q_2-Gaussian, for $\kappa = 2$.

According to the fractal analysis of complex signals [189], the time series profile $Y(i) = \sum_{k=1}^{i}(N_Q(t_k) - \langle N_Q(t_k)\rangle$ is divided into N_n segments of length n. The standard deviation $F_2(\mu, n) = \frac{\sum_{i=1}^{n}[Y((\mu-1)n+i) - y_\mu(i)]^2}{n}$ around the local trend $y_\mu(i)$ is computed at each segment $\mu = 1, 2, \cdots N_n$. Then the average over all segments exhibits a scaling law

$$F_2(n) = (1/N_n)\sum_{\mu=1}^{N_n} F_2(\mu, n) \sim n^H \tag{3.4}$$

with respect to the varied segment length n, where H is Hurst exponent. While $H = 1/2$ characterises random fluctuations, the values $H \in (1/2, 1)$ indicate the persisten fluctuations of the fractional Gaussian noise signal [190, 189, 191].

III. Mapping time series onto graphs. The sequence of events captured by the time series of the number of tunnelings represents a manifold in the state space of the underlying nanonetwork. Dealing with a fractal signal (see Results), we expect a more complex connections among these states to exist beyond the actually realised sequence. To reveal such complexity of the system's phase space, we use the

mapping of time series to mathematical graphs recently developed [180, 179]. Here, we apply the 'natural visibility' mapping [180, 181], which is particularly suitable in the case of persistent fluctuations [181]. The mapping procedure is illustrated in Fig. 3.25. Each data point of the time series is represented by a node of the graph, here termed *QTS-graph to indicate the charge fluctuations time series*. The node is connected by undirected links with all other data points that are visible from that data point, where the vertical bars are considered as non-transparent. Note that by varying the mapping procedure different graphs can be obtained. However, here we use the same mapping procedure for different time series (i.e., charge fluctuations in different nanonetworks) and perform a *comparative analysis of the structure of the resulting graphs*. For this purpose, we map equal parts (2000 data points) of the time series in each of the considered nanoparticle systems; skipping the initial 8000 points ensures that the regime $V \geq V_T$ has been reached.

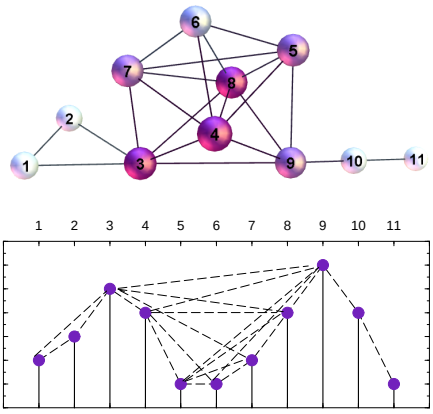


Figure 3.25: The sequence of data points in the lower pannel is mapped onto a graph in the upper pannel; each data point becomes a node of the graph, while the graph's edges are inserted according to the 'natural visibility' between data points, indicated by broken lines.

IV. Algebraic topology measures of QTS-graphs. Beyond standard graph theory measures [192, 193], the algebraic topology of graphs [194] is utilised to determine the higher-order structures of a graph, i.e., simplicial objects and their collections, simplicial complexes, closed under the inclusion of faces. Using the paradigm of Taylor expansion, such combinatorial spaces correspond to higher-order terms while the graph itself represents the linear term. Recently, the computational topology techniques have been applied to analyse the hierarchical organisation in online social networks [195], study of grain connectedness in trapped granular flow [196], and to reveal the changes in the topological structure of state space across the traffic jamming regime [181].

The method that we use [197, 198] identifies simplexes as maximal complete subgraphs or cliques. Thus, $q = 0$ represent an isolated node, $q = 1$ two nodes connected with a link, $q = 2$ is a triangle, $q = 3$ a tetrahedron and so on until $\sigma_{q_{max}}$, which identifies the highest order clique present in the graph. Faces of a simplex σ_q are the subsets $\sigma_r < \sigma_q$. The method determines cliques of all orders and identifies the nodes that belong to each clique. Utilising this rich information, we characterise the topological complexity of the graph at the global graph's level, i.e., by defining various structure vectors, as well as the level of each node [195, 181]. In particular, having identified the graph's topology layers $q = 0, 1, 2, \cdots q_{max}$, we describe the number of cliques and how they are interconnected via shared faces at each level from $q_{max} - 1, ..., 1$:

- Three structure vectors of graph have the components G_q, n_q and $\hat{G}_q = 1 - G_q/n_q$, which determine, respectively, the number of connected components at the level q, the number of simplexes from the level q upwards, and the degree of connectedness between the simplexes at q-level.

- The node's structure vector is defined [195] by the components G_q^i, the number of simplexes of order q to which the node i participates. Then $dim(G_i) = \sum_q G_q^i$ is the node's topological dimension.

- The topological entropy $S_G(q)$ and the "response" function f_q are defined [181] using the above quantities. Namely, the probability that a particular node i contributes to the occupation of the topological level q is $p_q^i = G_q^i / \sum_i G_q^i$. Then the graph's entropy is

$$S_G(q) = -\frac{\sum_i p_q^i log_{10} p_q^i}{log_{10} \sum_i (1 - \delta_{G_q^i,0})} \ , \tag{3.5}$$

  where the sum in the denominator indicates the number of nodes with a nonzero entry at the considered topology level. The q-level component f_q is defined as the *number of simplexes and shared faces at the topology* level q.

Notice that that the topological "response" f_q is different from the component n_q of the above defined second structure vector. The study in [181] have shown that the function f_q precisely captures the topology shifts occurring in the underlying time series due to changed driving conditions.

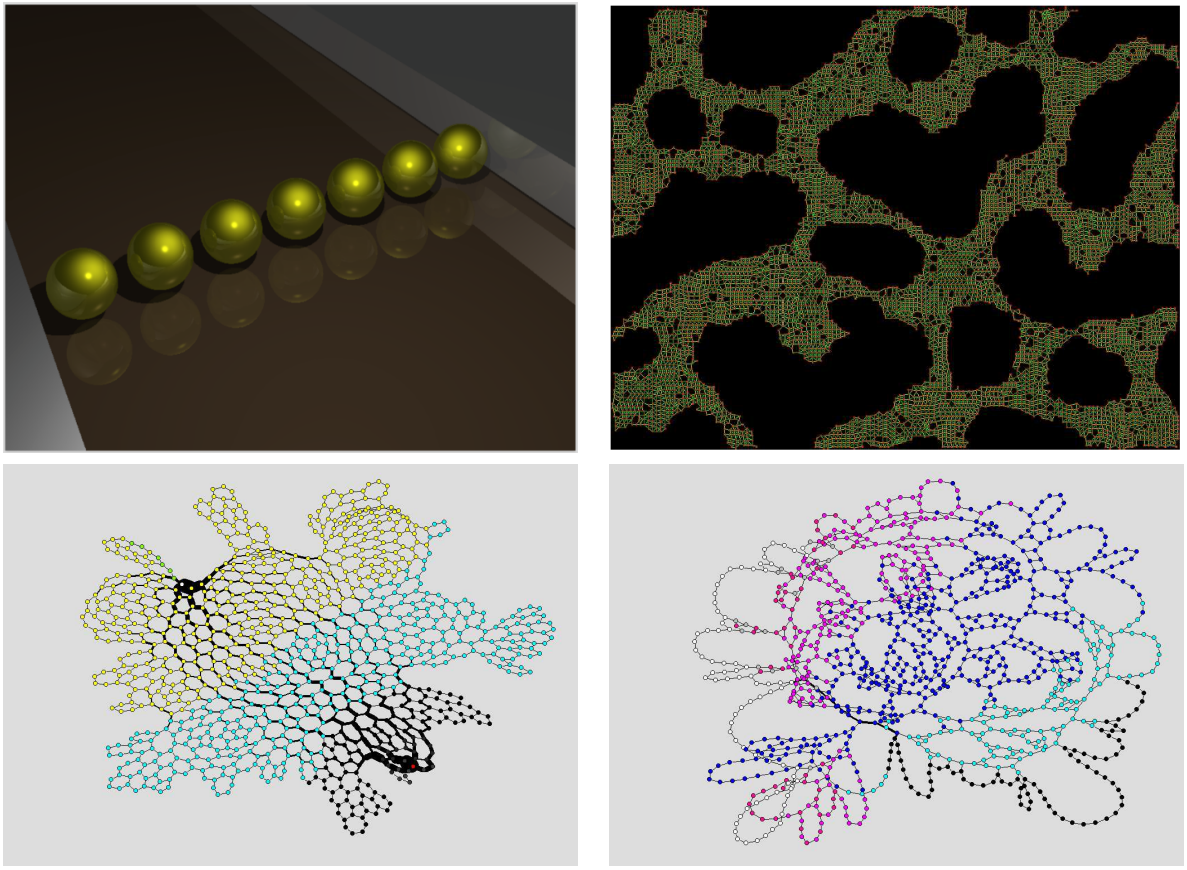


Figure 3.26: Top left: a 3-dimensional rendering of a chain of nanoparticles on a substrate and the attached electrodes. Top right: A 2-dimensional projection of the nanoparticle network NNET with 5000 nanoparticles self-assembled by the evaporation methods (data are from Ref. [157]). The lines indicate the tunneling junctions while a nanoparticle is found at the intersections. Bottom: CNET with a point-size electrodes (left) and SF22 with the extended electrodes (right), both grown by the cell aggregation [177]. The voltage drop across the sample is indicated by the color of the network's areas with high, intermediate and low potential.

3.4.3 Results and Discussion

Nanonetworks with different structural components and the origin of cooperative charge transport

For a comparative analysis of charge transport, the simulations of SET are performed in four nanoparticle assemblies with different structural characteristics. Specifically, two self-assembled nanoparticle arrays on a substrate, which are represented by the nanoparticle networks in the top row of Fig. 3.26, and two computer-generated structures, in the bottom row. For this work, these structures are identified as follows. The one-dimensional chain of nanoparticles between the electrodes, where we assume the presence of charge disorder $x_i \in [0, 1]$ as a uniform random number, is

named 1Dwd. Further, NNET is a strongly inhomogeneous structure of nanoparticles, which is self-assembled on the substrate by the evaporation process [157]. The occurrence of empty areas of different sizes as well as the regions where the nanoparticles appear to be densely packed is the marked characteristics of this assembly. CNET, shown in the lower left, and SF22 in the lower right panel, are grown by the cell aggregation process [177]. In these networks, we control the appearance of two relevant structural elements—branching and the extended linear segments—that appear statistically in the above described self-assembled structures. Namely, CNET has nearly regular hexagonal cells, resembling the dominant shapes in the dense areas of the NNET, while SF22 exhibits voids of different sizes; the cell sizes are taken from a power-law distribution with the exponent 2.2 [199]. Moreover, both CNET and SF22 have a fixed degree of internal nodes $k = 3$, which reduces the branching possibilities for the tunnelings, in contrast with the NNET, where a broader degree distribution of nodes is found [157, 155]. Furthermore, the electrodes are set to the left and right edge of the NNET, thus touching the closest layer of the nanoparticles. Similarly, the extended positive and negative electrode are set each along a quarter of the periphery nodes, indicated by white and black nodes in SF22 structure in Fig. 3.26. In contrast, the point-size electrodes are attached to two periphery nodes in CNET and two nanoparticles at the opposite ends of the 1Dwd chain.

The situation with point-size electrodes in CNET in Fig. 3.26 readily illustrates the importance of local structure for the collective charge transport of the assembly. Increasing voltage bias permits tunnelings between the electrode and the first layer, consisting of two connected nanoparticles. Then further tunnelings can occur by breaking the Coulomb blockade along the junctions towards nearest neighbor nanoparticles. The preferred direction of tunnelings follows the potential drop, which is indicated by the different color of nodes. Apart from the long-range electrostatic interactions, the local balance between the charge and potential at each nanoparticle depends on its neighbourhood. At $V \sim V_T$ the number of charges in the system is large enough to allow the first conduction channel to form along the shortest path between the electrodes. For $V \gtrsim V_T$ the conditions for the appearance of another next-to-shortest path are met. Here, the branching possibilities for the charge flow play an important role. Consequently, the process can involve several paths thus making a river-like structure that drains at the last layer of nanoparticles. These paths often share some central nodes and junctions. The most used junctions (indicated by thick lines of CNET make the main conduction channels [155], whose geometry crucially depends on the local structure. Draining along the conduction

paths can cause a cascade of tunnelings, where each event satisfies the local Coulomb blockade threshold and charge–potential balance. Thus, in this range of voltages, the SET represents a dynamical percolation problem [171] where several conduction paths contribute to the measured current, resulting in the nonlinear $I(V)$ curves [63]. The quantitative analysis presented in the remaining part of this section confirms this picture. On the other hand, for $V \gg V_T$, the local potential exceeds the Coulomb blockade potential yielding Ohmic conduction and the crossover to a linear $I(V)$ dependence.

Temporal correlations and the evidence of collective charge fluctuations

The number of tunnelings per the identified unit interval is recorded from the SET simulations in four nanonetworks of Fig. 3.26. The resulting time series are shown in Fig. 3.27. The corresponding power spectra displayed in top panel of Fig. 3.27, indicate the occurrence of temporal correlations that depend on the structure of the underlying nanonetwork. Specifically, the Eq. (3.2) applies within a different range of frequencies and different exponent ϕ, which is shown in the legend. Apart from slight variations in the exponent, the power spectrum reveals the certain similarity between the charge fluctuations in the nanonetworks SF22 and 1Dwd, on one side, and NNET and CNET, on the other. Hence, the dominance of the linear elements, which is obvious to the chain structure, seems to play a role in the SF22 too. The spectrum exhibits the power-law decay (3.2) in the entire range of ν. On the other hand, plenty of branching possibilities in CNET and NNET lead to the occurrence of multiple paths; a typical scale appears, which leads to the peak in the spectrum between the correlated high-frequency part and the rest of the spectrum resembling a white noise.

A further similarity between these two groups of nanostructures is found considering the nature of fluctuations and the statistics of avalanches, which are displayed in Fig. 3.28. Computing the standard deviations around a local trend on the segment of length n of the time series, we find the scaling regions according to Eq. (3.4) and determine the corresponding Hurst exponent. Note that, according to Eq. (3.4), the temporal scale of the collective fluctuations can be identified as the range of the segments n for which the scaling occurs (straight sections of the fluctuations curves in Fig. 3.28). Similarly, the spatial scale $\ell \sim s_0^{1/D_f}$ is determined by the cut-off size s_0 in the distributions of avalanches in Fig. 3.28, where D_f stands for the fractal dimension of the avalanches. The values of the exponents for all time series, which are listed in the legend, are in the range $H > 0.5$, suggesting *persistent* fluctuations of charge transport in all considered nanoparticle networks. The fluc-

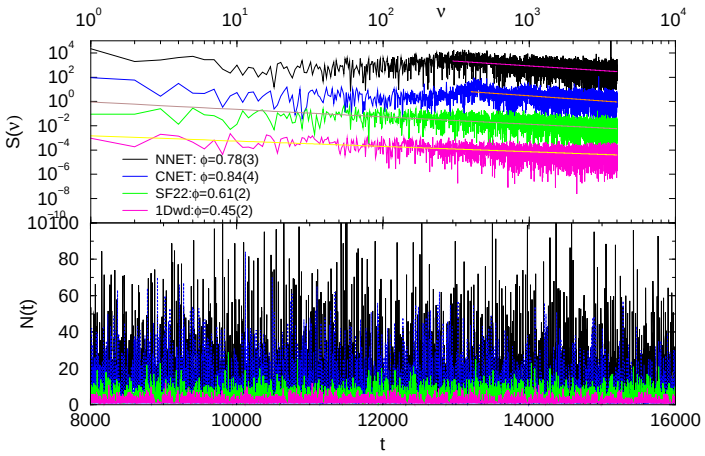


Figure 3.27: Time series of the number of tunnelings (bottom panel) and the corresponding power spectrum (top panel) in four nanoparticle networks from Fig. 3.26. The legend and color apply to both panels. The signal for NNET is scaled by 1/5 to fit the scale.

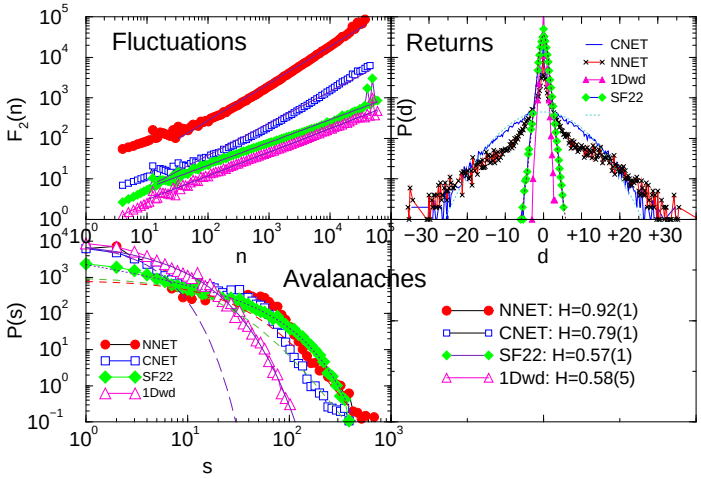


Figure 3.28: For four nanonetworks from Fig. 3.26: Double logarithmic plot of the fluctuations $F_2(n)$ against the segment size n, top left, and the scaling regions indicated by a straight line on each curve; The distributions $P(s)$ of the avalanche sizes s, bottom left, and the distribution $P(d)$ of the first returns d of the signal.

tuations in the linear chain of nanoparticles and the SF22 structure have a similar value of the Hurst exponent, again suggesting the relevance of the extended linear segments in these nanonetworks. Notably, the values of H for these nanostructures are slightly larger than the case of the randomized time series, where $H = 0.5$ within error bars. However, the two structures CNET and NNET, which allow the formation of more communicating paths, exhibit much stronger fluctuations and consequently larger values of the Hurst exponent. Furthermore, the statistics of the returns $d \equiv N_Q(t_k) - N_Q(t_{k-1})$ and the sizes s of clustered events (avalanches) also

suggests a similar grouping of these nanostructures. On one side, the charge fluctuations in SF22 and 1Dwd structures has the exponential distribution of avalanches $P(s) \sim s^{\tau} \exp(-s/s_0)$ (with a larger cut-off s_0 in the case of SF22) and Gaussian distributions of the returns. On the other hand, non-Gaussian fluctuations are found in the case of CNET and NNET. In this case, the distribution of avalanche sizes exhibits a power-law tail compatible with the expressions (3.3) with $\kappa = 1$ and $q_1 \sim 1.33$. Similarly, the tails of the distribution of the returns can be fitted with (3.3), where $\kappa = 2$, i.e., q_2-Gaussian, which is often found in complex signals [188]. It should be stressed that the part of the data for small returns and small avalanches $s < X_0$ in Eq. (3.3) in these two nanonetworks virtually coincides with the ones in the chain and SF22 structures. These weak fluctuations correspond to sporadic tunnelings away from the main conduction paths. This feature is particularly pronounced in the case of CNET, where the most used area of the nanonetwork is reduced due to the point-size electrodes (cf. Fig. 3.26).

Topology of phase space manifolds related with the collective charge fluctuations

The time series of charge fluctuations $N_Q(t_k)$ are converted into QTS-graphs, as described in Section 3.4.2. With analysis of these graphs, we provide a robust topological description of the collective charge fluctuations in different nanoparticle assemblies. For a better comparison, we map an equal segment of each time series in the considered nanoparticle assemblies. Hence, we generate four QTS-graphs of 2000 nodes; each QTS-graph clearly relates with the underlying nanoparticle structure in Fig. 3.26, from which the time series is taken. The utilized mapping rules account for the strength of fluctuations; consequently, the QTS-graphs of different structure are obtained. Fig. 3.29 displays the adjacency matrices of all four QTS-graphs. The pronounced block-diagonal structure in the adjacency matrices for QTS-graphs of CNET and 1Dwd indicates the occurrence of communities of dense links. While in the case of QTS-graphs of NNET and SF22, a sparse structure of connections among the diagonal blocks appears, which is compatible with a hierarchical organization of communities identifiable at the graph level. The ranking distributions of the node's degree and topological dimension of all QTS-graphs are shown in Fig. 3.30. Separate fits for small and large rank according to the discrete generalized beta function are provided, except for the case of FS22, where a satisfactory fit is obtained by a power-law decay with a cut-off (see legends).

A collection of standard graph-theoretic measures of all QTS-graphs is given in Table 3.6, exhibiting small differences at the graph level. However, the higher-

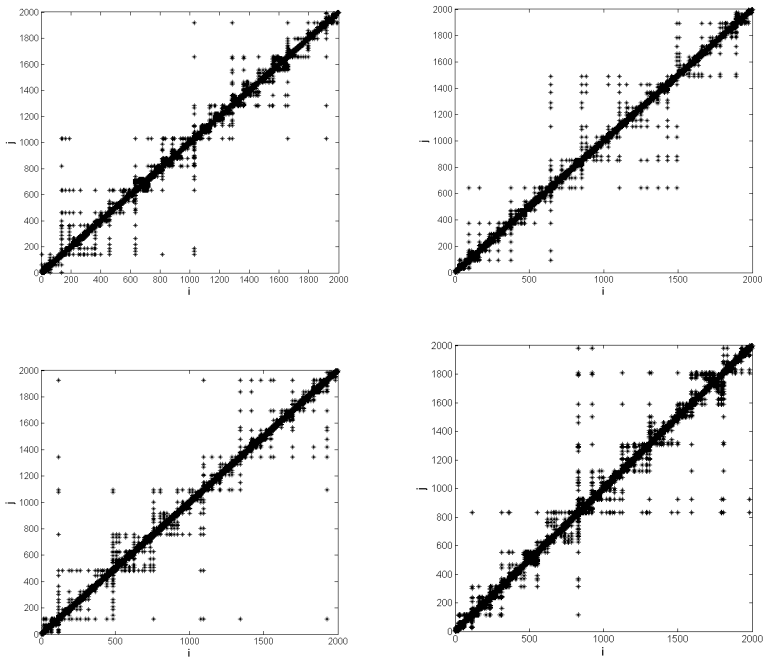


Figure 3.29: Adjacency matrix of the QTS-graph related with the charge fluctuations in the nanoparticle assemblies: 1Dwd and NNET (top row), CNET and SF22 (bottom row) from left to right.

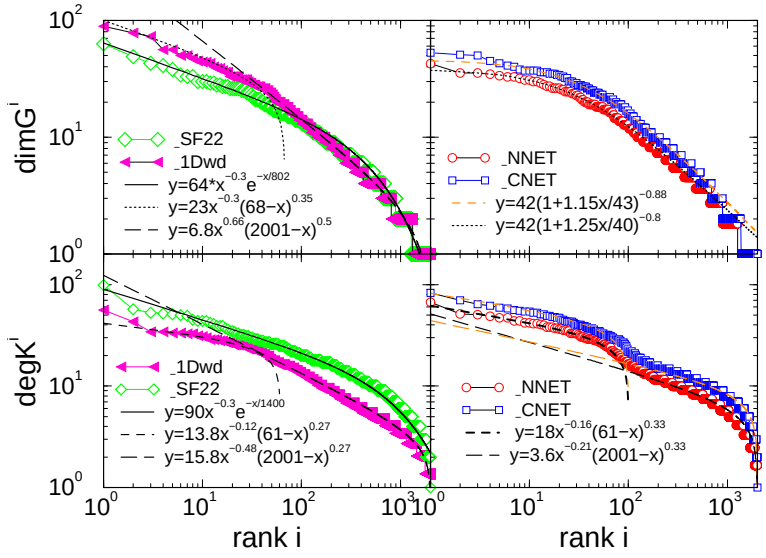


Figure 3.30: Ranking distributions of the node's degree (lower panels) and topological dimension (upper panels) in the studied QTS graphs of charge fluctuations in four nanonetworks from Fig. 3.26.

order structures of these graphs, which are described by the algebraic topology measures, can be considerably different; the results are displayed in Figs. 3.31-3.32 and illustrated by Fig. 3.33.

QTS-graph	d	$\langle k \rangle$	$\langle \ell \rangle$	Cc	No.triang	modul
_1Dwd	10	6.98	5.14	0.77	10314	0.907
_SF22	10	7.27	4.89	0.76	10987	0.897
_CNET	9	8.91	4.66	0.81	17369	0.906
_NNET	11	8.35	4.86	0.79	15418	0.909

Table 3.6: Standard graph theory measures (d–diameter, $\langle k \rangle$–average degree, $\langle \ell \rangle$–average path length, Cc–clustering coefficient, the number of triangles, modularity) of the QTS-graphs representing collective charge fluctuations in nanonetworks of Fig. 3.26.

QTS_	_NNET			_CNET			_SF22			_1Dwd			_randomised		
q	G_q	n_q	$\hat{G}_q$	G_q	n_q	$\hat{G}_q$	G_q	n_q	$\hat{G}_q$	G_q	n_q	$\hat{G}_q$	G_q	n_q	
0	1	1458	0.99	1	1550	0.99	1	1684	0.99	1	1721	0.99	1	1403	0
1	10	1458	0.99	6	1550	0.99	7	1684	0.99	7	1721	0.99	13	1403	0
2	265	1457	0.82	193	1549	0.88	455	1683	0.73	503	1721	0.71	904	1400	0
3	497	1326	0.63	399	1473	0.73	673	1410	0.52	733	1448	0.49	766	873	0
4	570	1025	0.44	594	1238	0..52	638	935	0.31	517	893	0.42	265	283	0
5	430	628	0.32	512	826	0.38	346	418	0.17	268	433	0.38	67	69	0
6	252	313	0.20	351	438	0.20	103	119	0.13	145	179	0.19	12	12	
7	86	105	0.18	128	137	0.07	25	25	0	29	37	0.22	1	1	
8	33	36	0.08	18	18	0				9	9	0			
9	8	8	0	1	1	0									
10				1	1	0									

Table 3.7: Components of three structure vectors of the graphs representing charge-fluctuation time series in nanonetworks from Fig. 3.26.

The higher-order structures of the QTS-graphs are revealed by determining the components of three structure vectors, which are defined in section 3.4.2. The results are displayed in Fig. 3.31. Their numerical values are summarized in Table 3.7. For a comparison, we also show the results for a graph, which is obtained from a randomized time series; for this purpose, such series is obtained by interchanging randomly-selected pairs of data points in the time series from NNET.

Notably, the randomized time series results in a graph exhibiting much simpler structure than the other graphs, which represent the fractal time series. The topological complexity of these QTS-graphs manifests in the occurrence of a vast number of topological levels $q = 0, 1, 2 \cdots q_{max}$ in accord with the number of combinatorial spaces and their interconnections at higher topological levels. As an example, in Fig. 3.33 we display the QTS-graph of NNET assembly. In two separate plots we also demonstrate the complexity of its top topological levels, $q = 8$ and $q = 9$. Specifically, it contains eight 10-cliques; three 10-cliques are separated, and the other five are interconnected via 9-cliques at the lower level $q = 8$. Here, according to the

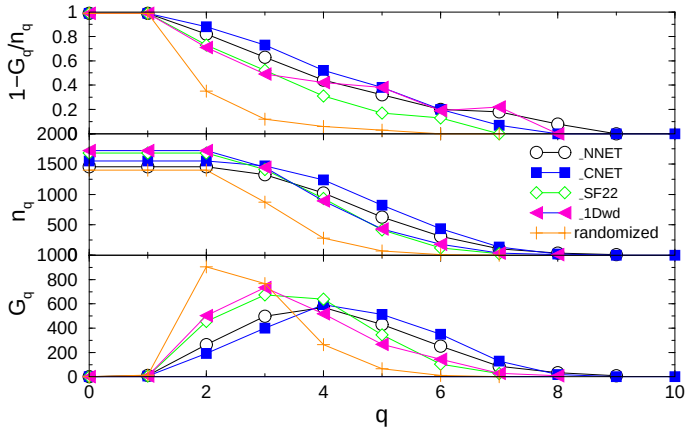


Figure 3.31: Components of the 1st, 2nd and 3rd structure vector, G_q, n_q and $\hat{G}_q = 1 - G_q/n_q$, respectively, plotted against the topology level q in the QTS-graphs related with the time series of charge fluctuations in the underlying nanoparticle structures, which are indicated in the legend, and a randomized time series.

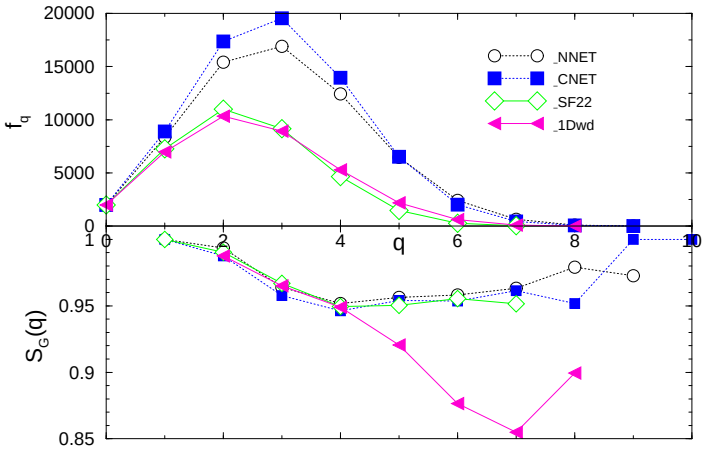


Figure 3.32: The entropy $S_Q(q)$ and the number of simplexes and shared faces f_q at the topological level q of the QTS graphs of the charge fluctuations in the indicated nanoparticle assemblies.

table 3.7, 36-8=28 additional 9-cliques exist and are interconnected such that 33 components occur at the level $q = 8$. That is, 33-28=5 components are identifiable at the upper level, cf. Fig. 3.33 bottom left. Considering the level below, $q = 7$, gives 105-36=69 new 8-cliques and 86 components, which suggests that maximally 86-69=17 groups of nodes can be distinguished at the level $q = 8$. Among these, five groups contain higher cliques, leaving at most 12 groups that are visually distinct in the top right figure 3.33. In contrast, the single 11-clique in QTS-graph of CNET remains isolated down to the level $q = 7$ (cf. table 3.7). Together with 17 other 9-cliques that appear at the level $q = 8$, they make 18 components. At the level below, we identify 137-18 =119 new 8-cliques. Then comparing the number of com-

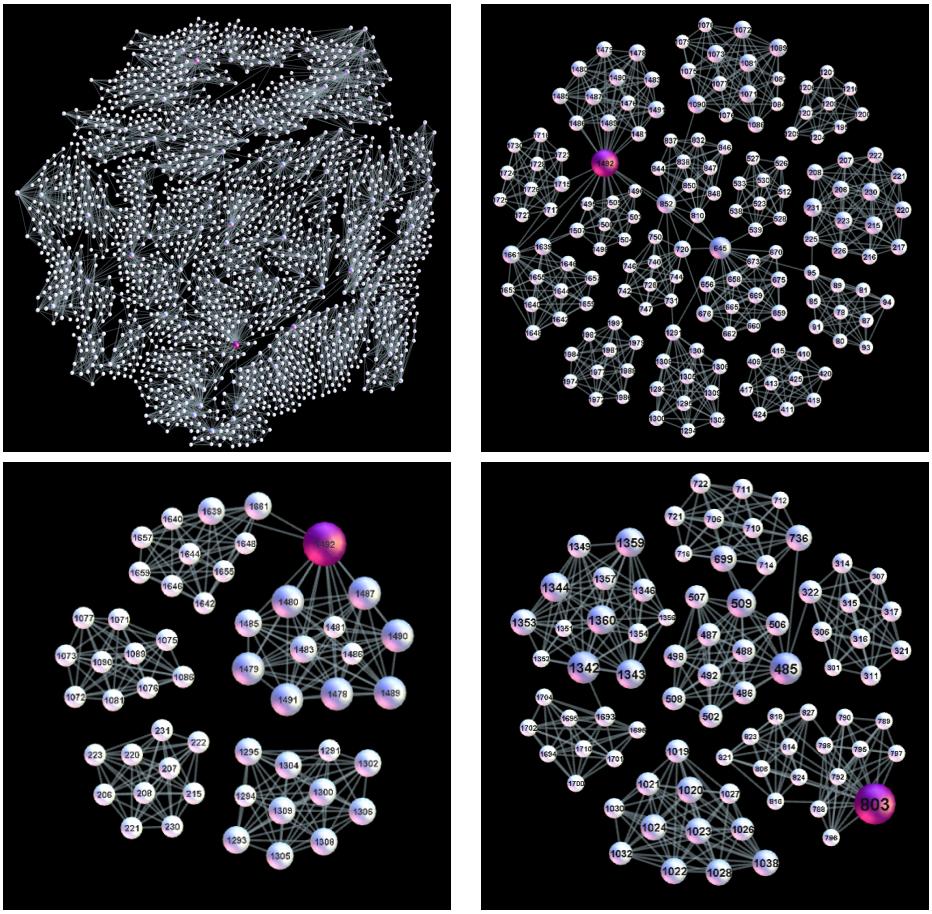


Figure 3.33: The structure of connections of the phase space manifolds corresponding to the charge fluctuations in the nanoparticle network NNET: complete QTS_NNET graph (top left) and its highest topological layers $q = 8$, top right, and $q = 9$, bottom left. For comparison, the layer $q = 8$ of the QTS_CNET graph is displayed, bottom right. In figures of the higher topological layers, ID of each node refers to the index of the time interval t_k of the time series. All edges between these nodes that exist in the adjacency matrix are shown.

ponents, 128-119=9 suggests that the cliques at the upper level $q = 8$ make at most nine groups, one of which refer to the 11-clique of the top levels. This structure is also shown in Fig. 3.33 bottom right.

According to Fig. 3.31, the second structure vectors of QTS-graphs of SF22 and 1Dwd exhibit certain similarity. Interestingly enough, the presence of charge disorder in the linear chain of nanoparticles leads to a more complex QTS-graph than the SF22 assembly. Namely, the charge disorder blocks the only existing path in the linear chain, which results in a sequence of zero entries in the time series, which is mapped onto a large clique linked to two or more near non-zero entries. In contrast, blocking a linear segment in SF22 structure causes tunnelings along the alternative paths. Hence, no abrupt changes in the time series occur. Apart from the high

topology levels, the similarity between the QTS-graphs of NNET and CNET, on one side, and SF22 and 1Dwd, on the other, is apparent at lower q-levels, cf. Fig. 3.31. These findings also apply to the topological response function f_q and the entropy $S_Q(q)$, shown in Fig. 3.32. The occurrence of mutually isolated cliques, i.e, at $q = 1$ or $q = q_{max}$, leads to the occupation probability $p_q^i \lesssim 1$ and the entropy close to zero. The simplicial complexes that occur at the intermediate q-levels share the faces at the level $q - 1$. Hence, same nodes participate in the identified complexes, leading to a lower occupation probability of the level and the entropy drops, as shown in Fig. 3.32. The 2-dimensional nanoparticle assemblies exhibit complex QTS-graphs in which a large number of simplicial complexes, sharing many nodes, is observed. In contrast, a small number of such compounds occur in the case of the linear chain structure with charge disorder. Thus, a reduced number of nodes are involved; this situation manifests in a more pronounced entropy minimum at $q_{max} - 1$ than in the other graphs.

3.5 Multi-brain connectivity networks and functional patterns during spoken communication

Human behavior in various circumstances mirrors the corresponding brain connectivity patterns, which are suitably represented by functional brain networks. While the objective analysis of these networks by graph theory tools deepened our understanding of brain functions, the multi-brain structures and connections underlying human social behavior remain largely unexplored. In this study, we analyze the aggregate graph that maps coordination of EEG signals previously recorded during spoken communications in two groups of six listeners and two speakers. Applying an innovative approach based on the algebraic topology of graphs, we analyze higher-order topological complexes consisting of mutually interwoven cliques of a high order to which the identified functional connections organize. Our results reveal that the topological quantifiers provide new suitable measures for differences in the brain activity patterns and inter-brain synchronization between speakers and listeners. Moreover, the higher topological complexity correlates with the listener's concentration to the story, confirmed by self-rating, and closeness to the speaker's brain activity pattern, which is measured by network-to-network distance. The connectivity structures of the frontal and parietal lobe consistently constitute distinct clusters, which extend across the listener's group. Formally, the topology quantifiers of the multi-brain communities exceed the sum of those of the participating individuals and also reflect the listener's rated attributes of the speaker and the

narrated subject. In the broader context, the presented study exposes the relevance of higher topological structures (besides standard graph measures) for characterizing functional brain networks under different stimuli.

3.5.1 Introduction

In the past few years, a big leap in understanding the structure and function of the human brain has been provided with both advances in brain imaging techniques [200, 201] as well as the use of complex networks perspective to analyze the emerging empirical data [202]. Currently, active research differentiates two aspects of brain networks, representing anatomic and functional connections between distinct brain regions [203]-[206]. Anatomical connections are chiefly investigated by diffusion tensor imaging. The functional brain connectivity, on the other hand, can be detected at different spatial and temporal scales. In this regard, functional magnetic resonance imaging (fMRI) captures synchronization among blood-oxygenation-level dependent signals at a good spatial resolution and low frequency. Much shorter time scales can characterize the brain connections related to different brain function, for instance, information processing, integration or segregation, cognitive control, empathy, and other. Therefore, electroencephalography (EEG) imaging has received an increased interest in functional brain research [207]-[211]. In this case, the functional connections are often reconstructed from EEG signals recorded at many scalp locations. In contrast to fMRI imaging, which measures spatially specific cortical or subcortical regions, the signal registered by an electrode at a particular scalp location (i.e., above a cortical region of interest) is spatially less specific, containing the average electric neuronal activities of all voxels belonging to that area [207, 211]. Nevertheless, regarding the generalized synchronization, the recognizable patterns of positively correlated EEG signals suitably reflect the macroscopic organization of the brain network [202]. Thus, the underlying brain activity corresponding to a variety of situations has been analyzed through EEG-based connections, for example, processing (un)pleasant music [212], the objective identification of emotions [209] or the pathological changes in the context of epilepsy [210], anesthetic agents induction [213], and other.

Brain anatomical connections are suitably represented by weighted networks. In this case, there is a growing consensus about the confidence level that a particular link is present as well as its weight [214]. On the other hand, a variety of functional connections have been observed, closely reflecting a particular brain activity, that map to a different functional network [203]. Such examples of the brain networks include the recently studied functional paths in integration and segregation of in-

formation [215], inter-regional communication [216], convergence of information in hippocampus [217], stimulus selection [218], cognitive control circuits [219], as well as the effects of different stimuli [218, 220, 221], learning [222], perception of time, numbers and languages [223, 224], the presence of a mental disease [225] and more. Although the anatomical connections lay the basis, the functional brain networks often appear to have a richer structure, which is attributed to dynamical factors: the appearance of longer paths as well as the avalanches of the cascading activity propagation. Thus, clearly distinguishing between the brain activity patterns related to particular mental processes remains a challenging task.

In the neuroscience research, a central problem is how to explain the brain function from microscopic, biochemical processes, on one side, and its impact on human behavior, on the other. Recently, concerns on how the brain mediates social interactions, which is closely related both to cognitive and affective neuroscience, lead to the development of the social neurology. In this respect, the research of the social impact on the processes in the brain [226]-[228] as well as the brain processes that underly an effective social behavior [229]-[237] is becoming a subject of increasing interest. New approaches using simultaneous scanning of groups of participants are being developed to study social cognition [238, 239]. The studies of face-to-face communications in dyads reveal a significant degree of synchronization in particular brain areas, depending on the performed task [233, 234, 235, 237]. Whereas, other types of communications seem to involve different mechanisms [234]. Furthermore, a higher brain activation level characterizes the leader compared to its follower, the game builder compared to its partner, as well as the same individual performing the cooperation compared to the competition role [236]. In contrast to the extensive study of functional brain connections, as stated above, little attention has been devoted [237, 239] to analyze multi-brain graphs and to identify the social impact onto the functional brain networks.

In this work, we study a complex network of brains of a group of individuals during spoken communications; we map an aggregate data of EEG signals, previously recorded in the experiment described in [240]. In the experiment, two different narrations of the speakers are superimposed and presented to the two groups of six listeners at the same time, while a group's aim was to focus to a particular speaker. The EEG signals were recorded simultaneously at all listeners during the session, while the recording was performed independently for each speaker during his/her narration, see 3.5.2. Note also that, in contrast to the platforms build on the face-to-face communications [238], in this experimental set-up the interaction is unidirectional from speakers to listeners. The self-rated experiences of the listeners

collected after each completed session, indicate wide variations of the listener's concentration to the story, correlating to the previous knowledge, interest, as well as the speaker's attractiveness and narrative quality.

The previous study of the data [240] focused to the statistically significant group-averaged features, in particular of the speaker-listener coordination. The aim of the present analysis is entirely different. We reveal the fine structure of the aggregate multi-brain network considering the location of each EEG electrode as a network node. Thus our approach, based on mapping the correlations among EEG signals, allows us to analyze the networks of connections elicited by the presented stimulus in each listener's brain as well as the speakerlistener and listener listener coordination, and the emerging cross-brain structures. We hypothesized that, while involving the similar brain areas, the brain activity patterns of each participant (depending on its role in the session, cognitive and emotional content communicated by the speakers, and other factors) results in different connectivity in the corresponding functional networks and brain-to-brain connections. To analyze these significant differences between brain networks, we develop an approach based on the algebraic topology of graphs, which identifies higher-order structures containing cliques of a large order and their aggregates beyond the standard network parameters. The applied methodology offers a new perspective in the analysis of inter-brain synchronization during social communications and other functional brain networks besides the social brain problems.

3.5.2 Materials and Methods

EEG data acquisition and preprocessing

We use the empirical data of two sessions, each session containing a different stimulus, from the previously recorded data set of [240]. A stimulus consists of two superimposed audiovisual recordings of a similar duration (about 4 minutes) where two speakers, female and male, narrate different stories to the video cameras. As described in [240], a specific software was used to superimpose onto each other the two faces of the speakers and the soundtracks of their voices. Thus obtained video was adjusted so that both speakers appeared equally prominent. The stimulus is then presented to twelve listeners. While the input was the same to all listeners, one group of six listeners was instructed to attend to one speaker and the other group to the other speaker. The attended speaker was introduced to each panel in the first 5 seconds of the stimulus. The schematic illustration of the experimental setup is shown in Fig 3.34. In each stimulus, EEG recordings from two speakers

are made during freely narrating stories, while the EEG scanning of all listeners is performed simultaneously during the session. We consider data of two different stimuli, in particular,

- Stimulus1 consists of the storytelling by a male speaker (S1) and a female speaker (S2), and the corresponding two groups of six listeners. Recalled by the speaker's own words, the narration is based on written versions of international fairy tales.

- Stimulus11 contains an ad-hoc invented narration by a female speaker S3, attended by the listener's group 2, and a free narration from the favorite book or movie of the speaker S1 (the same speaker as in stimulus1), attended by the group 1. For these narrations, no written text was available to the speakers.

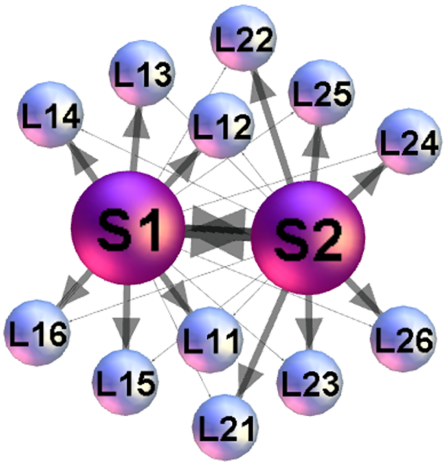


Figure 3.34: The listeners $k = 1, 2...6$ of each group L_{1-k} and L_{2-k} are instructed to follow the narration of a particular speaker, indicated by the heavy lines, while having the narration of the other speaker simultaneously accessible, shown by the thin lines.

The EEG signals were acquired from 63 scalp locations positioned according to the International 10/20 System. A detailed description of the EEG acquisition and preprocessing is given in the original paper [240]. According to [240], the preprocessing included the necessary steps in which (i) EEG recordings were aligned with the corresponding audio signal for each participant; (ii) the noise was reduced by applying 50Hz notch filter; (iii) artifactual components (i.e., due to speaking, eye movement) were removed using an advanced technique based on the independent component analysis, which can separate the signal components of different origin, and rating. Moreover, the data were transformed to the average reference and temporally trimmed to the overlapping time segments.

The data also contain a questionnaire, where we use the part of information related to the analysed stimulus1 and stimulus11. Specifically, we consider the listener's self-rating of the concentration, prior knowledge, interest and understanding of the story, as well as the sympathy to the appointed speaker, the speaker's narrative quality and attractiveness.

Mapping EEG data to network and filtering relevant links

Different methods are in use for mapping the brain activity signals to a graph, see review articles [202, 241]. Each recording method and, consequently, the networks extracted from it have certain limitations. For example, fMRI has good spatial resolution, and the network nodes have anatomical locations in the brain, but the low temporal resolution to detect the neuronal activity. On the other hand, the EEG recordings are at the right temporal resolution, but the network's nodes represent the locations of the electrodes on the scalp. However, by invoking a general operational principle of complex networks analysis [202], both of these microscopically distinct networks contain relevant information about the macroscopic organization of brain function.

Here we apply the methods of correlation matrix that was widely applied for mapping EEG signals [207, 242], stock market data [243], gene expression data [43, 44] and traffic signals [244]. In this case, correlations among EEG signals describe functional connectivity patterns between different brain areas, whose activity is recorded in 63 points at the scalp. Thus, the data consists of 882 time series. The length each time series is 83499 time steps in stimulus1, and 120911, in stimulus11, where one time step corresponds to 1/500 sec. According to [240], the interesting correlations are expected when the delay 12.5s (6250 time steps) is considered between speaker's and listener's time. In this case, we skip the first 6250 points in the speaker's time series and the last 6250 time points in the listener's. Else, the speakerspeaker and listenerlistener correlations are determined without any delay. First, the Pearson's coefficient is computed for each pair (A_i, B_i) of time series

$$C_{AB} = \frac{1}{N-1} \sum_{i=1}^{N} (\frac{A_i - \mu_A}{\sigma_A})(\frac{B_i - \mu_B}{\sigma_B}),$$

where μ and σ are the mean and standard deviation of the corresponding time series, and N is the length of time series. Further, to separate the strong positive correlations, which are relevant in almost every context, there is way to use the filtering algorithm described in [43, 44, 45]. The algorithm enhances those matrix elements C_{ij} that have a similar correlation pattern with the rest of the

matrix elements while diminishes those with a dissimilar patterns. First we map C_{ij} to the range $[0, 1]$ using $CP_{ij} = (C_{ij} + 1)/2$. Then, each element is multiplied $CP_{ij} \rightarrow F_{ij}CP_{ij}$ by the corresponding factor F_{ij}, which is computed as Pearson's coefficient of the rearranged matrix elements from row i and column j as follows: $\{C_{ij}, C_{i1}, C_{i2}, ..., C_{iN}\}\{C_{ji}, C_{1j}, C_{j2}, ..., C_{jN}\}$. The resulting filtered correlation matrix is also transfered to a binary adjacency matrix of the graph by retaining the correlations larger than a threshold value and inserting units for the retained edges.

Standard graph-theory measures & network communities

In graph theory [245], a graph is a mathematical object consisting of nodes and edges connecting the nodes; to characterise the graph structure, some measures [241, 246] are computed, here termed the standard graph measures to emphasise the distinction between the algebraic-topology measures, defined below. Given the specific type of graphs, for this work we determine the diameter of the graph d, the graph density ρ, the average degree $\langle k \rangle$ and path length $\langle l \rangle$, and the clustering coefficient Cc, as well as the community structure.

Communities on a network are identified as densely connected subgraphs; these are groups of nodes in which each node has more connections with the other members of the group than with the nodes outside of the group. To determine the network's community structure [247, 248] in these particular type of networks, we use the appropriate method based on the maximum modularity [249]

Comparing SB networks: Links overlap

Statistical measures of (dis)similarity between a considered pair of functional single-brain (SB) networks are quantified. In particular, we identify the fraction of the links E that are identical in both networks relative to the total number of connections in both networks. That is, for a pair of the listener's single-brain-networks, the overlap measure is given by

$$O(L_{ix}, L_{jy} = \frac{E(L_{ix}) \cap E(L_{jy})}{E(L_{ix}) \cup E(L_{jy}},$$

where $i, j \in \{1, 2\}$ denotes the group to which the listener belongs, $x, y \in 1, ..., 6$ and denotes the listener's identity number. Here L stands for listener; the analogous expression applies to any pair of L–S and S–S networks.

Comparing SB networks: Graph-edit distance (GED) GED counts the number of links that have to be deleted so that the two

GED counts the number of links that have to be deleted so that the two matrices become equal [250]. These counts are then used to position all listeners relatively to the two speakers in a 2Dcoordinate space. We set S_1 in the origin $(0, 0)$, while S_2 on the coordinate $(s_{12}, 0)$, where s_{12} is the GED between S_1 and S_2. The (x, y) coordinates of a listener are then calculated using trigonometric functions, where the length of the sides of the triangle are computed GED between the listener and both speakers.

Randomization of SB networks

We apply two procedures to randomize the network connections. In Random-K, the procedure preserves the node's degree. Starting from the original list of links, considered as oriented, for each link we find a randomly selected link in the list of all links and cross switch the outlinking between the corresponding pairs of nodes. We repeat the process until each link is switched at least once. We also apply the fully randomised procedure, here termed Random-L, which preserves the total number of connections in the network. In this case, while cutting a link from the original list, we insert a new link among a randomly selected disconnected pair of nodes.

3.5.3 Results

Correlation matrix and networks of speakers' and listeners' EEG signals

We analyse two different stimuli from the corpus of the previously recorded and pre-processed data of Ref. [240]. A full description of the analysed datasets is given in Methods. In each stimulus, we consider EEG signals recorded during the session at 63 locations on the scalp of two speaker and 12 listeners, in total 882 signals. The correlations among each pair of signals are computed as Pearson's coefficient. Between the speaker's and listener's events, we take into account the characteristic delay of 12.5s, which was observed and related to the processing of semantic content in the original work [240]. The equal-time correlations are considered among signals of the pairs of listeners. The correlation matrix is filtered (see Methods) to diminish the potentially redundant correlations. Finally, a threshold value (in this work $w_0 = 0.06$) is applied to remove weak correlations that are assumed to be normally distributed, Fig 3.35a and 3.35b. To select the right threshold, we observe the criteria that the modular decomposition, as chief feature of functional brain networks [247, 251], shows consistency in the multi-brain networks for the values above the selected

threshold. Also starting from this threshold value, each single-brain connectivity splits into a frontal and parietal cluster, the structure characteristic to awaken state according to Ref [213], whereas they may join to a single cluster when the threshold is lowered, see Fig 3.36. Moreover, we checked that the crossbrain connections around this threshold form a sparse network but sufficiently connected to ensure nontrivial structures, as the example in Fig 3.35c; the occurrence of large simplicial complexes tested by the node's topological dimension stabilises around the selected threshold value (see Fig 3.35a,top). The adjacency matrix of a binary multi-brain graph is then constructed, i.e., $A_{ij} = 1$ when the matrix element $w_{ij} > w_0$ and zero otherwise, and shown in Fig 3.35d. In this aggregate multi-brain network, the single list of indexes indicates the standard abbreviations of the 63 EEG scalp locations belonging to, respectively, speaker S_1, speaker S_2, then two groups of listeners L_{1-k} and L_{2-k}, $k = 1, 2, ..., 6$. Thus, the diagonal blocks in the adjacency matrix represent the 63×63 connectivity matrices related to each brain in the above-defined order. Whereas, each off-diagonal block contains the corresponding brain-to-brain correlations.

Here, we first analyze the structure of each diagonal block and some selected cross-correlations in the aggregate network in Fig 3.35d. Then we turn to the analysis of the whole multi-brain network, which allows for quantifying the speaker's impact onto the listener's brain activity patterns. Each diagonal block of the aggregate adjacency matrix in Fig 3.35d represents a functional brain network related to a particular participant, as described below. In the present context, these separate subgraphs are termed single-brain-networks (SBN) to distinguish them from the studied multi-brain structures. Fig 3.36 shows examples of SBN obtained at two different thresholds.

Single-brain networks (SBN) of speakers and listeners: Quantifying the topological differences

In Fig 3.37 the single-brain networks representing the correlation of the EEG signals on the scalp of both speakers and the listeners in both groups are shown. As we stated in the Introduction,the focus is on the occurrence of higher-order structures in SBN and other relevant subgraphs of the multi-brain network. In this respect, the topology levels and the corresponding vectors defined in Methods are first computed for all SBN, as the relevant units of the multi-brain graph. Before turning to the topology analysis, for comparison, we also determine the standard graph-theoretic measures for all SBN; they are summarized in Table 3.38. Further, we analyze the (dis)similarity between these functional connections in all pairs of SBN.

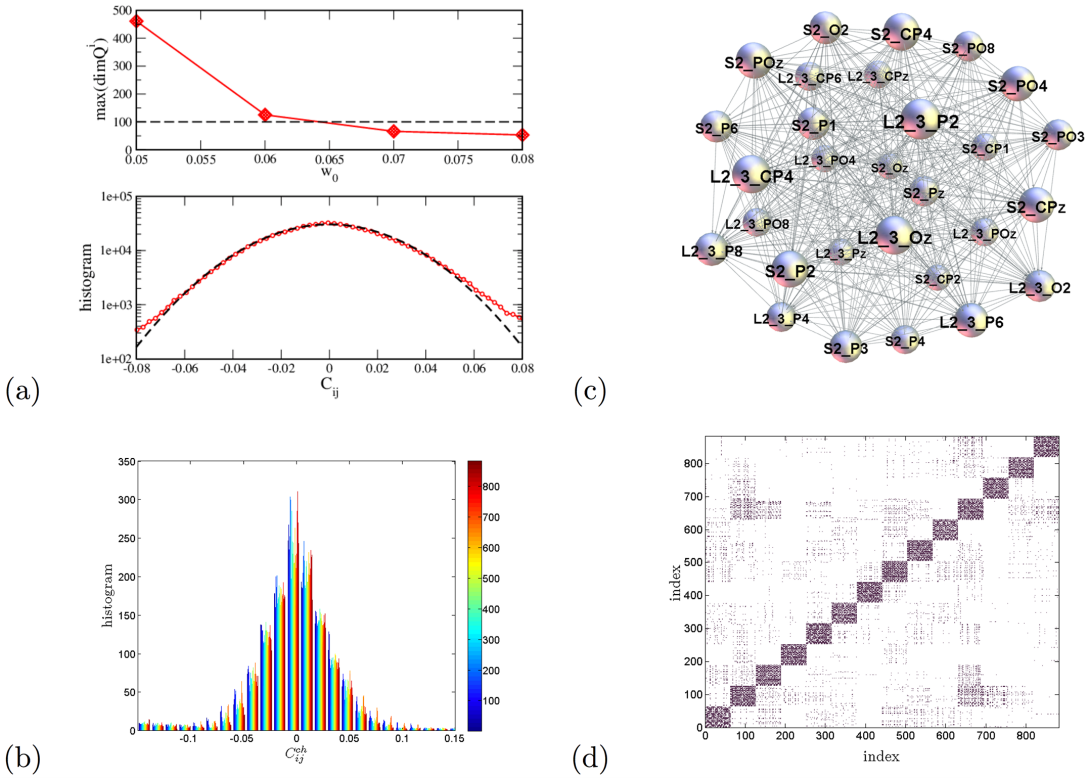


Figure 3.35: Maximum topological dimension plotted against various thresholds w_0 (top panel) stabilises near the selected threshold 0.06. The middle part of the histogram of correlation coefficients averaged over 882 channels (lower panel); the fit by the normal distribution, dotted line, deviates from the data for the correlations larger than 0.06. (b) The central part of the histograms for different channels. For a particular channel, the correlations with other 881 channels are marked by the corresponding colour indicated in the colour map; the presence of colours over different bins suggests that all channels obey a similar distribution. (c) An example of higher-order structure involving the signal locations at two scalpsthe speaker's S_2 and the listener's L_{2-3}. (d) The adjacency matrix of the multi-brain network of the two speakers and 12 listeners with the threshold $w_0 = 0.06$. The order of indexes is as explained in the text.

Comparing two networks is widely used in the literature to uncover the nodes and links that are responsible for a particular pattern, e.g., disease-related connections [252], or to infer the relevance of a particular link or a node [253]. In the present context, we expect that speakers and listeners use different activation of the brain areas to process and perceive the semantic contents during the communication. Thus, the corresponding brain activity patterns, reflected in several positively correlated EEG channels, can be differentiated through the differences in the SBN's topology. First, we compare the pairs of SBN by examining the presence/absence of a particular link (the correlated pair of channels). In this regard, the differences in the activity patterns are observed among both speakers and, evoked by the stimulus, among the

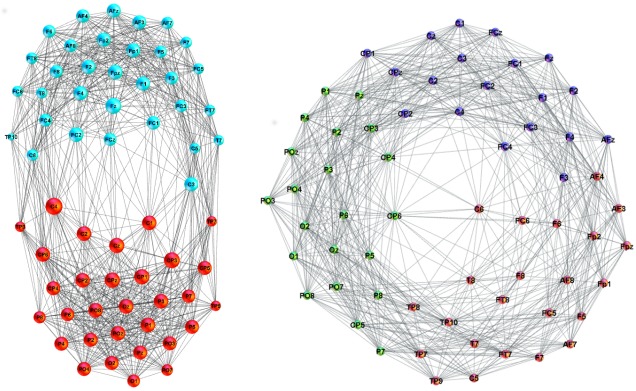


Figure 3.36: (left) An example of network mapping the functional brain connections recorded by EEG signals on the speaker's scalp (labels). The applied threshold $w_0 = 0.06$. The colour of nodes indicates two identified communitiesfrontal (F) and parietal (P). (right) Loss of the F/P community structure at a lower threshold $w_0 = 0.05$ in the correlation matrix.

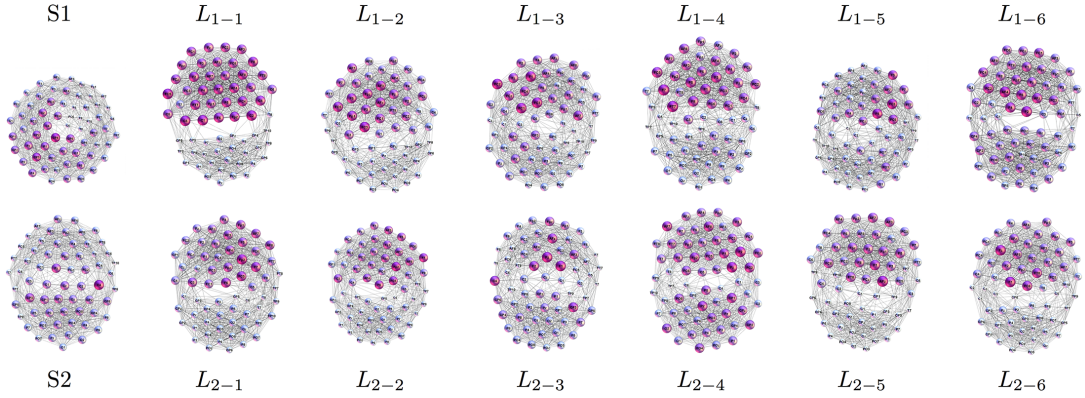


Figure 3.37: In each network, nodes represent different scalp locations (labels) while each link indicates the positive correlation that exceeds the threshold among the corresponding pair of EEG signals. Remarkably, each connectivity network visually splits into two clusters, which can be identified by the location labels as frontal (upper) and parietal (lower), also confirmed by the community detection analysis. See an example of a larger picture in Fig 3.36.

listeners within the same group as well as across the groups. Here, we quantify these differences using two methods of graph comparisons. First, we determine the extent to which the edges overlap (see Methods) in each pair of single-brain networks. The resulting statistics is displayed in Figs 3.39 and 3.40 for the situation in stimulus1 and simulus11, respectively. Furthermore, the occurrence of the excess links leading to a finite distance between the pairs of single-brain graphs in the topological space is expressed by graph-edit-distance, as described in Methods.

Link overlap statistics. The link-overlap statistics expresses the degree of similarity between each compared pair of SBNs. In Fig 3.39, we show the outcomes for

the overlaps for the listeners of both groups in stimulus1 with both speakers, as well as the overlaps among the pairs of listeners in both groups, exposed to the same input. Similarly, the overlaps for stimulus11 are shown in Fig 3.40.

As the Fig 3.39 shows, the brain activity patterns of both groups of the listeners in stimulus1 express a larger similarity to speaker S_2 than to speaker S1. It is also interesting to notice the overlaps among pairs of listeners in each group. Noticeably, the overlap between SBN of the two speakers is low (minimal overlap in the entire set). The situation is just opposite in stimulus11, where the same speaker S_1 features, however, narrating a different subject. While the overlap between the two speakers is larger than in stimulus1, four listeners in group 2 have a better overlap with speaker S_1 than with the attended speaker S_3. This situation manifests in a specific heterogeneity of the overlaps between the listeners in group 2. It is important to stress that, while the edges overlap varies among different pairs, it always stays significantly above the corresponding values for the randomized models.

SBN's distance between listeners and speakers. The excess links, which are present in an SBN but do not overlap with another SBN, represent a measure of distance between these networks in the topology space. Applying the graph edit distance (GED), as described in Methods, we first compute the distance between SBNs of speaker S_1 and speaker S_2. Then the distances of each listener from both speakers are calculated and presented by a point in the distance plane in Fig 3.41; two panels are for the stimulus1 and stimulus11, respectively. For a better comparison, both groups of listeners are plotted on the same graph.

The following features of the distance graphs are interesting. First, in the stimulus1, the majority of the listeners from both groups are closer, suggesting a larger similarity in the brain connectivity patterns, to the speaker S_2, than to the speaker S1. Excluding the listeners L_{1-1}, L_{2-2} and, to some extent, the listener L_{1-3}, the listeners of both groups form a cluster in the distance plane. Notice that, by definition of GED, the closeness of two listeners in this graph is referring to the similar fraction of the removed links. However, by comparing the exact links, the two listeners may have a finite distance from each other.

These properties of the distance plots are in good agreement with the histograms of the pairwise overlap in Fig 3.39. Namely, for the stimulus1, except for L_{2-2} and L_{1-3}, the speaker's S1 overlaps with both groups are worse than the overlaps of the speaker S_2 with the listeners in both groups. Moreover, the listeners in the centre of the cluster, for instance, L_{1-2} in group 1 and L_{2-6} in group 2, have similar overlaps with the remaining members of the group. Further comparisons of the listeners, for instance, L_{2-3} and L_{1-4}, who are quite close in the distance space but have different

SBN	S1	S2	L_{1-1}	L_{1-2}	L_{1-3}	L_{1-4}	L_{1-5}	L_{1-6}	L_{2-1}	L_{2-2}	L_{2-3}	L_{2-4}	L_{2-5}	L_{2-6}
<k>	25.8	27.7	29.0	28.1	27.2	27.3	26.9	27.7	27.4	27.2	26.1	27.6	28.0	27.7
<t>	1.64	1.63	1.71	1.62	1.63	1.66	1.65	1.70	1.66	1.65	1.69	1.65	1.63	1.67
Cc	0.676	0.743	0.847	0.745	0.717	0.749	0.726	0.785	0.759	0.757	0.736	0.759	0.752	0.763
ρ	0.416	0.448	0.468	0.453	0.440	0.441	0.435	0.447	0.442	0.441	0.421	0.444	0.452	0.447
mod.	0.307	0.368	0.412	0.346	0.346	0.370	0.355	0.414	0.382	0.382	0.379	0.390	0.352	0.373
n.c.	3	2	2	2	3	2	3	2	2	2	2	2	3	3
q_{max}	16	19	29	23	19	23	20	22	18	22	20	20	22	24
$n_{q_{max}}$	1	6	1	1	2	1	4	1	1	1	2	5	6	1

doi:10.1371/journal.pone.0166787.t001

Figure 3.38: Graph-theoretic characteristics of single-brain connectivity networks. Standard graph measures denoted in the first column (see Methods) for each SBN of the two speakers and two groups of listeners listed in the top row; data are for the case of stimulus1. For comparison, we also show the number of topology levels q_{max} and the number of cliques of the highest order in the corresponding network.

appointed speakers, reveals different cross-brain correlations, see later.

In the stimulus11, the same speaker S_1, here narrating a different type of story, attracts more attention of the listeners in both groups than the new speaker S_3. In this context, some striking examples are the listener L_{2-6} and L_{1-6}, cf Fig 3.41. Note that in this case the listeners L_{1-6}, L_{1-3}, L_{1-4}, L_{2-3} and L_{2-5} also form a group in the distance plane. In the analogy with the above-discussed stimulus1, these findings of the distance graph are in agreement with the corresponding overlaps in Fig 3.40 for stimulus11. Hence, the impact of a speaker may strongly depend on the narrating subject; formally, its corresponding brain activity results in a different network (see more details in Discussion). These features of SBNs are also reflected in the appearance of higher organised structures, as discussed in the following sections.

It is interesting to compare these objective graph-theoretic measures with the subjective experience of each listener; the self-reported ratings of self-concentration. While in both stimuli, all listeners reported no prior knowledge of the story, their interest varies in correlation with the self-reported narrative quality of the speaker. The listener's report of weak sympathy to the speaker, low speaker's narrative quality and attractiveness correlates well with the increased distance between speaker-listener SBN and low overlap. Oppositely, self-reported sympathy to the speaker, qualifications as attractive and good narrator agree well with the increased SBN overlaps and reduced topological distance, Figs 3.39, 3.40 and 3.41.

Topological spaces of single-brain networks and inter-brain linking. As mentioned in the Introduction, we anticipate that the key features of the brain activation of an individual during social communication are contained in the hierarchical organization of links between involved brain areas, which leads to the occurrence of topological complexes. In the multibrain graphs, the occurrence of higher-order structures is manifested in two ways: (i) the appearance of hierarchical organisa-

tion along different topology levels in each SBN, and (ii) the inter-brain correlations leading to a nontrivial community structure of multi-brain graph

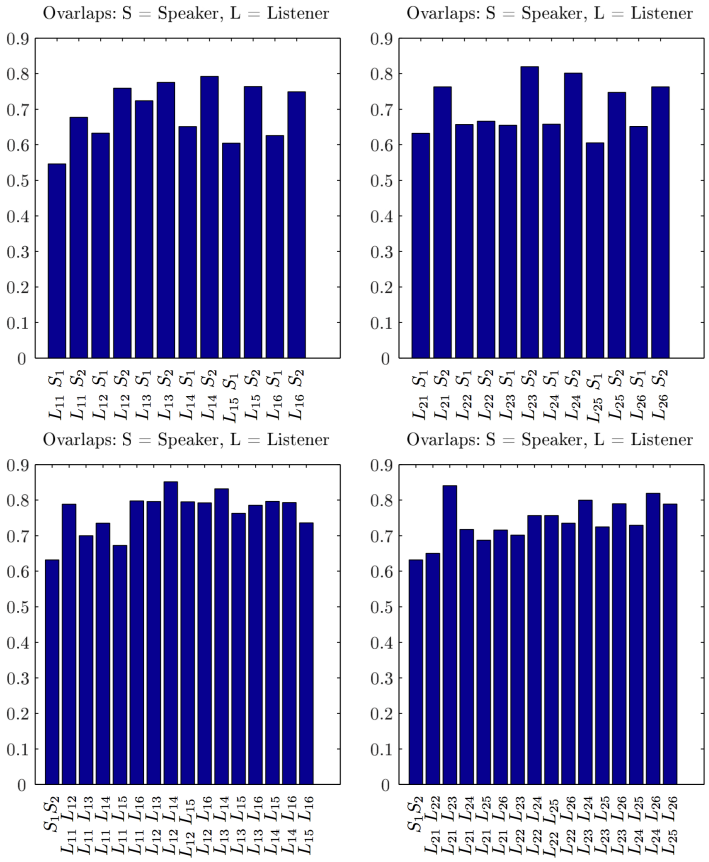


Figure 3.39: The similarity of single-brain networks in stimulus1. SBN overlaps between each listener in group 1 with speakers S_1 and S_2, the left panel, and each listener in group 2 with S_1 and S_2, right panel. Bottom left and right panels show the SBN overlaps between pairs of listeners in the group 1 and group 2, respectively. For a comparison, the SBN overlap between two speakers is also shown (the first bar in each bottom panel). Note that all overlaps are significantly larger than in the corresponding randomised models.

as well as the hierarchical organisation of these communities. In comparison with known heuristic approaches for hierarchical communities [251, 254], here we rely on mathematically strict definition of simplexes and simplicial complexes, as described in Methods. First, we compute the structure vectors defined in Methods to describe the higher-order structures in the individual brain connections both for speakers and listeners. In Fig 3.42, we show the results of the first and third structure vectors for each of 14 single-brain networks in the case of the stimulus1.

Performing the Q-analysis of each SBN, the components of the structure vectors Q_q and n_q are computed at each topology level $q = 0, 1, ...q_{max}$, where q_max is the order of the highest clique found in the corresponding brain connectivity graph.

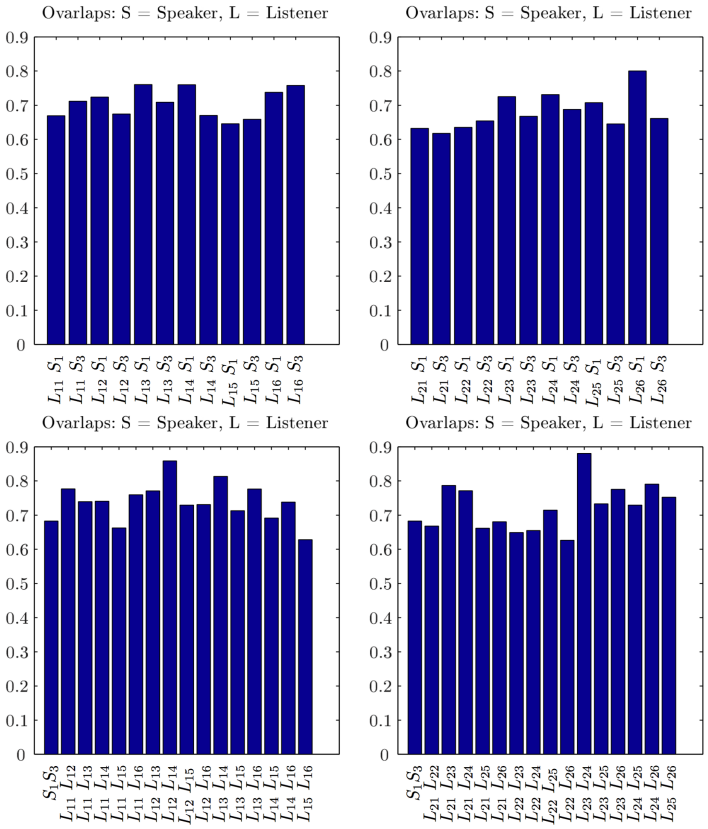


Figure 3.40: The similarity of single-brain networks in stimulus11. Link overlaps of the two listener's groups, explanations as for Fig 3.39, but with the speakers S_1 and S_3, respectively.

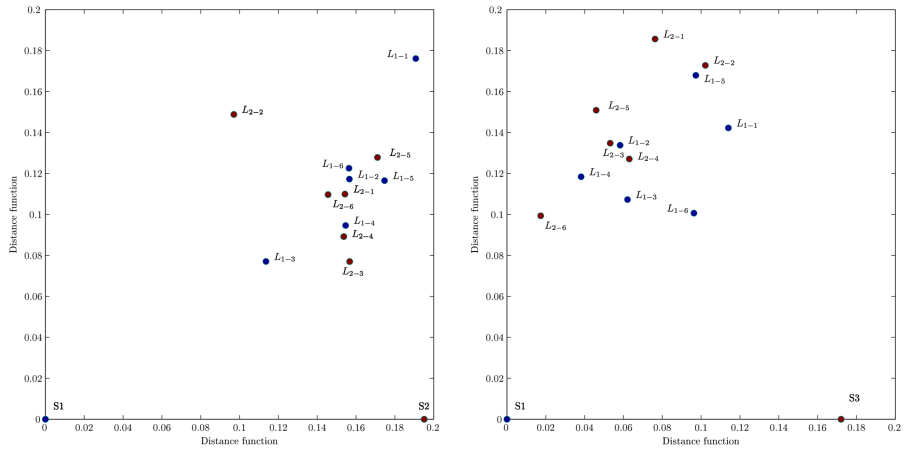


Figure 3.41: Graph edit distances between listeners and speakers. In stimulus1 (left) and stimulus11 (right), the speaker S_1 is placed in the origin and the speaker S_2 (S_3) at the corresponding distance along the x-axis while the coordinates of the listeners of both groups are shown in the distance plane. Concerning GED, both groups of listeners systematically appear closer to a "right" speaker (S_2 in stimulus1, S_1 in stimulus11) according to the listeners' subjective ratings.

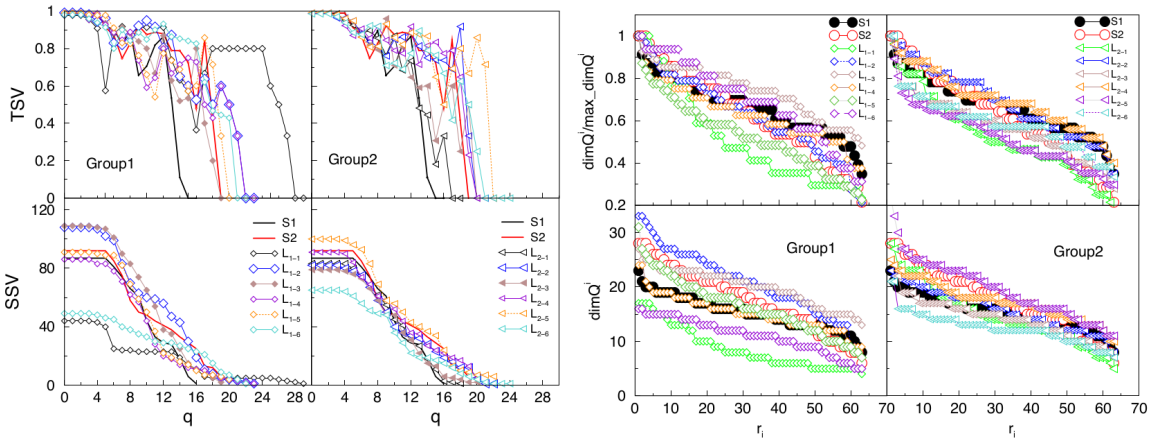


Figure 3.42: Topology measures of the SBNs of speakers and two groups of listeners for stimulus1. Components of the second (SSV) and third (TSV) structure vectors (left panels), and the ranking distributions of the nodes' topological dimensions (right panels). The full lines indicate the corresponding measures of the speakers.

Combined with Q_q, the components of the third structure vector are deduced, which quantify the connectivity among different cliques at each topology level. In this respect, both speakers, as well as the two groups of listeners, differ, as displayed in Fig 3.42. As a rule, the highest topological level in the case of listener's SBN connectivity exceeds one of the speaker's. Moreover, SBN of the speaker S_2 exhibits higher organisational complexity at all levels $q > 10$ than the speaker S_1. The results for the majority of the listeners in the group L_{2-k} are quite coherent and mainly follow the structures found in SBN of the speaker S_2. However, in the group L_{1-k} the listeners' structures exhibit larger deviations from the speakers' either at small or at large q. The striking example is the listener L_{1-1}, whose pattern of connections exhibit a much lower number of small complexes but also a certain number of vast complexes reaching at $q_{max} = 29$. This situation implies that, in contrast to all other participants in the stimulus1, in the case of the listener L_{1-1} a sizeable number of unusual connections are present, which enable the formation of cliques of the order from $24, ...30$, cf. SBNs in Fig 3.37. Note also that in the graph-distance measure, the listener L_{1-1} is far away from the both groups.

The analysis is complemented by the ranking distribution of 63 nodes in each SBN, according to the node's topological dimension, right panels in Fig 3.42. Again, the lines related to different listener's SBN suggest a higher heterogeneity of these networks for the listeners in group 1 than the group 2. These quantitative topology measures correlate well with the listeners' qualitative experience.

Apart from the diagonal blocks of the adjacency matrix in Fig 3.35, the off-diagonal matrices exhibiting the inter-brain connections provide valuable informa-

tion about the communication impact (speaker-listeners) as well as the brain-function synchronization under the same input stimulus (listener-listener correlations). These connections contribute to a nontrivial structure of the multi-brain network. In Fig 3.43, we show two representative examples visualizing differences in two-brain networks of a listener and a speaker. In one case, L_{2-3} well correlates with the speaker S_2. The corresponding two-brain network has 630 cross-links and an original structure with two communities, each of which contains the scalp locations of both individuals. This super-brain structure confirms a real focus of the listener L_{2-3} to the speaker's S_2 story, in full agreement with the corresponding distance and SBN overlap measures for L_{2-3} and S_2 discussed above. Oppositely, the two-brain network of the listener L_{1-4} and the speaker S_1 exhibits very few cross-links (57) and a community structure featuring separate brains. These results also agree well with the self-reported low concentration, uninteresting and confusing story, and bad qualities of the speaker.

Communities and topological spaces in multi-brain network

The interbrain synchronisation, which is often observed during social communications [230, 231, 237, 238], is also expected in the analysed spoken communication experiment; in the present context, it is embedded in the structure of the entire multi-brain network. Here, we perform a formal analysis of the multi-brain graph to describe the social impact on the brain activity of each participant. Moreover, we analyse the appearance of mesoscopic structures (communities) that involve scalp locations over several brains, as well as the hierarchical organisation of a particular community graph.

In general, the presence of communities is relevant for the synchronisation of stochastic processes taking part on the graph [251, 255]; the characteristic time scale of the coherence dynamics on different communities is directly related with the lowest eigenvalues of the Laplacian operator related with the graph's adjacency matrix while the corresponding eigenvectors localise on these communities [256]. Here, the activity patterns, involving different areas in the multi-brain graph, lead to the enhanced correlations and dense subgraphs or communities that involve scalp locations of several listeners and a speaker. The community structure of the multi-brain network both for simulus1 and stimulus11 are shown in Fig 3.44. In each case, there are several communities of different sizes. While some single-brain network (as the listener L_{1-1} and L_{2-2} in the case of stimulus1, and similarly, speaker S_3 in stimulus11) comprises a separate community, the majority of the identified communities are cross-brain type involving parts of the nodes in SBN of different

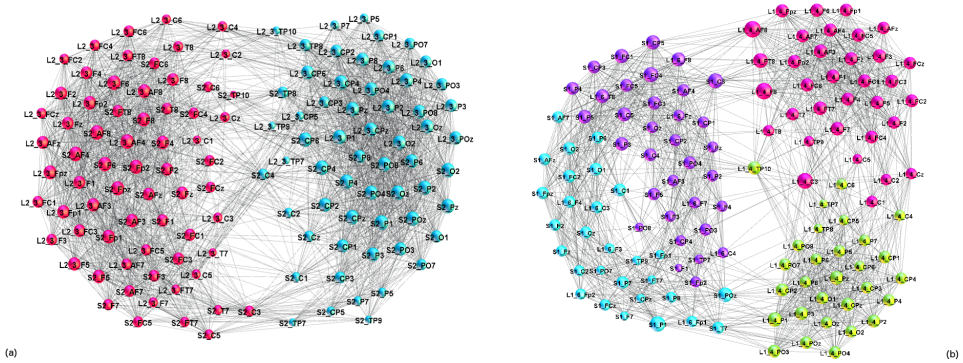


Figure 3.43: Structure of inter-brain linking. Superbrain network of the speaker S2 and the listener L_{2-3}, illustrating proper coordination, (left), and a weakly connected two-brain structure of the listener L_{1-4} and the speaker S1, corresponding to wrong coordination (right). Different colour of nodes indicates the identified functional communities. The node's labels belong to the unique list of 882 scalp locations of all participants; for example, $L_{2-3} - TP9$ and $S_2 - F7$ indicate the channel "TP9" on the scalp of the listener 3 in group 2, and channel "F7" of the scalp of the speaker S_2, respectively. See also the overlaps in Fig 3.39 and distances in Fig 3.41 for these pairs.

participants. Two such communities, related to the speaker S_2 in simulus1 are shown separately in Fig 3.45a and 3.45b. Similarly, examples of the communities related to the speaker S_1 in stimulus11 are shown in Fig 3.45c and 3.45d, respectively. It is interesting to stress that typically frontal scalp areas across different brains often form a separate community while parietal areas belong to another (here termed F- and P-community), cf. labels in Fig 3.45a-3.45d. A similar structure of the communities occurs in the two brain network in Fig 3.43a in a direct relation to a right speakerlistener coordination.

Synergy in the multi-brain communities. The results of algebraic topology analysis of the entire multi-brain network (MBN) and its largest communities are given in Fig 3.46. First, we compare the (additive) components of the first structure vectors FSV of the whole MBN with the sum of the components of all SBN. Remarkably, the MBN exhibits a more complex structure, i.e., higher values at all topology levels, which can be attributed to the contributions of inter-brain subgraphs, cf. the adjacency matrix in Fig 3.35. Hence, this feature of the MBN is a good quantifier of the social impact among the communicating brains. Similarly, the third structure TSV shows that the simplexes at all topology levels up to $q = 28$ are strongly interconnected in the MBN. In this context, the TSV of the corresponding cross-brain subgraphs suitably quantifies the speakerlistener coordination. The results of TSV for the cross-links in the two-brain network in Fig 3.43 show that the proper coordination among the listener L_{2-3} and the speaker S_2 corresponds to a topologically

rich structure; in contrast, a weaker or improper coordination between L_{1-4} and the speaker S_1 results in a much simple topology.

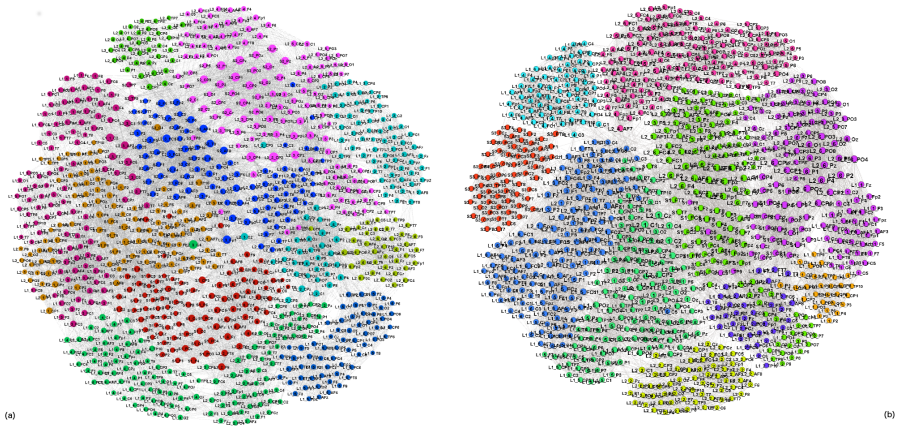


Figure 3.44: Multi-brain networks. Communities, marked by different color, of nodes in the whole multi-brain network in stimulus1 (a), and stimulus11 (b). The nodes' labels comprise the unique list of 882 scalp locations of all participants, as explained in the caption to Fig 3.43.

The two frontal- and parietal-communities from Fig 3.45 are associated with the speaker S_2 and involves several listeners. The topology analysis of these and some other communities (the SSV is shown in Fig 3.46) reveals an interesting structure. In general, in the situation of proper focus with a speaker (S_2 in simulus1, S_1 in stimulus11), a clear differentiation of F-based and P-based communities is found. Among these, P-based communities exhibit a richer structure, especially in the presence of higher order simplexes. In contrast, the situation of weak focus with the appointed speaker (S_1 in simulus1, S_3 in stimulus11), the speaker-related community involves a mixture of different SBN nodes of the speaker and the dedicated listeners. The SSV of such communities resembles an SBN the speaker's SSV. Also, the absence of a proper coordination with the speaker, several listeners appear to form a community, where also F- and P-based structures are present, but they are comparable in the topological complexity and much simpler than the speaker-based communities, cf. Fig 3.46.

Communities in multi-brain networks

By mapping a hyper-scanning dataset onto multi-brain network, we developed a systematic approach to quantify the differences in the brain activity patterns and inter-brain coordination during social communications. Our analysis of the representative spoken communications datasets (two stimuli, each consisting of the simultaneous EEG recordings of 12 listeners and two speakers narrating different

stories) confirms the leading idea of this work. Namely, the brain activation of each participant (depending on its role, communicated contents, and other factors) manifests in particular interconnections of the affected brain areas; these interconnection patterns lead to significant higher-order structures in the related brain network and interbrain graphs, which are adequately described by the algebraic topology measures. Precisely, the hierarchical structure of the scalp connectivity network and inter-brain graphs are quantified by the number of topology levels q_{max} given by order of the largest clique found in the network, and the ways that the cliques organise into larger complexes by sharing nodes at lower levels from $q = q_{max-1}, ..., 1$.

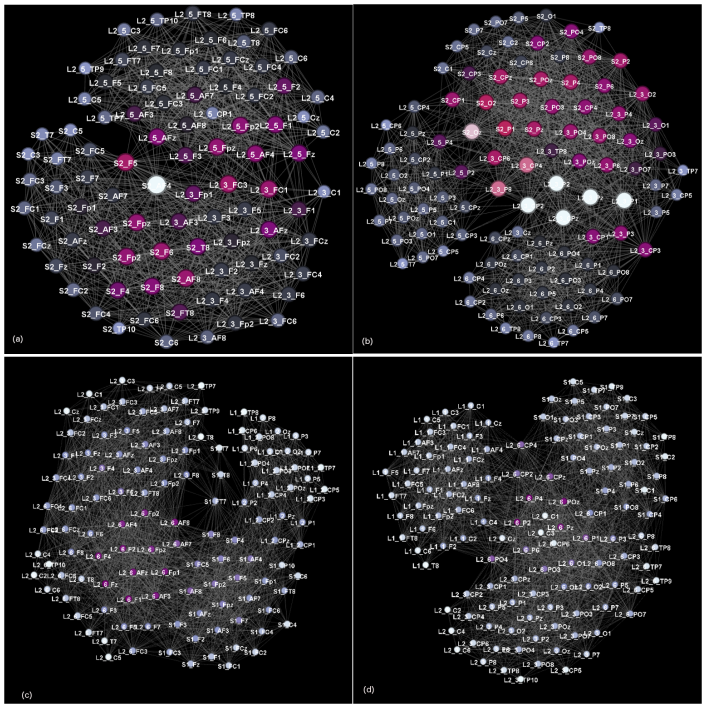


Figure 3.45: Speaker-related communities occurring in multi-brain network. Two communities dominated by frontal and parietal lobe locations are shown as separate graphs in stimulus1, (a) and (b), and in stimulus11 (c), and (d).

Assuming that the fluctuations of EEG signals on the scalp suitably reflect the underlying brain activity, the approach allows analyzing the fine-grain correlations (63 channels) of each participant, as well as cross-correlations between different brains, and the aggregate multibrain graph of two speakers and twelve listeners. The major advantage of this approach is that (even without knowing exact relationships between the correlations of the measured signals and potentially affected deeper brain areas), a comparative analysis of various networks provides a good measure of the differences in the underlying brain activity. This type of analysis combines well with the statistical features as well as with the listeners' self-rating experience during the communication.

According to these higher-order topology measures, we discovered some new features of the brain activity networks, which are not accessible to the conventional statistical methods and standard graph theory. In particular:

Significant differences between SBNs occur across the listeners and speakers, and are quantified by the topology vectors. Across the listeners' groups, the degree of heterogeneity strongly correlates with the increased distance to the appointed speaker. We also find $q^L_{max} > q^S_{max}$ in all studied examples, suggesting that the listener's brain activity results in a more complex architecture than the speaker's. In agreement with the statistical analysis in [240], this fact relates to the processing of semantic content in the presence of noise. A more detailed analysis reveals the excess links in the listener's SBN; these links correspond to the coherence between a set of different EEG channels, not occurring in the speaker's network. Moreover, these topology quantifiers accurately distinguish the patterns of the brain activity of the same speaker while narrating different subjects. Fig 3.47 displays the differences between corresponding EEG correlation networks as graphs and at each topology level for the speaker S1. Notably in contrast to the stimulus1, the number of big organised structures (for $16 < q < 24$) occur in the case of the speaker's narration in stimulus11, which also obtained higher ratings by the listeners.

The proper speakerlistener coordination is suitably quantified by the topological similarity of their SBNs and a rich structure of the corresponding two-brain network. At lower topology levels $1 < q < q^S_{max}$, the majority of listeners in both groups exhibit the brain activity patterns that are more similar, i.e., have better coordination, with one speaker than with the other. These topology findings are also supported by the statistics of the link overlaps and brain-to-brain distance measures, which consider $q = 1$ level, i.e., the presence and absence of each particular link in the compared networks. Interestingly enough, these findings compare well with the listener's self-rating of the sympathy to the speaker, the speaker's narrative quality and attractiveness, as well as the clarity and the interest of the story. It turns that a complex two-brain network suitably represents the case of a proper coordination. The strong frontalfrontal and parietalparietal connections between the brains appear as two communities in a super-brain structure, cf. Fig 3.47a. These features are absent in the case of weak coordination, as the example in Fig 3.43b. Further analysis of the inter-brain correlations concerning the issues of the experimental design, questionnaire, and the semantic contents is left for the future work. The coordination with a right speaker evolves over time, as measured by the distance between speakerlistener SBN constructed in a sequence of time intervals, cf. Fig 3.48. However, there is always a gap (minimal distance) between a speaker and anyone of the

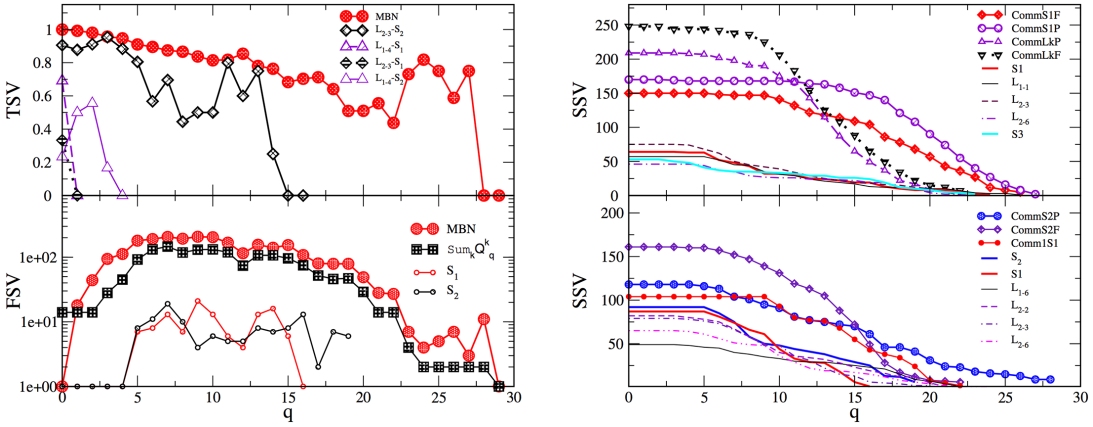


Figure 3.46: Topology vectors of multi-brain graphs. Left panels: Components of the first (FSV) and the third (TSV) structure vectors plotted against the topology level q for the whole multi-brain network and for some its subgraphs, as indicated in the corresponding legends. The additive components of the FSV allow a comparison of the whole MBN with the sum of the corresponding component of each participating SBN, the line is indicated by $\sum_k Q_q^k$, where k runs over all listeners and the two speakers in stimulus1. TSV of cross-graphs in two-brain networks from Fig 3.43 and their counterparts are shown. Right panels: Components of the second (SSV) structure vector of the largest four communities in stimulus11 (top) and three communities in stimulus1 (bottom). For comparison, the values obtained for the corresponding SBN of the speakers and listeners participating in these communities are also shown.

listeners, in agreement to the occurrence of higher structures and extra links in the listener's activity networks, mentioned above. It is also interesting to note that the listener's networks exhibit a high degree of similarity, perhaps suggesting similar initial brain activity patterns, before focusing to a particular speaker.

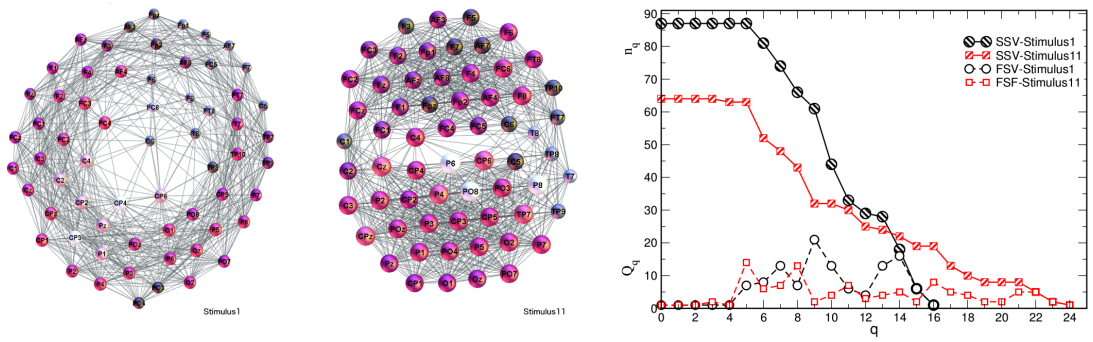


Figure 3.47: Subject-specific brain activity patterns of the speaker S_1. From left to right, SBNs represent EEG correlation patterns of the speaker S_1 narrating a fairy tale (in stimulus1) and giving instructions (in stimulus11), and the components of the first and the second topology vector of these SBNs.

F/P communities in the multi-brain networks reveal super-brain features. At the

aggregate graph level, the communities occur primarily related to a good narrator. These communities, as a rule, coordinate frontal brain areas of the speaker with the listener's frontal areas, and, similarly, parietal-to-parietal. These communities regarded as separate graphs exhibit a rich structure with a large number of high-order cliques and their complexes. It is interesting to note that in each identified community structure, the parietal-based community appears to be more complex than the frontal-based one. The occurrence of speaker-related communities suggests that a significant number of channels, which are coordinated in the speaker's brain also appear to be coordinated in the listener's brains. In contrast, when the appointed speaker is not followed (the case of speaker S1 in stimulus1 and speaker S_3 in stimulus11), some listeners appear to form communities, again connecting through frontal (parietal) channels. However, the corresponding topology of the listener's communities is quite simpler than the topology of the speaker-related communities. In this case, the listeners exposed to the same external input synchronise their activity patterns, without having any direct communication. The observed community structure is consistent when the stimuli and the threshold values are varied. Note that the presence of community structure was confirmed as an essential feature of this type of brain networks in [251] using the local algorithm.

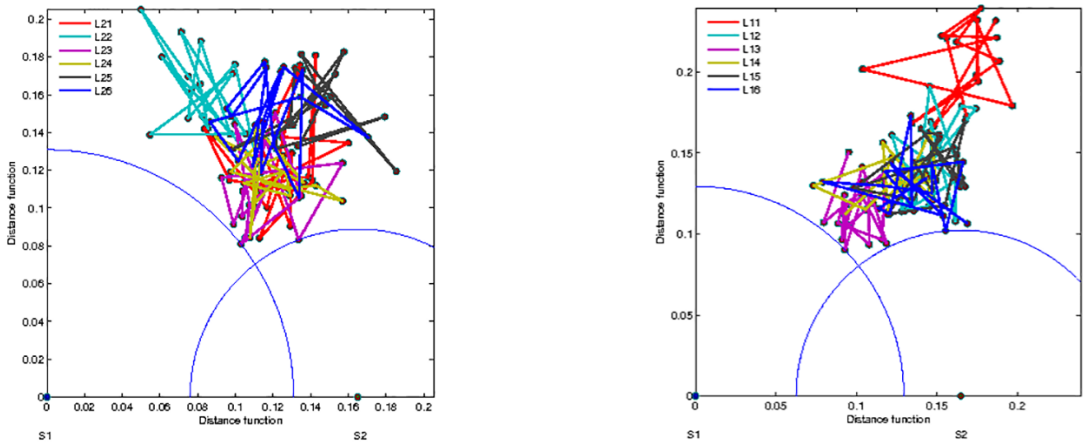


Figure 3.48: Evolution of the brain-to-brain distance. The timeline (16 frames) of the GED between brain activity networks of the listeners in group 1 (left) and in group 2 (right) from both speakers S_1 and S_2 is shown for the stimulus1. Each circle indicates the minimum distance from the corresponding speaker that occurred during the process.

Quantitatively, the aggregate graph structure shows new topology features as compared to the sum of the corresponding individual measures, i.e., in the case of the additive FSV. In this way, the new structure emanating from the cross-brain connections appropriately describes the social impact onto the individual brain

activity that can be captured by EEG signals.

3.5.4 Conclusions

We have considered the social brain structure in two concrete examples of simultaneous EEG recordings during spoken communications (of 12 listeners and two speakers narrating different subjects) by mapping the data onto the multi-brain network and applying the methods of algebraic topology of graphs. We have shown that the topology of higher-order complexes precisely quantifies the differences in the brain activation pattern between the participants during the social communication. Furthermore, the topology provides the accurate measure for the speakerlistener coordination and the speaker's impact onto a group of listeners. Our results also suggest that the mechanisms for super-brain functioning during spoken communications certainly involve strong frontal-to-frontal and parietal-to-parietal synchronization in dyads. In a more general context, the study of higher-order combinatorial structures by algebraic topology techniques provides a sensitive methodology to quantify the shifts in the functional brain networks, e.g., under changed activity or condition. By complementing the standard graph theory methods, the algebraic topology can contribute to a more in-depth analysis of other brain imaging data.

Chapter 4

Topological aspects of nonlinear dynamical systems

In the research of dynamical systems, physicists observe time evolution of a real world phenomenon as it evolve in time, where are observations are expressed like numbers recorded as they change over time. For example, let's consider the motion of the planets across the celestial firmament. During the daily motion of stars from East to West, planets distinguish themselves by moving among the fixed stars. Ancients discovered that by knowing a sequence of planets positions, latitudes and longitudes, its future position can be predicted. By tracking these coordinates on the celestial sphere, it is possible to calculate all possible values for positions and velocities of the planets and we form the *phase space* of the system [262][259]. Hence, in general, a fixed point in phase space, the state of physical system at certain point in time can be represented by a single point in an abstract space known as the *state space M*. As the evolution of system occurs, the system changes and so does representative point in state space and we assign *dynamics* to the evolution of such point. The phase space *trajectory* represents a set of states located within the total phase space, and represents a sequence of states of dynamical system, starting from any initial condition. *Evolution rule* is a function f^t which returns the value where the representative point is at time t. Then the pair (M, f) characterize *dynamical system*. Therefore, the term "dynamical system" refers to any physical or abstract singleton whose properties (from M) can be represented by the evolution rule given on set of numbers, *system variables*, at any given time, and whose properties in posterior time are uniquely determined by its present and past properties through f.

There are two types of transformation rules depending on "time", continuous-time systems and discrete-time systems [262][257]. For purpose of this thesis, focus

will be on discrete-time dynamical systems. For example, in discrete-time systems the evolution rules are expressed as equations:

$$\mathbf{x}(t+1) = \mathbf{F}(\mathbf{x}(t); \mathbf{p}, t),$$

where the vector $\mathbf{x}$ represents system variables, whereas the vector $\mathbf{F}$ represents a function of all system variables for fixed values of parameters $\mathbf{p}$. In discrete systems function $\mathbf{F}$ directly gives $\mathbf{x}$ at the next time step, and the function $\mathbf{F}$ is also known as a *map* that drive system from one time step to the next.

4.1 Nonlinear dynamical systems

A characteristic feature of nonlinear dynamical systems is that the equations of motion contain at least one nonlinear term, such as a square or higher power of the variable whose temporal or spatial, or spatiotemporal, evolution is followed, or some threshold function. Nonlinear systems are also characterized by multiple *attractors* [260][262], which are sets of numerical values toward which a system tends to evolve for a broad collection of initial conditions of the system. Nevertheless, posterior system states may be highly sensitive to its initial state and it gives a whole new set of phenomena that are associated with the way in which the pool of attraction shifts as parameters are changed. Nonlinearity may also affect the trajectory causing it to become irregular, eventually lying on the so called *strange attractor*, and then the trajectory is constrained to a region of state space where there are no fixed points or no stable limit cycles. When this occurs, time evolution of two states of the system that have close initial states rapidly diverge in later time. This phenomenon is known as the deterministic *chaos*. The presence and intensity of chaos can be measured via *Lyapunov exponents* λ_i [262], and the system is in the chaotic state if one of the exponents is positive. Stated differently, two trajectories of variable x with near initial states x_0 and x'_0 becomes widely separated in posterior time t, $\delta x = \delta x_0 e^{\lambda t}$, where $\delta x = |x - x'|$ so that $\lambda > 0$ indicates that the system is chaotic.

Classification of dynamical system based on the degree of randomness is as follows [262] [261][260]:

- no chaos
  - integrable systems
- deterministic chaos
  - ergodic systems

  - mixing systems
  - K-systems
  - C-systems
- chaos (not deterministic)
  - stochastic systems

An *ergodic* dynamical system is the one that has the same behavior averaged over time as averaged over space of all the system's states in its phase space. For continuous system, ergodicity implies

$$\lim_{x\to\infty}\frac{1}{T}\int_0^T f(\bar{x}(t))dt = \langle f(\bar{x})\rangle\,,$$

where function f is any smooth function of the phase space variable $\bar{x}$,$\bar{x}(t)$ is a trajectory in phase space and $\langle f(\bar{x})\rangle$ represents the average of $f(\bar{x})$ over the space. The last equation holds for almost every initial state considering the invariant measure. Ergodicity represents the weakest form of randomness and chaos does not necessarily appear in these systems.

A map is *mixing* if for any two measurable sets A and B the measure of the intersection of $T^{-n}A$ with B converges to the product $\mu(A)\mu(B)$, where T is measurable self transformation of a set and μ is probability measure. Roughly this means that after a long time every set A will spread evenly over the entire space (its contribution in every set B will be nearly proportional to the size of B). Every mixing system is ergodic. The notions of ergodicity and mixing (and many others) mimic the behavior first observed in physical systems.

We say that the system is the K-system if there occurs invariant sets with positive metric entropy. C-system is the one which is both chaotic and hyperbolic at every point in phase space.

4.1.1 Recurrence

In the terminology of dynamical systems *recurrence* [21] indicates a time after which the trajectory returns to very close to a location where it has been before (but with constraints to a dynamical systems with finite volume). The recurrence of states in nature has been known for a long time and discussed by many physicists, but the most crucial theorem is given by Henri Poincaré. Poincaré recurrence theorem [268] says that system will, after sufficiently long and finite time, return very close to the initial state(time varies depending on the degree of closeness and the exact initial

state). The Poincaré recurrence time τ is the time it takes for the trajectory to return close to the phase space point which was visited at some previous occasion.

For the study of recurrence the Takens Embedding Theorem [267] is also of great practical importance. The theorem shows how a time series of measurements of a single observable can be efficiently used to reconstruct qualitative features of the phase space of the system under study. The method relies on the use time delays of the time series entries in order to reconstruct the trajectory of the system in phase space. Suppose that a measurement of the system generates a time series $y(t_1), y(t_2), ...y(t_N)$ which lies on a d-dimensional attractor of an n-th order dynamical system. The initial stage of the application of the theorem is to obtain an *embedding* from the time series. This is achieved by using time delay coordinates for which a delay vector has the following form:

$$\boldsymbol{y(k)} = [y(k), y(k-\tau), ...y(k-(d_E-1)\tau]^T,$$

where d_E is the embedding dimension, τ is the delay time and k is an integer. Takens has shown that embedding with $d\, 2n+1$ is sufficient generically, so that there is a smooth map f:$\mathbb{R}^d \rightarrow \mathbb{R}$, such that $\boldsymbol{y(k+1) = f(y(k))}$**.** The practical aspect of the theorem relates to the estimation of the embedding dimension and the delay time, which will be explained in the further text.

4.2 Topological view on Recurrence

As a continuation of the analysis of nonlinear dynamical systems from the point of view of computational topology [272][273], the focus is shifted on the recurrency as one of the characteristic properties of dynamic systems. Many natural processes exhibit either regular cyclic behavior, e.g. periodic behavior or behavior characterized by irregular cycles. The distinctive feature of deterministic dynamical systems, and in particular of nonlinear and chaotic systems, is the recurrence of states in phase space in the sense that states become arbitrarily close after certain amount of time. Phase space is often high dimensional and the only way to visualize it is by projecting it on a two or three dimensional sub-spaces. In order to visualize the recurrence of states $\overrightarrow{x}_i$ in a high dimensional phase space a tool known as the *recurrent plots* was introduced [266]. Here the state of the system is a phase space trajectory usually reconstructed using time series $u_i(t)$ of only one dynamical variable [263], [264]

$$\overrightarrow{x}_i = (u_i, u_{i+\tau}, ..., u_{i+(m-1)\tau})^T,$$

where m is the embedding dimension and τ is the delay time ($\tau = \tau\ \Delta t$). After Takens theorem the topological structure of the original trajectory is preserved provided that $m \geq 2d + 1$, where d is the dimension of the attractor. This approach provides a means to study trajectories in a phase space, say with dimension m, by means of a two-dimensional representation of the recurrent states. The recurrence of a state i at certain subsequent time j is represented in the two-dimensional recurrent plot as a square matrix with, in the original formulation, black and white dots, with black dots indicating recurrence and with both axis representing time. Mathematical expression of recurrence of the state $\overrightarrow{x}_i$ with respect to the state $\overrightarrow{x}_j$ is:

$$R_{ij} = \Theta(\epsilon - \|y_i - y_j\|),$$

where $\Theta(\cdot)$ is the Heaviside function, ϵ is the threshold distance and R_{ij} is the *recurrence plot*. So if $y_i \approx y_j$, $R_{ij} \equiv 1$ and we call this the *recurrence point*.

Choice of parameters τ and m is based on methods originating from phase space reconstruction like the method of false nearest neighbors and mutual information. These methods ensure determination of the whole covering of all free parameters, and the avoidance of autocorrelated effects [265].

The presence of threshold distance ϵ in the above expression implies that the recurrence is defined as a state $\overrightarrow{x}_j$ which is sufficiently close to the initial state $\overrightarrow{x}_i$ since the state of the nonlinear dynamic system rarely recurs at the exact initial state but approaches it arbitrarily close. The state $\overrightarrow{x}_j$ that belong to a neighborhood of $\overrightarrow{x}_i$ of radius ϵ_i which may be, for example, in the form of a m-dimensional sphere for the L_2-norm or a box for the L_∞ norm, are called recurrent points. The properties of the dynamical system are inferred from the recurrent plot by considering all recurrent points in their totality and not from a single recurrent point which, by its definition, does not involve any information about the current states at times i and j. Recurrence plots of dynamical systems exhibit specific structural properties in the form of large scale and small scale patterns [21]. Careful inspection of the expression 4.2 suggests that the recurrence matrix $\boldsymbol{R}$ may be interpreted in several different ways. First, the recurrence matrix $\boldsymbol{R}$ may be viewed as the adjacency matrix A of an unweighted complex network. Second, it may be interpreted as an incidence matrix M of a neighborhood (simplicial) complex associated with the dynamics of the system.

The choice of the threshold distance requires special consideration, since ideally it should be as small as possible, however in practice one has to put in use a criterion which would reflect the dynamics of the system as well as the effects of noise.

Choosing the right parameter ϵ is crucial moment and special attention is required for its choice. If ϵ is chosen to small, almost no recurrence point will be shown in recurrence plot and we will not be able to learn anything about the structure of underlying system. On the other hand, if ϵ is too large, then the large number of recurrence points will be in ϵ neighborhood including consecutive points on the trajectory. In the course of solving these issues, there are several criteria given for different cases of dynamical systems [269, 270, 271].

Considering the homology of recurrence plot, it is shown that there is no need that embedding dimension is $m > 2n$ [273]. Namely, the persistence of the rank of homology groups occurs on dimensions lower than $2n$, and the embedded phase space is homology invariant to underlying dynamical system.

In order to make the dimension of homology groups computable in case of a large data set, we will also use homology invariant *witness complex* [272], more precisely fuzzy witness complex. We determine two data sets $W \subset \mathbb{R}^m$, formed by *witnesses*, and associated set $L \subset \mathbb{R}^m$, formed by *landmarks*, which may overlap with the set W. Landmarks represent vertices of the complex and connection between them depends on geometrical relationships between elements of $w \in W$ and $l \in L$. So, the point w is a witness of a q-dimensional simplex $\sigma = \{l_1, l_2, ..., l_{q+1}\}$ if $w \in r$-neighborhood of every $l_i \in \sigma$. Simpler implementation of this type is *fuzzy* or *lazy* witness complex. The fuzzy witness set of point $l \in L$ is a set of witnesses

$$W_\omega(l) = \{w \in W : \|w - l\| \leq \underset{l' \in L}{min}(\|w - l'\| + \omega)\}$$

A simplex σ_i in a neighborhood complex represent the density of points representing states within "epsilon" environment of a state i in phase space of the underlying dynamical system. Connectivity chain in neighborhood complex describes density of states of the phase space at trajectory's vicinity in the underlying dynamical system. Hence, the obtained witness complex represents the topological coarse graining of the neighborhood complex, since structural features are preserved.

Chapter 5

Application and results 2

Simulations for Lorenz and Rössler dynamical systems as well as for stochastic system, are performed for 7000-8000 time steps for different sets of parameters of these dynamical systems. In order to compare the properties of recurrent simplicial complexes of these systems computed for the same ϵ-neighborhood, it is necessary to normalize the amplitudes of time series to the same scale $[0, 1]$. Then, the *mutual information* method is used for finding τ, and relying on the calculations of the [273] *false nearest neighbors* method, we set the embedding dimensions to $m = \{3, 4\}$. When systems become stabilized, 2200-2300 points of the underlying time series were taken for further reconstruction of the phase space features. Results for the Lorenz and Rössler attractors and for the stochastic process will be presented in the subsequent Sections. Since there is a large fluctuation in amplitudes of these signals in order to compare the same results relating to the same ϵ, the normalization is performed thus scaling the amplitudes scaled to $\{0, 1\}$. For the Lorenz and Rössler systems only variable was used. We have chosen to represent the results pertaining to the x-variable while the results are similar for the other two variables, y and z, when we keep all other parameters the same.

5.1 Lorenz attractor

The Lorenz system, originally developed by Edward Lorenz, is a system of ordinary differential equations which represent a simplified model of atmospheric convection. It is notable for having a chaotic solutions for certain parameter values and initial conditions. The system of three ordinary differential equations of the model is:

$$\frac{dx}{dt} = \sigma(y - x),$$

$$\frac{dy}{dt} = x(\rho - z) - y,$$

$$\frac{dx}{dt} = xy - \beta z,$$

where $\rho, \sigma, \beta \in \mathbb{R}$ are the parameters of Lorenz system which is invariant under the transformation $(x, y, z) \to (-x, -y, z)$. Lorenz system has sensitive dependance on initial conditions for the set of parameter $\sigma = 10, \beta = 8/3$, and $\rho = 28$, forming the well-known butterfly-shaped Lorenz attractor, which is probably the most recognizable example of chaotic attractor. In order to study the recurrent dynamics only the parameter ρ is changed while parameters $\sigma = 10$ and $\beta = 8/3$ are kept fixed. Variation parameter ρ is taking value from the set $\{10, 13.93, 16, 24.05, 24.73, 28\}$. Three values of considerable importance for further analysis are those corresponding to the chaotic regime of the Lorenz system, i.e. for $\rho \geq 24.05$:

1. Heteroclinic bifurcation occurs at $\rho = 24.05$,
2. Hopf bifurcation at $\rho = 24.73$ and
3. butterfly at $\rho = 28$.

The analysis starts with simulation of Lorenz system for various ρ.

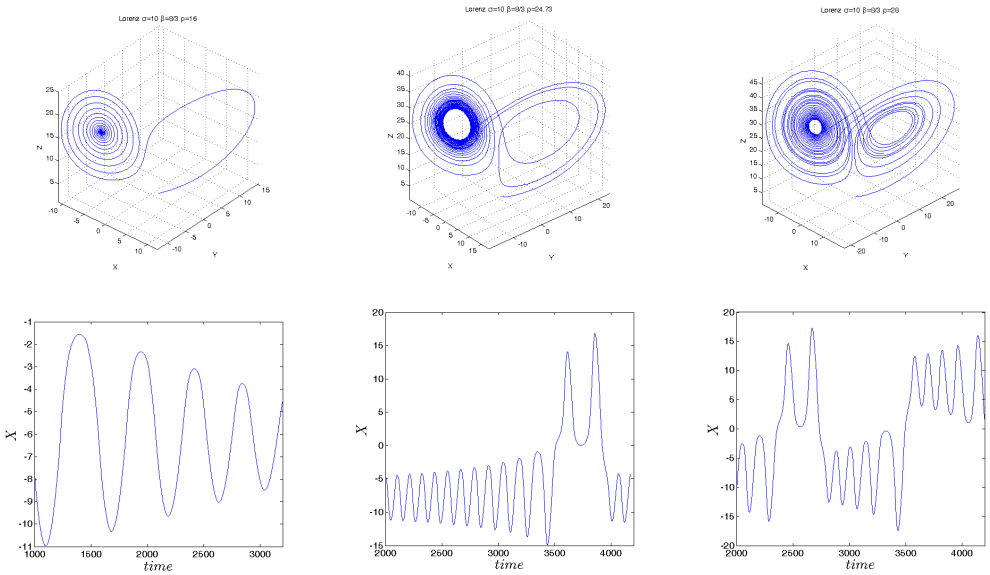


Figure 5.1: From the left to right: Lorenz system for parameter $\rho = \{16, 24.73, 28\}$. Top three figures represents Lorenz attractor for different ρ , bottom three figures represent time series of X axis respectively to the top figures

Using method presented in the previous section, we reconstruct the normalized time series forming two-dimensional adjacency matrix, i.e. the recurrence plot.

Using reconstruction parameters $m = 3$, $\epsilon = 0.1$, $L = 200$, the Lorenz attractor 5.1 is reconstructed in adjacency matrices presented in 5.2.

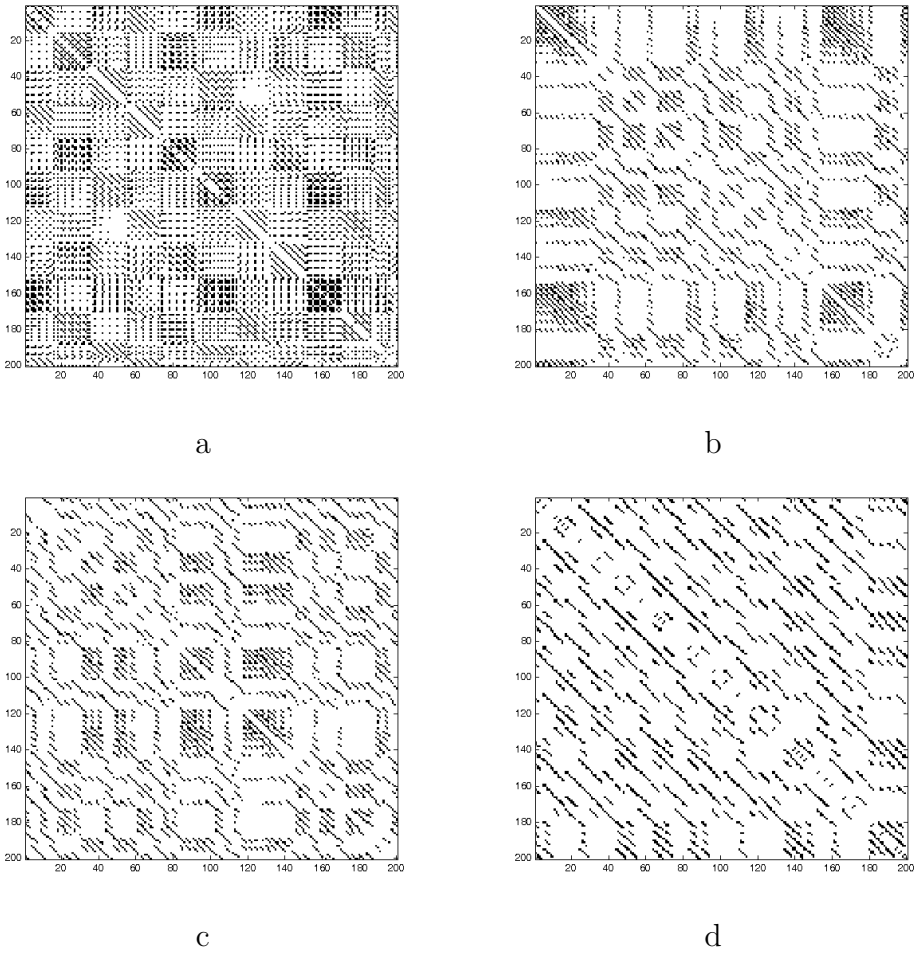
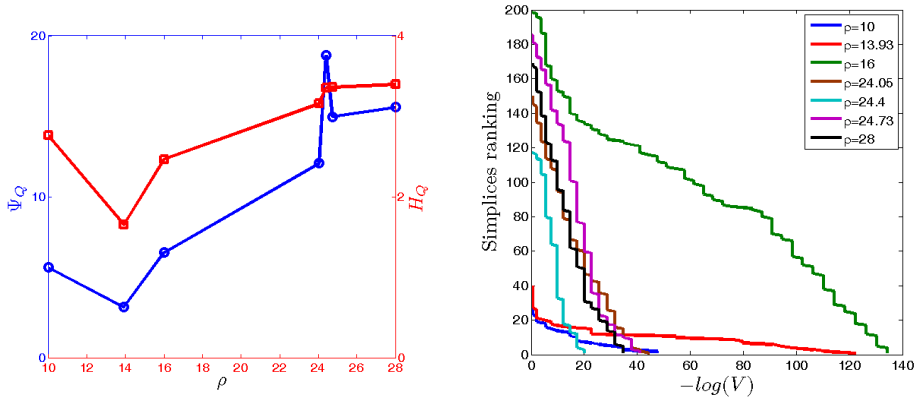


Figure 5.2: Recurrence plot of witness complex (200 landmarks) of the Lorenz attractor for $\rho = \{16, 24.05, 24.73, 28\}$ respectively

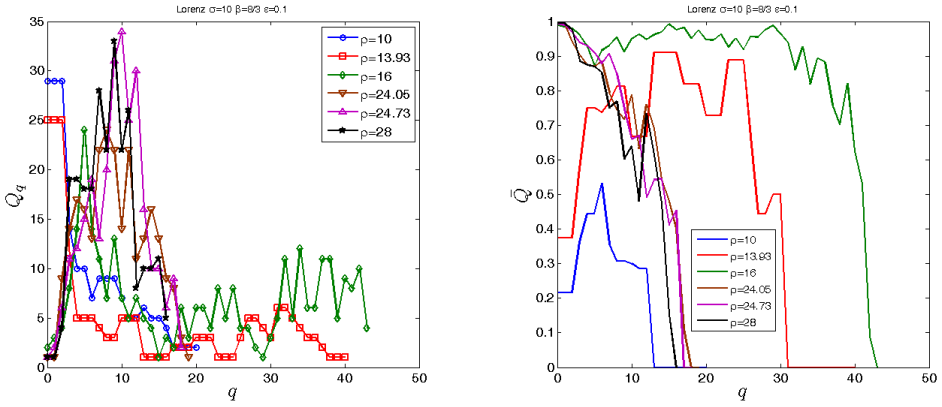
Simplicial complex analysis provide an information about the topological structures of obtained witness complexes, in the form of structure vectors 5.4. When approaching chaotic behavior, at bifurcation points above $\rho = 24.05$, the structure of simplicial complex displays different organization with respect to the dynamics near other bifurcation points, such as homoclinic bifurcation ($\rho = 13.93$) or where perturbation occurs ($\rho = 16$) [274]. After transition to chaos happens at heteroclinic bifurcation, all structures of a given simplicial complexes follow the same behavior. They create more complicated complexes according to formula $\Psi_Q = \frac{2}{(D+1)(D+2)} \sum_{i=0}^{D} (i+1)Q_i$ presented at Fig 5.3a, where Ψ_Q of simplicial complex Σ is representing simplicial complexity [275], $D = dim(\Sigma)$ is dimension of simplicial complex Σ. Simplicial complexes describing system in chaotic regime are creating

large number of clusters in terms of connectivity components in structure vectors compared to complexes that are describing pre-chaotic systems. Since Q-vector in chaotic regimes has higher component values than in pre-chaotic regime, phase space is less dense and more replete in the chaotic regime.



a b

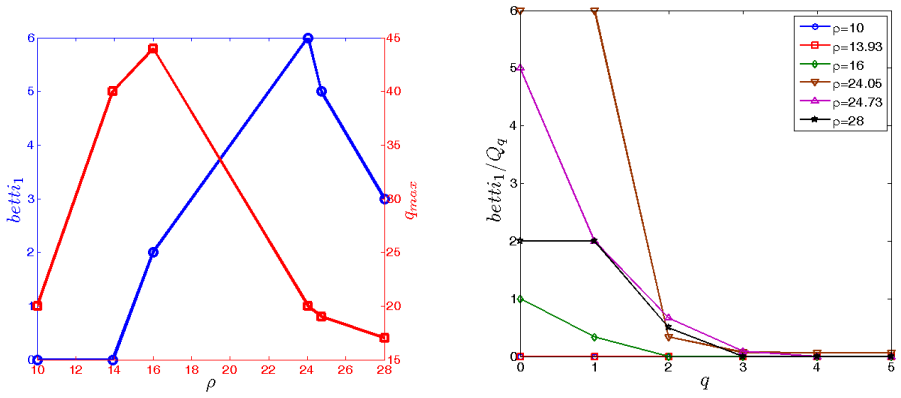
Figure 5.3: (a) The complexity Ψ_Q and the entropy H_Q and (b) Simplices volume distribution of witness complex of the Lorenz attractor



a b

Figure 5.4: Structure vector of witness complex of the Lorenz attractor for various ρ.

From results presented in Fig 5.5, we can observe that all simplicial complexes constructed on the butterfly attractor preserve two 1-dimensional holes. Simplicial complexes reconstructed from attractors at the heteroclinic point and the Hopf bifurcation point have almost the same topological structure as simplicial complex obtained from butterfly attractor, nevertheless the dimension of homology is signif-



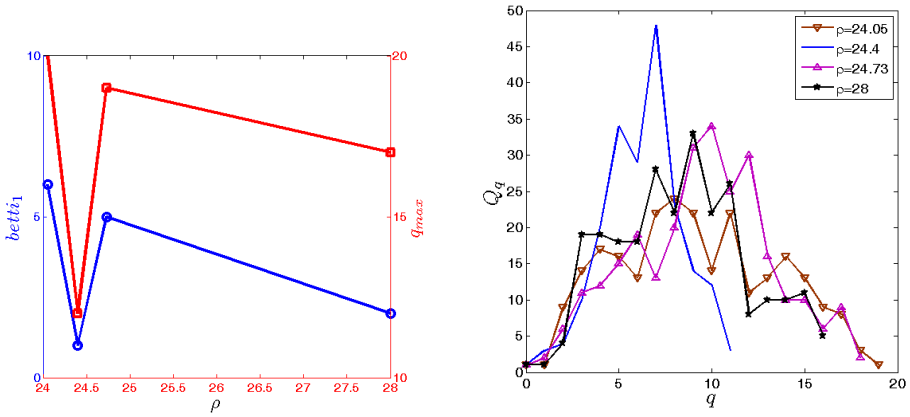
a b

Figure 5.5: (a) $Betti_1$ and the dimension of simplicial complexes for various ρ, and (b) $Betti_1$ number per connectivity classes $\frac{Betti_1}{Q_q}$ for different q dimensions of witness complex of the Lorenz attractor

icantly higher compared to butterfly attractor and decay toward lower values of ρ. Hence, when comparing both complexity and homology of these regimes, we can conclude that more complex simplicial complexes are formed near and in the regime of chaotic dynamics. Though this conclusion is rather expected, its significance lies in the inherent topological origins of such behavior, not seen previously by common methods. Preservation of holes within connectivity class at higher topological dimension does not occur. Additional holes in the system at bifurcation point disappear at $q = 2$ and the system become more similar to the chaotic one. These results are suggesting a way to differentiate various regimes in unknown dynamical system. At $q = 2$ all chaotic systems follow the same behavior and the same happens for pre-chaotic regimes.

We also explored differences in the structure, the simplicial complexity and homology when system is at the bifurcation point in transition to chaos, and when system is close to the bifurcation point that is also in chaotic regime as presented in Fig. 5.6. From the topological structure of complexes we found that they still follow the same behavior, the complexity is similar, but the main difference is in the dimension of homology group at dimension 1. From Fig 5.6a we can see that the homology group dimension at the specified simplicial complex dimension is higher for states of bifurcation compared to chaotic regimes near bifurcation point, and hence we are able to differentiate whether the system is at the bifurcation point or not, when it is already in the chaotic regime.

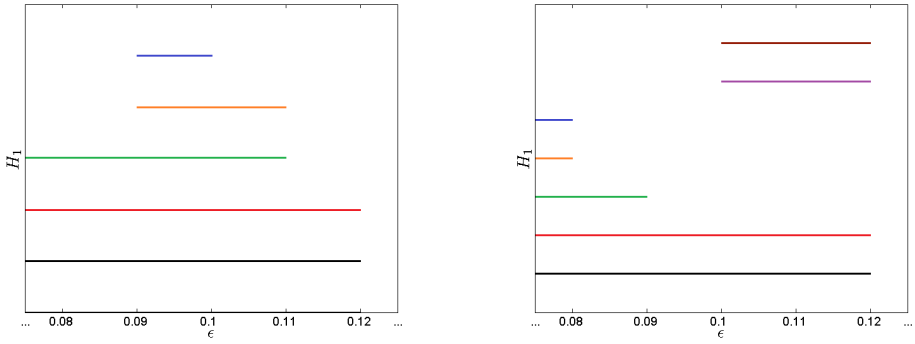
In figure 5.7 is presented different, but more common, graphical display of di-



a b

Figure 5.6: (a) $Betti_1$ and dimension of simplicial complexes for ρ above heteroclinic bifurcation and (b) the first structure vector for ρ above heteroclinic bifurcation for the witness complex of the Lorenz attractor

mension of the first homology group.

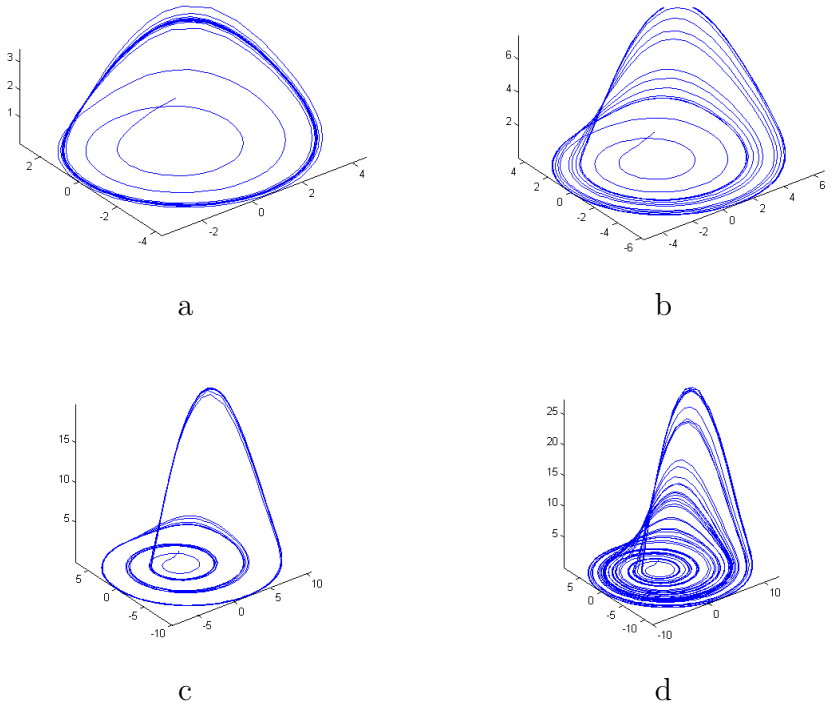


a b

Figure 5.7: $Betti_1$ bar-code of witness complex of the Lorenz attractor

Simplicial volume $V = \frac{\sqrt{n+1}}{n!\sqrt{2^n}}$ represents subvolume of phase space where simplices lives and its relative value can serve as the probability of finding the dynamical system in that part of initial phase space. Information-like entropy evaluated of the first structure vector H_Q is defined as $H_Q = -\sum_{i=0}^{i=q_{max}} Q_i log Q_i$. We can observe from Fig 5.3a that complexity ψ_Q has qualitatively similar behavior as H_Q. Simplices volume distribution is shown at Fig 5.3b. Following the same behavior as in previously analyzed properties, chaotic and nearly chaotic system showing us similar distribution of simplicial volume, but heteroclinic bifurcation shows interesting distribution that diverges from all other regimes.

5.2 Rössler attractor



a b

c d

Figure 5.8: Attractors of Rössler systems for $a = b = 0.2$ and for $c = \{2.3, 3.3, 5.3, 6.3\}$ respectively

The Rössler system is defined by three nonlinear differential equations, originally studied by Otto Rössler:

$$\frac{dx}{dt} = -y - z,$$
$$\frac{dy}{dt} = x + ay,$$
$$\frac{dx}{dt} = b + z(x - c).$$

where a, b, c are parameters, and x, y, z are three variables of the system. For certain values of parameters, this system is noticeable for the specific shape of the Rössler attractor. In this study we fixed parameters $a = b = 0.2$ and vary parameter c. For $c = 2.3$ 5.8 (a) we get one limit cycle attractor, for $c = 3.3$ 5.8 (b) two limit cycle attractor for $c = 5.3$ 5.8 (c) and three limit cycle chaotic attractor for $c = 6.3$ 5.8 (d). Using method presented in the previous section, we reconstructed the normalized time series into two-dimensional adjacency matrix, i.e. the recurrence plot. Using reconstruction parameters $m = 3$, $\epsilon = 0.1$, $L = 200$, the Rössler

attractor 5.8 is presented as a recurrence plot in 5.9.

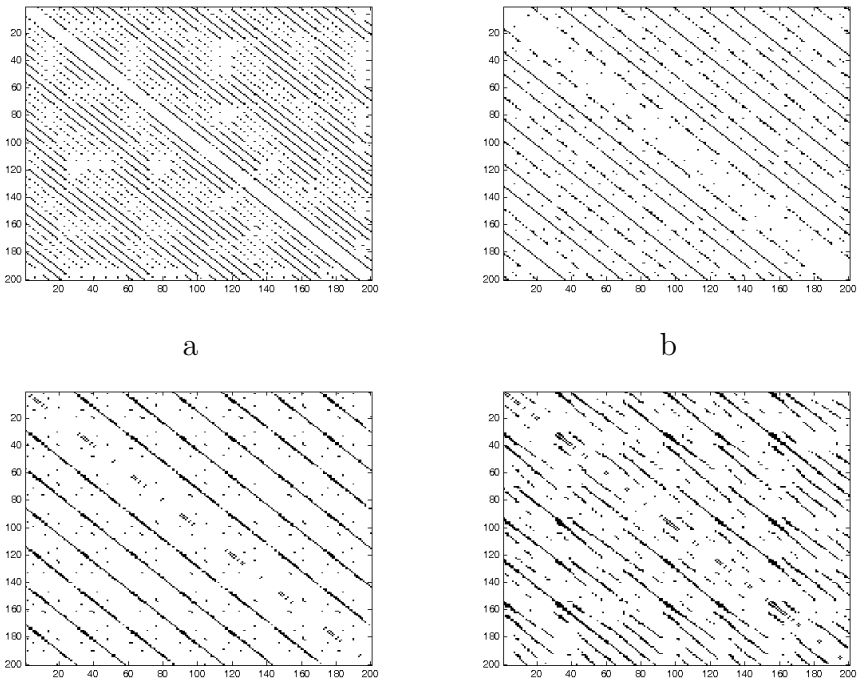
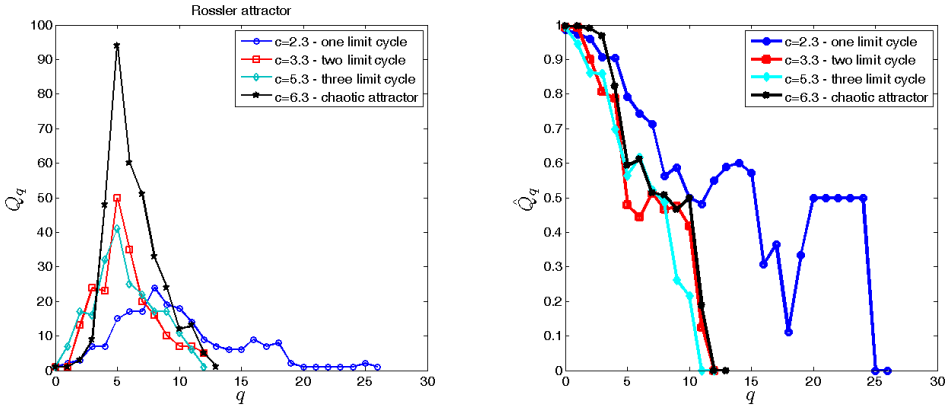


Figure 5.9: Recurrence plot of witness complex (200 landmarks) of the Rössler attractor for $a = b = 0.2$ and different values of $c = \{2.3, 3.3, 5.3, 6.3\}$

From the aspect of topological structure a similar behavior is found in the Rössler attractor as in the case of the Lorenz attractor. The homology of obtained simplicial complexes and its complexity are presented in Figs 5.10 and 5.11. Also structures increase in complexity, from the topological point of view, with the increasing chaos, and the dimension of homology group is preserved longer in connectivity classes than in regimes before the chaotic regime occurs. The same qualitative behavior is seen when analyzing the simplices volume distribution, presented in Fig. 5.11 (c), where it is especially interesting to note that the same features are for the one limit cycle attractor as for the heteroclinic bifurcation in the Lorenz system - far before the chaos occurs.

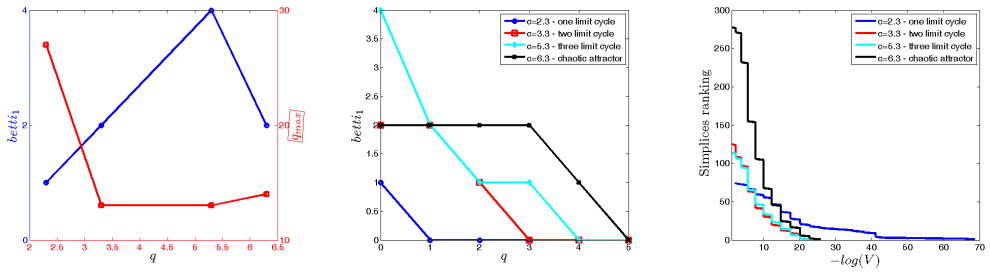
5.3 Stochastic signal versus deterministic chaos

In order to compare topological features between deterministic chaos and stochastic process, we decide to take the time series of a stochastic signal, presented in Fig. 5.12(a), which represents the velocity fluctuations as a function of time in fully developed turbulence [276].



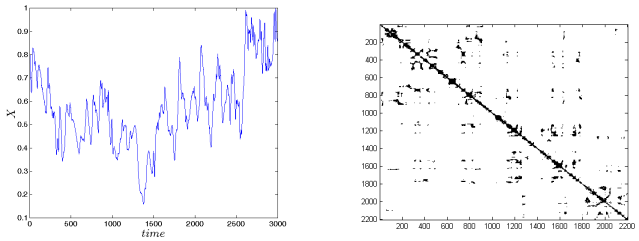
a b

Figure 5.10: First (a) and third (b) structure vectors of witness complex of the Rössler attractor



a b c

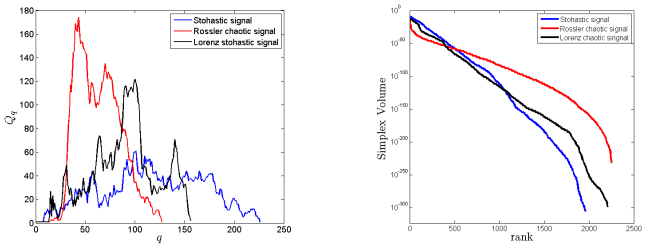
Figure 5.11: (a) $Betti_1$ and dimension of simplicial complexes for various c , (b) $Betti_1$ number for different q dimension, and (c) simplices volume distribution, for witness complex of the Rössler attractor



a b

Figure 5.12: (a) Time series of stochastic signal and (b) recurrence plot of witness complex (200 landmarks) obtained from time series of stochastic signal

Comparing first structure vectors, presented in Fig. 5.13(a), we concluded that topology of simplicial complex obtained from stochastic signal has lower complexity than in Lorenz and Rössler attractors in chaotic regimes. From our experience, this



a b

Figure 5.13: (a) First structure vectors and (b) volume distribution of recurrent simplicial complexes of stochastic signal and Lorenz and Rössler attractors in chaotic regimes

is expected behavior since first structure vector behaves as in other random processes which topological complexity turned to has the lowest complexity compared to other systems. Therefore, the most intense chaos has lower topological complexity than deterministic chaos.

Chapter 6

Conclusion and future perspectives

We found the proper way to recognize simplicial communities in emotion propagating network and find "key players" for switching emotions about the current subject. When emergence of knowledge occurs in temporal complex network, crucial role for creating collective knowledge is in the innovation channel containing high order simplices in it and its growth and the increased topological complexity over time provides the evolution pattern of the entire network. It systematically obeys the sensible connections of contents.

The detailed structure of the phase-space manifolds revealed by the algebraic topology technique contains vital knowledge about the systems collective behavior. The methods introduced here can yield insights into a variety of dynamical regimes occurring in complex systems, in particular, systems exhibiting a phase transition, percolation, explosive percolation, and others.

Regarding the social neurology problem, the study of higher-order combinatorial structures by algebraic topology techniques provides a sensitive methodology to quantify the shifts in the functional brain networks, e.g., under changed activity or condition. By complementing the standard graph theory methods, the algebraic topology can contribute to a more in-depth analysis of other brain imaging data.

Main results regarding the nonlinear dynamical systems showed us that complexity, both from topological and combinatorial aspects, of obtained simplicial complexes from chaotic attractors is more pronounced than in pre-chaotic attractors. Also, we found behavioral pattern which shows longer perseverance of homology group within connectivity classes in chaotic than in pre-chaotic dynamical regimes. Also, the dimension of homology group is highest at the bifurcation points. Using Takens embedding theorem and adapting algebraic topology tools for analyzing

nonlinear dynamical systems we are able to detect an unknown regime of dynamical system from a single time series, regardless of its inherent dimension.

Future perspectives for the studies which rely on topological methods are very broad since the methods may be applied to a large number of different complex systems from research areas such as sociology, biology, economics and others. The use of simplicial complex approach is applicable in every system that is defined in a discrete form.

All the results obtained by the use of algebraic topology, in particular by studying the properties of simplicial complexes formed from the available data, offer deep insight on hidden topology of complex systems which remain unnoticed by standard methods of analysis of complex systems, particularly the ones based on the analysis of time-series or graph theory.

Bibliography

[1] J. Ladyman, J. Lambert, K. Wiesner, *What is a complex system?,* Euro. Jour. Phil. Sci. June 2012.

[2] D. J. Watts, S. H. Strogatz, *Collective dynamics of small-world networks,* Nature 393, 440, 1998.

[3] D. Kozlov, *Combinatorial Algebraic Topology,* Algorithms and Computationin Mathematics, Springer-Verlag, Berlin Heidelberg, 2008.

[4] R. H. Atkin, *From cohomology in physics to q-connectivity in social sciences,* Int. J. Man-Machine Studies 4, 341, 1972.

[5] R. H. Atkin, *Combinatorial Connectivities in Social Systems,* Birkhauser Verlag, Base und Stuttgart, 1977.

[6] C. H. Dowker, *Homology groups of relations,* Annals of Mathematics 56, 84, 1952.

[7] R.H. Atkin, *Mathematical structure in human affairs,* Heinemann, London 1974.

[8] H. Barcelo, R. Laubenbacher, *Perspectives on A-homotopy theory and its applications,* Discrete Mathematics 298, 3961, 2005

[9] M. R. Muldoon, R. S. Mackay, J. P. Huke, D. S. Broomhead, *Topology from time series,* Physica D 65, 1-16, 1993

[10] C. W. Misner, K.S. Thorne, J. A. Wheeler, *Gravitation,* W.H. Freeman, San Francisco 1973.

[11] J. Frauendiener, *Discrete differential forms in general relativity,* Classical and Quantum Gravity, 23(16), S369-S385, 2006.

[12] A. Bossavit, *Computational Electromagnetism : Variational Formulations, Complementarity, Edge Elements,* Academic Press, 1998.

[13] S. H. Christiansen, T. G. Halvorsen, *A simplicial gauge theory,* J. Math. Phys. 53, 033501, 2012.

[14] A. Yavari, *On geometric discretization of elasticity,* Journal of Mathematical Physics 49(2), 022901-1-36, 2008.

[15] E. Gawlik, P. Mullen, D. Pavlov, J. E. Marsden, M. Desbrun, *Geometric, variational discretization of continuum theories,* Physica D 240, 1724, 2011.

[16] S. Maletić, M. Rajković, D. Vasiljević, *Simplicial Complexes of Networks and Their Statistical Properties,* Lecture Notes in Computational Science 5102(II), 568-575, 2008.

[17] D. Horak, S. Maletić, M. Rajković, *Persistent Homology of Complex Networks,* J. of Stat. Mech. 03, P03034, 2009.

[18] S. Maletić, L. Stamenić, M. Rajković, *Statistical mechanics on simplicial complexes,* Atti Semin. Mat. Fis. Univ. Modena Reggio Emilia 58, 23, 2011.

[19] S. Maletić, M. Rajković, *Simplicial Complex of Opinions on Scale-Free Networks,* Studies in Computational Intelligence, Springer 207, 127, 2009

[20] S. Maletić, M. Rajković, *Consensus formation on simplicial complex of opinions,* Physica A, in the procedure of review

[21] Norbert Marwan, M. Carmen Romano, Marco Thiel, Jurgen Kurths, *Recurrence plots for the analysis of complex systems* Physics Reports 438, 237 329, 2007

[22] L. Euler, *Solutio problematis ad geometriam situs pertinentis,* Comment. Acad. Sci. U. Petrop. 8, 128-140, 1736. Reprinted in Opera Omnia Series Prima 7, 1-10, 1976.

[23] P. Erdős, A. Rényi, *On random graphs,* Publicationes Mathematicae 6, 290, 1959.

[24] P. Erdős, A. Rényi, *On the evolution of random graphs,* Publications of Mathematical Institute of the Hungarian Academy of Sciences 5, 17, 1960.

[25] S. Milgram, *The small world problem,* Psychology Today 2, 60, 1967.

[26] R. Albert, A.-L. Barabási, *Statistical mechanics of complex networks,* Rev. Mod. Phys. 74, 47, 2002.

[27] D. J. Watts, S. H. Strogatz, *Collective dynamics of small-world networks,* Nature 393, 440, 1998.

[28] A.-L. Barabási, R. Albert, H. Jeong, *Mean-field theory for scale-free random networks,* Physica A 272, 173, 1999.

[29] J. R. Munkres, *Elements of Algebraic Topology,* Addison-Wesley Publishing, California, 1984.

[30] R. Atkin, *An Algebra of Patterns on a Complex I,* Int. J. Man-Machine Studies 6, 285, 1974.

[31] R. Atkin, *An Algebra of Patterns on a Complex II,* Int. J. Man-Machine Studies 8, 483, 1974.

[32] J. H. Johnson, *Some structures and notation of Q-analysis,* Environment and Planning B 8, 73, 1981.

[33] -s48 S. Maletić, M. Rajković, D. Vasiljević, *Simplicial Complexes of Networks and Their Statistical Properties,* Lecture Notes in Computational Science 5102(II), 568-575, 2008.

[34] -s49 D. Horak, S. Maletić, M. Rajković, *Persistent Homology of Complex Networks,* J. of Stat. Mech. 03, P03034, 2009.

[35] L. Lovász, *Knesers Conjecture, Chromatic Numbers and Homotopy,* J. Comb. Th. A 25, 319, 1978.

[36] F. G. Arenas, M. L. Puertas, *The Neighborhood Complex of an Infinite Graph,* Divulgaciones Matematicas 8, 69, 2000.

[37] -s10 D. Kozlov, *Combinatorial Algebraic Topology,* Algorithms and Computation in Mathematics, Springer-Verlag, Berlin Heidelberg, 2008.

[38] M. Andjelković, B. Tadić, S. Maletić, M. Rajković, *Hierarchical sequencing of online social graphs,* Physica A 436, 582, 2015.

[39] X. Dong, M. L. Wachs, *Combinatorial Laplacian of the matching complex,* Electronic Journal of Combinatorics 9, R17, 2002.

[40] A. Hatcher, *Algebraic Topology,* Cambridge University Press, Cambridge, 2002.

[41] V. de Silva, G. Carlsson, *Topological estimation using witness complexes,* In Symp. on Point-Based Graphics, pages 157, 2004.

[42] J. Garland, E. Bradley, J. D. Meiss, *Exploring the Topology of Dynamical Reconstructions,* arXiv:1506.01128v1

[43] J. Živković, M. Mitrović, B. Tadić, *Correlation Patterns in Gene Expressions along the Cell Cycle of Yeast,* Springer, Complex Networks pp 23-34,2009

[44] A. Madi, Y. Friedman, D. Roth, T. Regev, S. Bransburg-Zabary, E. B. Jacob, *Genome Holography: Deciphering Function-Form Motifs from Gene Expression Data,* PLoS ONE 3(7), e2708, 2008. doi: 10.1371/journal.pone.0002708 PMID: 18628959

[45] I. Baruchi, E. Ben-Jacob, *Functional holography of recorded neuronal networks activity,* Neuroinformatics 2(3), 333-351, 2004. doi: 10.1385/NI:2:3:333 PMID: 15365195

[46] F. Ballesteros, F. Luque, L. Lacasa, B. Luque, J. C. Nuno, *From time series to complex networks: the visibility graphs,* Proc. Natl. Acad. Sci. USA, 105:4972, 2008.

[47] M. Andjelković, N. Gupte, B. Tadić, *Hidden geometry of traffic jamming,* Phys. Rev. E, 91:052817, 2015.

[48] S. Maletić, M. Rajković, *Combinatorial Laplacian and entropy of simplicial complexes associated with complex networks,* Eur. Phys. J. Special Topics 212, 7797, 2012.

[49] K. Y. Degtiarev, *Systems Analysis: Mathematical Modeling and Approach to Strucutral Complexity Measure Using Polyhedral Dynamics Approach,* Complexity International 7, 1-22, 2000.

[50] M. Andjelkovic, B. Tadic ,S. Maletic, M. Rajkovic, *Hierarchical sequencing of online social graphs,* Physica A: Statistical Mechanics and its Applications 436, pp. 582-595, 2015

[51] -s48 S. Maletić, M. Rajković, D. Vasiljević, *Simplicial complexes of networks and their statistical properties,* in: Computational Science ICCS 2008, in: Lecture Notes in Computer Science, vol. 5102, Springer, Berlin Heidelberg, 2008, pp. 568575.

[52] -s75 J. R. Munkres, *Elements of Algebraic Topology,* Addison-Wesley Publishing, California, 1984.

[53] -s13 C. H. Dowker, *Homology groups of relations,* Annals of Mathematics 56, 84, 1952.

[54] -s10 D. Kozlov, *Combinatorial Algebraic Topology,* Algorithms and Computation in Mathematics, Springer-Verlag, Berlin Heidelberg, 2008

[55] T. E. Goldberg, *Combinatorial Laplacians of Simplicial Complexes,* Annandale-on-Hudson, New York, 2002.

[56]

[57] A. M. Duval, V. Reiner, *Shifted Simplicial Complexes are Laplacian Integral,* Transactions of the American Mathematical Society 354(11), 4313, 2002.

[58] -schom22 T. E. Goldberg, *Combinatorial Laplacians of Simplicial Complexes,* Annandale-on-Hudson, New York, 2002.

[59] W. V. D. Hodge, *The Theory and applications of harmonic integrals,* Cambridge at the University Press, 1952.

[60] J. Friedman, *Computing Betti numbers via combinatorial Laplacians,* Proc. 28th Annual ACM Symposium, Theory and Computations, 386, 1996

[61] M. Andjelkovic ,B. Tadic , M. Mitrović-Dankulov , M. Rajkovic, R. Melnik, Topology of innovation spaces in the knowledge networks emerging through questions-and-answers, PLoS ONE 11 (5), e0154655 (2016)

[62] M. Andjelkovic, N. Gupte, B. Tadic , *Hidden geometry of traffic jamming,* Physical Review E - Statistical, Nonlinear, and Soft Matter Physics 91 (5), 052817 2015

[63] B. Tadic, M. Andjelkovic, M. Suvakov, *The influence of architecture on collective charge transport in nanoparticle assemblies revealed by the fractal time series and topology of phase space manifolds,* Journal of Coupled Systems and Multiscale Dynamics, 4, 1, pp. 30-42(13), 2016

[64] B. Tadic ,M. Andjelkovic ,B.M. Boshkoska.Z. Levnajic, *Algebraic topology of multi-brain connectivity networks reveals dissimilarity in functional patterns during spoken communications,* PLoS ONE 11 (11), e0166787 2016

[65] B. Tadić, V. Gligorijević, M. Mitrović, M. Šuvakov, *Co-evolutionary mechanisms of emotional bursts in online social dynamics and networks,* Entropy 15 50845120 2013.

[66] M. Šuvakov, M. Mitrović, V. Gligorijević, B. Tadić, *How the online social networks are used: dialogues-based structure of myspace,* Journal of the Royal Society Interface 10 20120819, 2012.

[67] M. Szell, S. Thurner, *Measuring social dynamics in a massive multiplayer online game,* Social Netw. 32, 313329, 2010.

[68] S. Thurner, M. Szell, R. Sinatra, *Emergence of good conduct, scaling and Zipf laws in human behavioral sequences in an online world,* PLOS ONE 7 e29796, 2012.

[69] V. Gligorijević, M. Skowron, B. Tadić, *Structure and stability of online chat networks built on emotion-carrying links,* Physica A 392 538543, 2013.

[70] J. C. Nacher, T. Akutsu, *Analysis of critical and redundant nodes in controlling directed and undirected complex networks using dominating sets,* J. Complex Networks 2 394412, 2014.

[71] P. Csermely, A. London, L.-Y. Wu, B. Uzzi, *Structure and dynamics of core/periphery networks,* J. Complex Networks 1 93123, 2013.

[72] M. Szell, R. Lambiotte, S. Thurner, *Multirelational organization of large-scale social networks,* Proc. Natl Acad Sci. USA 107 1363613641, 2010.

[73] J. Gómez-Gardeñes, I. Reinares, A. Arenas, L. M. Floría, *Evolution of cooperation in multiplex networks,* Sci. Rep. 2 620, 2012.

[74] K.-M. Lee, J. Y. Kim, W.-K. Cho, K.-I. Goh, I.-M. Kim, *Correlated multiplexity and connectivity of multiplex random networks,* New J. Phys. 14 033027, 2012

[75] A. Cardillo, J. Gómez-Gardeñes, M. Zanin, M. Romance, D. Papo, F. del Pozo, S. Boccaletti, *Emergence of network features from multiplexity,* Scientific Reports 3 1344, 2013.

[76] E. Estrada, E. Vargas-Estrada, *How peer pressure shapes consensus, leadership, and innovations in social groups,* Scientific Reports 3 2905, 2013.

[77] S. Maletić, M. Rajković, *Consensus formation on a simplicial complex of opinions,* Physica A 397 111120, 2014.

[78] S. Maletić, D. Horak, M. Rajković, *Cooperation, conflict and higher-order structures of social networks,* Adv. Complex Syst. 15 (Suppplement 1) 1250055, 2012.

[79] B. Tadić, G. J. Rodgers, S. Thurner, *Transport in complex networks: Flow, Jamming and optimization,* Internat. J. Bifur. Chaos 17 23632385, 2007.

[80] N. Pržulj, *Biological network comparison using graphlet degree distribution,* Bioinformatics 23 e177e183, 2007.

[81] L. da Fontoura Costa, R. Fernandes Silva Andrade, *What are the best concentric descriptors for complex networks?,* New J. Phys. 9 311, 2007.

[82] E. Estrada, *Topological structural classes of complex networks* Phys. Rev. E 75 016103, 2007.

[83] N. Malod-Dognin, N. Pržulj, *Gr-align: fast and flexible alignment of protein 3d structures using graphlet degree similarity,* Bioinformatics 30 12591265, 2014.

[84] T. Milenković, N. Pržulj, *Uncovering biological network function via graphlet degree signatures,* Cancer Inform. 6 257273, 2008.

[85] E. Estrada, *Generalization of topological indices,* Chem. Phys. Lett. 336 248252, 2001.

[86] M. Mitrović, B. Tadić, *Spectral and dynamical properties in classes of sparse networks with mesoscopic inhomogeneities,* Phys. Rev. E 80 026123, 2009.

[87] J. Jonsson, *Simplicial Complexes of Graphs,* in: Lecture Notes in Mathematics, Springer-Verlag, Berlin Heidelberg, 2008.

[88] M. Kivela, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, M. A. Porter, *Multilayer networks,* J. Complex Networks 2 203271, 2014.

[89] G. Paltoglou, S. Gobron, M. Skowron, M. Thelwall, D. Thalmann, *Sentiment analysis of informal textual communication in cyberspace,* in: Proc. Engage 2010, Springer LNCS State-of-the-Art Survey, pp. 1325, 2010.

[90] J. A. Russell, *A circumplex model of affect,* J. Pers. Soc. Psychol. 39 11611178, 1980.

[91] http://www.cyberemotions.eu/ Collective emotions in cyberspace.

[92] M. Šuvakov, D. Garcia, F. Schweitzer, B. Tadić, *Agent-based simulations of emotion spreading in online social networks,* http://arxiv.org/abs/1205.6278, 2012.

[93] B. Tadić, M. Šuvakov, D. Garcia, F. Schweitzer, *Agent-based simulations of emotional dialogs in online social network MySpace,* in: Janusz Holyst (Ed.), Cyberemotions: collective emotions in Cyberspace, in: Springer Complexity Series, 2015.

[94] J. Sylvan Katz, *Indicators for complex innovation systems,* Res. Policy 35 893909, 2005.

[95] J. Živkovic, B. Tadić, N. Wick, S. Thurner, *Statistical indicators of collective behavior and functional clusters in gene expression network of yeast* European Phys. J. B 50 255258, 2006.

[96] F. Font-Clos, G. Boleda, Á. Corral, *A scaling law beyond Zipfs law and its relation to Heaps law,* New J. Phys. 15 093033, 2013.

[97] M. E. J. Newman, *Mixing patterns in networks,* Phys. Rev. E 67 026126, 2003.

[98] M. Šuvakov, B. Tadić, *Collective emotion dynamics in chats with agents, moderators and bots,* Condens. Matter Phys. 17 3380133812, 2014.

[99] V. Latora, V. Nicosia, P. Panzarasa, J. Stat. Phys. 151 745764, 2013.

[100] B. Tadić, *Dynamics of directed graphs: the world-wide Web,* Physica A 293 273284, 2001.

[101] G. C. Bowker, S. Leight Star, W. Turner, L. Gasser, *Social science, technical systems, and cooperative,* work. Psychology Press, New York, 2014.

[102] J. I. M. Carpendale, U. Müller, editors, *Social Interactions and the Development of Knowledge,* Lawrence Erlbaum Associates, Inc. Mahwah, New Jersey, 2013.

[103] J. Kimmerle, U. Kress, Ch. Held, *The interplay between individual and collective knowledge: technologies for organisational learning,* Knowledge Management Research & Practice 8, 3344, 2010.

[104] R. F. Kitchener, *Piagets Social Epistemology,* in Social Interactions and the Development of Knowledge, editors J. I. M. Carpendale, U. Müller, Lawrence Erlbaum Associates, Inc. Mahwah, New Jersey, pp. 4566, 2013.

[105] M. Mitrović Dankulov, R. Melnik, B. Tadic, *The dynamics of meaningful social interactions and the emergence of collective knowledge,* Scientific Reports 5, 12197, 2015

[106] K. Boudreau, P. Gaule, K. R. Lakhani, Ch. Riedl, A. Woolley, *From crowd to collaborators: initiating effort and catalyzing interactions among online creative workers,* Harvard Business School, working paper, 2014.

[107] K. R. Lakhani, E. von Hippel, *How open source software works: free user-to-user assistance,* Res. Policy 32, 923943, 2003

[108] G. Fortino, S. Galzarano, R. Gravina, W. A. Li, *A framework for collaborative computing and multi-sensor data fusion in body sensor networks,* Infor. Fusion 22, 5070, 2015

[109] S. Sloan, K. Bodey, R. Gyrd-Jones, *Knowledge sharing in online brand communities,* Quantitative Market Research 18(3), 320345, 2015

[110] M. Seraj, *We create, we connect, we respect, therefore we are: intellectual, social, and cultural value in online communities,* J. Interactive Marketing 26, 209222, 2012

[111] L. Blasco-Arcas, B. I. Hernandez-Ortega, J. Jiminez-Martinez, *Collaborating online: the roles of interactivity nad personalization,* The Service Industries J. 34(8), 677698, 2014

[112] I. Kotsireas, R. V. N. Melnik, B. West, editors, *Advances in Mathematical and Computational Methods: Addressing Modern Challenges of Science, Technology and Society,* American Institute of Physics (2011) Vol. 1368.

[113] A. E. Thessen, C. Sims Parr, *Knowledge extraction and semantic annotation of text from the encyclopedia of life,* PLoS ONE 9, e0089550, 2014

[114] G. Leibon, D. N. Rockmore, *Orienteering in knowledge spaces: The hyperbolic geometry of wikipedia mathematics,* PLoS ONE 8, e67508, 2012

[115] S. H. Uddin, A. Khan, L. A. Baur, *A framework to explore the knowledge structure of multidisciplinary research fields,* PLOS ONE 10, e0123537, 2015

[116] S. Fortunato, F. Radicchi, A. Vespignani, in book Citation Networks. Springer-Verlag Berlin Heidleberg, 2012.

[117] T. Kuhn, M. Perc, D. Heilbing, *Inheritance patterns in citation networks reveal scientific memes,* Phy. Rev. X 4, 041036, 2014.

[118] -myspace25 J. Jonsson, *Simplicial Complexes of Graphs,* Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2008.

[119] -sc 8 M. Andjelković, B. Tadić, S. Maletić, M. Rajković, *Hierarchical sequencing of online social graphs,* Physica A: Statistical Mechanics and its Applications 436, 582595, 2015

[120] -qnoise 33 M. Andjelković, N. Gupte, B. Tadić, *Hidden geometry of traffic jamming,* Phys. Rev. E 91, 052817, 2015

[121] C. Bron, J. Kerbosch, *Finding all cliques of an undirected graph,* Comm. ACM 16, 575577, 1973

[122] H. J. Bandelt, V. Chepoi, *Metric graph theory and geometry,* a surveyin J. E. Goodman; J. Pach; R. Pollack, editors. Surveys on discrete and computational geometry: Twenty years later. Contemporary Mathematics 453, 4986, 2008

[123] M. Diananti, *Unwinding the hairball graph: pruning algorithms for weighted complex networks,* Phys. Rev. E 93, 012304, 2016

[124] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, *Fast unfolding of communities in large networks,* Journal of Statistical Mechanics: Theory and Experiment 10, P10008, 2008

[125] A. Lancichinetti, M. Kivela, J. Saramaki, S. Fortunato, *Characterizing the community structure of complex networks,* PLoS ONE 5(8), e11976, 2010

[126] R. Lambiotte, J.-C. Delvenne, M. Barahona, *Laplacian Dynamics and Multiscale Modular Structure in Networks,* IEEE Transactions on Network Science and Engineering 1(2), 7690, 2015

[127] S. Dorogovtsev, *Lectures on Complex Networks,* Oxford University Press, Inc., New York, NY, USA, 2010.

[128] F. Tria, V. Loreto, V. D. P. Servedio, S. H. Strogatz, *The dynamics of correlated novelties,* Scientific Reports 4, 5890, 2014

[129] M. A. P. Sloot, R. Quax, *Information processing as a paradigm to model and simulate complex systems,* Journal of Computational Science 3, 247249, 2012

[130] B. Tadić, V. Gligorijević, M. Mitrović, M. Šuvakov, *Co-evolutionary mechanisms of emotional bursts in online social dynamics and networks,* Entropy 15(12), 50845120, 2013

[131] E. Estrada, J. Gómez-Gardeñes, *Communicability reveals a transition to coordinated behavior in multiplex networks,* Phys. Rev. E 89, 042819, 2014

[132] F. Shahandeh, J. Sperling, and W. Vogel, Phys. Rev. Lett. 113, 260502, 2014.

[133] A . Corral, Phys. Rev. E 69, 026107, 2004.

[134] N. Malod-Dognin and N. Przulj, Bioinformatics 30, 1259, 2014

[135] M. Kramar, A. Goullet, L. Kondic, and K. Mischaikow, Phys. Rev. E 90, 052203, 2014

[136] M. Andjelkovic, B. Tadic, S. Maletic, M. Rajkovic, Physica A, 2015.

[137] B. Tadic, G. J. Rodgers, and S. Thurner, Int. J. Bifurcation Chaos Appl. Sci. Eng. 17, 2363, 2007.

[138] A. Arenas, A. Daz-Guilera, and R. Guimera, Phys. Rev. Lett. 86, 3196 2001; G. Mukherjee and S. S. Manna, Phys. Rev. E 71, 066108 2005; B. Tadic and S. Thurner, Physica A 346, 183 2005; L. Zhao, Y.-C. Lai, K. Park, and N. Ye, Phys. Rev. E 71, 026125 2005; S. Mukherjee, N. Gupte, andG.Mukherjee, ibid. 81, 046109 2010; M. Garavello and B. Picolli, Traffic Flow on Networks, Applied Mathematics Vol. 1 (AIMS, Springfield, USA, 2012.

[139] B. Tadic, S. Thurner, and G. J. Rodgers, Phys. Rev. E 69, 036102, 2004.

[140] W.Min and L.Wynter, *The Dynamics of Complex Urban Systems,* Physica-Verlag HD, Heidelberg, Germany, 2008.

[141] W. Whitt, *Stochastic Process Limits: An Introduction of Stochastic Process Limits and Its Application to Queues*, Springer Series in Operations Research Springer, New York, 2001.

[142] R. A. da Costa, S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Phys. Rev. Lett. 105, 255701, 2010; A. D. Kachhvah and N. Gupte, Phys. Rev. E 86, 026104 2012.

[143] A. S. L. O. Campanhero, M. I. Sirer, R. D. Malmgren, F. M. Ramos, and L. A. N. Amaral, PLoS One 6, e23378 2011; P. N. Papanicolau, A. K. Charakopoulos, T. E. Karakasidis, and A. Liakopoulos, Chaos 24, 024408 2014; J. Zhang and M. Small, Phys. Rev. Lett. 96, 238701 2006; T. Weng, Y. Zhao, M. Small, and D. D. Huang, Phys. Rev. E 90, 022804 2014; S. Hempel, A. Koseska, J. Kurths, and Z. Nikoloski, Phys. Rev. Lett. 107, 054101 2011.

[144] F. Ballesteros, F. Luque, L. Lacasa, B. Luque, and J. C. Nuno, Proc. Natl. Acad. Sci. USA 105, 4972 2008.

[145] L. Lacasa, F. Ballesteros, and R. Robledo *In the case of stochastic unimodalmaps all stationary trajectories need to be identified and mapped into graphs,* PLoS One 6, e22411 (2011).

[146] -tag22 H.-J. Bandelt and V. Chepoi, Metric Graph *Theory and Geometry*, Contemporary Mathematics Vol. 453 AMS, Newport, RI, 2008.

[147] P. K. Dixon, L.Wu, S. R. Nagel, B. D.Williams, and J. P. Carini, Phys. Rev. Lett. 65, 1108 1990; P. Lukenheimer and A. Loidl, in Broadband Dielectric Spectroscopy, edited by F. Kremer and A. Schonhals (Springer-Verlag, Berlin, pp. 131169, 2002.

[148] S. Meloni, A. Arenas, and Y. Moreno, Proc. Natl. Acad. Sci. USA 106, 16897 2009.

[149] I. Kotsireas, R. V. N. Melnik, B. West, editors, *Advances in Mathematical and Computational Methods: Addressing Modern Challenges of Science, Technology and Society,* American Institute of Physics, Vol. 1368, 2011.

[150] P. Moriarty, *Nanostructured materials,* Reports on Progress in Physics, 64:297–381, 2001.

[151] J. Liao, S. Block, S. J. van der Molen, S. Diefenbach, A. W. Holleitner, C. Schönenberg, A. Vladyka, M. Calame, *Ordered nanoparticle arrays interconnected by molecular linkers: electronic and optoelectronic properties,* Chem. Soc. Review, 44:999–1014, 2015.

[152] J. Živković, B. Tadić, *Nanonetworks: The graph theory framework for modeling nanoscale systems,* Nanoscale Systems MMTA, 2:30–48, 2013.

[153] D. K. Ferry, S. M. Goodnick, *Transport in Nanostructures,* Cambridge University Press, 1997.

[154] I. S. Beloborodov, A. V. Lopatin, V. M. Vinokur, K. B. Efetov, em Granular electronic systems, Rev. Mod. Phys., 79:469, 2007.

[155] M. Šuvakov, B. Tadić, *Modeling collective charge transport in nanoparticle assemblies,* J. Phys. Condens. Matt., 22(16):163201, 2010.

[156] H. Fan, K. Yang, D. M. Boye, T. Sigmon, K. J. Malloy, H. Xu, G. P. Lopez, C. J. Brinker, *Self-assembly of ordered, robust, three-dimesionalgold nanocrystal/silica arrays,* Science, 304:567, 2004.

[157] M. O. Blunt, M. Šuvakov, F. Pulizzi, C. P. Martin, E. Pauliac-Vaujour, A. Stannard, A. W. Rushforth, B. Tadić, P.Moriarty, *Charge transport in cellular nanoparticle networks: meandering through nanoscale mazes,* Nano Letter, 7(4):855, 2007.

[158] M. O. Blunt, A. Stannard, E. Pauliac-Vaujour, C. P. Martin, I. Vancea, M. Šuvakov, U. Thiele, B. Tadić, P. Moriarty, *Patterns and pathways in nanoparticle self-organization,* Oxford Univ. Press, Oxford, UK, 2010.

[159] D. Joung, L. Zhai, S. I. Khondaker, *Coulomb blockade and hopping conduction in graphene quantum dots,* Phys. Rev. B, 83, 2011.

[160] D. Joung and S. I. Khondaker., *Structural evolution of reduced graphene oxyde of varying carbon* sp^2 *frcations investigated via Coulomb blockade transport,* J. Phys. Chem. C, 117:26776-26782, 2013.

[161] K. P. Loh, Q. Bao, M. Chhowalla, *Graphene oxide as a chemically tunable platform for optical applications,* em Nat. Chemistry, 2:1015-1024, 2010.

[162] H. Kohno, S. Takeda, *Non-Gaussian fluctuations in the charge transport of Si nanonchains,* Nanotechnology, 18:395706, 2007.

[163] I. S. Beloborodov, A. V. Lopatin, V. M. Vinokur, *Coulomb effects and hopping transport in granular metals,* Phys. Rev. B., 72:125121, 2005.

[164] T. B. Tran, I. S. Beloborodov, X. M. Lin, T. P Bigioni, V. M. Vinokur, H. M. Jaeger, *Multiple cotunneling in large quantum dot arrays,* Phys. Rev. Lett., 95:076806, 2005.

[165] W. A. Schoonveld, J. Wildeman, D. Fichou, P. A. Bobbert, B. J. van Wees, T. M. Klapwijk, *Coulomb-blockade transport in single-crystal organic thin film transistors,* Nature, 404:977–980, 2000.

[166] J. Kane, M. Inan, R. F. Saraf, *Self-assembled nanoparticle necklaces network showing single-electron switching at room temperature and biogating current by living microorganisms,* ACS Nano, 4:317–323, 2010.

[167] A. A. Middleton, N. S. Wingreen, *Collective transport in arrays of small metallic dots,* Phys. Rev. Lett., 71(19):3198–3201, 1993.

[168] A. J. Rimberg, T. R. Ho, J. Clarke, *Scaling behavior in the current-voltage characteristic of one- and two-dimensional arrays of small metallic islands,* Phys. Rev. Lett., 74(23):4714–4717, 1995.

[169] M. N. Wybourne, L. Clarke, M. Yan, S. X. Cai, L. O. Brown, J. Hutchison, J. F. W. Keana, *Coulomb-blockade dominated transport in patterned gold-cluster structures,* Jpn. J. Appl. Phys., 36:7796–7800, 1997.

[170] R. Parthasarathy, X. Lin, H. M. Jaeger, *Electronic transport in metal nanocrystal arrays: The effect of structural disorder on scaling behavior,* Phys. Rev. Lett., 87(18):186807, 2001.

[171] R. Parthasarathy, X. Lin, K. Elteto, T. F. Rosenbaum, H. M. Jaeger, *Percolating through networks of random thresholds: Finite temperature electron tunneling in metal nanocrystal arrays,* Phys. Rev. Lett., 92(7):076801, 2004.

[172] T. S. Basu, S. Ghosh, S. Gierolta ka, M. Ray, *Collective charge transport in semiconductor-metal hybrid nanocomposites,* Appl. Phys. Lett., 102, 053107, 2013.

[173] C.H. George, I. Szleifer, M. Ratner, *Multiple-time-scale motion in molecularly linked nanoparticle arrays,* ACS NANO, 7:108–116, 2013.

[174] T. B. Tran, I. S. Beloborodov, J. Hu, X. M. Lin, T. F. Rosenbaum, H. M. Jaeger, *Sequential tunneling and inelastic cotunneling in nanoparticle arrays,* Phys. Rev. B., 78:075437, 2008.

[175] A. Zabet-Khosousi, A.-A. Dihrani, *Charge transport in nanoparticle assemblies,* Chem. Rev., 108:4072–4124, 2008.

[176] J. K. Y. Ong, Ch. Van Nguyen, S. Sayood, R. F. Saraf, *Imaging electroluminescence from individual nanoparticles in an array exhibiting room temperature single electron effect,* ACS NANO, 7:7403–7410, 2013.

[177] M. Šuvakov, B. Tadić, *Topology of cell-aggregated planar graphs,* In V. Alexandrov *et al.*, editor, *Computational Science — ICCS 2006*, volume 3993 of Lecture Notes in Computer Science, pages 1098–1105, Berlin, 2006. Springer.

[178] M. Šuvakov, B. Tadić, *Simulation of the electron tunneling paths in network of nano-particle films,* In Y. Shi *et al.*, editor, Computational Science - ICCS 2007, volume 4488 of *Lecture Notes in Computer Science*, pages 641–648, Berlin, 2007.

[179] B. Luque, L. Lacasa, F. Ballesteros, A. Robledo, *Feigenbaum graphs: A complex network perspective of chaos,* PLOS One, 6, e22411, 2011.

[180] -qnoise 32 F. Ballesteros, F. Luque, L. Lacasa, B. Luque, J. C. Nuno, *From time series to complex networks: the visibility graphs,* Proc. Natl. Acad. Sci. USA, 105:4972, 2008.

[181] -qnoise 33 M. Andjelković, N. Gupte, B. Tadić, *Hidden geometry of traffic jamming,* Phys. Rev. E, 91:052817, 2015.

[182] T. Narumi, M. Suzuki, Y. Hidaka, T. Asai, S. Kai, *Active Brownian motion in threshold distribution of a Coulomb blockade model,* Phys. Rev. E, 84:051137, 2011.

[183] T. Narumi, M. Suzuki, Y. Hidaka, S. Kai, *Size dependence of current–voltage properties in Coulomb blockade networks,* J. Phys. Soc. Japan, 80:114704, 2011.

[184] B. Tadić, *Dynamic criticality in driven disordered systems: role of depinning and driving rate in Barkhausen noise,* Physica A: Statistical and Theoretical Physics, 270:125, 1999.

[185] M. Mitrović Dankulov, R. Melnik, B. Tadić, *The dynamics of meaningful social interactions and the emergence of collective knowledge,* Scientific Reports, 5:12197, 2015.

[186] C. Tsallis, *Introduction to nonextensive statistical mechanics : approaching a complex world,.* Springer, N.Y., 2009.

[187] A. Pluchino, A. Rapisarda, C. Tsallis, *Noise, synchrony, and correlations at the edge of chaos,* Phys. Rev. E, 87:022910, 2013.

[188] G. P. Pavlos, L. P. Karakatsanis, M. N. Xenakis, E. G. Pavlos, A. C. Iliopoulos, D. V. Sarafopoulos, *Universality of non-extensive Tsallis statistics and time series analysis: Theory and applications,* Physica A: Statistical Mechanics and its Applications, 395:58 – 95, 2014.

[189] A. N. Pavlov, V. S. Anishchenko, *Multifractal analysis of complex signals,* Physics–Uspekhi, 50:819–834, 2007.

[190] S. B. Lowen, M. C. Teich, *Estimation and simulation of fractal stochastic point processes,* Fractals, 3:183–210, 1995.

[191] M. Sadegh Movahed, G. R. Jafari, F. Ghasemi, S. Rahvar, M. R. Rahimi Tabar, *Multifractal detrended fluctuation analysis of sunspot time series,* Journal of Statistical Mechanics: Theory and Experiment, 2006(02):P02003, 2006.

[192] B. Bollobas, *Modern Graph Theory,* Springer, New York, 1998.

[193] -tag27 S. Dorogovtsev, *Lectures on Complex Networks,* Oxford University Press, Inc., New York, 2010.

[194] -myspace25 J. Jonsson, *Simplicial Complexes of Graphs,* Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2008.

[195] -sc8 M. Andjelković, B. Tadić, S. Maletić, M. Rajković, *Hierarchical sequencing of online social graphs,* Physica A: Statistical Mechanics and its Applications, 436:582 – 595, 2015.

[196] I. Kondić, A. Goullet, C. S. O'Hern, M. Kramar, K. Mishaikow, R. P. Behringer, *Topology of force networks in compressed granular media,* Eur. Phys. Lett., 97:54001, 2012.

[197] -tag21 C. Bron and J. Kerbosch, *Finding all cliques of an undirected graph,* Comm. ACM, 16:575-577, 1973.

[198] -tag22 H. J. Bandelt, V. Chepoi, *Metric graph theory and geometry,* a survey, in J. E. Goodman, J. Pach, R. Pollack, editors. "Surveys on discrete and computational geometry: Twenty years later". Contemporary Mathematics, 453:49–86, 2008.

[199] M. Šuvakov, B. Tadić, *Transport processes on homogeneous planar graphs with scale-free loops,* Physica A: Statistical Mechanics and its Applications, 372:354–361, 2006.

[200] F. Babiloni, F. Cincotti, C. Babiloni, F. Carducci, D. Mattia, L. Astolfi, et al., *Estimation of the cortical functional connectivity with the multimodal integration of high-resolution EEG and fMRI data by directed transfer function,* NeuroImage 24(1), 118-131, 2005

[201] J. A. McNab, B. L. Edlow, T. Witzel, S. Y. Huang, H. Bhat, Heberlein Kea, *The human connectome project and beyond: Initial applications of 300 mT/m gradients,* NeuroImage 80, 24-245, 2013

[202] E. Bullmore, O. Sporns, *Complex brain networks: graph theoretical analysis of structural and functional systems,* Nature Rev Neurosci 10, 186-198, 2009

[203] M. Rubinov, O. Sporns, *Complex network measures of brain connectivity: Uses and interpretations,* NeuroImage 52(3), 1059-1069, 2010.

[204] X. Shen, X. Papademetris, R. T. Constable, *Graph-theory based parcellation of functional subunits in the brain from resting-state fMRI data,* NeuroImage 50(3), 1027-1035, 2010

[205] A. Zalesky, A. Fornito, I. H. Harding, L. Cocchi, M. Yocel, C. Pantelis, et al., *Whole-brain anatomical networks: Does the choice of nodes matter?,* NeuroImage 50(3), 970-983, 2010

[206] P. Martijn, J. van den Heuvel, O. Sporns, *Network hubs in the human brain,* Trends in Cognitive Sciences 17(12), 683-696, 2013.

[207] M. Hassan, O. Dufor, I. Merlet, C. Berrou, F. Wendling, *EEG Source Connectivity Analysis: From Dense Array Recordings to Brain Networks,* PLoS ONE 9(8), e105041, 2014

[208] D. S. Bassett, M. A. Porter, N. F. Wymbs, S. T. Grafton, J. M. Carlson, P. J. Mucha, *Robust detection of dynamic community structure in networks,* Chaos , 23(1), 013142, 2013

[209] B. Garcia-Martinez, A. Martinez-Rodrigo, R. Zangrniz Cantabrana, J. M. Pastor Garcia, R. Alcaraz, *Application of Entropy-Based Metrics to Identify Emotional Distress from Electroencephalographic Recordings,* Entropy 18(6), 221, 2016.

[210] C. Gaier, K. Lhnertz, S. Bialonski, *Time-dependent degree-degree correlations in epileptic brain networks: from assortative to dissortative mixing,* Front Hum Neurosci 9, 462, 2015

[211] S. Sockeel, D. Schwartz, M. Pelegrini-Issac, H. Benali, *Large-Scale Functional Networks Identified from Resting-State EEG Using Spatial ICA,* PLoS ONE 11(1), 1-18, 2016

[212] D. Sammler, M. Grigutsch, T. Fritz, S. Koelsch, *Music and emotion: Electrophysiological correlates of the processing of pleasant and unpleasant music,* Psychophysiology 44, 293-304, 2007.

[213] E. C. Padovani, *Characterisation of the Community Structure of Large-Scale Functional Brain Networks During Ketamine-Mdetomidine Anesthetic Induction,* arxiv:q-bio 2016

[214] C. S. Parker, F. Deligianni, M. J. Cardoso, P. Daga, M. Modat, M. Dayan, et al., *Consensus between Pipelines in Structural Brain Networks,* PLoS ONE 9(10), 1-10, 2014

[215] G. Deco, G. Tononi, M. Boly, M. L. Kringelbach, *Rethinking segregation and integration: contributions of whole-brain modelling,* Nat Rev Neurosci 16(7), 430-439, 2015

[216] B. Mišić, O. Sporns, A. R. McIntosh, *Communication efficiency and congestion of signal traffic in largescale brain networks,* PLoS Comp Biol 10, e1003427, 2014

[217] B. Mišić, J. Goñi, R. F. Betzel, O. Sporns, A. R. McIntosh, A network convergence zone in the hippocampus, PLoS Comp Biol 10, e1003982, 2014

[218] T. Womelsdorf, S. Everling, *Long-Range Attention Networks: Circuit Motifs Underlying Endogenously Controlled Stimulus Selection,* Trends in Neurosciences 38(11), 682-700, 2015

[219] G. Shi, F. Pasqualetti, M. Cieslak, Q. K. Telesford, A. B. Yu, A. E. Kahn, et al., *Controllability of structural brain networks,* Nature Communications 6, 8414, 2015

[220] R. J. Krauzlis, A. Bollimunta, F. Arcizet, L. Wang, *Attention as an effect not a cause,* Trends in Cognitive Sciences 18(9), 457-464, 2016

[221] J. W. Bisley, M. E. Goldberg, *Attention, Intention, and Priority in the Parietal Lobe,* Annu Rev Neurosci 33, 1-21, 2010

[222] D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, S. T. Grafton, *Dynamic reconfiguration of human brain networks during learning,* Proceedings of the National Academy of Sciences 108 (18), 7641-7646, 2011

[223] M. Riemer, N. Diersch, F. Bublatzky, T. Wolbers, *Space, time, and numbers in the right posterior parietal cortex: Differences between response code associations and congruency effects,* NeuroImage 129, 72-79, 2016

[224] S. Y. Kim, T. Qi, X. Feng, G. Ding, L. Liu, F. Cao, *How does language distance between L1 and L2 affect the L2 brain network?,* An fMRI study of Korean, Chinese, English trilinguals. NeuroImage 129, 25-39, 2016

[225] L. L. Zeng, H. Shen, L. Liu, L. Wang, B. Li, P. Fang, et al., *Identifying major depression using whole-brain functional connectivity: a multivariate pattern analysis,* Brain 14 , 1-10, 2012

[226] R. Adolphs, *The Social Brain: Neural Basis of Social Knowledge,* Annu Rev Psychol 60, 693-716, 2009

[227] D. M. Amodio, C. D. Frith, *Meeting of minds: the medial frontal cortex and social cognition,* Nature Rev Neurosci 7, 268-277, 2006

[228] L. Q. Uddin, M. Iacoboni, C. Lange, J. P. Keenan, *The self and social cognition: the role of cortical midline structures and mirror neurons,* Trends in Cognitive Sciences 11(4), 153-157, 2007

[229] I. Konvalinka, A. Roepstorff, *The two-brain approach: how can mutually interacting brains teach us about social interactions?,* Front Hum Neurosci 6, 215, 2012

[230] K. Yun, K. Watanabe, S. Shimojo, *Interpersonal body and neural synchronization as a marker of implicit social interaction,* Sci Rep 2, 959, 2012

[231] G. Dumas, J. Nadel, R. Soussignan, J. Martinerie, L. Garnero, *Inter-brain synchronization during social interaction,* PLoS ONE 5, e12166, 2010

[232] F. Krueger, K. McCabe, J. Moll, N. Kriegerskorte, R. Zahn, et al., *Neural correlates of trust,* PNAS 104, 20084-20089, 2007

[233] F. De Vico Fallani, V. Nicosia, R. Sinatra, L. Astolfi, F. Cincotti, et al., *Defecting or not defecting: how to "read" human behavior during cooperative games by EEG measurements,* PLoS ONE 5, e14187, 2010

[234] J. Jiang, B. Dai, D. Peng, C. Zhu, L. Liu, et al., *Neural Synchronization during Face-to-Face Communication,* J Neurosci 32, 16064-16069, 2012

[235] J. Jiang, C. Cheng, B. Dai, G. Shi, G. Ding, L. Liu, et al., *Leader emergence through interpersonal neural synchronization,* PNAS 12(14), 4274-4279, 2015

[236] N. Liu, C. Mok, E. E. Witt, A. H. Pradhan, J. E. Chen, A. L. Reiss, *NIRS-Based Hyperscanning Reveals Interbrain Neural Synchronization during Cooperative Jenga Game with Face-to-Face Communication,* Frontiers in Human Neuroscience 10(82), 1-11, 2016

[237] V. Müller, J. Sänger, U. Lindenberger, *Intra- and Inter-Brain Synchronization during Musical Improvisation on the Guitar,* PLoS ONE 8(9), e73852, 2013

[238] L. Duan, W. J. Liu, R. N. Dai, R. Li, C. M. Lu, Y. X. Huang, C. Z. Zhu, *Cross-Brain Neurofeedback: Scientific Concept and Experimental Platform,* PLoS ONE 8(5), e64590, 2013

[239] R. N. Dai, X. Xiao, P. P. Sun, Z. Li, C. Z. Zhu, *Cluster imaging of multi-brain networks (CIMBN): a general framework for hyperscanning and modeling a group of interacting brains,* Frontiers Neurosci 9, 267, 2015

[240] A. K. Kuhlen, C. Allefeld, J. D. Haynes, *Content-specific coordination of listeners' to speakers' EEG during communication,* Frontiers in Human Neuroscience 6(266), 1-15, 2012.

[241] F. D. V. Fallani, J. Richardi, M. Chavez, S. Achard, *Graph analysis of functional brain networks: practical issues in translational neuroscience,* 2014 arxiv

[242] M. Chavez, M. Valencia, V. Navarro, V. Latora, J. Martinerie, *Functional modularity of background activities in normal and epileptic brain networks,* Phys Rev Lett 104, 118701, 2010

[243] R. N. Mantegna, E. H. Stanley, *An introduction to econophysics: correlations and complexity in finance,* Cambridge University Press; 2000

[244] B. Tadić, M. Mitrović, *Jamming and Correlation Patterns in Traffic of Information on Sparse Modular Networks,* European Physical Journal B 71(4), 631-640, 2009

[245] -bb-book B. Bollobas, *Modern Graph Theory,* Springer-Verlafg, Berlin Heidelberg; 1998

[246] -tag27 S. Dorogovtsev, *Lectures on Complex Networks,* New York, USA: Oxford University Press, Inc.; 2010

[247] D. Meunier, R. Lambiotte, E. T. Bullmore, *Modular and hierarchically modular organization of brain networks,* Frontiers Neurosci 4, 200, 2010

[248] S. Fortunato, *Community detection in graphs,* Physics Reports 486(3-5), 75-174, 2010.

[249] -tag 25 A. Lancichinetti, M. Kivela, J. Saramaki, S. Fortunato, *Characterizing the Community Structure of Complex Networks,* PLoS ONE 5(8), e11976, 2010.

[250] X. Gao, B. Xiao, D. Tao, *A survay of graph edit distance,* Pattern Anal Applic 13, 113-129, 2010

[251] G. Gronchi, A. Guazzini, E. Massaro, F. Bagnoli, *Mapping cortical functions with a local community detection algorithm,* Journal of Complex Networks 2, 637-653, 2014

[252] J. Menche, A. Sharma, M. Kitsak, S. D. Ghiassian, M. Vidal, J. Loscalzo, et al., *Uncovering disease-disease relationships through the incomplete interactome,* Science 347, 1527601, 2015

[253] J. Li, P. Luo, C. Wu, *A new network node similarity measure method and its applications,* 2014 arxiv

[254] H. N. Phillips, A. Blenkmann, L. Hughes, T. A. Bekinschtein, J. J. Rowe, *Hierarchical organization of frontotemporal networks for the prediction of stimuli across multiple dimansions,* J Neurosci 35, 9255-9264, 2015

[255] A. Arenas, A. Díaz-Guilera, C. J. Pérez-Vicente, *Synchronization Reveals Topological Scales in Complex Networks,* Physical Review Letters 96(11), 114102, 2006

[256] M. Mitrović, B. Tadić, *Spectral and dynamical properties in classes of sparse networks with mesoscopic inhomogeneities,* Phys Rev E 80(2), 026123, 2009

[257] J. E. S. Socolar, *Nonlinear Dynamical Systems* in book Complex Systems Science in Biomedicine, Springer, Biomedical engineering, 115-140, 2006

[258] J.A. Sanders F. Verhulst J. Murdock *Averaging Methods in Nonlinear Dynamical Systems* Springer, Applied Mathematical Sciences 59, 2007

[259] D. D. Nolte, *The tangled tale of phase space* Physics Today. 63 (4): 3338, 2010

[260] P. Walters, *An Introduction to Ergodic Theory,* Springer, 1982

[261] D. Kerr, H. Li, *Ergodic Theory,* Springer, 2017

[262] E. Ott, *Chaos in dynamical systems,* Cambridge: Cambridge Univ. Press, 1993

[263] N. Packard, J. Crutchfield, D. Farmer, and R. Shaw, *Geometry from a time series,* Phys. Rev. Lett. 45, 712, 1980

[264] F. Takens *Detecting strange attractors in turbulence,* In: Rand D., Young LS. (eds) Dynamical Systems and Turbulence, Warwick 1980, Lecture Notes in Mathematics, vol 898. Springer, 1981

[265] H. Kantz, T. Schreiber *Nonlinear Time Series Analysis* Cambridge: Cambridge Univ. Press, 2004

[266] J.P. Eckman, S. Oliffson Kamphorst, D. Ruelle: *Recurrence Plots of Dynamical Systems,* Europhysics Letters 4 , 973-977, 1987

[267] T. D. Sauer *Attractor reconstruction* Scholarpedia, 1(10):1727, 2006

[268] H. Poincaré *Sur le probléme des trois corps et les équations de la dynamique* Acta Math. 13: 1270, 1890

[269] M. Koebbe, G. Mayer-Kress, *Use of recurrence plots in the analysis of time-series data,* Proceedings of SFI Studies in the Science of Complexity, vol. XXI, Redwood City, Addison-Wesley, Reading, MA, pp. 361378, 1992.

[270] J.P. Zbilut, C.L. Webber Jr., *Embeddings and delays as derived from quantification of recurrence plots,* Phys. Lett. A 171 (34) 199203, 1992

[271] J.P. Zbilut, J.M. Zaldvar-Comenges, F. Strozzi, *Recurrence quantification based Liapunov exponents for monitoring divergence in experimental data,* Phys. Lett. A 297 (34) 173181, 2002

[272] J. Garlanda, E. Bradleya,b, J. D. Meissc, *Exploring the Topology of Dynamical Reconstructions* Physica D: Nonlinear Phenomena 334, 1, 49-59, 2016

[273] S. Maletić, Y. Zhao, M. Rajković *Persistent topological features of dynamical systems* Chaos 26(5), 172, 2015

[274] E. J. Doedel, B. Krauskopf, H. M. Osinga, *Global organization of phase space in the transition to chaos in the Lorenz system* Nonlinearity, 28, 11, 2015

[275] J.L. Casti. *Polyhedral dynamics - II: Geometrical structure as a basis for decision making in complex systems,* IIASA Report, Laxenburg, Austria, 1976

[276] Ch. Renner, J. Peinke, R.Friedrich, O. Chanal, B. Chabaud, *Universality of small scale turbulence,* Phys. Rev. Lett. 89, 124502, 2002

Biografija

Miroslav Anđelković je rođen 27.05.1987. godine u Kraljevu. Osnovnu školu završio je u Trsteniku. Srednju školu, matematičku gimnaziju, završio je u Kraljevu. Upisao je osnovne akademske studije 2006/2007 godine na fizičkom fakultetu, Univerzitet u Beogradu, na smeru teorijska i eksperimentalna fizika. Osnovne studije završio je 2010/2011 godine, sa prosečnom ocenom 8.76. Upisao je master studije 2011/2012 godine i završio sa prosečnom ocenom 10.00. Master rad pod naslovom "O uticaju prenosa informacije i topologije mreže agenata na model dvostruke kontinualne aukcije finansijskog tržišta" odbranio je sa ocenom 10 pod mentorstvom dr Vladimira Miljkovića. Doktorske studije je upisao 2013/2014 godine na smeru fizika kondenzovane materije i statistička fizika, pod mentorstvom dr Milana Rajkovića. Od 2015. godine zaposlen je kao istrživač-pripravnik, a od 2018. godine kao istraživač-saradnik u Institutu za nuklearne nauke "Vinča".

Изјава о ауторству

Име и презиме аутора Мирослав Анђелковић

Број индекса 8012/2013

Изјављујем

да је докторска дисертација под насловом Algebraic topology of complex networks and topological aspects of nonlinear dynamical systems (Алгебарска топологија комплексних мрежа и тополошки аспекти нелинеарних динамичких система)

- резултат сопственог истраживачког рада;
- да дисертација у целини ни у деловима није била предложена за стицање друге дипломе према студијским програмима других високошколских установа;
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио/ла интелектуалну својину других лица.

Потпис аутора

У Београду, 8.3.2019.

Мирослав Анђелковић

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора Мирослав Анђелковић

Број индекса 8012/2013

Студијски програм Физика кондензоване материје и статистичка физика

Наслов рада Algebraic topology of complex networks and topological aspects of nonlinear dynamical systems (Алгебарска топологија комплексних мрежа и тополошки аспекти нелинеарних динамичких система)

Ментор др Милан Рајковић

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла ради похрањена у **Дигиталном репозиторијуму Универзитета у Београду.**

Дозвољавам да се објаве моји лични подаци везани за добијање академског назива доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

Потпис аутора

У Београду, 8.3.2019.

Мирослав Анђелковић

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

Algebraic topology of complex networks and topological aspects of nonlinear dynamical systems (Алгебарска топологија комплексних мрежа и тополошки аспекти нелинеарних динамичких система) која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигиталном репозиторијуму Универзитета у Београду и доступну у отвореном приступу могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство (CC BY)

2. Ауторство – некомерцијално (CC BY-NC)

3. Ауторство – некомерцијално – без прерада (CC BY-NC-ND)

4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)

5. Ауторство – без прерада (CC BY-ND)

6. Ауторство – делити под истим условима (CC BY-SA)

(Молимо да заокружите само једну од шест понуђених лиценци.
Кратак опис лиценци је саставни део ове изјаве).

Потпис аутора

У Београду, 8.3.2019.

Мирослав Антић

1. **Ауторство**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце, чак и у комерцијалне сврхе. Ово је најслободнија од свих лиценци.

2. **Ауторство – некомерцијално**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела.

3. **Ауторство – некомерцијално – без прерада**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела. У односу на све остале лиценце, овом лиценцом се ограничава највећи обим права коришћења дела.

4. **Ауторство – некомерцијално – делити под истим условима**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца не дозвољава комерцијалну употребу дела и прерада.

5. **Ауторство – без прерада**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца дозвољава комерцијалну употребу дела.

6. **Ауторство – делити под истим условима**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца дозвољава комерцијалну употребу дела и прерада. Слична је софтверским лиценцама, односно лиценцама отвореног кода.