UNIVERSITY OF BELGRADE
FACULTY OF PHYSICS

Slobodan V. Maletić

SIMPLICIAL COMPLEXES AND COMPLEX NETWORKS: THE INFLUENCE OF HIGHER-ORDER (SUB)STRUCTURES ON NETWORK PROPERTIES

Doctoral Dissertation

Belgrade, 2013

UNIVERZITET U BEOGRADU

FIZIČKI FAKULTET

Slobodan V. Maletić

SIMPLICIJALNI KOMPLEKSI I KOMPLEKSNE MREŽE: UTICAJ (POD)STRUKTURA VIŠEG REDA NA KARAKTERISTIKE MREŽE

doktorska disertacija

Beograd, 2013

Podaci o mentoru i članovima komisije:

Mentor:

dr Milan Rajković
naučni savetnik Instituta za nuklearne nauke "Vinča"

Članovi komisije:

dr Milan Knežević
redovni profesor Fizičkog fakulteta Univerziteta u Beogradu

dr Sunčica Elezović-Hadžić
redovni profesor Fizičkog fakulteta Univerziteta u Beogradu

dr Vladimir Miljković
docent Fizičkog fakulteta Univerziteta u Beogradu

Datum odbrane:

Acknowledgments

"Alas, a few years ago, I should have said 'my universe': but now my mind has been opened to higher views of things."[1]

Writing the dissertation has come as a summation and arranging of the results, views and concepts which emerged during the last few years of my research. Nevertheless, such research could not be possible without a guidance and encouragement of my mentor dr Milan Rajković. I express gratitude to Milan for learning me the various skills of academic research, like how to surf through the various scientific disciplines which helped me to realize how the ideas and concepts can overlap and complement each other. With this approach I finally understood the true beauty and challenges of the complex systems research. In finding the adequate solution of some problem he thought that it is sometimes necessary to pose a question "Why not?", and that the main-stream approaches have to be neglected. I also thank him for going through my publication drafts, and correcting my presentations.

I wish to thank the members of my dissertation committee: prof. dr Milan Knežević, prof. dr Sunčica Elezović-Hadžić, and dr Vladimir Miljković, 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. Specially, I am thankful to my colleague and friend Nikola Mirkov: for mutual understanding of the hard times of the beginning-researcher was comforting, and for discussions about everything, from conspiracy theory to topology.

The process of doctoral dissertation research, as well as writing it, is rather stressful, but I went through this process much easier with training which gave me the rowing coach Vladan Milosavljević from the Red Star Rowing Club. Thanks to him, the rowing philosophy became the way of access to the life's struggles.

I am grateful to my parents Veljko and Mirjana, and brother Spasoja, for all their "investments" in me: I would not become the person I am today without them. Thank you for everything.

Last, but the most important, I dedicate this dissertation to my wife Dijana. She is the pillar of my life, and somehow, she had tolerance and patience for all my doubts and discouragements. To say "thank you" is too weak.

[1] Citation taken from *Flatland: A Romance in Many Dimensions* by E. A. Abbott.

Title:

SIMPLICIAL COMPLEXES AND COMPLEX NETWORKS:
the influence of higher-order (sub)structures on network properties

Abstract

In modern theoretical physics (quantum gravity, computational electromagnetism, gauge theories, elasticity, to name a few) simplicial complexes have become an important objects due to their computational convenience and power of algebraic topological concepts. On the other hand, physics (and mathematics) of complex systems formed by the large number of elements interacting through pairwise interactions in highly irregular manner, is the most commonly restricted to concepts and methods of the graph theory. Such systems are called complex networks and notions of graph and complex network are used interchangeably. The achievements of the complex networks research are important for modern world and largely reshape our notion of a large class of complex phenomena, primarily because seemingly random and disorganized phenomena display meaningful structure and organization. The same stands also for the aggregations of complex network's elements into communities (modules or clusters), which as a major drawback has that they are restricted to the collections of pairwise interactions.

In this thesis to the notions of structure and substructure of complex systems, exemplified by complex networks, are given a new meaning through the changing the notion of community, by defining a simplicial community. Unlike the common notion of community, simplicial community is characterized by higher-order aggregations of complex network's elements. Namely, starting from typical properties of complex systems it was shown that the natural substructure of complex networks emerges like the aggregations of a multidimensional simplices. It was further shown that simplicial complexes may be constructed from complex networks in several different ways, indicating the possible different hidden organizational patterns leading to the final structure of complex network and which are responsible for the network properties. In this thesis two simplicial complexes obtained from complex networks are studied: the neighborhood and the clique complex.

Relying on the combinatorial algebraic topology concepts a unified mathematical framework for the study of their properties is proposed. The topological quantities, like structure vectors, Betti numbers, combinatorial Laplacian operator are calculated for diverse models real-world networks. Properties of spectra of combinatorial

Laplacian operator of simplicial complexes are explored, and the necessity of higher order spectral analysis is discussed and compared with results for ordinary graphs. The relationship of properties resulting from combinatorial Laplacian spectra with connectivity properties stored in the Q-vector is analyzed and discussed. The basic statistical features of complex networks are preserved by algebraic topological quantities of simplicial complexes, indicating possible presence of the so far unknown generic mechanisms in the complex networks formation. The spectral entropy is proposed as a measure of complexity which is determined by the eigenvalues of combinatorial Laplacian. All results support the necessity of developing a novel research field, called statistical mechanics of simplicial complexes as a unifying theory of the complex systems represented by simplicial complexes.

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

Scientific field:
physics

Specific scientific field:
statistical physics

UDC number: 533.9(043.3)

Naslov:

SIMPLICIJALNI KOMPLEKSI I KOMPLEKSNE MREŽE: uticaj (pod)struktura višeg reda na karakteristike mreže

Rezime

U savremenoj teorijskoj fizici (na primer, kvantnoj gravitaciji, računskom elektromagnetizmu, gejdž teoriji, elastičnosti) simplicijalni kompleksi su postali važni objekti zbog njihove računske pogodnosti i moći koncepata algebarske topologije. Sa druge strane, fizika (i matematika) kompleksnih sistema formiranih od velikog broja elemenata koji interaguju parnim interakcijama na izrazito neregularan način, najčešće je ograničena na koncepte i metode teorije grafova. Takvi sistemi se nazivaju kompleksne mreže i pojmovi graf i kompleksna mreža se poistovećuju. Doprinosi istraživanja kompleksnih mreža su važni za savremeni svet i umnogome preoblikuju naše poimanje velike klase kompleksnih fenomena, pre svega zbog toga što naizgled slučajni i neuređeni fenomeni pokazuju smislenu strukturu i organizaciju. Isto važi i za agregacije elemenata kompleksne mreže u zajednice (module ili klastere), koje kao najveći nedostatak imaju osobinu da su ograničene na kolekcije parnih interakcija.

U ovoj tezi pojmovima strukture i podstrukture kompleksnog sistema, kroz primer kompleksne mreže, dato je novo značenje menjanjem pojma zajednice, definisanjem simplicijalne zajednice. Za razliku od uobičajenog pojma zajednice, simplicijalna zajednica je karakterisana sa agregacijama višeg reda elemenata mreže. Naime, pošavši od tipičnih osobina kompleksnih sistema pokazano je da se kao prirodna podstruktura kompleksne mreže pojavljuju agregacije multidimenzionalnih simpleksa. Pokazano je, dalje, da se simplicijalni kompleksi mogu iz kompleksnih mreža konstruisati na nekoliko različitih načina, ukazujući na postojanje različitih skrivenih organizacionih obrazaca koji vode do konačne strukture kompleksne mreže i koji su odgovorni za osobine mreže. U ovoj tezi su razmatrana dva simplicijalna kompleksa dobijena iz kompleksne mreže: kompleks susedstva i klika kompleks.

Oslanjajući se na koncepte kombinatorijalne algebarske topologije predložen je objedinjeni matematički okvir za proučavanje njihovih osobina. Topološke veličine, kao što su strukturni vektori, Betti brojevi, operator kombinatorni laplasijan, računate su za različite modele realnih mreža. Ispitivane su osobine spektra operatora kombinatorni laplasijan simplicijalnog kompleksa, i razmatrana je neophodnost spektralne analize višeg reda koja je poređena sa rezultatima za obične grafove. Analizirana je i razmatrana veza osobina dobijenih iz spektra kombinatorijalnog laplasijana sa osobinama povezanosti sadržanih u Q-vektoru. Osnovne statističke osobine kompleksnih mreža su očuvane kod simplicijalnih kompleksa kroz veličine algebarske topologije, ukazujući na moguće postojanje do sada nepoznatih generičkih mehanizama u formiranju kompleksne mreže. Kao mera kompleksnosti predložena je spektralna entropija koja je definisana preko svojstvenih vrednosti kombinatorijalnog laplasijana. Svi rezultati podržavaju neophodnost razvoja novog polja istraživanja, nazvanog statistička mehanika simplicijalnih kompleksa, kao objedinjujuće teorije kompleksnih sistema predstavljenih kao simplicijalni kompleksi.

Ključne reči:
statistička mehanika, kompleksni sistemi, graf, kompleksne mreže, kombinatorna algebarska topologija, simplicijalni kompleksi, topološka invarijanta, kombinatorni laplasijan, entropija

Naučna oblast:
fizika

Uža naučna oblast:
statistička fizika

UDK broj: 533.9(043.3)

Contents

Chapter 1

Introduction

Complexity is a very intriguing concept which attracts attention and hence the research interest of scientists from different disciplines. Some kind of common purpose unified their efforts in formulation of a unique, up to some specific differences, theory of complex systems, which could be applied to diverse systems appearing in physical, social, biological, technological, informational, and many other phenomena. This approach is rather familiar in physics, particulary in the field of critical phenomena where the concept of "universality" [1] means that physical systems formed by the different elements will have the same behavior near critical point if they have the same numerical value of the so called critical exponents. Hence, it is not coincidence that the statistical physics had the largest influence on the sudden burst of important results, which reshaped our image about complex systems, and especially about complex networks.

However, if we want to begin the formulation of complex systems theory by introducing the definitions upon which we can build further theory, and primarily, definition of what complex system is, we face the problems: there is not a unique and widely accepted definition of a complex system [2]. This seemingly essential obstacle does not prevent researchers to adapt concepts and methods from their disciplines on a specific complex system, and as a result give informative and practical description of their behavior. Nevertheless, one large subset of complex systems singles out due to the specific mathematical framework within which problems related to those complex systems can be tackled. The characteristic of those systems is that they are formed by the large number of elements which interact among themselves through pairwise interactions. The easiest way to represent mathematically such complex system is by a graph, associating the elements with the nodes or vertices, and their interactions with the links or edges of a graph forming a complex network [3]. Although the underlying assumptions of research that led to the results in this thesis did not rely on any rigorous definition of a complex system, we have accepted some qualitative and most common properties which characterize complex systems in general. Based on these properties we have developed a characterization of complex system by a suitable mathematical framework.

In general, the line of reasoning was the following: complex networks are essentially a complex systems, and as such they should obey qualitative properties of complex systems. First property is that complex systems, despite their great irregularity, display some sort of organization which can be termed as an "organized complexity" [4], and we assume that this is true for complex networks. The last assumption is proved in numerous research articles through the analysis of complex networks by representing them as a graphs (for example, [5], [6], [7]). Second property is that any collection of elements of a complex system which are related by certain rules displays a qualitatively different behavior than the mere sum of those elements [8], and this has to be captured in mathematically rigorous fashion. Here is a very simple example: hot coffee is a collection of, say, four ingredients coffee powder, water, milk, and sugar, and knowing the flavor of each ingredient can not help us a lot in knowing the flavor of their mix in the hot coffee. Moreover, the same is true for any collection of three ingredients, or for any collection of two ingredients. Hence, such phenomena have to be mapped onto a suitable mathematical framework which captures these properties. This second property has to be further explained in the context of complex networks. Consider, say four people seated at the table and engaged in a group conversation so that each person may hear everyone else. A complete graph consisting of four nodes would be immediately assumed as an adequate mathematical framework for this relationship. Consider now the situation in which each person may only whisper in the ear of another person. Again, the only applicable graph model is the complete four node graph although the situation is completely different. The model is not adequate since the conversation of the group as an entity is not captured by the (1-dimensional) graph model which can not make distinction between a four-person *group* conversation and a set of six pairwise conversations. Hence, the behavior of a complex network can not be anticipated by knowing only the pairwise relationships between its elements, but it is to the large extent influenced by the structure built by those elements and their relations. The network of streets underlies the traffic carried by vehicles, and it is often hard to predict occurrence of, for example, traffic jams. Nevertheless, knowledge of the structure, or connectivity, or the hidden organizational patterns of the streets can be very informative in the prediction of possible jams. Of course, following the same reasoning, the similar problems can occur in the cases like the Internet (as a physical connections between computers) or the Power Grids or the network of neurons connected by synapses. By eliminating or adding new elements and/or connections, the structure may change, and it further affects the traffic through the networks. Finally, the third property is that the complex system possess an intrinsic hidden hierarchical organization which is responsible for the appearance of the system as it is [9]. Again, we can consider an example in order to give a clarification of this assumption: take for example a large company which is divided into sectors which are further divided into subsectors, which are divided into subsubsectors, and so on. Furthermore, the building blocks of a complex system are arranged in an irregular yet meaningful way which is revealed through the hidden organizational patterns which appear on different hierarchical levels characterized by aggregations of complex system's elements.

These typical properties of complex systems can be easily captured by simplicial complexes, i.e., a set of connected polyhedra which build a higher dimensional discrete geometrical space. An example of the formation of the simplicial complex is presented in Figure 1.1 by gluing simplices (polyhedra).

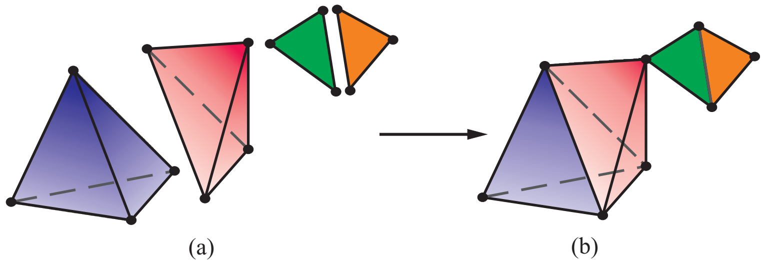


Figure 1.1: An example of gluing polyhedra (a) along common faces in forming the simplicial complex (b).

Therefore, we will place the problem of the hierarchical organization of intrinsic substructures of complex networks into suitable mathematical framework most adequate for the analysis of simplicial complexes, that is, combinatorial algebraic topology [10]. The idea of modeling complex systems by analyzing its elements represented by simplices is not a new one. Namely, Ron Atkin [11], [12], following the ideas of Dowker [13] of building a simplicial complex from the relations between the elements of two sets (or the same set), have introduced the method of Q-analysis [14]. Researchers have sporadically used the methods of Q-analysis for the analysis of specific systems, often with small number of elements. Such cases span from studying qualitative and quantitative structure of television program [15], analysis of the content of newspaper stories [16], social networks [17], [18], [19], [20], urban planning [21], [22], relationships among geological regions [23], distribution systems [24], decision making [25], diagnosis of failure in large systems [26], to mention a few.

From this short overview of the applications we can see the wide range of systems to which Q-analysis can be applied. Recently Atkin's methodology received a further development in the work of Barcelo and Laubenbacher [27], naming their theory an A-homotopy theory in honor of Atkin.

In modern theoretical physics simplicial complexes are recognized as important and convenient objects [28], [29], [30] due to their analytical and computational convenience. The language of modern physics is based on the calculus on manifolds which are discretized using simplicial complexes. Reversely, simplicial complexes can be used for the study of topological properties of a manifold obtained from experimental data [31]. Also, the use of simplicial complexes in discretization of exterior differential forms is extremely important. It is now widely recognized that geometry and topology are at the foundation of many physical theories such as general relativity [32], [33], electromagnetism [34], gauge theory [35], elasticity [36], to mention a few. For example, the development of simplicial quantum gravity [37] depends on the results of the Regge calculus [38], which, in turn, was developed by approximating smooth 4-dimensional manifold by rigid simplices. On the other hand, in the computational electromagnetism [39], [40], [41] Maxwell's equations can be directly expressed in terms of discrete differential forms which are defined as cochains on simplicial complexes. Generally, 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 so that 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 may be expressed in terms of exterior derivative of differential forms [42], many physical laws may be expressed in terms of differential forms. Discretization of differential forms using finite differences, for example, and using their coordinate values leads to numerical invalidation of some basic theorems (Stokes, for example) making traditional discretization methods futile. It turns out that proper discretization of differential forms that preserves all the fundamental differential properties is possible only on simplicial complexes [43]. As a curiosity let us mention that there are some attempts to formulate physical theories in a completely discrete fashion [44], [45], [46], [47], with emphasis on simplicial complexes. Considering the importance of simplicial complexes in the fundamental theories, we may consider simplicial complexes as universal tools for comprehensive and wide encompassing study of complex systems.

This "simplexication of sciences" and "complexication in sciences" naturally led to the ideas of relating concepts of statistical mechanics and algebraic topology [48],

[49], [50], [51], [52]. Namely, it turns out that essential characteristics of complex systems exemplified by complex networks, like the behavior of degree distribution is preserved when we calculate distributions of algebraic topological quantities of simplicial complex obtained from complex network. This further implies that we have to look for the additional rules which lead to the structure of complex network as it is, and relying on the substructures revealed through the algebraic topology concepts. Therefore, we have to turn our attention to those substructures and verify their validity within the structure of a complex network. So far the substructures in complex networks called communities (or modules, or clusters) [53] are considered as aggregations of pairwise interactions of network's elements based on certain rule, with the most general requirement that the density of the node connectivity within the community is higher then the connectivity among communities. The lack of a precise and widely accepted definition of what community is resulted in a diversity of definitions mostly as a consequence of a detection algorithm [54]. The detection of communities and their definitions includes node-based methods [55], link-based methods [56], [57], clique-based methods [58], [59], [60], to mention the most important ones. The link-based methods [56], [57] and clique-based methods [58], [59], [60] have been successful in capturing important properties of substructures related to the sharing of common nodes (and links) between different communities, that is the overlapping property. Recently, the attention has been shifted from the community detection to the so called structural group [61] defined as the aggregation of nodes with common structural properties.

In the approach introduced in this thesis the notion of community has an essentially different meaning. The definition of a simplicial community depends on the context determined by relationship between network elements including the process of the simplicial complex (i.e. the network). Introducing the general notion of simplicial community in the framework of combinatorial algebraic topology, we define and show that different substructures emerge in complex networks by mapping a complex networks into diverse discrete topological spaces. In some cases the approach is an extension and upgrade of the present community definitions, such as the case of k-clique communities [59].

Partitioning a simplicial complex, constructed from the initial complex network into mesoscale structures which capture relationships between simplices and their overlappings (called faces) and arranging them into levels defined by their corresponding dimensions [48], [49], we obtain a hierarchical description of simplicial complex connectivity. This representation is then compared, analyzed and interpreted by the spectrum of the (higher dimensional) combinatorial Laplacian of the corresponding simplicial complex [62], [63], which also has a hierarchical structure defined by the dimension of the simplicial complex [64]. Since the graph itself is a 1-dimensional simplicial complex the combinatorial Laplacian of the underlying graph is the $0^{\underline{th}}$ order combinatorial Laplacian of the corresponding simplicial (clique) complex. Hence, the combinatorial Laplacian of simplicial complex represents generalization of the graph Laplacian. Properties of the spectrum of the graph Laplacian are well known (see for example [65]), and they are important for the study of dynamic processes taking place on the complex network [66], among other applications. Although much less known, combinatorial Laplacian of a simplicial complexes represents an area of active mathematical research [62], [67], [68].

In order to measure "complexity" of complex networks several entropic measures has been proposed following information theoretical and statistical mechanics formalisms related to the structure and organization of complex networks. The Shannon entropy was derived under given structural constraints [69], [70], [71]. The von Neumann entropy (spectral entropy) proportional to the logarithm of eigenvalues of the density matrix ρ which quantifies the degree of mixedness of quantum states was defined in [72] and [73]. A density matrix ρ is said to be pure if $Tr(\rho^2) = 1$, or equivalently if $\rho = \rho^2$, and mixed otherwise. Laplacian matrix of a graph scaled by the sum of graph degrees is positive, symmetric with trace equal to 1 and hence equal to the density matrix. In this thesis we define the spectral entropy of the simplicial complex derived from the underlying graph as a vector whose components are spectral entropies defined for each dimension q of the simplicial complex. This entropy is proportional to the logarithm of the probability of the appearance of the eigenvalue of the q^{th} higher dimensional combinatorial Laplacian. With adequate normalization the q^{th} entropy quantifies the difference of the simplicial subcomplex at each dimension from the set of disconnected q-simplices [74].

Chapter 2

Properties of a simplicial complex

This chapter is devoted to definitions, concepts, and quantities related to simplicial complexes. Although this chapter is purely mathematical, rigorous definitions and proofs are avoided where unnecessary.

2.1 Definition of simplicial complex

Let us start with a finite set $B = \{b_1, b_2, \ldots, b_m\}$, whose elements we call *vertices*. A convex hull of $q+1$ elements $\{b_{\alpha_0}, b_{\alpha_1}, ..., b_{\alpha_q}\}$ of the set B is called a *q-dimensional simplex*, or just a *q-simplex* [75]. Geometrical realization of a q-simplex is through the polyhedra embedded in d-dimensional $\mathbb{R}^d$ space, where $q \leq d$, and the rigorous proof of this realization can be found in [75]. Hence, simplices can be understood as a higher-dimensional generalizations of a point, a line, a triangle, a pyramid, and so on. Figure 2.1 left, illustrates geometrical realization of q-simplices for various q, of the set $B = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14\}$ with subsets: $\{1, 2, 3, 4, 5\}$, $\{2, 3, 5, 6\}$, $\{6, 8\}$, $\{2, 7\}$, $\{7, 8, 9\}$, $\{8, 9, 10\}$ and $\{11, 12, 13, 14\}$.

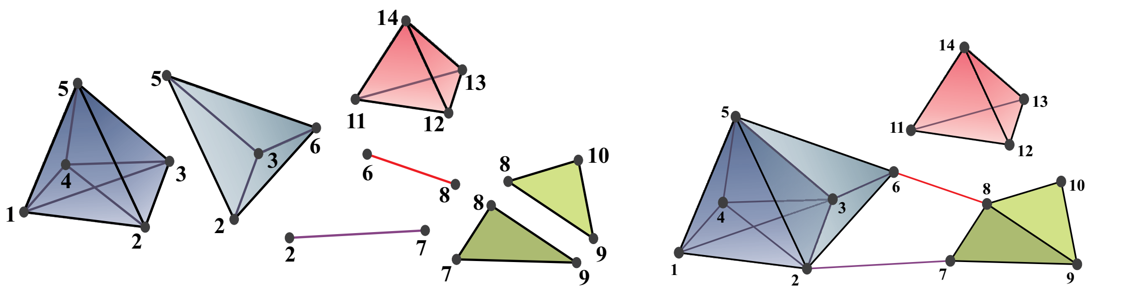


Figure 2.1: Formation of simplicial complex on the right from simplices on the left part of the picture.

In the following we will label a q-dimensional simplex as σ_q. A p-simplex σ_p is a *p-face* of a q-simplex σ_q, denoted by $\sigma_p \leq \sigma_q$, if every vertex of σ_p is also a vertex of σ_q. Therefore, if two simplices σ_q and σ_r share $p+1$ common vertices, then they share a p-face. From the Figure 2.1 left we can see that subset $\{2,3,5\}$ of a simplex $\{1,2,3,4,5\}$ is also a subset of a simplex $\{2,3,5,6\}$, meaning that these two simplices share a 2-face. Also, the definition of a simplex implies that a 2-face $\{2,3,5\}$ is also a simplex. Defined in this way, simplices are the maximal subsets (simplices), in the sense that they are not face of any other simplex.

Collection of simplices together with all their faces is called a *simplicial complex*. In more formal terms a simplicial complex K on a finite set $B = \{b_1, b_2, \ldots, b_m\}$ of vertices is a nonempty subset of the power set of B, such that K is closed under the formation of subsets [75]. The maximal dimension of a simplex in K determines the dimension of the whole simplicial complex, $D = dim(K)$. Figure 2.1 (right) illustrates how simplices on the left side form the 4-dimensional simplicial complex.

The above definition of simplicial complex is the abstract onelacking the meaning of vertex aggregations into subsets of B. Namely, for practical purposes when we are dealing with a concrete elements $b_1, b_2, \ldots, b_m$ we must have some rule according to which we aggregate elements into subsets which form simplices, and we must know what these simplices actually represent. Let us introduce a new set $A = \{a_1, a_2, \ldots, a_n\}$ and a binary relation λ, which together with the set $B = \{b_1, b_2, \ldots, b_m\}$ contribute to the formation of two simplicial complexes [13]. We will introduce these two simplicial complexes leaning on the concepts of Q-analysis developed by R. Atkin [14], [11], [76], [77], and further developed by J. Johnson [78], [79], [80]. The binary relation λ by some rule or property assigns to every element in A one or more elements in B, i.e., for every $a_i \in A$ there exists $b_j \in B$ such that $a_i \lambda b_j$. The set A and the relation λ determine the subset K of the power set of B and we label each element $\{b_{\alpha_0}, b_{\alpha_1}, ..., b_{\alpha_q}\} \in K$ $(q \leq m)$ by the element $a_i \in A$ for which $a_i \lambda b_{\alpha_0}$, $a_i \lambda b_{\alpha_1}$, ..., $a_i \lambda b_{\alpha_q}$. To distinguish the element a_i from the set A and its associated element from the set K due to the relation λ, the element of the set K will be labeled as $\sigma(a_i)$. Therefore, the notation $\sigma_q(a_i) = \langle b_{\alpha_0}, b_{\alpha_1}, b_{\alpha_2}, \ldots, b_{\alpha_q} \rangle$ [78] means that an element a_i of the set A is λ-related to q elements $\{b_{\alpha_0}, b_{\alpha_1}, b_{\alpha_2}, \ldots, b_{\alpha_q}\}$ of the set B. The elements of the set B are called *vertices*, whereas the elements of the set K are called *q-dimensional simplices* or just *q-simplices*. Further, an element a_i is λ-related to any subset of the set $\{b_{\alpha_0}, b_{\alpha_1}, b_{\alpha_2}, \ldots, b_{\alpha_q}\}$, and hence, every subset of $\{b_{\alpha_0}, b_{\alpha_1}, b_{\alpha_2}, \ldots, b_{\alpha_q}\}$ is also a simplex, meaning that any such subset is a face of simplex, due to the definition of q-face. Since each $a_i \in A$ identifies a q-simplex $\sigma_q(a_i)$ (for some q) together with all its faces, this collection of simplices is called a simplicial complex K, which we will denote $K_A(B, \lambda)$ [12].

To illustrate the construction of simplicial complex from two sets let us introduce a set $A = \{a, b, c, d, e, f, g\}$, together we previously introduced set $B = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14\}$, and suppose that the elements of the set A are λ-related to the elements of the set B. For example, the letters from the set A may correspond to the individuals the numerals from the set B may correspond to the diverse interests of each individual and the relation λ may correspond to the property "person a has an interest in 1". As another example, we may assume that the letters from the set A correspond to patients, whereas numerals from the set B correspond to diverse clinical symptoms, and the relation λ corresponds to the property "patient a has a symptom 1". Or, the letters from the set A may correspond to the city streets whereas the numerals from the set B may correspond to the diverse junctions and the relation λ may correspond to the property "street a contains a junction 1". As another example consider that letters from the set A correspond to the TV shows whereas the numerals from the set B correspond to the diverse subjects, covered by the show and the relation λ corresponds to the property "TV show a has a subject 1". In the context of social issues the letters from the set A may correspond to the social groups, the numerals from the set B may correspond to the diverse persons, and the relation λ may correspond to the property "social group a has as a member person 1". As a final example, the letters from the set A may correspond to the geological regions, whereas the numerals from the set B may correspond to the diverse rock types, and the relation λ may correspond to the property "geological region a has a rock type 1", and so on.

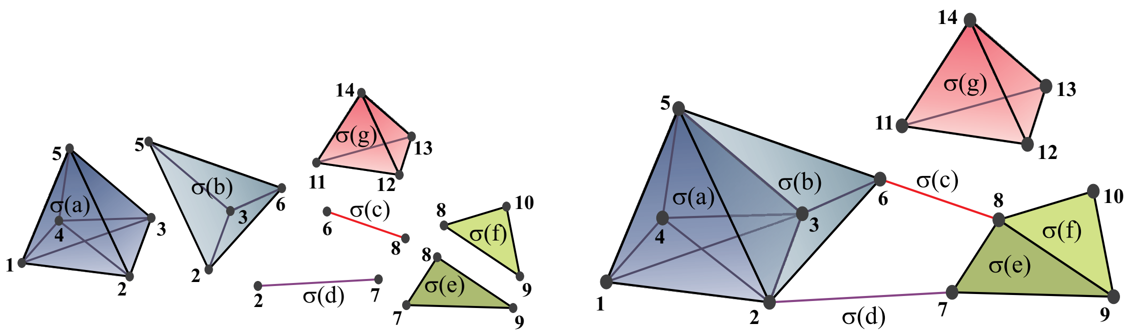


Figure 2.2: Formation of the simplicial complex from Figure 2.1, but now the simplices are labeled.

Figure 2.2 left illustrates polyhedral representation of simplices obtained by the elements from the set A, which are λ-related to the elements of the set B. For example, an element a is λ-related to the elements $\{1, 2, 3, 4, 5\}$. The obtained simplices are:

$\sigma(a) = \langle 1, 2, 3, 4, 5 \rangle$

$\sigma(b) = \langle 2, 3, 5, 6 \rangle$

$\sigma(c) = \langle 6, 8\rangle$
$\sigma(d) = \langle 2, 7\rangle$
$\sigma(e) = \langle 7, 8, 9\rangle$
$\sigma(f) = \langle 8, 9, 10\rangle$
$\sigma(g) = \langle 11, 12, 13, 14\rangle$.
Figure 2.2 illustrates the simplicial complex formed by "gluing" simplices along their shared faces. By simple inspection of left and right sides of Figures 2.1 and 2.2 we can see that simplicial complexes are the same with a slight difference: the simplices from Figure 2.2 are labeled and a meaning is attached to them. For example, consider a simplicial complex of streets (simplices) and junctions (vertices) in an urban area. Then from the above example, street a contains junctions 1, 2, 3, 4 and 5, whereas street b contains junctions 2, 3, 5 and 6, and these two streets share junctions 2, 3 and 5. Hence, it is easy to comprehend how simplicial complex from Figure 2.2 can capture the complicated relationships between streets through common junctions. Obviously, if we do not assign a street name with the corresponding junction, we would lose an important information.

Since the relation λ relates the elements of the set A with elements of the set B, there must be some relation which does the reverse, i.e., relates the elements of the set B with the elements of the set A. That role is taken by the *inverse relation* λ^{-1} [14], [78] of λ which relates the elements of the set B with the elements of the set A: $1\lambda^{-1}a$, $2\lambda^{-1}a$, $2\lambda^{-1}b$, $2\lambda^{-1}d$, $3\lambda^{-1}a$, $3\lambda^{-1}b$, and so on. Following the same procedure, we form a simplicial complex $K_B(A, \lambda^{-1})$ on the vertex set A defined by the relation λ^{-1}, represented in Figure 2.3. Note that the elements of sets A and B have changed their roles, and in the complex $K_B(A, \lambda^{-1})$ simplices are from the set B whereas the vertices are from the set A. Now we can generalize these results to the simplicial complex defined by two arbitrary sets $A = \{a_1, a_2, \ldots, a_n\}$ and $B = \{b_1, b_2, \ldots, b_m\}$ and relation λ. Then simplicial complex $K_B(A, \lambda^{-1})$ defined on the sets $A = \{a_1, a_2, \ldots, a_n\}$ and $B = \{b_1, b_2, \ldots, b_m\}$ by the inverse relation λ^{-1} of the relation λ is called *the conjugate complex* of the simplicial complex $K_A(B, \lambda)$ [14], [78]. In order to clarify the importance of the simplicial complex and its conjugate, let us consider an example where the elements of the set A are patients, and the elements of the set B are clinical symptoms. Then the simplicial complex represents a collection of patients sharing the symptoms, whereas its conjugate complex represents a collection of clinical symptoms sharing the patients which have them.

Finally we would like to emphasize, that the simplicial complex can be created on a single set, that is, following the above notation $A = B$, and, hence, $K_A(A, \lambda)$. In this case the simplicial complex and its conjugate complex are the same.

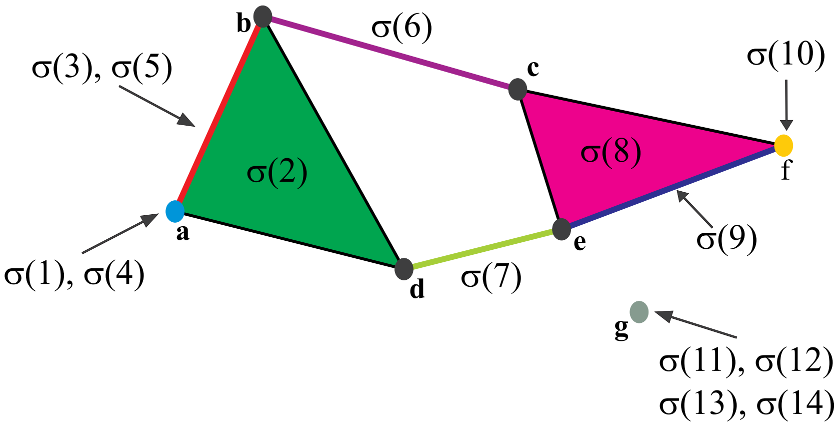


Figure 2.3: Conjugate complex of the simplicial complex from Figure 2.2.

The geometrical representation of the simplicial complex is not the practical way to represent the relation between two sets. The more practical representation is by the so called *incidence matrix* [14], [78] Λ. The rows of this matrix are associated with the simplices and columns are associated with the vertices, and the matrix entry $[\Lambda]_{ij}$ is equal to 1 if simplex $\sigma(i)$ contains a vertex j and otherwise it is equal to 0. Hence, for the above example rows correspond to the elements of the set A, columns correspond to the elements of the set B and a matrix element $[\Lambda]_{ij}$ is equal to 1 if an element $a_i \in A$ is λ-related to the element $b_j \in B$:

$$\begin{array}{cc} & \begin{array}{cccccccccccccc} \lambda & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \end{array} \\ \Lambda = & \begin{array}{c} a \\ b \\ c \\ d \\ e \\ f \\ g \end{array} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \end{array}$$

The matrix representation of the conjugate complex $K_B(A, \lambda^{-1})$ of the simplicial complex $K_A(B, \lambda)$ is the transpose matrix of Λ (Λ^T). The matrix that captures the relationships between simplices, and hence, the properties of simplicial complex is the so called *connectivity matrix* defined as:

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

where Λ is the incidence matrix, and Ω is matrix with all entries equal to 1. Rows and columns of the matrix Π are associated to the simplices, the diagonal elements represent the dimension of simplices, whereas the non-diagonal elements represent the dimensionality of faces which simplices share. By convention, the entry $[\Pi]_{ij} = -1$ $(i \neq j)$ means that two simplices do not share face. For the example of simplicial complex in Figure 2.1, the connectivity matrix has the following form:

$$
\Pi = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \\ g \end{array}
\begin{array}{c}
\begin{array}{ccccccc} a & b & c & d & e & f & g \end{array} \\
\begin{pmatrix}
4 & 2 & -1 & 0 & -1 & -1 & -1 \\
2 & 3 & 0 & -1 & -1 & -1 & -1 \\
-1 & 0 & 1 & -1 & 0 & 0 & -1 \\
0 & -1 & -1 & 1 & 0 & -1 & -1 \\
-1 & -1 & 0 & 0 & 2 & 1 & -1 \\
-1 & -1 & 0 & -1 & 1 & 2 & -1 \\
-1 & -1 & -1 & -1 & -1 & -1 & 3
\end{pmatrix}
\end{array}
$$

2.2 Chains of connectivity and structure vectors

So far we have introduced the dimension of the simplex and the relationship (or adjacency) between two simplices through the shared common 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 (see Figure 2.4), and hence, they are also $(q-1)$-, $(q-2)$-, ..., 1- and 0-near.

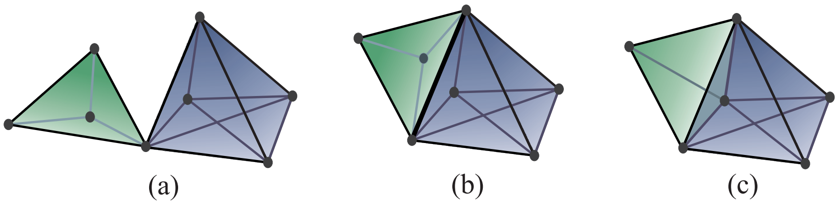


Figure 2.4: An example of q-nearness between simplices. Two simplices are: (a) 0-near, (b) 1-near, and (c) 2-near.

The collection of simplices in which any pair of simplices is connected by a sequence of simplices where a pair of successive simplices is q-near is called the *q-connected component*. More formally, two simplices σ and ρ are $q-$ *connected*

[11] if there is a sequence of simplices $\sigma, \sigma(1), \sigma(2), ..., \sigma(n), \rho$, such that any two consecutive ones share at least a q-face. As an example of q-connectivity see Figure 2.5. Note that if two simplices σ_p and σ_r are q-connected, they are also $(q-1)-$, $(q-2)-$, ..., 1, 0-connected in K.

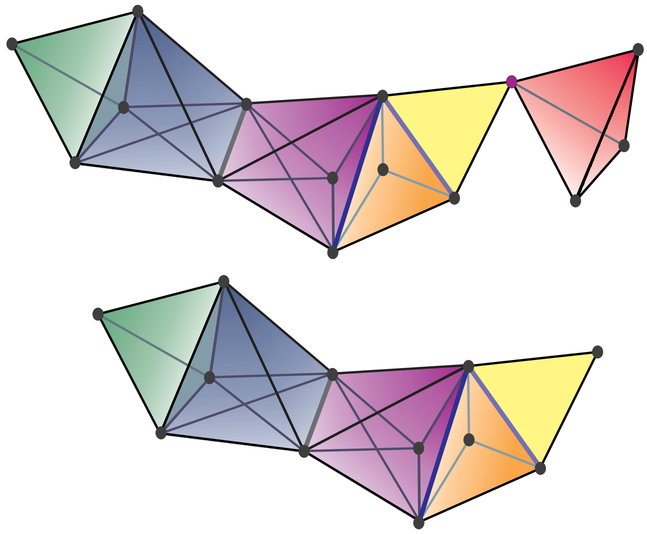


Figure 2.5: An example of q-connectedness: (a) green and red simplices are 0-connected, and (b) green and yellow simplices are 1-connected.

The q-connectivity between simplices induces an equivalence relation on simplices of a complex K, since it is reflexive, symmetric, and transitive. This equivalence relation will be denoted by γ_q so that

$$(\sigma(i), \sigma(j)) \in \gamma_q \text{ if and only if } \sigma(i) \text{ is q-connected to } \sigma(j).$$

Let K_q be the set of simplices in K with dimension greater than or equal to q. Then γ_q partitions K_q into equivalence classes of q-connected simplices. These equivalence classes are members of the quotient set K_q/γ_q and they are called the *q-connected components* of K. Every simplex in a q-component is q-connected to every other simplex in that component, but no simplex in one q-component is q-connected to any simplex on a distinct q-connected component. The cardinality of K_q/γ_q is denoted Q_q and is the number of distinct q-connected components in K. The value Q_q is the q^{th} entry of the so called *Q-vector* [74] (*first structure vector* [78]), an integer vector with the length $dim(K) + 1$. The values of the Q-vector entries are usually written starting from the number of connected components for the largest dimension in descending order, i.e.:

$$\mathbf{Q} = \{Q_{dim(K)} \quad Q_{dim(K)-1} \quad \dots \quad Q_1 \quad Q_0\}.$$

An example illustrating the partitioning of the simplicial complex into q-connectivity classes and Q-vector for the example from Figure 2.1 is presented in Figure 2.6, with Q-vector entries:

$$\mathbf{Q} = \{1 \quad 3 \quad 4 \quad 5 \quad 2\}.$$

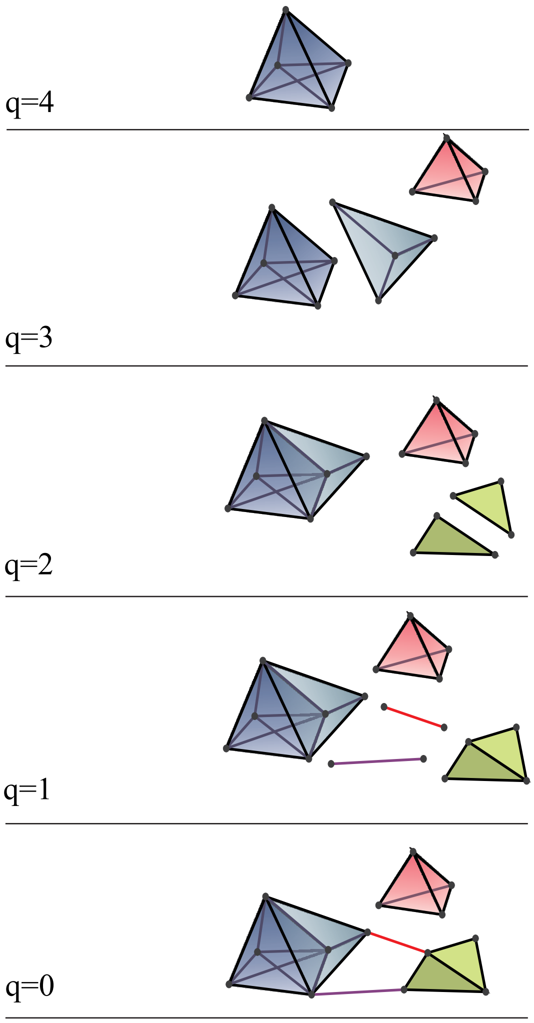


Figure 2.6: Q-vector of simplicial complex from Figure 2.1.

Another vector-based quantity is the so called *second structure vector* [78]

$$\mathbf{n} = \{n_{dim(K)} \quad n_{dim(K)-1} \quad \dots \quad n_1 \quad n_0\},$$

which is an integer vector with $dim(K)+1$ components, like the Q-vector, and the q-th entry, n_q, is equal to the number of simplices with dimension larger or equal to q, that is, it is equal to the number of simplices at the q-level. For the example of simplicial complex from the Figure 2.1, whose Q-vector components are presented at the Figure 2.6, the second structure vector goes is:

$$\mathbf{n} = \{1 \quad 3 \quad 5 \quad 7 \quad 7\}.$$

Finally, the entries of the third structure vector $\overline{Q}_q$ are defined in the following way [81]:

$$\overline{Q}_q = 1 - \frac{Q_q}{f_q}\,, \tag{2.2}$$

where Q_q is q-th entry of the first structure vector, and f_q is q-th entry of the second structure vector. The third structure vector measures the degree of connectedness on each q-level, or in other words, it measures the number q-connected components per number of simplices.

2.3 Homology groups and Betti numbers

So far, structural properties of simplicial complex have been explored only through the connectivity of simplices deduced from the relationship between two sets. We now concentrate on the topological properties of simplicial complex and take into account a key property of the simplicial complex definition - that the power set of the set on which simplicial complex is defined is closed under the formation of subsets. In other words, every subsimplex, that is the face, is also a simplex in simplicial complex. Hence, when we say "q-simplices", we mean "all maximal q-dimensional simplices and all q-dimensional faces".

Let us start again with a finite vertex set $B = \{b_1, b_2, \ldots, b_m\}$. An arbitrary ordering of vertices $\{b_{\alpha_0}, b_{\alpha_1}, ..., b_{\alpha_q}\}$ of a simplex defines an oriented q-simplex which we denote $[b_{\alpha_0}, b_{\alpha_1}, ..., b_{\alpha_q}]$, and we say that simplicial complex K is oriented if all simplices in K are oriented. Note that an unoriented simplex was denoted as $\langle b_{\alpha_0}, b_{\alpha_1}, ..., b_{\alpha_q}\rangle$. An example of oriented 0-, 1-, 2-, and 3-simplices is illustrated in Figure 2.7, and by convention 0-simplex does not have an orientation.

Let $C_q(K)$ (for each $q \geq 0$) be the vector space whose bases is the set of all q-simplices of an oriented simplicial complex K, and the elements are the linear combinations of bases vectors, called *chains*. Accordingly $C_q(K)$ is called a *chain group* [75] (the term chain group is accepted for traditional reasons, regardless of the vector space properties of $C_q(K)$, nevertheless $C_q(K)$ is still a group). The dimension of $C_q(K)$ is equal to the q^{th} entry of an important topological invariant, the f-vector, $f = (f_0, f_1, ...f_q...f_n)$. In this expression f_q is equal to the number of q-dimensional simplices of the simplicial complex K, i.e. f_0 represents the number of vertices, f_1 number of edges and so on. For q larger than the dimension of K,

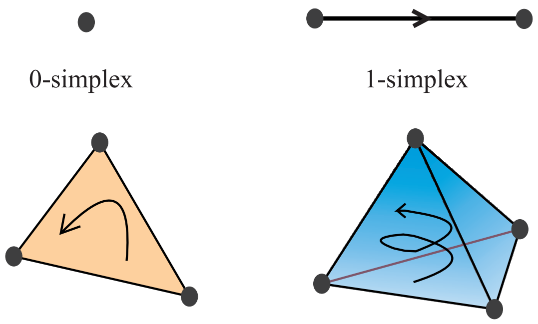


Figure 2.7: Examples of orientation of 0-, 1-, 2-, and 3-simplex.

vector space $C_q(K)$ is trivial and equals to 0. For a set of vector spaces $C_q(K)$ with $0 \leq q \leq dim(K)$ the linear transformation $\partial_q : C_q(K) \rightarrow C_{q-1}(K)$ called *boundary operator* acts on the bases vectors $\langle v_{\alpha_0}, v_{\alpha_1}, ..., v_{\alpha_q} \rangle$ in the following way [75]

$$\partial_q \langle v_{\alpha_0}, v_{\alpha_1}, ..., v_{\alpha_q} \rangle = \sum_{i=0}^{q} (-1)^i \langle v_{\alpha_0}, ..., v_{\alpha_{i-1}}, v_{\alpha_{i+1}}, ..., v_{\alpha_q} \rangle \, .$$

An example of the action of the boundary operator on a 3-simplex and its subsimplices from Figure 2.1 is illustrated in Figure 2.8.

Taking a sequence of chain groups $C_q(K)$ connected through the boundary operators ∂_q the so-called *chain complex* is defined in the following way

$$\emptyset \rightarrow C_q \stackrel{\partial_q}{\rightarrow} C_{q-1} \stackrel{\partial_{q-1}}{\rightarrow} .... \rightarrow C_1 \stackrel{\partial_1}{\rightarrow} C_0 \stackrel{\partial_0}{\rightarrow} \emptyset,$$

with $\partial_q \partial_{q+1} = \emptyset$ for all q. The kernel of ∂_q is the set of q-chains with empty boundary while a q -cycle, denoted by Z_q, is a q-chain in the kernel of ∂_q. The image of ∂_q is the set of $(q-1)$-chains which are boundaries of q-chains with a q-boundary, denoted by B_q, being a q -chain in the image of ∂_{q+1}. The q^{th} *homology group* [75] is defined as

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

The rank of the q^{th} homology group $\beta_q = rank(H_q)$ or $\beta_q = dim(H_q)$ is topological invariant called the q^{th} *Betti number* and is equal to the number of q-dimensional holes in simplicial complex. Since it is a topological invariant it is used to distinguish topological spaces one from another. For example, the value of β_0 is the number of connected components of simplicial complex, β_1 is the number of tunnels, β_2 is the number of voids, etc. In the simplicial complex presented in Figure 2.1 right we can see that there are two connected components, hence $\beta_0 = 2$, and one 1-dimensional hole bounded by 1-dimensional simplices $[2, 6]$, $[2, 7]$, $[7, 8]$, and $[7, 8]$,

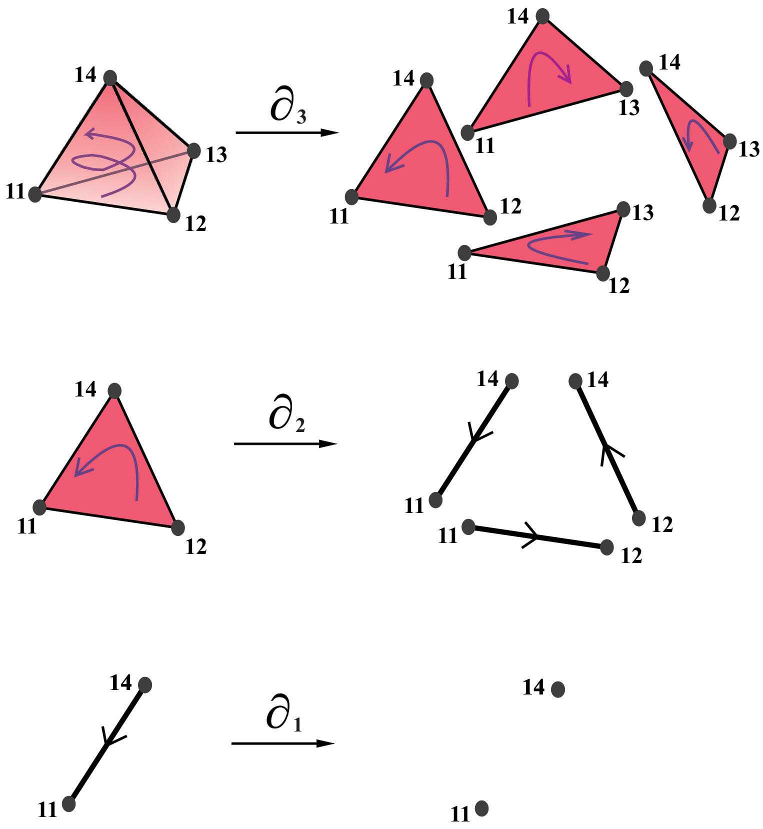


Figure 2.8: The action of boundary operator on the 3-, 2-, and 1-simplices.

hence $\beta_1 = 1$. As R. Atkin pointed out [12] the zeroth Betti number is equal to Q_0, nevertheless, the higher-order Betti numbers are not equal to the higher order Q-vector entries. Therefore, the analysis presented in the previous section (Q-vector) gives a generalization of the zeroth order Betti number, although different from the homology theory. The values of Betti numbers of simplicial complex from Figure 2.2 are preserved for its conjugate complex (see Figure 2.3), and Dowker [13] have proved that the homology groups of simplicial complex and its conjugate complex are isomorphic.

Each boundary operator ∂_q has its matrix representation B_q with respect to bases of vector spaces $C_q(K)$ and $C_{q-1}(K)$, with rows associated with the number of $(q-1)$-simplices and the columns associated with the number of q-simplices. To each boundary operator ∂_q corresponds an adjoint operator $\partial_q^* : C_{q-1}(K) \rightarrow C_q(K)$ with the associated matrix representation equal to the transpose of matrix representation of boundary operator ∂_q, that is B_q^T. It is important to mention that the q^{th} adjoint boundary operator is in fact the same as the q^{th} coboundary operator $\delta_q : C^{q-1}(K) \rightarrow C^q(K)$ [63], whereas, their matrix representations coincide when proper scalar products are chosen for the definition of ∂_q^*.

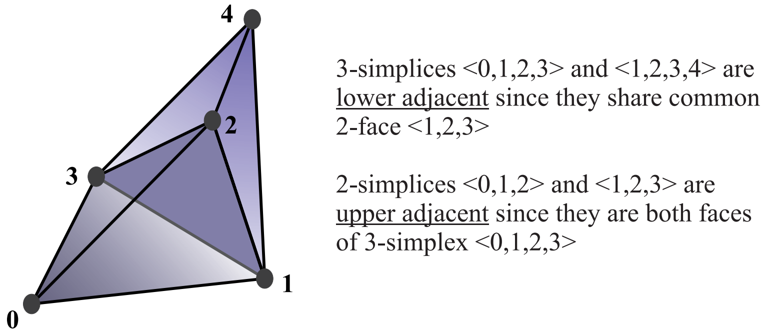


Figure 2.9: An example of adjacency between two 3-simplices.

2.4 Combinatorial Laplacian

Since we have defined an oriented simplicial complex and boundary operator, we are prepared to introduce new concepts. For two q-simplices $\sigma(i)$ and $\sigma(j)$ of an oriented simplicial complex K we say that they are *upper adjacent*, denoted $\sigma(i) \sim_U \sigma(j)$, if they are both faces of some $(q+1)$-simplex in K. The *upper degree* of a q-simplex σ in K, denoted $deg_U(\sigma)$, is the number of $(q+1)$-simplices in K of which σ is a face. If oriented q-simplices $\sigma(i)$ and $\sigma(j)$ are upper adjacent and have a common $(q+1)$-simplex τ, we say that $\sigma(i)$ and $\sigma(j)$ are *similarly oriented* if orientations of $\sigma(i)$ and $\sigma(j)$ agree with the ones induced by τ. For two q-simplices $\sigma(i)$ and $\sigma(j)$ of an oriented simplicial complex K we say that they are *lower adjacent*, denoted $\sigma(i) \sim_L \sigma(j)$, if they have common $(q-1)$-face (that is $(q-1)$-simplex as a face). Hence, the *lower degree* ($deg_L(\sigma)$) of a q-simplex is defined as the number of $(q-1)$-faces in σ, which is always equal to $q+1$. Example of upper/lower adjacency is illustrated in Figure 2.9.

Defining the boundary operator and its adjoint we have provided necessary conditions for the definition of combinatorial Laplacian of simplicial complex. Namely, for a simplicial complex K and an integer $q \geq 0$, the q^{th} *combinatorial Laplacian* is linear operator (since the composition of linear maps is a linear map) defined as $L_q : C_q \to C_q$ and given by [62]

$$L_q = \partial_{q+1} \circ \partial_{q+1}^* + \partial_q^* \circ \partial_q \,.$$

A convenient notation to use is

$$L_q^{UP} = \partial_{q+1} \circ \partial_{q+1}^* \quad \text{and} \qquad L_q^{DN} = \partial_q^* \circ \partial_q,$$

where L_q^{UP} is referred to as the upper combinatorial Laplacian and L_q^{DN} is the down combinatorial Laplacian. Corresponding matrix representation relative to some ordering of the standard bases for C_q and C_{q-1} for the q^{th} *Laplacian matrix* of K 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 may use the following notation for convenience

$$\mathcal{L}_q^{UP} = B_{q+1}B_{q+1}^T \qquad \text{and} \qquad \mathcal{L}_q^{DN} = B_q^T B_q \,.$$

Clearly, graph represents a 1-dimensional simplicial complex since links (1-dim simplices) connect nodes (0-dimensional simplices) and the largest dimension of a simplex in the complex is 1. We now apply this fact to obtain the combinatorial Laplacian of a graph. The 0-dimensional combinatorial Laplacian of simplicial complex K is a linear map $L_0 : C_0(K) \to C_0(K)$, and since the maps ∂_0 and ∂_0^* are assumed to be zero maps, it follows that

$$L_0 = \partial_1 \circ \partial_1^* \,,$$

where the boundary operator $\partial_1 : C_1(K) \to C_0(K)$ maps edges to vertices. Since in matrix representation B_1 of boundary operator ∂_1 the rows are associated with edges and the columns are associated with vertices, it is obvious that the matrix B_1 is equal to the incidence matrix of an oriented graph. Therefore matrix representation of combinatorial Laplacian is $\mathcal{L}_0 = B_1 B_1^T$, and the matrix elements are

$$(\mathcal{L}_0)_{ij} = \begin{cases} deg(v_i), & \text{if } i = j \\ -1, & \text{if } v_i \sim v_j \\ 0 \,, & \text{otherwise} \end{cases} \tag{2.3}$$

where $deg(v_i)$ is vertex degree (that is number of neighbors of a vertex v_i) and the relation $v_i \sim v_j$ is the adjacency relation between vertices v_i and v_j, and is the same as upper adjacency $v_i {\sim_U} v_j$. Clearly, the entries of the 0-dimensional combinatorial Laplacian are the same as the graph Laplacian entries defined in the usual way via expression $L_{qraph} = D - A$, where diagonal entries of matrix D are equal to the vertex degrees ($D_{ii} = deg(v_i)$) and nondiagonal entries are zeros, and the entries of matrix A are $(A)_{ij} = 1$ if $v_i \sim v_j$, $(A)_{ij} = 0$ if vertices v_i and v_j are not neighbors, and $(A)_{ii} = 0$ (undirected, unweighted, without loops and multiple edges graph) [65].

For the general case let us assume that K is an oriented simplicial complex, q is an integer with $0 < q \leq dim(K)$, and let $\{\sigma^1, \sigma^2, ..., \sigma^n\}$ denote the q-simplices of complex K, then it is not difficult to deduce from $\mathcal{L}_q = \mathcal{L}_q^{UP} + \mathcal{L}_q^{DN}$ that

$$(\mathcal{L}_q)_{ij} = \begin{cases} deg_U(\sigma^i) + q + 1, & \text{if } i = j \\ 1, & \text{if } i \neq j \text{ and } \sigma^i \text{ and } \sigma^j \text{ are not upper adjacent but have} \\ & \text{a similar common lower simplex} \\ -1, & \text{if } i \neq j \text{ and } \sigma^i \text{ and } \sigma^j \text{ are not upper adjacent but have} \\ & \text{a dissimilar common lower simplex} \\ 0\ , & \text{if } i \neq j \text{ and } \sigma^i \text{ and } \sigma^j \text{ are upper adjacent or are not} \\ & \text{lower adjacent} \end{cases} \tag{2.4}$$

since $(\mathcal{L}_q^{UP})_{ii} = deg_U(\sigma^i)$ and $(\mathcal{L}_q^{DN})_{ii} = deg_L(\sigma^i)$. Detailed proof of the above expression is straightforward [63]. For later use it would be useful to notice that $(\mathcal{L}_q)_{ii} = deg_U(\sigma^i) + deg_L(\sigma^i) = deg_U(\sigma^i) + q + 1$ since every simplex of dimension $q > 0$ has exactly $q + 1$ $(q - 1)$-faces. Clearly, for $q = 0$ Laplacian matrix of general simplicial complex reduces to graph Laplacian.

Let us focus now on the eigenvalues and eigenvectors of q^{th} combinatorial Laplacian L_q. For an oriented simplicial complex K and an integer q with $0 \leq q \leq dim(K)$, the q^{th} *Laplacian spectrum* is denoted as $S(L_q(K))$. It represents set of eigenvalues of $L_q(K)$ together with their multiplicities and is independent on the choice of orientation of q-simplices in the complex K. Since the q^{th} Laplacian matrix is positive semidefinite, all its eigenvalues are nonnegative. The null space of $N(L_q(K))$ is the eigenspace of $L_q(K)$ and corresponds to the zero eigenvalues. The combinatorial Hodge theorem states that the q^{th} homology group $H_q(K)$ is isomorphic to the null space of q^{th} combinatorial Laplacian [82], that is

$$H_q(K) \cong N(L_q(K))\,,$$

for each integer q with $0 \leq q \leq dim(K)$. Therefore, the multiplicity of zero eigenvalues of q^{th} combinatorial Laplacian is equal to the number of the q-dimensional holes in a simplicial complex, i.e. a Betti number. This is a very useful expression providing a practical method for calculation of Betti numbers [83].

In the following we introduce some properties of the spectra of the q^{th} combinatorial Laplacian which will be useful for the analysis and interpretation of our results. If simplicial complex K consists of disconnected components which are themselves simplicial complexes $K_1, K_2, ..., K_n$, then the spectra of q^{th} combinatorial Laplacian $L_q(K)$ of K for each q with $0 \leq q \leq dim(K)$ are equal to the union of spectra of each $L_q(K_i)$ for $i = 1, ..., n$ separately [63], that is

$$S(\mathcal{L}_q(K)) = S(\mathcal{L}_q(K_1)) \cup S(\mathcal{L}_q(K_2)) \cup ... \cup S(\mathcal{L}_q(K_n))\,.$$

Another very important property is that if simplicial complex K is formed by gluing two simplicial complexes K_1 and K_2 along a q-face, then the spectrum S is the union of spectra of K_1 and K_2, i.e. $S(L_i(K)) = S(L_i(K_1)) \cup S(L_i(K_2))$ for all $i \geq q+2$ [63]. Since we are dealing here with the simplicial complex formed from the cliques of a graph, we want to emphasis that the spectrum of a single k-clique, denoted by G, is $S(L_0(G)) = \{0, [k]^{k-1}\}$ [85], which is equivalent to $S(L_0(G)) = \{0, [k]^{f_0-1}\}$, and $S(L_i(G)) = \{[k]^{f_{i-1}}\}$, where $i = 2, ..., k$, and $f_0, f_1, ..., f_{k-1}$ are the entries of f-vector, and the exponent of $[k]$ means the multiplicity of an eigenvalue k. These properties are consequences of (2.3) and (2.4). Namely, every vertex in a k-clique G has upper degree $k-1$ and every pair of distinct vertices has a dissimilar common lower simplex (an edge), hence for $q = 0$ from (2.3) implies that

$$\mathcal{L}_0(G) = \begin{pmatrix} k-1 & -1 & ... & -1 \\ -1 & k-1 & ... & -1 \\ ... & ... & ... & ... \\ -1 & -1 & ... & k-1 \end{pmatrix}$$

so solving the eigenvalue problem of $\mathcal{L}_0(G)$ implies that $S(L_0(G)) = \{0, [k]^{k-1}\}$, and since the 0^{th} entry of f-vector is equal to the number of vertices in a complex $f_0 = k$, we can write the general expression $S(L_0(G)) = \{0, [k]^{f_0-1}\}$. For $(k-1) \geq q > 0$, every q-simplex σ^i in G has upper degree equal to $\deg_U(\sigma^i) = (k-1) - q$ and every pair of distinct q-simplices σ^i and σ^j are upper adjacent, hence from (2.4) implies that for $q > 0$

$$\mathcal{L}_q(G) = \begin{pmatrix} k & 0 & ... & 0 \\ 0 & k & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & k \end{pmatrix}_{f_q \times f_q}\,.$$

The eigenvalue spectra has only one eigenvalue $\lambda = k$ with multiplicity equal to the number of q-simplices which is equal to f_q, the q^{th} entry of f-vector, so that $S(L_q(G)) = \{[k]^{f_q}\}$. For a single $(k-1)$-simplex f_q is equal to the number of q-dimensional faces, that is

$$f_q = \frac{k!}{(k-1-q)!(q+1)!}\,.$$

An example of the properties of combinatorial Laplacian spectra is illustrated in Figure 2.10 where Sq denotes the q^{th} component of the spectrum. In a) the complex consisting of two disjointed simplices is presented with the corresponding spectra and in b) through d) the two simplices are first joined along a 0-dimensional face (case (b)), followed by attachment along a 1-dimensional face (case (c)) and completing the process with attachment along a 2-dimensional face.

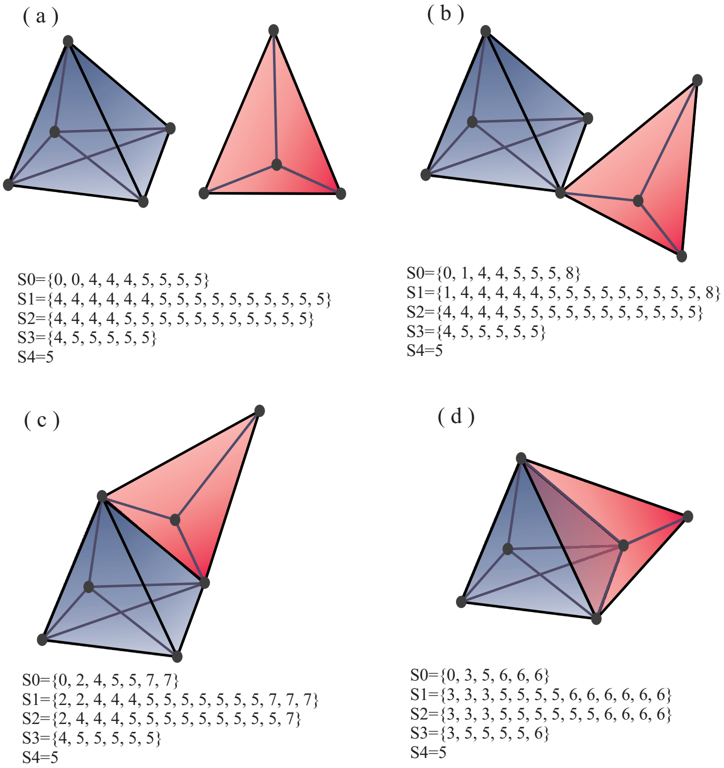


Figure 2.10: Spectrum of simplicial complex formed by 4-simplex and 3-simplex, when they share: (a) (-1)-face; (b) 0-face; (c) 1-face; (d) 2-face.

We must utilize some more practical methods in order to compare eigenvalue spectra of combinatorial Laplacian for different simplicial complexes, and the most transparent mode is visualization. To avoid problems which emerge from histogram or relative frequency plots due to the choice of the number of bins and their size and since we are dealing with $dim(K)+1$ eigenvalue spectra for a single simplicial complex (a fairly large number), we need a visualization method which depends on a single valued parameter unique for all the plots. For that purpose we use the convolution of the spectral density represented by Dirac delta function $\sum_i \delta(\lambda, \lambda_q^i)$ with a smooth kernel $g(x, \lambda)$ so that the density function [84]

$$f(x) = \int g(x, \lambda) \sum_i \delta(\lambda, \lambda_q^i) d\lambda = \sum_i g(x, \lambda_q^i) \,,$$

has advantageous visual properties. In the above expression λ_q^i is i^{th} eigenvalue of the q^{th} combinatorial Laplacian. Many kernels may be rendered useful in forming the density function, such as the Cauchy-Lorentz distribution $\frac{1}{\pi}\frac{\gamma}{(\lambda-x)^2+\gamma^2}$ or the Gaussian distribution $\frac{1}{\sqrt{2\pi\sigma}}exp\left(-\frac{(x-m_x)^2}{2\sigma^2}\right)$. Our choice here is the Cauchy-Lorentz kernel yielding the following density function

$$f(x) = \sum_i \frac{\gamma}{(\lambda_q^i - x)^2 + \gamma^2} \, ,$$

where γ is a fixed parameter which regulates the resolution (the level of detail in the plot) so that a too high value blurs the spectrum while too low value disguises it. In all spectra presented here the value $\gamma = 0.03$ was used chosen after careful consideration of a number of different γ-values.

Chapter 3

From complex networks to simplicial complexes and back

This section is devoted to the definitions of simplicial communities in complex networks. It will be shown there are various types of simplicial communities depending on the simplicial complex representation of the complex network. The term "simplicial community" is defined with reference to the definition of q-connectivity classes of simplicial complex.

3.1 Simplicial complexes of complex networks

From the simplicial complex's properties, introduced in the previous chapter, we notice that relationships as well as the aggregations of simplices are strongly dependent on various dimensions. The aggregations of simplices at various dimensions, namely the q-levels, we will call q-dimensional simplicial communities. In this way the q-dimensional simplicial communities are identified by the q-connectivity classes, although the term "community" will be made more clear when different types of simplicial complexes that may be constructed from complex networks are presented in detail. Our aim is to show that the formation, identification, overlapping and merging of simplicial communities are captured by the structure vectors and combinatorial Laplacian.

The versatility of simplicial complex representations of complex networks enables us to investigate topological properties of different substructures emerging from their relationships and whose interconnectidness forms the overall structure of complex networks. As a result we have an insight into the impact of aggregations of simplicial communities on the overall structure of the complex network. There are several types of simplicial complexes that may be constructed from graphs. We mention here the most important ones:

• *Clique complex* [10]: the vertices of clique complex are nodes of the underlying graph G, and simplices are all maximal cliques (together with all their subcliques);

• *Neighborhood complex* [48], [49]: the vertices of neighborhood complex are nodes of the underlying graph G, and to each vertex v of graph G corresponds a simplex which contains a vertex v and all of its neighboring vertices, that is simplices are all the subsets of the vertex set of G that have a common neighbor;

• *Independence complex* [10]: the vertices of independence complex are nodes of the underlying graph G, whereas simplices are maximal cliques (together with all their subcliques) of the complement graph of G (a graph in which two nodes are adjacent if they are not adjacent in graph G, and vice versa), that is simplices are all the independent sets (anticliques) of G;

• *Matching complex* [86]: the vertices of matching complex are the edges of the underlying graph G and simplices are sets of edges of G with no two edges having a common vertex; in other words, the matching complex is a clique complex of the complement graph of the line graph of G.

Since there are different simplicial complex representations of the complex network, different substructures emerge. In the present paper we focus on the properties of two simplicial complex representations of complex networks: the clique complex and the neighborhood complex. In the case of the clique complex we are actually "filling" a k-clique (complete graph with k vertices) and form a $(k-1)$-dimensional polyhedra (embedded in a $(k-1)$-dimensional space) The formation of simplicial communities has similarity with the formation of k-clique communities [59]. Nevertheless, there are differences. In k-clique communities cliques are adjacent if they share $k-1$ vertices. In the case of clique simplicial communities, in a q-connectivity component, clique simplices have dimension larger or equal than q and two clique simplices are adjacent if they share $q+1$ vertices. However some pairs of clique simplices may share $q+2$, $q+3$, ..., $dim(K)$ vertices. Consequently, k-clique communities are contained in the clique simplicial communities, and the overlapping between two clique simplicial communities is encoded in the transition from the q-level to the $(q-1)$-level. In Fig. 3.1, the appearance of clique simplicial communities at three q-levels ($q = 4, 3, 2$) of the coauthorship network of scientists working on network theory and experiment, as compiled by M. Newman [87] is presented. Due to the lack of space not all simplicial communities (including those which contain a single simplex) on either q-level are presented, but only those which participate in the formation of a large clique simplicial community. At the 4-level two clique simplices of different dimensions (one has dimension $q = 5$ and the other $q = 7$)

form a clique simplicial community. At the 3-level a 4-dimensional clique simplex (which appears at the 4-level) joins the clique simplicial community by sharing four vertices (i.e. 3-dimensional face) with one of the clique simplices in simplicial community. Finally, at the 2-level 3-dimensional simplex (appearing as single simplicial community at 3-level) joins a clique simplicial community by sharing three vertices (i.e. 2-face) with one of the clique simplices in simplicial community. From this simple example we can see the these kind of communities formed by cliques could not be detected by the Clique percolation method, and the restriction to the fixed k in k-clique communities is not necessary.

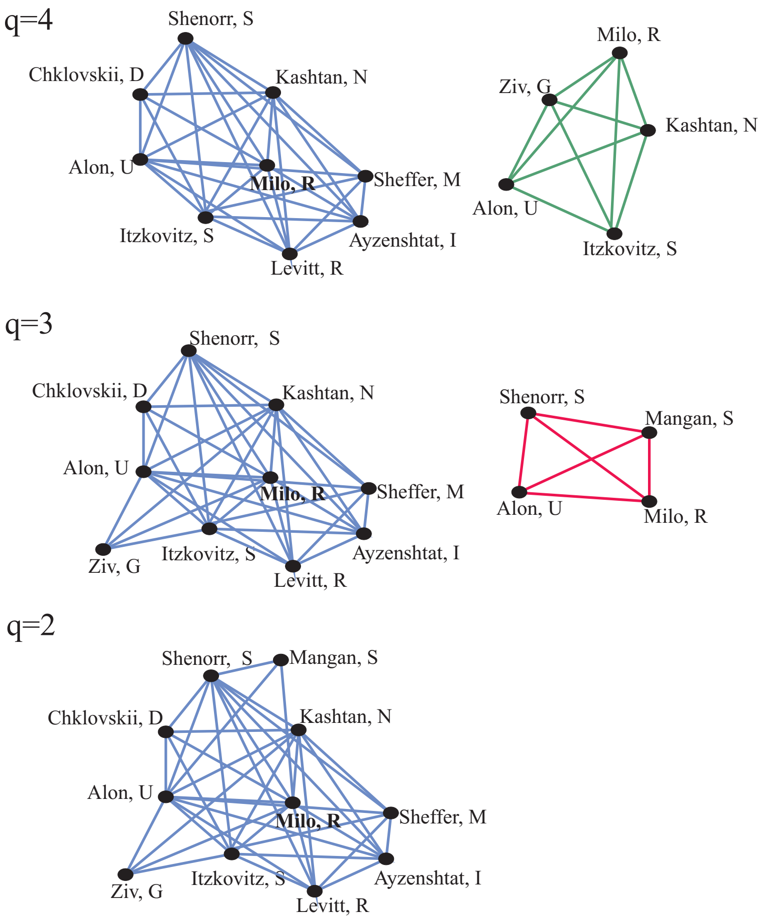


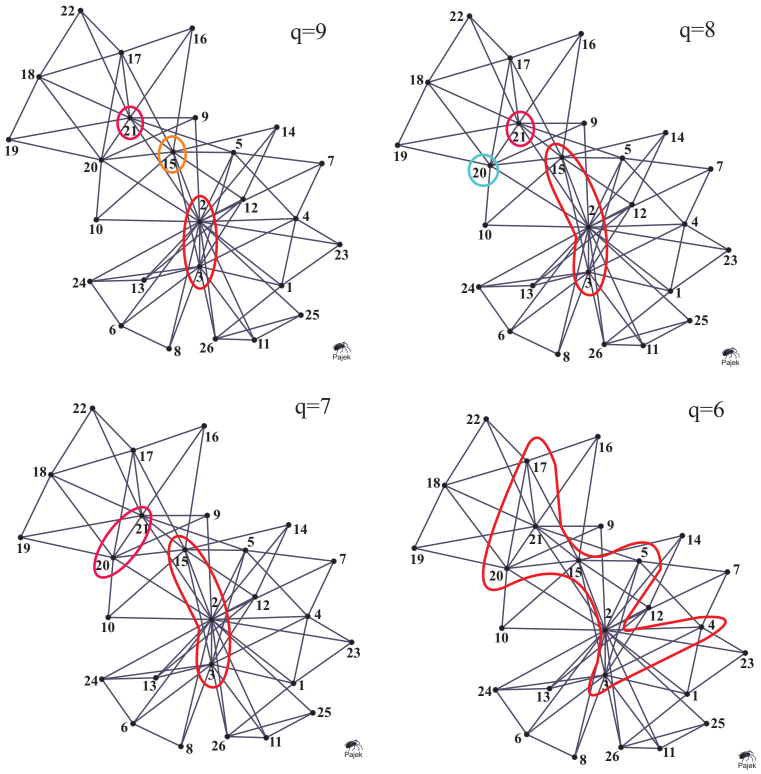
Figure 3.1: A sample of the clique simplicial communities at three q-levels ($q = 4, 3, 2$) of the coauthorship network.

Another advantage of the clique simplicial community approach as compared with the k-clique communities is in the complementary nature of simplicial complex and its conjugate. In the case of social network, simplicial communities of the conjugate complex are formed by the polyhedra which represent agents, whereas the vertices polyhedra are graph (network) vertices which are associated with network cliques to which agents belong. From the definition of q-nearness, it is obvious that two agent-polyhedra are adjacent if they belong to the same cliques. Such representation of substructures (related to cliques) have been useful in the analysis of cooperation and conflict in social networks [64]. For the above coauthorship network simplices in conjugate clique complex are associated to the authors, and vertices are associated to the papers which theycoauthored, as in Fig. 2.3. Among many simplicial communities appearing at q-levels, for example at the 4-level, we have detected two simplicial communities, one formed by simplices of different dimensions {*Barabási, A.-L., Jeong, H, Oltvai, Z*}, and the other {*Krapivski, P, Redner, S*}, containing the names of authors who appear jointly in a number of publications.

The polyhedra $\sigma(i)$ in the neighborhood complex is formed by "filling" the space between the corresponding node i and the neighboring nodes and two polyhedra $\sigma(i)$ and $\sigma(j)$ in the neighborhood complex are q-near if their corresponding nodes (i and j, respectively) have q common neighbors. Since the simplicial complex analysis strongly depends on dimension of simplices and of their faces, the analysis of higher order structure properties of the neighborhood complex are particulary interesting, since important quantities characterizing complex networks depend on degrees of the nodes. Namely, dimension q_i of simplex $\sigma(i)$ is equal to the degree k_i of the node i in the corresponding complex network.

As an example of the neighborhood complex we use the so called Brain network [88] in which nodes represent brain areas and links communication between them. This network[1] (Fig. 3.2) was obtained using the Planar Maximally Filtered Graph (PMFG) [89] from the correlation matrix of the time series collected by fMRI measurements of the brain areas activity while the people are asked to do two different tasks assessing short-term, that is episodic, memory. We avoid the detailed analysis of the methods for calculation of the correlation matrix and focus only on the obtained final network. In Fig. 3.2 we present the aggregation of nodes in the neighborhood simplicial communities at four q-levels ($q = 9, 8, 7, 6$) using the following procedure: brain network → neighborhood complex → neighborhood simplicial communities → brain network. At the 9-level two brain areas *Superior Frontal Gyrus* and *Middle Frontal Gyrus* form a simplicial community, and their functions (among others) are self-awareness, and a role in sustaining attention and working memory, respectively. Shifting to the next 8-level the another brain area *Superior Parietal Lobule*, responsible for the spatial orientation and receives a large portion of visual input, joins this simplicial community. At the 7-level another neighborhood simplicial community is formed containing *Inferior Occipital Gyrus* and *Middle Occipital Gyrus*, both parts of the larger *Occipital Lobe*, responsible for processing

[1] I thank J.-P. Schmidt for providing the data.



1-Anterior Cingulate; 2-Superior Frontal Gyrus; 3-Middle Frontal Gyrus; 4-Inferior Frontal Gyrus; 5-Medial Frontal Gyrus; 6-Insular Cortex; 7-Parahippocampnal Gyrus; 8-Superior Temporal Gyrus; 9-Inferior Temporal Gyrus; 10- Middle Temporal Gyrus; 11-Cingulate Gyrus; 12-Postcentral Gyrus; 13- Precentral Gyrus; 14-Inferior Parietal Lobule; 15-Superior Parietal Lobule; 16-Precuneus; 17-Cuneus; 18-Lingual Gyrus; 19-Fusiform Gyrus; 20-Inferior Occipital Gyrus; 21-Middle Occipital Gyrus; 22-Posterior Cingulate; 23-Unicate Fasciculus; 24-Lenticular Nucleus; 25-Thalamus; 26-Caudal

Figure 3.2: Neighborhood simplicial communities at four q-levels ($q = 9, 8, 7, 6$) of the brain network.

visual information. Finally, at the 6-level the two neighborhood communities merge into the large one, together with three more brain areas, namely *Inferior Frontal Gyrus*, *Medial Frontal Gyrus* and *Cuneus*. Even from this simple analysis we can anticipate the importance of aggregation of brain areas at different levels in order to perform tasks related to the short-term (or analogously long-term) memory.

Chapter 4

Results - preliminary and illustrative

We have found that for some characteristic networks topological quantity (Q-vector) of clique (and its conjugate) complex and neighborhood complex representation satisfy the statistical invariance, in the sense that it follows the behavior of degree distribution of the underlying complex network. For finding all maximal cliques we have used Bron-Kerbosch algorithm [90]. The values of the entries of Q-vector of Barabási-Albert complex network [91] clique complex representation is presented in Figure 4.1 left, whereas the right part of Figure 4.1 represents the values of Q-vector of conjugate clique complex and neighborhood complex.

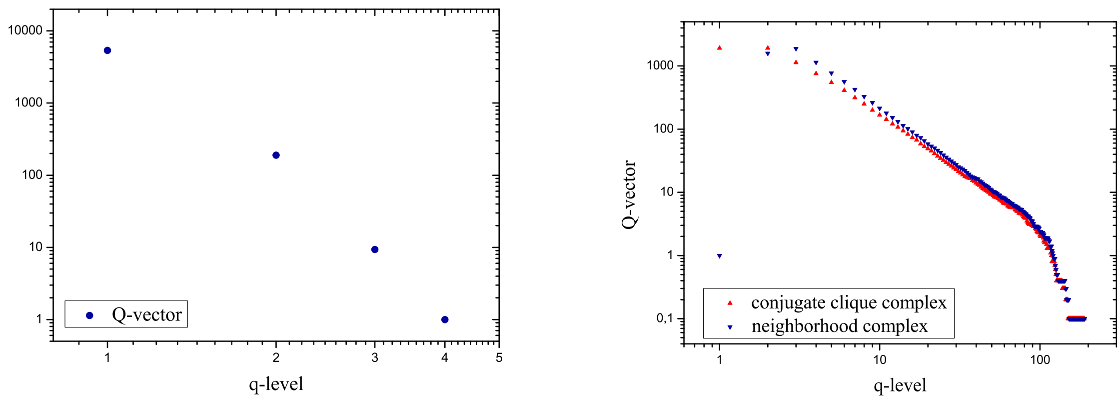


Figure 4.1: Q-vector entries of Barabási-Albert scale-free network: clique complex representation (left), and conjugate clique complex and neighborhood complex (right).

In Figure 4.2 left we have presented on double logarithmic scale the values of Q-vector entries of clique complex of protein-protein interaction complex network in yeast *S. cerevisiae* [92], whereas on right part of Figure 4.2 Q-vector values of conjugate clique complex and neighborhood complex on double logarithmic scale are presented. The degree distribution of protein-protein interaction complex network in yeast displays a power-law.

Since these two networks satisfy statistical invariance, we have tested this behavior on several other generated and real-world networks, and it turns out that in each case Q-vector follows the behavior of the degree distribution.

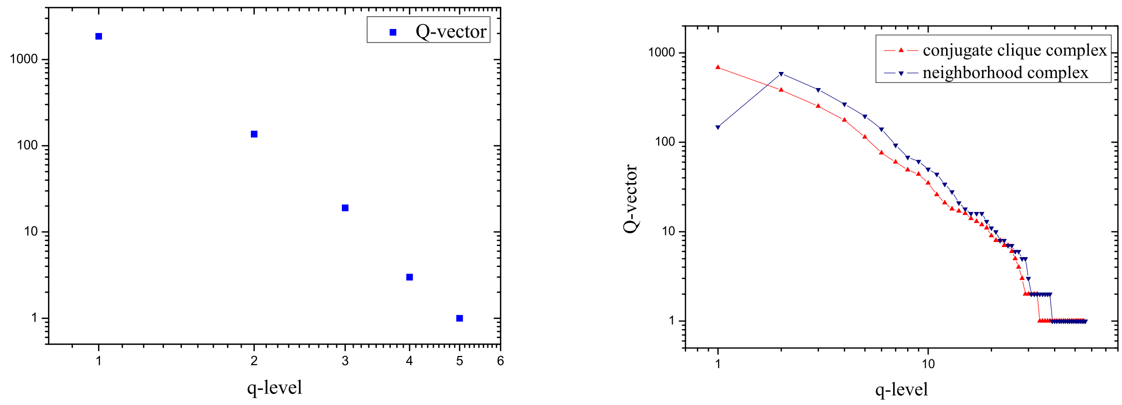


Figure 4.2: Q-vector entries of protein-protein interaction complex network in yeast: clique complex representation (left), and conjugate clique complex and neighborhood complex (right).

Betti numbers calculation of clique complexes of Barabási-Albert and protein-protein interaction complex network revealed that nonzero values appear only for dimensions $q = 0, 1, 2$ due to the small clique complex dimension (see left parts of Figures 4.1 and 4.2). Nevertheless, since the β_1 is related to the 1-dimensional "islands" of obstructions in discrete topological space defined through the clique complex of corresponding network, the determination of these "islands" is important for the traffic which flows through the network. The calculation of Betti numbers can be done by solving of the eigenvalue problem of higher-order combinatorial Laplacians [64] (and references therein), and the determination of "islands" of obstruction can be detected by considering eigenvectors of higher-order combinatorial Laplacian [64]. To track the changes of the number of tunnels (or "islands"), that is the values of β_1, depending on the scaling exponent γ of networks displaying power-law degree distribution, we have generated several generalized random networks [93] for different γ and represented them as clique complex. The dependence of number of 1-dimension holes (β_1) on the scaling exponent γ is presented in Figure 4.3, displaying the decrease in the number of 1-dimensional holes by increasing γ. Since the large number of real-world networks display power-law degree distribution with the exponent in the range $\gamma \in [2, 3]$, it is important to notice that the decrease of β_1 is

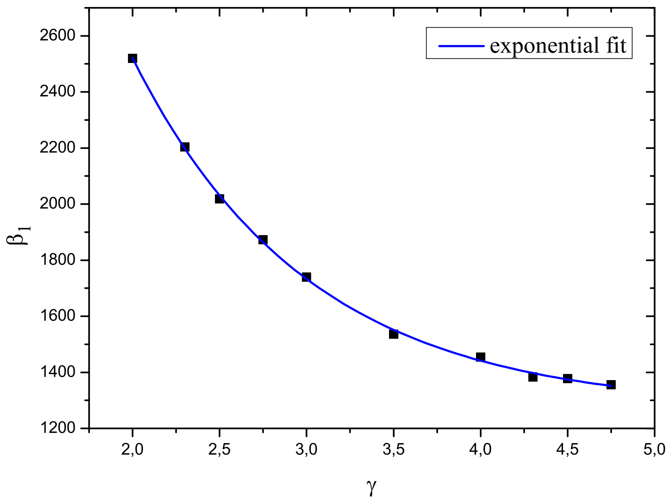


Figure 4.3: The dependence of number of 1-dimension holes (β_1) on the scaling exponent γ for clique complex of generalized random network with power-law degree distribution.

more rapid in this range then for $\gamma > 3$.

Chapter 5

Results - The neighborhood complex

This chapter is devoted to presentation of results related to the topological properties of the neighborhood complex of various complex networks [48], [50]. They are confirming the conclusions of the previous chapter.

5.1 Random network

The simplest type (and model) of a network, despite of its inadequacy for explaining the real world networks, is examined from the aspect of the simplicial complex representation, more concretely, from the aspect of the associated neighborhood complex NC_1. This is done for comparison as well as an illustration of the concepts and measures defined in the previous sections. The random network under study consists of 2000 nodes, with probability $p = 0.005$ that two nodes have a link. As mentioned before there is straightforward relationship between the degree of the node and the dimension of the corresponding simplex. That implicates the equivalence between degree distribution and dimension distribution. Furthermore, we expect that dimension distribution follows the well known bell-shaped form, which is characteristic of random networks. This property is presented at the Figure 5.1.

A random network has a characteristic scale in its node connectivity reflected by the peak of the distribution which corresponds to the number of nodes with the average number of links. Because of the equivalence of the distributions of degrees and dimensions, this property holds also for the corresponding simplicial complex representation.

The distribution of vector valued measures is illustrated by distributions of the

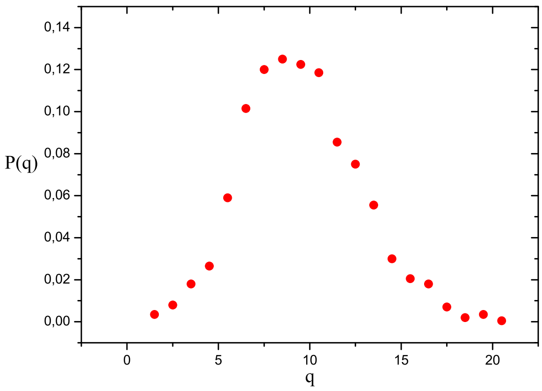


Figure 5.1: Distribution of dimensions of random network with $N = 2000$ nodes and probability $p = 0.005$ that two nodes have a link.

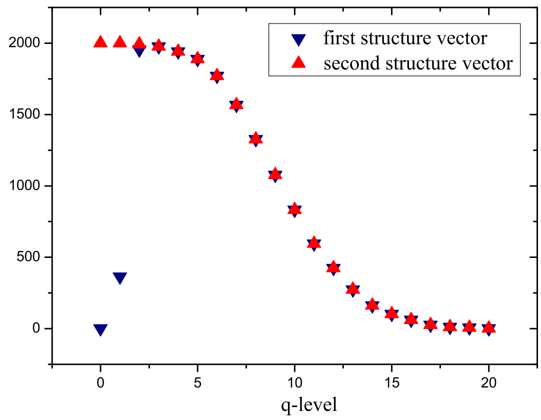


Figure 5.2: Values of first and second structure vectors for random network with $N = 2000$ nodes and linkage probability $p = 0.005$.

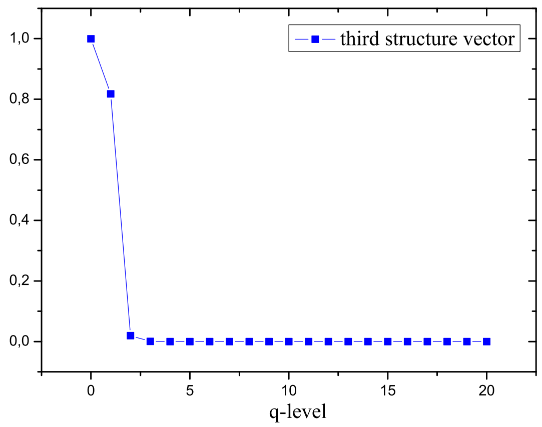


Figure 5.3: Values of third structure vector for random network with $N = 2000$ nodes and linkage probability $p = 0.005$.

first and second structure vector (Figure 5.2), as well as third structure vector (Figure 5.3).

The graphics of structure vectors indicate an interesting property of the structure of simplicial complex representation of random networks. From the highest up to the third level the structure is not connected, and then there is a jump in the connectivity of structure. The structure, observing it globally, is homogenous and there is not any preferential pattern of formation of connectivity classes.

5.2 Barabási-Albert model of scale-free networks

Following the algorithm introduced in [91], the scale-free network is generated. At this moment we will repeat just important features of this algorithm. Starting with m_0 randomly connected nodes, at each time step we add one new node which can be linked to m nodes already present in the network. The probability of connection to some old node i depends on its number of links k_i as

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} .$$

In this way two properties of real world network are captured: growth and preferential attachment. In this paper we have chosen $m_0 = 5$ and $m = 3$ values of parameters, and following the above procedure the network with $N = 5000$ nodes was generated. As it is already mentioned there is equivalence between degree distribution of complex network and dimension distribution of its corresponding

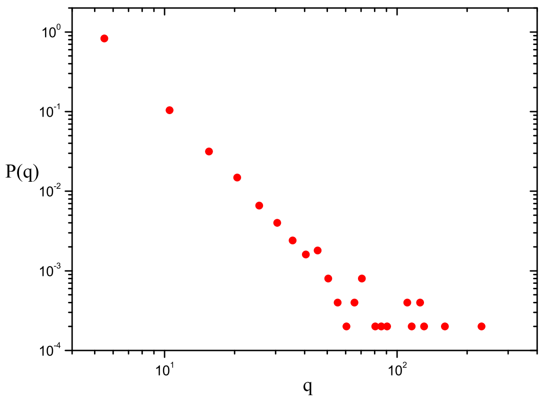


Figure 5.4: Dimension distribution of Barabási-Albert scale-free network type with $N = 5000$ nodes and parameters $m_0 = 5$ and $m = 3$.

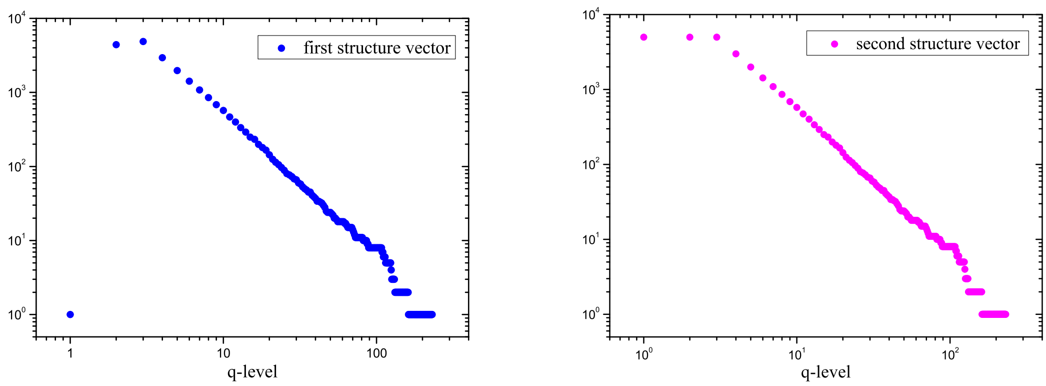


Figure 5.5: First (left) and second (right) structure vector values for Barabási-Albert scale-free network type with $N = 5000$ nodes and parameters $m_0 = 5$ and $m = 3$.

neighborhood complex NC_1. In Figure 5.4 the dimension distribution is presented.

The vector valued distributions of the first, the second, and the third structure vectors are presented in Figure 5.5 left, Figure 5.5 right, and Figure 5.6, respectively.

We can notice that the first and the second structure vectors in Figure 5.5 follow power-law behavior over few decades. The discrepancy from the power-law behavior for dimension distribution over the whole range comes from the finiteness of the network, as well as because of the randomness of the linking process. We can assume that these features have influence on the behavior of the first and the second structure vectors.

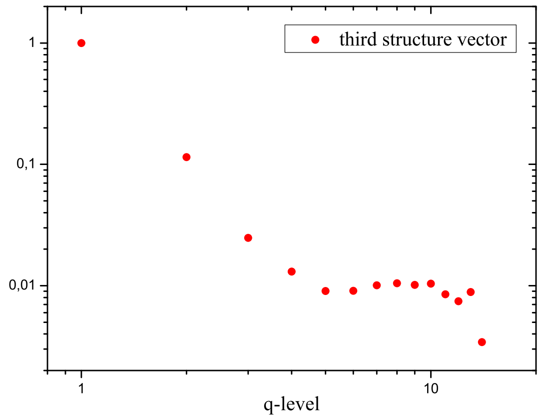


Figure 5.6: Third structure vector values for Barabási-Albert scale-free network type with $N = 5000$ nodes and parameters $m_0 = 5$ and $m = 3$.

By definition the length of the third structure vector is equal to the length of the first and the second, hence, from Figure 5.6 we can conclude that from the highest to the 14th level the structure is disconnected.

5.3 Exponential network

It is already mentioned that the majority of the real world networks have power-law degree distribution as its main characteristic. Nevertheless, there are some networks which have exponential degree distribution. An illustrative example of this type of network is the US Power Grid [3]. The nodes of this network are generators, transformers, and substations, and links are high-voltage transmission lines. We analyzed a US power grid network of the western United States which consists of 4941 nodes [94]. The dimension distribution, as well as the vector valued measures (normalized values of the 1st- and the 2nd-structure vectors) are illustrated in Figure 5.7 left, and the third structure vector is presented in Figure 5.7 (right). As can be seen from these figures all four measures are well fitted to the exponential function.

5.4 Scale-free networks

As an example of network with power-law degree (i.e. dimension) distribution an information type of network will be considered. We use *epa* to label this network, and it represents pages linking to www.epa.gov, and consists of $N = 4772$ nodes [94]. "This graph was constructed by expanding a 200-page response set to a search

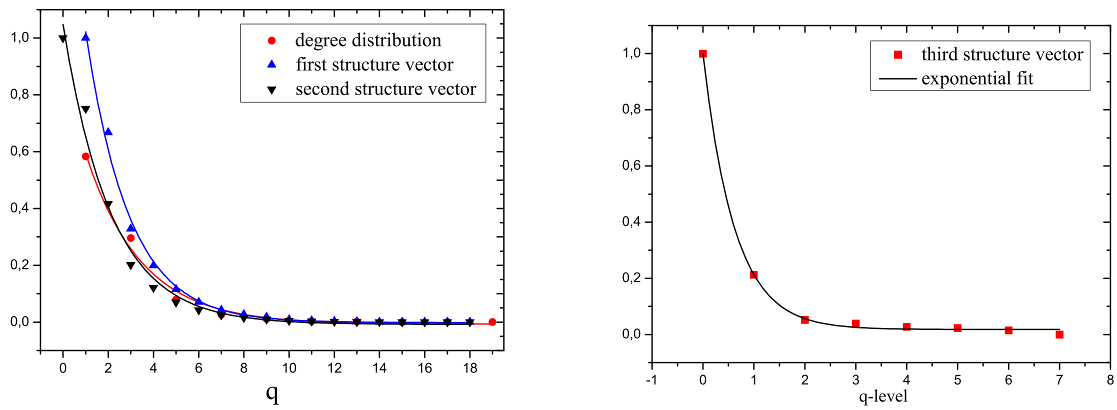


Figure 5.7: Degree (dimension) distribution, first, and second structure vectors (left) and third structure vector (right) for exponential US Power grid network. The exponential fit is indicated.

engine query, as in the hub/authority algorithm." [94]

The distribution of dimensions of epa network is presented in Figure 5.8 left.

Vector valued measures of the 1st and the 2nd structure vectors are illustrated in Figure 5.8 right, and the 3rd structure vector is presented in Figure 5.9. The connectivity levels are filled with simplices but they are not q-connected up to a certain value of the q-level. This level is rather high compared to other types of simplicial complex representations of complex networks studied in previous sections.

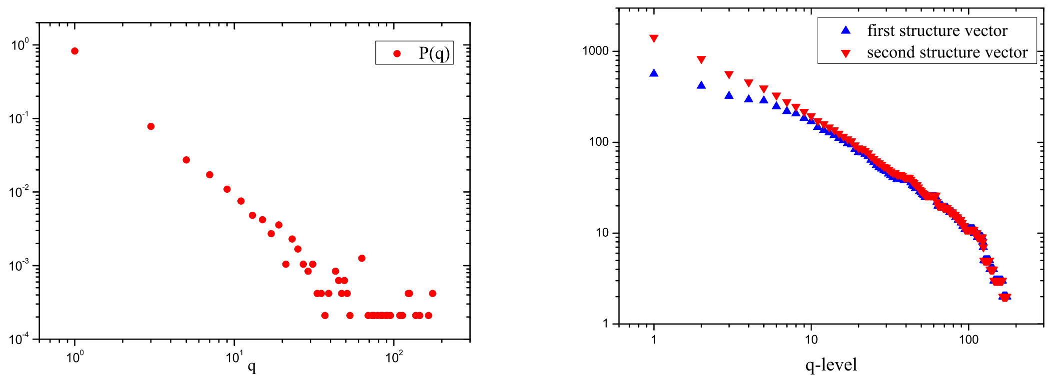


Figure 5.8: Dimension distribution of scale-free *epa* network (left) and values of the first and the second structure vectors for scale-free *epa* network (right).

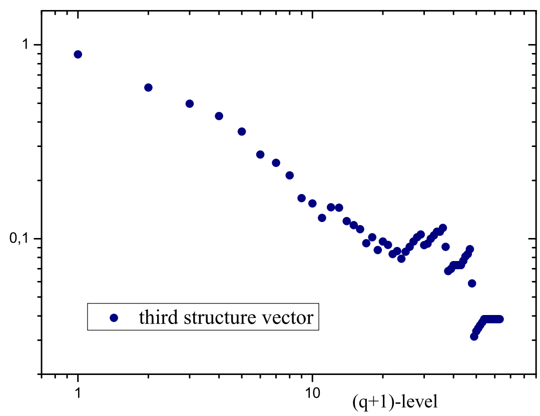


Figure 5.9: Values of the third structure vector for scale-free *epa* network.

Chapter 6

Results - The clique complex

The results of this chapter highlight the properties of the clique complex representation of complex networks [64] and their advantages in revealing the complexity of the underlying complex network [74].

6.1 The Zachary karate club

This network is formed by the members of karate club whose activities and relationships were observed by Zachary [95] over the time period of two years. During this time period the conflicting situation between the Instructor and the Administrator caused partition of club members into two fractions, each fraction supporting one of them. The Instructor and the Administrator are not in direct contact since their corresponding nodes are not nearest neighbors. Nevertheless, we may assume that they have been connected before the conflict. This network represents thus, a good example of a social network in which conflict causes changes in network's topology. The following analysis will show some properties of simplicial communities formed by the members of the club (clique complex analysis) and simplices associated to the members of the club formed by the cliques to which they belong (conjugate clique complex analysis).

6.1.1 Analysis of the clique complex

Simplices of the clique complex created from the Zachary karate club network are formed by members of the club together with all their subcliques. Initially all maximal cliques were found using the Bron-Kerbosch algorithm [90] and Q-vector components were determined whose graphical representation is presented in Figs. 6.1 and 6.2. The 0-level component is omitted since only one connectivity class

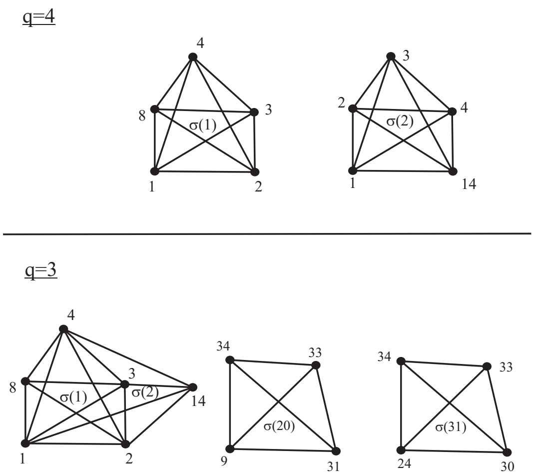


Figure 6.1: Graphical representation of the connectivity classes at 4- and 3-levels for clique complex of the Zachary karate club.

represented by the underlying graph exists at that level. At the highest q-level two 4-simplices (5-cliques) σ^1 and σ^2 appear (Fig. 6.1) and the Instructor (node labeled by 1) is part of each of them. Going to the next q-level (3-level) two more simplices σ^{20} and σ^{31} appear (Fig. 6.1), and the Administrator (node labeled by 34) is part of each of them. However, simplices σ^1 and σ^2 merge into a single connectivity class, indicating the existence of a strong group of Instructor's supporters since simplices σ^1 and σ^2 share the face of dimension one less than their simplex dimensions. At the 2-level (Fig. 6.2 left) many 2-simplices appear formed by supporters of either the Instructor or the Administrator (clustering), irrsepective of whether they contain either the Instructor or the Administrator vertex. The important transition occurs from 2-level to 1-level (Fig. 6.2), where simplices σ^1, σ^2, σ^{20} and σ^{31} merge into a single connectivity class together with majority of 2-simplices. Visual inspection also suggest that 2-simplices of supporters accumulate around the Instructor and the Administrator. At 1-level (Fig. 6.2 right) there is one more connectivity class, disconnected from the large one, of the Instructor's supporters formed by simplices σ^8, σ^9, σ^{10}, σ^{11} and σ^{19}.

Note that any subsimplex (face) of a simplex is also simplex and therefore any subclique of a maximal clique is also a clique. The analogous concept in the context of communities is nested communities, or communities inside communities. In the settings of the karate club this means that there are relationships between various subcliques of supporters. These subtle relationships can not be seen from the properties of q-connectivity and the Q-analysis, however, information about nested cliques (communities) is stored in the matrix of the q^{th} combinatorial Laplacian. Let us recall that diagonal elements $\mathcal{L}_{ii}$ are equal to the total number of $(q-1)$-simplices

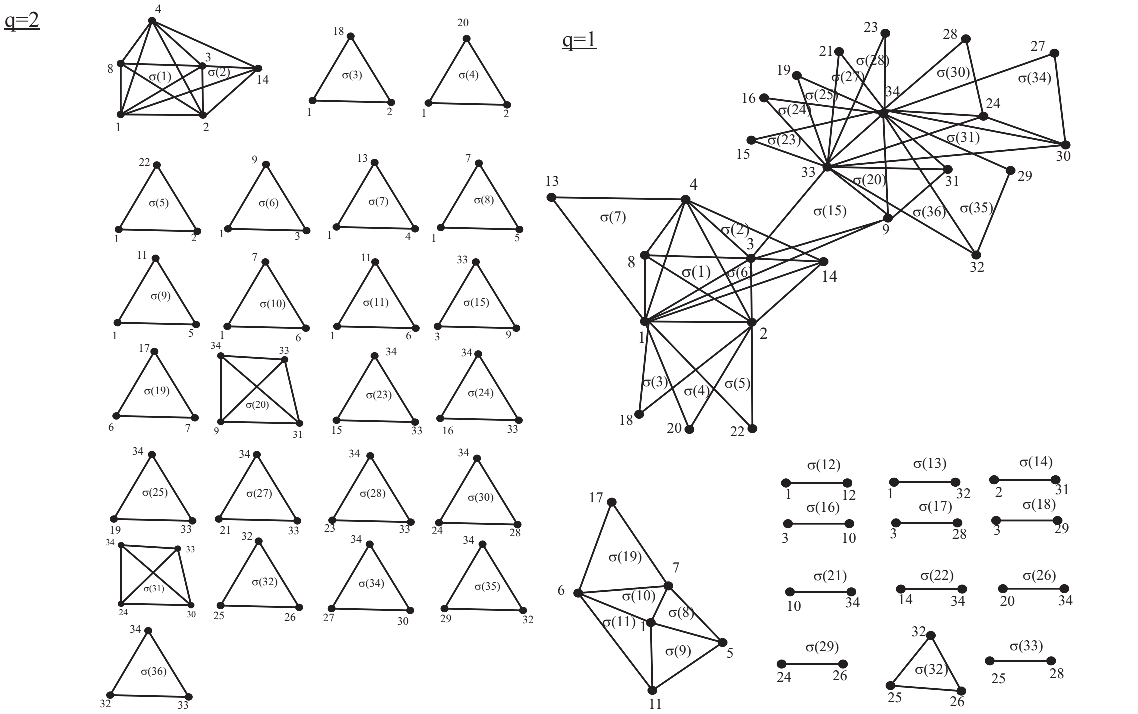


Figure 6.2: Graphical representation of the connectivity classes at 2- and 1-levels for clique complex of the Zachary karate club.

and $(q+1)$-simplices attached to the q-simplex i, and offdiagonal elements $\mathcal{L}_{ij}$ are nonzero if i and j q-simplices have a common $(q-1)$-simplex. Consequently, information about relationship between cliques (simplices) is reflected in the eigenvalues of the q^{th} combinatorial Laplacian.

The plots of combinatorial Laplacian eigenvalues of the clique complex obtained from the Zachary karate club network are presented in Figs 6.3, 6.4 and 6.5 for q-dimensions arranged in descending order from $q = 4$ to $q = 0$. In order to understand the meaning of the eigenvalues we will use the plots of simplices in Figs. 6.1 and 6.2. Consider two simplices at the highest 4-level (Fig. 6.1) and observe two eigenvalues of the 4-dimensional combinatorial Laplacian in Fig. 6.3 (left plot corresponding to $q = 4$). They share one 3-face ($\langle 1, 2, 3, 4 \rangle$) and the result is the formation of a connectivity class at 3-level, as reflected in two eigenvalues $\lambda_4^1 = 4$ and $\lambda_4^2 = 6$ in Fig. 6.3 left (as mentioned in Chapter 2 subscript marks dimension of the simplex and superscript represents the index). If these two simplices were not sharing a 3-face we would see only one peak at eigenvalue $\lambda_4 = 5$ as a consequence of the property of the spectrum formed by disconnected cliques of the same size (see Chapter 2). At the next lower q-level (3-level) we see that there are three connectivity classes: one formed by two 4-simplices (σ^1 and σ^2) sharing a 3-face and two classes formed by each of the simplices σ^{20} and σ^{31}. Since simplices σ^{20} and

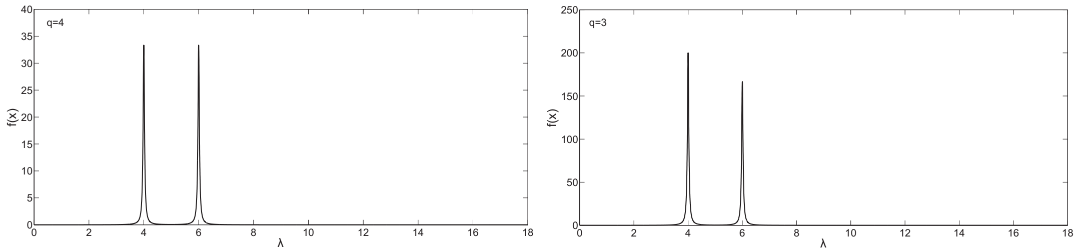


Figure 6.3: Spectral plots of clique complex of Zachary karate club for dimensions $q = 4$ and $q = 3$.

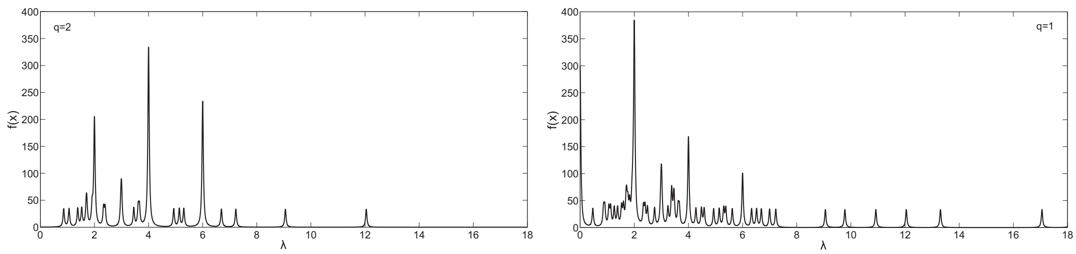


Figure 6.4: Spectral plots of clique complex of Zachary karate club for dimensions $q = 2$ and $q = 1$.

σ^{31} do not share 2-face at the 2-level they still form separate connectivity classes, and they contribute to the eigenvalue spectra of 3-dimensional combinatorial Laplacian by a single eigenvalue $\lambda_3^1 = 4$. Subsimplices of dimension 3 (3-simplices) of 4-simplices σ^1 and σ^2 mutually share 2-face, contributing to the $q = 3$ spectrum with eigenvalues $\lambda_3^2 = 4$ and $\lambda_3^3 = 6$. Consequently, the multiplicity of eigenvalues 4 and 6 is increased as seen in the increased heights of two peaks in Fig 6.3 right. The same type of analysis can be continued to other lower q-levels all the way to level 0. At lower connectivity levels the analysis becomes complicated due the increasing number of lower-dimensional simplices and their mutual connectivities. In spite of that, persistent presence of eigenvalues 4 and 6 is easily noticed in Figs 6.3 and 6.4, corresponding to q-dimensions $4, 3, 2$ and 1, originating from the 4- and 3-simplices mentioned above. *However, for $q = 0$ these two obviously important eigenvalues do not exist in the spectrum indicating that consideration of only the graph (0^{th} order combinatorial) Laplacian no information about larger communities and (overlapping) subcommunities is available.* Furthermore, the origin of the eigenvalue $\lambda_0 = 2$ which is dominant at the 0^{th}-dimension (Fig. 6.5) is more clear if we observe what is happening at dimensions $q = 2$ and $q = 1$. Its appearance at dimension $q = 2$ (Fig. 6.4 left) stems from 1-faces shared between (sub)simplices of dimension $q = 2$, while at 1-level its multiplicity originates due to appearance of 1-simplices.

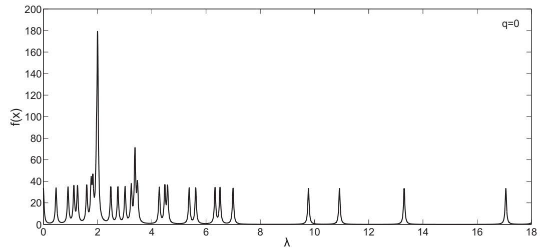


Figure 6.5: Spectral plots of clique complex of Zachary karate club for dimension $q = 0$.

6.1.2 Analysis of the conjugate complex

In the conjugate Zachary karate club clique complex (high dimensional) individuals are defined by the cliques to which each of them belongs in contrast to the original approach were cliques were composed of individuals from the club. From the simplicial clique complex K we have formed its conjugate complex K^{-1} in which simplices and vertices exchange roles, i.e. simplices are the individuals defined by cliques (vertices) to which they belong. For example, simplices τ_{12}^{1} and τ_{13}^{34} (we use τ to mark simplices in the conjugate complex) are the 12-simplex and 13-simplex associated to the Instructor and the Administrator, respectively, and they are defined by the maximal cliques to which they belong. Hence, the simplex τ_{13}^{34} associated to the Administrator is defined by 14 vertices $\{20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 34, 35, 36\}$ which are associated to the cliques $\{\sigma^{20}, \sigma^{21}, \sigma^{22}, \sigma^{23}, \sigma^{24}, \sigma^{25}, \sigma^{26}, \sigma^{27}, \sigma^{28}, \sigma^{30}, \sigma^{31}, \sigma^{34}, \sigma^{35}, \sigma^{36}\}$, and similarly for the Instructor 12-simplex.

We omit the graphical presentation of the Q-vector due to its large number of connectivity levels, however we present it componentwise in Table 1, where each simplex label, as before, is represented as a superscript.

The spectral plots corresponding to the conjugate clique complex K^{-1} are presented in Figs. 6.6, 6.7, 6.8 and 6.9 for dimensions $q = 13$ to $q = 0$.

The hierarchical structure of the Q-vector and high dimensional combinatorial Laplacian of the conjugate complex bear qualitative resemblance to the original complex. However, the conjugate simplicial complex may detect possible structural and q-connectivity sources of conflict between individual τ^{1}, labeled as the Instructor, and individual τ^{34}, labeled as the Administrator, features that may remain beyond the reach of the original complex. As previously mentioned, conflict arises between these two persons causing Instructor to leave the club and start up a new one. From the structure of Q-vector (Table 1) and spectral plots (Figs. 6.6, 6.7, 6.8 and 6.9) it is clear that at dimensions 13 through 8 the Instructor and the Administrator are the only simplicies due to their association in the highest number of cliques. At $q = 3$ level both are parts of two connectivity classes but at level $q = 2$ the connec-

Table 6.1: Components of the Q-vector of the conjugate clique complex constructed from the Zachary karate club network

$q = 13$	τ^{34}
$q = 12$	τ^{1}, τ^{34}
$q = 11$	τ^{1}, τ^{34}
$q = 10$	τ^{1}, τ^{34}
$q = 9$	τ^{1}, τ^{34}
$q = 8$	τ^{1}, τ^{33}, τ^{34}
$q = 7$	τ^{1}, $\tau^{(33,34)}$
$q = 6$	τ^{1}, τ^{3}, $\tau^{(33,34)}$
$q = 5$	τ^{1}, τ^{2}, τ^{3}, $\tau^{(33,34)}$
$q = 4$	$\tau^{(1,2)}$, τ^{3}, $\tau^{(33,34)}$
$q = 3$	$\tau^{(1,2)}, \tau^{3}, \tau^{32}, \tau^{(33,34)}$
$q = 2$	$\tau^{(1,2,3,4)}, \tau^{6}, \tau^{7}, \tau^{9}, \tau^{24}, \tau^{28}, \tau^{32}, \tau^{(33,34)}$
$q = 1$	$\tau^{(1-9,11,24,30,32,33,34)}, \tau^{10}, \tau^{14}, \tau^{20}, \tau^{25}, \tau^{26}, \tau^{28}, \tau^{29}, \tau^{31}$
$q = 0$	{all simplices}

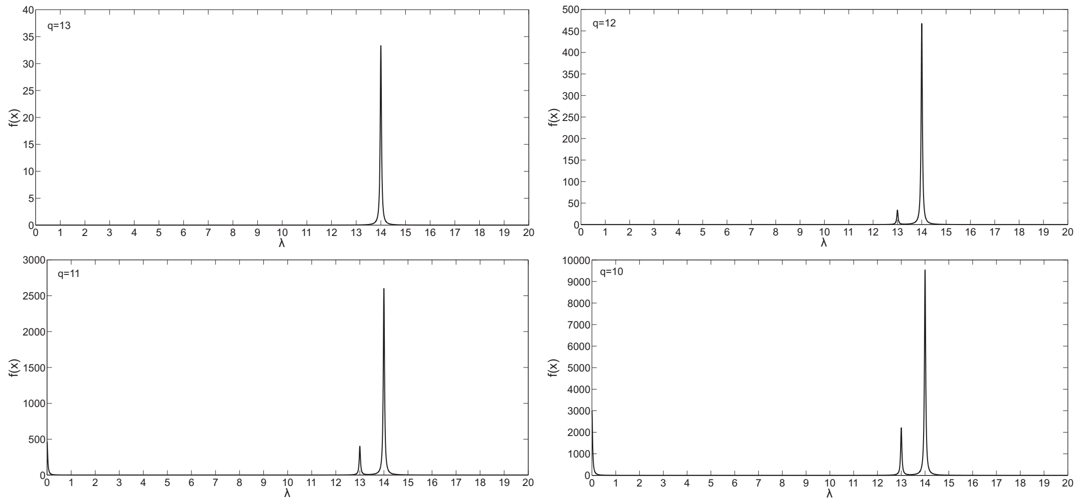


Figure 6.6: Spectral plots of conjugate clique complex of Zachary karate club network for dimensions from $q = 13$ through $q = 10$.

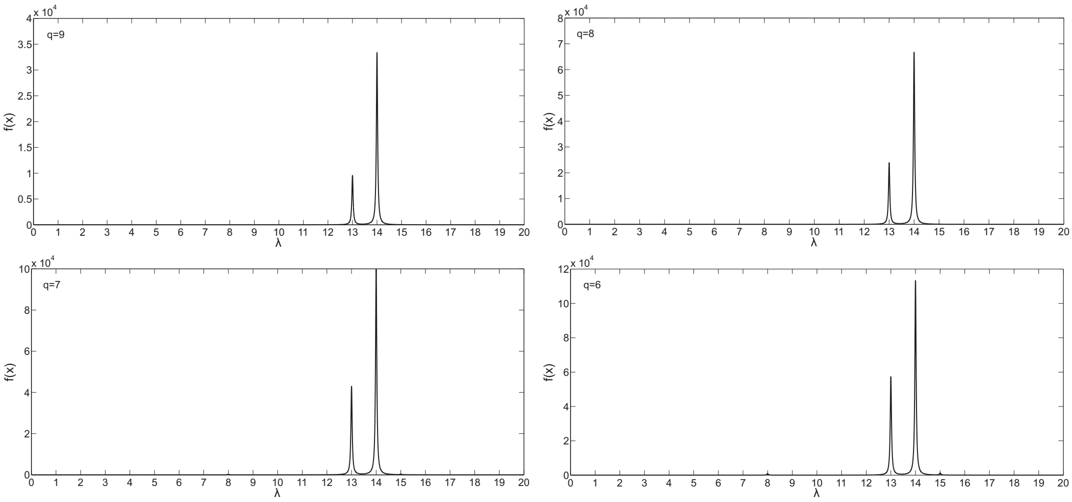


Figure 6.7: Spectral plots of conjugate clique complex of Zachary karate club network for dimensions from $q = 9$ through $q = 6$.

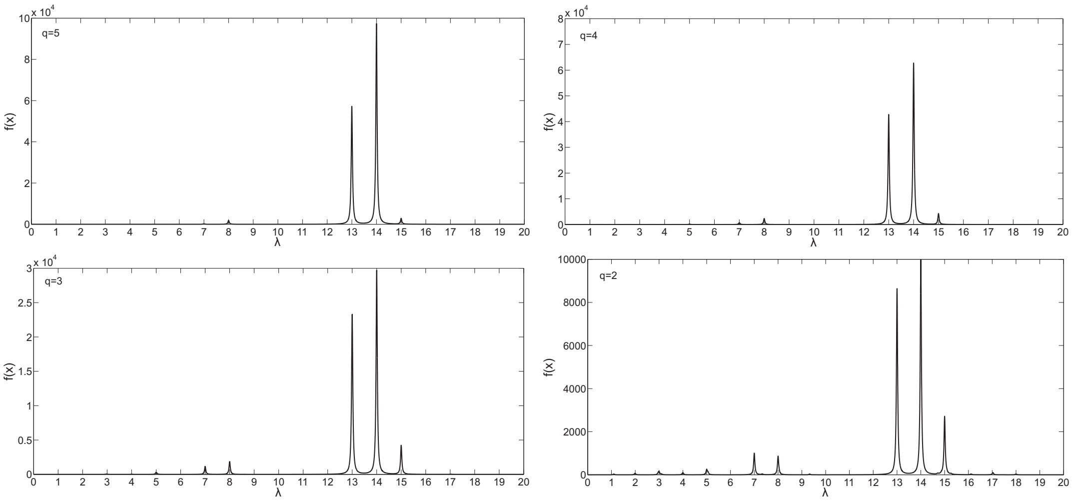


Figure 6.8: Spectral plots of conjugate clique complex of Zachary karate club network for dimensions from $q = 5$ through $q = 2$.

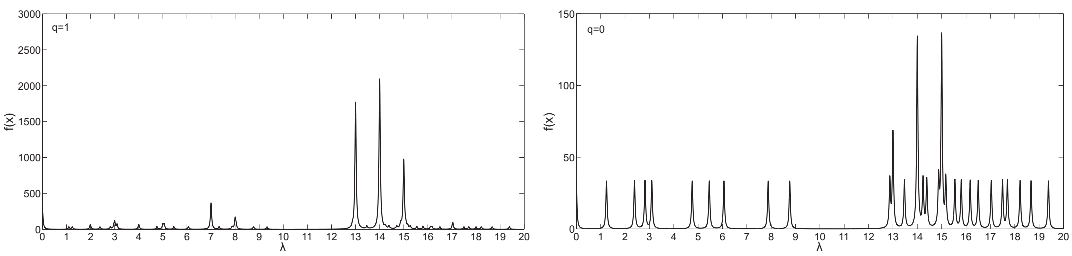


Figure 6.9: Spectral plots of conjugate clique complex of Zachary karate club network for dimensions from $q = 1$ through $q = 0$.

tivity class with the Instructor grows due to an increasing number of new members. At $q = 1$ level connectivity classes containing τ^1 and τ^{34} merge into one large connectivity class. All other simplices appear either as a separate connectivity class or as part of the connectivity classes to which either τ^1 or τ^{34} belong to, signaling the polarization of the club members around the Instructor and the Administrator. Hence, it is clear that τ^1 and τ^{34} stand out at each level of connectivity as individuals who cause social disintegration in the club. The corresponding eigenvalues of the combinatorial Laplacian are 13 and 14 which dominate the spectrum even for $q = 0$.

6.2 Les Miserables

This social network represents a good example of mixed cooperation/conflict interaction between the network elements. It is formed of 33 key characters extracted from the Victor Hugo's novel *Les Miserables* due to their co-appearances [96]. By mixed cooperation/conflict interactions we mean that a single character is in the cooperation (or friendship) interaction with one person and in the conflict interaction with another. The example of these mixed interactions is illustrated in Fig. 6.10. In the following we will use the term cooperation instead of the term friendship.

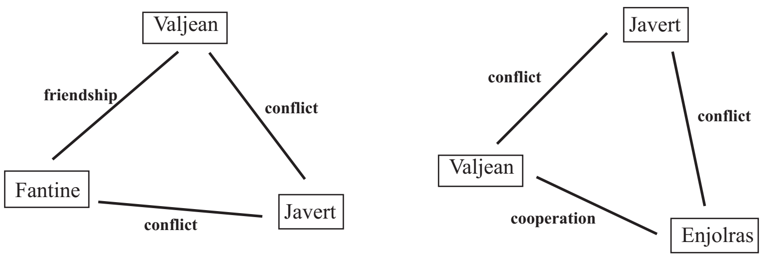


Figure 6.10: An example of mixed cooperation/conflict interaction between the characters of *Les Miserables*.

6.2.1 Analysis of the Les Miserables clique complex

Simplicial communities of the Les Miserables network are the cliques formed by interactions of the novel's characters according to their co-appearance. As in the Zachary karate club case, we first extract all maximal cliques using Bron-Kerbosch algorithm [90]. In Table 2 all connectivity classes of Q-vector are presented where each simplex has an integer label represented by the superscript. We omit graphical representation due to practical reasons.

Table 6.2: Components of the Q-vector of the clique complex constructed from the Les Miserables network

$q = 5$	$\{\sigma^{16}\}$
$q = 4$	$\{\sigma^{16}\}$, $\{\sigma^{20}\}$
$q = 3$	$\{\sigma^{11}\}$, $\{\sigma^{13}\}$, $\{\sigma^{14}\}$, $\{\sigma^{16}, \sigma^{20}\}$, $\{\sigma^{17}\}$, $\{\sigma^{19}\}$, $\{\sigma^{22}\}$
$q = 2$	$\{\sigma^{1}\}$, $\{\sigma^{3}\}$, $\{\sigma^{4}\}$, $\{\sigma^{5}\}$, $\{\sigma^{6}\}$, $\{\sigma^{7}\}$, $\{\sigma^{8}\}$, $\{\sigma^{10}\}$, $\{\sigma^{11}\}$, $\{\sigma^{13}\}$, $\{\sigma^{14}\}$, $\{\sigma^{16}, \sigma^{17}, \sigma^{19}, \sigma^{20}, \sigma^{22}\}$, $\{\sigma^{18}\}$
$q = 1$	$\{\sigma^{1}\}$, $\{\sigma^{2}\}$, $\{\sigma^{3}, \sigma^{4}, \sigma^{5}, \sigma^{6}, \sigma^{7}, \sigma^{8}, \sigma^{10}, \sigma^{16}, \sigma^{17}, \sigma^{19}, \sigma^{20}, \sigma^{22}\}$, $\{\sigma^{9}\}$, $\{\sigma^{11}\}$, $\{\sigma^{12}\}$, $\{\sigma^{13}\}$, $\{\sigma^{14}\}$, $\{\sigma^{15}\}$, $\{\sigma^{18}\}$, $\{\sigma^{21}\}$
$q = 0$	$\{\sigma^{1}, ..., \sigma^{22}\}$,

Dominant simplicies at levels (Table 2) from $q = 5$ to $q = 2$ include σ^{16}, σ^{20}, σ^{17}, σ^{19}, and σ^{22} whose vertices are members of the revolutionary group called The Friends of the A B C, (fr. Abaisse, the debased), that is to say, the depraved people. Thus, each pairwise connection between the individuals is characterized by cooperation. If we consider these simplices as "cells" of an underground organization, we see that at the 2-level they all aggregate into a single connectivity class through clustering (aggregation of 2-simplices). Simplex σ^{14} which appears at the 3-level represents a street gang of murderers and robbers called Patron-Minette, so the connection between vertices is through cooperation. Two simplices emerging from male and female friendships, σ^{11} and σ^{13} respectively, appear at the 3-level and can be generally assumed as composed of cooperating individuals (vertices). The simplices formed by exlusively cooperative individuals who are also cooperative as a group are labeled as cooperation simplices. Note that cooperation simplex assumes the concept of group cooperation, as well. In a similar manner we define conflict simplices though no such simplices exist in this network. At levels $q = 5$ through $q = 3$ only cooperation simplices appear. In 2-simplices in which conflict among individuals is present there is always one cooperation interaction and two conflict interactions (see example at Fig. 6.10). Hence, cooperation groups are formed by at least 4 persons, and conflict between certain characters entails that the groups to which they are associated are formed by 3 persons.

In order to gain a deeper insight into the structure of the network we will analyze the combinatorial Laplacian spectra, whose plots are presented in Fig. 6.11.

We focus on eigenvalues $\lambda^1 = 4$ and $\lambda^2 = 6$ which dominate all spectra, although not jointly at each dimension. Emergence of the eigenvalue $\lambda^2 = 6$ at dimension $q = 5$ is caused by a single cooperation simplex (Table 2) and it may be explained by the property of the clique combinatorial Laplacian spectrum (Chapter 2). Both eigenvalues, $\lambda^1 = 4$ and $\lambda^2 = 6$ appear at dimension $q = 4$ and one more eigenvalue $\lambda^1 = 4$ emerges from the 3-dimensional subsimplex shared by simplices σ^{16} and

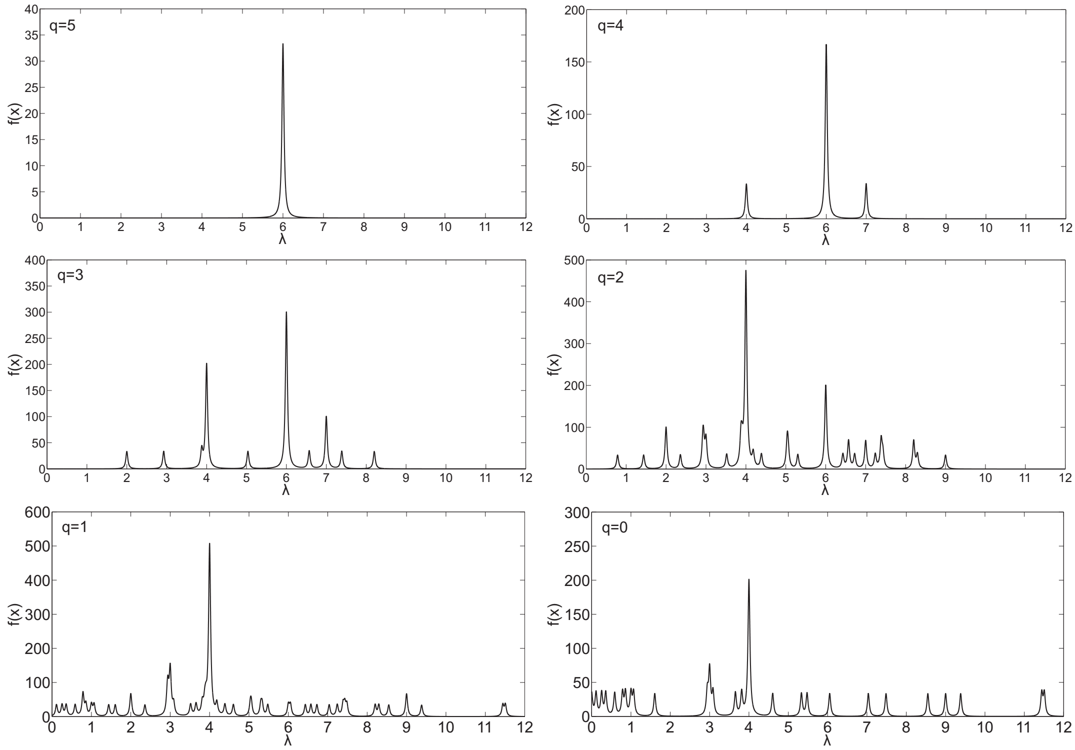


Figure 6.11: Spectral plots of the clique complex of Les Miserables network for dimensions from $q = 5$ through $q = 0$.

σ^{20}. Multiplicity of eigenvalue $\lambda^1 = 4$ increases at level $q = 3$ due to appearance of cooperative 3-simplices σ^{11}, σ^{13}, σ^{14}, σ^{17}, σ^{19} and σ^{22}. If solely the graph Laplacian ($0^{\underline{th}}$ order combinatorial Laplacian) was considered the origin of this important eigenvalue would remain inaccessible.

6.2.2 Analysis of the Les Miserables conjugate clique complex

We have reversed the roles of characters and cliques to which they belong, by associating simplices to individuals and vertices to cliques. Using the Bron-Kerbosch algorithm [90] all maximal cliques were found and the Q-vector components were determined (Table 3). The simplex τ_8^4 (τ is used to mark simplices in conjugate complex) is associated with the main character of the novel, Valjean and he is present in the largest number of cliques, however, he is connected to another person through shared cliques at the 2-level. All cliques to which Valjean belongs are 2-simplices which, as mentioned earlier, consist of mixed cooperation/conflict relationships while the simplex τ_2^{16} belonging to the same connectivity class as Valjean is represented by Inspector Javert, another person present exclusively in the cooperation/conflict

Table 6.3: Components of the Q-vector of the conjugate clique complex constructed from the Les Miserables network

$q = 8$	$\{\tau^4\}$
$q = 7$	$\{\tau^4\}$
$q = 6$	$\{\tau^4\}$
$q = 5$	$\{\tau^4\}$, $\{\tau^{23}\}$
$q = 4$	$\{\tau^4\}$, $\{\tau^{22}\}$, $\{\tau^{23}\}$, $\{\tau^{24}, \tau^{27}\}$
$q = 3$	$\{\tau^4\}$, $\{\tau^{22}\}$, $\{\tau^{23}\}$, $\{\tau^{24}, \tau^{27}\}$
$q = 2$	$\{\tau^4, \tau^{16}\}$, $\{\tau^{14}\}$, $\{\tau^{22}\}$, $\{\tau^{23}, \tau^{24}, \tau^{27}\}$
$q = 1$	$\{\tau^1\}$, $\{\tau^4, \tau^{13}, \tau^{14}, \tau^{15}, \tau^{16}, \tau^{19}, \tau^{22}, \tau^{23}, \tau^{24}, \tau^{26}, \tau^{27}, \tau^{29}, \tau^{30}\}$, $\{\tau^8\}$, $\{\tau^9\}$, $\{\tau^{12}\}$
$q = 0$	$\{\tau^1, ..., \tau^{33}\}$

simplices. Simplex τ_5^{23} at the 5-level is associated to Enjolras, the leader of the revolutionary group "The Friends of the A B C", who joins the connectivity class formed by other members of this group at the 2-level.

Spectral plots of the combinatorial Laplacian for dimensions $q = 8$ through $q = 0$ are presented in Fig. 6.12. A clear dominance of the eigenvalue $\lambda = 9$, originating from the simplex τ_8^4, that is Valjean, is easily recognized. For dimension $q = 0$ the dominating eigenvalue is $\lambda = 5$, originating from simplices associated to the members of the large cooperation simplicial community (members of The Friends of the A B C).

6.3 Spectral entropy

The above example illustrates that analysis using both the Q-vector and the spectra of the higher order combinatorial Laplacian are powerful complementary tools for the analysis of complex networks in spite of large number of eigenvalues that appear in the spectrum of ordinary graphs (networks). A careful analysis through various connectivity levels of the corresponding clique complex reveals *simplex communities* as overlapping or disjoint entities. We introduce here another quantity which measures the degree of overlapping of simplices in the complex in each dimension. Let λ_q^i be the eigenvalues of q^{th} combinatorial Laplacian and $i \in \{1, 2, ..., f_q\}$, where f_q is the q^{th} entry of f-vector, that is the number of q-simplices (not maximal). Then the q^{th} *spectral entropy* H_q is defined as

$$H_q = -\frac{1}{\log(f_q)} \sum_{i=1}^{f_q} p(\lambda_q^i) \log p(\lambda_q^i) \,, \tag{6.1}$$

where $p(\lambda_q^i) = \frac{\lambda_q^i}{\sum_{j=1}^{f_q} \lambda_q^j}$ is the eigenvalue probability, which may be understood as the contribution of the eigenvalue λ_q^i to the whole spectrum of the q^{th} combinatorial Laplacian. For $\lambda_q^i = 0$ clearly $H_q = 0$ and $\frac{1}{\log(f_q)}$ is the normalization constant which restricts the entropy values between 0 and 1 where f_q denotes the number of q-dimensional simplices in the simplicial complex. Namely, if a simplicial complex is a single vertex, then the spectral entropy is minimal and equal to zero, whereas when simplicial (sub)complex is formed by the q-simplices (where q is the same for all simplices) and no pairs of different q-simplices share $(q-1)$-face then q^{th} spectral entropy is maximal and equal to 1. Therefore, the maximum of combinatorial entropy H_q for a particular dimension q corresponds to the disconnectivity of q-simplices at dimension q. Of course, at dimensions smaller than $q-1$ any pair of q-simplices may share $q-2$, $q-3$, ..., 0 face, and henceforth H_{q-1}, H_{q-2}, H_{q-3}, ..., is different than 1. When simplicial complex is formed by only one q-simplex (that is one $(q+1)$-clique) the combinatorial entropy H_q is equal to 1 for all q, since it is disconnected from any other q-simplex at all dimensions. Hence, any deviation from $H_q = 1$ for specific q indicates that internal (sub)structures of simplicial complex at dimension q are more or less overlapped.

The dependance of the q^{th} spectral entropy on q-dimensions for the Zachary karate club clique complex and its conjugate complex are presented in Figure 6.13 (a) and (b), respectively. Inspection of Figure 6.13 (a) shows that for dimension $q = 4$ there are two simplices which share a 3-face, hence $H_4 \neq 1$ and in the case that they share $\{2, 1, 0, -1\}$-face instead of 3-face the 4^{th} spectral entropy would be $H_4 = 1$. Moving to the next (lower) dimension $(q = 3)$, two new simplices are added. Nevertheless, since they are not sharing 2-faces the spectral entropy is increased but it is not equal to 1 since two simplices from the 4-level share 3-face, and hence their subsimplices (3-simplices) share 2-face. Moving to the next lower dimension new simplices are added and the q^{th} spectral entropy $(q < 3)$ decreases and thus diverges from the disconnected cliques behavior. From Figure 6.13 (b) we see that the spectral entropy of the conjugate simplicial complex $H_q \approx 1$ or $H_q = 1$ for dimensions $2 < q \leq 13$ meaning that at these dimensions the complex behaves like a set of disconnected cliques. This is the consequence of the polarizing effect of two simplices, the Instructor and the Administrator. The origin of this property is obvious from the spectral plots (Figures 6.6, 6.7, 6.8 and 6.9) in which eigenvalue $\lambda = 13$ appears from the Instructor simplex and eigenvalue $\lambda = 14$ arises from the Administrator simplex, both easily recognized at higher dimensional levels. Furthermore, unlike in the clique complex case, these two dominating eigenvalues persist for all q, including $q = 0$.

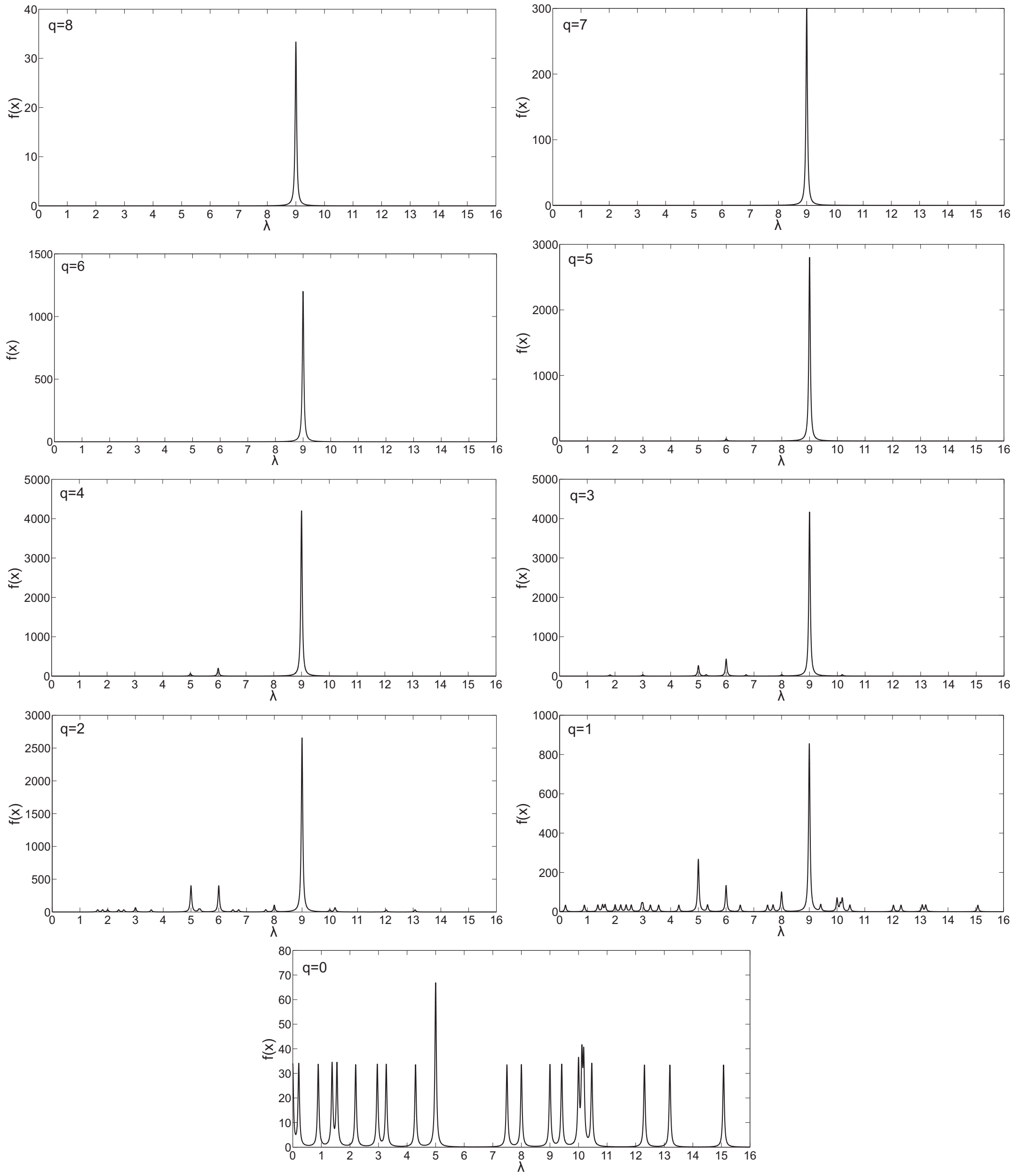


Figure 6.12: Spectral plots of conjugate clique complex of Les Miserables network for dimensions from $q = 8$ through $q = 0$.

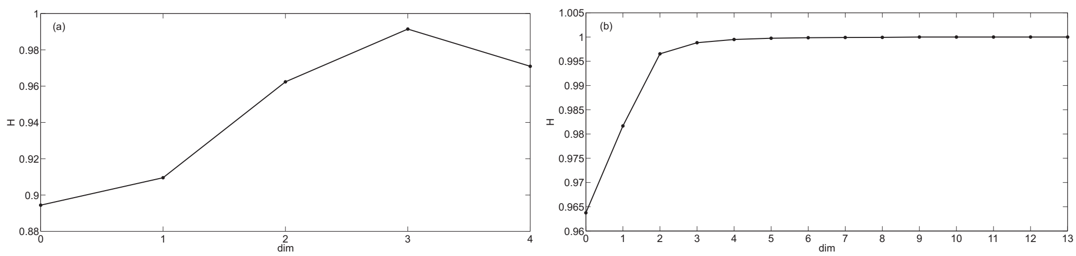


Figure 6.13: Spectral entropies of clique complex (a) and its conjugate (b) of the Zachary karate club network.

Chapter 7

The role of simplicial complexes in the study of complex systems and perspectives for future research

In this chapter we discuss some of the less familiar applications of simplicial complexes in physics in general and in the field of complex systems in particular. The list of applications is certainly not exhaustive as, among other reasons, new areas of utilization of topology in physics and other sciences appear continually. The intent is to be rather informative than to present a rigorous treatment of the subject.

7.1 Differential forms and simplicial complexes

We have already emphasised in the Introduction the importance of discrete differential forms in the modern theoretical physics research and their relationships with simplicial complexes [30]. Let us recall the main highlights: simplicial complexes are discretized manifolds, and cochains on such simplicial complexes are discrete differential forms. Since coboundary operator on simplicial complex can be related to the discrete exterior derivative, higher-order combinatorial Laplacian represents a discretized Laplace-Beltrami operator on a manifold. Simplicial complexes, and as well higher-order combinatorial Laplacians, exist independently on the discretization procedure, as we have shown in preceding chapters, it would be insightful to examine the possibility of the reverse procedure. Namely, the reverse procedure would imply the reconstruction of a manifold and obtaining discrete differential forms from the knowledge of the simplicial complex and its structure. From graph theory, the graph Laplacian captures the structural properties of the complex network, and proves to be convenient for the examination of dynamical processes taking place on complex networks [66]. Likewise, we can anticipate that higher-order combinatorial Laplacians, acting as operators on the discrete differential forms, should be convenient for the examination of dynamical processes on higher-order (sub)structures of complex networks captured by diverse simplicial complexes. It is important to point that some attempts in this course already present [97].

7.2 Percolations on simplicial complexes

The percolation theory [98] occurs as an adequate framework for the examination of the critical phenomena taking place on complex networks [99]. Since complex networks are highly irregular, the computation of percolating clusters is not simple, though some algorithms are developed [100]. A new algorithm could be developed by calculating a Q-vector of the neighborhood complex obtained from complex network, in which case the connectivity classes at the 0-connectivity level represent percolative clusters. Hence, the procedure should be as followed: calculate the Q-vector of a neighborhood complex before occupying nodes, and then, whenever the nodes of a network are occupied, the Q-vector entries are updated. The advantage of this procedure is twofold: first, the algorithm for the percolating clusters originates from the well defined mathematical quantity, and second, we have an insight into the changes of topological properties depending on the occupation probability.

The above example of the usage of the Q-vector entries is limited to percolations on complex networks. Nevertheless, the types of percolation processes on simplicial complexes are far more richer, to mention a few:

- vertices can be occupied;
- simplices can be occupied and vertices can be left unoccupied, unless all simplices which posses them are occupied;
- simplices and vertices are occupied simultaneously, possible with different probabilities;
- the same as the above, but for a conjugate complex.

These rather complex processes is not easy to analyze, and one possible formalism for their analysis can be the so called "algebra of patterns" introduced by Atkin [76], [77], [78] for the examination of changes on and of simplicial complexes.

7.3 Time evolution of vector-valued quantities

In Chapter 2 the vector-valued quantities like Q-vector, the second and the third structure vectors, have been introduced for the characterization of the topological properties of simplicial complexes. Nevertheless, since real-world complex networks are changing, like adding or deleting nodes and/or links, simplicial complexes obtained from complex networks experience changes also. Therefore, vector-valued quantities experience changes too and it would be useful to explore eventual change patterns through the q-levels. This problem is in a sense related to the previous one (percolations) since adding or deleting nodes and/or links can be treated as percolation process.

7.4 Foundations of statistical mechanics of simplicial complexes

In the preceding chapters we have seen that some properties of complex networks are preserved, such as the behavior of the degree distribution, with algebraic topological properties of simplicial complexes obtained from those networks. Whereas attempts for precise formulation of statistical mechanics of complex networks are successful [5], [66], [101] the field of statistical mechanics of simplicial complexes is still in its development. The multidimensional and sophisticated structure of simplicial complexes requires different approach than the one for the complex networks and in turn offers many advantages and more sophisticated analysis of complex phenomena.

7.5 Construction of simplicial complex from correlation matrix

Reconstructing a graph from the correlation matrix corresponding to the dynamics of a certain complex system is rather arbitrary, and is based on some threshold criterion on the values of the correlation matrix entries. The general requirement for building a graph from correlation matrix entries is the one which maximizes the sum of the correlations over the connected edges [102]. The simplest method for extracting a connected graph with all nodes involved is the so called Minimum Spanning Tree (MST) method [102], by which with respect to the above requirement the resulting graph has a tree structure, i.e., does not have cycles, or triangles. As a generalization of MST method the so called Planar Maximally Filtered Graph (PMFG) [89], which captures more edges, contains cliques, and contains a MST as a subgraph, was proposed. In the Chapter 3 we have used as an example of neighborhood complex brain networks obtained from correlation matrix using the PMFG. In either case the simplicial complex is obtained from the correlation matrix indirectly, from a graph. Hence, for capturing algebraic topological properties of a complex system whose information is stored in the correlation matrix, it is important to develop a method for construction of simplicial complex directly from a correlation matrix, depending on the correlation matrix entries.

7.6 Weighted simplicial complexes

Related to the previous section is the problem of topological properties of the weighted simplicial complex. Namely, a large number of real world complex networks, and as well simplicial complexes, are characterized by association of certain numerical values to the nodes and/or the links of a network, and hence to the vertices, the faces and the simplices of the obtained simplicial complex. Although there are some attempts in the course of weighted combinatorial Laplacians [67], [68] or weighted simplicial homology development [103], the methodology applied to the wide class of complex systems is still lacking.

Chapter 8

Concluding remarks

Starting from a typical properties of complex systems, the notion of (sub)structure of complex networks was redefined and combinatorial algebraic topology was proposed as an adequate mathematical framework for its study. Within the context of algebraic topology simplicial complexes have been used for the modeling of the structure of complex networks, and accordingly, simplicial communities have been introduced to capture substructure properties. It turns out that thanks to the versatility of simplicial complexes which can be constructed from a single network, different hidden organizational patterns formed by the relationships between simplicial communities within complex network can be detected.

The neighborhood complex obtained from diverse modeled and real-world complex networks shows that statistical properties of complex networks have been preserved when we do the transition from the graph to simplicial complex representation. Since the simplices in neighborhood complex are the collections of nodes and the dimension of a simplex is equal the degree of corresponding node, it indicates the importance of the taking into account such substructures in complex networks' evolution models.

The usefulness and adequacy of simplicial complexes to model qualitative features of complex networks is demonstrated in the case of clique complexes of two well known social networks, the Zachary karate club and the co-appearance network of characters in the Victor Hugo's novel Les Miserables. It appears that an approach using both the initial and the conjugate clique complex is particularly informative for the analysis of social networks. The spectra of the combinatorial Laplacian of simplicial complexes are deeply related to the connectivity properties of simplicial complexes and hence not related to their geometrical features. Based on the concept of simplicial community, communities and their interlacing is well reflected in the spectra and provides precise information on the occurrence of such complex structures. Connectivity properties, well captured by the Q-vector, complement the spectra providing hierarchical description of simplicial complexes and complex networks. Finally, it has been demonstrated that although extensively studied and used in many areas of science, properties of ordinary graph Laplacian are in certain cases not adequate to capture higher-order information that resides in complex networks.

At the end we can conclude that the results of thesis indicate that striving toward the consistent and precise theory of complexity can without a doubt be supported by the formalism coming from the algebraic topology using simplicial complexes as working objects. Furthermore, the rigorous and sometimes seemingly rigid and abstract approach of combinatorial algebraic topology proved to be an advantage rather than the obstacle, and that the ”simplexication” of sciences far transcends physics.

Bibliography

[1] H. E. Stanley, Scaling, universality, and renormalization: Three pillars of modern critical phenomena, Rev. Mod. Phys. **71**, S358, 1999

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

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

[4] W. Weaver, Science and complexity, American Scientist **36**, 536-544, 1948

[5] R. Albert, A.-L. Barabási, Statistical mechanics of complex networks. Rev. Mod. Phys. **74**, 4797, 2002

[6] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Complex networks: Structure and Dynamics, Phys. Rep. **424**, 175-308, 2006

[7] M. E. J. Newman, The structure and function of complex networks. SIAM Rev. **45**, 167256, 2003

[8] P. W. Anderson, More is different, Science **177**, 393-396, 1972

[9] H. A. Simon, The Architecture of Complexity, Proceedings of the American Philosopical Society **106**, 467, 1962

[10] D. Kozlov, Combinatorial Algebraic Topology, (Algorithms and Computation in Mathematics, Springer-Verlag, Berlin Heidelberg, 2008)

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

[12] R. H. Atkin, Combinatorial Connectivities in Social Systems (Birkhäuser Verlag, Base und Stuttgart, 1977)

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

[14] R.H. Atkin, Mathematical structure in human affairs, (Heinemann, London 1974)

[15] P. Gould, J. Johnson, G. Chapman, The Structure of Television, (Pion Limited, London, 1984)

[16] T. L. Jacobson, W. Yan, Q-Analysis Techniques for Studying Communication Content, Quality & Quantity **32**, 93-108, 1998

[17] S. B. Seidman, Rethinking backcloth and traffic: prespectives from social network analysis and Q-analysis, Environment and Planning B **10**, 439-456, 1983

[18] L. C. Freeman, Q-analysis and the structure of friendship networks, Int. J. Man-Machine Studies **12**, 367-378, 1980

[19] P. Doreian, Polyhedral Dynamics and Conflict Mobilization in Social Networks, Social Networks **3**, 107-116, 1981

[20] P. Doreian, Leveling coalitions as network phenomena, Social Networks **4**, 27-45, 1982

[21] R. H. Atkin, J. Johnson and V. Mancini, An analysis of urban structure using concepts of algebraic topology, Urban Studies **8**, 221-242, 1971

[22] J. H. Johnson, The q-analysis of road intersections, Int. J. Man-Machine Studies **8**, 531-548, 1976

[23] J. C. Griffiths, Geological Similarity by Q-Analysis, Mathematical Geology **15**, 85, 1983

[24] L. Duckstein, Evaluation of the Performance of a Distribution System by Q-Analysis, Applied Mathematics and Computation **13**, 173-185, 1983

[25] L. Duckstein, S. A. Nobe, Q-analysis for modeling and decision making, European Journal of Operational Research **103**, 411-425, 1997

[26] Y. Ishida, N. Adachi, H. Tokumaru, Topological approach to failure diagnosis of large-scale systems, IEEE Trans. Syst., Man and Cybernetics **15**, 327-333, 1985

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

[28] H. Flanders, Differential Forms with Applications to the Physical Sciences, (New York: Academic Press, 1963)

[29] T. Frankel, The Geometry of Physics: An Introduction (Cambridge University Press, Cambridge, 1997)

[30] H. Eschrig, Topology and Geometry for Physics, (Lecture Notes in Physics, 822 Springer-Verlag Berlin Heidelberg 2011)

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

[32] C. W. Misner, K.S. Thorne and J. A. Wheeler, Gravitation, (W.H. Freeman, San Francisco 1973)

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

[34] A. Bossavit, Computational Electromagnetism : Variational Formulations, Complementarity, Edge Elements, (Academic Press, 1998)

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

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

[37] H. W. Hamber, Simplicial quantum gravity, in: Critical phenomena, random systems, gauge theories, (North-Holland, Amsterdam, 1986)

[38] T. Regge, General relativity without coordinates, Nuovo Cimento **19**, 558, 1961

[39] G. A. Deschamps, Electromagnetics and Differential Forms, Proceedings of IEEE **69**, 676, 1981

[40] F. L. Teixeira. Geometric aspects of the simplicial discretization of Maxwells equations. In F. L. Teixeira, editor, Geometric Methods for Computational Electromagnetics, chapter 7. EMW Publishing, 2001

[41] E. Tonti. Finite formulation of electromagnetic field. IEEE Trans. Mag. **38**, 333336, 2002

[42] R. W. Sharpe, *Differential Geometry: Cartan's genralization of Klein's Erlangen Program*, (Springer-Verlag, New York, 1997)

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

[44] R. H. Atkin, Abstract physics, Nuovo cimento **38**, 496, 1965

[45] R. H. Atkin and T. Bastin, A Homological Foundation for Scale Problems in Physics, International Journal of Theoretical Physics **3**, 449, 1970

[46] E. Tonti, The reason for analogies between physical theories, Appl. Math. Modelling **1**, 37, 1976

[47] E. Tonti, A direct discrete formulation of field laws: The cell method, Comput. Model. Eng. Sci. **2**, 237258 2001

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

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

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

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

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

[53] M. Girvan, M. E. Newman, Community structure in social and biological networks. Proc. Natl. Acad. Sci. USA **99**, 78217826, 2002

[54] S. Fortunato, Community detection in graphs, Phys. Rep. **486**, 75, 2010

[55] M. E. J. Newman, Communities, modules and large-scale structure in networks, Nature Physics **8**, 25, 2012

[56] Y.-Y. Ahn, J. P. Bagrow, S. Lehmann, Link communities reveal multiscale complexity in networks, Nature **466**, 761, 2010

[57] T. S. Evans, R. Lambiotte, Line graphs, link partitions, and overlapping communities, Physical Review E **80**, 016105, 2009

[58] T. S. Evans, Clique graphs and overlapping communities, Journal of Statistical Mechanics: Theory and Experiment P12037, 2010

[59] G. Palla, I. Derényi, I. Farkas, T. Vicsek, Uncovering the overlapping community structure of complex networks in nature and society, Nature **435**, 814, 2005

[60] G. Palla, A.-L. Barabási, T. Vicsek, Quantifying social group evolution, Nature **446**, 664, 2007

[61] T. Nishikawa, A. E. Motter, Discovering Network Structure Beyond Communities, Scientific Reports **1**, 151, 2011

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

[63] T. E. Goldberg, Combinatorial Laplacians of Simplicial Complexes, (Annandale-on-Hudson, New York, 2002)

[64] S. Maletić, D. Horak, M. Rajković, Cooperation, conflict and higher-order structures of complex networks, Advances in Complex Systems **15**, 1250055, 2012

[65] B. Mohar, The Laplacian spectrum of graphs, Graph Theory, Combinatorics, and Applications, **2**, 871, 1991

[66] M. E. J. Newman, Networks: An Introduction, (Oxford University Press, USA, 2010)

[67] D. Horak and J. Jost, arXiv: 1105.2712.

[68] D. Horak, arXiv: 1111.1836.

[69] G. Bianconi, Entropy of randomized network ensembles, Europhys. Lett. **81**, 28005, 2008

[70] G. Bianconi, A.C.C. Coolen and C. J. Perez Vicente, Entropies of complex networks with hierarchically constrained topologies, Phys. Rev. E **78**, 016114, 2008

[71] G. Bianconi, Entropy of network ensembles, Phys. Rev. E **79**, 036114, 2009

[72] S. L. Braunstein, S. Ghosh, S. Severini, The Laplacian of a Graph as a Density Matrix: A Basic Combinatorial Approach to Separability of Mixed States, Annals of Combinatorics **10**, 291, 2006

[73] F. Passerini, S. Severini, The von Neumann entropy of networks, arXiv:0812.2597v1

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

[75] J. R. Munkres, Elements of Algebraic Topology, (Addison-Wesley Publishing, California, 1984)

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

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

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

[79] J. H. Johnson, q-Transmission in simplicial complexes, Int. J. Man-Machine Studies **16**, 1982, 351-377

[80] J. H. Johnson, Hypernetworks of Complex Systems, Complex 2009, Part I, LNICST **4**, 2009, 364-375

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

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

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

[84] A. Banerjee, J. Jost, Graph spectra as a systematic tool in computational biology, Discrete Applied Mathematics **157**, 2425, 2009

[85] F. R. K. Chung, *Spectral Graph Theory* ((CBMS Regional Conference Series in Mathematics, No. 92) American Mathematical Society, 1996)

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

[87] M. E. J. Newman, Finding community structure in networks using the eigenvectors of matrices, Phys. Rev. E **74**, 036104, 2006

[88] O. Sporns, Networks of the Brain, (The MIT Press, Cambridge, Massachusetts, 2011)

[89] M. Tumminello, T. Aste, T. Di Matteo, and R. N. Mantegna, A tool for filtering information in complex systems, PNAS **102**, 1042110426, 2005

[90] C. Bron, J. Kerbosch, Finding all cliques of an undirected graph, Comm. ACM **16**, 575, 1973

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

[92] H. Jeong, S. P. Mason, A.-L. Barabási, Z. N. Oltvai, Lethality and centrality in protein networks, Nature **411**, 41 2001

[93] M. E. J. Newman, S. H. Strogatz, D. J. Watts, Random graphs with arbitrary degree distributions and their applications, Phys. Rev. E **64**, 026118, 2001

[94] http://vlado.fmf.uni-lj.si/pub/networks/data

[95] W. W. Zachary, An information flow model for conflict and fission in small groups, Journal of Anthropological Research **33**, 452, 1977

[96] D. E. Knuth, The Stanford GraphBase: A Platform for Combinatorial Computing, (Addison-Wesley, Reading, MA, 1993)

[97] A. Muhammad, M. Egerstedt, Control Using Higher Order Laplacians in Network Topologies, Proceedings of the 17th International Symposium on Mathematical Theory and Systems, Kyoto, Japan, 1024, 2006

[98] D. Stauffer, Introduction to Percolation Theory, (Taylor and Francis, London and Philadelphia, 1985)

[99] S. N. Dorogovtsev, A. V. Goltsev, J. F. F. Mendes, Critical phenomena in complex networks, Rev. Mod. Phys. **80**, 12751335, 2008

[100] A. Al-Futaisi, T. W. Patzek, Extension of Hoshen-Kopelman algorithm to non-lattice environments, Phys. A **321**, 665, 2003

[101] J. Park and M. E. J. Newman, Statistical mechanics of networks, Phys. Rev. E **70**, 066117, 2004

[102] R. N. Mantegna, Hierarchical structure in financial markets, Eur. Phys. J. B **11**, 193, 1999

[103] R. J. M. Dawson, Homology of weighted simplicial complexes, Cahiers de Topologie et Géométrie Différentielle Catégoriques **31**, 229, 1990

Biografija

Slobodan Maletić rođen je u Beogradu 18. 08. 1976. godine, gde je završio osnovnu i srednju školu (Gimnazija, prirodno-matematički smer). Diplomirao je 2005. godine na Fizičkom fakultetu Univerziteta u Beogradu, smer Teorijska fizika, odbranivši diplomski rad pod nazivom "Otvorene hamiltonovske šetnje na fraktalnim rešetkama". Poslediplomske studije upisuje 2005. godine na Fizičkom fakultetu Univerziteta u Beogradu, smer Teorijska fizika kondenzovanog stanja materije. Magistrirao je 2009. godine odbranivši magistarsku tezu pod nazivom "Statistička mehanika simplicijalnih kompleksa" i iste godine upisuje treću godinu doktorskih studija na Fizičkom fakultetu Univerziteta u Beogradu. U maju 2012. godine Nastavno-naučno veće Fizičkom fakultetu Univerziteta u Beogradu odobrilo mu je izradu doktorske teze pod nazivom "Simplicijalni kompleksi i kompleksne mreže: uticaj (pod)struktura višeg reda na karakteristike mreže". Od 2006. godine zaposlen je kao istraživač-pripravnik, a od 2009. godine kao istraživač-saradnik Laboratorije za termotehniku i energetiku Instituta za nuklearne nauke "Vinča".

Прилог 1.

Изјава о ауторству

Потписани мр Слободан Малетић

број индекса D7/2009

Изјављујем

да је докторска дисертација под насловом

Симплицијални комплекси и комплексне мреже: утицај (под)структура вишег реда на карактеристике мреже

- резултат сопственог истраживачког рада,
- да предложена дисертација у целини ни у деловима није била предложена за добијање било које дипломе према студијским програмима других високошколских установа,
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио интелектуалну својину других лица.

Потпис докторанда

У Београду, 27. мај 2013. год.

Прилог 2.

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора_____мр Слободан Малетић

Број индекса ________________D7/2009____________________________

Студијски програм ____

Наслов рада: Симплицијални комплекси и комплексне мреже: утицај (под)структура вишег реда на карактеристике мреже

Ментор ___др Милан Рајковић___

Потписани ___мр Слободан Малетић

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/~~ла~~ за објављивање на порталу **Дигиталног репозиторијума Универзитета у Београду.**

Дозвољавам да се објаве моји лични подаци везани за добијање академског звања 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 докторанда

У Београду, __27. мај 2013. год.__

____

Прилог 3.

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

Симплицијални комплекси и комплексне мреже: утицај (под)структура вишег реда на карактеристике мреже

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигитални репозиторијум Универзитета у Београду могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство

2. Ауторство - некомерцијално

3. Ауторство – некомерцијално – без прераде

4. Ауторство – некомерцијално – делити под истим условима

5. Ауторство – без прераде

6. Ауторство – делити под истим условима

(Молимо да заокружите само једну од шест понуђених лиценци, кратак опис лиценци дат је на полеђини листа).

Потпис докторанда

У Београду, _27. мај 2013. год.__

____

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 прераде. 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 прераде. 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.