Polytechnic University of Catalonia

Department of Statistics and Operational Research

PhD Thesis

Mathematical programming based approaches for classes of complex network problems

Economical and sociological applications

Author:

Stefano Nasini

Supervisor:

Jordi Castro Perez

December 10, 2014

To my parents

Contents

Acknowledgements

Foremost, this thesis would not have been possible without the aid of several people, whose continuous support allowed me a constant commitment to my Ph.D study and research.

I would like to express my sincere gratitude to my supervisor, Dr. Jordi Castro, who shared with me a lot of his knowledge and scientific expertise. His guidance helped me in all the time of research and writing of this thesis.

A special thanks must be addressed to Prof. Pau Fonseca i Casas, whose technical support gave rise to substantial mathematical and computational advances in several parts of the thesis.

Besides my advisor and tutor, I would like to thank Prof. Francisco Saldanha da Gama, for offering me the summer internship opportunities in his group and leading me working on diverse exciting projects.

I wish to thank to the whole research group of numerical optimization and modeling (GNOM), which provided me with remarkably useful technical support during the long journey of my Ph.D. project. It was a pleasure to work with, and to be part of, this group.

I thank my colleagues and officemate in the Polytechnic University of Catalonia: Francesc Lopez Ramos, Hajar Nasr and Nima Rabei, for the stimulating discussions, for the sleepless nights we were working together, and for all the fun we have had in the last years.

Most importantly, a special thanks must be addressed to my entire family for providing a loving environment for me: my parents Flavio Nasini and Simonetta Ferracuti, for giving birth to me at the first place and supporting me spiritually throughout my life; my wife Yolima del Rocio Marin Valencia, who was really patient during these years and inciting me into a full commitment to this thesis.

In conclusion, I recognize that this research would not have been possible without the financial assistance of the FPI-UPC grant and the grants MTM2012-31440 of the Spanish research program and SGR-2009-1122 of the Government of Catalonia.

Symbols and abbreviations

List of symbols

- $\mathbb{R}$ The set of real numbers;
- $\mathbb{Z}$ The set of integer numbers;
- $\mathbb{N}$ The set of natural numbers;
- $\mathbb{R}_{+\cup\{0\}}$ The set of non negative real numbers;
- $\mathbb{Z}_{+\cup\{0\}}$ The set of non negative integer numbers;
- $\mathcal{G}$ A graph or network;
- $\mathcal{V}$ A set of nodes;
- $\mathcal{E}$ A set of edges;
- χ A set of networks;
- $||\mathbf{x}||_p$ The p-norm of a vector $\mathbf{x}$;
- $||A||_p$ The p-norm of a matrix A;

List of abbreviations

- AM Adjacency Matrix;
- B&B Branch and Bound;
- CN Complex Network;
- DN Directed Network;
- ECDN Edge-colored directed networks;
- ECUN Edge-colored undirected networks;
- ER Elementary Reallocation;

- ERGM Exponential Random Graph Model;
- ERP Elementary Reallocation Problem;
- IP Interior Point;
- IPM Interior Point Method;
- LP Linear Program;
- MSS Mean Similarity Spectrum;
- SER Sequence of Elementary Reallocations;
- VSS Variance Similarity Spectrum;
- UN Undirected Network;
- PCG Preconditioned Conjugate Gradient;

Preface

This Ph.D. thesis contains the main results of the mathematical and computational investigation carried out at the Department of Statistics and Operation Research of the Polytechnic University of Catalonia (Barcelona) by the last years. It was realized within the field of Complex Network Theory and its applications to Social and Economic Sciences, resulting in a fruitful attempt to conjugate novel mathematical and computational methodologies into the analysis of economic and social phenomena.

The mathematical analysis of Complex Networks (CN from now on) and its sociological translation into Social Network Analysis (SNA from now on) has recently gained a great attention among mathematicians, economists and sociologists, partially due to the powerful methodologies derived from the Graph Theory, the Statistical Inference, the Computer Simulation and the Numerical Optimization. These classes of methodologies bring together mathematical and computational tools, whose applications resulted highly effective when dealing with transportation networks, routing, logistic, epidemiology, etc. During the past three decades, however, the scientific interest substantially moved toward applications of those methodologies outside these traditional areas, leading to a new cross-disciplinary effort in the analysis of large-scale economical and sociological phenomena (social segmentation, economical stratification, dissemination of culture and social learning, etc.).

In the very beginning of the Ph.D. project, my main concern was the existence of several heterogeneous (often poorly related) methodologies of SNA; the lack of a common methodological view in the state–of–the–art publications suggested me the possibility of analyzing and developing a systematic and comprehensive framework, which mathematically unifies many of those miscellaneous methods. Part of this intention has been captured in the design of a common Mathematical Programming based approach for both *stochastic* and *strategic* models of network formation, as discussed in the second and third parts of the thesis. The evolving of the researches and the coming up of interesting computational results, during the second year of the Ph.D. project, led this analytic effort toward a substantial prominence of the Mathematical Programming approaches, among the aforementioned classes of methodologies. In this respect, the appearance of algebraical relations between many classes of CN problems casted a light into the possibility of reformulating traditional and well-studied methods, with the aim of establishing cross-disciplinary connections between far apart areas of the SNA. The first preliminary results in this direction suggested the existence of a fruitful and poorly-explored line of research concerning the development and formulation of mathematical programming based approaches for many classes of CN problems, which have been traditionally studied since the very beginning of Social and Economic Sciences.

Thinking in terms of networks and in terms of discrete relation between objects is a relatively modern conceptual approach in science, whose legendary beginning is related to be the famous Knigsberg bridge problem, proposed by Euler [87] in the 1736: *Does there exist a walk crossing each of the seven bridges of Knigsberg exactly once?* It took two centuries before the first book

on Graph Theory was written by Konig in 1936 [143] and only after the World War II the use of CNs to model large-scale economical and sociological phenomena established itself as a powerful approach for mapping and measuring relationships and flows between intelligent agents, such as animals, organizations and other biological entities.

One of the first contributions in this direction came from Rashevsky [195, 196], who developed a program in *mathematical biophysics* at the University of Chicago during the 1940s, and by Harary [113], who provided a strong mathematical support to the study of negative and positive social relations[1].

The spring of another fundamental chapter in the history of CNs and SNA started with the publication of Granovetter's seminal paper, *'The Strength of Weak Ties'* [108], which will be discussed in different parts of the thesis. He focused on the study of common properties observed in real-world social networks and suggested a flexible and general approach for the analysis of valued networks. *'The Strength of Weak Ties'* resulted in a highly influential sociological paper, with over 29.000 citations according to Google Scholar (by September 2014), which ushered a prominent chapter of SNA, dealing with valued connections and interactions between social agents [19–21, 82, 129, 141, 150, 164].

More recently, the study of the small-world phenomenon – which is somehow contiguous to the analysis introduced by Granovetter – has gained a great attention among mathematicians, economists and sociologists, starting from the seminal work of Watts and Strogatz [226]. Their studies made use of a range of probabilistic approaches to explain the emergence of the small-world phenomenon in real-world social networks and proposed different classes of random graph generators whose underlying processes somehow emulated the structural properties of the small-world phenomenon (high transitivity and high connectivity).

Classes of CN problems related to these two lines of research – the one introduced by Granovetter [108] and the other introduced by Watts and Strogatz [226] – will be taken into account in the first part of this thesis (chapters 1 and 2), by means of specialized methodologies based on Spectral Theory and Optimization. In the original formulation of this approach, our intention was mainly oriented toward the descriptive analysis of structural properties and the problem of low-dimensional representation of network structures. However, the coming up of computational results during the third and fourth year of my Ph.D. research, led to the application of the aforementioned Spectral Theory based approach to the problem of assessing the goodness of fit of random network models, as discussed in Chapter 5.

In the classical literature of SNA, the study of structural properties has been usually carried out by applying two conceptually different modeling approaches: *random models of network formation* and *strategic models of network formation*. Jackson [121] highlighted the fact that while Random Graphs are helpful in algebraically capturing certain features of CNs, they lack the explanation of local decisions of independent agents which entail the emergence of the overall network structure.

The study of random models of network formation will be taken into account in the second part of this thesis (chapters 3, 4 and 5), by means of a "massive" application of methodologies and techniques coming from Combinatorial Optimization, Linear Programming and Interior Point Methods. On the other hand, strategic models of network formation – which can also be thought as *optimization-based models of network formation*, in the sense that they strictly take into account the optimization criteria that given observed networks verify – will be studied in the third part of this thesis (Chapter 6).

Those applications will be mainly oriented to the development of statistical simulation pro-

[1] Harary's Structure Theorem says that if a network of interrelated positive and negative ties is balanced, e.g. as illustrated by the principle that *'my friend's enemy is my enemy'*, then it consists of two subnetworks such that each has positive ties among its nodes and there are only negative ties between nodes in distinct subnetworks.

cedures to sample from "highly combinatorial" spaces. The need of statistical simulation in the study of random networks is due to the complex properties of these discrete mathematical objects, which might be hardly analyzed without the aid of efficient computational methods. The construction and development of general simulation procedures represent indeed the main purpose of the second part of this Ph.D. thesis. From this outlook, two aspects will appear of particular importance:

i. the algebraic characterization of families of networks with fixed structural properties by systems of linear constraints;

ii. the construction of Mathematical Programming methods to provide random basic feasible solutions of those systems.

A lot of researches are currently struggling to find polynomial algorithms to deal with big network problems [1], particularly in transportation networks, routing, logistic, epidemiology, etc. Nonetheless, most of them are not concerned with application in statistical simulation, which represents a fundamental tool for random models of network formation [165, 216]. The second part of this thesis will try to fill this gap, by initiating a novel area of application of the general Optimization Theory in different classes of probabilistic models of CNs.

To summarize, the thesis is arranged in the following three parts.

Part I *The fine-grained structure of complex networks: theories, models and methods;*

Part II *Mathematical Programming based approaches for random models of network formation;*

Part III *Strategic models of network formation.*

The computer implementation of the described numerical procedures has been realized by Java, MatLab, R and AMPL. The runs were carried out on a Fujitsu Primergy RX300 server with 3.33 GHz Intel Xeon X5680 CPUs (24 cores) and 144 GB of RAM, under a GNU/Linux operating system (Suse 11.4), without exploitation of multithreading capabilities.

The computer implementations toke up a remarkable part of this thesis and perhaps they might been further improved in the future, by using low level languages. Preferably, I would have liked to have studied more computational courses, with the aim of getting a higher methodological freedom and fluency when dealing with computer implementations. However, what I have learned since the very beginning of my education is that computational methods are indispensable but they can't make sense of observed data; only our modeling strategies can, as they take place in the mind of the researcher, not in the computer. Because of this, the main efforts have been addressed to the development of a rigorous mathematical programming based framework, which is able to capture many classes of CN problems, whereas the computer implementation have been carried out by high level programming languages, such as Java, MatLab, R and AMPL.

The researches in the context of this Ph.D. thesis gave rise to the following publications in peer-reviewed journals, scientific conferences and research reports

Peer-reviewed publications

- Castro J., Nasini S., (2014), On geometrical properties of preconditioners in IPMs for classes of block-angular problems, to be submitted to *Mathematical Programming*.
  - Corresponding to Chapter 4.

- Nasini S., Castro J., Fonseca P., (2015), A Mathematical programming approach for different scenarios of bilateral bartering, accepted to *SORT-Statistics and Operation Research Transactions*.
  - Corresponding to Chapter 6.
- Nasini S., Castro J., Fonseca P., (2013), Novel representation of network structures by spectral theory consideration, under review in *Journal of Social Networks*.
  - Corresponding to Chapter 2.
- Castro J., Nasini S., (2013), Mathematical programming approach for classes of random network problems, under review in *European Journal of Operation Research*.
  - Corresponding to chapters 3 and 5.

Scientific conferences

- Nasini S., Castro J., Specialized interior point methods for classes of random network problem, *20th Conference of the International Federation of Operational Research Societies IFORS*, Polytechnic University of Catalonia, Barcelona, Catalonia, July 2014. Invited presentation.
  - Corresponding to chapters 4 and 5.
- Nasini S., Castro J., Preconditioning IPMs for block-angular problems with "almost linearly dependent" constraints, *International Conference on Applied Mathematical Programming and Modelling APMOD*, University of Warwick, Warwick, United Kingdom, April 2014. Invited presentation.
  - Corresponding to Chapter 4.
- Nasini S., Castro J., Generating random networks by linear programming approaches, *Joint International 26th European Conference on Operational Research (EURO 2013)-INFORMS*, Rome, Italy, July 2013. Invited presentation.
  - Corresponding to chapters 3 and 5.
- Nasini S., Castro J., Generating conditional uniform random networks by optimization procedures, *International Network Optimization Conference 2013*, Tenerife, Spain, May 2013.
  - Corresponding to Chapter 3.

Research reports

- Nasini S., (2014), Maximizing graph probability under conditionally exponential models, *Arxiv-Cornell University Library*, available at `http://arxiv.org/abs/1409.5476`.
  - Corresponding to sections 2.4 and 5.5.
- Nasini S., Castro J., Fonseca P., (2014), Bartering integer commodities with exogenous prices, accepted by *Arxiv-Cornell University Library*, `http://arxiv.org/abs/1401.3145`.
  - Corresponding to Chapter 6.

I hope this thesis might be able to give the reader a clear idea about the great variety of the existing modeling possibilities, when dealing with social and economic phenomena. Preferably, it would have been much easier for me to have followed a more canonical and confirmatory research line, rather than such an explorative journey into the great variety of methodological possibilities that Mathematical Programming can provide to the field of CNs. Nonetheless, the main advances and progresses in Science are achieved when many small attempts of deviating the canonical ways of thinking result in a unified and consistent paradigm. Because of this, some degree of cross-disciplinary and exploratory strategy is absolutely needed in Science to avoid getting stuck into "stagnant local optima".

Part I

The fine-grained structure of complex networks: theories, models and methods

Abstract

This part of the thesis includes two chapters, whose primary scope is to establish methodological and empirical connections between different fields of complex network theory, laying the foundation of the mathematical programming-based approaches described and used throughout the thesis. From an economical outlook, Chapter 1 investigates different methodological approaches to deal with the seminal problem of *structure and agency* in social and economic theory (the interdependence between the individual actions and the macroscopical social phenomena), taking into account the effect that a complex structure of social relations might have on relevant economical phenomena, such as commodity production and distribution. A critical cross-disciplinary overview of the state–of–the–art will be provided and compared with the proposed mathematical programming framework. Based on these considerations, Chapter 2 introduces a novel methodological approach for the problem of *low-dimensional representations of networks* and proposes a general Spectral Theory based method to summarize structural network properties, which is valid for both cases of binary and valued networks. The numerical validation of the correctness of the described methodologies is based on both optimization-based models and probabilistic models of network formation, resulting in a cross-fertilization between areas of network analysis, which are usually treated as separated research fields: structural similarity, network centralization, assortative mixing, triadic closure and community structure.

Keywords: Complex Networks, Social Network Analysis, Mathematical Programming, Spectral Theory, Structural Similarity.

Chapter 1

Overview and preliminaries

1.1 Complex Networks: theories, models and methods

1.1.1 The starting problems of network-based models in economical and social sciences

The past few decades have seen an increasing interest in the application of the Theory of Networks (or Graph Theory), outside the traditional areas of transportation, routing, logistic and epidemiology, modeling a great variety of large-scale economical and sociological phenomena (social segmentation, economical stratification, dissemination of culture and social learning, etc.).

In its basic notion, the interconnected objects of a network (or graph) $\mathcal{G}$ are represented by a finite set of elements, say $\mathcal{V}(\mathcal{G})$, conventionally named vertices or nodes, and the links that connect some pairs of vertices, called edges, say $\mathcal{E}(\mathcal{G}) \subseteq \mathcal{V}(\mathcal{G}) \times \mathcal{V}(\mathcal{G})$. Generally, these mathematical objects are graphically depicted as a collection of dots for the nodes, joined by lines for the edges.

Extensions of this basic framework, designed to accommodate complex real-world situations, are many and varied. One possibility is to avoid the assumption of dichotomous relationships by allowing edges to carry numerical or categorical labels [108, 150, 164]. Another proposed extension was to introduce multilateral relationships (such as group memberships), as in the case of hyperedges involving arbitrarily many nodes [224]. It must be stressed that the increase of complexity often resulted in less robust models, with high sensitivity on the researcher's assumptions, or poorly informative about the specific question that the researcher wanted to address. In fact, it is precisely the naive nature of graphical structures that has facilitated the achievement of a high mathematical development [30] and technological applications [171] of network-based approaches.

One of the main line of application, which is of particular interest in the context of this thesis, is given by different classes of multi-agent simulation, where social and biological agents (individuals, groups, countries, etc.) are linked in complex connection structures.

Suppose, for instance, we had information about trade-flows of m different commodities among n nations of the world. Here, the n nations can be thought of as nodes, and the amount of each commodity exported from each nation to each of the other $n-1$ can be thought of as

the strength of a directed tie from the focal nation to the other. A social scientist might be interested in studying what are the most significative factors which effect trades among nations and which particular structure of connections has been built by the different nations. The answer might probably depend on the mathematical formalization of the problem and specifically by the formal definition of nations and trades.

As argued by Butts [45], two question should be addressed before dealing with applications in this area:

i. *when is a node a node?*

ii. *when is an edge an edge?*

The answer of these questions has not only a theoretical importance but also effective consequence on modeling decisions and interpretation of the results of networks analysis. In SNs the nodes might be either individual humans [120, 224] or aggregates such as groups, households, or organizations [18, 58, 145]. Any change of the node set can substantially influence the size, density and topology of the resulting network, with considerable implications for subsequent analysis, as different aggregation decisions can produce networks with very different structural features.

A similar reasoning is valid when taking into account the set of dyadic links, as noted by Borgatti et al. [37]. In the above example concerning trades among nations, the inclusion of the different components of the balances of payments might result in different definition of dyadic links. This fact will be clear in section 5.4 and 6.3, where different definition of dyadic links will give rise to different numerical results in the statistical analysis.

Besides this problem of node and edge definition, an open issue (a prominent line of research) which must be mentioned is associated to the mathematical and computational difficulties of dealing with network data sets. Social networks are usually big in size and the problem of describing and managing the resulting information can easily become computationally intractable. A lot of researches are currently devoted in finding polynomial algorithms to efficiently deal with problems involving large networks and some chapters of this thesis represent a further effort in this line of research.

The computational difficulties of dealing with network data sets appear clearer when the observed data set are regarded as a sample of an unknown probability distribution we wish to infer. In this context the absence of closed-form mathematical results for most of the probabilistic models of networks implies the need of massive statistical simulation of random networks, as discussed in chapters 3, 4 and 5.

1.1.2 Mathematical framework and notation

The preliminary requisites this thesis demands to a reader is: i) a robust knowledge of Linear Algebra and Mathematical Programming; ii) a first course in Probability and Statistical Inference, iii) a general comprehension of Microeconomics and sociological theories.

The concept of network (or graph) will be continuously used throughout the thesis and requires a proper and formal definition. A graph $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$, with $|\mathcal{V}(\mathcal{G})| = n$ and $\mathcal{E}(\mathcal{G}) \subseteq \mathcal{V}(\mathcal{G}) \times \mathcal{V}(\mathcal{G})$, is normally represented in terms of a $n \times n$ binary matrix $X(\mathcal{G})$, called *adjacency matrix* (AM from now on), whose (i, j)-entry, x_{ij}, are equal to 1 if there is a link between the corresponding row and column elements and 0 otherwise. A simple graph has no loop, so that the diagonal elements of $X(\mathcal{G})$ are null.

A valued graph is a graph in which each link is given a numerical value. The set $\mathcal{E}(\mathcal{G}) \subseteq \mathcal{V}(\mathcal{G}) \times \mathcal{V}(\mathcal{G})$ of pairs of nodes is replaced by the set of values associated to each couple of nodes

$\omega : \mathcal{V}(\mathcal{G}) \times \mathcal{V}(\mathcal{G}) \longrightarrow \mathbb{R}$ and the valued adjacency matrix became $X(\mathcal{G}) \in \mathbb{R}^{n \times n}$, whose (i, j)-entry, x_{ij}, is equal to the value (or tie strength) between i and j.

When the network we are talking about is clear, the label $\mathcal{G}$ will be omitted in denoting the node set, so that we will write $\mathcal{V}$, $\mathcal{E}$ and X, instead of $\mathcal{V}(\mathcal{G})$, $\mathcal{E}(\mathcal{G})$ and $X(\mathcal{G})$. The number of nodes will be denoted either by n or $|\mathcal{V}|$.

The AM is an element of the set of binary matrices

$$\begin{aligned} & \chi = \{x_{ij} \in \{0,1\}, (i,j) \in \mathcal{H}^2\}, \\ \text{where} \quad & \mathcal{H}^2 = \{(i,j) : 1 \leq i \leq n-1, i < j \leq n\} \qquad \text{for undirected graphs} \\ \text{or} \quad & \mathcal{H}^2 = \{(i,j) : 1 \leq i \leq n, 1 \leq j \leq n, i \neq j\} \qquad \text{for directed graphs.} \end{aligned} \tag{1.1}$$

This representation of set of networks by solutions of system of algebraic equation will be particularly important in chapters 3, 4 and 5.

When X is regarded as an outcome of an underlying unknown random process, χ will represent a sample space of networks and $(\chi, P_\chi, \Im)$ a probability space, where $\Im$ is a σ-algebra and P_χ a probability measure used to model the uncertainty on the network structure.

Throughout the thesis we denote the vector of variables associated to the components of the AM as either $\mathbf{x}^T = \left[x_{12}, \ldots, x_{1n}, x_{23}, \ldots, x_{(n-1)n}, x_{21}, \ldots, x_{n(n-1)}\right]$ (i.e., the rowwise upper triangle of AM followed by its columnwise lower triangle) for directed graphs, or $\mathbf{x}^T = \left[x_{12}, \ldots, x_{1n}, x_{23}, \ldots, x_{(n-1)n}\right]$ (only the rowwise upper triangle of AM) for undirected graphs.

Another interesting matrix representation of a network is the Laplacian matrix, as will be seen in Chapter 2. Let f_i be the degree of a vertex i, that is $f_i = \sum_{j \in \mathcal{V}(\mathcal{G})} x_{ij}$ and $D(\mathcal{G}) = \text{diag}(f_1, \ldots, f_n)$. The Laplacian matrix of $\mathcal{G}$ is defined as $L(\mathcal{G}) = D(\mathcal{G}) - X(\mathcal{G})$[1]. The sequence $\lambda_1^{(L}, \ldots, \lambda_n^{(L}$, being the multiset of eigenvalues of $L(\mathcal{G})$, is called graph spectrum. Two nonisomorphic graphs can share the same spectrum and an open area of research is the one dealing with the relation between graph properties and the eigenvalues of the Laplacian matrix.

The present chapter keeps the discussion somewhat broad and define, in general and conceptual form, the central mathematical programming problems that will concern us in the rest of this thesis.

1.1.3 Complex networks in the sociological literature

From a sociological viewpoint, a nifty definition of *interpersonal tie* would consider interaction-carrying connections between individual agents. As noted in Subsection 1.1.1, the nature of these connections might involve many overlapping properties, as the number of interactions performed in a period, the length of the interacting period, the content of those interactions, etc. A comprehensive social network theory should take into account such a multivariate nature and construct on its bases explanatory models of social phenomena.

However, despite the presence of multiple properties, the most important features of interpersonal ties appear to be highly correlated in real SNs and a formal characterization by means of a numerical or categorical label might sometimes represent a reasonable reduction of redundancy and facilitate a parsimonious low dimensional representation of our data[2]. The *strength* of an interpersonal tie has been often adopted by network scientists [108, 150, 164] to this end.

[1] A set of n non negative numbers $f_1, \ldots, f_n$ might represent the row (or column) sums of an adjacency matrix if and only if it verifies $[\sum_{i=1}^{k} f_i \leq k(k-1) + \sum_{i=k+1}^{n} \min(f_i, k)]$.(for more details see Erdos [86]).

[2] The curse of dimensionality is the name used to refer to various phenomena that arise when analyzing data in high-dimensional spaces that do not occur in low-dimensional settings. The reduction of the dimensionality prior to any modeling decision is often used to avoid the effects of data redundancy and the curse of dimensionality.

> *'Most intuitive notions of the 'strength' of an interpersonal tie should be satisfied by the following definition: the strength of a tie is a (probably linear) combination of the amount of time, the emotional intensity (mutual confiding), and the reciprocal services which characterize each tie'.* (See Granovetter [108], page 1361.)

Granovetter's seminal article *The Strength of Weak Ties* [108] was a highly influential paper, which engendered increasing applications of the theory of valued graphs in the analysis of SNs [19–21, 82, 129, 150, 164]. By considering interpersonal ties of three discrete varieties (strong, weak and absent ties), Granovetter observed that the macroscopic and mesoscopic structural pattern of many networks result to be somehow interrelated with the distribution of the tie strength among couple of nodes.

Consider a network represented by several highly-connected communities, where the internal connections are mostly composed of strong ties and some nodes are located in the interface of two or more communities[3]. Most of the members of the communities are strongly related with each other, whereas only weak connections among members of different communities are present. Granovetter speculated that the structural peculiarity of these weak ties translates into differences in the resources that a node which is located in the interface of two or more communities might use. A *bridge* between far apart source of information, opinion, wealth and whatever other process or flow is carried out in the network, is established, resulting in substantial implications for the overall structure.

In graph theory the idea of a *bridging* property of a dyadic link is indeed captured by considering edges whose deletion increases the number of connected components of the network. Thus, a *bridge* is an edge which is literally the only path between its endpoints. To extend this notion Watts [227] introduced the more comprehensive concept of *local bridge of order* r.

Definition 1 (Local bridge)**.** *The order of a local bridge* $(i,j) \in \mathcal{E}$ *is the length of the shortest path between* i *and* j *in the absence of a direct link between them.*

Granovetters theory rests on the idea that local bridges tend to be weak ties, whereas strong ties might only be found within community members. As a consequence, nodes which are strongly connected are also likely to share many common contacts inside the community, so that a *positive association between structural similarity (overlapping) and strength of a tie* must appear. This claim might be regarded as *Granovetter's first hypothesis*[4].

The statistical measures of *structural similarity* between two nodes operationalize the idea of how similar their respective patterns of connections with the rest of the network are, translating the association between structural similarity and tie strength into a clear mathematical object. In binary networks structural similarity coincides with the idea of neighborhoods overlap, as first illustrated by Lorrain and White [151]. They labeled as *structurally equivalent* any two nodes which are related in the same ways to the same other nodes. Of course, in real SNs hardly anybody might ever be structurally equivalent to anybody, so that a suitably weaker concept replacing the idea of *equivalence* with the one of *similarity* might take into account the cosine between column vectors of the AM, or other measure of their statistical correlation. In the next section we shall investigate the properties of the following measure of structural similarity:

$$\Pi(x_i, x_j) = \frac{x_i^T x_j \; + \; x_{ij}(2\Delta_{\max} - (f_i + f_j))}{\Delta_{\max}^2}, \tag{1.2}$$

[3] We are assuming to be dealing with a network with a reasonable level of community structure, which is a property implying a straightforward detection of (potentially overlapping) subgraphs, such that each set of nodes is densely connected internally.

[4] Note that the positive association between structural similarity (overlapping) and strength of a tie directly entails a substantial level of community structure, as nodes which are structurally similar are likely to be strongly connected and by the transitivity principle of similarity this must be true for all the nodes inside a community.

where x_i is the i^{th} column vector of the AM and and $\Delta_{\max}$ is the highest degree in the network ($\Delta_{\max} = \max\{f_1, \ldots, f_n\}$). In the case of valued networks the idea is pretty much the same and the measure in (1.2) still provides information about the similarity of the corresponding column vectors of symmetric positive square matrices.

If we are dealing with random networks, the claim of positive association between structural similarity and strength of a tie might be formulated by saying that for every three nodes $i, j, k \in \mathcal{V}$, if $u > v$ then

$$E\left[x_j^T x_k \middle| x_{jk} = u\right] \geq E\left[x_j^T x_k \middle| x_{jk} = v\right], \tag{1.3}$$

where $E[Y|Z]$ is the conditional expectation of Y given the value of Z. The inequality (1.3) postulates that the expected overlap between two arbitrary nodes i and j cannot decreases when the value of their mutual ties increases. In the case of binary networks this is similar to the claim that the expected embeddedness[5] of an edge (i, j) cannot decrease when passing from $x_{ij} = 1$ to $x_{ij} = 0$.

The reverse reasoning would state that the expected values of a tie x_{ij} cannot decrease when none of the links with a third contact k decreases, which is indeed a particular way of rearranging the hypothesis of the presence of triadic closure in random networks. A random model of network exhibit triadic closure if for every three nodes $i, j, k \in \mathcal{V}$:

$$E[x_{ij} \mid x_{ik} = w, x_{jk} = t] \geq E[x_{ij} \mid x_{ik} = w, x_{jk} = r] \tag{1.4}$$

for all $w, t, r \in \mathbb{R}_{+\cup\{0\}}$ with $t \geq r$. A strong version of this claim says that if a strong tie exists between i and k and between j and k, there must be either a weak or strong tie between i and j, which should be regarded as *Granovetter's second hypothesis*[6].

It can be seen that for quite general probabilistic models – verifying very soft conditions – the hypothesis of triadic closure in (1.4) implies the association between structural similarities and tie strength in (1.3). To see why this is the case, consider the probability space $(\chi, P, \Im)$, where χ is the sample space of valued networks, $\Im$ is a σ-algebra and P a probability measure, verifying pairwise independence[7] of links in the networks, i.e. $P(x_{ji} = a, x_{ki} = b) = P(x_{ji} = a)P(x_{ki} = b)$ and non negative range, i.e. $P(x_{ij} < 0) = 0$ for every $(i, j) \in \mathcal{V} \times \mathcal{V}$. Applying the properties of joint and conditional probability, we obtain that (1.4), that is the hypothesis of the presence of triadic closure in a random network model specified by $(\chi, P, \Im)$, might be written as

$$\int z\, P(x_{ij} = z \mid x_{ik} = w,\, x_{jk} = t)\, dz \geq \int z\, P(x_{ij} = z \mid x_{ik} = s,\, x_{jk} = r)\, dz \tag{1.5a}$$

$$\int z\, \frac{P(x_{ij} = z,\, x_{ik} = w,\, x_{jk} = t)}{P(x_{ij} = w,\, x_{jk} = t)}\, dz \geq \int z\, \frac{P(x_{ij} = z,\, x_{ik} = w,\, x_{jk} = r)}{P(x_{ij} = r,\, x_{jk} = r)}\, dz \tag{1.5b}$$

$$\int z\, \frac{P(x_{ij} = z,\, x_{ik} = w, \mid x_{jk} = t)}{P(x_{ik} = w, \mid x_{jk} = t)}\, dz \geq \int z \frac{P(x_{ij} = z,\, x_{ik} = w, \mid x_{jk} = r)}{P(x_{ik} = w, \mid x_{jk} = r)}\, dz \tag{1.5c}$$

We are using the Lebesgue integral as a generalization of summation for continuous spaces, that is to say, spaces of valued networks, where χ is the set of valued AMs. Since the probability measure P verifying pairwise independence, then

$$\int z\, P(x_{ij} = z,\, x_{ik} = w, \mid x_{jk} = t)\, dz \geq \int z P(x_{ij} = z,\, x_{ik} = w, \mid x_{jk} = r)\, dz. \tag{1.6}$$

[5] The embeddedness of an edge is the number of common neighbors the two endpoints have. Under Granovetter's first hypothesis strong ties should also be highly embedded.

[6] Granovetter's second hypothesis is a hard formulation of triadic closure, which is too extreme to hold across very large valued networks, but it might be sometimes a useful simplification of reality which helps to understand social interactions.

[7] Note that a pairwise independent collection of random variables are not mutually independent.

As w is a non negative constant, both sides of the inequalities might be multiplied by w without affecting the direction of the inequality:

$$w \int zP(x_{ij} = z, x_{ik} = w, \mid x_{jk} = t)dz \geq w \int zP(x_{ij} = z, x_{ik} = w, \mid x_{jk} = r)dz, \tag{1.7}$$

$$\iint wzP(x_{ij} = z, x_{ik} = w, \mid x_{jk} = t)dz\,dw \geq \iint wzP(x_{ij} = z, x_{ik} = w, \mid x_{jk} = r)dz\,dw. \tag{1.8}$$

Since the last inequality is true for all $i \in \mathcal{V}$, the following inequalities must also hold:

$$E\left[x_j^T x_k \middle| x_{jk} = t\right] \geq E\left[x_j^T x_k \middle| x_{jk} = r\right]. \tag{1.9}$$

We obtain that (1.4) implies (1.3), providing a conceptual (and mathematical) relationship between the community structure (caught by the level of association between structural similarity and tie strength) and the triadic closure [226, 227]. (Note that we are using the Lebesgue integral as a generalization of the summation for continuous spaces, that is to say, spaces of valued networks, where χ is a set of valued AMs. In the cases of zero-one-networks, the integral can be replaced by a summation.)

An immediate problem in this respect is to find out whether (1.4) and (1.3) are empirically verified by observed social networks. As we will see in the second part of this thesis, statistical methodologies to test these hypothesis requires a proper specification of the probability space $(\chi, P, \Im)$ and the knowledge of a well defined sampling probability[8].

An interesting study aiming at testing the accomplishment of Granovetter's hypothesis of observed social networks has been carried out in 1980 by Friedkin [95]. Based on a social network of biological scientists, the study supports the positive association between structural similarities and tie strengths: the contact circles of two scientists tend to overlap more as the strength of the tie between the two scientists increases.

If strong ties mostly appear inside densely connected communities (as argued by Granovetter [108]), then sparse networks are likely to exhibit a substantial level of community structure, suggesting a possible partition of edge set in weak local bridges and strong highly embedded ties: *'the stronger the tie between two individuals, the larger the proportion of people to which they are both tied'* [108]. Table 1.1 illustrates a taxonomy of edges in accordance with this reasoning.

	strong	weak
high embedded	within communities	
low embedded		local bridges

Table 1.1: Taxonomy of edges.

As Easley and Kleinberg noted [84], a long line of research in sociology, catalyzed by the influential work of Coleman [68, 69] has argued that if two individuals are connected by an embedded edge, then this makes it easier for them to trust one another, and to have confidence in the integrity of the transactions (social, economic, or otherwise) that take place between them. Indeed, the presence of mutual friends puts the interactions between two people *on display* in a social sense. No similar kind of deterring threat exists for edges with zero embeddedness (local

[8] A sampling distribution is the probability distribution of a given statistic, based on a random sample. In our case the interest would be in the probability distribution of a statistic involving (1.4) and (1.3).

bridges), whose endpoints are subject to distinct norms and expectations from different groups they are associates with [67].

As influentially argued by Burt [43], although being connected with the rest of the network by means of highly embedded edges has the aforemention advantages, a related line of research in sociology has argued that nodes who are located at the ends of multiple local bridges have also many fundamental advantages. Empirical studies of managers in large corporations has correlated individual's success within a company to their access to local bridges [42, 43].

Most edges of networks verifying Granovetter hypothesis are in the main diagonal of Table 1.1. Although weak local bridges and strong highly embedded ties are the only needed ingredients to ensure high transitivity in sparse networks, this conditions do not guarantee high network connectivity, in the general sense of short distance between pares of nodes. If strong ties mostly appear inside densely connected communities, the overall structure might result to be poorly connected[9].

To increase the connectivity while keeping high transitivity in sparse networks, one should zoom inside diagonal cell of Table 1.1 with the aim of considering the structural position of endpoints of the two type of edges. Table 1.2 take into account a possible taxonomy of the endpoints of strong highly embedded ties and weak local bridges, in accordance with the nodal embeddness and degree.

	low degree	high degree
high embeded	conservative	
low embeded		adventurer

Table 1.2: Taxonomy of nodes.

Following the Boorman reasoning [35], any person has a limited amount of energy/time to invest in maintaining contacts across the organization, and a node incident with multiple local bridges is investing his/her energy efficiently by reaching out to different groups rather than basing all her contacts in the same group: $\sum_j t_{ij}x_{ij} \leq T$, where T is the energy/time to invest in maintaining contacts across the organization – which is approximately homogeneous for all people[10]– and t_{ij} is the energy/time required to node i to keep a relation to actor j. A nodal decision would be to establish the optimal distribution of a limited amount of energy/time between this two kind of edges[11].

In the context of networks verifying Granovetter's hypothesis, the possible strategies of a node would be to distribute energy/time either among few highly embedded strong ties, or among many weak local bridges. We call the two class of nodes *conservatives* – embedded in a tightly-knit communities – and *adventurer* or *hub* – located in the interfaces of far apart communities – respectively.

Sparse valued networks where most edges are in the main diagonal of Table 1.1 and most

[9] When we say that a graph is a *small world*, we mean, informally, that almost every pair of nodes is connected by a path with an extremely small number of steps.

[10] Note that in the last half a century, social structures exhibited a progressive increase in density, due to the improvement of machineries and technological facilitating a lower consumption of time and energy in keeping alive social contacts.

[11] If we look at the network from a temporal point of view, how long these local bridges last before triadic closure produces short-cuts around them, and the extent to which people in an organization are consciously, strategically seeking out local bridges and trying to maintain them, is less well understood; it is a topic of ongoing research [44], [138].

nodes in the main diagonal of Table 1.2 are clear cases of highly transitive and highly connected structures, as the presence of local bridges of a reasonably high order associated to poorly embedded hubs reduces the distances between nodes of different communities without affecting the global community structure. Other configurations can be obtained when edges and nodes are assigned to different combinations of the aforementioned edge and node taxonomies.

The presence of hubs and conservatives implies certain variation at individual level, which has been often regarded to be one of the most characterizing features of humans social networks. Some people have few friends whereas others have many. Some people are embedded in tightly-knit groups where everyone knows each other, whereas others belong to many different groups where there is little overlap between friends.

To summarize, highly connected valued networks verifying Granovetter's hypothesis and Boorman principle (ensuring sparsity) might be captured by the following requirements:

- *weakness of local bridges;*
- *association between structural similarities and tie strengths;*
- *heterogenity/polarization of structural positions.*

Quite clearly, the three statements are not rigorously formulated in mathematical terms. The choice of the network statistics which best represent those qualitative statements might be a matter of scientific subjectivity.

The operativization of network properties by computable network statistics has been studied since the very beginning of the analysis of SNs [96, 224], but few researchers have been devoted to the generalization of network statistics to the case of valued networks. Recently there have been some attempts to generalize the average local clustering coefficient proposed by Watts [227] to the case of valued connections. Equations (1.10) show the valued clustering coefficients for node $i \in \mathcal{V}$ proposed by Lopez [150] and Barrat [19], respectively.

$$CC_{LF}(i) = \frac{1}{s_i(s_i-1)} \sum_{i,j \in N(i)} x_{hj} \tag{1.10a}$$

$$CC_{B}(i) = \frac{1}{f_i(f_i-1)} \sum_{h,j \in N(i)} (x_{ih} + x_{ij}) I(x_{hi} > 0) I(x_{ij} > 0) I(x_{hj} > 0) \tag{1.10b}$$

Here, $f_i = \sum_{j \in \mathcal{V}} x_{ij}$ is the total value of ties in the neighborhood $N(i)$ of node i, $s_i = \sum_{j \in \mathcal{V}} I(x_{ij} > 0)$, and $I(e)$ is a boolean function which is equal to 1 when the expression inside the parenthesis is true and 0 in the opposite case. The measure of clustering proposed by Barrat et al. combines topological information with the distribution of values in the network.

We remark that both clustering coefficients suffer from the drawback that they require an underlying binary network and, in the case this is not known, it must be presumably obtained by discretizeing the valued edges, on the base of some thresholding criterion. (A detailed analysis of valued clustering coefficients has been performed by Kalna and Higham [129].)

In the Chapter 2 the operationalization of the three aforementioned network concepts (weakness of local bridges, association between structural similarities and tie strengths, heterogeneity of structural positions) is done by considering the distribution of the following nodal quantities:

- *relative sociability*: $1 - f_i \Delta_{\max}^{-1}$, for $i \in \mathcal{V}$
- *relative dispersion*: $\Delta_{\max}^{-2} \sum_{j \in \mathcal{V}} x_{ij}^2$, for $i \in \mathcal{V}$

- *relative association between structural similarity and tie strength*: $\Delta_{\max}^{-1}\sum_{j\in\mathcal{V}} x_{ij}\Pi(x_i,x_j)$, for $i\in\mathcal{V}$

I will be shown that the three nodal quantities are algebraically related to graph eigenvalue, structural similarity, network connectivity and transitivity. Their properties will allow for a formal mathematical treatment of the sociological studies of open and closed systems.

From this view, a star network $\mathcal{G}_\star$ might be associated to open systems, based on the average profile of the aforementioned nodal quantities[12] – nodes and edges are not embedded, structural positions are polarized, connectivity is maximized – as follow:

$$\frac{1}{n}\sum_{i\in\mathcal{V}}\left(1-f_i\Delta_{\max}^{-1}\right)=\frac{n-2}{n}\quad\underset{n\to\infty}{\longrightarrow}\quad 1 \tag{1.11}$$

$$\frac{1}{n}\sum_{i\in\mathcal{V}}\left(\Delta_{\max}^{-2}\sum_{j\in\mathcal{V}}x_{ij}^2\right)=\frac{2}{n(n-1)}\quad\underset{n\to\infty}{\longrightarrow}\quad 1 \tag{1.12}$$

$$\frac{1}{n}\sum_{i\in\mathcal{V}}\left(\Delta_{\max}^{-1}\sum_{j\in\mathcal{V}}x_{ij}\Pi(x_i,x_j)\right)=\frac{n-2}{n(n-1)^2}\quad\underset{n\to\infty}{\longrightarrow}\quad 0 \tag{1.13}$$

As we will discuss Section 1.2 this properties of a star as open system reflects in the processes of exchanges material, energy, capital and information within its internal subsystems. In contrast, closed network structures results in highly embedded nodes, facilitating internal processes requiring highly transitive interactions.

1.1.4 Computational methods in the analysis of complex networks

A significant amount of recent researches in the field of Network Theory has focused on the development of algorithms and computational methods for:

i. graph clustering and community detection;

ii. spectral graph theory;

iii. node centrality;

iv. statistical inference and simulation of random graphs;

v. network flow optimization problems;

vi. game theory and multi-agent systems.

In this subsection the basic literature and state of the art of these four families of network problems are investigated along with some introductory explanation of our main contribution in these fields.

[12] Beside the relative sociability and the relative dispersion, the centralization of the edges distribution in a star can be assessed by looking at the Gini concentration $\mathrm{Gini}(\mathcal{G}_\star)$, which is defined, for the distribution of an arbitrary variable y among a population of n subjects, as $1-2(\sum_{i=1}^n(n+1-i)y_i)/(\sum_{i=1}^n y_i)+1/n$, where $y_1,\ldots,y_n$ have been placed in a non decreasing order. We see that for a star with n vertices the Gini concentration of the degree sequence is

$$\mathrm{Gini}(\mathcal{G}_\star)=1-\frac{1}{n(n-1)}\left(\sum_{i=1}^{n-1}(n+1-i)+(n-1)+\frac{1}{n}\right)=\frac{1}{2}-\frac{1}{n}\quad\underset{n\to\infty}{\longrightarrow}\quad\frac{1}{2}.$$

i. Blockmodels, graph clustering and community detection. To start with, consider the problem of constructing subgraphs in large real-world graphs based on some principles of internal homogeneity. This problem represented a computationally difficult task since the very beginning o the Network Theory. As far as I know, the first application of this problem in the analysis of SN is related to the studies of blockmodelling [39, 151, 224] and the ones of statistical cluster analysis [232].

A blockmodel is a node partition into discrete positions, called blocks, consisting of square submatrices of structurally equivalent nodes or structurally similar nodes (nodes with very similar, if not identical, relations with the rest of the network).

Definition 2 (Blockmodel)**.** *A blockmodel $BM = (P, B)$ of a graph $\mathcal{G}$ consists of two parts:*

- *A partition $P = (P_1, \ldots, P_L)$ of V into L disjoint subsets called the positions of $\mathcal{G}$;*
- *A density matrix representing relations among positions*[13]*.*

Blockmodels reproduce relations among nodes occupying structurally equivalent positions (blocks) in networks. A blockmodel could be built by requiring the positions to represent structurally equivalent blocks, such as the workers of different industries.

CONCOR (a short form of *CONvergence of iterated CORrelations*) has been for three decades the most popular method to deal with the problem of partitioning nodes based on their structural similarities. Starting from the Pearson product-moment correlation coefficient among couples of rows (call it $C_1 \in [-1, 1]^{n \times n}$) of the AM, the CONCOR obtains the correlation matrix of the rows of C_1 and iteratively calculates correlations of correlations. This process is supposed to converge to a steady-state correlation matrix, call it C_*, whose entries are either -1 or $+1$. The function

Algorithm 1 CONCOR

Set $k = 0$ and C_0 as the initial matrix
repeat
  $C_{k+1} = \mathrm{Corr}(C_k)$; $k = k + 1$.
until Convergence

$\mathrm{Corr} : \mathbb{R}^{n \times n} \longrightarrow \mathbb{R}^{n \times n}$ returns a matrix whose components are the Pearson product-moment correlation coefficient between the columns of the matrix, passed as input.

The rows of C_* can be partitioned into two groups so that every $+1$ occurs between rows and columns in the same group, and every -1 occurs between vertices in different groups. Repeating these procedures, CONCOR can split each of the two initial blocks into two more blocks, and so on. The decision about where to stop the division process will determine the ultimate number of obtained blocks.

We mentioned in the previous section that the statistical measures of *structural similarity* between two nodes operationalize the idea of how similar their respective patterns of connections with the rest of the network are. A direct consequence of Granovetters theory describen in Section 1.1.3 is that nodes which are strongly connected are also likely to share many common contacts inside the community, so that a positive association between structural similarity (overlapping) and strength of a tie must appear, turning into a substantial level of community structure and supporting the previously proved implication between (1.4) and (1.3).

From this outlook the problem of detecting highly connected communities in networks might be seen as a generalization of blockmodelling problem, when the within-group similarity criterion is not specified, so that the within-group homogeneity might be based on arbitrary nodal features

[13] A density matrix has groups rather than individual people as its rows and columns, so that the partitioning leads to the problem of deciding how the subsets relate to each other.

and/or pattern of connectivity. In fact,the studies of blockmodelling and the ones of statistical cluster analysis started overlap in several novel researches by the last decade [81, 208], resulting in a substantial interdisciplinary connection between graph theory and statistics.

In its basic framework, the problem of detecting communities involves an unknown number of clusters which are supposed to have an internally high connectivity and/or to be similar with respect to some nodal feature. An objective function can be chosen to capture this intuition, resulting in well defined Mathematical Programming problems, which are typically NP-hard to be optimized exactly [208]. When the number of communities are a priori known, an integer programming problem to assign nodes to communities might be:

$$\begin{array}{lll} \text{minimize} & \psi\big(z_{ik},\ i \in \mathcal{V}, k \in \mathcal{K}\big) & \\ \text{subject to} & \sum_{k \in \mathcal{K}} z_{ik} = 1 & i \in \mathcal{V}, \\ & z_{ik} \in \{0,1\} & i \in \mathcal{V}, k \in \mathcal{K}, \end{array} \tag{1.14}$$

where $\mathcal{K}$ is the set of communities, $\mathcal{V}$ is the set of nodes, $\mathcal{E}$ is the set of edges. The decision variable z_{ik} denotes whether the node i belongs to community j; function ψ is a real-valued function representing the cost of adding a node to a group and might be defined in term of an optimization problem or as a closed-form expression, such as $\psi = \sum_{(i,j)\in\mathcal{E},k\in\mathcal{K}} \delta_{ij} z_{ik} z_{jk}$ or $\psi = \max\{\delta_{ij} z_{ik} z_{jk} \ : \ (i,j) \in \mathcal{E}, k \in \mathcal{K}\}$, where δ_{ij} is a measure of distance between node i and node j. In the case δ_{ij} is the geodesic distance in the graph (the shortest path between node i and node j), then the optimal solution of (1.14) is the assignment of nodes to communities in such a way that the largest diameter (the largest distance within the community) is minimized. A similar idea is valid for arbitrary measure δ_{ij}, such as the ones associated to the spaces of nodal attributes.

Heuristics or approximation algorithms for community detection [174] have been successfully developed [191], ushering in the beginning of a massive research in mathematical programming based approaches for this prominent field of complex networks.

ii. Spectral graph theory. The main achievements in spectral graph theory, as the science which studies mathematical connections between spectral properties of graph matrices and topological network features, have also occurred during the eighteen's and nineteen's [32, 106, 161, 217], resulting in a great amount of heterogeneous applications is social sciences and engineering [66].

To provide a brief description of such mathematical connections between spectral properties of graph matrices and topological network features, let $\lambda_1^{(L}, \ldots, \lambda_n^{(L}$ be the multiset of eigenvalues of L and $\{\lambda_1^{(A}, \ldots, \lambda_n^{(A}\}$ the multiset of eigenvalues of X. The following statements illustrate some important facts about $\{\lambda_1^{(L}, \ldots, \lambda_n^{(L}\}$ and $\{\lambda_1^{(A}, \ldots, \lambda_n^{(A}\}$.

i. The number of times 0 appears as an eigenvalue in the Laplacian is the number of connected components in the graph.

ii. The second smallest eigenvalue of L is the algebraic connectivity of $\mathcal{G}$, whose magnitude reflects how well connected the overall graph is.

iii. In the binary case, the summation of the squares of the adjacency matrix eigenvalues $(\lambda_1^{(A}, \ldots, \lambda_n^{(A})$ is equal to the network absolute density: $\sum_i (\lambda_i^{(A})^2 = \sum_{ij} x_{ij}$.

The advantage of looking at the eigenvalues is that they can in fact be efficiently computed by polynomial algorithms and provides important information concerning the structure of the network.

The QR-algorithm is one of the best known numerical procedure to calculate the eigenvalues and eigenvectors of a matrix. The basic idea is to perform a QR-decomposition and to iteratively multiply the factors in the reverse order. Similarly, power iteration is one of many eigenvalue algorithms that may be used to find the dominant eigenvector.

Chapter 2 will be entirely devoted to the description of novel representation of complex networks based on the multiset of eigenvalue of the of L. Its application will be clear in Chapter 5, when the proposed spectral-based representation will be apply to the problem of assessing the goodness of fit of random network models.

iii. Node centrality. Another popular network problem where the use of computational methods resulted in successful engineering applications is associated to the construction of web-page ranking [41]. The PageRank is perhaps one of the most popular method to rank web-sites in search engine results [184]. It was first developed by Larry Page and Sergey Brin in 1996 with the aim of estimating the relative importance of a web-page within the internet.

Algorithms for web-page ranking are typically based on nodal centrality, measuring the relative importance of nodes within a network [32, 96, 130]. Usual taxonomies of the main measures of centrality consider four classes, associated with the *degree*, the *betweenness*, the *closeness*, and the *eigenvector*. (To have a more detailed explanation of centrality indexes, see the glossary of complex networks in Appendix A.) Formally, the centrality of node i is defined in term of the centralities of the other nodes:

$$c_i = \frac{1}{\lambda_{\max}} \sum_{j \in \mathcal{V}} x_{ij} c_j \tag{1.15}$$

where $\lambda_{\max}$ is a constant corresponding to the greatest eigenvalue of the AM, so that $X\mathbf{c} = \lambda_{\max}\mathbf{c}$.

A generalized eigenvector centrality has been proposed by Bonacich [32, 33]:

$$c_i(\alpha, \beta) = \sum_{j \in \mathcal{V}} (\alpha + \beta c_j(\alpha, \beta)) x_{ij}, \tag{1.16}$$

where $\mathbf{e}$ is a vector of ones, β reflects the degree to which a node's centrality is a function of the centralities of its neighbors and α is a scaling parameter. In matrix form we have $\mathbf{c}(\alpha, \beta) = \beta X^T \mathbf{c}(\alpha, \beta) + \alpha X \mathbf{e}$, so that $\mathbf{c}(\alpha, \beta) = \alpha(I - \beta X)^{-1} X \mathbf{e}$.

Bonacichs approach is base on the differentiation between power and centrality: a node is central when it has many connections within its local network (degree), whereas its power depends on the power of its neighbors (recursive definition). The goodness of being connected with powerful or powerless neighbors depends on the type of commodity flowing within the network. If the utility of nodes are related to the amount of obtained information, the non rival nature of information suggest a positive association between the power of a node and the power of its neighbors. To stress this point Bonacich [32] claimed that

> *a set of exchange relations is positive if exchange in one relation is contingent on exchange in others and negative if exchange in one relation precludes exchange in others. In communication networks, exchanged information is usually received from others, and so the system is positive, but, when exchanging a commodity with one person precludes exchange with another, the relation is negative. These would be modeled with positive and negative values of β respective.* (See Bonacich [32], page 1171.)

The main idea behind the eigenvector centrality (both is its original version and in Bonacich's generalization) is to somehow take into account indirect connections. Another variant of eigenvector centrality which is also oriented to balance the contribution of directed and indirected

links is the Katz centrality [135]:

$$c_i(\alpha) = \sum_{k=1}^{\infty} \sum_{j \in \mathcal{V}} \alpha^k (X^k)_{ij} \tag{1.17}$$

where $(X^k)_{ij}$ is the (i,j) component of the k^{th} power of the AM. It can be seen that the limit of Katz centrality as α approaches $1/\lambda_{\max}$ is the principal eigenvector, so that Katz centrality can be similarly regarded a generalization of the original eigenvector centrality. A similar argument is valid for PageRank, one of the most popularized centrality measures to rank internet websites. PageRank satisfies the following system:

$$c_i(\alpha) = \alpha \sum_{j \in \mathcal{V}} c_j(\alpha) \frac{x_{ij}}{f_j} + \frac{1-\alpha}{n} \tag{1.18}$$

where f_j is the degree of node j. PageRank might also be seen as the probability that an internet user randomly clicking on links will arrive at any particular page. In this sense, α represents a damping factor, that is, the probability, at any step, that the internet user will eventually continue clicking.

iv. Statistical inference and simulation of random graph models. Random models are usually adopted to capture qualitative properties observed in large-scale network data. This area of NA is sometimes quite disconnected from the previously described research fields and begins with the seminal work of P. Erdös and A. Rainyi [86], who considered a fixed set of nodes and an independent and equal probability of observing edges among them. There are two closely related variants of the Erdös-Rainyi model:

- *the* $\mathcal{G}(n,p)$ *model*, where a network is constructed by connecting nodes randomly with independent probability p;
- *the* $\mathcal{G}(n,m)$ *model*, where a network is chosen uniformly at random from the collection of all graphs with n nodes and m edges.

Both models possess the considerable advantage of being exactly solvable for many of their average properties: clustering coefficient, average path length, giant component, etc. (For more details about network properties, see Bollobas [29], and Wasserman and Faust [224].)

Nonetheless, most of random graph models have very few analytical results and the simulation of large random samples of networks is required to obtain empirical distributions of their average properties. Typical approaches for network simulation – mainly developed within the fields of mathematical sociology – were based on fixing topological properties, named *conditional random networks* [29]. These models are very useful when assessing the hypothesis of whether a particular network property is likely to appear under a uniform distribution of all networks verifying given constraints.

Simulating from this models generally results in highly combinatorial procedures, such as the ones to generate uniform random directed networks with given in- and out-degree sequences [62, 194, 203, 216, 222]. The analytical study of these models involves binary matrices with specified properties and some combinatorial results have been obtained by Ryser [202].

The construction and development of general simulation procedure which are valid both for binary and valued networks, represents indeed the main purpose of the second part of this Ph.D. thesis, where mathematical programming models of several families of networks are introduced.

Consider the probability space $(\chi, P_\chi, \Im)$, where χ is the set of all AMs of directed networks – denoted in vector form $\mathbf{x}^T = [x_{12}, x_{13}, \ldots, x_{1n}, x_{23}, x_{24}, \ldots, x_{2n}, \ldots, x_{(n-1)n}, x_{21}, x_{31}, x_{32},$

$x_{41}, \ldots, x_{n1}]$ – verifying a specified collection of linear constraints $A\mathbf{x} = \mathbf{b}$, $\Im$ is a σ-algebra on χ and P_χ a probability measure. For a given random cost vector $\mathbf{c} \sim f_C$ a way to simulate from $(\chi, P_\chi, \Im)$ is to solve the LP:

$$\begin{array}{rl} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \in \{0,1\}^n \end{array} \tag{1.19}$$

where $\mathcal{H} = \{(i,j) | i = 1 \ldots n-1; i < j \leq n\}$. The continuous relaxation of χ, name it $CR(\chi)$, is obtained by replacing $x_{ij} \in \{0,1\}$ by $x_{ij} \in [0,1]$, in (1.19). If A is totally unimodular (TU, from now on) then all extreme points of $CR(\chi)$ are binary vectors, representing networks[14]. As shown in [115], the next theorem provides sufficient conditions for a matrix to be TU:

Theorem 1. *Let $A \in \{-1,0,1\}^{m\times n}$ be a matrix obtained by elementary operations of $B \in \boldsymbol{Z}^{m\times n}$ and consider a partition of the rows of A in two disjoint sets $\mathcal{J}_1$ and $\mathcal{J}_2$. The following three conditions together are sufficient for B to be TU:*

1. *Every column of A contains at most two non-zero entries, which are either 1 or -1.*
2. *If two non-zero entries in a column of A have the same sign, then the row of one is in $\mathcal{J}_1$, and the other in $\mathcal{J}_2$.*
3. *If two non-zero entries in a column of A have opposite signs, then the rows of both are either in $\mathcal{J}_1$ or $\mathcal{J}_2$.*

Therefore, it is possible to generate a bunch of graphs by merely solving linear programs (LP) with random gradients in the objective function, or by non-degenerated simplex pivoting, starting from a given initial extreme point [182]. Beyond that, they can be generated in polynomial time if interior point methods are used [231].

If the gradient of the objective function $\mathbf{c}$ is a properly defined random vector of density function $f_C(\mathbf{c})$, then the solution of (1.19) is also a random vector whose probability distribution can be computed as

$$P_\chi(\mathbf{x}) = \int_\Omega f_C(\mathbf{c})\ d\mathbf{c}, \tag{1.20}$$

where Ω is the set of gradients for which $\widehat{\mathbf{x}}$ is an optimal solution. The gradient and optimal solutions are related through the KKT conditions of the LP and this fact will be exploited in Chapter 5 to derive a closed-form probabilistic relation between f_C – the probability distribution of the objective costs – and P_χ – the probability distribution of networks verifying the specified system of linear constraints –.

These properties will be extensively used in next chapters 3, 4 and 5, where random network generators will be introduced, along with specialized interior point methods to algorithmically deal with this class of problems.

v. Network flow optimization problems. Network flow optimization problems stand in the frontier that separates continuous and integer optimization. The boundary between these two fields results from the characterization of the constraint polyhedron by the convex hull of its extreme points. In this class of problems, it is usually possible to identify a subset of the system of linear constraints for which the extreme points of the associated polyhedron are integer and represent solutions of combinatorial problems that are seemingly unrelated to linear

[14] If $B \in \mathbf{Z}^{m\times n}$ is a TU matrix and b is integer, then polyhedrons of the form $\varphi = \{y \in \mathbb{R}^n : By = b;\ y \geq 0\}$ have only integer extreme points, as every non singular $m \times m$ submatrix of B has integer inverse. (For more details on Unimodularity in Integer Programming see [209], [142] and [144]).

programming [1]. Mathematically, a flow in a network is a real-valued function $\omega : \mathcal{V} \times \mathcal{V} \longrightarrow \mathbb{R}$, which is denoted as $x_{ij} = \omega(i, j)$, for $(i, j) \in \mathcal{V} \times \mathcal{V}$, with the following properties – known as flow conservation or balance equations – for all nodes $i \in \mathcal{V}$:

$$\sum_{(j,h)\in\mathcal{E}} x_{jh} - \sum_{(h,u)\in\mathcal{E}} x_{hu} = b_h, \qquad h \in \mathcal{V} \tag{1.21}$$

where b_i is the demand/supply of flow in node i, that is to say, the total flow departing from node i less the total flow arriving at i. The matrix structure associated to the balance equations (1.21) is known as the node-arc incidence matrix and has the nice property of being TU. A feasibility condition is that $\sum_{h\in\mathcal{V}} b_h = 0$. A well known network flow optimization problem is to maximize the total flow from an origin, say $v_o \in \mathcal{V}$, to a destination, say $v_d \in \mathcal{V}$, subject to the edge capacities. This is known as the maximum flow problem[15].

In Chapter 4, a particular attention will be given in this thesis to multicommodity network flow problems, which have been shown to be challenging for state-of-the-art solvers [51, 55]. The basic idea is the simultaneous conservation of flows of a set $\mathcal{C}$ of commodities, resulting in $k = |\mathcal{C}|$ systems of balance equations: $\sum_{(j,h)\in\mathcal{E}} x^i_{jh} - \sum_{(h,v)\in\mathcal{E}} x^i_{hv} = b^c_h$, for $i \in \mathcal{C}$, $h \in \mathcal{V}$. Sometimes there may be capacities on the total flows of arcs:

$$\sum_{i\in\mathcal{C}} x^i_{jh} \le q_{jh}, \qquad (h, j) \in \mathcal{E}, \tag{1.22}$$

or on the total in-out-flow of nodes:

$$\sum_{i\in\mathcal{C}} \left(\sum_{j\in\mathcal{V}} x^i_{hj} + \sum_{j\in\mathcal{V}} x^i_{jh} \right) \le q_h, \qquad h \in \mathcal{V}. \tag{1.23}$$

Given a directed network of n' arcs and n nodes, the general formulation for multicommodity network flow problem is:

$$\min \quad \sum_{i\in\mathcal{C}} ({c^i}^T x^i + {x^i}^T Q_i x^i) \tag{1.24a}$$

$$\text{subject to} \quad \begin{bmatrix} N & & & & 0 \\ & N & & & 0 \\ & & \ddots & & \vdots \\ & & & N & 0 \\ L & L & \dots & L & I \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^k \\ x^0 \end{bmatrix} = \begin{bmatrix} b^1 \\ b^2 \\ \vdots \\ b^k \\ q \end{bmatrix} \tag{1.24b}$$

$$0 \le x^i \le u^i \qquad i = 0 \dots k. \tag{1.24c}$$

Vectors $x^i \in \mathbb{R}^{n'}$, $i \in \mathcal{C}$ are the flows for commodity i. In the case of dealing with capacities on the total flows of arcs $x^0 \in \mathbb{R}^{n'}$ are the slacks of arc capacities constraints, whereas $x^0 \in \mathbb{R}^n$ are the slacks of nodal capacities constraints when total out-flow of nodes is considered. The node-arc incidence matrix of the directed graph is $N \in \mathbb{R}^{n'\times(n-1)}$. The arc capacities for all the commodities are $u \in \mathbb{R}^{n'}$, while $u^i \in \mathbb{R}^{n'}, i = 1, \dots, k$, are the individual capacities per commodity; u^0 are the upper bounds of slacks x^0; vectors $b^i \in \mathbb{R}^{n-1}, i \in \mathcal{C}$, are the node supply/demands for each commodity. Vectors $c^i \in \mathbb{R}^{n'}, i \in \mathcal{C}$, are the arc linear costs per

[15] Other examples of network flow optimization problems include shortest path problems, minimum cost flow problems, minimum cost equal-flow problems, multicommodity flows problems, etc. They can be used to model several complex systems in which some entity travels through a network of nodes, such as traffic in a road system, fluids in pipes, electrical circuit, etc.

commodity, and the diagonal positive semidefinite matrices $Q_i \in \mathbb{R}^{n' \times n'}, i \in \mathcal{C}$, denote the arc quadratic costs. The linking constraints matrix L is either a n'-dimensional identity matrix with the associated right hand term $q \in \mathbb{R}^{n'}$, when edge capacities are taken into account, or defined as abs(N) with the associated right hand term $q \in \mathbb{R}^n$, when nodal capacities are considered, where abs(N) is the matrix derived from N by switching all negative signs to positive.

Some of the solution strategies can be broadly classified into four main categories: simplex-based methods [57, 159], decomposition methods [11, 90], approximation methods [27], and interior point methods [11, 50]. Some of the approaches for linear multicommodity flows were compared in [61]. Significant advances have also been made for nonlinear multicommodity flows. Among them we find active set methods [57], ACCPM approaches [10], interior point methods for quadratic problems [52], proximal point algorithms [178], and bundle-type decomposition [147]. A description and empirical evaluation of additional nonlinear multicommodity algorithms can be found in the survey [179].

vi. Graphical games and multi-agent system. Several recent works have introduced graph-theoretic frameworks into multi-agent systems [136, 177], so that each node represents a single agent, and the edges represent pairwise interaction between agents. Game theory is typically used to model multi-agent interactions, and the global states of interest are the so-called Nash equilibria, in which no agent has a unilateral incentive to deviate.

Graphical games are games consisting of n players, each with a finite set of *actions available*, along with a specification of the *utility* to each player. Let Ξ_i be the space of actions of player i, and $v_i \in \Xi_i$. A pure strategy profile is a vector of strategies to players, that is an n-tuple $\mathbf{v} = [v_1 \dots v_n]$, such that $v_1 \in \Xi_1, \dots, v_n \in \Xi_n$. Let $\boldsymbol{\Xi} = \Xi_1 \times \dots \times \Xi_n$. The preference relation $\preceq_i$ of player $i \in \mathcal{V}$ on $\boldsymbol{\Xi}$ is represented by the utility function $u_i : \boldsymbol{\Xi} \to \mathbb{R}$, resulting from the joint action: $u_i(\mathbf{v} \mid \mathcal{G})$, where $\mathcal{G}$ is the network structure, which is regarded as a parameter of the problem. (Introductory concepts of Game Theory are provided in Appendix A.)

In a graphical game [136], each player i is represented by a node in an undirected graph $\mathcal{G}$ and what is relevant for the utility of agent i is the only action of its neighbors in $\mathcal{G}$. We use $N(i) \subseteq \mathcal{V}$ to denote the *neighborhood* of player i in $\mathcal{G}$ — that is, those vertices j such that the edge (i, j) appears in $\mathcal{G}$. By convention $N(i)$ always includes i itself as well. If $\mathbf{v}$ is a joint action, $\mathbf{v}_{N(i)}$ is used to denote the induced vector of actions just on the players in $N(i)$, so that the utility function of player i results to be $u_i(\mathbf{v}_{N(i)} \mid \mathcal{G})$.

The analogy between graphical games and general multi-agent systems become clear when the latter are practically applied for the numerical computation of stable equilibria[16]. A multi-agent system is a dynamical process involving entities (called agents) who are required to take specified actions within an environment.

Perhaps one of the main referential multi-agent system in the analysis of social interaction is the Axelrod's simulation [6] of the dissemination of culture. In its modern version agents are associated to nodes of a graph $\mathcal{G}$, and the individual attributes of each node are defined by a set of F features, each taking one of q possible traits. This means that each agent i has a cultural state vector $[v_1^i, v_2^i, ..., v_F^i]$. Each component (cultural features) v_f^i $(f = 1, \dots, F)$ takes any of the q values, which are initially assigned to each agent independently and with an equal probability of $1/q$. Culture satisfies two simple premises:

- agents are more likely to interact with others who share many of their cultural attributes;
- interactions between two agents tend to increase the knowledge they share.

[16] The problem of computing exact Nash equilibria in sparse graphs is computationally difficult, as suggested by its combinatorial nature. In a recent paper Kearns et al. [136] described an exact algorithm for computing Nash equilibria of graphical games.

In accordance with this assumptions, the discrete time dynamics of the model is governed by the principle that the probability that a cultural feature is transmitted from one agent to another increases with the number of features that they already share. Adjacent sites i and j cannot interact if either they share all traits ($l_{ij} = F$) or none of them ($l_{ij} = 0$) and the bond between i and j is regarded as inactive. When two agents linked by an active bond interact, one of their different features switches in accordance with the state of the neighbor. When all bonds are inactive the configuration is absorbing and the process stops evolving[17]. The pseudo-code for the simulation is the following.

Algorithm 2 Axelrod's multi-agent system for the dissemination of culture.

repeat
  Select at random a site i and any of its neighbors j
  Calculate the number of common features: $l_{ij} = \sum_{f=1}^{F} \delta v_f^i v_f^j$. Edge ij is active if $l_{ij} < 0$.
  if the edge is active and there exist a non-common cultural feature **then**
    Change the value of one non common feature of agent j to that of agent i with probability l_{ij}/F
  end if
until No active bond exist

If cultural features are randomly assigned to nodes of a network, in accordance with a principle of homophyly, the cultural profiles of adjacent nodes become more similar and then more likely to interact in the next iteration of the dynamic process. Axelrod used the relative size of the largest homogeneous domain S_{max}/n, as an order-disorder measure and observed a phase transition of S_{max}/n as function of the number of traits q.

Note that equilibria of graphical games as well as steady state of multi-agent systems are not necessarily optimal under a global point of view. More precisely, if a global objective function is defined to aggregate the individual utilities $\phi(u_1(\mathbf{v}_{N(1)} \mid \mathcal{G}), \ldots, u_n(\mathbf{v}_{N(n)} \mid \mathcal{G}))$ then an interesting problem might be to compare in term of ϕ the equilibrium solution with the optimal solution. In this context the *price of anarchy* [177] is a popular measure of inefficiency of an equilibrium. It is defined as the ratio between the worst objective function value of an equilibrium of the game and that of an optimal outcome (social optimum).

Sometimes the set of strategies to be played by player i is given by the set of nodes to be included in $N(i)$, that is to say, i is selecting which nodes he wishes to be linked with. This class of game are generally known as *games of network formation*, as first proposed by Aumann and Myerson [9] and popularized by Jackson [121]. In the original version, players sequentially propose links which are then accepted or rejected.

Myerson [167] suggests to formulate the strategy space of each player as list of other players, so that they simultaneously announce which other players they wish to be connected to. If $\mathbf{s} \in S_1 \times \ldots \times S_n$ is the set of strategies played, then link (i, j) forms if and only if both $j \in \mathbf{s}_j$ and $i \in \mathbf{s}_i$. This simultaneous move game captures the notion of Nash equilibrium by a small computational cost, but with generally large multiplicity of equilibria, as no player possibly suggests any links under the correct expectation that no other player reciprocate.

The contribution of this thesis to the analysis multi-agent systems based on graph-theoretic frameworks is exclusively concerned with the comparative goodness of different network structures with respect to the optimal aggregates utilities (from a point of view of a global planner). The use of a aggregates utilities also allows to model the network formation, as it will be studied in Chapter 6.

[17] The Axelrod's model is a particular version of the *stepping stone model*, first proposed by Kimura [137] in 1953. It constitute an absorbing Markov chain, formalized by an $n \times n$ array of squares, where each square is initially any one of k different colors. For each step, a square is chosen at random and one of its eight neighbors is randomly "contaminated", by assuming the color of that neighbor

1.2 Models of economical interaction

With some explanations and definitions in hand, we introduce in this section a problem which will be extensively analyzed in the third part of this thesis.

Starting from the 1970's, many studies of computer simulation have attempted to reproduce the process of interactions which leads to the emergence of macroscopic social phenomena, such as the spread of rumors and information, the mutual influence [6, 137], the allocation of goods and services [229], the establishment of norms [157, 230], etc. The mathematical and computational results obtained in the context of the previously described graphical games and multi-agent systems casted a light into the ways in which the structure of interpersonal relationship affects the performance of those macroscopic social phenomena, as well as social tasks and processes.

Such an effect of the network structure relates to the famously difficult-to-define term of *social capital* [68, 149, 190, 192], which reflects the *network externality* and the benefit obtained from the existence of a structure of social interactions. Based on the classical taxonomy of goods, *social capital* is (partially) *excludible*, as it might often be possible to keep non-payers from its *consumption*, by marginalization and social exclusion, and is *non-rivalrous*, as the use of the social capital by one agent does not reduce the access, participation and consumption of others (generally the participation and affiliation in social activities and groups neither prevent the membership of others nor damage the capacity of a community to carry out social tasks and processes, but this fact must be precisely evaluated for each specific social task).

	Excludable	Non-excludable
Rivalrous	Private goods	Common goods
Non-rivalrous	Club goods	Public goods

Table 1.3: Taxonomy of commodities.

The application of the classical taxonomy of commodities in accordance with the *rivalrousness* and *excludability*, as shown in Table 1.3, to social capital – in the sense we conceive it – might be further formalized by the aid of the game-theoretic-framework introduced in Subsection 1.1.4.

In that context the payoff of agent $i \in \mathcal{V}$ has been defined in term of the actions of its neighbors in the network $\mathcal{G}$, that is $u_i(\mathbf{v}_{N(i)} \mid \mathcal{G})$. The $\mathcal{G}$-parameterized utility reflects the use/consumption of the relational resource. In the example of economical trades between nodes of a network, it might be seen that the presence of many agents who take part in and advantage of the mutual trades of commodities increases the possibility of trading and finding better opportunities for majority of agents.

To clarify this notion, consider the classical scenario proposed by Harrison and Hirshleife [114], where homogeneous nodes must decide whether to produce (call it action 1) or not (call it action 0) a non-rivalrous good. The action of each node $i \in \mathcal{V}$ represents an effort and his/her utility depends on the aggregate efforts of him/herself and that of his/her neighbors, minus some cost for his/her own effort, as summarized by the utility function:

$$u_i(\mathbf{v} \mid \mathcal{G}) = \begin{cases} 1-c & \text{if } v_i = 1 \\ 1 & \text{if } v_i = 0 \text{ and } v_j = 1 \text{ for some } j \in N(i) \\ 0 & \text{if } v_i = 0 \text{ and } v_j = 0 \text{ for all } j \in N(i) \end{cases}, \tag{1.25}$$

where $1 > c > 0$, $N(i)$ is the set of neighbors of node i and $\mathbf{v}$ is strategy profile of nodes in the specified exogenous network $\mathcal{G}$. It might be easily observed that a Nash equilibrium of this game is such that for any link (i, j), not both nodes i and j decide to produce the commodity; but for any non producing node i (i.e. $v_i = 0$) there must exist a node $j \in N(i)$ such that $v_j = 1$.

If the network structure $\mathcal{G}$ is a star, one of the two equilibria gives rise to the highest possible payoff, which is the one obtained under a global maximization – the central node plays one and the others play zero.

A sensitivity analysis of the effects of adding links to a network on the equilibrium strategy profile has been carried out by Galeotti et al. [99]. He considered an example in which a disconnected network made of two stars with n nodes were in equilibrium, so that the two centers choose one (buying the public good), while the $2(n-1)$ peripheral nodes choose zero (acting as free-riders). The aggregate utility in this equilibrium is $2n - 2c$. If this state is perturbed by adding a link between the centres of the two stars, the old strategy profile is no longer in equilibrium and the two centres has an incentive to act as a free-rider with respect to the other. It might be shown that in any of the equilibria associated with the addition of a link, the aggregate utilities are lower than in the starting equilibrium.

Nonetheless, despite this surprising results of Galeotti et al. [99], if the global maximization is adopted as a solution concept, rather than the one of Nash equilibrium, we see that under no circumstances the solution of the problem of maximizing the aggregate utilities reduces when adding links to the existing network, as the effect of this changes only perturbs the myopic local behavior of nodes, not the one of a global planner (solver).

Using pairwise stability as a solution concept, other games of network formation has been studied by Jackson [121–123]. Well studied cases where the network effect are related to diffusion processes [148]. Suppose the utilities of agents reflect their amount of knowledge. A person-to-person information network is defined, by considering the existence of a non-rivalrous good (the information and knowledge) circulating within the network and giving rise to an *epidemic process*, reaching progressively larger area of the networks.

As noted in Subsection 1.1.4, multi-agent systems, such as the Axelrod's process for the dissemination of culture, represent general approaches to study the diffusion thought networks. Klemm et al. [139] studied the effect of a randomized perturbation of the network structure to the diffusion of cultural attributes in the Axelrod's model. Starting from a 2-dimensional lattice as an initial network configuration, they considered a parameter π as the probability of rewriting a link of the original configuration – in the case $\pi = 0$ we have the 2-dimensional lattice and in the case $\pi = 1$ we have an Erdos Reni Random Graph –. For different values of π they have analyzed the expected phase transaction of S_{max}/n (the relative size of the largest homogeneous domain in Algorithm 2), showing that the more the relational structure is ordered the lower is the phase transaction from polarization to globalization. In other words, the phase transaction increases with the amount of spatial disorder.

Despite the diffusion, circulation and sharing of public goods among network agents has been the center of the attention of several network analysts and game theorists for more than two decades [114, 121–123, 139, 148], the effect of the structure of interaction on the exchange and allocation of private goods certainly received a much more modest considerations [8, 13, 229].

In Chapter 6 the effect of the network structure on the final allocation of private goods is taken into account from the optimization point of view and several barter processes to allocate fixed quantities among self-interested agents will be analyzed.

Consider a collection of m types of commodities, call it $\mathcal{C}$, a commodity space Ξ_i (representing the feasible bundle of commodity that agent $i \in \mathcal{V}$ may hold and usually given by a subset of the nonnegative orthant in $\mathbb{R}^m$), the initial endowments $q_j^i \in \Xi_i$ of agent $i \in \mathcal{V}$ for each commodity $j \in \mathcal{C}$ (corresponding a budget constraints) and utility functions $u^i : \boldsymbol{\Xi} \longmapsto \mathbb{R}$, representing

preference relation $\preceq_i$ on $\boldsymbol{\Xi}$, where $\boldsymbol{\Xi} = \Xi_1 \times \ldots \times \Xi_n$.

Since the very beginning of the Economic Theory [85, 127], this mathematical framework has generally been adopted for the study of markets of excludable and rivalrous goods. When agents attempt to simultaneously maximize their respective utilities, conditioned to balance constraints, the resulting problems are $\max[u^i(\mathbf{v}) : i \in \mathcal{V}]$ subject to $\sum_{i\in\mathcal{V}} v_j^i = \sum_{i\in\mathcal{V}} q_j^i$ for $j \in \mathcal{C}$, where $[v_1^i \ldots v_m^i] \in \Xi_i$ is a bundle of commodities demanded by agent i (the superindex shall denote the agent and the subindex shall denote the commodity). The characterizing element of this problem is that many allocations might be simultaneously suitable for all the agents.

Definition 3. *Let $\mathcal{U} \subset R^n$ be the space of allocations of an n-agents bargaining problem. Points in $\mathcal{U}$ can be compared by saying that $\overline{\mathbf{u}} \in \mathcal{U}$ strictly dominates $\mathbf{u} \in \mathcal{U}$ if each component of $\overline{\mathbf{u}}$ is not less than the corresponding component of $\mathbf{u}$ and at least one component is strictly greater, that is, $u^i \leq \overline{u}^i$ for each i and $u^i < \overline{u}^i$ for some i. This is written as $\mathbf{u} \prec \overline{\mathbf{u}}$. Then, the Pareto frontier is the set of points of $\mathcal{U}$ that are not strictly dominated by others.*

Edgeworth [85] called the Pareto optimal solutions of this problem the *final settlements* and a long-standing line of research focused on axiomatic approaches for the determination of a unique solution (for details, see Nash [168] and Rubinstein [201]).

A parallel line of research is the one concerned with the search of competitive prices. Arrow and Debreu [80] showed that under certain economic conditions (convex preferences, perfect competition and demand independence) there must be a vector of prices $\widehat{P} = (\widehat{p}_1, \widehat{p}_2, \widehat{p}_3, \ldots, \widehat{p}_m)^T$, such that aggregate supplies will equal aggregate demands for every commodity in the economy.

As studied by Dreze [83], when prices are regarded as fixed, markets do not clear and the imbalance between supply and demand is resolved by some kind of quantity rationing. The system of linear constraints associated a n-consumer-m-commodity market with fixed prices exhibits a block angular structure with rank $m + n - 1$:

$$\begin{bmatrix} p_1\, p_2\, \cdots\, p_m & & & & \\ & p_1\, p_2\, \cdots\, p_m & & & \\ & & \ddots & & \\ & & & p_1\, p_2\, \cdots\, p_m & \\ I & I & \ldots & I & \end{bmatrix} \mathbf{v} = \begin{bmatrix} p_1 q_1^1 + \ldots + p_m q_m^1 \\ p_1 q_1^2 + \ldots + p_m q_m^2 \\ \vdots \\ p_1 q_1^n + \ldots + p_m q_m^n \\ \mathbf{q}^1 + \ldots + \mathbf{q}^n \end{bmatrix}, \qquad (1.26)$$

where $p_1, \ldots, p_m$ are relative prices between commodities, $\mathbf{q}^i = (q_1^i, \ldots, q_m^i)^T$, and $\mathbf{v} = (v_1^1, \ldots, v_m^1, \ldots, v_1^n, \ldots, v_m^n)^T$. The constraints matrix of (1.26) could also be written as $\begin{pmatrix} I \otimes P \\ \mathbf{1} \otimes I \end{pmatrix}$, where $P = (p_1, p_2, p_3, \ldots, p_m)$ and $\otimes$is the Kronecker product between two matrices. Note that the linking constrains (i.e., the conservation of commodities $(\mathbf{1} \otimes I)\mathbf{v} = \mathbf{q}^1 + \ldots + \mathbf{q}^n$) are implied by the balance equations of a network flow among the agents.

All the feasible allocations lay in a $(m + n - 1)$ dimensional hyperplane defined by the prices (always containing at least one solution, which is represented by the vector of initial endowments $\mathbf{q}$), and restricted to the fact that agents are rational: $u^i(\mathbf{v}) \geq u^i(\mathbf{q})$, for $i \in \mathcal{V}$.

From a multi-objective optimization point of view, a suitable technique to generate the Pareto efficient allocations is the ε-constraint method, which is based on converting all but one objectives into constraints [111]. Recently, efficient algorithms to find non-dominated Pareto allocations of bargaining problems associated to markets with not infinitesimally divisible goods and fixed exogenous prices have been studied by Vazirani et al. [221] and by Ozlen, Azizoglu and Burton [180, 181], who developed a general approach to generate all nondominated objective vectors, by recursively identifying upper bounds on individual objectives using problems with

fewer objectives[18].

Proposition 1 below shows that an asymptotic approximation of an upper bound of the number of nonnegative solutions of (1.26) is $\mathcal{O}(\frac{n^{(mb)}}{b^m})$, where b is the average amount commodity available, i.e., $b = \frac{\sum_{j=1}^{m}(\sum_{h=1}^{n} v_j^h)}{m}$.

Proposition 1. *Let Ξ be the set of nonnegative integer solutions of (1.26), i.e., the allocation space of a problem of bargaining integer amounts of m commodities among n agents with fixed prices. If the allocation space satisfies the mild conditions $b_j = \sum_{h=1}^{n} v_j^h \geq n$ (where b_j is the overall amount of commodity j in the system, which is a fix quantity, associated to the rhs of (1.26)), then $|\Xi| \in \mathcal{O}(\frac{n^{(mb)}}{b^m})$.*

Proof. The set of nonnegative solutions of (1.26) is a subset of the union of bounded sets, as $\Xi \subset \bigcup_{j=1}^{m}\{(v_j^1 \dots v_j^n) \in \mathbb{R}^n : v_j^1 + \dots + v_j^n = q_j^1 + \dots + q_j^n; v_j^1 \dots v_j^n \geq 0\}$. Therefore, Ξ is a finite set, as it is the intersection between $\mathbb{Z}$ and a bounded subset of $\mathbb{R}^{mn}$. Let Ξ' be the set of nonnegative solutions of (1.26), without considering the price constraints, i.e., the n diagonal blocks $p_1 v_1^h + p_2 v_2^h + \dots + p_m v_m^h = p_1 q_1^h + p_2 q_2^h + \dots + p_m q_m^h$, for $h = 1, \dots, n$. We know that $|\Xi'| \geq |\Xi|$. However, $|\Xi'|$ can be easily calculated, as the number of solutions of m independent Diophantine equations with unitary coefficients. The number of nonnegative integer solutions of any equation of the form $\sum_{h=1}^{n} v_j^h = b_j, j = 1, \dots, m$, might be seen as the number of distributions of b_j balls among m boxes: $\frac{(n+b_j-1)!}{(n-1)!\,b_j!}$. Since we have m independent Diophantine equations of this form, then the number of possible solutions for all of them is $\prod_{j=1}^{m} \frac{(n+b_j-1)!}{(n-1)!\,b_j!}$. Thus, we know that $|\Xi| \leq \prod_{j=1}^{m} \frac{(n+b_j-1)(n+b_j-2)\dots n}{b_j!} \leq \prod_{j=1}^{m} \frac{(n+b_j-1)^{b_j}}{b_j!} \leq \frac{\prod_{j=1}^{m}(n+b_j-1)^{b_j}}{b^m}$, where the last inequality holds because $b_j \geq n \geq 2$. Finally, we conclude that $\frac{\prod_{j=1}^{m}(n+b_j-1)^{b_j}}{b^m} \leq \frac{\mathcal{O}(n)^{bm}}{b^m} \leq \mathcal{O}(\frac{n^{(mb)}}{b^m})$. □

The n-objective optimization problem of maximizing utilities $\max[u^i(\mathbf{v})\ i \in \mathcal{V}]$ subject to $\mathbf{v} \in \Xi$ will be extensively studied in Chapter 6.

An important modeling extension in this context is obtained by introducing a network structure, restricting the flows of commodities between couples of agents. Since no production is allowed, the aggregated stock of commodities stays constant, as expressed by the conservation of commodities $\mathbf{v}^1 + \mathbf{v}^2 + \dots + \mathbf{v}^n = \mathbf{q}^1 + \mathbf{q}^2 + \dots + \mathbf{q}^n$. The balance equations of flow circulating through the network guarantees the conservation of the aggregated stock of commodities:

$$\sum_{k:(k,v)\in\mathcal{E}} x_j^{kv} - \sum_{h:(v,h)\in\mathcal{E}} x_j^{vh} = v_j^v - q_j^v, \qquad c \in \mathcal{C},\ v \in \mathcal{V} \tag{1.27}$$

(Note that in the definition of multi-commodity network flow problems in Subsection 1.1.4, we used subindexes to refer to nodes and superindexes to refer to commodities. Here the opposite convention is assumed.) This formulation allows introducing in the agents' decision the pattern of exchanges with other agents, represented by the flows of commodities: $u^i(\mathbf{v}, \mathbf{x})$, for $i \in \mathcal{V}$, where $\mathbf{v} \in \Xi$, $\mathbf{x} \in \chi - \Xi$ is the space of feasible allocations and χ the space of feasible flows –. The variables of the problem are now v_j^i, which again represent the amount of commodity $j \in \mathcal{C}$ hold by agent $i \in \mathcal{V}$ and x_j^{hk} which constitute the flow of commodity j from agent h to agent k. The set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ in (1.27) represents the set of possible interactions $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$ – a structure

[18] In the case of particularly difficult combinatorial constraints, the use of evolutionary algorithms has been quite popular in the last decades. A review of the wide range of approaches proposed to solve multi-objective integer optimization problems can be found in [79].

of geographical proximity and reachability –, which can be expressed in term of an incidence matrix $N(\mathcal{G})$ and the resulting mathematical programming formulation is the following:

$$\max \quad [u^i(\mathbf{v},\mathbf{x}), i = 1,\dots,n] \tag{1.28a}$$

subject to

$$\left[\begin{array}{ccc|ccc} P & & & & & \\ & \ddots & & & & \\ & & P & & & \\ \hline & & & N(\mathcal{G}) & & \\ & I & & & \ddots & \\ & & & & & N(\mathcal{G}) \end{array}\right] \left[\begin{array}{c} \mathbf{v} \\ \mathbf{x} \end{array}\right] = \left[\begin{array}{c} b^1 \\ \vdots \\ b^n \\ \mathbf{q} \end{array}\right] \tag{1.28b}$$

$$\begin{array}{l} u^i(\mathbf{v},\mathbf{x}) \geq u^i(\mathbf{q},\mathbf{0}) \quad i = 1\dots,n \\ \mathbf{v} \in \mathbb{Z}^{mn} \geq 0, \qquad \mathbf{x} \in \mathbb{Z}^{mn(n-1)} \geq 0, \end{array} \tag{1.28c}$$

where $u^i : \Xi \times \chi \to \mathbb{R}$, $P \in \mathbb{Q}^{1\times m}$, $b^i \in \mathbb{Q}$, $i = 1,\dots,n$, $N(\mathcal{G}) \in \{-1,0,1\}^{n\times n(n-1)}$, and $\mathbf{q} \in \mathbb{Q}^{nm}$.

The conditions $\tilde{u}^i(\mathbf{y},\mathbf{x}) \geq \tilde{u}^i(\mathbf{q},\mathbf{0}), \quad i = 1\dots,n$, guarantee that no agent gets worse under a feasible reallocation, which is known in general bargaining literature as the *disagreement point* ($\tilde{u}^i(\mathbf{q},\mathbf{0})$ is the utility function of agent i evaluated in the initial endowments $\mathbf{q}$ with null flow).

From this view, the creation of interpersonal ties, as a production of a quasi-public-goods, results in positive or negative externality, since the marginal utility of a consumer depends on the interpersonal ties of other agents in the economy.

This understanding of the social capital in terms of the network effect, where the utility of an agent is dependent on the others connected in the network, helps internalizing the analysis of the complex structure of social interaction within the general framework of Microeconomic Theory and Operation Research.

The effect of the network structure $\mathcal{G}$ on the problem of allocating private goods among a population of self-interested agents has been computationally investigated by Wilhite [229], taking into account three economic outputs of the system: price volatility, number of trades, number of searches (how many times possible trades have been negotiated but not performed). A sequence of bilateral interactions among agents with CobbDouglas utility functions has been performed; forcing agents to exclusively get into trades with adjacent neighbors. The result obtained by Wilhite's computational experiments are shown in the Table 1.4.

$\mathcal{G}$	Price (standard deviation)	Total trades	Total searches
Complete graph	1.0046 (0.00168)	1953	2.015.960
Local disconnected	1.0396 (0.2771)	1728	31.590
Local connected	1.0048 (0.0146)	93.976	2.734.270
Small World	1.0045 (0.00724)	45.945	1.236.954

Table 1.4: Each of the three economic outputs has been studied with respect to four different network structure with 500 nodes: i) a complete graph, ii) a local disconnected graph, iii) a local connected graph, iv) a small-world network.

In the complete network every agent can negotiate and trade with any other agent in the population. Thus, the most advantageous trades occur first and a global equilibrium is reached

quickly. On average, there was little variation around the equilibrium price in the global networks. Specifically, the average standard deviation was 0.00168 for an equilibrium price of 1.0046. This price was reached quickly, in 8.08 rounds of trading, on average, and required about 1953 trades. This speed of convergence, however, comes at a cost. In this open network each trade follows an extensive search involving all 500 agents, that is, every agent negotiates a unique price with every other agent in the economy.

In Chapters 6 the effect of the network structure on the allocation and production of private and public goods will be taken into account from the point of view of global the optimization.

1.3 Objectives and contributions

After the description of the state of the art approaches for the analysis for CNs held in the previous sections, it remains to address a precise list of objectives and motivations which links the contribution of this thesis to the existing and previously described lines of research.

Foremost, our primary goal is to apply methodologies coming from Combinatorial Optimization, Linear Programming and Interior Point Methods to the development of efficient statistical simulation procedure to sample from "highly combinatorial" spaces. From this outlook, two aspects will appear of particular importance:

- the algebraic characterization of families of networks with fixed structural properties by systems of linear constraints (Chapter 3);
- the construction of Mathematical Programming methods to provide random basic feasible solutions of those systems (chapters 4 and 5).

In the respect, the following contributions are worth to be mentioned:

- In Chapter 3 an efficient random graph generator will be described, allowing drawing samples from families of networks with complex combinatorial properties by means of LP-based methods.
- A probability density function of the primal-dual solution (as in (1.20)) will be derived from the KKT conditions of an LP, as shown in Section 5.3.
- A specialized interior point approach to deal with primal-block angular LPs will be studied in Chapter 4, allowing to increase the efficiency of the network generation procedures described in Chapter 3.

As previously mentioned in Subsection 1.1.4, plenty of researches are currently dealing with big network problems [1], particularly in transportation networks, routing, logistic and epidemiology, but very few of them are concerned with applications in statistical simulation, which represents a necessary mathematical and computational tool when working with random models of network formation [165, 216]. The goal of chapters 3, 4 and 5 is to to fill this gap, by initiating a novel area of application of the general Optimization Theory in different classes of probabilistic models of CNs.

A secondary goal is to extend the range of application of the described approaches to the studies of strategic models of network formation (Chapter 6). In the respect, the following contributions are worth to be mentioned:

- a mathematical linkage between optimization-like properties and multi-agents properties of a strategic model of network formation has been established in Chapter 6;

- in the subsections 5.5.1 and 6.2.5 the analyzed model will included within the probabilistic framework, opening various possibilities of futures investigations;
- an application of this modeling framework to answer specific questions of computational economics will be addressed in Section 6.3, when solving problems related to (1.28).

Finally, specialized optimization-based methodologies will be also applied to the descriptive analysis of structural properties and the problem of low-dimensional representation of network structures, resulting in a successful approach to deal with the problem of assessing the goodness of fit of random network models, as discussed in Chapter 2 and 5.

An intense application of optimization methods will constitute a binding element of the different parts of the thesis, allowing the formulation of a quite general and comprehensive mathematical programming based framework for network analysis.

Chapter 2

Novel representation of network structures by spectral theory considerations

2.1 Purposes and preliminary overview

In the previous chapter, highly connected valued networks verifying Granovetter's hypothesis and Boorman principle have been described in terms of

- *weakness of local bridges,*
- *association between structural similarities and tie strengths,*
- *heterogenity/polarization of structural positions.*

To deal with these concepts the following nodal quantities have been introduced:

- *relative sociability*: $1 - f_i\Delta_{\max}^{-1}$,
- *relative dispersion*: $\Delta_{\max}^{-2}\sum_{j\in\mathcal{V}} x_{ij}^2$,
- *association between structural similarity and tie strength*: $\Delta_{\max}^{-1}\sum_{j\in\mathcal{V}} x_{ij}\Pi(x_i, x_j)$,

for $i \in \mathcal{V}$, where x_i is the i^{th} column vector of the AM and and $\Delta_{\max}$ is the highest degree in the network ($\Delta_{\max} = \max\{f_1, \ldots, f_n\}$). Based on general properties of spectral graph theory, it will be shown that these nodal quantities are algebraically related to the structural similarity, the network connectivity and the triadic closure, allowing for a formal mathematical treatment of the sociological studies of open and closed systems.

In their construction and development, the main results of this chapter follow a *bottom–up* design, in the sense that they have been obtained starting from a numerical evidence and then providing theoretical reasons to such observations. Nonetheless, the way in which these results are here explained follows an opposite approach: spectral network properties are algebraically deduced and then numerical results are introduced to support such theoretical statements.

The advantage of looking at spectral properties (eigenvectors and eigenvalues) is that they can in fact be efficiently computed by polynomial algorithms and provides important information

concerning the structure of the network. For instance, Takaguchi and Miyazaki [218] studied spectral properties of a random walk on $\mathcal{V}(\mathcal{G})$ for the case of binary networks exhibiting the small-world property. Let W be a transaction probability matrix of a random walk on $\mathcal{V}(\mathcal{G})$ and let $\{\lambda_1^{(W}, \ldots, \lambda_n^{(W}\}$ be the multiset of eigenvalues of W, which are arranged in descending order of their real parts. Takaguchi and Miyazaki defined $\lambda_{(i,i+1)}^{(W)}$, the nearest neighbor eigenvalue spacing (NNES) between λ_i^W and λ_{i+1}^W, for $i = 1, \ldots, n-1$, and observed the $\lambda_{(1,2)}^{(W)}$ is strictly related with the rewriting probability of a Watt-Strogatz process [226]. Numerical evidence showed that the expected $\lambda_{(1,2)}^{(W)}$ linearly depends on the rewriting probability of a Watt-Strogatz process.

Another example on this line of research has been proposed by Golender [106], who introduced a doubly stochastic matrix representation based on the inverse of $I - L(\mathcal{G})$. In the case of simple graphs, the determinant of $I - L(\mathcal{G})$ is an interesting graph invariant, representing the number of spanning trees of the graph obtained by adding to the original graph a vertex connected with all vertices. Since the number of spanning trees of such a redefined network must be positive, $I - L(\mathcal{G})$ cannot be singular. Golender showed that the doubly stochastic matrix $\Gamma = (I - L(\mathcal{G}))^{-1}$ has many nice properties:

i. Γ is symmetric and entrywise positive iff the underlying network is connected;

ii. $\Gamma_{ii} > \Gamma_{ij}$ for each $i \neq j$ (diagonal maximality);

iii. the metric $\rho(i,j) = \Gamma_{ii} + \Gamma_{jj} - 2\Gamma_{ij}$ verifies the triangular inequality;

iv. for any three different vertices i, j, k, $\Gamma_{ij} > \Gamma_{ik}$ iff the underlying network contains a path from i to j and each path from i to k includes j.

In this chapter, we confine ourselves to undirected valued and binary networks and claim that the sequence $\{\sum_i (1 - (\max_i f_i)^{-1} \lambda_i^{(L)})^k\}_{k=1}^{\infty}$ provides a highly relevant information to study the pattern of transitivity and connectivity characterizing the trade off between open and closed network structures.

2.2 A new doubly-stochastic representation of binary and valued networks

Consider the symmetric doubly stochastic matrix $\Lambda(\mathcal{G}) = I - \Delta_{\max}^{-1} L(\mathcal{G})$. For whatever binary or valued network $\mathcal{G}$, matrix $\Lambda(\mathcal{G})$ belongs to the family of all doubly stochastic matrices of the form $I - L(\mathcal{G})\text{diag}(\mathbf{f} + \mathbf{t})^{-1}$, where $\mathbf{d} = (f_1, \ldots, f_n)^T$ is the row marginal of the AM (degree sequence in the binary case) and $\mathbf{t} = (t_1, \ldots, t_n)^T \in \mathbb{R}^n$ verifies $t_i - t_j = f_j - f_i$, $t_i \geq 0$, for $(i,j) \in \mathcal{V} \times \mathcal{V}$.

We claim that $\Lambda(\mathcal{G})$ represents the unique doubly stochastic matrix of the form $I - L(\mathcal{G})\text{diag}(\mathbf{f} + \mathbf{t})^{-1}$, for a specific value of $\mathbf{t}$, providing maximal information about $\mathcal{G}$, in the sense of minimizing the diagonal elements of the matrix power $(I - L(\mathcal{G})\text{diag}(\mathbf{f} + \mathbf{t})^{-1})^{\tau}$, for every non-negative integer τ.

To see why this is the case, let $\mathfrak{T} = \{\mathbf{t} \in \mathbb{R}^n : t_i - t_j = f_j - f_i, (i,j) \in \mathcal{V} \times \mathcal{V}\} = \mathfrak{T} = \bar{\mathbf{t}} + \mathcal{N}(T)$, where $\bar{\mathbf{t}}$ is an arbitrary point in $\mathfrak{T}$ and $\mathcal{N}(T)$ is the null space of matrix $T \in R^{n(n-1)/2 \times n}$, associated to the linear system defining $\mathfrak{T}$. By Gaussian elimination on T, it might be easily seen that, $(\Delta_{\max} - f_1, \Delta_{\max} - f_2, \ldots, \Delta_{\max} - d_n) \in \mathfrak{T}$ and that the null space of T is simply $\{(\theta, \theta, \ldots, \theta) \in \mathbb{R}^n : \theta \in \mathbb{R}\}$. Thus, we have

$$\mathfrak{T} = \left\{ \begin{bmatrix} \Delta_{\max} + \theta - f_1 \\ \Delta_{\max} + \theta - f_2 \\ \vdots \\ \Delta_{\max} + \theta - f_n \end{bmatrix} ; \theta \in \mathbb{R} \right\}, \tag{2.1}$$

and the space of doubly stochastic matrices $I - L(\mathcal{G})diag(\mathbf{f} + \mathbf{t})^{-1}$, for $\mathbf{t} \in \mathfrak{T}$ is

$$\left\{ \begin{bmatrix} 1 - \dfrac{f_1}{(\Delta_{\max} + \theta)} & \dfrac{x_{12}}{(\Delta_{\max} + \theta)} & \cdots & \dfrac{x_{1n}}{(\Delta_{\max} + \theta)} \\ \dfrac{x_{12}}{(\Delta_{\max} + \theta)} & 1 - \dfrac{f_2}{(\Delta_{\max} + \theta)} & & \\ \vdots & & \ddots & \\ \dfrac{x_{1n}}{(\Delta_{\max} + \theta)} & & & 1 - \dfrac{f_n}{(\Delta_{\max} + \theta)} \end{bmatrix} \in \mathbb{R}^{n \times n} \, : \, \theta \in \mathbb{R} \right\}. \tag{2.2}$$

To ensure that $1 - f_i(\Delta_{\max} + \theta)^{-1} \geq 0$ and $x_{ij}(\Delta_{\max} + \theta)^{-1} \geq 0$, for $i, j = 1, \ldots, n$, it is necessary that $\theta \geq \max\{d_i - \Delta_{\max};\ i = 1, \ldots, n\} = 0$.

Let $\widehat{D}_\theta = \text{diag}(\Delta_{\max} + \theta - f_1, \ \ldots, \ \Delta_{\max} + \theta - d_n)$ and $\Lambda_\theta(\mathcal{G}) = (I - L(\mathcal{G})\widehat{D}_\theta^{-1})$. Expanding the matrix power $\Lambda_\theta(\mathcal{G})^\tau$ we have that for $\tau \geq 2$ the i^{th} diagonal element has the form

$$\Lambda_\theta(\mathcal{G})^\tau_{ii} = (\Lambda_\theta(\mathcal{G})_{ii})^\tau + \sum_{h=0}^{\tau-2} (\Lambda_\theta(\mathcal{G})_{ii})^h \sum_{j \neq i} \Lambda_\theta(\mathcal{G})_{ij} \Lambda_\theta^{(\tau-h-1)}(\mathcal{G})_{ij} = \tag{2.3}$$

$$= \left(1 - \frac{f_i}{\Delta_{\max} + \theta}\right)^\tau + \sum_{h=0}^{\tau-2} \left(1 - \frac{f_i}{\Delta_{\max} + \theta}\right)^h \sum_{j \neq i} \frac{x_{ij}}{\Delta_{\max} + \theta} \Lambda_\theta^{(\tau-h-1)}(\mathcal{G})_{ij}. \tag{2.4}$$

Applying the limit for θ going to infinity, we find $\lim_{\theta \to \infty} \Lambda_\theta(\mathcal{G})^\tau_{ii} = 1$. Since $\Lambda_\theta(\mathcal{G})^\tau$ is doubly stochastic for every τ, then $\lim_{\theta \to \infty} \Lambda_\theta(\mathcal{G})^\tau_{ij} = 0$, for $i \neq j$. In other words, in the large limit of θ, matrix $\Lambda_\theta(\mathcal{G})^\tau$ tends to the identity matrix, for each τ and (more importantly) for each $\mathcal{G}$. The main consequence is that the trace of powers of $\Lambda_\theta(\mathcal{G})^\tau$ becomes less and less dependent (informative) on the structure of $\mathcal{G}$ when θ increases, so that $I - \Delta_{\max}^{-1} L(\mathcal{G}) = \Lambda_\theta(\mathcal{G})|_{\theta=0}$ results to be the most informative doubly stochastic representation of $\mathcal{G}$ within the family $\{I - L(\mathcal{G})diag(\mathbf{d} + \mathbf{t})^{-1} \in \mathbb{R}_+^{n \times n} : \mathbf{t} \in \mathfrak{T}\}$.

Comparing the two doubly stochastic representations $(I - L(\mathcal{G}))^{-1}$ and $I - \Delta_{\max}^{-1} L(\mathcal{G})$ in terms of their diagonal components, we realize that, while $(I - L(\mathcal{G}))^{-1}$ produces a kind of diagonal maximality, as noted by Golender [106], which guarantees that $(I - L(\mathcal{G}))^{-1}_{ii} > (I - L)^{-1}_{ij}$, for all $i \neq j$, matrix $I - \Delta_{\max}^{-1} L(\mathcal{G})$ gives rise to a diagonal minimality inside the family $\{I - L(\mathcal{G})diag(\mathbf{f} + \mathbf{t})^{-1} \in \mathbb{R}_+^{n \times n} : \mathbf{t} \in \mathfrak{T}\}$, as just shown.

A fundamental result about Markov chains is that any ergodic chain has a unique stationary distribution π and, regardless of the initial state, the time-t distribution of the chain converges to π as t tends to infinity. Thus, given a connected valued network represented by $L(\mathcal{G})$, the doubly stochastic matrix $I - \Delta_{\max}^{-1} L(\mathcal{G})$ represents an ergodic Markov process, whose stationary distribution is a vector with components $\pi_i = 1/n$, for $i = 1 \ldots n$.

Definition 4 (Relative pointwise distance and mixing time)**.** *Consider the Markov chain represented by the doubly stochastic matrix $\Lambda = I - \Delta_{\max}^{-1} L(\mathcal{G})$ and its stationary distribution π. The relative pointwise distance between the transaction probability at step τ and the stationary probability is*

$$||\Lambda(\mathcal{G})^\tau - \pi|| = \sum_{i,j \in \mathcal{V}} \frac{|\Lambda(\mathcal{G})^\tau_{ij} - \pi_j|}{\pi_j} \tag{2.5}$$

The mixing time is the number of steps until the total variation distance between the current distribution and the stationary distribution has decreased to ε:

$$\mathcal{T}_{mix}(\varepsilon) = \{t > 0 : ||\Lambda(\mathcal{G})^{\tau} - \pi|| < \varepsilon\} \tag{2.6}$$

The rate of convergence of $||\Lambda(\mathcal{G})^{\tau}_{ij} - \pi||$ depends on $\lambda_1^{(\Lambda}, \ldots, \lambda_n^{(\Lambda}$, the eigenvalues of $\Lambda(\mathcal{G})$.

Tools for proving the mixing time include also arguments based on conductance, which is away to measure how difficult it is to leave a small subgraph. This measure of communication between subsystems relates with our discussion concerning the level of closeness of the network structure. Let $Q_{i,j} = \pi_i \Lambda(\mathcal{G})_{ij} = (\Delta_{\max} K_i)^{-1} x_{ij}$, where K_i is the number of nodes in the connected component in which i is located. If $\mathcal{G}$ is connected, $Q_{i,j} = \pi_i \Lambda(\mathcal{G})_{ij} = \pi_j \Lambda(\mathcal{G})_{(ij)} = (\Delta_{\max}|\mathcal{V}|)^{-1} x_{ij}$ (the chain is time-reversibe). The conductance of a subset of nodes $\mathcal{V}^* \subseteq \mathcal{V}(\mathcal{G})$ of the chain represented by $I - \Delta_{\max}^{-1} L(\mathcal{G})$ is

$$\phi_\Lambda(\mathcal{V}^*) = \frac{\sum_{i \in \mathcal{V}^*, j \in \overline{\mathcal{V}^*}} Q_{i,j}}{\sum_{i \in \mathcal{V}^*} \pi_i} = \frac{1}{\Delta_{\max}|\mathcal{V}^*|} \sum_{i \in \mathcal{V}^*, j \in \overline{\mathcal{V}^*}} x_{ij}, \tag{2.7}$$

where $\overline{\mathcal{V}^*}$ is the complement of $\mathcal{V}^*$ with respect to $\mathcal{V}(\mathcal{G})$. The conductance of the overall chain is defined in (2.8), as the minimum conductance among all subsets of nodes $\mathcal{V}^* \subseteq \mathcal{V}(\mathcal{G})$:

$$\phi_\Lambda = \min_{\mathcal{V}^* \subset \mathcal{V}(\mathcal{G})} \frac{\sum_{i \in \mathcal{V}^*, j \in \overline{\mathcal{V}^*}} Q_{i,j}}{\sum_{i \in \mathcal{V}^*} \pi_i} = \min_{\mathcal{V}^* \subset \mathcal{V}(\mathcal{G})} \frac{1}{\Delta_{\max}|\mathcal{V}_*|} \sum_{i \in \mathcal{V}^*, j \in \overline{\mathcal{V}^*}} x_{ij} \tag{2.8}$$

Theorem 2 (Sinclair and Jerrum [217])**.** *Let $\lambda_s^{(\Lambda}$ be the second absolute greatest eigenvalue of $\Lambda(\mathcal{G})$, i.e. $\lambda_s^{(\Lambda} = \max\{|\lambda_2^{(\Lambda}|, |\lambda_n^{(\Lambda}|\}$. The following three inequality must hold:*

$$\sum_{i,j \in \mathcal{V}} ||n\Lambda^{\tau}(\mathcal{G})_{ij} - 1|| \leq n(\lambda_s^{(\Lambda})^{\tau} \tag{2.9a}$$

$$1 - 2\phi_\Lambda \leq \lambda_s^{(\Lambda} \leq 1 - \frac{(\phi_\Lambda)^2}{2} \tag{2.9b}$$

$$(1 - 2\phi_\Lambda)^{\tau} \leq \sum_{i,j \in \mathcal{V}} ||n\Lambda^{\tau}(\mathcal{G})_{ij} - 1|| \leq n\left(1 - \frac{(\phi_\Lambda)^2}{2}\right)^{\tau} \tag{2.9c}$$

From theorem 2 and from the discussion of the previous chapter, we learned that a high conductance leads to faster convergence but also to a poor association between structural similarity and tie strength (community structure) and low level of triadic closure. Beside, the mixing time of $(I - \Delta_{\max}^{-1} L(\mathcal{G}))^{\tau}$, and the association between structural similarity and tie strength, can be algebraically linked by considering that, for every couple of nodes $i, j \in \mathcal{V}(\mathcal{G})$, the measure of structural similarity (1.2) represents the transaction probability of moving from i to j in two steps: $\Pi(x_i, x_j) = \Lambda^2(\mathcal{G})_{ij}$.

This fact suggests the possibility of considering structural similarities of higher order, that is to say, how similar nodes are with respect to their pattern of structural similarity itself. This idea is captured by further exponentiation of $\Lambda(\mathcal{G})$.

For $\tau \geq 0$, let $\Pi_{(\tau)}(x_i, x_j) = \Lambda^{\tau+1}(\mathcal{G})_{ij}$, $\Lambda^{2\tau}(\mathcal{G})_{ij} = \sum_k \Pi_{(\tau-1)}(x_i, x_k) \cdot \Pi_{(\tau-1)}(x_j, x_k)$ and $\Lambda^{\tau}(\mathcal{G})_{ij} = \sum_k \Pi_{(\tau)}(x_i, x_k) \cdot \Lambda(\mathcal{G})_{jk}$. We have that $\Pi_{(\tau)}(x_i, x_j)$ is a convex combination of

$\Pi_{(\tau-1)}(x_i, x_1), \ldots, \Pi_{(\tau-1)}(x_i, x_n)$, where the loads of the combination are indeed the j's normalized strengths of the ties: $x_{j1}/\Delta_{\max}, \ldots, 1 - d_j/\Delta_{\max}, \ldots, x_{jn}/\Delta_{\max}$.

Given a node $i \in \mathcal{V}(\mathcal{G})$, the information provided by $\Pi_{(\tau)}(x_i, x_i)$, i.e. the i's diagonal component of $\Lambda^\tau(\mathcal{G})$, has a straightforward interpretation:

- $\Lambda(\mathcal{G})_{ii}$ tells about the range of the degree distribution (the presence of hubs, which make the ratio between f_i and $\Delta_{\max}$ go to zero);
- $\Lambda^2(\mathcal{G})_{ii}$ constitutes a measure of dispersion of the strength of i' ties, i.e. if i has few strong friends or plenty of weak acquaintances;
- $\Lambda^3(\mathcal{G})_{ii}$, which is the scalar product between $[\Pi(x_i, x_1) \ldots \Pi(x_i, x_n)]$ and $[x_{i1}/\Delta_{\max}, \ldots, 1 - f_i/\Delta_{\max} \ldots x_{in}/\Delta_{\max}]$, measures the association between the tie strengths of node i and its structural similarities, which is Granovetter's main hypothesis.

In general, the i^{th} diagonal component of $\Lambda^\tau(\mathcal{G})$ gives information about the association between the $(\tau-1)$-step walks linking i with its neighbors and the strength of the respective ties with them, which is the association between structural similarities of order $\tau - 1$ and the tie strengths. Expanding the matrix power $\Lambda^\tau(\mathcal{G})$ as in (2.3) and defining $\Pi_{(0)}(x_i, x_j) = \Delta_{\max}^{-1} x_{ij}$, for $\tau \geq 2$ we observe that

$$\Lambda^\tau(\mathcal{G})_{ii} = (1 - \Delta_{\max}^{-1} f_i)^\tau + \Delta_{\max}^{-1} \sum_{h=0}^{\tau-2} (1 - \Delta_{\max}^{-1} f_i)^h \sum_{j \neq i} x_{ij} \Pi_{(\tau-h-2)}(x_i, x_j). \tag{2.10}$$

When dealing with homogeneous row marginals of the AM it reduces to

$$\Lambda^\tau(\mathcal{G})_{ii} = \Delta_{\max}^{-1} \sum_{j \neq i} x_{ij} \Pi_{(\tau-2)}(x_i, x_j), \tag{2.11}$$

which is the association between the structural similarities of order $\tau-2$ and the tie strengths. We call IMARSIM (a short form for *Iterative MARkovian SIMilarities*) this repeated exponentiation of $\Lambda(\mathcal{G})$.

Let $S_{(\tau-h-2)}(i) = \Delta_{\max}^{-1} \sum_{j \in \mathcal{V}} x_{ij} \Pi_{(\tau-h-2)}(x_i, x_j)$ and $r(i) = (1 - \Delta_{\max}^{-1} f_i)$. For a fixed τ, we have that $\sum_{h=0}^{\tau-2} r^h(i) S_{(\tau-h-2)}(i)$ is a linear combination of the association between the strengths of the ties and the structural similarities of order $\tau - h - 2$, for $h = 0, \ldots, \tau - 2$. The loads of this combination decreases when the order of the similarity decreases ($r^h(i)$ goes to zero as h increases if the network has non uniform row marginals of the AM), suggesting that similarities of lower order will have smaller and smaller effect in the level of $\Lambda^\tau(\mathcal{G})_{ii}$. Given the equality between the trace of a matrix and the sum of its eigenvalues and recognizing in $\sum_{h=0}^{\tau-2} r^h(i)$ a geometric series, we find that

$$\begin{aligned} \sum_{i \in \mathcal{V}(\mathcal{G})} (1 - \Delta_{\max}^{-1} \lambda_i^{(L)})^\tau &= \sum_{i \in \mathcal{V}(\mathcal{G})} \Lambda(\mathcal{G})_{ii}^\tau \\ &\leq \sum_{i \in \mathcal{V}(\mathcal{G})} \left(r^\tau(i) + \frac{1 - r^{\tau-1}(i)}{1 - r(i)} \right) \\ &\leq \sum_{i \in \mathcal{V}(\mathcal{G})} \frac{1 + (1 - \Delta_{\max}^{-1} f_i)^{\tau-1}}{\Delta_{\max}^{-1} f_i} \\ &\leq n \frac{\Delta_{\max}}{\Delta_{\min}} \left(1 + \left(\frac{\Delta_{\max} - \Delta_{\min}}{\Delta_{\max}} \right)^\tau \right), \end{aligned} \tag{2.12}$$

which goes to $n\Delta_{\max}/\Delta_{\min}$ in the large limit of τ and the speed of convergence depends on the relative range $(\Delta_{\max} - \Delta_{\min})/\Delta_{\max}$ ($\Delta_{\min}$ is the minimum row marginal of the AM).

A similar reasoning can also be applied to deduce a mathematical connection between the range $(\Delta_{\max} - \Delta_{\min})$ and the generalized eigenvector centrality measure proposed by Bonacich [32, 33], as introduced in Subsection 1.1.4, i.e. $c(\alpha, \beta) = \alpha(I - \beta X(\mathcal{G}))^{-1} X(\mathcal{G})\mathbf{e}$, where $\mathbf{e}$ is a vector of ones, β reflects the degree to which a node's centrality is a function of the centralities of its neighbors and α is a scaling parameter.

Since $X(\mathcal{G}) = \Delta_{\max}(\Lambda(\mathcal{G}) - I) + D(\mathcal{G})$ and that $\text{diag}(\Lambda(\mathcal{G})) = I - \Delta_{\max}^{-1} D(\mathcal{G})$, we have

$$\begin{aligned} \alpha(I - \beta X(\mathcal{G}))^{-1} X(\mathcal{G})\mathbf{e} &= \alpha \left(\sum_{k=1}^{\infty} \beta^{k-1} X(\mathcal{G})^k \right) \mathbf{e} \\ &= \alpha \left(\sum_{k=1}^{\infty} \Delta_{\max}^k \beta^{k-1} (\Lambda(\mathcal{G}) - \text{diag}(\Lambda(\mathcal{G})))^k \right) \mathbf{e}. \end{aligned} \tag{2.13a}$$

The infinite sum in (2.13a) converges only if $|\beta| < 1/\lambda_1^{(A}$. By the definition of $\text{diag}(\Lambda(\mathcal{G}))$ and $\Lambda(\mathcal{G})$, it might be proved that $(\Lambda(\mathcal{G}) - \text{diag}(\Lambda(\mathcal{G})))^k \mathbf{e} \leq (I - \text{diag}(\Lambda(\mathcal{G}))^k)\mathbf{e}$, for each $k \geq 1$[1]. We find

$$\begin{aligned} c(\alpha, \beta) &\leq \alpha \left(\sum_{k=1}^{\infty} \Delta_{\max}^k \beta^{k-1} (I - \text{diag}(\Lambda(\mathcal{G}))^k) \right) \mathbf{e} \\ &= \frac{\alpha}{\beta} \left(\sum_{k=1}^{\infty} (\beta \Delta_{\max} I)^k - \sum_{k=1}^{\infty} (\beta \Delta_{\max} \text{diag}(\Lambda(\mathcal{G})))^k \right) \mathbf{e} \\ &= \frac{\alpha}{\beta} \left(\frac{\beta \Delta_{\max}}{(1 - \beta \Delta_{\max})} I - \widehat{D} \right) \mathbf{e} \end{aligned} \tag{2.13c}$$

Matrix $\widehat{D}$ is diagonal with elements $\beta(\Delta_{\max} - f_i)/(1 - \beta(\Delta_{\max} - f_i))$, for $i = 1, \ldots, n$. Thus, we obtain that the centrality of the i^{th} node is

$$\begin{aligned} c_i(\alpha, \beta) &\leq \alpha \frac{(1 - \beta \Delta_{\max})\Delta_{\max} - (\Delta_{\max} - f_i)(1 - \beta \Delta_{\max})}{(1 - \beta \Delta_{\max})(1 - \beta(\Delta_{\max} - f_i)} \\ &= \alpha \frac{f_i}{1 - \beta \Delta_{\max}(\Delta_{\max} - f_i)} \\ &\leq \alpha \frac{\Delta_{\max}}{1 - \beta \Delta_{\max}(\Delta_{\max} - \Delta_{\min})} \end{aligned} \tag{2.13d}$$

The upper bound of the generalized eigenvector centrality measure [32, 33] depends again on the range $\Delta_{\max} - \Delta_{\min}$. As mentioned in Subsection 1.1.3, in sparse networks high range of the degree sequence facilitates the coexistence of triadic closure and high connectivity.

A high range of the degree sequence might be the result of the concentration of nodes in the diagonal cells of the nodal taxonomy in Tables 1.2, where the endpoints of strong highly embedded ties and weak local bridges are grouped, in accordance with the nodal embeddedness and degree. Consider the partition $\mathcal{V} = \Gamma_{11} \cup \Gamma_{12} \cup \Gamma_{21} \cup \Gamma_{22}$, in accordance with the nodal

[1] We prove it by induction. For $k = 1$ we have $\Delta_{\max}^{-1} f_i \leq \Delta_{\max}^{-1} f_i$. For the inductive step, note that $(\Lambda(\mathcal{G}) - \text{diag}(\Lambda(\mathcal{G})))^k = \Delta_{\max}^{-k} X(\mathcal{G})^k$ and that each element of $X^k(\mathcal{G})\mathbf{e}$ is bounded from above by $\Delta_{\max}^k \mathbf{e}$. We claim that, if $\Delta_{\max}^{-k} X(\mathcal{G})^k \mathbf{e} \leq (I - \text{diag}(\Lambda(\mathcal{G}))^k)\mathbf{e}$, for an arbitrary positive integer k, then $\Delta_{\max}^{-(k+1)} X(\mathcal{G})^{k+1} \mathbf{e} \leq (I - \text{diag}(\Lambda(\mathcal{G}))^{k+1})\mathbf{e}$. To see why this is the case, note that

$$\begin{aligned} X(\mathcal{G})^{k+1}\mathbf{e} &\leq \Delta_{\max} X(\mathcal{G})^k \mathbf{e} \\ \Delta_{\max}^{-(k+1)} X(\mathcal{G})^{k+1}\mathbf{e} &\leq \Delta_{\max}^{-(k+1)} \Delta_{\max} X(\mathcal{G})^k \mathbf{e} \\ \Delta_{\max}^{-(k+1)} X(\mathcal{G})^{k+1}\mathbf{e} &\leq \Delta_{\max}^{-k} X(\mathcal{G})^k \mathbf{e}. \end{aligned} \tag{2.13b}$$

It turns out that, if $\Delta_{\max}^{-k} X(\mathcal{G})^k \leq (I - \text{diag}(\Lambda(\mathcal{G}))^k)\mathbf{e}$ for an arbitrary positive integer k, we have $\Delta_{\max}^{-(k+1)} X^{k+1}(\mathcal{G})\mathbf{e} \leq \Delta_{\max}^{-k} X^k(\mathcal{G})\mathbf{e} \leq (I - \text{diag}(\Lambda(\mathcal{G}))^k)\mathbf{e} \leq (I - \text{diag}(\Lambda(\mathcal{G}))^{k+1})\mathbf{e}$. This means that $(\Lambda(\mathcal{G}) - \text{diag}(\Lambda(\mathcal{G})))^k \mathbf{e} \leq (I - \text{diag}(\Lambda(\mathcal{G}))^k)\mathbf{e}$, for any positive integer k.

taxonomy of Tables 1.2, where Γ_{11} is the set of *conservative* nodes (low degree, high embedness) and Γ_{22} is the set of *adventurer* (high degree, low embedness).

- For each $i \in \Gamma_{11}$, the sequence $\{\Lambda^\tau(\mathcal{G})_{ii}\}_{\tau=1}^{\infty}$ can be approximated by the upper bound $1+\sum_{h=0}^{\tau-2} S_{\tau-h-2}(i)$ (since $r^\tau(i) \sim 1$), which only reflect the association between structural similarities and tie strength.

- For each $i \in \Gamma_{22}$, the sequence $\{\Lambda^\tau(\mathcal{G})_{ii}\}_{\tau=1}^{\infty}$ can be approximated by the lower bound $\{r^\tau(i)\}_{\tau=2}^{\infty}$ (since $S_{\tau-h-2}(i) \sim 0$), which behave as a geometric sequence.

We mentioned that a network with high community structure and weak local bridges might exhibit high connectivity only if most nodes belong to Γ_{11} and Γ_{22}, since the presence of few highly connected hubs increases the network connectivity and might support the appearance of triadic closure even in very sparse networks. In other words, small world valued networks verifying Granovetter's hypothesis and Boorman principle are such that the *adventurer* and the *conservative* nodes are associated to sequences $\{\Lambda(\mathcal{G})_{ii}\}_{\tau=1}^{\infty}$ which quickly decrease starting from a value closed to one. However, for conservative nodes this decreasing pattern depends on the vanishing association between structural similarity of order τ and tie strength, as it could be seen by deducing from (2.10) the following difference:

$$\Lambda(\mathcal{G})_{ii}^{\tau} - \Lambda(\mathcal{G})_{ii}^{\tau+1} \leq \sum_{h=0}^{\tau-2} \sum_{j \neq i} \frac{x_{ij}\left[\Pi_{(\tau-h-2)}(x_i, x_j) - \Pi_{(\tau-h-1)}(x_i, x_j)\right]}{\Delta_{\max}} + g(f_i), \tag{2.14}$$

where $g(f_i) \leq 1$ is a function which only depends on f_i. Thus, under a vanishing association between structural similarity of order τ and tie strength, the expected consequence for small world valued networks verifying Granovetter's hypothesis and Boorman principle is that the sequence $\mathrm{Tr}(\Lambda(\mathcal{G}))$, $\mathrm{Tr}(\Lambda^2(\mathcal{G}))$, ..., $\mathrm{Tr}(\Lambda^\tau(\mathcal{G}))$ is supposed to drop off quickly starting from a value closed to n.

Let $\mathcal{G}_{\mathcal{K},\kappa}$ be the family of valued networks composed by $\mathcal{K}$ disconnected cliques with κ nodes per clique, such that each within-clique connection has value $u > 0$ and the between-clique connections have zero value. It can be seen that, for each network $g_{\mathcal{K},\kappa} \in \mathcal{G}_{\mathcal{K},\kappa}$, we have $\Lambda(g_{\mathcal{K},\kappa})_{ij}$ is either 0 (when i and j belong to different groups) or $(\kappa-1)^{-1}$ (when i and j belong to the same group) for $i \neq j$. By repeated matrix product we observe that

$$\Lambda^\tau(g_{\mathcal{K},\kappa})_{ij} = \frac{\kappa-2}{\kappa-1}\Lambda^{\tau-1}(g_{\mathcal{K},\kappa})_{ij} + \frac{1}{\kappa-1}\Lambda^{\tau-1}(g_{\mathcal{K},\kappa})_{ii} \tag{2.15}$$

$$\Lambda^\tau(g_{\mathcal{K},\kappa})_{ii} = \Lambda^{\tau-1}(g_{\mathcal{K},\kappa})_{ij}, \tag{2.16}$$

for $j \neq i$ belonging to the same group. Solving the difference equation system we find

$$\begin{bmatrix} \Lambda^{\tau+1}(g_{\mathcal{K},\kappa})_{ij} \\ \Lambda^{\tau+1}(g_{\mathcal{K},\kappa})_{ii} \end{bmatrix} = \begin{bmatrix} \frac{\kappa-2}{\kappa-1} & \frac{1}{\kappa-1} \\ 1 & 0 \end{bmatrix}^\tau \begin{bmatrix} \frac{1}{\kappa-1} \\ 0 \end{bmatrix}, \tag{2.17}$$

and by singular value decomposition

$$\begin{bmatrix} \Lambda^{\tau+1}(g_{\mathcal{K},\kappa})_{ij} \\ \Lambda^{\tau+1}(g_{\mathcal{K},\kappa})_{ii} \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{\kappa-1} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{\kappa-1} \end{bmatrix}^\tau \begin{bmatrix} 1 & -\frac{1}{\kappa-1} \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{\kappa-1} \\ 0 \end{bmatrix}. \tag{2.18}$$

This implies

$$\begin{aligned} \Lambda^{\tau+1}(g_{\mathcal{K},\kappa})_{ij} &= \frac{1}{\kappa} + \frac{(-1)^\tau(\kappa-1)^{-\tau}+1}{\kappa(\kappa-1)}, \\ \Lambda^{\tau+1}(g_{\mathcal{K},\kappa})_{ii} &= -\frac{(-1)^\tau(\kappa-1)^{-\tau}-1}{\kappa} \end{aligned} \tag{2.19}$$

$$\mathrm{Tr}(\Lambda^{\tau}(g_{\mathcal{K},\kappa})) - \mathrm{Tr}(\Lambda^{\tau+1}(g_{\mathcal{K},\kappa})) = (-1)^{\tau}\frac{\mathcal{K}\kappa}{(\kappa-1)^{\tau}} \tag{2.20}$$

The network plots in Fig. 2.3 shows three cases of $\mathcal{K}$ disconnected cliques with $\kappa = 2, 3, 4$ nodes per clique.

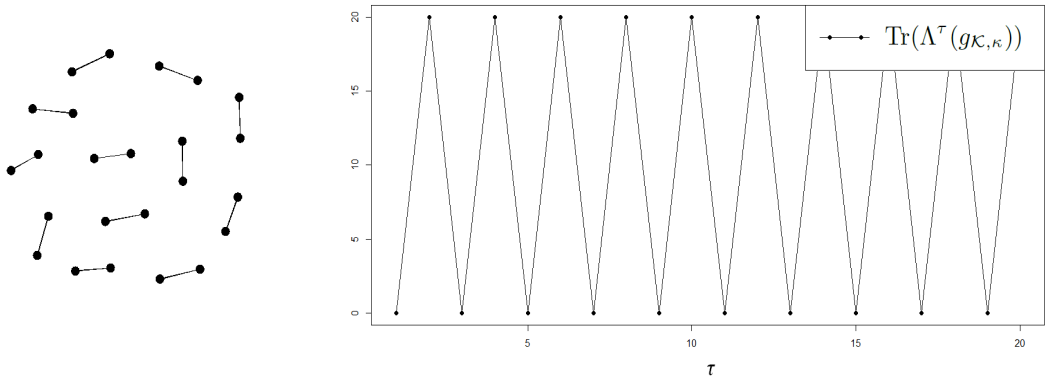


Figure 2.1: Sequence $\mathrm{Tr}(\Lambda^{\tau}(g_{\mathcal{K},\kappa}))$ for $\tau = 1, \ldots, \mathcal{K}\kappa$, associated to disconnected cliques with $(\mathcal{K}, \kappa) = (10, 2)$.

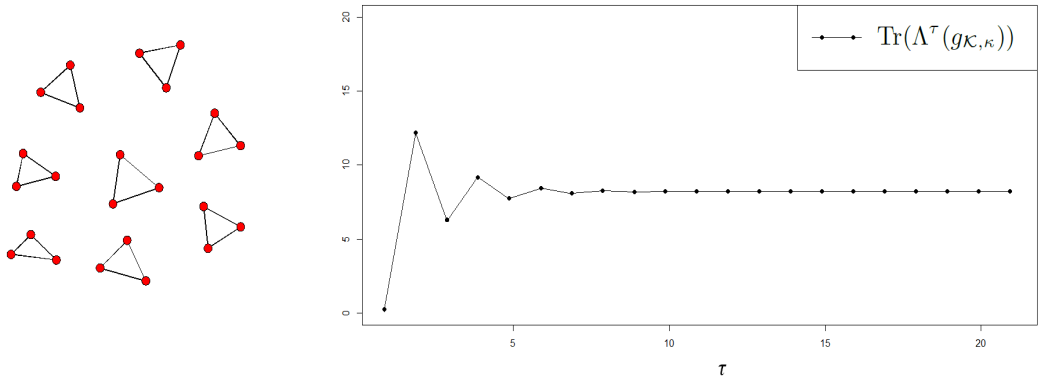


Figure 2.2: Sequence $\mathrm{Tr}(\Lambda^{\tau}(g_{\mathcal{K},\kappa}))$ for $\tau = 1, \ldots, \mathcal{K}\kappa$, associated to disconnected cliques with $(\mathcal{K}, \kappa) = (7, 3)$.

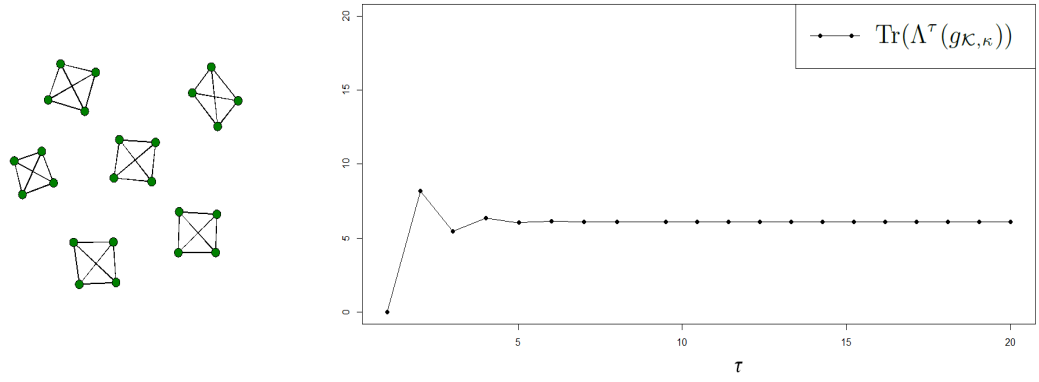


Figure 2.3: Sequence $\mathrm{Tr}(\Lambda^{\tau}(g_{\mathcal{K},\kappa}))$ for $\tau = 1, \ldots, \mathcal{K}\kappa$, associated to disconnected cliques with $(\mathcal{K}, \kappa) = (5, 4)$.

The main idea is that $\Lambda(\mathcal{G})_{ii}^{\tau+1}$ measures how $\tau-$similar the i^{th} node is with the rest of the

network and the recursive definition of $\tau-$similarity comes from the cosine similarity of nodes with respect to the their $(\tau-1)-$similarities.

In the '70s the question about the similarity of similarities motivated the studies of Breiger, Boorman and Arabie [39], who proposed the CONCOR method (a short form of *CONvergence of iterated CORrelations*). As described in Subsection 1.1.4, starting from the Pearson product-moment correlation coefficient among couples of rows of the AM (call it $C_1 \in [-1,1]^{n\times n}$), CONCOR obtains the correlation matrix of the rows of C_1 and iteratively calculates correlations of correlations, converging to a steady-state correlation matrix C_*, whose entries are either -1 or $+1$.

Both CONCOR and IMARSIM are iterative procedures which take a square real matrix as an input and repeatedly apply a measure of similarity between columns vectors: The function

Algorithm 3 Iterative similarity

Set $k=0$ and C_0 as the initial matrix
repeat
$C_{k+1} = \mathrm{Sim}(C_k)$; $k = k+1$.
until Convergence

$\mathrm{Sim} : \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}^{n\times n}$ returns a matrix whose components represent a measure of similarity between the columns of the matrix, passed as input.

CONCOR was originally proposed without a mathematical justification and a spontaneous question could be: what is likely to happen to the sequence of matrices $C_0, C_1, \ldots$ when the measure of similarity used is somehow replaced? The input matrix of CONCOR is $C_0 = X(\mathcal{G})$ and the measure of similarity is the Pearson correlation. In the case of $\Lambda^\tau(\mathcal{G})$ the measure of similarity used is the cosine similarity and the precise mathematical meaning of this iterative process is well known by means of the Markovian Theory, as long as $C_0 = \Lambda(\mathcal{G})$ is adopted as input matrix, representing the network structure.

For each positive integer τ, the set $\{\Lambda(\mathcal{G})^\tau_{ii} \mid i \in \mathcal{V}(\mathcal{G})\}$ represents a distribution of the values obtained by the IMARSIM among the set of nodes. The statistical moments of the associated frequency distribution might be used to summarize the shape of this set of points. For each positive integer τ, we consider the first two moments, giving rise to the average and variance profile of the distribution of the nodal quantities $\{\Lambda(\mathcal{G})^\tau_{ii} \mid i \in \mathcal{V}(\mathcal{G})\}$. The resulting sequences are

$$\mu[\Lambda^\tau(\mathcal{G})] = \frac{1}{n}\,\mathrm{Tr}(\Lambda^\tau(\mathcal{G})) = \frac{1}{n}\sum_{i\in\mathcal{V}(\mathcal{G})} (1-\Delta^{-1}_{\max}\lambda_i^{(L)})^\tau \qquad \tau \geq 1 \qquad (2.21)$$

$$\sigma^2[\Lambda^\tau(\mathcal{G})] = \frac{1}{n}\sum_{i\in\mathcal{V}(\mathcal{G})} (\Lambda^\tau_{ii})^2 - \left(\frac{1}{n}\sum_{i\in\mathcal{V}(\mathcal{G})} \Lambda^\tau_{ii}\right)^2 \qquad \tau \geq 1 \qquad (2.22)$$

representing the means and variances of the diagonal components of $\Lambda^\tau(\mathcal{G})$, so that $\mu[\Lambda^\tau(\mathcal{G})] \in [0,1]$ and $\sigma[\Lambda^\tau(\mathcal{G})] \in [0,1]$, for $\tau \geq 1$. (The square root of the variance is taken to obtain the same unit of measurement as the one used in the original scale.) For their construction in term of structural similarities, the two sequences shall be respectively called *mean similarity spectrum* (MSS from now on) and *variance similarity spectrum* (VSS from now on).

As already mentioned in the previous section, one of the most characterizing features of humans social networks is the rich variation at the individual level. The VSS might capture these heterogeneities and variation of structural positions. For instance, when $\mu[\Lambda(\mathcal{G})]$ is close to one and $\sigma[\Lambda(\mathcal{G})]$ closed to zero, the distribution of the row marginals of the AM must exhibits a hub structure, where most of nodes have low row summation and very few of them have a

particularly high sociability. This is observed in Figure 2.4, for the case of a star (left plot) and a ring (right plot) with 60 nodes.

For the case of a star, $\{\mu[\Lambda^\tau]\}_{\tau=1}^n$ exhibits a slow decrease, starting from $(n-2)/n$, outlining a poor change in structural similarity from order $\tau-2$ to $\tau-1$.

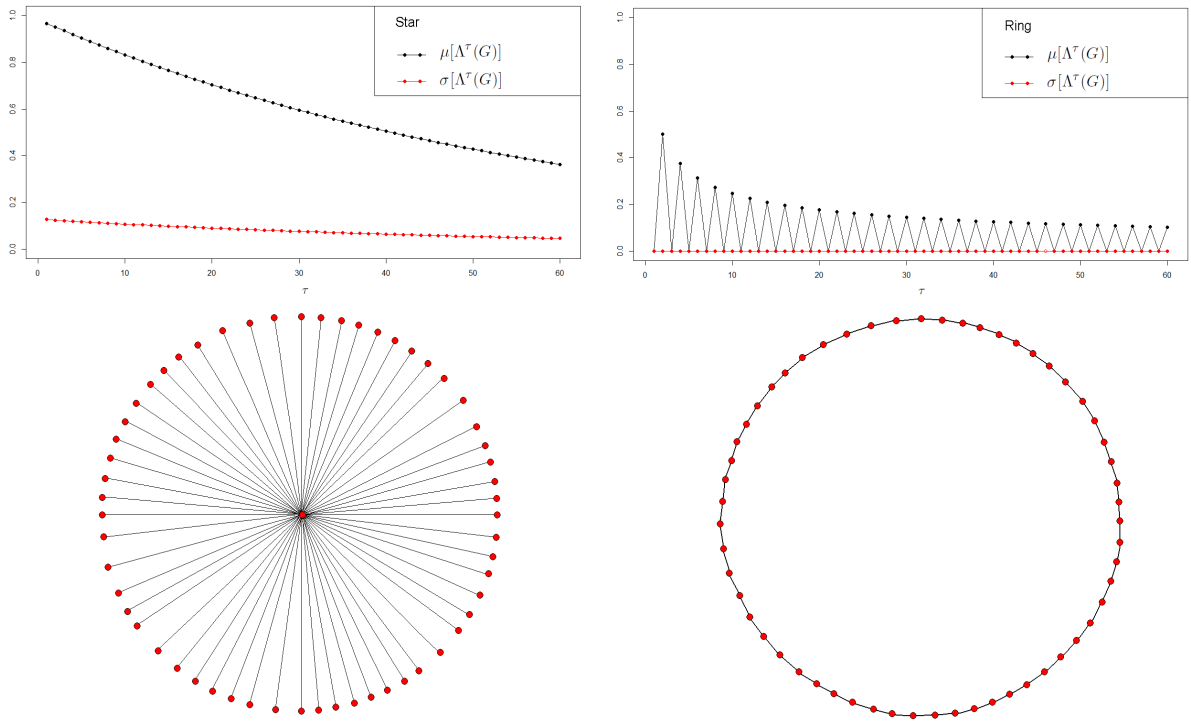


Figure 2.4: The MSS and VSS of a star and a ring with 60 nodes.

As numerically shown in the next section, the MSS and VSS can help to capture the fundamental features of open and closed network structures and operationalize *the weakness of local bridges*, *the association between structural similarities and tie strengths* and the *heterogeneity of structural positions*, as introduced in Subsection 1.1.3. A synthetical overview of the degree of openness/closeness, which reflects the pattern of connectivity, structural similarity and variation at the individual level will be captured in a graphical representation.

2.3 The MSS and VSS of observed network structures

As already mentioned in Section 2.1, the construction and development of the main results of this chapter followed a *bottom–up* design, in the sense that they have been obtained starting from a numerical evidence and then providing theoretical reasons to such observations. Nonetheless, the way in which we decided to present them follows the opposite approach: from the algebraical deduction – as carried out in Section 2.2 – to the empirical observation.

This section provides a detailed analysis of eight popular and well known real world networks, which are available online from the Ucinet [34] database. They are here used to numerically validate and support the effectiveness of the MSS and VSS in capturing important network properties associated to degree of openness/closeness. In particular, we argue that the observed shapes of the MSS and VSS levels, are characterized by the following properties:

- *hub structure of the row marginals, as suggested by the high starting value of the MSS and the low starting value of the VSS*;

- *fast convergence of the sequence* $\mu[\Lambda(\mathcal{G})]$, $\mu[\Lambda^2(\mathcal{G})]$, ..., $\mu[\Lambda^\tau(\mathcal{G})]$, *suggesting he presence of local bridges and supporting a vanishing association between structural similarity of order* τ *and tie strengths*;

- *high variation at individual level, resulting in a slowly decreasing VSS when* τ *increase.*

Table 2.1 provides commonly used network statistics applied to the network data sets studied in the following subsections. The first column reports the name of the data set; the second columns contains the density, computed as the total amount of valued components in the AM divided by the maximal amount. The third and fourth columns show the standard deviation of the AM marginals and one of the Bonacich eigenvector centrality (with $\beta = 1$) respectively.

Dataset	density	edge	AM marginals (sd)	centrality (sd)
Berdnard and Killworth fraternity	0.03778	valued	84.86742	0.47475
Berdnard and Killworth office	0.04792	valued	16.11312	0.40583
Kapferer Tailor shop (first period)	0.21323	binary	4.83297	7.28318
Kapferer Tailor shop (second period)	0.30094	binary	5.50977	3.49072
Dolphins' social network	0.08408	binary	2.95587	3.97711
Newman scientific collaboration	0.00144	valued	3.20884	18.9310
Florentine Reminiscence families	0.16667	valued	2.46885	0.42915
Kapferer's mine	0.23809	binary	1.75933	4.279519

Table 2.1: Descriptive statistics of five network data sets.

The observed MSS and VSS will be compared with their empirical distribution, under conditionally uniform random networks with the corresponding fixed density[2]. The idea is to graphically support the substantial differentiation between the similarity spectra of real-world networks and the ones obtained under the conditionally uniform random model. Such a similarity will be precisely discussed, in accordance with the described structural features introduced in Subsection 1.1.3. The high variation at the individual level will result as a common feature of most of the observed MSS and VSS.

2.3.1 Berdnard and Killworth fraternity

In a series of researches in the 1970s, Bernard, Killworth and Sailer [16–18] considered the accuracy of retrospective sociometric surveys in a religious fraternity with 58 occupants. To quantify the connections among members, an unobtrusive researcher observed the interactions in the fraternity every 15 minutes, 21 hours a day, for five days. The aggregate of these data provided frequency counts represented by a valued network.

Figures 2.5 and 2.6 show respectively the 58 nodes valued network plot and the MSS and VSS associated to the record of the unobtrusive researcher frequency counts.

[2]Methods to generate conditionally uniform random networks will be discussed in chapters 3 and 5. Here we apply the Q-kernel method for valued networks with fixed number of edges, as described in Section 3.3.3.

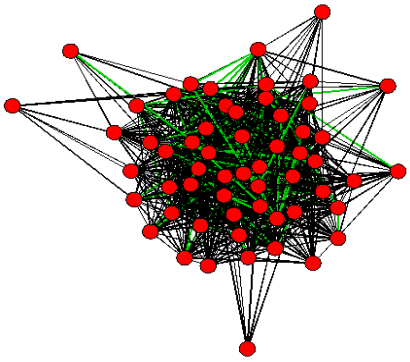


Figure 2.5: Network plot of the Berdnard and Killworth fraternity. The green color represents strong ties (whose value is higher then the median of positive connections), whereas the black color represent weak ties.

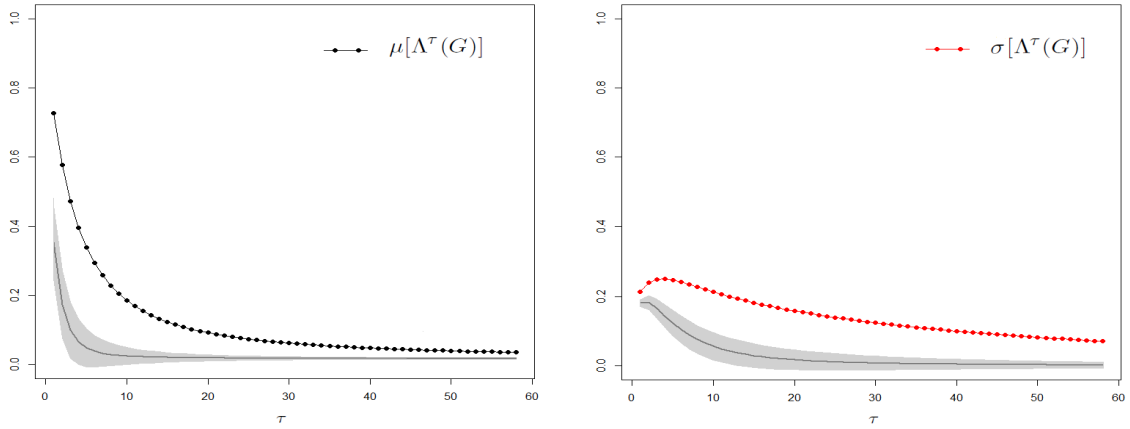


Figure 2.6: MSS and VSS of the Berdnard and Killworth fraternity. The gray area reports the envelop of the MSS and VSS under a simulation of 500 valued networks with fixed sum of the AM elements.

The first interpretation of the MSS and VSS is that they suggest the presence of weakly connected hubs, due to the high starting value of the MSS and low starting value of the VSS. It must be also noted that VSS follows an increasing pattern, for $\tau = 1, \ldots, 4$, entailing high variation at the individual level, as far as the *relative dispersion* and *association between structural similarities and tie strengths* are concerned. High variation at the individual level is suggested by the persistency of the VSS for high values of τ.

2.3.2 Berdnard and Killworth office

Bernard, Killworth and Sailer [16–18] collected network data concerning interactions in a small business office, recorded by an *unobtrusive* observer. The data set contains the observed frequency of interaction, made as the observer patrolled a fixed route through the office every fifteen minutes during two four-day periods.

Figure 2.5 shows the 58 nodes valued network plot and Figure 2.8 the MSS and VSS associated to the record of the unobtrusive researcher frequency counts.

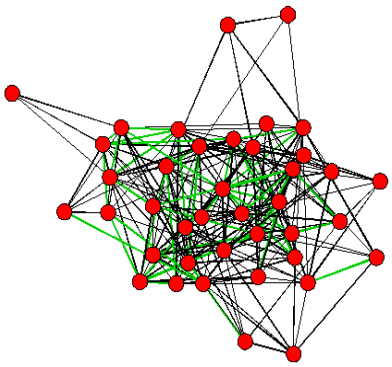


Figure 2.7: Network plot of the Berdnard and Killworth office. The green color represents strong ties (whose value is higher then the median of positive connections), whereas the black color represent weak ties.

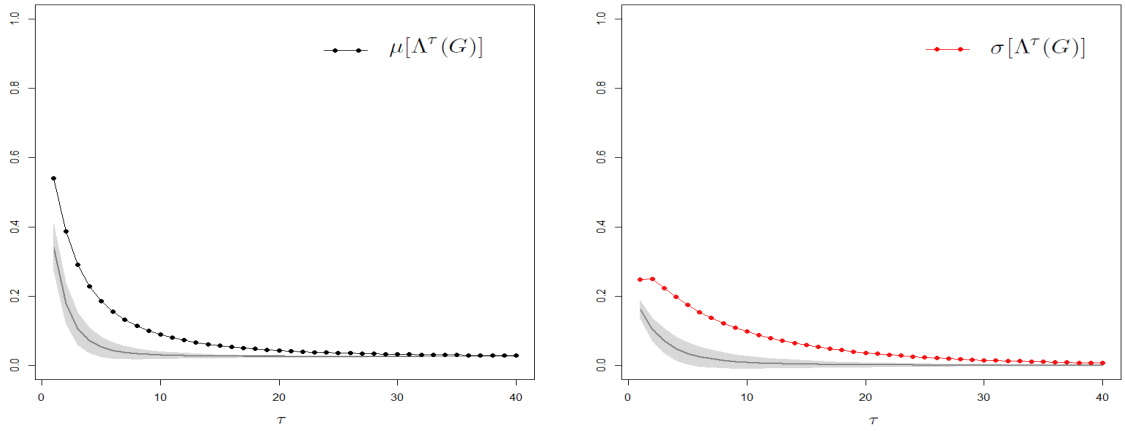


Figure 2.8: MSS and VSS of the Berdnard and Killworth office. The gray area reports the envelop of the MSS and VSS under a simulation of 500 valued networks with fixed sum of the AM elements.

The MSS and VSS suggest the presence of a high variability of the nodes's sociability and *relative dispersion* of the tie strengths within nodes. On the one hand, the highly positive gap $\mathrm{Tr}(\Lambda(\mathcal{G})) - \mathrm{Tr}(\Lambda^2(\mathcal{G}))$ entails a remarkable *relative dispersion* within nodes, as the nodal connections are likely to be overspread almost uniformly, instead of being concentrated in few contacts. On the other hand, the dispersion pattern has a high variation among nodes.

2.3.3 Kapferer Tailor shop

This well-known social network dataset, collected in Zambia in 1965 by Bruce Kapferer [132, 133], recorded interactions among 39 workers in a tailor shop. In this particular dataset, he considered two kinds of interactions: *sociational* (sharing of gossip and the enjoyment of a drink together), as opposed to *instrumental* (work- and assistance-related). The data are recorded at two different times (seven months apart) over a period of one month. Here we only consider sociational interactions, as shown in Figure 2.2, for the network plots of the two respective periods, and figures 2.9 and 2.10 corresponding MSS and VSS.

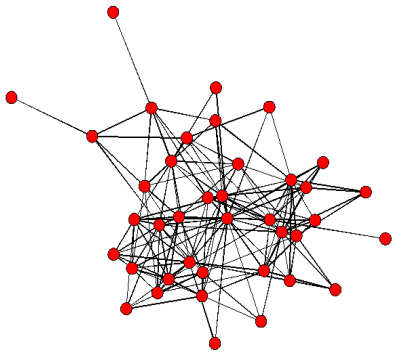
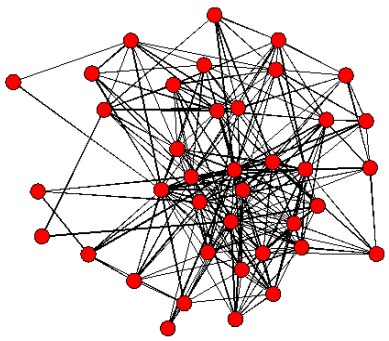
First period	Second period
	

Table 2.2: Network plots of the Berdnard and Killworth office in the first and second periods.

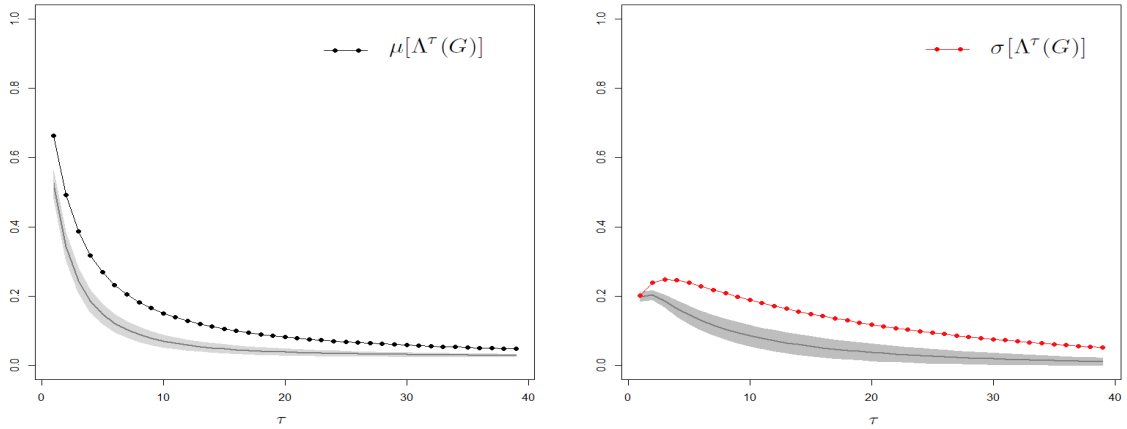


Figure 2.9: MSS and VSS of the Berdnard and Killworth office in the first period. The gray area reports the envelop of the MSS and VSS under a simulation of 500 valued networks with fixed sum of the AM elements.

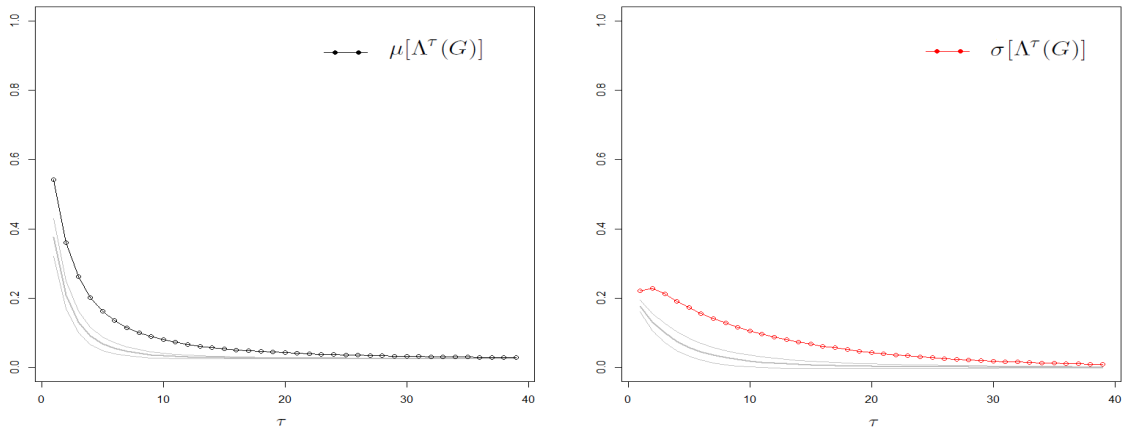


Figure 2.10: MSS and VSS of the Berdnard and Killworth office in the second period. The gray area reports the envelop of the MSS and VSS under a simulation of 500 valued networks with fixed sum of the AM elements.

An increase in density between the first and second period is graphically observed, but it does not seem to correspond to a reduction in the variation at the individual level, as shown by the VSS in figures 2.9 and 2.10. The MSS and VSS associated to the two Kapferer Tailor shop networks quite likely resemble the oners of the Berdnard and Killworth fraternity. In the first period, the positive gap $\mathrm{Tr}(\Lambda(\mathcal{G})) - \mathrm{Tr}(\Lambda^2(\mathcal{G})) = \sum_{i\in\mathcal{V}(\mathcal{G})}[f_i/\Delta_{\max} - (f_i/\Delta_{\max})^2] - \Delta_{\max}^2 D$ in binary networks entails small density and the presence of hubs, whose degree is much higher than the average degree.

2.3.4 Lusseau's dolphins network

Data concerning the frequent associations between dolphins in Doubtful Sound (New Zealand) were collected in a long-term research program of the University of Otago and first analyzed by Lusseau [152] and by Lusseau and Newman [153] [3]. The data represent a binary symmetric network of 62 nodes, encoding the frequent associations between dolphins. Figure 2.11 shows the network plot and Figure 2.12 the associated MSS and VSS.

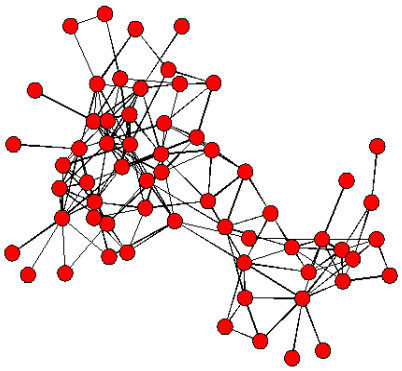


Figure 2.11: Network plot of the Lusseau's dolphins network.

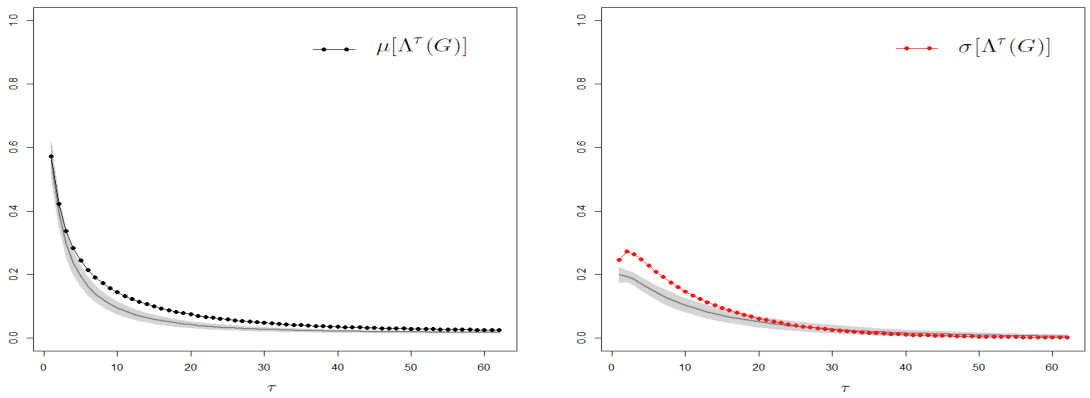


Figure 2.12: MSS and VSS of the Lusseau's dolphins network. The gray area reports the envelop of the MSS and VSS under a simulation of 500 valued networks with fixed sum of the AM elements.

[3] The analysis of animal social networks can offer substantial insights into management strategies, suggesting striking similarities to human networks [71].

The low density of the dolphins' social network justifies starting value of the MSS and VSS, which are close to the expected under randomness, suggesting the absence of a hub structure of the degree sequence. This is similar to what we observed in the Kapferer tailor shop in the second period. Nonetheless, once again, the slowly decreasing pattern of the VSS, supports high variation at the individual level, particularly in the cases of $\tau = 2$ (relative dispersion) and $\tau = 3$ (association between structural similarities and tie strengths).

2.3.5 Newman's scientific collaboration

This dataset consists of 1589 nodes undirected valued network, representing connections between scientists of a specific field. Two scientists are considered connected if they have authored a paper together. Here we are considering the largest connected component of this network, which consists of 372 nodes. Vertices represent authors of scientific paper; and the value of an edge corresponds to how many time two names appear on the same paper.

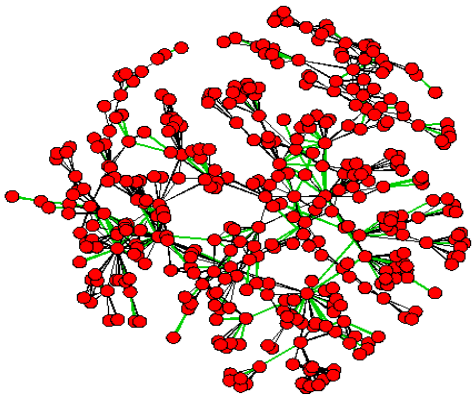


Figure 2.13: Network plot of the Newman's scientific collaboration network. The green color represents strong ties (whose value is higher then the median of positive connections), whereas the black color represent weak ties.

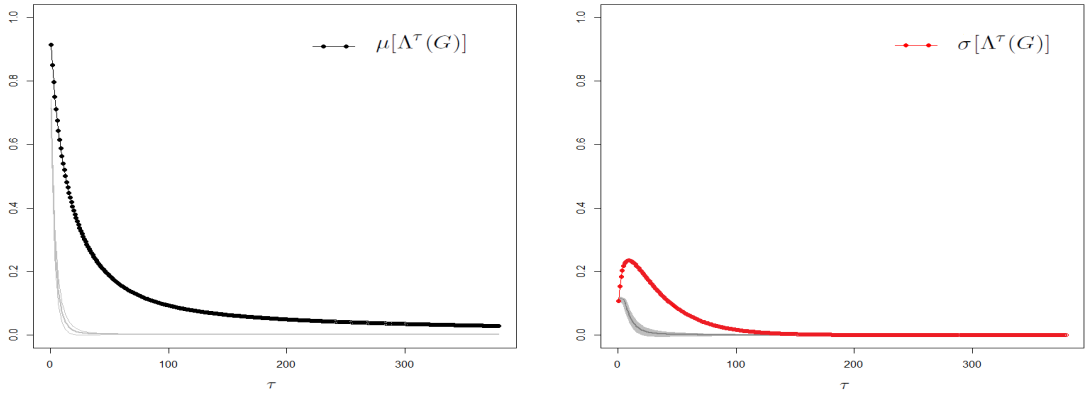


Figure 2.14: MSS and VSS of the Newman's scientific collaboration network. The gray area reports the envelop of the MSS and VSS under a simulation of 500 valued networks with fixed sum of the AM elements.

Newman [169] provided a detailed analysis of this network, after reducing edge in binary values. He showed that randomly chosen pairs of scientists are typically separated by only a short path of intermediate acquaintances and demonstrate the presence of clustering. Figures 2.13 and 2.14 respectively show the network plot and the associated MSS and VSS.

What clearly emerges from figures 2.13 and 2.14 is a hub structures of the row marginals, as suggested by the high starting value of the MSS and the low starting value of the VSS. We also note a fast convergence of the sequence $\mu[\Lambda(\mathcal{G})]$, $\mu[\Lambda^2(\mathcal{G})]$, ..., $\mu[\Lambda^\tau(\mathcal{G})]$, suggesting he presence of local bridges and supporting a vanishing association between structural similarity of order τ and tie strengths.

2.3.6 Renaissance Florentine families

This is a data set of marriage and business relations among Renaissance Florentine families, provided by Padgett [183]. The set of nodes is given by 15 Florentine families. Two sets of edges are takeing into account to represent the two kind of relations: marriage and business. This qualitative relations are here converted into valued ties by regarding as strongly connected those couple of families with both business and marriage relations. Instead, ties among couple of families with either business or marriage relation (but not both) are coded as weak. Formally, for each $(i, j) \in \mathcal{V} \times \mathcal{V}$, $i \neq j$, we have:

$$x_{ij} = \begin{cases} 0 & \text{neither marriage nor business between } i \text{ and } j \\ 1 & \text{ether business interaction or marriage alleance between } i \text{ and } j \\ 2 & \text{both business interaction and marriage alleance between } i \text{ and } j \end{cases} \tag{2.23}$$

The network plots in Table 2.3 shows the two type of ties, as well as the defined valued ties.

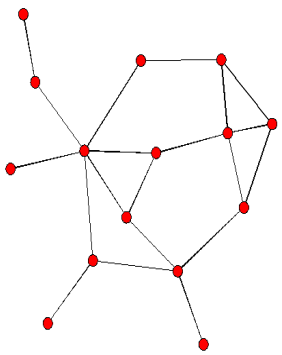
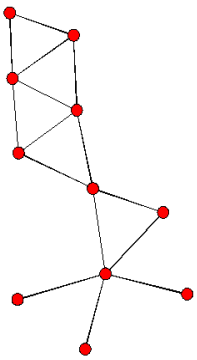
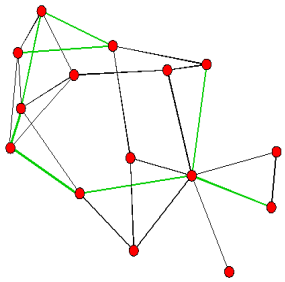
Marriage connections	Business connections	Valued network
		

Table 2.3: From left to right, the three networks in this figure show the marriage alliances, the business interactions and the valued ties with the codification 2.23. The green color represents strong ties (whose value is higher then the median of positive connections), whereas the black color represents weak ties.

The MSS and VSS of the resulting valued network, as numerically codified in (2.23) are shown in Figure 2.15. Likewise the previous network data sets, the VSS of the Renaissance Florentine families exhibit a persistent variation at individual level.

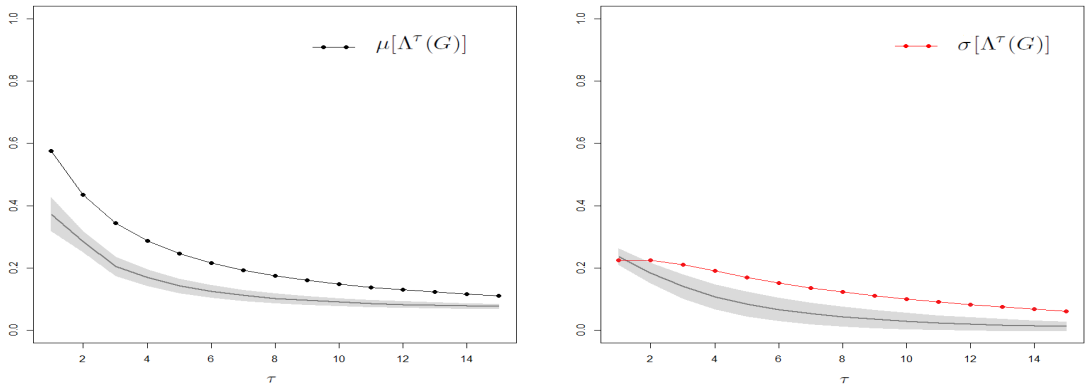


Figure 2.15: MSS and VSS of the network of Renaissance Florentine families. The gray area reports the envelop of the MSS and VSS under a simulation of 500 valued networks with fixed sum of the AM elements.

2.3.7 Kapferer's mine

Bruce Kapferer [132] collected data during a mining operation in Zambia, considering workers involved in several types of interactions: conversation, joking, job assistance, etc. Figure 2.16 reports the network plots associated to one type of relationship.

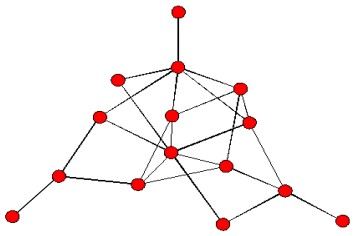


Figure 2.16: Network plot of the Kapferer's miner network.

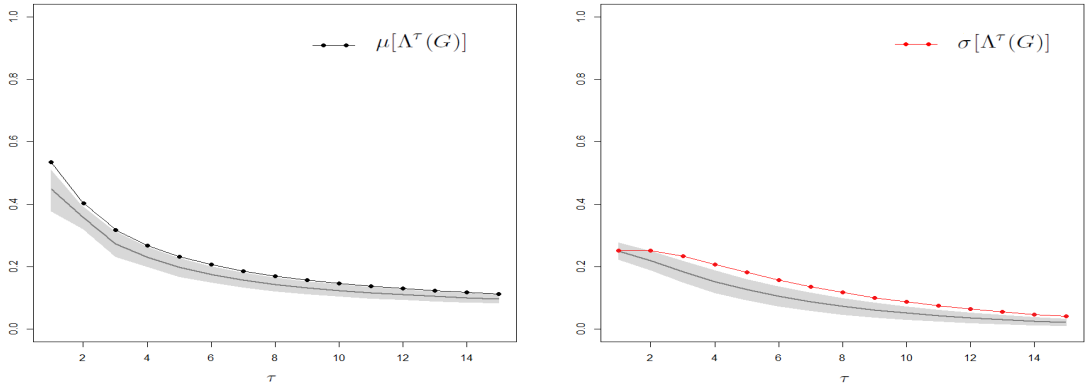


Figure 2.17: MSS and VSS of the Kapferer's mine network. The gray area reports the envelop of the MSS and VSS under a simulation of 500 valued networks with fixed sum of the AM elements.

The MSS and VSS in Figure 2.17 seems to be comparatively closer to the random case, than the ones of the other observed data sets. Nonetheless, even in this small network a hub structures of the row marginals slightly appears, as suggested by the high starting value of the MSS and the low starting value of the VSS. A substantial variation at the individual level is also supported by the VSS.

2.4 The MSS and VSS of simulated networks

In this section we take into account the MSS and VSS of different topological structures generated both by random and deterministic processes of network formation. As already discussed in Subsection 1.1.4, the analysis of network formation tries to explain how specific generation processes operate. The R codes to reproduce the numerical results of this section are available at http://www-eio.upc.edu/~nasini/R_Programs/R_chaper_2.txt.

As a first family of models, let us consider spatially embedded networks, such as random geometric graphs [188]. They consist in random undirected graphs drawn on a bounded region $\mathcal{S}$, in accordance with the following generation mechanism:

- a set of nodes $\mathcal{V}$ is associated with uniform distribution to points in the unitary space $\mathcal{S}$;
- pairs of vertices $i, j \in \mathcal{V}$, with $j \neq i$, are connected if and only if the distance between them $\delta(i, j)$ is at most a threshold r, where δ is given by some metric on $\mathcal{S}$.

Probabilistic results of this model are well known [188], allowing interesting mathematical relations between network properties and the threshold r.

Our interest is the change in the MSS and VSS when r varies. To do so we consider $n = 60$ nodes placed in the 2-dimensional unit cube $\mathcal{S} = [0, 1]^2$ and their Euclidean distances $\delta(i, j)$, for $i, j \in \mathcal{V}$. Figures 2.18, 2.19 and 2.29 show the MSS and VSS of three randomly generated geometric networks, for $r = 0.2$, $r = 0.3$ and $r = 0.4$ respectively.

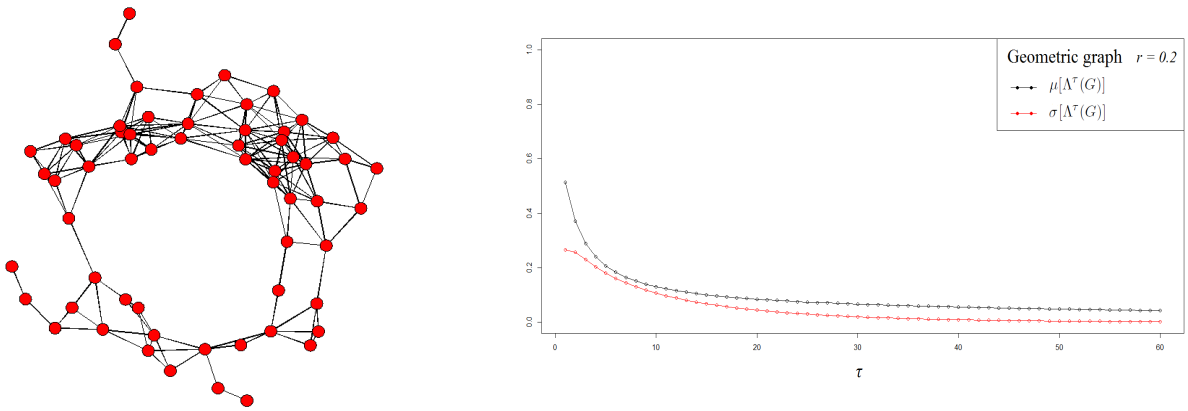


Figure 2.18: Geometric graphs with $n = 60$ and $r = 0.2$.

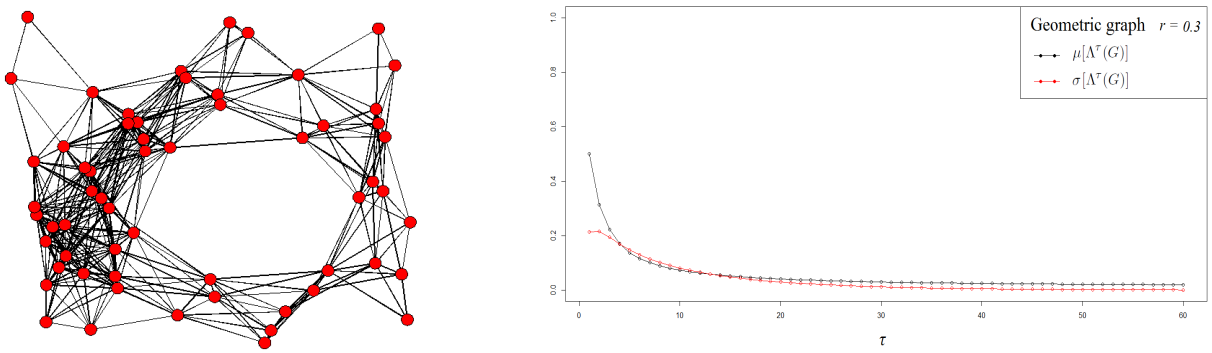


Figure 2.19: Geometric graphs with $n = 60$ and $r = 0.3$.

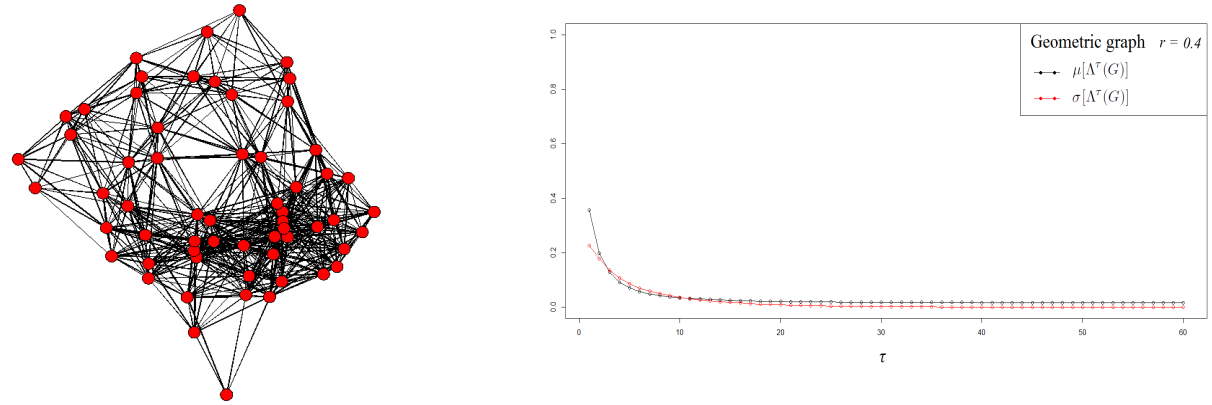


Figure 2.20: Geometric graphs with $n = 60$ and $r = 0.4$.

It might be seen a quite accurate resemblance between the MSS and VSS of the two randomly generated geometric networks in figures 2.19 and 2.19 ($r = 0.2$ and $r = 0.3$, correspondingly) and the ones of the dolphin's social network in Figure 2.12 and the Berdnard and Killworth fraternity in Figure 2.6. The density, the clustering coefficient and average path length of the three samples are reported in Table 2.4.

Dataset	Density	CC	APL
$r = 0.2$	0.10734	0.59254	4.23446
$r = 0.3$	0.22034	0.67550	3.54858
$r = 0.4$	0.34915	0.69655	1.88531

Table 2.4: Density, clustering coefficient and average path length of the three instances of geometric graphs in figures 2.18, 2.19 and 2.29.

The ability of this model to capture spatial relationships – as edges reflect closeness in space – supports the observed substantial transitivity and community structure of the generated networks, as shown in Table 2.4. This is coherent with the fact that $\mathrm{Tr}(\Lambda(\mathcal{G})) - \mathrm{Tr}(\Lambda^2(\mathcal{G})) > \mathrm{Tr}(\Lambda^2(\mathcal{G})) - \mathrm{Tr}(\Lambda^3(\mathcal{G}))$, for the three instances in figures 2.18, 2.19 and 2.29 – a positive *association between structural similarities and tie strengths* – .

The main source of differentiation between the geometrical graphs and the observed networks described in Section 2.3 is the appearance of a Poisson degree distribution under this spatial based connection mechanism, as also suggested by the low starting value of the MSS and high starting value of the VSS. Moreover, for $r = 0.4$, the generated network has high density and low variation of the degree sequence, which does not result in most network structures described in Section 2.3.

To bypass this problem, other models of network formation might be considered. A flexible class of models, which might be capable to reproduce both the community structure of the geometric graphs and the high variation of the degree sequence, is the Exponential Random Graph Models (ERGM from now on), which will be studied in Chapter 5.

For the comparative scope of this section, a process of network formation with poorly embedded nodes and high communication between subgraphs is described: the preferential attachment mechanism [12, 175]. According to this generation process (resembling open structures), we begin with an initial set of nodes associated to an arbitrary internal topology; a new node is added to this initial configuration at a time and connected to m of the existing nodes; the probability of connected with an existing node $i \in \mathcal{V}$ is proportional to $(f_i)^\gamma$ – the degree of i to the power of γ –, so that new nodes have a preference to attach themselves to *hubs*, who tend to accumulate even more ties during the process. Networks resulting from the preferential attachment mechanism exhibit power law degree distribution and low clustering coefficient, giving rise to high connections between subgraphs and poor embeddedness of nodes in communities.

The preferential attachment mechanism has recently been extended to the case of valued networks by means of different generation procedures [19–21]. Among them an interesting scale-free network generator has been proposed by Dorogovtsev and Mendes [82]. They described a process of attachment of new vertices to preferentially chosen valued edges; the values of the selected edges are sequentially updated by a fixed parameter, say δ, resulting in networks with scale-free distributions of the edge values, of node's number of contacts, and row marginals of the AM.

The graphical illustrations in figures 2.21, 2.22 and 2.23 show three networks generated by Dorogovtsev and Mendes process with $\delta = 0.5$, 3 and 10, respectively.

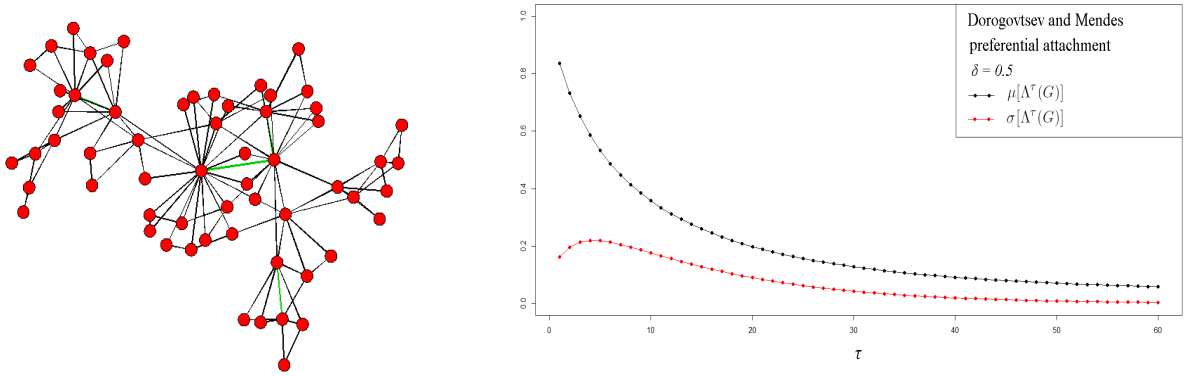


Figure 2.21: Dorogovtsev and Mendes model with $n = 60$ and $\delta = 0.5$.

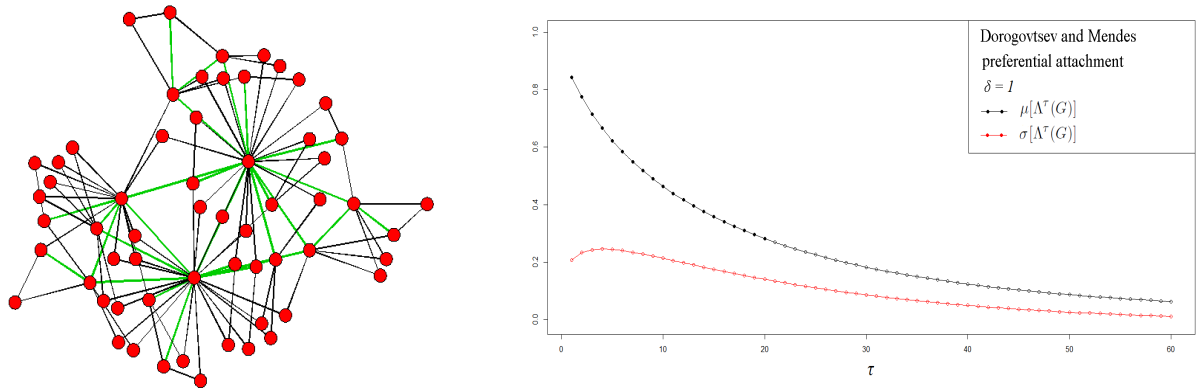


Figure 2.22: Dorogovtsev and Mendes model with $n = 60$ and $\delta = 3$.

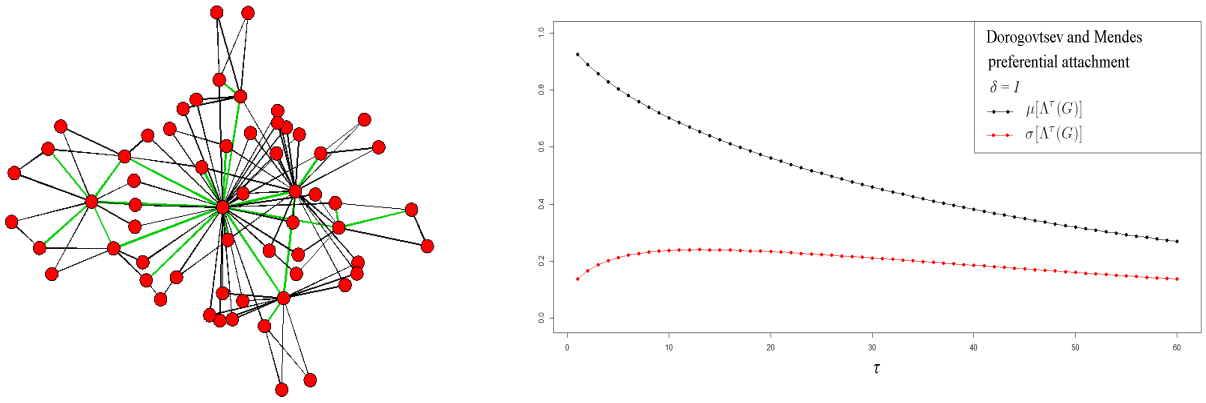


Figure 2.23: Dorogovtsev and Mendes model with $n = 60$ and $\delta = 10$.

The main source of differentiation between the networks simulated from the Dorogovtsev and Mendes model and the real-world networks in Section 2.3 is the poor level of community structure, associated to the graph conductance and to the speed of convergence of the underlying MSS. This is also suggested by the modest clustering coefficients reported in Table 2.5, in comparison with the ones of the geometric graph in Table 2.4[4]. It must also be noted that the average path length keeps low despite the small number of connections in the network, as it generally results in open structures.

[4]The computation of the clustering and average path length is carried out ignoring the edge values.

Dataset	Density	CC	APL
$r\delta = 0.5$	0.05589	0.20232	2.85367
$r\delta = 3$	0.02825	0.21428	2.83220
$r\delta = 10$	0.01624	0.20446	2.77062

Table 2.5: Density, clustering coefficient and average path length of the networks in figures 2.21, 2.22 and 2.23, which have been generated by the Dorogovtsev and Mendes process.

Random models of network formations are of great utility when trying to reproduce certain features of CNs, but they lack the explanation of both local decisions of independent agents and global decisions of an optimization-based planner. Jackson [121] highlighted that strategic models of network formation allow getting into this deeper level of understanding, by considering the optimality of the interaction structures, as introduced in Section 1.2.

From this point of view, the models we are introducing next in this section allow takeing into account the emerging properties from the point of a global planner, who wish to allocate connections among nodes in such a way as to optimize a specified criterion. Hence, the interpretation of the previous results might be comparatively integrated with the analysis of the MSS and VSS of networks resulting from such a modeling framework. Moreover, this level of analysis can bring the discussion back to the problem of the optimality of the network structures, introduced in Section 1.2 and again retroposed from a different point of view in Section 5.5.

Consider the families of all connected binary undirected networks and find the subset of them which maximize the number of closed triangles and minimize the density. The max–min mathematical programming formulation associated to the two–objective problem is:

$$\begin{aligned}
&\max \quad g \\
&\text{subject to} \quad g \leq \alpha \sum_{(i,j,k)\in\mathcal{H}^3} w_{ijk} \\
&\qquad\qquad g \leq (1-\alpha)\left(\frac{n(n-1)}{2} - \sum_{(i,j)\in\mathcal{H}^2} x_{ij}\right)
\end{aligned} \tag{2.24a}$$

$$\begin{aligned}
&1 - z_{ijk} \leq x_{ij} \leq y_{ijk} && (i,j,k)\in\mathcal{H}^3 \\
&1 - z_{ijk} \leq x_{jk} \leq y_{ijk} && (i,j,k)\in\mathcal{H}^3 \\
&1 - z_{ijk} \leq x_{ik} \leq y_{ijk} && (i,j,k)\in\mathcal{H}^3 \\
&y_{ijk} - z_{ijk} \leq w_{ijk} \leq 1 - z_{ijk} && (i,j,k)\in\mathcal{H}^3 \\
&z_{ijk} \leq 3 - (x_{ij} + x_{jk} + x_{ik}) && (i,j,k)\in\mathcal{H}^3
\end{aligned} \tag{2.24b}$$

$$\begin{aligned}
&f_{ij} + f_{ji} \leq (n-1)x_{ij} && (i,j)\in\mathcal{H}^2 \\
&\sum_{j=1}^{n} f_{1j} - \sum_{j=1}^{n} f_{j1} = n-1 \\
&\sum_{j=1}^{n} f_{kj} - \sum_{j=1}^{n} f_{jk} = -1 && k\in\mathcal{V}, k\neq 1
\end{aligned} \tag{2.24c}$$

$$\begin{aligned}
&x_{ij}\in\{0,1\}, f_{ij}\geq 0, f_{ji}\geq 0 && (i,j)\in\mathcal{H}^2 \\
&y_{ijk}, z_{ijk}, w_{ijk} && (i,j,k)\in\mathcal{H}^3
\end{aligned} \tag{2.24d}$$

where x_{ij} is the binary indicator of a tie, for $(i,j)\in\mathcal{H}^2$, associated to the $n(n-1)/2$ upper-diagonal components of the AM.

If $\widehat{t} = \sum_{i<j<k\in\mathcal{V}} x_{ij}x_{jk}x_{ij}$, then the constraints (2.24b) are used to linearize this property, by introducing auxiliary variables $y_{ijk}\in\mathcal{H}^3$ and $z_{ijk}\in\mathcal{H}^3$ to characterize the state of a triad, where $\mathcal{H}^3 = \{(i,j,k) : 1\leq i\leq n-2,\ i<j\leq n-1,\ j<k\leq n\}$. Variables w_{ijk} are the binary

indicators of transitivity in the connections among $(i,j,k) \in \mathcal{H}^3$. Constraints (2.24c) make use of an artificial flow of $n-1$ units departing from one node and arriving to each of the remaining nodes. The existence of such flow is a sufficient and necessary condition for the network to be connected. Finally, (2.24a) represents the max-min program and (2.24d) is a declaration of the domain of the decision variables.

The direct computation of the optimal solution of problem (2.24) by branch and bound algorithm (B&B from now on) is time consuming when $n \geq 50$ and a strong lower bound of the optimal solution can be used to reduce the number of explored B&B nodes.

Proposition 2. *The number of closed triangles in the optimal solution of problem (2.24) is bounded from below by*

$$h = \min\left\{(n-1),\ \alpha\frac{(n-2)(n-1)}{2}\right\} \tag{2.25}$$

and the number of edges is bounded from below by $(n-1)+h$.

Proof. Since χ is the set of all connected graphs, we must have $\sum_{(i,j)\in\mathcal{H}^2} x_{ij} \geq n-1$. Every connected graph with $n-1$ edges – a tree – is associated to $h=0$ in (2.24), as no closed triangle exist in it.

Consider a star as a feasible solution of (2.24) with $h=0$ and add h edges between pairs of nodes at distance two. Since we have $n-1$ pairs of nodes at distance two in a star, $h \leq n-1$ additional edges might be included, creating h new triangle and resulting in an increase of the value of the objective function up to $\min\left\{(1-\alpha)h,\ \alpha\left(\frac{n(n-1)}{2}-(n-1)-h\right)\right\}$.

Thus, we can keep increasing h as long as $(1-\alpha)h \leq \alpha\left(\frac{n(n-1)}{2}-(n-1)-h\right)$, up to $n-1$, so that the number of closed triangles and the number of edges in the optimal solution will be bounded from below by

$$h = \min\left\{(n-1),\ \alpha\frac{(n-2)(n-1)}{2}\right\} \text{ and } (n-1)+h, \text{ respectively.} \tag{2.26}$$

□

Stronger valid inequalities can be numerically found by heuristic methods, such as local search, tabu search, ant colony, etc[5]. We implemented a first-improve local search to find a lower bound to the optimal solution. It adds and removes one edge in each iteration up to converge to a local optimum, in which no single edge can be added or removed with an improvement of the objective function[6].

Using the heuristic solution as a lower bound on g, three instances of problem (2.24) are solved for $n=60$ and $\alpha = 0.7,\ 0.5,\ 0.3$; the MSS and VSS of the three optimal solutions are shown in figures 2.24, 2.25 and 2.26 respectively.

[5]Another interesting approach which might be used to solve (2.24) is the Lagrangian relaxation of the connectivity constraints (2.24c) and the iterative numerical solution of the resulting Lagrangian dual problem.

[6]The AMPL implementation of the model and the local search to provide a lower bound of problem (2.24) is available at www-eio.upc.edu/~nasini/Thesis_programs/AMPL_thesis_codes/AMPL_Chapter2/MaxMin_TriadsDensity/MaxMin_TriadsDensity.mod and www-eio.upc.edu/~nasini/Thesis_programs/AMPL_thesis_codes/AMPL_Chapter2/MaxMin_TriadsDensity/MaxMin_TriadsDensity.run, for the model and data generator script corresponding to figures 2.24, 2.25 and 2.26.

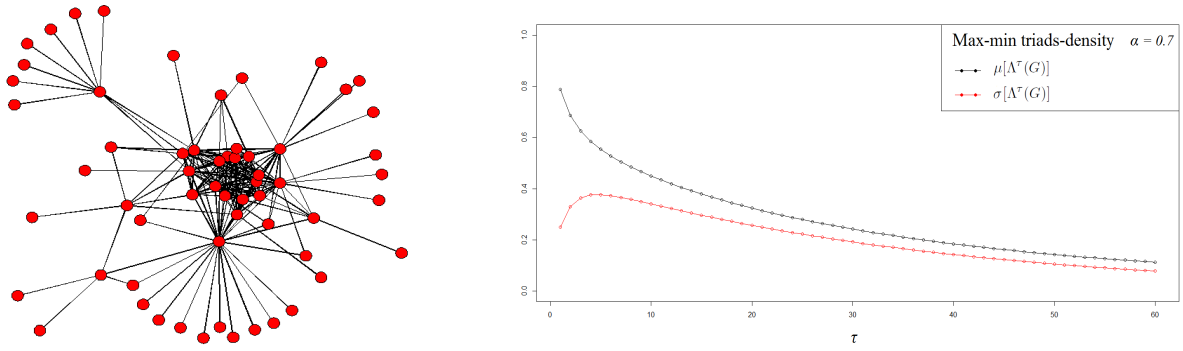


Figure 2.24: Optimal solution of problem (2.24), for $n = 60$ and $\alpha = 0.7$.

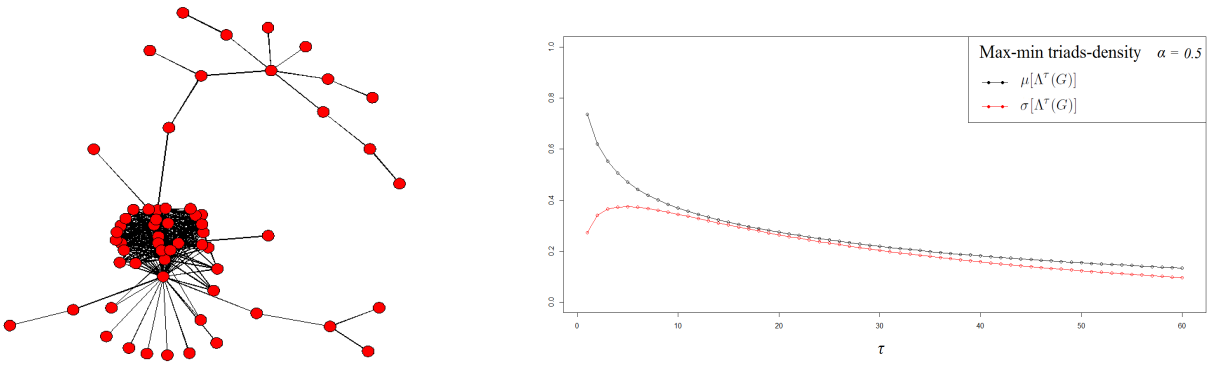


Figure 2.25: Optimal solution of problem (2.24), for $n = 60$ and $\alpha = 0.5$.

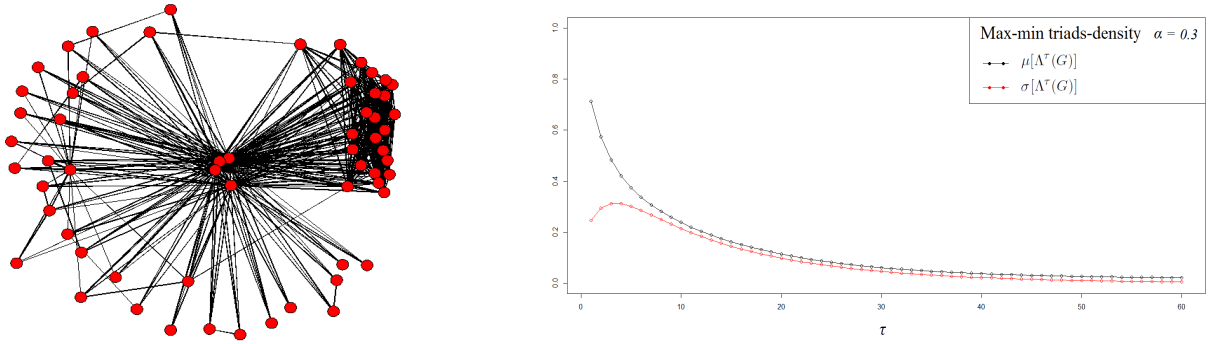


Figure 2.26: Optimal solution of problem (2.24), for $n = 60$ and $\alpha = 0.3$.

The fast convergence of the sequence $\mu[\Lambda(\mathcal{G})]$, $\mu[\Lambda^2(\mathcal{G})]$, ..., $\mu[\Lambda^\tau(\mathcal{G})]$, for the three instances in figures 2.24, 2.25 and 2.26, supports the presence of local bridges and a vanishing association between structural similarity of order τ and tie strengths. This is not surprising, as the transitivity was a characteristic that problem (2.24) sought to maximize. By contrast, despite the objective only involves the minimum between triads and sparsity, the resulting networks in figures 2.24, 2.25 and 2.26 exhibit a remarkably high connectivity (the shortest path lengths shown in Table 2.6), mainly due to the presence of highly connected hubs who are incident to many local bridges.

Dataset	Density	CC	APL
$\alpha = 0.7$	0.11921	0.61378	2.49265
$\alpha = 0.5$	0.17966	0.66311	3.17344
$\alpha = 0.3$	0.28192	0.63821	1.71808

Table 2.6: Local clustering coefficient and average path length of the geometric graph.

This results provide a substantial insight on the optimization reasons behind the appearance of small-world network structures and on the quite accurate resemblance of the associated MSS and VSS with the ones of the observed data set of Section 2.3. Beside, they seem to be particularly coherent with the ones of Mathias and Gopa [158], who showed that the small-world topology arises as a consequence of a tradeoff between maximal connectivity and minimal wiring.

> *Consider a toy model of the brain. Let us assume that it consists of local processing units, connected by wires. What constraints act on this system? On the one hand, one would want the highest connectivity (shortest path length) between the local processing units, so that information can be exchanged as fast as possible. On the other hand, it is wasteful to wire everything to everything else.* (See Mathias and Gopa [158], page 2.)

The model of Mathias and Gopa [158] tries to merge the intuition of a spatial distribution of nodes, as in the previously studied geometric random graph, with the optimization based criterion of edge formation. They pointed out that two types of distances are meaningful in networks: the minimal number of connections between any two vertices of the graph (which is the APL) and the physical distance between nodes which are spatially embedded, as in the case of a geometric graph. From this point of view, their problem was to connect spatially embedded nodes in such a way that the resulting graph has minimum APL and minimum physical distance (wire) between connected nodes.

From our point of view, the major criticism of their approach is the lack of a mathematical programming model of the problem to be solved, limiting the possibility of an exact evaluation of the obtained heuristic solutions.

Consider a reformulation of Mathias and Gopa model, where the degree of separation between nodes is measured by the cost of carrying a flow between them in an undirected network. In accordance with the spatially embedded random graph model, the set of nodes $\mathcal{V}$ is randomly associatedto points in the unitary space $\mathcal{S}$ and the distance between them δ is given by some metric on $\mathcal{S}$. The optimal network structure is defined in such a way that the total cost is minimized, along with the physical distance between connected nodes.

This mathematical programming model requires the definition of n types of flow (one per node) and $n-1$ unit of each type of flow departing from the corresponding node and reaching all the remaining nodes. Minimizing the circulating flow would result in a complete graph, where every node is one step away from all the others. By contrast, the minimization of the physical distances between connected nodes would result in spanning trees, where the flow must travel

along the network before being delivered. The goal of a global planner is to find an optimal topology, combining the two objectives. A min–max mathematical programming formulation is:

$$\begin{aligned} \min \quad & g \\ \text{subject to} \quad & g \geq \alpha \sum_{(i,j)\in\mathcal{H}^2} \delta_{ij} x_{ij} \\ & g \geq (1-\alpha) \sum_{(i,j)\in\mathcal{H}^2} \sum_{k=1}^{n} f_{ij}^k + f_{ji}^k \end{aligned} \tag{2.27a}$$

$$\begin{aligned} & \sum_{k=1}^{n} f_{ij}^k + f_{ji}^k \leq n(n-1)x_{ij} \quad (i,j) \in \mathcal{H}^2 \\ & \sum_{j=1}^{n} f_{kj}^k - \sum_{j=1}^{n} f_{jk}^k = n-1 \qquad k \in 1 \ldots n \\ & \sum_{j=1}^{n} f_{tj}^k - \sum_{j=1}^{n} f_{jt}^k = -1 \qquad t \in \mathcal{V},\ k \in 1 \ldots n,\ t \neq k \end{aligned} \tag{2.27b}$$

$$x_{ij} \in \{0,1\}, f_{ij}^k \geq 0, f_{ji}^k \geq 0 \quad (i,j) \in \mathcal{H}^2,\ k \in \mathcal{V} \tag{2.27c}$$

where x_{ij} is the binary indicator of a tie, for $(i,j) \in \mathcal{H}^2$ and f_{ij}^k is the flow of type k from i to j, for $k \in 1 \ldots n$, $(i,j) \in \mathcal{H}^2$.

The direct computation of the optimal solution of problem (2.27) by CPLEX is time consuming (more than three days of CPU time) when $n \geq 50$ and a strong lower bound of the optimal solution can be used to reduce the number of branch and bound nodes[7].

Three instances of problem (2.27) are solved for $n = 60$ and $\alpha = 0.7, 0.5, 0.3$; the MSS and VSS of the three optimal solutions are shown in figures 2.27, 2.28 and 2.29 respectively.

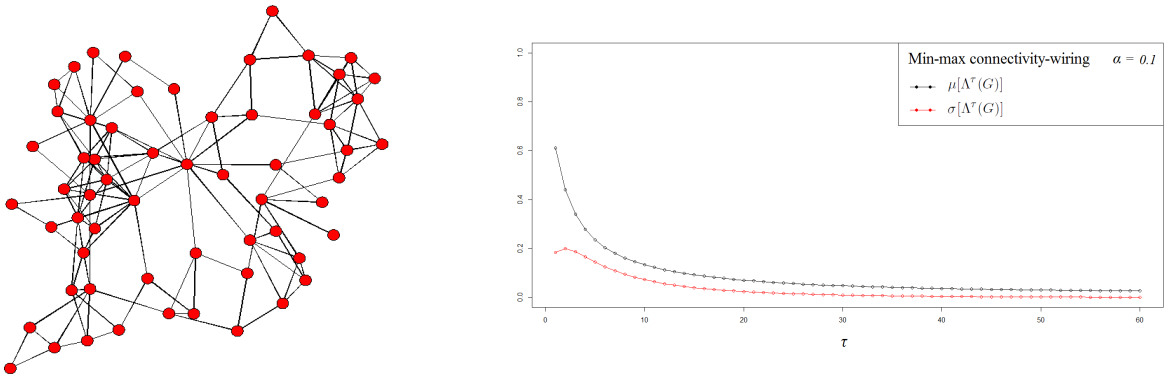


Figure 2.27: Optimal solution of problem (2.27), for $n = 60$ and $\alpha = 0.1$.

[7]This problem allows a quite straightforward application of decomposition techniques, such as the Benders decomposition, which separately and iteratively solves the problem of finding the minimum cost network structure and the one of setting the minimum circulating flow within it. We solve problem (2.27) by the branch and bound algorithm, providing a lower bound to the optimal solution by a heuristic method.

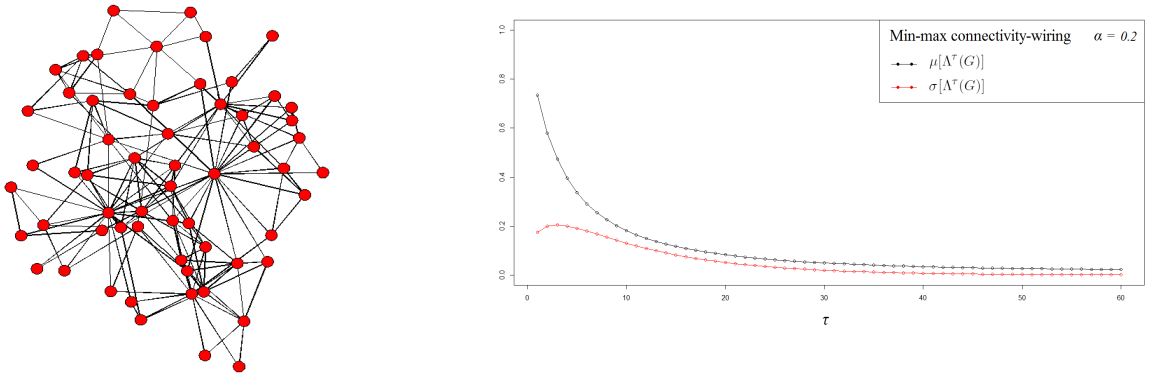


Figure 2.28: Optimal solution of problem (2.27), for $n = 60$ and $\alpha = 0.2$.

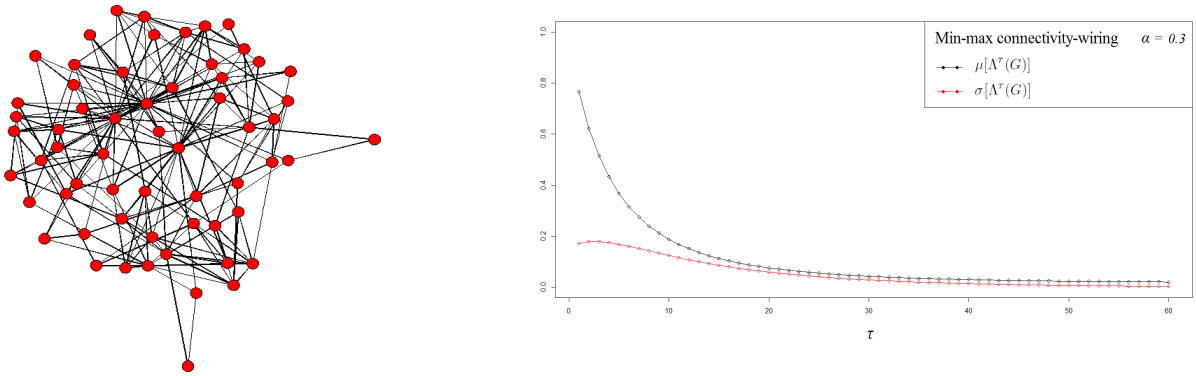


Figure 2.29: Optimal solution of problem (2.27), for $n = 60$ and $\alpha = 0.3$.

Although problem (2.27) only requires to minimize the maximum between circulating total flows and total physical distance between connected nodes, the resulting networks shown in figures 2.27, 2.28 and 2.29 exhibit a remarkably high transitivity (high clustering coefficient), as shown in Table 2.7.

Dataset	Density	CC	APL
$\alpha = 0.1$	0.07910	0.41945	3.24633
$\alpha = 0.2$	0.09944	0.27842	2.49380
$\alpha = 0.3$	0.28192	0.23901	2.16554

Table 2.7: Local clustering coefficient and average path length of the geometric graph.

This result was already observed by Mathias and Gopa [158], who argued that high transitivity emerges from the combined optimization of network distance and physical distance. The fast convergence of the sequence $\mu[\Lambda(\mathcal{G})]$, $\mu[\Lambda^2(\mathcal{G})]$, ..., $\mu[\Lambda^\tau(\mathcal{G})]$ also supports the presence of local bridges and a vanishing association between structural similarity of order τ and tie strengths.

Table 2.8 shows standard deviation of two centrality measures for the twelves networks analyzed in this section: the AM marginals and the Bonacich eigenvector centrality (with $\beta = 1$) respectively.

Dataset	edge	AM marginals (sd)	centrality (sd)
Geometric network ($r = 0.2$)	binary	0.54671	15.6037
Geometric network ($r = 0.4$)	binary	0.42653	0.40458
Geometric network ($r = 0.8$)	binary	0.35224	1.92021
Preferential attachment ($\delta = 0.2$)	valued	5.46239	1.88501
Preferential attachment ($\delta = 0.1$)	valued	8.81864	1.29460
Preferential attachment ($\delta = 1.8$)	valued	15.8561	1.92524
Max–min triads–density ($\alpha = 0.7$)	binary	1.16833	26.5424
Max–min triads–density ($\alpha = 0.5$)	binary	1.02996	1.66313
Max–min triads–density ($\alpha = 0.2$)	binary	0.85976	0.46820
Min–max connectivity–wiring ($\alpha = 0.1$)	binary	2.21449	5.01740
Min–max connectivity–wiring ($\alpha = 0.2$)	binary	3.84649	1.45383
Min–max connectivity–wiring ($\alpha = 0.3$)	binary	5.11881	2.20503

Table 2.8: Descriptive statistics of 12 simulated networks.

Comparing the values in Table 2.8 with the ones associated to the observed network data sets in Table 2.1, we realized that most of the observed network structures exhibit a higher variation at the individual level in comparison with the generated networks. This is particularly true when considering the networks corresponding to the Kapferer Tailor shop, the Dolphins' social network and network of scientific collaboration. However the valued preferential attachment and the min–max connectivity–wiring result in consistent variation of the two centrality indices, in agreement with the ones observed in Table 2.1.

As a final consideration, it must be stressed that the overall results of problems 2.24 and 2.27 provides a substantial insight on the optimization reasons behind the appearance of small-world network structures and on the quite accurate resemblance of the associated MSS and VSS with the ones of the observed data set of Section 2.2.

Part II

Mathematical Programming based approaches for random models of network formation

Abstract

The use of random simulation is quite common when statistically studying properties of highly combinatorial sets. In many of those cases, closed-form expressions are hard to be found and the availability of *efficient* and *correct* simulation procedures might be of remarkable importance. On the basis of Mathematical Programming approaches, the following three chapters are entirely devoted to the construction of efficient methods for the statistical simulation of (binary and valued) networks belonging to specified families. Chapter 3 will described a collection of Mathematical Programming models to characterize families of networks by systems of linear constraints and proposes efficient optimization-base methods to obtain random solution of the described systems. Chapter 4 will propose a specialized interior point method to deal with systems of linear constraints with primal block angular structures. Chapter 5 will derive important probabilistic properties of the described optimization-based mechanism.

Keywords: Random Graphs, Linear Programming, Interior point methods, Preconditioning, Complex Networks, Simulation and numerical modeling.

Chapter 3

Families of networks as systems of linear constrains

3.1 Purposes and preliminary overview

A particular interest has been payed in the previous chapters to the construction of a unified methodology, capable of analyzing both binary and valued networks without conceptual discontinuities. This chapter keeps carrying out this effort by showing that combinatorial objects, such as binary networks, and continuous objects, such as valued networks, might be algebraically characterized by a unified approach based on well-defined systems of linear constraints. These approach is capable to open novel methodologies for statistical simulation of random networks based on Mathematical Programming methods.

Let x_{ij} be entries of the AM of either a directed or undirected graph with no loops or multiples edges. The AM is an element of the set of binary matrices

$$\begin{aligned} & \chi = \{x_{ij} \in \{0,1\}, (i,j) \in \mathcal{H}^2\}, \\ \text{where} \quad & \mathcal{H}^2 = \{(i,j) : 1 \leq i \leq n-1, i < j \leq n\} \qquad \text{for undirected graphs} \\ \text{or} \quad & \mathcal{H}^2 = \{(i,j) : 1 \leq i \leq n, 1 \leq j \leq n, i \neq j\} \qquad \text{for directed graphs.} \end{aligned} \tag{3.1}$$

The continuous relaxation of χ, name it $CR(\chi)$, is obtained by replacing $x_{ij} \in \{0,1\}$ by $x_{ij} \in [0,1]$, in (3.1)). Clearly, all extreme points of $CR(\chi)$ are integer. If we consider a conditional graph by adding extra linear constraints to χ, then $CR(\chi)$ may contain fractional extreme points, unless its constraints matrix is totally unimodular (TU).

Theorem 1, introduced in Subsection 1.1.4, will be extensively used in this chapter[1] to show that when the constraints matrix of $CR(\chi)$ is TU, each extreme point of $CR(\chi)$ must represent a graph. Based on this fact, the generation of a bunch of graphs might be carried out by merely solving linear programs (LP) with random gradients in the objective function, or by non-degenerated simplex pivoting, starting from a given initial extreme point [182]. Beyond that, they can be generated in polynomial time if interior-point methods are used [231].

We will differentiate between two cases:

[1] If $B \in \mathbf{Z}^{m \times n}$ is a TU matrix and b is integer, then polyhedrons of the form $\varphi = \{y \in \mathbb{R}^n : By = b;\ y \geq 0\}$ have only integer extreme points, as every non singular $m \times m$ submatrix of B has integer inverse. (For more details on Unimodularity in Integer Programming see [209], [142] and [144]).

- families of networks which are bijectively associated to the set of extreme points of specified polytopes;
- families of networks which are associated to subsets of the extreme points of specified polytopes.

In the first case, the solution of any linear program in the feasible region of those polytopes gives rise to a network with the prescribed properties. In the second case, some linear programs might not result in a valid network with the prescribed properties, as fractional components are free to appear.

The next section introduces families of networks which are associated to the set of extreme points of specified polytopes and to the characterization of the convex hull of the AMs of those families by systems of linear constraints.

3.2 Families of binary networks and extreme points of polytopes

Let χ be the set of AMs of a family of either directed or undirected networks with n nodes, and let $CR(\chi)$ be its continuous relaxation. For several families of networks the extreme points of $CR(\chi)$ can be seen to be integer.

Next Proposition 3, which provides a sufficient condition for the existence of a bijection between extreme points of $CR(\chi)$ and the set of feasible networks, will be useful to show that some constraints matrices are TU.

Proposition 3. *For a given family of either directed or undirected networks with n nodes, let $F \in \mathbb{R}^{l\times m}$, be a matrix of $l \leq m$ linear constraints characterizing the family of networks under consideration, where $m = n(n-1)$ or $m = n(n-1)/2$ for, respectively, directed and undirected networks. Let $CR(\chi) = \{\mathbf{x} \in [0,1]^m : F\mathbf{x} = \mathbf{b}\}$ be the continuous relaxation of the constraints. If $\mathbf{b}$ is integer and F can be reduced by elementary row operations to a matrix, call it F', with a unique unitary element (either $+1$ or -1) per column and all the elements of the same row with the same sign, then there is a bijection between the extreme points of $CR(\chi)$ and the set of networks under consideration.*

Proof. In extended matrix form, the system of linear constraints associated to $CR(\chi)$ is

$$\left[\begin{array}{c|c} I & I \\ \hline F' & \end{array}\right] \left[\begin{array}{c} \mathbf{x} \\ \hline \mathbf{s} \end{array}\right] = \left[\begin{array}{c} \mathbf{e} \\ \hline \mathbf{b} \end{array}\right], \qquad \left[\begin{array}{c} \mathbf{x} \\ \hline \mathbf{s} \end{array}\right] \geq 0. \tag{3.2}$$

From Theorem 1, the constraints matrix of (3.2) is TU by considering the following partition of rows: set the first m rows (associated to the identities) in $\mathcal{J}_1$; if elements of row i of F' are negative, then set this row in $\mathcal{J}_1$; otherwise, if they are positive, set the row in $\mathcal{J}_2$. Therefore, all extreme points of $CR(\chi)$ are integer and they correspond to the AM of a network. In addition, no integer point may be located in the interior of $CR(\chi)$ since it is a subset of the unit hypercube, completing the proof. □

The acronyms UN, DN, ECUN and ECDN will be used to denote undirected networks, directed networks, edge-colored undirected networks, edge-colored directed networks respectively.

The following families of networks are associated to systems of linear constraints reducible to the case of Proposition 3:

- UNs conditioned to the density (i.e., number of edges);

- UNs conditioned to the within-&-between-group densities;
- UNs conditioned to the within-group densities;
- UNs conditioned to the between-group densities;
- UNs conditioned to the within-group densities and the density;
- UNs conditioned to the between group densities and the density;
- UNs conditioned to the lower bound of the nodes degree range and density;
- ECUNs conditioned to the within-color densities;
- ECUNs conditioned to the between-color density;
- ECUNs conditioned to the within-color and within-group densities;
- ECUNs conditioned to the within-color and between-group densities;
- ECUNs conditioned to the within-color and within-&-between-group densities;
- ECUNs conditioned to the within color densities and lower bound of the degree range.

Similarly, the ten families of networks associated to the directed version of the aforemention classes are also reducible to the case of Proposition 3.

In the rest of this section we prove total unimodularity of systems of linear constraints. The readers not interested in the proofs can directly jump to Section 3.3 which describes three LP procedures to randomly generate conditional networks. Acceptation-rejection procedures for the families of networks whose constraints matrices are not TU are empirically evaluated in Section 3.4, by computing the proportion of fractional solutions.

3.2.1 Basic models of networks conditioned to linear constraints

One of the simplest cases is that of networks conditioned to the density d. The following result is immediate by noting that the system of linear constraints characterizing $CR(\chi)$ verifies the hypotheses of Proposition 3:

Proposition 4. *Let $CR(\chi) = \{\mathbf{x} \in [0,1]^m : \sum_{(i,j)\in\mathcal{H}^2} x_{ij} = d\}$, where $\mathcal{H}^2$ defined in (3.1) either relates to a directed or undirected graph. Then there is a bijection between the extreme points of $CR(\chi)$ and the set of graphs with n nodes and density d.*

In some situations, nodes might be partitioned into g different groups, $\gamma_1, \ldots, \gamma_g$ and our interest might be to keep the within-group densities fixed when simulating random networks. Let Γ be the set of such groups and consider a function, $\theta : \mathcal{V} \times \mathcal{V} \to \Gamma \times \Gamma$, associating to each pair of nodes the pair of groups they belong to. The density constraint between group γ_k and γ_h, $1 \leq k \leq g$, $k \leq h \leq g$, is $\sum_{(i,j)\in\mathcal{H}^2:\theta(i,j)=(\gamma_k,\gamma_h)} x_{ij} = d_{kh}$, where d_{kh} is a non-negative integer. Note that when $k = h$ we have a within group density constraint, otherwise a between group density constraint. This system of linear constraints verifies the hypotheses of Proposition 3.

Proposition 5. *Let $CR(\chi) = \{\mathbf{x} \in [0,1]^m : \sum_{(i,j)\in\mathcal{H}^2:\theta(i,j)=(\gamma_k,\gamma_h)} x_{ij} = d_{kh}, 1 \leq k \leq g, k \leq h \leq g\}$. Then there is a bijection between the extreme points of $CR(\chi)$ and the set of graphs with n nodes and within and between group densities d_{kh}.*

A widely studied family of networks is the undirected networks with fixed degree sequence [170, 171, 173], whose associated set of AMs is $\chi = \{\mathbf{x} \in \{0,1\}^m : \sum_{j=1}^{i-1} x_{ji} + \sum_{j=i+1}^{n} x_{ij} = f_i, i = 1 \dots n\}$, where f_i is the degree of node i. Denoting these linear constraints as $F\mathbf{x} = \mathbf{f}$, we see that each column of F has two $+1$, thus it does not verify the hypothesis of Proposition 3, and the constraints matrix is not TU. However, if we only add the constraints associated to the degrees of two particular nodes—with optionally the constraint associated to the number of edges (density)—the resulting matrix is TU by Proposition 3. The information we are conditioning in this case can be seen as a lower bound of the distance between the maximum and minimum degrees. Next Propostion summarizes this result.

Proposition 6. *Let $i_1, i_2 \in \mathcal{V}$ be two nodes with degrees f_{i_1} and f_{i_2}, $\widetilde{f}_{i_1}$ and $\widetilde{f}_{i_2}$ their degrees without considering the arcs (i_1, i_2) and (i_2, i_1), and d the total number of edges in the network. Let $J(k,h) = \{j : 1 \leq j \leq n, j \neq k, j \neq h\}$ and $CR(\chi) = \{\mathbf{x} \in [0,1]^m : \sum_{(i,j)\in\mathcal{H}^2} x_{ij} = d; \sum_{j\in J(i_1,i_2)} x_{ij} = \widetilde{f}_i, i = i_1, i_2\}$. Then, there is a bijection between the extreme points of $CR(\chi)$ and the set of graphs with n nodes, d edges and degree range greater than or equal to $|f_{i_1} - f_{i_2}|$. This same result holds if the density constraint $\sum_{(i,j)\in\mathcal{H}^2} x_{ij} = d$ is removed from $CR(\chi)$.*

The researcher might sometimes be interested in studying networks whose edges are associated to a categorical value (color), generally known under the name of edge-colored networks[2].

Let $\mathcal{C}$ be a given set of colors, $|\mathcal{C}| \geq 2$. Formally, an edge-colored graph is a tuple $G_{\mathcal{C}} = (\mathcal{V}, E, \tau)$, where $\mathcal{V}$ and E are the sets of nodes and edges, respectively, and $\tau : E \longrightarrow \mathcal{C}$ a function assigning a color to each edge. They are, in some sense, related to multicommodity networks. Edge-colored graphs can be modeled as

$$\begin{aligned} &\sum_{c=1}^{|\mathcal{C}|} x_{ij}^c \leq 1 && (i,j) \in \mathcal{H}^2 \\ &x_{ij}^c, x_{ji}^c \in \{0,1\} && (i,j) \in \mathcal{H}^2, c = 1, \dots, |\mathcal{C}| \end{aligned} \tag{3.3}$$

where x_{ij}^c is 1 if an arc with color c from node i to node j exists, and 0 otherwise, and $\mathcal{H}^2$ was defined in (3.1). The first set of constraints of (3.3) —multicommodity or generalized upper bounding constraints— complicates the structure of the constraints matrix for some structural properties, such as the total number of edges, the number of edges per color, and the lower bound of the degree range.

In the case of edge-colored networks conditioned to having d_c edges per color c, the constraints $\sum_{(i,j)\in\mathcal{H}^2} x_{ij}^c = d_c, c = 1, \dots, |\mathcal{C}|$ should be included.

Proposition 7. *By adding within-color densities constrains the coefficient matrix associated to system (3.3) becomes*

$$\left[\begin{array}{cccc|cccc} I & & & & I & & & \\ & I & & & & I & & \\ & & \ddots & & & & \ddots & \\ & & & I & & & & I \\ \hline F & & & & & & & \\ & F & & & & & & \\ & & \ddots & & & & & \\ & & & F & & & & \\ \hline G & G & \dots & G & I & & & \end{array}\right] \left[\begin{array}{c} \mathbf{x}^1 \\ \mathbf{x}^2 \\ \vdots \\ \mathbf{x}^{|\mathcal{C}|} \\ \hline \mathbf{s}^1 \\ \mathbf{s}^2 \\ \vdots \\ \mathbf{s}^{|\mathcal{C}|} \end{array}\right] = \left[\begin{array}{c} \mathbf{e} \\ \mathbf{e} \\ \vdots \\ \mathbf{e} \\ \hline \mathbf{d} \\ \hline \mathbf{e} \end{array}\right]. \tag{3.4}$$

[2]The study of edge-colored networks (i.e., graphs with different types of edges) has given rise to important developments during the last few decades. From the point of view of applicability, problems arising in molecular biology are often modeled using edge-colored graphs [189], and the problem of interpersonal ties in social networks might be also modeled considering different types of arcs [109], as discussed in Chapter 1.

where $\mathbf{d} \in \mathbb{Z}^{|\mathcal{C}|}$ *is the vector of within-color densities,* $F = \mathbf{e}^T$ *and* $G = I$*. As matrix* $F' = [\mathbf{e}\ I]^T$ *verifies the conditions of Proposition 3, the coefficient matrix of (3.4) is TU.*

Note that the slack of the inequality of $\sum_{c=1}^{|\mathcal{C}|} x_{ij}^c \leq 1$ is obtained in (3.4) by defining an *auxiliary color with no specified within-color density,* c_a and an extended set of colors $\mathcal{C}^* = \mathcal{C} \bigcup \{c_a\}$.

The two models of networks described in propositions 5 and 7 can be combined to obtain another family of network which is also characterized by a TU system. Consider an edge-colored network where nodes are partitioned into g different groups: $\gamma_1, \ldots, \gamma_g$. Since the connections within members of the same group might have different colors, our interest is to keep the within-color and between-group (or within-group) densities fixed when simulating random networks. The number of c-color edges between group γ_k and γ_h, is $\sum_{(i,j)\in\mathcal{H}^2:\theta(i,j)=(\gamma_k,\gamma_h)} x_{ij}^c = d_{kh}^c$, where d_{kh}^c is a non-negative integer, for $c \in \mathcal{C}$, $1 \leq k \leq g$, $k \leq h \leq g$. Note that when $k = h$ we have a within group density constraint, otherwise a between group density constraint. This system of linear constraints verifies the hypotheses of Proposition 3.

Proposition 8. *Consider the edge-colored undirected graph conditioned to the within-color and within-group densities. Let* $CR(\chi)$ *be the subset of the* m*-dimensional unitary cube (that is* $\mathbf{x} \in [0,1]^m$*), verifying the following system*

$$\begin{aligned}
\sum_{c\in\mathcal{C}} x_{ij}^c &\leq 1 && (i,j) \in \mathcal{H}^2 \\
\sum_{\theta(i,j)=(\gamma_k,\gamma_k)} x_{ij}^c &= d_k^c && c \in \mathcal{C}, 1 \leq k \leq g \\
\sum_{(i,j)\in\mathcal{H}^2} x_{ij}^c &= d^c && c \in \mathcal{C} \\
x_{ij}^c, x_{ji}^c &\in \{0,1\} && (i,j) \in \mathcal{H}^2, c \in \mathcal{C}
\end{aligned} \tag{3.5}$$

whose corresponding extended matrix form is equal to (3.4), except for the fact that $\mathbf{d} \in \mathbb{Z}^{|\mathcal{C}|(g+1)}$ *is now the vector of within-color and within-group densities and* $F \in \{0,1\}^{g+1\times m}$ *is a matrix which can be reduced by elementary row operations to contain a unique unitary element of the same sign per column. As* $F' = [F\ I]^T$ *verifies the conditions of Proposition 3 the described system is TU.*

Another case in which (3.4) is TU is obtained when the linking constraints are associated to the within-color degree of two nodes and within-color densities, as the family of network described in Proposition 6, so that $F \in \{0,1\}^{3\times n(n-1)/2}$.

On the contrary, when the total within-color degree of all the nodes are introduced, the resulting $F \in \mathbb{R}^{n\times n(n-1)/2}$ do not verifies the properties described in Proposition 6, entailing the possible existence of fractional extreme points.

3.3 Specialized computational procedures

3.3.1 Sequential r-blocks algorithm

Let χ be one of the families of networks associated to extreme points of polytopes of the form $CR(\chi) = \{\mathbf{x} \in [0,1]^{n'} : A\mathbf{x} = \mathbf{b}\}$, $A \in \mathbb{R}^{m'\times n'}$. As we noted in Section 3.1, given a vector $\mathbf{c} \in \mathbb{R}^{n'}$, we can compute a network by solving $\min_{\mathbf{y}} \mathbf{c}^T\mathbf{y}$, s.to $\mathbf{y} \in CR(\chi)$. Similarly, if we have

a given extreme point $\mathbf{x}^k$ of $CR(\chi)$, we can obtain another extreme point $\mathbf{x}^{k+1}$ by fixing $n'-t$ variables and optimizing, with a given objective cost vector $\mathbf{c} \in \mathbb{R}^t$, the remaining t variables.

Formally, if we partition the set of variables in r blocks of dimensions t_i, $i = 1, \ldots, r$, $\sum_{i=1}^{r} t_i = n'$, and denote by $\mathbf{x}_{F_i} \in \mathbb{R}^{n'-t_i}$ and $\mathbf{x}_{C_i} \in \mathbb{R}^{t_i}$ the fixed and *changing* components of $\mathbf{x}$ associated to block i, and by A_{F_i} and A_{C_i} the submatrices of A associated to $\mathbf{x}_{F_i}$ and $\mathbf{x}_{C_i}$, the new extreme point is obtained by solving

$$\begin{array}{ll} \min & \mathbf{c}^T\mathbf{y} \\ \text{s.to} & \\ & A_{C_i}\mathbf{y} = \mathbf{b} - A_{F_i}\mathbf{x}_{F_i} \\ & 0 \leq \mathbf{y} \leq 1 \end{array} \tag{3.6}$$

for some random vector $\mathbf{c} \in \mathbb{R}^{n'-t_i}$ and $i \in \{1, \ldots, r\}$. Algorithm 4 shows how to obtain $\bar{k}$ random networks by iteratively applying this procedure. For small r values, the r-blocks algorithm generates less dependent networks at the expense of solving from scratch many linear optimization problems. In the extreme case, for $r = 1$, the cost vectors generated are non-correlated.

Algorithm 4 r-blocks

1: Let $k = 0$, $\mathbf{x}^0$ be an initial extreme point;
2: **repeat**
3: Randomly select $i \in \{1, \ldots, r\}$ and $\mathbf{c} \in \mathbb{R}^{t_i}$;
4: Let $\mathbf{x}^k_{F_i}$ and $\mathbf{x}^k_{C_i}$ be the vectors of fixed and changing components respectively;
5: Let $A^k_{F_i}$ and $A^k_{C_i}$ be the associated coefficient matrices;
6: Solve (3.6) and let $\mathbf{y}^*$ be its optimal solution;
7: $\mathbf{x}^{k+1}_{C_i} = \mathbf{y}^*$, $\mathbf{x}^{k+1}_{F_i} = \mathbf{x}^k_{F_i}$;
8: $k := k + 1$;
9: **until** $k \geq \bar{k}$

The r-blocks algorithms give rise to a Markov chain on the set of extreme points of the analyzed polytopes.

3.3.2 Sequential s-pivots algorithm

Considering again the polytope $CR(\chi) = \{\mathbf{x} \in [0,1]^{n'} : A\mathbf{x} = \mathbf{b}\}$, where $A \in \mathbb{R}^{m' \times n'}$, $m' < n'$, we know from Linear Programming that there is an equivalence between extreme points and basic solutions, which can be written as $\mathbf{x}^T = [\mathbf{x}_B^T, \mathbf{x}_N^T]$, $\mathbf{x}_B \in \mathbb{R}^{m'}$, $\mathbf{x}_N \in \mathbb{R}^{n'-m'}$, $A = [B\ N]$, $B \in \mathbb{R}^{m' \times m'}$, $N \in \mathbb{R}^{m' \times (n'-m')}$, for a suitable permutation of the variables.

Denoting by e_q the q-th column vector of the identity matrix, and by B_k and N_k the basic and nonbasic submatrices of A, given a basic solution $\mathbf{x}^k$ we can obtain another one by moving along the simplex-like direction

$$\Delta_k(q) = \begin{bmatrix} -B_k^{-1}N_k e_q \\ e_q \end{bmatrix}. \tag{3.7}$$

If the nonbasic variable q is 0, then the iteration performed is $\mathbf{x}^{k+1} = \mathbf{x}^k + \lambda\Delta_k(q)$, for some non-negative step-length λ. On the other hand if $\mathbf{x}^k_{N_q} = 1$ then we apply $\mathbf{x}^{k+1} = \mathbf{x}^k - \lambda\Delta_k(q)$. It can be easily verified that in both cases $A\mathbf{x}^{k+1} = A\mathbf{x}^k = b$, i.e., the new point satisfies the linear constraints. In addition, since the constraints matrices of Section 3.2 are TU, the step-lengths λ—computed by a ratio test—are always either 0 or 1. It is thus possible to generate a

new basic solution (i.e., a new random graph) by randomly selecting $q \in \{1, \ldots, n' - m'\}$ and computing $\mathbf{x}^{k+1} = \mathbf{x}^k \pm \lambda \Delta_k(q)$. A sample of $\bar{k}$ networks can be obtained by iteratively applying this procedure.

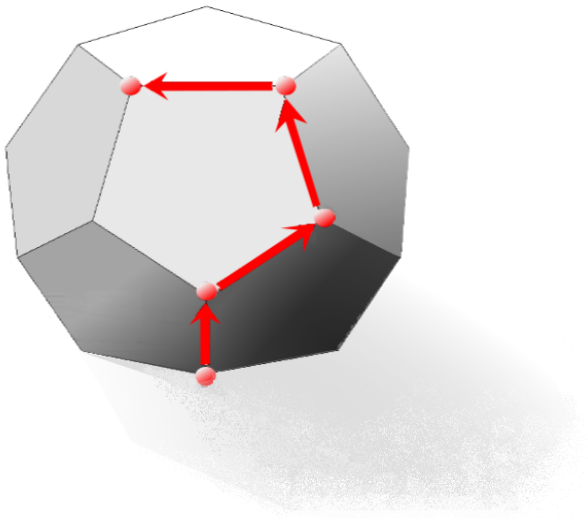


Figure 3.1: Random pivoting.

Since the resulting sample may be claimed to be quite local, every s iterations we can jump to an independent extreme point of the polytope by generating some random cost vector, and solving the associated LP. Algorithm 5 summarizes this procedure. For more technical details about the s-pivot method, see Appendix D.

Algorithm 5 s-pivots

Let $\mathbf{x}^0$ be an extreme point computed with some initial cost vector $\mathbf{c}$; $k = 0$;
repeat
  if $((k+1) \bmod s) > 0$ **then**
    repeat
      Randomly select $q \in \{1, \ldots, n' - m'\}$ (without replacement);
      Compute step-length $\lambda \in \{0, 1\}$ associated with $\pm\Delta_k(q)$;
    until $\lambda \neq 0$
    Compute $\mathbf{x}^{k+1} = \mathbf{x}^k \pm \lambda \Delta_k(q)$; update B_k and N_k; $k = k + 1$;
  else
    Generate cost vector $\mathbf{c} \in \mathbb{R}^{n'}$; compute a new extreme point $\mathbf{x}^k$;
    $k := k + 1$;
  end if
until $k \geq \bar{k}$

Pros and cons of the s-pivots algorithm. The main idea behind the s-pivots algorithm is to construct a non-homogeneous Markov chain on the state space of extreme points of $CR(\chi)$, whose transaction probabilities are defined in accordance with Algorithm 5. As long as the probabilities of selecting $q \in \{1, \ldots, n' - m'\}$ are non null, the irreducibility of the defined Markov chains can be proved, since all basic feasible solutions communicate with one another.

Three drawbacks of this procedure are: (1) many iterations may be degenerate, i.e., $\lambda = 0$, so no new point is obtained; (2) the sample of networks obtained may be highly correlated if s is large because of its proximity in the feasible polytope; (3) only extreme points might

be generated, so that the s-pivots method cannot deal with valued networks. On the other hand, this procedure may be very efficient, since it only requires simplex pivots to obtain a new network.

3.3.3 Sequential q-kernel algorithm

The method introduced in this subsection allows dealing with linear systems not characterizing the convex hull of the required AM set.

Considere a family of networks defined as the set of integer points of a polytope and its continuous relaxation $CR(\chi) = \{\mathbf{x} \in [0,u]^{n'} : A\mathbf{x} = \mathbf{b}\}$, where $A \in \mathbb{R}^{m' \times n'}$, $\mathbf{b} \in \mathbb{R}^{m'}$ $m' < n'$, $u \in \mathbb{N}$. As just mentioned, the s-pivots algorithm produces a partition of the columns of $A = [B\ N]$, where $B \in \mathbb{R}^{m' \times m'}$, $N \in \mathbb{R}^{m' \times (n'-m')}$, for a suitable permutation of the variables. Here we propose a different bipartition of variables, making easier to deal with integer valued networks (i.e. $u > 1$), when the simplex-like direction forbids to explore integer points in the interior of the polytope.

Let $\mathbf{x}^T = [\mathbf{x}_{\mathcal{Q}}, \mathbf{x}_{\mathcal{Q}^c}]$, $\mathbf{x}_{\mathcal{Q}} \in \mathbb{R}^q$, $\mathbf{x}_{\mathcal{Q}^c} \in \mathbb{R}^{n'-q}$ and $A = [A_{\mathcal{Q}}\ A_{\mathcal{Q}^c}]$, for a suitable permutation of the variables. Given a network $\mathbf{x}^k \in \chi$, another network $\mathbf{x}^{k+1} \in \chi$ can be obtained by selecting q decision variables, indexed by the set $\mathcal{Q} \subset \{1,\dots,n'\}$ and finding a point in $\varphi_{\mathcal{Q}} = \{\mathbf{y} \in \{0,u\}^q : A_{\mathcal{Q}}\mathbf{y} = \mathbf{b} - A_{\mathcal{Q}^c}\mathbf{x}^k_{\mathcal{Q}^c}\}$.

Denoting by $B_{\mathcal{Q}} \in \mathbb{R}^{r \times l}$ the bases for the null space of $A_{\mathcal{Q}}$, the sequence $\mathbf{x}^{k+1}_{\mathcal{Q}} = \mathbf{x}^k_{\mathcal{Q}} + B_{\mathcal{Q}}\boldsymbol{\lambda}$ is a collection of networks belonging to the specified family, where $\boldsymbol{\lambda} \in \Lambda(\mathbf{x}^k_{\mathcal{Q}}, B_{\mathcal{Q}}) = \{\boldsymbol{\theta} \in \mathbb{R}^\ell \mid -\mathbf{x}^k_{\mathcal{Q}} \le B_{\mathcal{Q}}\boldsymbol{\theta},\ u - \mathbf{x}^k_{\mathcal{Q}} \ge B_{\mathcal{Q}}\boldsymbol{\theta},\ \boldsymbol{\theta} \neq 0\}$ and ℓ is the dimension of the null space of $A_{\mathcal{Q}}$.

When q is small the resulting sample may be claimed to be quite local, so that a possible jump to an independent solution might be considered by generating some random cost vector, and solving the associated LP, as discussed in Subsection 3.3.2 and illustrated in Algorithm 6.

Algorithm 6 q-kernel.

Let $\mathbf{x}^0$ be an extreme point computed with some initial cost vector $\mathbf{c}$; $k = 0$;
repeat
  if $((k+1) \bmod s) > 0$ **then**
    Randomly select $\mathcal{Q} \subset \{1,\dots,n'\}$;
    if $\Lambda \neq \varnothing$ **then**
      Randomly select a $\boldsymbol{\lambda} \in \Lambda(\mathbf{x}^k_{\mathcal{Q}}, B_{\mathcal{Q}})$:
      if $\Lambda \neq 0$ **then**
        Compute $\mathbf{x}^{k+1}_{\mathcal{Q}} = \mathbf{x}^k_{\mathcal{Q}} + e_{\mathcal{Q}}B_{\mathcal{Q}}\boldsymbol{\lambda}$;
        $k := k+1$;
      end if
    end if
  else
    Generate cost vector $\mathbf{c} \in \mathbb{R}^{n'}$ and compute a new extreme point $\mathbf{x}^k$; $k = k+1$;
  end if
until $k \ge \bar{k}$

The efficiency of this method relies on the availability of a closed–form expression for the basis of the null space of $A_{\mathcal{Q}}$.

Specific cases where the selection of $\mathcal{Q}$ and the associated $B_{\mathcal{Q}}$ might be efficiently performed are:

- UNs conditioned to the density d;
- UNs conditioned to the degree sequence $f_1,\dots,f_n$;

- ECUNs conditioned to the density d;
- ECUNs conditioned to have d_c edges per color $c \in \mathcal{C}$;
- ECUNs conditioned to have degree sequence $f_1^c, \ldots, f_n^c$, for each color $c \in \mathcal{C}$.

Note that the values of the edges, i.e. whether binary, integer or fractional, only affect the definition of $\Lambda(\mathbf{x}_{\mathcal{Q}}^k, B_{\mathcal{Q}})$. In the rest of this subsection we are assuming integer valued edges, though this might be straightforwardly generalized to real valued edges. The number of decision variables will be denoted by m, which is equal to either $n(n-1)$ or $m = n(n-1)/2$ for, respectively, directed and undirected networks.

UNs conditioned to densities. Let $\chi = \{\mathbf{x} \in \{0, u\}^m : \sum_{(i,j)\in\mathcal{H}^2} x_{ij} = d\}$. For every $\mathbf{x} \in \chi$, consider three randomly selected nodes $i, j, k \in \mathcal{V}$ and let $\mathbf{x}_{\mathcal{Q}}^T = [x_{ij},\, x_{ik},\, x_{jk}]$. We have

$$\mathbf{x}_{\mathcal{Q}} + B_{\mathcal{Q}}\boldsymbol{\lambda} = \begin{bmatrix} x_{ij} \\ x_{ik} \\ x_{jk} \end{bmatrix} + \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{\lambda}, \tag{3.8}$$

where $\boldsymbol{\lambda} \in \mathbb{Z}^2$ must verify

$$\begin{array}{lcl} x_{ij} - u \leq & \lambda_1 + \lambda_2 & \leq x_{ij} \\ -x_{ik} \leq & \lambda_1 & \leq u - x_{ik} \\ -x_{jk} \leq & \lambda_2 & \leq u - x_{jk} \end{array} \tag{3.9}$$

When $u = 1$ this approach gives rise to a combinatorial procedure, which allows to assign values in $\{-1, 0, 1\}^2$ to $\boldsymbol{\lambda}$ for each arrangement of the selected x_{ij}, x_{ik}, and x_{jk}, as shown in the following table:

x_{ij}	x_{ik}	x_{jk}	(λ_1, λ_2)
0	0	0	$(0,0)$
0	0	1	$(0,0)$ $(0,-1)$ $(1,-1)$
0	1	0	$(0,0)$ $(-1,0)$ $(-1,1)$
0	1	1	$(0,0)$ $(-1,0)$ $(0,-1)$ $(-1,1)$
1	0	0	$(0,0)$ $(1,0)$ $(0,1)$
1	0	1	$(0,0)$ $(0,-1)$ $(1,-1)$ $(0,1)$
1	1	0	$(0,0)$ $(0,1)$ $(-1,1)$ $(0,-1)$
1	1	1	$(0,0)$

When $u > 1$ an enumeration of the possible values of λ is also possible, though much more tedious.

UNs conditioned to the degree sequence. Let $\chi = \{\mathbf{x} \in \{0, u\}^m : \sum_{j=1}^{i-1} x_{ji} + \sum_{j=i+1}^{n} x_{ij} = f_i, i = 1 \ldots n\}$, where f_i is the degree of node i. For every $\mathbf{x} \in \chi$, consider four randomly selected nodes $i, j, k, h \in \mathcal{V}$ and let $\mathbf{x}_{\mathcal{Q}}^T = [x_{ij},\, x_{ik},\, x_{ih},\, x_{jk},\, x_{jh},\, x_{kh}]$. We have

$$\mathbf{x}_{\mathcal{Q}} + B_{\mathcal{Q}}\boldsymbol{\lambda} = \begin{bmatrix} x_{ij} \\ x_{ik} \\ x_{ih} \\ x_{jk} \\ x_{jh} \\ x_{kh} \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ -1 & -1 \\ -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{\lambda}, \tag{3.10}$$

where $\boldsymbol{\lambda} \in \mathbb{Z}^2$ must verify

$$\begin{array}{lcl} \max\{x_{ih},\, x_{jk}\} - u \leq & \lambda_1 + \lambda_2 & \leq \min\{x_{ih},\, x_{jk}\} \\ -\min\{x_{ik},\, x_{jh}\} \leq & \lambda_1 & \leq u - \max\{x_{ik},\, x_{jh}\} \\ -\min\{x_{ij},\, x_{kh}\} \leq & \lambda_2 & \leq u - \max\{x_{ij},\, x_{kh}\}. \end{array} \tag{3.11}$$

ECUNs conditioned to the total density. Let $\chi = \{\mathbf{x} \in \{0,u\}^{m|\mathcal{C}|} : \sum_{c=1}^{|\mathcal{C}|} x_{ij}^c \geq 1, (i,j) \in \mathcal{H}^2\}$, where $\mathcal{C}$ is the set of colors and $\mathcal{C}^* = \mathcal{C} \bigcup \{c_a\}$ the extended set of colors. For every $\mathbf{x} \in \chi$, a set of $\overline{n}$ nodes are randomly selected, along with two colors, $c_1, c_2 \in \mathcal{C}$. Let $\mathbf{x}_\mathcal{Q} \in \{0,u\}^{\overline{n}(\overline{n}-1)}$ be a vector with components $x_{12}^{c_1}$, ..., $x_{(\overline{n}-1)\overline{n}}^{c_1}$ $x_{12}^{c_2}$, ..., $x_{(\overline{n}-1)\overline{n}}^{c_2}$, associated to the selected nodes and colors. We have

$$\mathbf{x}_\mathcal{Q} + B_\mathcal{Q}\boldsymbol{\lambda} = \mathbf{x}_\mathcal{Q} + \begin{bmatrix} I \\ -I \end{bmatrix} \boldsymbol{\lambda}, \tag{3.12}$$

where $\boldsymbol{\lambda} \in \mathbb{Z}^{\overline{n}(\overline{n}-1)/2}$ must verify

$$\max\{-x_{ij}^{c_1},\ x_{ij}^{c_2} - u\} \leq \quad \lambda_{ij} \quad \leq \min\{-x_{ij}^{c_2},\ u - x_{ij}^{c_1}\} \qquad i,j = 1,\ldots,\overline{n}, i < j \tag{3.13}$$

ECUNs conditioned to the within-color densities. Let $\chi = \{\mathbf{x} \in \{0,u\}^{m|\mathcal{C}|} : \sum_{(i,j)\in\mathcal{H}^2} x_{ij}^c = d_c, c = 1,\ldots,|\mathcal{C}|+1, \sum_{c=1}^{|\mathcal{C}|+1} x_{ij}^c = 1, (i,j) \in \mathcal{H}^2\}$, where $\mathcal{C}$ is the set of colors and $\mathcal{C}^* = \mathcal{C} \bigcup \{c_a\}$ the extended set of colors. For every $\mathbf{x} \in \chi$, consider three randomly selected nodes $i,j,k \in \mathcal{V}$, two randomly selected colors $c_1, c_2 \in \mathcal{C}$ and let $\mathbf{x}_\mathcal{Q}^T = [x_{ij}^{c_1}, x_{ik}^{c_1}, x_{jk}^{c_1}, x_{ij}^{c_2}, x_{ik}^{c_2}, x_{jk}^{c_2}]$. We have

$$\mathbf{x}_\mathcal{Q} + B_\mathcal{Q}\boldsymbol{\lambda} = \begin{bmatrix} x_{ij}^{c_1} \\ x_{ik}^{c_1} \\ x_{jk}^{c_1} \\ x_{ij}^{c_2} \\ x_{ik}^{c_2} \\ x_{jk}^{c_2} \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \\ -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{\lambda}, \tag{3.14}$$

where $\boldsymbol{\lambda} \in \mathbb{Z}^2$ must verify

$$\begin{array}{lcl} \max\{-x_{ij}^{c_2},\ x_{ij}^{c_2} - u\} \leq & \lambda_1 + \lambda_2 & \leq \min\{u - x_{ij}^{c_1},\ x_{ij}^{c_2}\} \\ \max\{x_{ik}^{c_1} - u,\ -x_{ik}^{c_2}\} \leq & \lambda_1 & \leq \min\{x_{ik}^{c_1},\ u - x_{ik}^{c_2}\} \\ \max\{x_{jk}^{c_1} - u,\ -x_{jk}^{c_2}\} \leq & \lambda_2 & \leq \min\{x_{jk}^{c_1},\ u - x_{jk}^{c_2}\} \end{array} \tag{3.15}$$

ECUNs conditioned to the within-color degree sequence. Let $\chi = \{\mathbf{x} \in \{0,u\}^{m|\mathcal{C}|} : \sum_{j=1}^{i-1} x_{ji}^c + \sum_{j=i+1}^{n} x_{ij}^c = f_i^c, i = 1\ldots n,\ c = 1\ldots\mathcal{C}, \sum_{c=1}^{|\mathcal{C}|} x_{ij}^c \geq 1, (i,j) \in \mathcal{H}^2\}$. For every $\mathbf{x} \in \chi$, four nodes $i,j,k,h \in \mathcal{V}$ are randomly selected, along with two colors $c_1, c_2 \in \mathcal{C}$. Let $\mathbf{x}_\mathcal{Q}^T = [x_{ij}^{c_1}, x_{ik}^{c_1}, x_{ih}^{c_1}, x_{jk}^{c_1}, x_{jh}^{c_1}, x_{kh}^{c_1}, x_{ij}^{c_2}, x_{ik}^{c_2}, x_{ih}^{c_2}, x_{jk}^{c_2}, x_{jh}^{c_2}, x_{kh}^{c_2}]$. We have

$$\mathbf{x}_\mathcal{Q} + B_\mathcal{Q}\boldsymbol{\lambda} = \begin{bmatrix} x_{ij}^{c_1} \\ x_{ik}^{c_1} \\ x_{ih}^{c_1} \\ x_{jk}^{c_1} \\ x_{jh}^{c_1} \\ x_{kh}^{c_1} \\ x_{ij}^{c_2} \\ x_{ik}^{c_2} \\ x_{ih}^{c_2} \\ x_{jk}^{c_2} \\ x_{jh}^{c_2} \\ x_{kh}^{c_2} \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ -1 & -1 \\ -1 & -1 \\ 1 & 0 \\ 0 & 1 \\ 0 & -1 \\ -1 & 0 \\ 1 & 1 \\ 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \boldsymbol{\lambda}, \tag{3.16}$$

where $\boldsymbol{\lambda} \in \mathbb{Z}^2$ must verify

$$\begin{array}{rcl} \max\{x_{ih}^{c_1}-u, x_{jk}^{c_1}-u, -x_{ih}^{c_2}, -x_{jk}^{c_1}\} \leq & \lambda_1+\lambda_2 & \leq \min\{x_{ih}^{c_1}, x_{jk}^{c_1}, u-x_{ih}^{c_2}, u-x_{jk}^{c_2}\} \\ \max\{-x_{ik}^{c_1}, -x_{jh}^{c_1}, x_{ik}^{c_2}-u, x_{jh}^{c_2}-u\} \leq & \lambda_1 & \leq \min\{u-x_{ik}^{c_1}, u-x_{jh}^{c_1}, x_{ik}^{c_2}, x_{jh}^{c_2}\} \\ \max\{-x_{ij}^{c_1}, -x_{kh}^{c_1}, x_{ij}^{c_2}-u, x_{kh}^{c_2}-u\} \leq & \lambda_2 & \leq \min\{u-x_{ij}^{c_1}, u-x_{kh}^{c_1}, x_{ij}^{c_2}, x_{kh}^{c_2}\} \end{array} \tag{3.17}$$

Pros and cons of the q-kernel algorithm. To summarize, we have a Markov chain on the state space of elements of χ. Given a current state $\mathbf{x}^k \in \chi$, we jump to next state, say $\mathbf{x}^{k+1}$, by randomly selecting $\mathcal{Q} \subset \{1,\ldots,m\}$ and $\boldsymbol{\lambda} \in \Lambda(\mathbf{x}^k_{\mathcal{Q}}, B_{\mathcal{Q}})$, when $(k+1) \bmod s > 0$. The main advantage of this method is that the selected solution are not supposed to be extreme points, so that the case where $u > 1$ might also be considered.

We didn't try to find sufficient conditions for the irreducibility of the defined Markov chains in the general case. However, in the specific cases of the classes of networks conditioned to the density it can be easily proved that all states communicate with one another, as long as the probabilities of selecting $\mathcal{Q} \subset \{1,\ldots,m\}$ and $\boldsymbol{\lambda} \in \Lambda(\mathbf{x}^k_{\mathcal{Q}}, B_{\mathcal{Q}})$ are not null.

The main drawback of this procedure is that the generated sample might be highly correlated if q is small and s is large. On the other hand, this procedure may be very efficient. Note that the s-pivots, q-kernel and r-blocks methods are equivalent for $s = r = 1$. They will be computationally evaluated in the next two chapters.

3.4 Efficiency of the sampling procedures

The r-blocks, s-pivots and q-kernel procedures require the repeated solution of several LP problems, which may be computationally expensive for large networks.

In this section we study different aspects related to the solution of the resulting LPs and introduces methodological procedures which will be mathematically taken into account in the next Chapter.

The first fact which must be noted is that the computational performance of the LP solvers might be substantially affected by the degeneracy of the basic feasible solutions. This is particularly true for the s-pivots, whose fundamental operations are entirely based on sequences of change of bases. A basic feasible solution is said to be degenerate if there exists in it one or more basic variables with limit values (in our case, either 0 or 1). As previously mentioned, the integrality of the extreme points of the described polytopes entails that all basic variables are always at their limits. Degeneracy implies a null step-length λ in the movement $\mathbf{x}^{k+1} = \mathbf{x}^k \pm \lambda\Delta_k(q)$, for some nonbasic variable q.

Example 1 provides an illustrative case when degeneracy might affect the s-pivots method.

Example 1 (Four-node edge-colored networks)**.** *Consider an undirected binary four-node edge-colored networks with $|\mathcal{C}| = 2$ colors, conditioned to have d_c edges per color c, specifically $d = [3,2]^T$, and the associated system of linear constraints*

$$\begin{bmatrix} \mathbf{e}^T & & \\ & \mathbf{e}^T & \\ I & I & I \end{bmatrix} \mathbf{x} = \begin{bmatrix} 3 \\ 2 \\ \mathbf{e}_6 \end{bmatrix}, \qquad 0 \leq \mathbf{x} \leq 1 \tag{3.18}$$

where $\mathbf{e}_6$ is a vector of six ones. Let $\mathbf{x} = [1,0,1,0,0,1,0,1,0,1,0,0,0,0,0,0,1,0]$ be a basic feasible solution, corresponding to an edge-colored AM in vector notation. Applying a QR-decomposition, the coefficient matrix $A\mathbb{R}^{m'\times n'}$ of the system (3.18) is factorized into a product of an orthogonal matrix $Q \in \mathbb{R}^{8\times 8}$ and an upper triangular matrix $R \in \mathbb{R}^{8\times 12}$. Since $\operatorname{rank}(A) = m' - 1 = 7$, the first 7 columns of Q form an orthogonal basis for the column space of A. The indexes of the basic matrix are: 6, 8, 18, 1, 2, 3, 4, 5. Let us randomly select non basic index: 7, 8, 9, 10, 11, 13, 14, 15, 16, 17. The following table shows the values of the step-length λ resulting from the minimum ratio test for each non basic entering variable:

non basic index	λ
7	1
8	1
9	1
10	0
11	0
13	0
14	0
15	0
16	0
17	0

Since the non basic indexes $\{1, \ldots, n' - m'\}$ *are randomly selected without replacement, the maximum number of minimum ratio tests to be performed before jumping to a new extreme point is* 7*. Geometrically, the intuition behind this degenerate change of bases is that any movement in the direction* $\Delta_k(q)$ *breaks the feasibility, as it point outward the polytope. (For more details about the simplex pivoting for bounded variables see Appendix D).*

Consider the data set of 62 dolphins introduced in the Section 2.3 and the following families of networks:

- UNs conditioned to the density;
- UNs conditioned to the within and between group densities;
- UNs conditioned to the lower bound of the degree sequence range and density.

These three models are specified by the observed parameters of the dolphin's social network. The within group densities are obtained from the community structure of the observed network, computed by the *walk trap* community search algorithm of [174], as shown in Figure 3.2 with different colors[3].

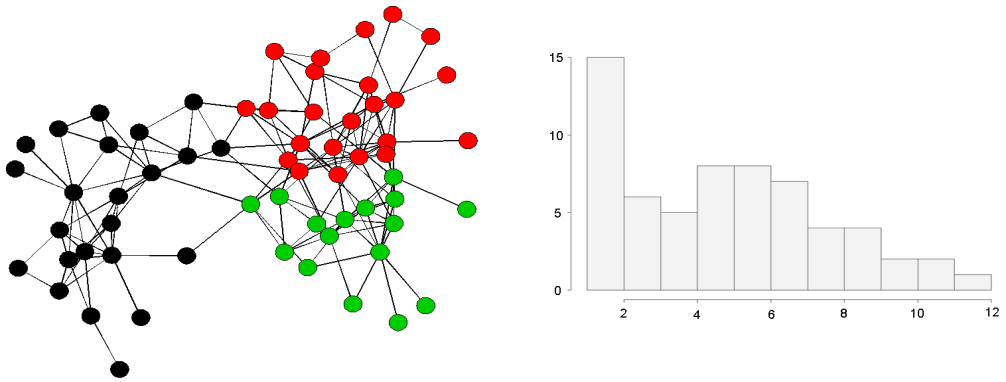


Figure 3.2: The network on the left part of the figure represents the structure of connections between the 62 dolphins studied by Lusseau. On the right the histogram of the nodes' degrees is shown. The `R` code needed to obtain the communities is is available at `http://www-eio.upc.edu/~nasini/Thesis_Programs/R_chaper_3.txt`

[3]Lusseau and Newman [153] showed that dolphins can be easily grouped into communities such that each of them are internally better connected. Many popular methods to find communities in networks are implemented in the `igraph` library for the `R` package. We applied to the dolphin network dataset two of the several search algorithm available in the `R` package: i) the *walk trap* community algorithm and ii) the fast greedy community algorithm. Both search algorithms find 4 communities and the resulting classifications only differ with respect to four nodes, which are indeed associated to communities of size two. To avoid taking into account groups with less then three nodes, we merged the two respective two-node-communities in the corresponding bigger groups.

On the right part of Figure 3.2, the histogram of nodes' degrees is shown. Table 3.1 reports the number of basic solutions explored by the s-pivots procedure to generate 100 different networks, and the required CPU time.

Graph conditioned to ...	bases	CPU time
total density	1166	1.7
within group densities	1388	0.7
lower bound of the degree range and density	1097	0.7
within group densities and total density	1184	0.8

Table 3.1: Number of bases visited to compute 100 extreme points with the s-pivots method for the three specified polytopes

The results confirm the high degeneracy of the basic feasible solutions. Therefore, although state-of-the-art implementations of the simplex method can be used, it is worth to exploit the problem structure whenever possible [70], by LP methods which are more robust against degeneracy [231].

It can be shown that, under a proper row and column permutation, most of the constraints matrices of Section 3.2 exhibit a primal block-angular structure such as

$$\begin{bmatrix} N_1 & & & & \\ & N_2 & & & \\ & & \ddots & & \\ & & & N_k & \\ L_1 & L_2 & \dots & L_k & I \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^k \\ x^0 \end{bmatrix} = \begin{bmatrix} b^1 \\ b^2 \\ \vdots \\ b^k \\ b^0 \end{bmatrix}. \tag{3.19}$$

Matrices N_i and L_i, $i = 1, \dots, k$, respectively define the block-diagonal and linking constraints, k being the number of blocks. Vectors x^i, $i = 1, \dots, k$, are the variables for each block. x^0 are the slacks of the linking constraints $\sum_{i=1}^{k} L_i x^i \leq b^0$ ($x^0 = 0$ if linking constraints are equalities). $b^i, i = 1, \dots, k$, is the right-hand side vector for each block of constraints, whereas b^0 is for the linking constraints.

A specialized interior-point method to deal with this particular matrix structure will be studied in the next chapter.

3.5 Dealing with fractional extreme points

Sometimes we are able to model a family of networks by a system of linear constraints, but such a system does not characterize the convex hull of the required AM set. In this case, the bijective relation between the associated set of the extreme points and the specified family of networks cannot be established.

This section describes some mathematical programming models of networks which does not characterize the convex hull of the required AM set and introduces a rejection method whose probabilistic properties will be further discusses in the next chapter.

Some of the most relevant families of networks which might be modeled by these kind of systems are shown in the following list:

- UNs conditioned to connectivity;
- UNs conditioned to the degree sequence;

- UNs conditioned to the number of triads;
- DNs conditioned to the dyads count;
- ECUNs conditioned to the between-color degree sequence;
- ECUNs conditioned to the within-color color degree sequences;
- ECUNs conditioned to the within-color color number of triads;
- ECUNs conditioned to the between-color number of triads;
- ECUNs conditioned to the within-color dyads count;
- ECUNs conditioned to the between-color dyads count;

The intersections of these classes of networks are also cases of fractional-polytopes (polytopes whose set of extreme points is not known to be integer), such as the UNs conditioned to the within-group densities and the degree sequence.

For this specific example, consider again the network of 62 dolphins introduced in Section 2.1, after fixing the internal densities of the three communities and the degree sequence in accordance with the detected community structure of the observed network, the latter class of network can be specified as $\chi = \{(x_{12}, \ldots, x_{n-1n}) \in \{0,1\}^m : \sum_{j=1}^{i-1} x_{ji} + \sum_{j=i+1}^{n} x_{ij} = f_i, i = 1 \ldots n; \sum_{i,j\in\gamma_1} x_{ij} = 48; \sum_{i,j\in\gamma_2} x_{ij} = 53; \sum_{i,j\in\gamma_3} x_{ij} = 32;\}$, where $f_1, \ldots, f_n$ is the degree sequence.

As long as no integer solution is supposed to exist in the interior of the defined polytope, the discussed injective relation between families of networks and extreme points of polytopes (i.e., any random network is associated to an extreme point, but not the opposite) still exists and a rejection approach might still be valid to avoid fractional solutions[4]. (As it will be discussed in the next chapter, the rejection of fractional solutions does not affect the probabilistic properties of the mathematical programming based generation procedures.)

A bunch of 100 networks is generated from the family χ of UNs conditioned to the within-group densities and the degree sequence. The r-blocks method is applied to $CR(\chi)$, starting from the observed dolphin social network.

r	Fractional networks	Loops	CPU time
4	44	0	112.0
5	38	1	103.2
6	26	5	89.1
8	10	16	61.4
10	8	110	68.7
13	3	372	74.6

Table 3.2: Number of fractional solutions and loops of the sequential r-blocks optimization method for different values of r, starting from the observed dolphin's social network.

Table 3.2 provides the numerical results for $r \in \{4, 5, 6, 8, 10, 13\}$. Column "Fractional networks" shows the number of rejected fractional solutions. If two consecutive LPs with different objectives provided the same solution (named "loops"), the repeated network is also rejected;

[4]If the edge values are required to be binary, the absence of integer solution in the interior of the described polytope can still be ensured, so that the problem of generating random networks still coincides with the one of generating random extreme points.

they are reported in column "Loops" of Table 3.2. Last column provide the CPU time in seconds, using the same computational environment than in previous sections. In accordance with these results, the number of fractional networks seems to increase with the number of optimized blocks r. Nonetheless, when r is small, consecutive networks are more likely to be the same, so that we face a trade-off between minimizing the number of loops and minimizing the number of fractional solutions. Moreover, the simple verification that two consecutive networks are not identical is time consuming in the overall generation process and avoiding the check out by increasing the number of optimizing blocks might often be worth.

Another system of linear constraints associated to fractional-polytope with no integer solution in the interior is the undirected edge-colored network with fixed within-colored densities and between-color degree sequence, when edges are allowed to have multiple colors. The general matrix structure of this problem is the same as in (3.4), with $F = \mathbf{e}^T$ and

$$G = \begin{bmatrix} \mathbf{e}_{n-1}^T & & & \\ I_{n-1} & \mathbf{e}_{n-2}^T & & \\ & I_{n-2} & \mathbf{e}_{n-3}^T & \\ & & I_{n-3} & \begin{matrix} \cdots & 1 \\ \ddots & 1 \end{matrix} \end{bmatrix}, \tag{3.20}$$

where $\mathbf{e}_q = [1 \dots 1]^T \in \mathbb{R}^q$ and I_q is the q-dimensional identity matrix. In this case $F' = [\mathbf{e}\ G^T]^T$ dose not verify the conditions of Proposition 3 and the coefficient matrix of (3.4) is not TU.

Let us consider the network dataset concerning the business and marriage relations among Florentine Renascence families and an edge-colored graph model introduced in Subsection 3.2.1, where the set of colors is $\mathcal{C} = \{\text{business}, \text{mariage}, \text{both}\}$, denoting the type of relation, i.e. either business or marriage or both; the number of each type relations is taken as a fixed quantity and the degree of each node with respect to both type of relations is also taken as a fixed quantity. The associated system of linear constraints in matrix form is (3.4), with matrix G defined as in (3.20).

Applying the r-block method, we obtain a sequence of 100 networks by solving LPs with matrix form (3.4) and (3.20), starting from the network data set associated to the business and marriage relations among Florentine Renascence families. Both the number of fractional solutions and the number of loops are observed in Table 3.3.

r	Fractional networks	Loops	CPU time (sec)
2	58	0	25.6
2	51	0	26.7
2	66	0	22.3
2	54	0	22.4

Table 3.3: Number of fractional solutions and loops of the sequential r-blocks optimization method for $r = 2$, starting from the observed dolphin's social network. The four runs have been randomized.

In accordance with the results in Table 3.3, the number of selected fractional extreme points seems to be particularly high if compared to the ones observed in Table 3.2. Nonetheless, no consecutive identical networks are obtained.

Another interesting model has been used in Chapter 2 to deal with network transitivity. The system (2.24b) characterized the set of UNs with fixed number of triads, that is to say, the number of times three nodes are connected by a cycle of length three. Let f^T be the number of triads, corresponding to the non-linear property $\sum_{i<j<k\in\mathcal{V}} x_{ij}x_{jk}x_{ij} = f^T$, where x_{ij} is the

binary indicator of a tie, for $(i,j) \in \mathcal{H}^2$, associated to the $n(n-1)/2$ upper-diagonal components of the AM. A system of linear constraints characterizing the family of UNs with n nodes and f^T triads is (2.24b,2.24d).

The final model considered in this section is the family of DNs with fixed number of mutual, asymmetric, and null connections, corresponding to a fractional-polytope with no integer solution in the interior.

Holland and Leinhardt [118] initiated the study and application of uniform models for directed networks constrained to the dyad —group of two people— count, i.e., the number of mutual, asymmetric, and null dyads.

Denoting by f^M and f^N the number of, respectively, mutual and null dyads, and considering (i) binary variables $x_{ij}, x_{ji},\ (i,j) \in \mathcal{H}^2$ associated to the $n(n-1)$ non-diagonal components of the AM, and (ii) binary variables y_{ij}^M and $y_{ij}^N,\ (i,j) \in \mathcal{H}^2$—which are 1 if nodes i and j are, respectively, a mutual or null dyad, and 0 otherwise—, this problem can be formulated as

$$x_{ij} \geq y_{ij}^M \quad (i,j) \in \mathcal{H}^2 \tag{3.21a}$$
$$x_{ji} \geq y_{ij}^M \quad (i,j) \in \mathcal{H}^2 \tag{3.21b}$$
$$y_{ij}^M \geq x_{ij} + x_{ji} - 1 \quad (i,j) \in \mathcal{H}^2 \tag{3.21c}$$
$$1 - x_{ij} \geq y_{ij}^N \quad (i,j) \in \mathcal{H}^2 \tag{3.21d}$$
$$1 - x_{ji} \geq y_{ij}^N \quad (i,j) \in \mathcal{H}^2 \tag{3.21e}$$
$$y_{ij}^N \geq 1 - (x_{ij} + x_{ji}) \quad (i,j) \in \mathcal{H}^2 \tag{3.21f}$$
$$\sum_{(i,j)\in\mathcal{H}^2} y_{ij}^M = f^M \tag{3.21g}$$
$$\sum_{(i,j)\in\mathcal{H}^2} y_{ij}^N = f^N \tag{3.21h}$$
$$x_{ij}, x_{ji}, y_{ij}^M, y_{ij}^N \in \{0,1\} \quad (i,j) \in \mathcal{H}^2. \tag{3.21i}$$

From (3.21c), $y_{ij}^M = 0$ implies that x_{ij} and x_{ji} cannot be both equal to 1, whereas $y_{ij}^M = 1$ implies that both x_{ij} and x_{ji} must be equal to 1 (constraints (3.21a)–(3.21b)). Similarly, when $y_{ij}^N = 0$ one has that x_{ij} and x_{ji} cannot be both equal to 0 (constraint (3.21f)), whereas, when $y_{ij}^N = 1$, both x_{ij} and x_{ji} must be equal to 0 (constraints (3.21d)–(3.21e)). Thus, (3.21g) and (3.21h) force the number of mutual dyads and null dyads to be f^M and f^N, respectively. Consequently, $n(n-1)/2 - \sum_{(i,j)\in\mathcal{H}^2}(y_{ij}^M + y_{ij}^N)$ represents the number of asymmetric dyads.

Chapter 4

Specialized interior point methodologies

4.1 Purposes and preliminary overview

The previous chapter dealt with the characterization of several classes of networks by systems of linear constraints, enabling the researcher to apply integer and linear programming techniques to generate feasible points in specified polytopes. We showed that the coefficient matrices associated to several of those systems are TU, establishing a bijective association between given families of networks and the sets of extreme points of algebraically defined polytopes.

The three numerical procedures for network generation presented in Section 3.3 – the r-blocks, the s-pivots and the q-kernel – require the solution of LPs which might be computationally expensive for large number of nodes. Although state-of-the-art implementations of the simplex method and polynomial time interior point algorithms can be used, it is worth to exploit the problem structure whenever possible [70]. We argued in the previous chapter that under a proper row and column permutation, most of the constraints matrices of Section 3.2 exhibit a primal block-angular structure such as (3.19).

This chapter is of particular importance in the methodological setting of this thesis, as it embraces the overall collection of algebraical techniques used to generate random networks with specified structural properties. As in the previous chapters, we denote the vector of variables associated to the components of the AM with the boldface character $\mathbf{x}^T = [x_{12}, x_{13}, \ldots, x_{1n}, x_{23}, x_{24}, \ldots, x_{2n}, \ldots, x_{45}, \ldots, x_{(n-1)n}, x_{21}, x_{31}, x_{32}, x_{41}, \ldots, x_{n(n-1)}]$, corresponding to the lexicographic ordering of the rows and columns of the upper triangular part of the AM.

4.2 Specialized interior point methods for some classes of primal-block angular problems

The amount of computation required to solve a LP is related to its constraints structure, apart from its size. It can be shown that, under a proper row and column permutation, most of LPs associated to the families of networks of Section 3.2 can be reformulated as a primal

block-angular problem of the form

$$\begin{aligned} \min \quad & \sum_{h=0}^{k} (c^h)^T x^h && \text{(4.1a)} \\ \text{subject to} \quad & \begin{bmatrix} N_1 & & & & \\ & N_2 & & & \\ & & \ddots & & \\ & & & N_k & \\ L_1 & L_2 & \dots & L_k & I \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^k \\ x^0 \end{bmatrix} = \begin{bmatrix} b^1 \\ b^2 \\ \vdots \\ b^k \\ b^0 \end{bmatrix} && \text{(4.1b)} \\ & 0 \le x^h \le u^h \quad h = 0, \dots, k. && \text{(4.1c)} \end{aligned}$$

Matrices $N_h \in \mathbb{R}^{m_h \times n_h}$ and $L_h \in \mathbb{R}^{l \times n_h}$, $h = 1, \dots, k$, respectively define the block-diagonal and linking constraints, k being the number of blocks. Vectors $x^h \in \mathbb{R}^{n_h}, h = 1, \dots, k$, are the variables for each block. $x^0 \in \mathbb{R}^l$ are the slacks of the linking constraints. $b^h \in \mathbb{R}^{m_h}, h = 1, \dots, k$, is the right-hand side vector for each block of constraints, whereas $b^0 \in \mathbb{R}^l$ is for the linking constraints. The upper bounds for each group of variables are defined by $u^h, h = 0, \dots, k$; in our problems $u^h = \mathbf{e}$, i.e., a vector of ones.

As already mentioned in Chapter 1, most of the solution strategies to deal with problems in the form of (4.1) can be broadly classified into four main categories: simplex-based methods [57, 159], decomposition methods [11, 90], approximation methods [27], and interior point methods [11, 50].

Here we are taking into account the specialized interior point method, first proposed by Castro [50, 52, 53]. (For more details about prima-dual interior point methods, see Appendix B). Briefly, the algorithm here considered requires the solution at each interior point iteration of the system

$$\begin{aligned} A\Theta A^T \Delta_y &= \left[\begin{array}{ccc|c} N_1\Theta_1 N_1^T & & & N_1\Theta_1 L_1^T \\ & \ddots & & \vdots \\ & & N_k\Theta_k N_k^T & N_k\Theta_k L_k^T \\ \hline L_1\Theta_1 N_1^T & \dots & L_k\Theta_k N_k^T & \Theta_0 + \sum_{h=1}^{k} L_h\Theta_h L_h^T \end{array}\right] \Delta_y \\ &= \begin{bmatrix} B & C \\ C^T & D \end{bmatrix} \begin{bmatrix} \Delta_{y_1} \\ \Delta_{y_2} \end{bmatrix} = g, \end{aligned} \qquad \text{(4.2)}$$

where Δ_y is the direction of movement for the dual variables, $\Theta_h = ((U - X_h)^{-1} W_h + X_h^{-1} Z_h)^{-1}$ $h = 0, \dots, k$, are diagonal matrices, and g is some right-hand side. By eliminating Δ_{y_1} from the first group of equations of (4.2), we obtain

$$\begin{aligned} (D - C^T B^{-1} C)\Delta_{y_2} &= g_1 && \text{(4.3a)} \\ B\Delta_{y_1} &= g_2, && \text{(4.3b)} \end{aligned}$$

for a proper partition of the right-hand side into g_1 and g_2. One of the most efficient interior point methods for this class of problems solves the normal equations by a combination of k Cholesky factorizations, for the system involving B, and preconditioned conjugate gradient for system involving $(D - C^T B^{-1} C)$ –the Schur complement of (4.2) with dimension l, the number

of linking constraints–[1]. Note that the inverse of $(D - C^T B^{-1} C)$ might be written as an infinite power series:

$$(D - C^T B^{-1} C)^{-1} = \left(\sum_{j=0}^{\infty} (D^{-1}(C^T B^{-1} C))^j \right) D^{-1}. \tag{4.4}$$

A preconditioner is thus obtained by truncating the infinite power series (4.4) at some term. The more the terms included, the better the preconditioner will be, at the expense of increasing the execution time of each preconditioned conjugate gradient iteration. If we only include the first term of the infinite power series (4.4) the resulting preconditioner will be D^{-1} and if we include the first and second term the resulting preconditioner will be $(I + D^{-1}(C^T B^{-1} C))D^{-1}$.

Although the expected performance for a general primal block-angular matrix is problem dependent, the effectiveness of the preconditioner obtained by by truncating the infinite power series (4.4) is governed by the spectral radius of $(D^{-1}(C^T B^{-1} C))$ and the structure of matrix D, as each preconditioner conjugated gradient iteration entails the solution of a system with matrix D.

Recently a splitting preconditioner based on LU factorization was introduced [28]. Given an LP in standard form with coefficient (full rank) matrix $A \in \mathbb{R}^{m' \times n'}$ and right-hand term $\mathbf{b} \in \mathbb{R}^{m'}$, where $m' < n'$, the splitting preconditioner for normal equations can be obtained as follows. Consider a partition of the columns of A into two groups of basic and non basic columns, forming the basic matrix A_B and A_N, respectively. Applying the same partition to Θ, the normal equations matrix can be rewritten as

$$A\Theta A^T = A_B \Theta_B A_B^T + A_N \Theta_N A_N^T \tag{4.5}$$

The symmetric application of the preconditioner $\Theta_B^{-1/2} A_B^{-1}$ to matrix $(A\Theta A^T)$ gives:

$$\begin{aligned} (\Theta_B^{-1/2} A_B^{-1})(A\Theta A^T)(\Theta_B^{-1/2} A_B^{-1})^T &= \Theta_B^{-1/2} A_B^{-1}(A_B \Theta_B A_B^T + A_N \Theta_N A_N^T) A_B^{-T} \Theta_B^{-1/2} \\ &= I + (\Theta_B^{-1/2} A_B^{-1} A_N \Theta_N^{1/2})(\Theta_B^{-1/2} A_B^{-1} A_N \Theta_N^{1/2})^T. \end{aligned} \tag{4.6}$$

Sufficiently close to an optimal solution, with a suitable choice of the columns of A_B, the diagonal entries of Θ_B^{-1} and Θ_N are very small and $(\Theta_B^{-1/2} B^{-1} A_N \Theta_N^{-1/2})$ approaches the zero matrix. To find matrix A_B, the first m' linearly independent columns of $A\Theta$ with smallest 1-norm are selected.

The main appeal of this class of preconditioners is that they work better near a solution of the linear programming problem. This is a very welcome feature since the linear system is known to be very ill-conditioned close to the optimizer, making difficult its solution by iterative methods. Thus, an hybrid preconditioner combines the power series preconditioner (either D^{-1} or $(I + D^{-1}(C^T B^{-1} C))D^{-1}$) in the initial primal-dual iterations and switches to the splitting preconditioner $\Theta_B^{-1/2} A_B^{-1}$ when the former becomes inefficient [28].

This algorithm has been recently implemented in a software package, called *BlockIP* [54], that will be used for the computational results of this chapter.

[1] The preconditioned conjugate gradient algorithm is a suitable method for solving linear systems with symmetric and positive-definite matrix. It can be applied to sparse systems that are too large to be handled by direct methods such as the Cholesky decomposition. Preconditioning is simply a manipulation of the original system in a way that improves the spectral radius of its coefficient matrix, and hence its rate of convergence. The preconditioner itself is nothing more than a matrix M such that $M^{-1}A$ has a better spectral radius. The obvious best choice is $M = A$.

4.3 A new preconditioner for block-angular problems

The main source of difficulty in solving a primal block-angular problem of the form (4.1) is the presence of linearly independent linking constraints, $L_1x_1 + L_2x_2 + \ldots + L_kx_k \leq b_0$. Consider the lucky case where A is such that for every $i = 1 \ldots k$ each row vector of L_i belong to the column space of N_i^T (linearly linking constraints). By the definition of C, B, C and D in (4.2), we may write

$$C^T B^{-1} C = \sum_{i=1}^{k} L_i \Theta_i N_i^T (N_i \Theta_i N_i^T)^{-1} N_i \Theta_i L_i^T. \tag{4.7}$$

For every $i = 1 \ldots k$ we have that $P_i = \Theta_i^{1/2} N_i^T (N_i \Theta_i N_i^T)^{-1} N_i \Theta_i^{1/2}$ is the projection operator onto $\mathcal{R}(\Theta_i^{1/2} N_i^T)$, the column space of $\Theta_i^{1/2} N_i^T$. If each column vector of $\Theta_i^{1/2} L_i^T$ belong to $\mathcal{R}(\Theta_i^{1/2} N_i^T)$ (which is equivalent to say that the row vectors of L_i might be written as a linear combination of the row vectors of N_i) then $P_i \Theta_i^{1/2} L_i^T = \Theta_i^{1/2} L_i^T$ and

$$\begin{aligned} D - C^T B^{-1} C &= D - \sum_{i=1}^{k} L_i \Theta_i^{1/2} P_i \Theta_i^{1/2} L_i^T \\ &= D - \sum_{i=1}^{k} L_i \Theta_i L_i^T = \Theta_0. \end{aligned} \tag{4.8}$$

On the contrary, if each column vectors of $\Theta_i^{1/2} L_i^T$ belong to $\mathcal{N}(N_i \Theta_i^{1/2})$, the null space of $N_i \Theta_i^{1/2}$, then $P_i \Theta_i^{1/2} L_i^T = 0$ and

$$D - C^T B^{-1} C = D - \sum_{i=1}^{k} L_i \Theta_i^{1/2} P_i \Theta_i^{1/2} L_i^T = D \tag{4.9}$$

Thus, we see that Θ_0 and D are the results of $D - C^T B^{-1} C$ in the two opposite cases when, for $i = 1 \ldots k$, each column vector of $\Theta_i^{1/2} L_i^T$ belong to either $\mathcal{R}(\Theta_i^{1/2} N_i^T)$ or $\mathcal{N}(N_i \Theta_i^{1/2})$ respectively[2].

4.3.1 Geometrical and spectral properties

The goodness of the approximation of Θ_0^{-1} and D^{-1} to $(D - C^T B^{-1} C)^{-1}$ might be measured by the principal angles between the range of $\Theta_i^{1/2} N_i^T$ and the range of $\Theta_i^{1/2} L_i^T$, for $i = 1, \ldots, k$ – as Θ_i changes in each primal-dual iteration, the angles between the subspaces also changes resulting in a dynamical approximation of Θ_0^{-1} and D^{-1} to $(D - C^T B^{-1} C)^{-1}$ –.

The principal angles provide information about the relative position of two subspaces of an inner product space. Consider the subspaces $\mathcal{L}_\Theta$ and $\mathcal{N}_\Theta$ of an inner product space $\mathcal{R}_\Theta$, with $\dim \mathcal{L}_\Theta = \dim \mathcal{N}_\Theta = l$. The principal angles between $\mathcal{L}_\Theta$ and $\mathcal{N}_\Theta$, $0 \leq \gamma_1 \leq \ldots \leq \gamma_l \leq \pi/2$ are given by [25]

$$\cos(\gamma_j) = \max \quad \langle u \cdot v \rangle \tag{4.10}$$

$$\text{subject to} \quad ||u|| = 1, \ ||v|| = 1 \tag{4.11}$$

$$u \in \mathcal{L}_\Theta, \ v \in \mathcal{N}_\Theta \tag{4.12}$$

$$\langle v_k \cdot v \rangle = 0, \ \langle u_k \cdot u \rangle = 0, \ k = 1 \ldots j-1 \tag{4.13}$$

[2]Note that matrix P_i is the identity operator of the subspace generated by the columns of $\Theta_i^{1/2} N_i^T$.

The vectors $\{u_1, \ldots, u_l\}$ and $\{v_1, \ldots, v_l\}$ are called *principal vectors*, associated to principal angles $\{\gamma_1, \ldots, \gamma_l\}$. The principal angles between subspaces can be graphically depicted as in Figure 4.1. They can be computed by the singular value decomposition.

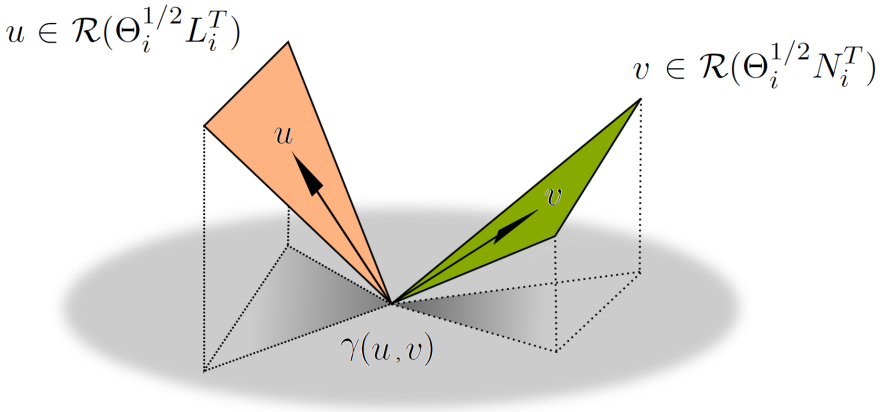


Figure 4.1: Angles between subspaces.

Theorem 3 (Bjorck and Golub [25])**.** *Let the columns of matrices Q_L and Q_N form orthonormal bases for the subspaces $\mathcal{L}$ and $\mathcal{N}$, correspondingly. Principal vectors $\widetilde{u}$ and $\widetilde{v}$ must verify $\widetilde{u} = Q_N^T u$ and $\widetilde{v} = Q_L v$, where u and v are left and right singular vectors of $Q_N Q_L^T$, associated to the singular value $\cos(\gamma(\widetilde{u}, \widetilde{v}))$, that is to say, $(Q_N Q_L^T)v = \cos(\gamma(\widetilde{u}, \widetilde{v}))\, u$.*

Using this singular value decomposition based equivalence, Proposition 9 below helps to understand how the goodness of the approximation of Θ_0^{-1} and D^{-1} to $(D - C^T B^{-1} C)^{-1}$ dynamically change along the interior point iterations and how the principal angles between $\mathcal{L}_i$ and $\mathcal{N}_i$ affect the principal angles between $\mathcal{L}_{\Theta_i}$ and $\mathcal{N}_{\Theta_i}$. (For more details about principal angles between subspaces and their relation with the singular value decomposition, see [25])

Proposition 9. *Consider the subspaces $\mathcal{L} \subseteq \mathbb{R}^n$ and $\mathcal{N} \subseteq \mathbb{R}^n$, with corresponding orthogonal bases Q_L and Q_N, and their image sets $\mathcal{L}_\Theta \subseteq \mathbb{R}^n$ and $\mathcal{N}_\Theta \subseteq \mathbb{R}^n$ under the linear transformation $\Theta^{1/2}$, where Θ is a positive diagonal matrix. (Q_L and Q_N are respectively obtained from L^T and N^T by QR decomposition.) Let $\gamma(\widetilde{u}, \widetilde{v})$ be a principal angle between $\mathcal{L}$ and $\mathcal{N}$, associated to principal vectors $\widetilde{u}$ and $\widetilde{v}$, and $\gamma(\Theta^{1/2}\widetilde{u}, \Theta^{1/2}\widetilde{v})$, the angle between the transformed vectors $\Theta^{1/2}\widetilde{u} \in \mathcal{N}_\Theta$ and $\Theta^{1/2}\widetilde{v} \in \mathcal{L}_\Theta$. The following inequalities must hold:*

$$\frac{1}{\cos(\gamma(\widetilde{u}, \widetilde{v}))} \min\left\{ \frac{\varrho_{\min}}{\Theta_{\max}}, \frac{\varrho_{\min}}{\Theta_{\min}} \right\} \leq \cos\left(\gamma(\Theta^{1/2}\widetilde{u}, \Theta^{1/2}\widetilde{v})\right), \tag{4.14a}$$

$$\frac{1}{\cos(\gamma(\widetilde{u}, \widetilde{v}))} \max\left\{ \frac{\varrho_{\max}}{\Theta_{\max}}, \frac{\varrho_{\max}}{\Theta_{\min}} \right\} \geq \cos\left(\gamma(\Theta^{1/2}\widetilde{u}, \Theta^{1/2}\widetilde{v})\right), \tag{4.14b}$$

where $\Theta_{\min}$ and $\Theta_{\max}$ are the least and greatest component of matrix Θ and $\varrho_{\min}$ and $\varrho_{\max}$ are the smallest and greatest eigenvalues of $Q_L Q_N^T Q_N \Theta Q_L^T$.

Proof. By the definition of cosine we have

$$\cos(\gamma(\widetilde{u}, \widetilde{v})) \geq \frac{\widetilde{u}^T \widetilde{v}}{||\widetilde{u}||\ ||\widetilde{v}||} \tag{4.15a}$$

and

$$\cos\left(\gamma(\Theta^{1/2}\widetilde{u},\Theta^{1/2}\widetilde{v})\right) \quad \geq \quad \frac{\widetilde{u}^T\Theta\widetilde{v}}{||\Theta^{1/2}\widetilde{u}||\ ||\Theta^{1/2}\widetilde{v}||} \tag{4.15b}$$

From Theorem 3 we know that principal vectors $\widetilde{u}$ and $\widetilde{v}$ must verify $\widetilde{u} = Q_N^T u$ and $\widetilde{v} = Q_L^T v$, where u and v are left and right singular vectors of $Q_N Q_L^T$, associated to the singular value $\cos(\gamma(\widetilde{u},\widetilde{v}))$, that is to say, $(Q_N Q_L^T)v = \cos(\gamma(\widetilde{u},\widetilde{v}))\, u$. (For more details about principal angles between subspaces and their relation with the singular value decomposition, see [25]) Applying this property we find

$$\begin{aligned}
\widetilde{u}^T\Theta\widetilde{v} &= (Q_N^T u)^T\Theta Q_L^T v \\
&= u^T Q_N\Theta Q_L^T v \\
&= \frac{((Q_N Q_L^T)v)^T Q_N\Theta Q_L^T v}{\cos(\gamma(\widetilde{u},\widetilde{v}))} \\
&= \frac{v^T\left(Q_L Q_N^T Q_N\Theta Q_L^T\right)v}{\cos(\gamma(\widetilde{u},\widetilde{v}))}
\end{aligned} \tag{4.16a}$$

Let $\varrho_{\min}$ and $\varrho_{\max}$ be the smallest and greatest eigenvalues of $Q_L Q_N^T Q_N\Theta Q_L^T$. We have

$$\widetilde{u}^T\Theta\widetilde{v} \quad \geq \quad \frac{1}{\cos(\gamma(\widetilde{u},\widetilde{v}))}\min_{||v||_2=1} v^T\left(Q_L Q_N^T Q_N\Theta Q_L^T\right)v = \frac{\varrho_{\min}}{\cos(\gamma(\widetilde{u},\widetilde{v}))} \tag{4.16b}$$

and

$$\widetilde{u}^T\Theta\widetilde{v} \quad \leq \quad \frac{1}{\cos(\gamma(\widetilde{u},\widetilde{v}))}\max_{||v||_2=1} v^T\left(Q_L Q_N^T Q_N\Theta Q_L^T\right)v = \frac{\varrho_{\max}}{\cos(\gamma(\widetilde{u},\widetilde{v}))} \tag{4.16c}$$

The result is that

$$\cos\left(\gamma(\Theta^{1/2}\widetilde{u},\Theta^{1/2}\widetilde{v})\right) \quad \geq \quad \frac{\varrho_{\min}}{\cos(\gamma(\widetilde{u},\widetilde{v}))\,||\Theta^{1/2}\widetilde{u}||\ ||\Theta^{1/2}\widetilde{v}||} \tag{4.16d}$$

and

$$\cos\left(\gamma(\Theta^{1/2}\widetilde{u},\Theta^{1/2}\widetilde{v})\right) \quad \leq \quad \frac{\varrho_{\max}}{\cos(\gamma(\widetilde{u},\widetilde{v}))\,||\Theta^{1/2}\widetilde{u}||\ ||\Theta^{1/2}\widetilde{v}||} \tag{4.16e}$$

Observing that $||\widetilde{u}|| = ||\widetilde{u}|| = 1$, the proof can then be completed by few algebraical manipulations:

$$\frac{1}{\cos(\gamma(\widetilde{u},\widetilde{v}))}\min\left\{\frac{\varrho_{\min}}{\Theta_{\max}},\frac{\varrho_{\min}}{\Theta_{\min}}\right\} \quad \leq \quad \cos\left(\gamma(\Theta^{1/2}\widetilde{u},\Theta^{1/2}\widetilde{v})\right), \tag{4.16f}$$

$$\frac{1}{\cos(\gamma(\widetilde{u},\widetilde{v}))}\max\left\{\frac{\varrho_{\max}}{\Theta_{\max}},\frac{\varrho_{\max}}{\Theta_{\min}}\right\} \quad \geq \quad \cos\left(\gamma(\Theta^{1/2}\widetilde{u},\Theta^{1/2}\widetilde{v})\right). \tag{4.16g}$$

□

An important consequence of Proposition 9 is that we have no information about the principal angles in the final iterations, as $\Theta_{\min}/\Theta_{\max}$ approaches the zero when the iterative process get close to the optimal solution. However, Proposition 9 helps us understand how the goodness of the approximation of Θ_0^{-1} and D^{-1} to $(D - C^T B^{-1} C)^{-1}$ dynamically change along the interior point iterations.

A downside of the use of principal angles between subspaces for this purpose appears when $\mathcal{R}(\Theta_i^{1/2} L_i^T) = \mathbb{R}^{n_i}$, for $i = 1, \ldots, k$. In such case the principal angles are always zeros, regardless

of what $\mathcal{R}(\Theta_i^{1/2}N_i^T)$ is. This happens in multi-commodity network flow problems and edge-colored network problems, where $L_i = I$, for $i = 1, \ldots, k$, as shown in Section 4.6.

The goodness of the dynamical approximation of Θ_0^{-1} and D^{-1} to $(D - C^TB^{-1}C)^{-1}$ along the interior point iterations is related to the changes in spectral radius of matrix $D^{-1}(C^TB^{-1}C)$ – which is always in $[0, 1)$ [50, Theorem 1] –. Since D constitute the first term of the power series (4.4), the farther away from 1 is the spectral radius of $D^{-1}(C^TB^{-1}C)$ the better is the quality of the approximation of the first few terms of (4.4), obtained by truncation with $h = 0$ or $h = 1$. Although the particular behavior of the spectral radius value is problem dependent, in general, it comes closer to 1 as we approach the optimal solution, because of the ill-conditioning of the Θ matrix.

Next result provides clear relationship between the spectral radius of $D^{-1}(C^TB^{-1}C)$ and the projection operators in the subspaces $\mathcal{L}_\Theta$ and $\mathcal{N}_\Theta$.

Proposition 10. *Let λ be an arbitrary eigenvalue of $(D^{-1}(C^TB^{-1}C))$. If each column vector of $\Theta_i^{1/2}L_i^T$ belong to $\mathcal{R}(\Theta_i^{1/2}N_i^T)$ then*

$$\lambda = \frac{r^T\left(\sum_{i=1}^k L_i\Theta_i L_i^T\right)r}{r^TDr}, \tag{4.17}$$

where r is the corresponding eigenvector associated to λ. On the contrary, if each column vectors of $\Theta_i^{1/2}L_i^T$ belong to $\mathcal{N}(N_i\Theta_i^{1/2})$, the null space of $N_i\Theta_i^{1/2}$, then

$$\lambda = 0. \tag{4.18}$$

Proof. Eigenvalue λ of $(D^{-1}(C^TB^{-1}C))$ satisfies $(C^TB^{-1}C)r = \lambda Dr$ for some eigenvector r. From the definition of C, B, D in (4.2) we have

$$\left(\sum_{i=1}^k L_i\Theta_i N_i^T\left(N_i\Theta_i N_i^T\right)^{-1}N_i\Theta_i L_i^T\right)r = \lambda Dr,$$

$$\Updownarrow$$

$$(1-\lambda)\,Dr = \left(\Theta_0 + \sum_{i=1}^k L_i\Theta_i L_i^T\right)r - \left(\sum_{i=1}^k L_i\Theta_i N_i^T\left(N_i\Theta_i N_i^T\right)^{-1}N_i\Theta_i L_i^T\right)r,$$

$$\Downarrow$$

$$(1-\lambda) = \frac{r^T\Theta_0 r + r^T\left(\sum_{i=1}^k L_i\Theta_i^{1/2}(I-P_i)\Theta_i^{1/2}L_i^T\right)r}{r^TDr}. \tag{4.19}$$

From this expression and applying the definition of $W_\mathcal{R}$ and $W_\mathcal{N}$

$$W_\mathcal{N} = \sum_{i=1}^k L_i\Theta_i^{1/2}(I-P_i)\Theta_i^{1/2}L_i^T \quad \text{and} \quad W_\mathcal{R} = \sum_{i=1}^k L_i\Theta_i^{1/2}(P_i)\Theta_i^{1/2}L_i^T, \tag{4.20}$$

we find the following two identities:

$$\lambda = \frac{r^T\left(\sum_{i=1}^k L_i\Theta_i L_i^T\right)r - r^TW_\mathcal{N}r}{r^TDr} \quad \text{and} \quad \lambda = \frac{r^TW_\mathcal{R}r}{r^TDr}. \tag{4.21}$$

that is to say, the matrices associated to the inner products between the row vectors of $L_i\Theta_i^{1/2}$ and their projections into $\mathcal{R}(\Theta_i^{1/2}N_i^T)$ and $\mathcal{N}(N_i\Theta_i^{1/2})$ respectively. It turns out that $r^TW_\mathcal{N}r = 0$

when each column vector of $\Theta_i^{1/2} L_i^T$ belongs to $\mathcal{R}(\Theta_i^{1/2} N_i^T)$ and $r^T W_{\mathcal{R}} r = 0$ when each column vector of $\Theta_i^{1/2} L_i^T$ belongs to $\mathcal{N}(N_i \Theta_i^{1/2})$. The proof is complete.

□

Note that a clear lesson from Proposition 10 is that the spectral radius λ might be small even in the case of almost collinearity between the row vectors of $L_i \Theta_i^{1/2}$ and their projections into $\mathcal{R}(\Theta_i^{1/2} N_i^T)$, when $r^T \Theta_0 r \gg r^T (\sum_{i=1}^k L_i \Theta_i L_i^T) r$. This is consistent with following results from Castro and Cuesta [55].

Theorem 4 (Castro and Cuesta [55]). *Let A be the constraint matrix of problem (4.1), with full row rank matrices $N_i \in \mathbb{R}^{m_i \times n_i}$ $i = 1, \ldots, k$, and at least one full row rank matrix $L_i \in \mathbb{R}^{l \times n_i}$, $i = 1, \ldots, k$. Then, the spectral radius ρ of $D^{-1}(C^T B^{-1} C)$ is bounded by*

$$0 \leq \rho \leq \max_{j \in \{1,\ldots,l\}} \frac{\gamma_j}{\left(\frac{u_j}{v_j}\right)^2 \Theta_{0j} + \gamma_j} < 1, \tag{4.22}$$

where u is the eigenvector (or one of the eigenvectors) of $D^{-1}(C^T B^{-1} C)$ for ρ; γ_j, $j = 1, \ldots, l$, and $V = [V_1 \ldots V_l]$, are respectively the eigenvalues and matrix of columnwise eigenvectors of $\sum_{i=1}^k L_i \Theta_i {L_i}^T$; $v = V^T u$; and, abusing of notation, we assume that for $v_j = 0$, $(u_j/v_j)^2 = +\infty$.

4.4 Numerical validation

Consider two full rank matrices $N \in \mathbb{R}^{m \times n}$ and $\Lambda \in \mathbb{R}^{l \times m}$ and let $L = \lambda N$. The rows of $L \in \mathbb{R}^{l \times n}$ are linear combinations of the rows of N and the principal angle between the subspaces generated by L^T and N^T is zero. Rotate each vector in $\mathcal{R}(L^T)$ around the i^{th} and j^{th} coordinate axes by an angle α, pre-multiplying L by the block matrix

$$R_{ij}(\alpha) = \begin{bmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & 0 & \cdots & 0 & \vdots & \ddots & \vdots \\ 0 & \cdots & \cos(\alpha) & 0 & \cdots & 0 & -\sin(\alpha) & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ldots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & -\sin(\alpha) & 0 & \cdots & 0 & \cos(\alpha) & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ldots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{4.23}$$

If we consider $H = \{(i,j) : 1 \leq i \leq n-1, i < j \leq n\}$, the set of all possible coordinate axes, the $n(n-1)/2$ distinct $R_{ij}(\alpha)$ rotation matrices may be concatenated in some order to produce a new rotation matrix such as $\prod_{(i,j) \in \rho} R_{ij}(\alpha)$, where $\rho \subseteq H$[3].

We are interested in showing the performance of the PCG method in each interior point iteration when changing the geometrical relations between the diagonal block matrices and the linking constraint matrices, in accordance with specified rotations. To do so we randomly draw $N_i \in \mathbb{R}^{m \times n}$ and $\Lambda_i \in \mathbb{R}^{l \times m}$ from a uniform probability distribution. Then $\rho \subseteq H$ is also

[3] Rotations in three dimensions (and higher) do not commute, so that different orderings give different rotations. Thus, $\rho \subseteq H$ is an ordered set and the number of ways it might be selected is $|H|!/(|H| - |\rho|)!$, where $|H| = n(n-1)/2$ and $|\rho| = \overline{r}$.

randomly selected to compute the rotation matrix ($\prod_{(i,j)\in\rho} R_{ij}(\alpha)$). Finally, obtain the linking constraint matrices $L_i = \Lambda_i N_i \left(\prod_{ij} R_{ij}(\alpha)\right)$.

We take into account the number $\overline{r} = |\rho|$ of concatenated rotation matrices and the angle of rotation α to evaluate the associated changes in the PCG iterations using Θ_0^{-1} and D^{-1} as preconditioners.

The computational results shown in tables 4.1 and 4.2 report the CPU time associated to the BlockIP algorithm (using D^{-1} and Θ_0^{-1} preconditioner) and the CPLEX available LP methods (Primal Simplex, Dual Simplex, Barrier Method). The first column show the number $\overline{r}$ of rotation matrices multiplied to obtain the linking constrain matrix $L\left(\prod_{ij} R_{ij}(\alpha)\right)$, whereas the second columns reports the angle α of each rotation. The remaining columns of tables 4.1 and 4.2 give the CPU time and number of iterations (in parentheses) for the three algorithms tested.

α	$\overline{r}$	CPLEX 12.5			BlockIP	
		Primal Simplex	Dual Simplex	Barrier	Θ_0^{-1}	D^{-1}
$\pi/14$	2	0.93 (7107)	1.23 (2190)	1.30 (19)	1.00 (26)	1.70 (26)
$\pi/10$	4	2.25 (6210)	1.81 (3171)	1.34 (20)	1.31 (32)	1.34 (31)
$\pi/6$	6	1.37 (8555)	1.89 (2869)	1.28 (19)	2.20 (49)	2.59 (55)
$\pi/2$	8	3.13 (10158)	1.57 (2326)	1.35 (20)	2.00 (39)	2.10 (47)

Table 4.1: CPU time of BlockIP (using D^{-1} and Θ_0^{-1} preconditioner) and CPLEX available LP methods (Primal Simplex, Dual Simplex, Barrier Method). The LPs have $l = 100$ linking constraints and $k = 100$ equal diagonal block matrices $N \in \mathbb{R}^{10\times 50}$.

α	$\overline{r}$	CPLEX 12.5			BlockIP	
		Primal Simplex	Dual Simplex	Barrier	Θ_0^{-1}	D^{-1}
$\pi/14$	2	913.95 (380190)	193.27 (50774)	323.44 (25)	233.93 (31)	615.43 (77)
$\pi/10$	4	1896.22 (413821)	295.37 (57418)	341.19 (25)	365.56 (44)	557.93 (61)
$\pi/6$	6	$>$ 3000.00 (477039)	359.95 (62830)	656.23 (51)	608.34 (60)	559.89 (63)
$\pi/2$	8	$>$ 3000.00 (453654)	315.85 (55586)	364.53 (28)	919.33 (79)	697.96 (75)

Table 4.2: CPU time of BlockIP (using D^{-1} and Θ_0^{-1} preconditioner) and CPLEX available LP methods (Primal Simplex, Dual Simplex, Barrier Method). The LPs have $l = 200$ linking constraints and $k = 1000$ equal diagonal block matrices $N \in \mathbb{R}^{20\times 200}$.

It can be seen from tables 4.1 and 4.2 that the CPU time associated to the dual simplex and barrier method are not correlated with the angles between the rows of L and the ones of N. On the contrary, the computational performances of the specialized interior point method strongly relies on those angles for both small instances, as in Table 4.1, and big instances, as in Table 4.2. What seems to be surprising is that the primal simplex algorithm resulted to be affected by the angles.

Tables 4.3 and 4.4 report the average PCG iteration for each IP iteration associated to D^{-1} and Θ_0^{-1} preconditioners, for the same LPs of tables 4.1 and 4.2.

α	$\overline{r}$	BlockIP	
		Θ_0^{-1}	D^{-1}
$\pi/14$	2	5.034483	13.172414
$\pi/10$	4	9.705882	12.794118
$\pi/6$	6	13.435897	12.615385
$\pi/2$	8	14.031250	11.375000

Table 4.3: Average PCG iteration for each IP iteration of D^{-1} and Θ_0^{-1} preconditioners, for LPs with $l = 100$ linking constraints and $k = 100$ equal diagonal block matrices $N \in \mathbb{R}^{10\times 50}$.

α	$\overline{r}$	BlockIP	
		Θ_0^{-1}	D^{-1}
$\pi/14$	2	3.3	19.2
$\pi/10$	4	15.4	32.9
$\pi/6$	6	28.9	28.6
$\pi/2$	8	30.2	27.3

Table 4.4: Average PCG iteration for each IP iteration of D^{-1} and Θ_0^{-1} preconditioners, for LPs with $l = 200$ linking constraints and $k = 200$ equal diagonal block matrices $N \in \mathbb{R}^{10\times 50}$.

A substantial effect of the geometrical relations between the diagonal block matrices and the linking constraint matrices clearly emerges, as the average number of PCG iterations appears to increase or decrease in accordance with the angles for the cases of Θ_0^{-1} and D^{-1} respectively.

The plots of figures 4.2, 4.3, 4.4 and 4.5 support this fact by illustrating the evolution of the PCG iterations for different level of the $\log \mu$ (the natural logarithm of the barrier parameter). The blue and green lines show polynomial curves of degree four which have been fitted to the observed PCG iterations associated to Θ_0^{-1} and D^{-1} respectively.

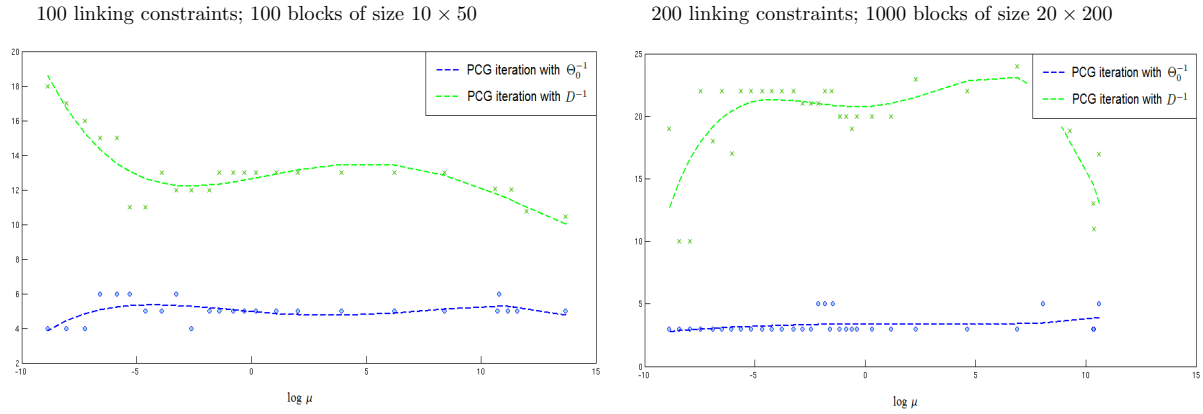


Figure 4.2: PCG iterations along the primal-dual algorithm for $\alpha = \pi/16$, $\overline{r} = 2$. The blue and green lines show the average PCG iterations associated to Θ_0^{-1} and D^{-1}.

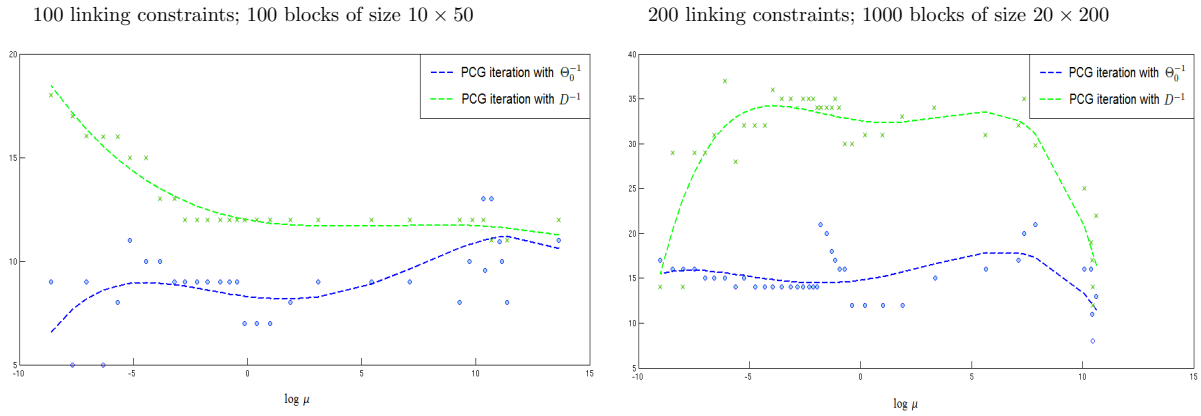


Figure 4.3: PCG iterations along the primal-dual algorithm for $\alpha = \pi/12$, $\overline{r} = 6$. The blue and green lines show the average PCG iterations associated to Θ_0^{-1} and D^{-1}.

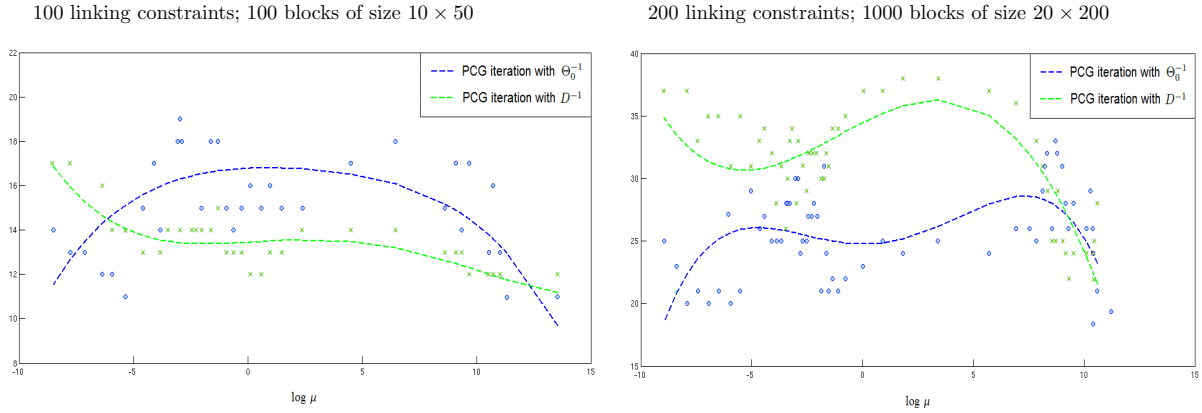


Figure 4.4: PCG iterations along the primal-dual algorithm for $\alpha = \pi/8$, $\overline{r} = 10$. The blue and green lines show the average PCG iterations associated to Θ_0^{-1} and D^{-1}.

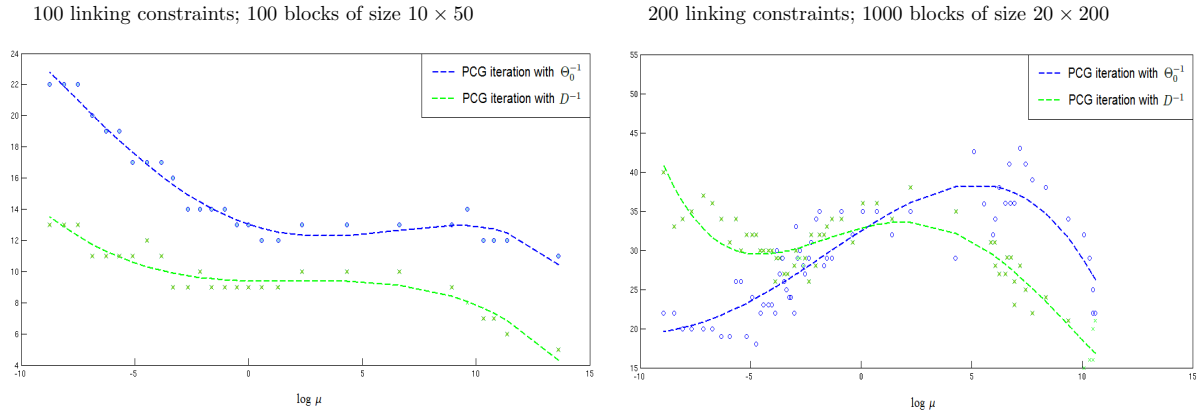


Figure 4.5: PCG iterations along the primal-dual algorithm for $\alpha = \pi/4$, $\overline{r} = 14$. The blue and green lines show the average PCG iterations associated to Θ_0^{-1} and D^{-1}.

4.5 Multicommodity network flow problem with nodal capacity

Network flow optimization problems have been introduced in Subsection 1.1.4. The general notion of network flow of a given commodity $c \in \mathcal{C}$ has been described as a real-valued function $\omega^i : \mathcal{V} \times \mathcal{V} \longrightarrow \mathbb{R}$, which is denoted as $x^c_{ij} = \omega^c(i,j)$, for $(i,j) \in \mathcal{V} \times \mathcal{V}$, with the property of flow conservation described in (1.23). Here a multicommodity network flow problem with nodal capacities (MNFPNC from now on) is considered. The nodal capacities consist of an upper limit on the total out-flow of nodes: $\sum_{i \in \mathcal{C}} \sum_{j \in \mathcal{V}} x^i_{hj} \leq q_h$, for $h \in \mathcal{V}$.

The general mathematical programming model for a network of n' arcs and n nodes is (4.1), where vectors $x^i \in \mathbb{R}^{n'}$, $i \in \mathcal{C}$ are the flows for commodity i; $x^0 \in \mathbb{R}^n$ are the slacks of nodal capacities constraints when total out-flow of nodes is considered. The node-arc incidence matrix of the directed graph is $N_i = N \in \mathbb{R}^{n' \times (n-1)}$, for $i \in \mathcal{C}$. Vectors $b^i \in \mathbb{R}^{n-1}$ are the node supply/demands and $c^i \in \mathbb{R}^{n'}$ are the arc linear costs for each commodity $i \in \mathcal{C}$. The linking constraints matrices $L_i = L \in \mathbb{R}^{n \times n'}$ are derived from N by switching all negative signs to positive.

Tables 4.5 and 4.6 report two computational experiments associated to MNFPNCs of different sizes. The eight instances of Table 4.5 correspond to problems where only 20% of nodes are constrained to have an out-flow capacity, whereas the eight instances of Table 4.6 correspond to problems where all the nodes are constrained to have an out-flow capacity.

n	k	n. var.	n. con.	CPLEX 12.5			BlockIP	
				Primal Simplex	Dual Simplex	Barrier	Θ_0^{-1}	D^{-1}
150	200	880000	30030	59.5 (997332)	5.9 (85732)	9.0 (18)	7.2 (38)	6.3 (34)
150	400	1803200	60030	246.7 (428032)	15.1 (163391)	22.9 (19)	15.1 (39)	13.5 (35)
150	600	2601000	90030	418.2 (444258)	28.1 (249271)	40.1 (21)	23.7 (37)	22.0 (38)
150	800	3850400	120030	685.5 (575356)	38.0 (329632)	49.6 (21)	34.3 (38)	34.5 (38)
300	200	3587400	60060	541.8 (4676110)	31.9 (203873)	54.2 (21)	40.8 (39)	42.0 (38)
300	400	7022400	120060	> 3000 (1902592)	75.5 (415156)	112.8 (22)	82.3 (38)	81.5 (38)
300	600	11245200	180060	> 3000 (2509984)	116.9 (571964)	229.4 (23)	128.2 (39)	129.1 (37)
300	800	14862400	240060	> 3000 (1077201)	219.4 (784104)	301.1 (23)	181.4 (39)	174.6 (38)

Table 4.5: CPU time and iteration (into parenthesis) of BlockIP (using D^{-1} and Θ_0^{-1} preconditioner) and CPLEX available LP methods (Primal Simplex, Dual Simplex, Barrier Method) when solving MNFPNCs. Only 20% of nodes are conditioned to have an out-flow capacity, so that $l = 0.2n$.

n	k	n. var.	n. con.	CPLEX 12.5			BlockIP	
				Primal Simplex	Dual Simplex	Barrier	Θ_0^{-1}	D^{-1}
150	200	880000	30030	76.1 (967177)	6.6 (86446)	16.7 (22)	7.3 (37)	7.1 (35)
150	400	1803200	60030	251.2 (607837)	17.6 (163574)	33.8 (22)	17.0 (41)	15.9 (38)
150	600	2601000	90030	431.5 (729617)	32.3 (249227)	49.3 (22)	28.8 (44)	24.3 (39)
150	800	3850400	120030	561.4 (476154)	51.8 (330616)	83.4 (24)	36.7 (41)	34.0 (38)
300	200	3587400	60060	682.9 (4310028)	28.3 (203780)	85.4 (16)	46.4 (43)	39.83 (38)
300	400	7022400	120060	> 3000 (1138428)	74.8 (413460)	190.7 (18)	97.9 (44)	88.10 (41)
300	600	11245200	180060	> 3000 (2410216)	122.5 (567590)	339.1 (19)	147.5 (42)	137.1 (40)
300	800	14862400	240060	> 3000 (1717037)	188.3 (779265)	478.1 (21)	197.69 (43)	178.4 (39)

Table 4.6: CPU time and iterations (into parenthesis) of BlockIP (using D^{-1} and Θ_0^{-1} preconditioner) and CPLEX available LP methods (Primal Simplex, Dual Simplex, Barrier Method) when solving MNFPNCs. All the nodes are conditioned to have an out-flow capacity, so that $l = n$.

It can be seen from tables 4.5 and 4.6 that the CPU time associated to the primal simplex is always far greater than the ones of the other solvers. For small instances the dual simplex is quite competitive and outperforms both the barrier method and the specialized interior point method. The network size n does not seem to have a substantial effect on the comparative efficiency of the five solvers. Instead the increase in the number of linking constraints (either $l = 0.2n$, in Table 4.5 or $l = 0.2n$, in Table 4.6) almost double the CPU time of the barrier method, whereas slightly affects the specialized interior point method.

Let us consider a modification of the MNFPNCs, obtained by introducing to each commodity a set of equal flow constraints, requiring that each arc in a specified set $\mathcal{R}_r$ must carry the same amount of flow: $x^i_{wj} = x^i_{kh}$, for $i \in \mathcal{C}$, $(i,j), (k,h) \in \mathcal{R}_r$, for every group of arcs $r \in R$.

This constraints arose while modeling some real-life problems, such as water resource system management [155]. We call this problem *multicommodity equal flow problem with nodal capacities* (MEFPNC from now on). Here we are considering MEFPNCs of different sizes with $R = 0.1n$ groups of arcs having the same flows per each commodity and $|\mathcal{R}_r| = 0.1n$ number of arcs in each group $r = 1 \dots R$.

Tables 4.7 and 4.8 report two computational experiments associated to these MEFPNCs.

n	k	n. var.	n. con.	CPLEX 12.5			BlockIP	
				Primal Simplex	Dual Simplex	Barrier	Θ_0^{-1}	D^{-1}
150	200	880000	30030	60.4 (1055443)	7.5 (108007)	10.8 (21)	8.0 (38)	7.7 (38)
150	400	1803200	60030	254.2 (530866)	18.3 (199104)	26.6 (22)	18.1 (40)	18.3 (40)
150	600	2601000	90030	413.4 (700019)	35.2 (331808)	46.4 (27)	28.4 (43)	26.5 (40)
150	800	3850400	120030	610.3 (685684)	52.6 (431006)	73.9 (29)	44.1 (48)	35.1 (41)
300	200	3587400	60060	549.1 (4710833)	40.9 (268745)	77.8 (31)	50.0 (41)	50.2 (41)
300	400	7022400	120060	> 3000.0 (2363994)	114.6 (557715)	177.7 (37)	102.1 (42)	109.0 (45)
300	600	11245200	180060	> 3000.0 (920761)	222.5 (945861)	289.6 (35)	171.8 (48)	173.13 (46)
300	800	14862400	240060	> 3000.0 (1007948)	343.6 (1337210)	415.1 (43)	199.3 (40)	206.2 (42)

Table 4.7: CPU time and iterations (into parenthesis) of BlockIP (using D^{-1} and Θ_0^{-1} preconditioner) and CPLEX available LP methods (Primal Simplex, Dual Simplex, Barrier Method) when solving MNFPNCs. Only 20% of nodes are conditioned to have an out-flow capacity, so that $l = 0.2n$.

n	k	n. var.	n. con.	CPLEX 12.5			BlockIP	
				Primal Simplex	Dual Simplex	Barrier	Θ_0^{-1}	D^{-1}
150	200	880000	30030	72.9 (956279)	8.4 (102700)	22.5 (27)	10.2 (41)	9.7 (41)
150	400	1803200	60030	265.4 (603381)	29.2 (243553)	36.8 (21)	21.4 (43)	19.3 (40)
150	600	2601000	90030	580.1 (944494)	49.9 (312280)	87.7 (30)	35.2 (43)	31.2 (39)
150	800	3850400	120030	792.2 (1044690)	80.9 (432939)	109.6 (29)	47.0 (44)	42.3 (41)
300	200	3587400	60060	836.3 (4908387)	55.4 (331366)	127.1 (21)	67.0 (47)	62.1 (45)
300	400	7022400	120060	> 3000.0 (3413000)	148.8 (679719)	415.5 (38)	141.8 (50)	127.2 (46)
300	600	11245200	180060	> 3000.0 (3913672)	230.9 (932892)	622.6 (37)	193.1 (46)	175.8 (43)
300	800	14862400	240060	> 3000.0 (392948)	364.9 (1278588)	667.0 (27)	286.2 (51)	239.1 (44)

Table 4.8: CPU time and iterations (into parenthesis) of BlockIP (using D^{-1} and Θ_0^{-1} preconditioner) and CPLEX available LP methods (Primal Simplex, Dual Simplex, Barrier Method) when solving MNFPNCs. All the nodes are conditioned to have an out-flow capacity, so that $l = n$.

The eight instances of Table 4.7 correspond to problems where only 20% of nodes are required to have a finite out-flow capacity, whereas the eight instances of Table 4.8 correspond to problems where all the nodes have a finite out-flow capacity.

The inclusion of equal flow constraints negatively affects the computational performance of all the considered solvers, though in different proportions. The dual simplex almost double its CPU times with respect to the MNFPNCs, whereas the ones associated to the specialized interior point method slightly increases. This is particularly true when medium and big instances are taken into account. In fact, for small number of commodities the dual simplex is quite competitive and outperforms the other solvers, but the specialized interior point method becomes the most efficient algorithm for this class of problems even for $k = 400$. Also for the MEFPNCs, as it was for the MNFPNCs, the network size n does not seem to have a substantial effect on the ranking of the five solvers.

The use of Θ_0^{-1} as a preconditioner results slightly more effective than D^{-1} in solving MEFPNCs with $l = 0.2n$, as shown in Table 4.7. Instead D^{-1} results to be slightly a better preconditioner when $l = n$. In both cases BlockIP outperforms the Cplex available LP methods.

The information of the average principal angles between the subspaces generated by the columns of L^T and N^T, for each instance of the analyzed multi-commodity network flow problems with nodal capacities, are reported in Table 4.9.

n	average principal angles			
	Table 4.5	Table 4.6	Table 4.7	Table 4.8
150	0.8774	0.8319	0.7958	0.8225
300	0.8544	0.8297	0.7732	0.8247

Table 4.9: Average principal angles between the subspaces generated by the columns of L^T and N^T, for each instance of multi-commodity network flow problems with nodal capacities in tables 4.5, 4.6, 4.7 and 4.8.

The first information we obtain from Table 4.9 is that the average principal angles are generally stable with respect to the number of nodes. Beyond that, we can observed a clear correspondence between the good computational performance of BlockIP with Θ_0 as a preconditioner and the smaller average principal angles associated to the instances of Table 4.7.

4.6 Edge-colored network problems

In this section computational results for the edge-colored network problems studied in Subsection 3.2.1 are provided, by applying the specialized interior point method to the primal-block angular problems in the form of (3.4). The importance of these results is due to the possibility of using interior point methods for the problem of random network generation described in Chapter 3, allowing a formal derivation of relevant probabilistic properties, as discussed next in Chapter 5.

Consider the following two network models:

- the edge-colored networks conditioned to the within-color densities in (3.4), where $L_i = I$ and $N_i = \mathbf{e}^T$, for $i = 1, \ldots, k$;
- the undirected edge-colored network where edges are allowed to have multiple colors and the number of edges per color is fixed, as well as the between color degree sequence. This represents a spacial case of (3.4), with linking constraints (3.20).

As discussed in Chapter 3, most of the primal block-angular LPs resulting from families of networks have the particular advantage of being associated to row vector blocks, i.e. $N_i \in \mathbb{R}^{1\times n_i}$.

This is also true for the two aforementioned network problems. If we let $\mathrm{Tr}(M)$ denote the trace of matrix M, for both models we have

$$B = \begin{bmatrix} \mathrm{Tr}(\Theta_1) & & \\ & \ddots & \\ & & \mathrm{Tr}(\Theta_C) \end{bmatrix},$$

A numerical comparison of the the dual simplex, the barrier method and the BlockIP, when applied to the aforementioned LPs are reported in tables 4.10 and 4.11. Columns n and k show the number of nodes and colors of the networks. Columns "n. var." and "n. constr." give the number of variables and constraints of the resulting LP problems. Note that the largest case has more than 45 million variables, and 900 constraints. The remaining columns give the CPU time and number of iterations (in parentheses) for the three algorithms tested: Cplex 12.5 dual simplex, Cplex 12.5 barrier (interior point), and the specialized interior point method of BlockIP.

n	k	n. var.	n. con.	CPLEX 12.5			BlockIP	
				Primal Simplex	Dual Simplex	Barrier	Θ_0^{-1}	D^{-1}
50	50	61250	100	0.3 (77502)	0.9 (907)	0.4 (18)	0.8 (28)	0.2 (28)
150	50	558750	200	1.0 (403039)	0.2 (1080)	1.5 (17)	3.5 (44)	1.3 (44)
450	50	5051250	500	3.2 (910328)	0.9 (1272)	5.7 (15)	10.1 (42)	3.2 (42)
50	150	183750	200	23.5 (594661)	1.5 (7106)	4.0 (14)	8.7 (27)	2.7 (27)
150	150	1676250	300	45.9 (5702201)	4.7 (8231)	19.4 (20)	54.2 (49)	11.5 (47)
450	150	15153750	600	104.6 (7724554)	16.2 (9481)	81.2 (17)	131.6 (53)	42.7 (53)
50	450	551250	500	1597.2 (5512088)	27.0 (64394)	45.0 (16)	281.2 (40)	67.9 (31)
150	450	5028750	600	6830.0 (20014384)	83.3 (70635)	184.1 (18)	534.6 (53)	140.7 (49)
450	450	45461250	900	18011.0 (38380791)	294.2 (73656)	1047.3 (24)	673.6 (92)	436.0 (91)

Table 4.10: CPU time and iterations (in parentheses) of four LP algorithms for LPs associated to the edge-colored networks conditioned to the within-color densities.

n	k	n. var.	n. con.	CPLEX 12.5			BlockIP	
				Primal Simplex	Dual Simplex	Barrier	Θ_0^{-1}	D^{-1}
50	50	61250	100	1.4 (79089)	0.7 (733)	0.3 (13)	0.4 (33)	0.3 (34)
150	50	558750	200	7.9 (275329)	5.0 (1435)	1.0 (14)	1.3 (41)	1.0 (40)
450	50	5051250	500	46.0 (865750)	5.2 (1956)	5.6 (15)	4.2 (44)	3.8 (45)
50	150	183750	200	35.0 (619349)	6.3 (934)	6.3 (26)	5.5 (42)	3.9 (43)
150	150	1676250	300	198.1 (2280889)	57.7 (2503)	23.5 (30)	16.9 (52)	12.4 (52)
450	150	15153750	600	1409.5 (8115972)	377.1 (4822)	101.1 (35)	53.2 (52)	38.9 (51)
50	450	551250	500	766.2 (6037183)	124.3 (2462)	110.3 (46)	80.2 (55)	45.7 (56)
150	450	5028750	600	4967.4 (17541823)	1149.8 (5243)	479.9 (53)	225.3 (61)	146.9 (58)
450	450	45461250	900	17967.4 (59881679)	10926.3 (13047)	1287.4 (52)	624.1 (62)	419.6 (54)

Table 4.11: CPU time and iterations (in parentheses) of four LP algorithms for LPs associated to the edge-colored networks conditioned to the within-color densities and between-color degree sequence.

From Table 4.11, the simplex method is clearly outperformed by the barrier algorithm, and the gap increases with the size of the instance. BlockIP, the specialized interior point algorithm, was two to three times faster than the Cplex barrier in the largest instances, resulting to be the most efficient approach for this kind of problems. By contrast, the dual simplex appears to be the most efficient method for the LPs associated to the edge-colored networks conditioned to the within-color densities, as shown in Table 4.10.

The CPU times shown in Tables 4.11 seem to agree with our discussion concerning the high degeneracy of basic solutions, as suggested by the poor performance of the primal simplex. When the number of the $kn(n-1)/2$ decision variables increases the dual simplex become highly inefficient, whereas the CPU times of the two interior point methods grow slowly.

When LP associated edge-colored network with fixed within color density and between color degree sequence are taken into account, the BlockIP results faster than the CPLEX available LP methods even for small problem sizes, as shown in Table 4.10. Instead, when $L_i = I$ and $N_i = \mathbf{e}^T$, for $i = 1, \ldots, k$, the dual simplex appears surprisedly faster than the others.

The information of the average principal angles between the subspaces generated by the columns of L^T and N^T, for each instance of the analyzed LPs associated to the edge-colored networks conditioned to the within-color densities, are reported in Table 4.12.

n	average principal angles	
	Table 4.11	Table 4.10
50	0	0
150	0	0
450	0	0

Table 4.12: Principal angles between the subspaces generated by each column of L^T and N^T, for different sizes of edge-colored networks in tables 4.11 and 4.10.

As previously mentioned in Subsection 4.3.1, a downside of the use of principal angles between $\mathcal{R}(\Theta_i^{1/2} L_i^T)$ and $\mathcal{R}(\Theta_i^{1/2} N_i^T)$ to measure the goodness of the approximation of Θ_0^{-1} and D^{-1} to $(D - C^T B^{-1} C)^{-1}$ appears when $\mathcal{R}(\Theta_i^{1/2} L_i^T) = \mathbb{R}^{n_i}$, for $i = 1, \ldots, k$. In such case the principal angles are always zeros, as in the case of the edge-colored network problems with fixed within color densities shown in Table 4.12. Hereby a much better measure of the goodness of the two preconditioners are the average principal angles between $\mathcal{R}(\Theta_i^{1/2} N_i^T)$ and $\mathcal{R}(\Theta_i^{1/2} \ell_{ij}^T)$, where ℓ_{ij} is the j^{th} row of matrix L_i.

n	average principal angles	
	Table 4.11	Table 4.10
50	1.5422	1.3694
150	1.5613	1.4550
450	1.5677	1.5041

Table 4.13: Average principal angles between the subspaces generated by each column of L^T and all the columns of N^T, for different sizes of of edge-colored networks in tables 4.11 and 4.10.

Table 4.13 shown that the average principal angles between $\mathcal{R}(N_i^T)$ and $\mathcal{R}(\ell_{ij}^T)$ are very close to orthogonality and this fact is quite coherent with the computational results of tables 4.7 and 4.8, where D^{-1} results as a much better preconditioner.

Note that the advantages of looking at the principal angles between $\mathcal{R}(N_i^T)$ and $\mathcal{R}(\ell_{ij}^T)$ is that the singular value of a column vector is the norm of the vector itself and the unique source of computational effort is to obtain an orthonormal bases for $\mathcal{R}(N_i^T)$.

Chapter 5

Efficiency and correctness of network simulation

5.1 Purposes and preliminary overview

Chapters 3 and 5 considered algebraic characterizations of different families of networks by means of systems of linear constraints and introduced a specialized interior-point approach to deal with this class of problems. Numerical results supported the suitability of these computational procedures for our specific application, but no insight about the statistical properties of those randomized methods has been mentioned so far.

This chapter tries to fill this gap by focusing on two prominent random models which have been particularly relevant in the analysis of social networks: i) conditionally uniform random models, ii) conditionally exponential random models. Sufficient conditions for the mathematical programming based approaches of chapters 3 and 5 to fit the probabilistic properties of these ensembles are hereby taken into account.

To set a brief introduction to this topics we should place ourself in the context of networks whose vertices or edges are generated by some random process, so that different random mechanisms of vertex and edge selection leads to different probability distributions on networks. From a statistical point of view, both binary and valued networks are random matrices with complicated patterns of dependence. Whereas statistical modeling is generally based on assumptions of independence, random networks make difficult the development of statistical models for these high dimensional systems with complicated dependence patterns.

The study of random graphs begins with the seminal work of P. Erdös and A. Rainyi [86], who considered a fixed set of nodes and an independent and equal probability of observing edges among them. There are two closely related variants of the Erdös-Rainyi model:

- *the* $G(n,p)$ *model*, where a network is constructed by connecting nodes randomly with independent probability p;
- *the* $G(n,m)$ *model*, where a network is chosen uniformly at random from the collection of all graphs with n nodes and m edges.

Both models possess the considerable advantage of being exactly solvable for many of their average properties: clustering coefficient, average path length, giant component, etc. (For more details about network properties, see Bollobas [29], and Wasserman and Faust [224].) In other

words, the expectation of many structural properties of networks generated by the Erdös-Rainyi processes is analytically obtainable.

The $G(n, m)$ model poorly fits the most relevant structural properties of real-world networks. This fact has been carefully analyzed by Newman [171], who reported theoretical and empirical results for several observed networks, as shown in Table 5.1. The disagreement between the clustering coefficient of the real-world networks and the one expected under the Erdös–Rainyi model is quite evident for all the analyzed data sets.

	clustering coefficient	
network	measured	random graph
Internet	0.240	0.000600
World-wide web	0.110	0.000230
Power grid	0.080	0.000540
Biology collaborations	0.081	0.000010
Math collaborations	0.150	0.000015
Actor collaborations	0.200	0.000250
Company directors	0.590	0.001900
Word co-occurrence	0.440	0.000150
Neural network	0.280	0.049000
Metabolic network	0.590	0.090000
Food web	0.220	0.065000

Table 5.1: Clustering coefficient of observed networks reported by Newman [171].

Based on this lack of fit, other conditionally uniform random model might be taken into account, before switching the analysis to non-uniform models. Nonetheless, when the conditioning information is not necessarily the number of edges but whatever arbitrary network feature, very few analytical results are available and simulation is in order to obtain empirical distributions of average network properties. This difficulty in obtaining closed-form solutions of statistical properties of interest represents indeed the main motivation of the research proposed in chapters 3 and 5 and will be further handled in this chapter.

Conditionally uniform models have been studied in the second half of the twenty century [216] [222] [194], but seems to have fallen into disuse by the last few years. We intend to provide a general methodological framework to generate networks with constraints, representing structural features the researcher wishes to control for.

5.2 Conditionally uniform random networks

Conditional uniform models can be seen as a generalization of the $G(n, m)$ model, when the conditioning information is not necessarily the number of edges but whatever other arbitrary network property. Unfortunately, in this case we have very few analytical results and simulation is required to obtain empirical distributions of their average properties.

Suppose we want to check whether the observed average path length is *ordinary* for the type of network we would like to consider, for example networks with fixed structural features, such as density, degree sequence, number of mutual links, etc. Then we might generate a bunch of networks uniformly at random from all networks having the same specified structural features and count how often an average path length at least as extreme as ours occurs. In practice these

types of tests are the most used tests in network analysis. They are called *conditional uniform graph tests*.

Drawing uniformly at random from complex networks is not as easy as it sounds, as numerical procedures may not explore the whole sample space with the same probability. Although other Operations Research tools have been used in the context of social networks [14], as far as we know, this work is the first attempt to use integer optimization for the generation of several classes of conditional random graphs. Previous approaches, developed within the fields of *mathematical and computational sociology*, were ad hoc procedures for some particular types of networks, in general difficult to generalize and not very efficient.

Holland and Leinhardt [118] first introduced an extension of $G(n, m)$, conditioning the set of directed networks to the dyad counts (i.e., group of two nodes/people). Shortly afterward Karlberg [134] compared the transitivity of observed networks to the one obtained under the uniform distribution of all networks conditioned to the node degrees, allowing to find out whether there is a tendency toward transitivity in the observed networks.

Likewise, the distribution of all networks conditioned to the nodes in- and out-degree have difficult combinatorial properties, as its analytical study involves binary matrices with fixed marginal rows and marginal columns. In this respect, some combinatorial results have been obtained by Ryser [202], who derived necessary and sufficient conditions for two vectors of non-negative integers to constitute the row sums and column sums of some zero-one matrix. On the other hand, ways to simulate uniform random networks with given degree distribution were developed by Snijders [216], Rao [194], Roberts [203], and Verhelst [222].

Snijders [216] considered a complete order $\prec$ in the $n \times n$ cells of the AM. The vector of entries of all cells preceding a given cell (i, j) is called the initial segment of that cell and has the form $(x_{i'j'} : (i', j') < (i, j))$. When considering a given initial segment of a given cell (i, j), either the element X_{ij} is completely determined by the initial segment, or both values 0 and 1 are possible; in these two cases, respectively, the cell is either determined or free.

The enumeration tree for the family of digraphs with 4 nodes and uniform in-degree and out-degree distribution equal to 1, as reported by Snijders [216], is shown below in Figure 5.1.

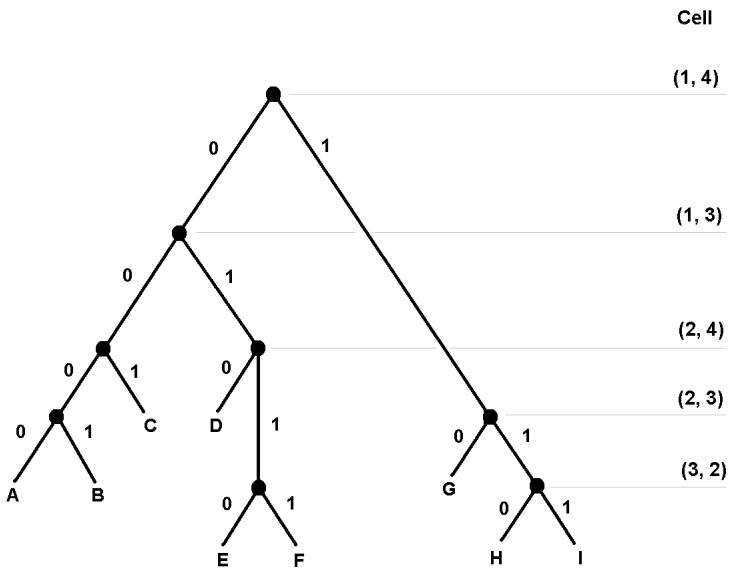


Figure 5.1: Enumeration tree

The cells of the AM are ordered from top to bottom and from right to left, while the

diagonal cells are omitted; thus, the ordering is $(1,4) \prec (1,3) \prec (1,2) \prec (2,4) \prec \ldots \prec (4,1)$. A lot of computation is necessary to determine whether the next cell is determined or free, relative to the initial segment. Moreover, given that enumeration is intractable for larger AM, Snijders [216] proposed a Monte Carlo simulation approach to simulate independent rows such that for each row, its corresponding number of ones is places at random in the admissible cells. An implementation of these methods written in Turbo Pascal, called ZO, is available on-line (http://stat.gamma.rug.nl/stocnet/).

Few years later Rao [194] proposed a Markov Chain Monte Carlo approach to generate AM with given in-degrees and out-degrees. His procedure is based on the iterative selection of two rows i_1 and i_2 and two columns j_1 and j_2 of the AM. A random rectangle S defined by the intersection of these rows and columns. Let X_t and S_t be respectively the AM and the selected random rectangle at iteration t. Then the movement described in (5.1) is applied.

$$\text{if } S_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ then replace it with } \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{5.1a}$$

$$\text{if } S_t = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ then replace it with } \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{5.1b}$$

$$otherwise\, X_{t+1} = X_t \tag{5.1c}$$

In most cases, it will be possible to pass from any feasible network to another by a sequence of such movements on different submatrices S. However, as proved by Rao [194], updating rectangles in this manner is not sufficient to ensure that the Markov chain is irreducible. He showed that this problem can be bypassed considering at each iteration t, also submatrices with identical row and column indices, call them C_t, and applying the following movement:

$$\text{if } C_t = \begin{bmatrix} - & 1 & 0 \\ 0 & - & 1 \\ 1 & 0 & - \end{bmatrix} \text{ then replace it with } \begin{bmatrix} - & 0 & 1 \\ 1 & - & 0 \\ 0 & 1 & - \end{bmatrix} \tag{5.2a}$$

$$\text{if } C_t = \begin{bmatrix} - & 0 & 1 \\ 1 & - & 0 \\ 0 & 1 & - \end{bmatrix} \text{ then replace it with } \begin{bmatrix} - & 1 & 0 \\ 0 & - & 1 \\ 1 & 0 & - \end{bmatrix} \tag{5.2b}$$

$$otherwise\, X_{t+1} = X_t \tag{5.2c}$$

Note that in any directed network, the former movements keep untouched both the degrees of nodes and the number of mutual choices. That is why this procedure has been used to simulate both the uniform random network conditional to the degree sequence and the uniform random network conditioned to the number of mutual connections. However, as Robert [203] showed, a particular disadvantage of this procedure is the particularly high rate of rejections.

In the case of undirected networks an easy way to generate instances with fixed degree sequence is as follows:

1. give to each node i a number d_i of edge-ends emerging from it;
2. choose pairs of these edge-ends belonging to different nodes uniformly at random and join them together
3. when all edge-ends have been used up, the resulting network is a random member of the ensemble of network with the desired degree sequence.

As already mentioned in the previous chapter, not every sequence of positive integers is a degree sequence of an undirected network. Let $f_1, \ldots, f_n$, be a sequence of n non-negative integers verifying $\sum_{i=1}^{k} f_i \leq k(k-1) + \sum_{i=k+1}^{n} \min(f_i, k)$, for all $k = 1, \ldots, n$. We say that $f_1, \ldots, f_n$ is a degree sequence of a simple graph with n nodes if it represents the row (or column) sums of the associated adjacency matrix. (For more details see Erdos [86].)

The particular importance of the uniform random network model conditioned to the degree sequence is mainly due to the fact that the expectation of the clustering coefficient and the characteristic path length are analytically solvable for uniform undirected network with given degree distribution can be obtained without any computational effort. (For more details about network properties, such as the clustering coefficient and the characteristic path length, see Bollobas [29] and Wasserman and Faust [224].) Newman [171] derived a closed form for the clustering coefficient:

$$C == \frac{E[k]}{n}\left[\frac{E[k^2] - E[k]^2}{E[k]^2}\right]^2 = \frac{E[k]}{n}\left[c_v^2 + \frac{E[k]-1}{E[k]}\right]^2 \tag{5.3}$$

where $E[k]$ is the expected degree and c_v^2 the squared of its coefficient of variation. Expression 5.3 shows that if the expected degree does not depend on the network size, the clustering will end up to zero, no matter the degree distribution we are conditioning on. This fact might be interpreted as a lack of information on the structure of the network when its size increases, i.e. the degree distribution does not provide so much information on the structure of the network when its size is very big.

The need of adding more condition comes exactly from the attempt of giving a better characterization of the network structure. Robert [203] studied a conditionally uniform model where vertices of the network are viewed as belonging to one of two sets, γ_1 and γ_2. One might wish to know the distribution of a particular network parameter, when the number of ties between members of γ_1 and γ_2 and within members of γ_1 and γ_1 are kept constant. The mathematical programming model associated to this family of networks has been described in Proposition 5.

In practice one would like to go even further in conditioning, which however leads to self-defeating attempts because of combinatorial complexity. That is why a pressing theoretical need is to prove irreducibility of Markov Chains in the other situations beyond the well known models introduced in this section. However, if the Markov chains in these cases are not irreducible, it would be useful to identify the conditions under which they are reducible, as it still might be possible to use the approach in most realistic situations.

5.3 The probability of networks as primal-dual solutions

This section provides analytical results concerning the probabilities of networks generated in accordance with the mathematical programming based methods introduced in chapter 3. The core idea lies on the probabilistic relationship between the set of networks with given topological properties and the random generation of parameters of the associated LPs.

Let $CR(\chi) = \{\mathbf{x} \in [0,1]^N : A\mathbf{x} = \mathbf{b}\}$ be a polytope whose set of extreme points is bijectively related to a given family of networks and consider the associated (feasible and bounded) LP

$$\min \mathbf{c}^T\mathbf{x} \text{ s. to } A\mathbf{x} = \mathbf{b},\ 0 \leq \mathbf{x} \leq 1, \tag{5.4}$$

where $A \in \mathbb{R}^{M\times N}$ is a full row rank matrix, $\mathbf{b} \in \mathbb{R}^M$ and $\mathbf{c} \in \mathbb{R}^N$. By adding slacks, (5.4) can be written in standard form as

$$\min \widehat{\mathbf{c}}^T\widehat{\mathbf{x}} \text{ s. to } \widehat{A}\widehat{\mathbf{x}} = \widehat{\mathbf{b}},\ \widehat{\mathbf{x}} \geq 0, \tag{5.5}$$

where $\widehat{A} \in \mathbb{R}^{\widehat{M}\times\widehat{N}}$ ($\widehat{M} = M + N$ and $\widehat{N} = 2N$) is a full row rank matrix, $\widehat{\mathbf{b}} = [\mathbf{b}^T \quad \mathbf{e}^T]^T \in \mathbb{R}^{\widehat{M}}$ and $\widehat{\mathbf{c}} = [\mathbf{c}^T \quad \mathbf{0}^T]^T \in \mathbb{R}^{\widehat{N}}$. If the gradient of the objective function $\widehat{\mathbf{c}}$ is a properly defined random vector of density function $f_C(\widehat{\mathbf{c}})$, then the solution of (5.5) is also a random vector whose probability distribution can be computed as

$$P(\widehat{\mathbf{x}}) = \int_{\Omega} f_C(\widehat{\mathbf{c}})\, d\widehat{\mathbf{c}}, \tag{5.6}$$

where Ω is the set of gradients for which $\widehat{\mathbf{x}}$ is an optimal solution.

The gradient and optimal solutions are related through the KKT conditions of the LP, which can be written as

$$G(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) = \begin{bmatrix} \mathbf{0} \\ \widehat{\mathbf{c}} \\ \mathbf{0} \end{bmatrix} \quad \text{where} \quad G(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) \triangleq \begin{bmatrix} \widehat{A}\widehat{\mathbf{x}} - \widehat{\mathbf{b}} \\ \widehat{A}^T\widehat{\mathbf{y}} + \widehat{\mathbf{z}} \\ \widehat{X}\widehat{Z}\mathbf{e} \end{bmatrix} \quad \text{and} \quad (\widehat{\mathbf{x}}, \widehat{\mathbf{z}}) \geq 0, \tag{5.7}$$

$\widehat{\mathbf{y}}$ and $\widehat{\mathbf{z}}$ being respectively the Lagrangean multipliers of the equations and bounds of (5.5), and $\widehat{X}$ and $\widehat{Z}$ diagonal matrices built up with the components of $\widehat{\mathbf{x}}$ and $\widehat{\mathbf{z}}$. The vector function G, as defined in (5.7), is not injective, since (5.5) may have more than one solution (indeed, an infinite number of them) for some $\widehat{\mathbf{c}}$, that is, G may map different points into the same vector $[\mathbf{0}^T\ \widehat{\mathbf{c}}^T\ \mathbf{0}^T]^T$. However, it is possible to guarantee the bijectivity of G by restricting its domain to the set $\mathcal{I}_D = \{\mathbf{t} = (\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) : \widehat{\mathbf{x}} > 0, \widehat{\mathbf{z}} > 0, \widehat{A}\widehat{\mathbf{x}} = \widehat{\mathbf{b}}, \widehat{A}^T\widehat{\mathbf{y}} + \widehat{\mathbf{z}} = \widehat{\mathbf{c}}, \widehat{X}\widehat{Z}\mathbf{e} = \mu\mathbf{e}\}$ for some $\mu \in \mathbb{R}$, $\mu > 0$, and considering the KKT-μ perturbed conditions

$$G(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) = \begin{bmatrix} \mathbf{0} \\ \widehat{\mathbf{c}} \\ \mu\mathbf{e} \end{bmatrix}. \tag{5.8}$$

The codomain of G is thus $\mathcal{I}_C = \{\mathbf{s} = (\mathbf{0}, \widehat{\mathbf{c}}, \mu\mathbf{e}) : \mu \in \mathbb{R}, \mu > 0, \widehat{\mathbf{c}} \in \mathbb{R}^{\widehat{N}}\}$. For a fixed $\widehat{\mathbf{c}}$, the set of solutions $(\widehat{\mathbf{x}}_\mu, \widehat{\mathbf{y}}_\mu, \widehat{\mathbf{z}}_\mu)$ of (5.8) for $\mu > 0$ is an arc of feasible points known as the primal-dual central path [231, 233], which is widely used in interior-point methods. When $\mu \to 0$, the central path converges to an optimal solution of (5.5). If instead of a unique solution we have an optimal face, then the central path converges to the analytic center of this optimal face [233]. If the strictly feasible set of (5.5) is nonempty (i.e., $\mathcal{F}^0 = \{(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) : (\widehat{\mathbf{x}}, \widehat{\mathbf{z}}) > 0, \widehat{A}\widehat{\mathbf{x}} = \widehat{\mathbf{b}}, \widehat{A}^T\widehat{\mathbf{y}} + \widehat{\mathbf{z}} = \widehat{\mathbf{c}}\} \neq \emptyset$), then the central path exists and it is unique for each $\mu > 0$ (see [231, 233] for a proof of this result). This uniqueness guarantees that, for some $\widehat{\mathbf{c}}$ and $\mu > 0$, there is a single point $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}})$ satisfying (5.8).

The bijectivity of $G : \mathcal{I}_D \to \mathcal{I}_C$ allows a straightforward application of the theorem of change of variables for multidimensional integrals [97], in order to deduce a probability density function of $G^{-1}(\mathbf{s})$, for $\mathbf{s} \in \mathcal{I}_C$:

Lemma 1. *Let $G : \mathcal{I}_D \longrightarrow \mathcal{I}_C$ be a one-to-one and continuously differentiable map of the open set $\mathcal{I}_D$ into $\mathcal{I}_C$, such that $J(G(\mathbf{t}))$ is nonsingular for all $\mathbf{t} \in \mathcal{I}_D$. If $f : \mathcal{I}_C \longrightarrow \mathbb{R}$ is a nonnegative locally integrable function, then*

$$\int_{\mathcal{I}_D} f(G(\mathbf{t}))||J(G(\mathbf{t}))||d\mathbf{t} = \int_{\mathcal{I}_C} f(\mathbf{s})d\mathbf{s}, \tag{5.9}$$

where the symbol $||\ ||$ is used to denote the absolute value of the determinant and $\mathbf{t} = (\widehat{\mathbf{x}}\ \widehat{\mathbf{y}}\ \widehat{\mathbf{z}})$ and $\mathbf{s} = (\widehat{\mathbf{b}}\ \widehat{\mathbf{c}}\ \boldsymbol{\mu})$. By the inverse-function theorem, $\mathcal{I}_C$ is open and the inverse point mapping $G^{-1}(\mathbf{s})$ is continuously differentiable.

Lemma 1 requires

$$J(G(\widehat{\mathbf{x}},\widehat{\mathbf{y}},\widehat{\mathbf{z}})) = \begin{bmatrix} \widehat{A} & & \\ & \widehat{A}^T & I \\ \widehat{Z} & & \widehat{X} \end{bmatrix} \tag{5.10}$$

to be nonsingular. This is the matrix of the Newton system to be solved at each iteration of primal-dual interior-point methods, which is known to be non singular if $\hat{A}$ is full row rank and $(\widehat{\mathbf{x}},\widehat{\mathbf{z}}) > 0$ [231].

Considering the probability density functions $f_{\mathcal{I}_D} : \mathcal{I}_D \to \mathbb{R}$ and $f_{\mathcal{I}_C} : \mathcal{I}_C \to \mathbb{R}$ and applying Lemma 1 to $G : \mathcal{I}_D \to \mathcal{I}_C$, we obtain that for every open subset $\mathcal{D} \subseteq \mathcal{I}_D$ and $G(\mathcal{D}) \subseteq \mathcal{I}_C$ (the image set of $\mathcal{D}$ under the transformation G) the probability that $(\widehat{\mathbf{x}}\ \widehat{\mathbf{y}}\ \widehat{\mathbf{z}}) \in \mathcal{D}$ is equal to the probability that $(\mathbf{0}\ \widehat{\mathbf{c}}\ \mu\widehat{\mathbf{e}}) \in G(\mathcal{D})$ is

$$\int_{\mathcal{D}} f_{\mathcal{I}_D}(\mathbf{t})d\mathbf{t} \;=\; \int_{G(\mathcal{D})} f_{\mathcal{I}_C}(\mathbf{s})d\mathbf{s} = \int_{\mathcal{D}} f_{\mathcal{I}_C}(G(\mathbf{t}))||J(G(\mathbf{t}))||d\mathbf{t}. \tag{5.11}$$

Since (5.11) is true for all open subset $\mathcal{D} \subseteq \mathcal{I}_D$, we have

$$f_{\mathcal{I}_C}(\mathbf{s}) \;=\; \frac{f_{\mathcal{I}_D}(\mathbf{t})}{||J(G(\mathbf{t}))||}\bigg|_{\mathbf{t}=G^{-1}(\mathbf{s})} \tag{5.12}$$

and

$$f_{\mathcal{I}_D}(\mathbf{t}) = f_{\mathcal{I}_C}(G(\mathbf{t}))||J(G(\mathbf{t}))|| \;=\; f_{\mathcal{I}_C}\left(\begin{bmatrix} \widehat{A}\widehat{\mathbf{x}} \\ \widehat{A}^T\widehat{\mathbf{y}} + \widehat{\mathbf{z}} \\ \widehat{X}\widehat{Z}\mathbf{e} \end{bmatrix}\right)\left\|\begin{matrix} \widehat{A} & & \\ & \widehat{A}^T & I \\ \widehat{Z} & & \widehat{X} \end{matrix}\right\|. \tag{5.13}$$

Let μ be a positive number arbitrarily close to zero and f_C a $\widehat{N}$-dimensional probability density function. We assume $f_{\mathcal{I}_C}$ to have the following form:

$$f_{\mathcal{I}_C}\left(\begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{bmatrix}\right) = \begin{cases} f_C(\mathbf{r}_2) & \text{if } \mathbf{r}_1 = \mathbf{0} \text{ and } \mathbf{r}_3 = \mu\mathbf{e} \\ 0 & \text{otherwise} \end{cases} \tag{5.14}$$

Therefore, under the transformation G we have

$$f_{\mathcal{I}_D}\left(\begin{bmatrix} \widehat{\mathbf{x}} \\ \widehat{\mathbf{y}} \\ \widehat{\mathbf{z}} \end{bmatrix}\right) = \begin{cases} f_C\big(\widehat{A}^T\widehat{\mathbf{y}} + \widehat{\mathbf{z}}\big)\left\|\begin{matrix} \widehat{A} & & \\ & \widehat{A}^T & I \\ \widehat{Z} & & \widehat{X} \end{matrix}\right\| & \text{if } \widehat{A}\widehat{\mathbf{x}} = \widehat{\mathbf{b}} \text{ and } \widehat{X}\widehat{Z}\mathbf{e} = \mu\mathbf{e} \\ 0 & \text{otherwise} \end{cases} \tag{5.15}$$

Note that the support of $f_{\mathcal{I}_D}$ is $\mathcal{I}_D$, so that it only provides the probability of central path points for some particular μ. As $\mu \to 0$ the central path points converge to the solution of (5.5), and then (5.15) would approach the probability distribution of the solutions of the LP problem in terms of the probability distribution of the cost vector.

Theorem 5. *The limit of (5.15) when $\mu \to 0$ exists, that is,*

$$\lim_{\mu\to 0} f_{\mathcal{I}_D}(\widehat{\mathbf{x}}(\mu),\widehat{\mathbf{y}}(\mu),\widehat{\mathbf{z}}(\mu)) = f_{\mathcal{I}_D}(\widehat{\mathbf{x}}(0),\widehat{\mathbf{y}}(0),\widehat{\mathbf{z}}(0)). \tag{5.16}$$

Proof. By [233, Theorem 2.17], the points $(\widehat{\mathbf{x}}(\mu), \widehat{\mathbf{y}}(\mu), \widehat{\mathbf{z}}(\mu))$ on the central path are bounded, the central path converges to $(\widehat{\mathbf{x}}(0), \widehat{\mathbf{y}}(0), \widehat{\mathbf{z}}(0))$, and $\widehat{\mathbf{x}}(0)$ and $\widehat{\mathbf{z}}(0)$ are, respectively, the analytic centers of the primal and dual optimal faces (i.e., for any $\widehat{\mathbf{c}}$ the central path converges to a unique point—the analytic center of the optimal face—even if there are multiple solutions for this $\widehat{\mathbf{c}}$). Therefore, the determinant of (5.15) computed at $\hat{Z}(0)$ and $\hat{X}(0)$ is bounded, and $\lim_{\mu\to 0} f_{\mathcal{I}_D}(\widehat{\mathbf{x}}(\mu), \widehat{\mathbf{y}}(\mu), \widehat{\mathbf{z}}(\mu))$ exists by continuity of the determinant. □

The above results are illustrated by this small example.

Example 2. *Consider the small two-dimensional problem* $\min c_1x_1 + c_2x_2$ *s. to* $x_1 + Mx_2 = 1, (x_1, x_2) \geq 0$ *($M > 0$ being a given parameter), which matches (5.5) for* $\widehat{A} = [1\ M]$ *and* $\widehat{b} = 1$. *The feasible region of this problem is the segment between* $V_A = (1, 0)$ *and* $V_B = (0, 1/M)$. *When* $c_2 > Mc_1$, V_A *is the optimal extreme point; when* $c_2 < Mc_1$, *the solution is* V_B; *when* $c_2 = Mc_1$ *(which is unlikely if the cost vector is randomly generated) the whole segment is optimal. Unless* $M = 1$, *the probability of* V_A *and* V_B *being optimal is not uniform, even if the cost vector is randomly generated (in particular we have* $P(V_A) = 1 - M/2$ *and* $P(V_B) = M/2$*).*

The associated KKT-μ *perturbed conditions are*

$$\begin{aligned} x_1 + Mx_2 &= 1 \\ y + z_1 &= c_1 \\ My + z_2 &= c_2 \\ x_1z_1 &= \mu \\ x_2z_2 &= \mu \\ (x_1, x_2, z_1, z_2) &> 0 \end{aligned} \tag{5.17}$$

and, after a few manipulations from the dual feasibility and complementarity conditions, we obtain:

$$Mc_1 + \frac{M\mu}{1 - x_1} = c_2 + \frac{M\mu}{x_1} \qquad \text{and} \qquad Mc_1 + \frac{\mu}{x_2} = c_2 + \frac{M\mu}{1 - Mx_2}. \tag{5.18}$$

When $c2 \neq Mc_1$, *the central path is obtained by solving the two quadratic equations (5.18) with respect to* $x_1(\mu)$ *and* $x_2(\mu)$*:*

$$\begin{aligned} &x_1(\mu) = \frac{c_2 - Mc_1 - 2M\mu + \sqrt{\Delta}}{2(c_2 - Mc_1)},\ x_2(\mu) = \frac{c_2 - Mc_1 + 2M\mu - \sqrt{\Delta}}{2M(c_2 - Mc_1)}, \\ &\Delta = 4M^2\mu^2 + (c_2 - Mc_1)^2, \qquad z_1(\mu) = \frac{\mu}{x_1(\mu)}, \quad z_2(\mu) = \frac{\mu}{x_2(\mu)}, \qquad y(\mu) = c_1 - z_1(\mu). \end{aligned} \tag{5.19}$$

When $c2 = Mc_1$, *we directly have*

$$x_1(\mu) = \frac{1}{2}, \quad x_2(\mu) = \frac{1}{2M}, \quad z_1(\mu) = 2\mu, \quad z_1(\mu) = 2M\mu, \quad y(\mu) = c_1 - 2\mu. \tag{5.20}$$

For $c2 \neq Mc_1$, *the limit point of the central path is*

$$\begin{aligned} \lim_{\mu\to 0} x_1(\mu) &= 1 \\ \lim_{\mu\to 0} x_2(\mu) &= 0 \\ \lim_{\mu\to 0} z_1(\mu) &= 0 \\ \lim_{\mu\to 0} z_2(\mu) &= c_2 - Mc_1 \\ \lim_{\mu\to 0} y(\mu) &= c_1 \end{aligned} \quad \textit{for } c_2 > Mc_1, \qquad \textit{and} \qquad \begin{aligned} \lim_{\mu\to 0} x_1(\mu) &= 0 \\ \lim_{\mu\to 0} x_2(\mu) &= \frac{1}{M} \\ \lim_{\mu\to 0} z_1(\mu) &= \frac{Mc_1 - c_2}{M} \\ \lim_{\mu\to 0} z_2(\mu) &= 0 \\ \lim_{\mu\to 0} y(\mu) &= \frac{c_2}{M} \end{aligned} \quad \textit{for } c_2 < Mc_1.$$

For $c2 \neq Mc_1$, *the primal solution* $(x_1(\mu), x_2(\mu))$ *does not depend on* μ *and it provides the analytic center of the feasible primal segment, as expected. The dual limit point is* $z_1(0) = z_2(0) = 0$, $y(0) = c_1$.

By (5.15), the asymptotic value of the density function of primal-dual solutions is

$$f_{\mathcal{I}_D}(x_1(\mu), x_2(\mu), y(\mu), z_1(\mu), z_2(\mu)) = f_{I_C}(c_1, c_2)|\det(J)|, \tag{5.21}$$

where in this simple problem $\det(J) = M^2 x_2(\mu) z_1(\mu) + x_1(\mu) z_2(\mu)$. *Therefore, when* $\mu \to 0$ *we have that*

$$det(J(x_1(0), x_2(0), z_1(0), z_2(0)) = \begin{cases} (c_2 - Mc_1) & \text{if } c_2 > Mc_1 \\ (Mc_1 - c_2) & \text{if } c_2 < Mc_1 \\ 0 & \text{if } c_2 = Mc_1. \end{cases} \tag{5.22}$$

We see from (5.21) and (5.22) that the probability density of the primal solutions V_A *and* V_B *increase, respectively, with* $c_2 - Mc_1$ *and* $c_1 - Mc_2$, *which is consistent with the solution of the primal problem. In addition, the probability density of the analytic center solution* $(\frac{1}{2}, \frac{1}{2M})$ *is 0, which is consistent with the fact that this solution can only be obtained when the two random costs* c_1, c_2 *satisfy* $c_2 = Mc_1$ *(the probability of such an event being 0).*

It is worth to make some observations to (5.15) and Theorem 5:

- From (5.15), the probability density function of the primal-dual solutions depends on the randomness of $\widehat{\mathbf{c}}$, but also on the feasible polyhedron defined by the constraints matrix $\hat{A}$, which appears in the determinant of the Jacobian of G. This is coherent with the intuition.
- Given a primal-dual central path point$(\widehat{\mathbf{x}}(\mu), \widehat{\mathbf{y}}(\mu), \widehat{\mathbf{z}}(\mu))$ for some cost vector $\widehat{\mathbf{c}}$, equation (5.15), which is easily computed, provides the probability density $f_{\mathcal{I}_D}$ of this primal-dual point. However, to compute the probability density of the primal point $\widehat{\mathbf{x}}(\mu)$ we need to solve
$$f(\widehat{\mathbf{x}}(\mu)) = \int_{\widehat{\mathbf{z}}(\widehat{\mathbf{c}}), \widehat{\mathbf{y}}(\widehat{\mathbf{c}})} f_{\mathcal{I}_D}(\widehat{\mathbf{x}}(\mu), \widehat{\mathbf{y}}(\mu), \widehat{\mathbf{z}}(\mu)) d\widehat{\mathbf{z}} d\widehat{\mathbf{y}}. \tag{5.23}$$
Although (5.23) is a difficult integral, the expression of the density (5.15) is enough to compute primal-dual solutions with any desired distribution (as seen in below in this section).
- A closed form expression for (5.15) can not be computed, in general, because of the determinant of the Jacobian of G, which involves the constraints matrix $\hat{A}$ and the values of $(\widehat{\mathbf{x}}, \widehat{\mathbf{z}})$ in the central path (which do not admit a closed form expression, but for toy problems). However, this result may be valuable to perform, for instance, numerical simulations or experiments.
- Since the Jacobian is nonsingular (as stated above), the determinant is never zero. In addition, since the points of the central path are bounded [233], the determinant is bounded.
- A simpler expression for the determinant can be derived. Denoting by J the Jacobian, and adding to the first block-row of J the last block-row multiplied by the diagonal matrix $\widehat{X}^{-1}$ we have
$$\det(J) = \det(J_1) \prod_{i=1}^{n} \widehat{\mathbf{x}}_i \quad \text{where} \quad J_1 = \begin{bmatrix} -\widehat{X}^{-1}\widehat{Z} & \widehat{A}^T \\ \widehat{A} & \end{bmatrix}.$$
Since
$$J_1 = CDC^T \quad \text{where} \quad C = \begin{bmatrix} I & \\ -\widehat{A}\widehat{X}\widehat{Z}^{-1} & I \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} -\widehat{X}^{-1}\widehat{Z} & \\ & \widehat{A}\widehat{X}\widehat{Z}^{-1}\widehat{A}^T \end{bmatrix},$$
using that $\det(C) = 1$ we finally obtain
$$\det(J) = \det(D) \prod_{i=1}^{n} \widehat{\mathbf{x}}_i = (-1)^n \det(\widehat{A}\widehat{X}\widehat{Z}^{-1}\widehat{A}^T) \prod_{i=1}^{n} \widehat{\mathbf{z}}_i. \tag{5.24}$$
$\widehat{A}\widehat{X}\widehat{Z}^{-1}\widehat{A}^T$ is the symmetric and positive definite matrix—if $\widehat{A}$ has full row rank—of the normal equations system of interior-point methods [231].

It remains to investigate whether $f_{\mathcal{I}_D}$ will be random enough when f_C is generated from a uniform random distribution. The *differential entropy* of a continuous random variable ξ of density $f(\xi)$ and support Ξ can be used for this purpose, defined as [74]

$$H(f(\xi)) = -\int_{\Xi} f(\xi)\log f(\xi)d\xi. \tag{5.25}$$

The differential entropy measures how the probability density is overspread, such that uniform distributions are related to maximal entropies. Let us consider the cost vector $\widehat{\mathbf{c}}$ is randomly generated from a uniform distribution of support $[-c,c]^{\hat{N}}$. From (5.25), it is straightforward to see that the differential entropy of uniform distributions of support $[a,b]^{\hat{N}}$ is $\log(b-a)^{\hat{N}}$. If $|b-a|$ tends to 0—that is, there is no "randomness"—, the entropy tends to $-\infty$. In our case, the differential entropy of $\widehat{\mathbf{c}}$ is $\log(2c)^{\hat{N}}$.

Since $f_{\mathcal{I}_D}(\mathbf{t})$ and $f_{\mathcal{I}_C}(\mathbf{s})$ are related through (5.13), the differential entropy of $f_{\mathcal{I}_D}(\mathbf{t})$ can be computed as follows [74]

$$H_1[f_{\mathcal{I}_D}] = H_1[f_{\mathcal{I}_C}] - E\left[\log\left(\left\|\begin{array}{ccc}\widehat{A} & & \\ & \widehat{A}^T & I \\ \widehat{Z} & & \widehat{X}\end{array}\right\|\right)\right]. \tag{5.26}$$

where $E[\cdot]$ is the expectation operator. If $H(f_{\mathcal{I}_D}(\mathbf{t}))$ is not "very small" (or even tends to $-\infty$) we may claim is "random enough". If the determinant of the Jacobian was uniformly upper bounded for any μ and $\widehat{\mathbf{c}}$ then (5.26) would provide a lower bound for $H(f_{\mathcal{I}_D}(\mathbf{t}))$.

The importance of this results lays on the possibility to arbitrarily increases the min-entropy and the differential-entropy of the prima-dual solutions $[\widehat{\mathbf{x}}\ \widehat{\mathbf{y}}\ \widehat{\mathbf{z}}] \in \mathcal{I}_D$ by increasing the one of the objective gradient $\widehat{\mathbf{c}} \in \mathbb{R}^{\hat{N}}$, allowing a closer approximation to the uniform distribution. Moreover, since $XZ\mathbf{e} = \mu\mathbf{e}$, when μ approaches the zero (from the right), the expectation of the log determinant in (5.26) monotonically decreases and $H_1[f_{\mathcal{I}_D}]$ increases.

An exact calculation of (5.26) requires the expectation for all the possible solutions and it is valid for any LP. Therefore the analysis of whether $H(f_{\mathcal{I}_D}(\mathbf{t}))$ is lower bounded for some problem should be done numerically. To illustrate this fact, we considered two small instances of n nodes of the edge-colored networks described in Proposition 7 (Subsection 3.2.1), and solved them for two randomly generated $\widehat{\mathbf{c}}$, and a sequence of μ values. Figure 5.2 shows the results obtained for $n \in \{6,12\}$ and $c \in \{0.5,100\}$, where $\widehat{\mathbf{c}} \in [-c,c]^{\widehat{n}}$. The horizontal axis is μ and the vertical one shows $H(f_{\mathcal{I}_D}(\mathbf{t}))$. Every point of the plots is the average of 1000 LP problems obtained by randomly generating the cost vector. It can be seen that as μ approaches 0 (i.e., as the central path tends to the LP solution) the differential entropy increases. When $c = 0.5$, the differential entropy of $\widehat{\mathbf{c}}$ is $\log(2\cdot 0.5)^{\widehat{n}} = 0$, whereas that of the LP solutions of the two instances are 10^{-6} and 10^{-46} approximately; therefore, both the cost vector and the LP solutions have similar randomness. When $c = 100$, the differential entropy of the solutions do not tend to $-\infty$ (but to approximately 0 and 650), guaranteeing they are "random enough".

If our aim is to draw a large sample from the uniform distribution of primal-dual solutions, the availability of a computable probability density function (5.15) allows a straightforward application of the Metropolis-Hastings acceptance criterion [197]. This method consists in generating a stochastic sequence of primal-dual solutions, from an arbitrary starting point $\mathbf{t}^0 = (\widehat{\mathbf{x}}^0, \widehat{\mathbf{y}}^0, \widehat{\mathbf{z}}^0)$, and the following rule to go from a current state $\mathbf{t}^k = (\widehat{\mathbf{x}}^k, \widehat{\mathbf{y}}^k, \widehat{\mathbf{z}}^k)$ to a new state $\mathbf{t}^{k+1} = (\widehat{\mathbf{x}}^{k+1}, \widehat{\mathbf{y}}^{k+1}, \widehat{\mathbf{z}}^{k+1})$:

1. Initialization: choose an arbitrary point $\mathbf{t}^0$ to be the first sample and let (5.15) be the proposal distribution, which suggests a candidate for the next sample value $\mathbf{t}^{k+1}$, given

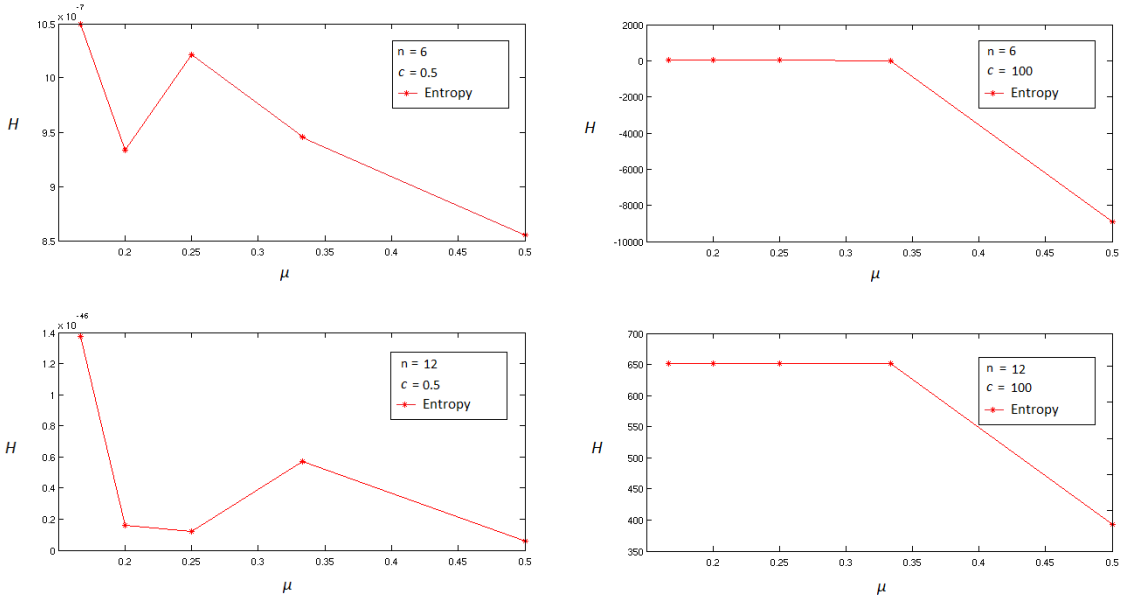


Figure 5.2: Evolution of the differential entropy $H(f_{\mathcal{I}_D}(\mathbf{t}))$ of the solutions for four edge-colored network LPs as μ approaches 0, for cost values and number of nodes $c \in \{0.5, 100\}$ and $n \in \{6, 12\}$, respectively.

the previous sample value $\mathbf{t}^k$. (In the case of (5.15) the transaction probabilities are independents.)

2. Candidate state: propose a candidate point $\mathbf{t}^{k+1}$ from a proposal distribution, (5.15) that may depend on the current $\mathbf{t}^k$. (In our case the proposal entails the solution of a LP.)

3. Acceptance criterion: accept the candidate state with probability

$$\alpha(\mathbf{t}^k, \mathbf{t}^{k+1}) = \min\left(1,\ \frac{f_{\mathcal{I}_D}(\mathbf{t}^k)}{f_{\mathcal{I}_D}(\mathbf{t}^{k+1})}\right). \tag{5.27}$$

First a candidate value is generated by solving the LP so to perturb the current state of the chain with a random innovation. Then the acceptance probability is computed, and the chain is updated appropriately depending on whether the proposed new value is accepted.

The correctness of this procedure to generate uniform primal-dual solutions can be directly assessed by checking that the acceptance criterion (5.36) verifies detailed balance and the proposal distribution (5.15) is aperiodic and irreducible. (For more details about Markov Chain Monte Carlo methods see [197] and Appendix E.)

The efficiency of Metropolis-Hastings method strongly depends on the number of rejections. When the ratio between $|\det(J(G(\widehat{\mathbf{x}}^k, \widehat{\mathbf{y}}^k, \widehat{\mathbf{z}}^k)))|$ and $|\det(J(G(\widehat{\mathbf{x}}^{k+1}, \widehat{\mathbf{y}}^{k+1}, \widehat{\mathbf{z}}^{k+1})))|$ is closed to zero the number of rejected networks increase. In tables 5.2 and 5.3 summarize the distribution of α for a sample of 50 primal-dual solutions of the LPs associated to the family of networks with fixed density and the family of networks with fixed degree sequence. The ratio between the $50(50-1) = 2450$ couples of determinants have been computed to quantify the probability of a rejection.

		$\mu = 0.5$			$\mu = 0.1$		
n	c	$\alpha < 0.9$	$\alpha < 0.7$	$\alpha < 0.5$	$\alpha < 0.9$	$\alpha < 0.7$	$\alpha < 0.5$
6	0.5	0.4878	0.4555	0.3971	0.4771	0.4457	0.3849
6	100	0.4955	0.4865	0.3971	0.4971	0.4898	0.3763
12	0.5	0.4878	0.4555	0.3971	0.4951	0.4698	0.3865
12	100	0.4988	0.4906	0.4284	0.4992	0.4290	0.3935

Table 5.2: The proportion of α smaller then 0.9, 0.7 and 0.5 has been computed under different parametrization of the cost vector c and the network size n for the family of networks with fixed density. The same results are obtained for two different values of the barrier parameter μ.

		$\mu = 0.5$			$\mu = 0.1$		
n	c	$\alpha < 0.9$	$\alpha < 0.7$	$\alpha < 0.5$	$\alpha < 0.9$	$\alpha < 0.7$	$\alpha < 0.5$
6	0.5	0.3760	0.1415	0.0006	0.4989	0.4034	0.3136
6	100	0.4992	0.4969	0.0062	0.4927	0.4152	0.3784
12	0.5	0.4028	0.2042	0.0548	0.4955	0.4547	0.3971
12	100	0.4997	0.4263	0.3879	0.4994	0.4283	0.3952

Table 5.3: The proportion of α smaller then 0.9, 0.7 and 0.5 has been computed under different parametrization of the cost vector c and the network size n for the family of networks with fixed degree sequence. The same results are obtained for two different values of the barrier parameter μ.

The numerical results in tables 5.2 and 5.3 support our conjecture about the likely uniform distribution of (5.15).

Note that computing the acceptance probability implies the evaluation of

$$\frac{f_{\mathcal{I}_D}(\widehat{\mathbf{x}}^k, \widehat{\mathbf{y}}^k, \widehat{\mathbf{z}}^k)}{f_{\mathcal{I}_D}(\widehat{\mathbf{x}}^{k+1}, \widehat{\mathbf{y}}^{k+1}, \widehat{\mathbf{z}}^{k+1})} = \eta \frac{\det(\hat{A}\hat{(}X^k)^2\hat{A}^T)\prod_{i=1}^n \widehat{\mathbf{z}}_i^k}{\det(\hat{A}\hat{(}X^{k+1})^2\hat{A}^T)\prod_{i=1}^n \widehat{\mathbf{z}}_i^{k+1}}. \tag{5.28}$$

where $\eta = \exp\left(\frac{1}{2\sigma^2}(\widehat{\mathbf{c}}^k - \widehat{\mathbf{c}}^{k+1})^T(\widehat{\mathbf{c}}^k - \widehat{\mathbf{c}}^{k+1})\right)$ in the case the objective costs are distributed as a non correlated multivariate normal with means zero and variances σ^2; or $\eta = 1$ in the case the objective costs are uniformly distributed. Other distributions of the objective costs might also be taken into account.

In practice we calculate the log-acceptance ratio to avoid multiplying a very large sequence of small numbers, where floating point underflow might represent a constant threat:

$$\begin{aligned}\min\left(0,\ \log\eta\frac{\det(\hat{A}\hat{(}X^k)^2\hat{A}^T)\prod_{i=1}^n \widehat{\mathbf{z}}_i^k}{\det(\hat{A}\hat{(}X^{k+1})^2\hat{A}^T)\prod_{i=1}^n \widehat{\mathbf{z}}_i^{k+1}}\right) &= \log\det(\hat{A}\hat{(}X^k)^2\hat{A}^T) - \log\det(\hat{A}\hat{(}X^{k+1})^2\hat{A}^T) \\ &\quad + \sum_{i=1}^n \log\widehat{\mathbf{z}}_i^k - \log\widehat{\mathbf{z}}_i^{k+1} + \log\eta\end{aligned}$$

Since the number of rejections is particularly small and the computation of the log-determinant can be quite troublesome for big matrices[1], approximations of the log-acceptance ratio might be considered[2].

[1] Computing $\det(\hat{A}\hat{(}X^k)^2\hat{A}^T)$ by Cholesky decomposition requires $\mathcal{O}(\widehat{M}^3)$ operation, where $\widehat{M}$ is the number of rows of $\hat{A}$. Since what is needed is the logarithm of the determinant rather than the determinant itself, an important implementation trick is to keep the computation at logarithm scale, so that we can effectively tackle the problem by computing sum-of-log rather than computing log-of-product.

[2] Barry and Pace [22] toke into account a power series expansion of the log-determinant, where all the terms in the expansion involve matrix traces; they proposed to approximate through a Monte Carlo method the traces instead of computing them explicitly.

5.3.1 Irreducibility and aperiodicity of the Markov chain of μ-solutions

We show in this subsection that the Metropolis-Hastings chain of primal-dual μ-solutions is irreducible. Consider the transaction probability from $\mathbf{t}^k = (\widehat{\mathbf{x}}^k, \widehat{\mathbf{y}}^k, \widehat{\mathbf{z}}^k)$ to $\mathbf{t}^{k+1} = (\widehat{\mathbf{x}}^{k+1}, \widehat{\mathbf{y}}^{k+1}, \widehat{\mathbf{z}}^{k+1})$ associated to the Metropolis-Hastings chain:

$$P(\mathbf{t}^{k+1} \mid \mathbf{t}^k) = f_{\mathcal{I}_D}(\mathbf{t}^{k+1})\alpha(\mathbf{t}^k, \mathbf{t}^{k+1}) \tag{5.29}$$

To prove the irreducibility of the Metropolis-Hastings chain it is sufficient to check the non negativity of the proposal probability for every μ-solutions, that is to say, $f_{\mathcal{I}_D}(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) > 0$ in $\{(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) \ : \ G(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) = [\mathbf{0}^T \ \widehat{\mathbf{c}}^T \ \mathbf{0}^T]^T$, for every $\widehat{\mathbf{c}} \in \mathbb{R}^{\widehat{N}}\}$. (Less strict conditions of irreducibility can be obtained; see Roberts and Tweedie [199]).

From the KKT conditions in (5.7), we see that for every point $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}})$, verifying $\widehat{A}\widehat{\mathbf{x}} = \widehat{\mathbf{b}}$ and $\widehat{X}\widehat{Z}\mathbf{e} = \mu\mathbf{e}$, there must exist a $\widehat{\mathbf{c}} \in \mathbb{R}^{\widehat{N}}$ for which $\widehat{A}^T\widehat{\mathbf{y}} + \widehat{\mathbf{z}} = \widehat{\mathbf{c}}$. Thus, as long as $f_C(\widehat{\mathbf{c}}) > 0$, for every $\widehat{\mathbf{c}} \in \mathbb{R}^{\widehat{N}}$ the proposal probability verifies $f_{\mathcal{I}_D}(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) > 0$ in $\{(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) \ : \ G(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) = [\mathbf{0}^T \ \widehat{\mathbf{c}}^T \ \mathbf{0}^T]^T$, for every $\widehat{\mathbf{c}} \in \mathbb{R}^{\widehat{N}}\}$ and the Metropolis-Hastings chain is irreducible.

As far as the aperiodicity is concerned, it is strait straightforward to check that the probability of the chain to remain in the same state $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}})$ is $f_{\mathcal{I}_C}(G(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}))$, which is strictly positive in $\{(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) \ : \ G(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) = [\mathbf{0}^T \ \widehat{\mathbf{c}}^T \ \mathbf{0}^T]^T$, for every $\widehat{\mathbf{c}} \in \mathbb{R}^{\widehat{N}}\}$ if $f_C(\widehat{\mathbf{c}}) > 0$.

5.3.2 Truncated distribution of primal-dual solutions

The literal meaning of truncation in the mathematical context is to *cut–off* elements from a given set. In statistics the notion of truncation of a random vector concerns the change of its density function after reducing its support, that is to say, assigning a zero probability to some elements of the sample space and redefining the density function of the remaining support.

If applied to our model, it might be shown that the Metropolis-Hastings mechanism works pretty well when the target distribution which does not have full support.

Let $\mathcal{I}_D = \{\mathbf{t} = (\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) : \widehat{\mathbf{x}} > 0, \widehat{\mathbf{z}} > 0, \widehat{A}\widehat{\mathbf{x}} = \widehat{\mathbf{b}}, \widehat{A}^T\widehat{\mathbf{y}} + \widehat{\mathbf{z}} = \widehat{\mathbf{c}}, \widehat{X}\widehat{Z}\mathbf{e} = \mu\mathbf{e}\}$ and $\varphi \subset \mathcal{I}_D$, representing a subset of primal-dual solutions in a ε-neighborhood of fractional extreme points, that is to say, $\varphi = \{(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) \in \mathcal{I}_D \ : \ \widehat{\mathbf{x}} - \text{round}(\widehat{\mathbf{x}}) > \varepsilon\}$. The ε-neighborhood of fractional extreme points are illustrated in Figure 5.3.

If the result of a sampling procedure is *cut–off* we obtain a *shorten* distribution. Since the density of the truncated prima-dual points is omitted from their probability density function, the remaining distribution is shifted upward so that the area beneath it is still one, in accordance with Theorem 6.

Theorem 6. *Let $\varphi \subset \chi$ and $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) \in \chi$ be a continuous random vector, with probability density function $f_{\mathcal{I}_D}(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}})$. If the distribution is truncated so that only the values in χ/φ are observed, then the probability density function of the truncated random vector is*

$$f^{\varphi}_{\mathcal{I}_D}(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) \ = \ \frac{f_{\mathcal{I}_D}(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}})}{1 - \displaystyle\int_{\varphi} f_{\mathcal{I}_D}(s)ds}. \tag{5.30}$$

Since $1 - \displaystyle\int_{\varphi} f_{\mathcal{I}_D}(s)ds$ is a constant value, then the Metropolis-Hastings acceptance ratio does not change. By repeatedly sampling from $f_{\mathcal{I}_D}$ until we obtain a point $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}) \notin \varphi$, we are actually implementing a rejection sampler to generate primal-dual solution truncated at φ.

Now, the truncated density has the same density as the untruncated density, apart from the differing support and a normalizing constant.

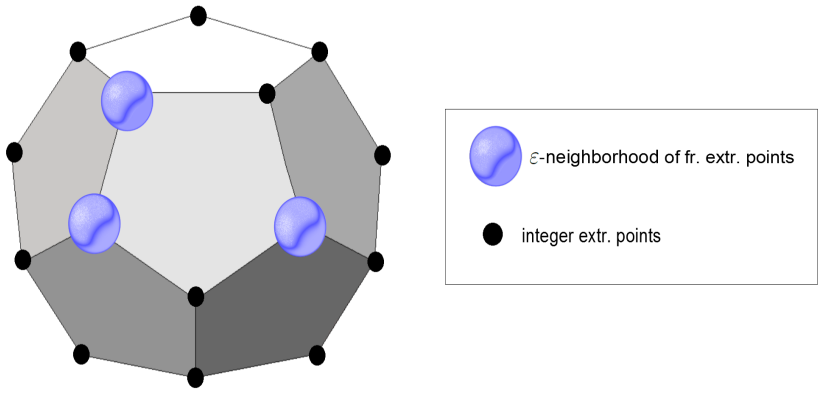


Figure 5.3: ε-neighborhood of fractional extreme points in a polytope.

A truncated distribution of primal-dual solutions can also be used to reject points which do not verify certain conditions. For example, in the case of dealing with connected networks the omitted subset φ correspond to the primal dual solutions which do not verify

- $x_{ij} - \text{round}(x_{ij}) < \varepsilon$, for each $(i,j) \in \mathcal{H}^2$;
- for each $(i,j) \in \mathcal{H}^2$ there exists $f_{ji} \geq 0$, such that $f_{ij} + f_{ji} \leq (n-1)x_{ij}$ and $\sum_{j=1}^{n} f_{1j} - \sum_{j=1}^{n} f_{j1} = n-1$, $\sum_{j=1}^{n} f_{kj} - \sum_{j=1}^{n} f_{jk} = -1$, for $k \in \mathcal{V}, k \neq 1$.

In practice such a rejection method might be useful as long as the number of solution to be rejected is not particularly high. In the case of dealing with connected networks, the researcher should consider whether the rejecting disconnected networks is more convenient than introducing the connectivity constraints into the mathematical programming model, so to uniquely reject fractional solutions.

5.3.3 Generating conditionally uniform random networks: numerical analysis

In this subsection, the r-blocks, s-pivots and q-kernel methods proposed in Chapter 3 are analyzed with respect of their correctness and efficiency. The r-blocks methods is applied by using the Metropolis-Hastings acceptance criterion, whereas the s-pivots and q-kernel methods are directly applied without rejections.

Three network data set described in Chapter 2 are here considered: the 62 node undirected graph, representing the social network of frequent associations between dolphins in a community living off Doubtful Sound, New Zealand [152, 153]; and two data sets corresponding to 39 node undirected graphs, representing alliances among workers during extended negotiations for higher wages in a tailor shop in Zambia at two different times (seven months apart) [133].

A uniform random network conditioned to the density is considered, along with the empirical distribution of two network features: clustering coefficient (CC) and assortativity coefficient (AC). The simulation is divided into pre- and post-convergence periods, where the pre-convergence part, known as *burn-in*, is discarded and the post-convergence part is used for

inference. The numerical results of this section have been obtained discarding the first 200 networks, so they only include a smaller sample size of 800 instances.

For the 62 node network of dolphins, four samples of 1000 networks have been obtained using the r-blocks, s-pivots and q-kernel methods with different r and s values, as shown in tables 5.4, 5.5 and 5.6. The mean and standard deviation of the CC and AC over the samples generated are reported, as well as the CPU time in seconds.

Theoretical results [29] state that the expected CC of a uniform random network with n nodes conditioned to d edges is $= 2d/n^2$, which in the case of the dolphin social network is $2 \cdot 159/62^2 = 0.08272633$. This is approximately what we obtained in tables 5.4, 5.5 and 5.6. The expected AC of a uniform random network with n nodes conditioned to d is approximately $-\frac{1}{n/2-1}$ (this approximation is based on the multivariate hypergeometric distribution of the degree vector), which in the case of the dolphin social network is $-\frac{1}{62/2-1} = -0.03333$, fitting reasonably well the simulated networks.

r	mean CC	std. CC	mean AC	std. AC	rejected	CPU time
12	0.0819	0.0185	-0.0353	0.0774	1226	58.4
24	0.0806	0.0109	-0.0313	0.0740	1659	47.4
48	0.0796	0.0185	-0.0195	0.0618	1779	61.5
96	0.0869	0.0181	-0.0562	0.0860	1865	55.3

Table 5.4: Numerical results using the r-blocks method for the dolphin data set.

s	mean CC	std. CC	mean AC	std. AC	CPU time
20	0.0846	0.0207	-0.0301	0.0751	20.4
40	0.0836	0.0192	-0.0301	0.0823	18.8
80	0.0801	0.0198	-0.0468	0.0816	17.9
160	0.0759	0.0121	-0.0265	0.0659	17.4

Table 5.5: Numerical results using the s-pivots method for the dolphin data set.

s	mean CC	std. CC	mean AC	std. AC	CPU-time
20	0.0791	0.0198	-0.0293	0.0715	11.7
40	0.0898	0.0386	-0.0280	0.0849	8.9
80	0.0962	0.0515	-0.0430	0.0765	7.5
160	0.1022	0.0676	-0.0354	0.0802	6.4

Table 5.6: Numerical results using the q-kernel method for the dolphin data set.

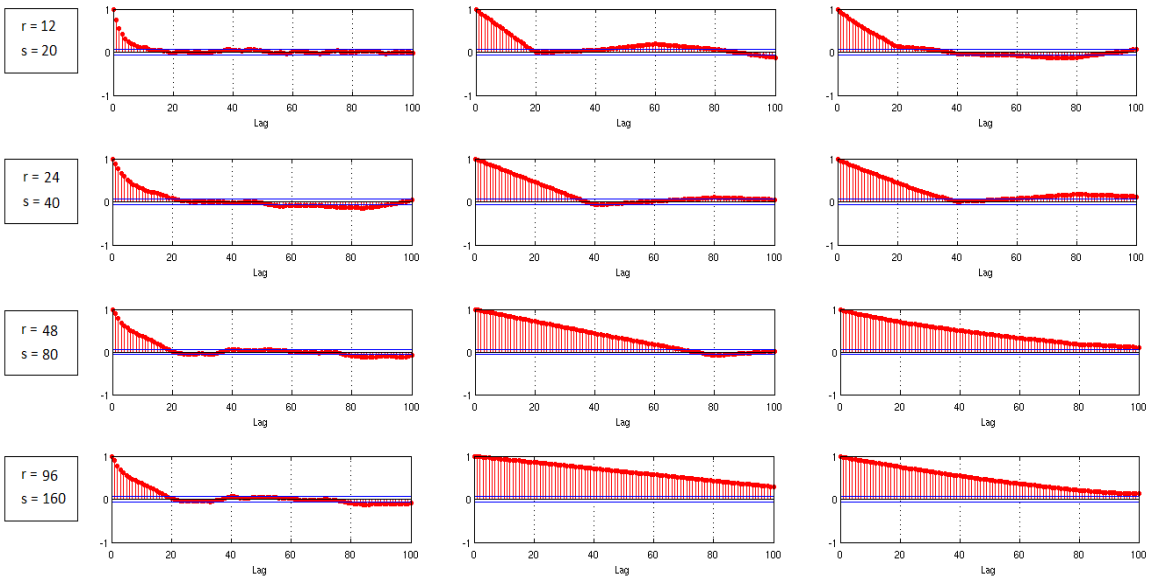


Figure 5.4: Autocorrelation function of CC for the r-blocks (left plots), s-pivots (center plots) and q-kernel (right plots) algorithm corresponding to the samples in tables 5.4, 5.5 and 5.6.

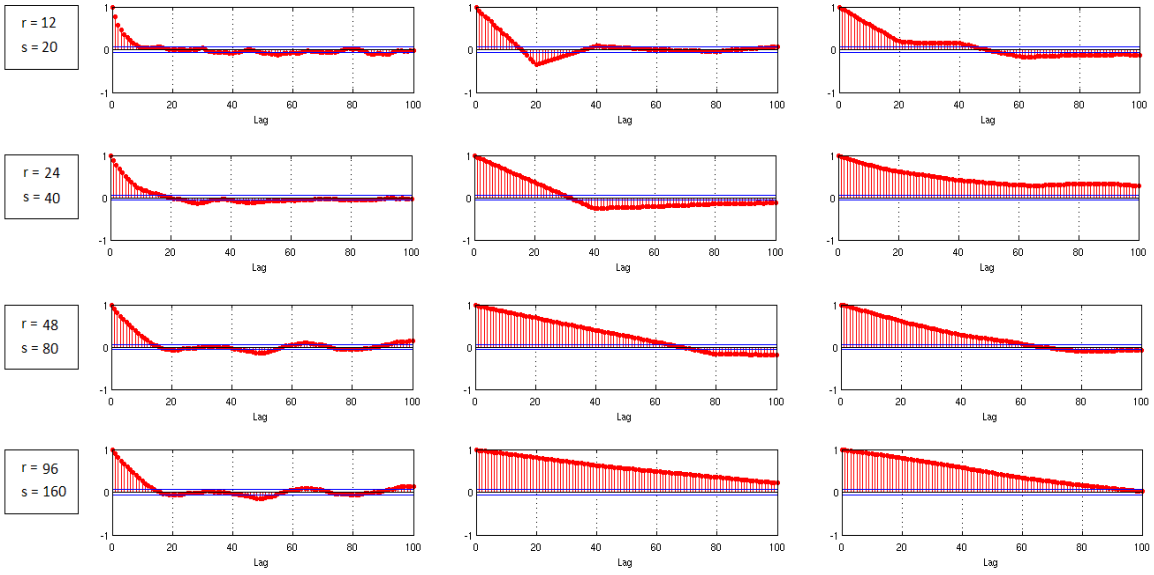


Figure 5.5: Autocorrelation function of AC for the r-blocks (left plots), s-pivots (center plots) and q-kernel (right plots) algorithm corresponding to the samples in tables 5.4, 5.5 and 5.6.

The plots in figures 5.4 and 5.5 show the respective autocorrelations function of CC and AC for the 1000 networks obtained in each of the runs of tables 5.4, 5.5 and 5.6, for increasing values of r and s, from top to bottom. The length of the burn-in period should reflect the autocorrelation of the resulting process, as well as the mixing-time of these chains. We observe a decreasing pattern when r and s are small.

Tables 5.7, 5.8 and 5.9 summarize the analogous information for network data set of 39 workers in a tailor shop observed in the first period. The numerical values of CC and AC also

resemble the theoretical values under the considered probabilistic model, which are $2 \cdot 158/39^2 = 0.2077581$ and $-\frac{1}{39/2-1} = -0.05405405$ respectively.

r	mean CC	std. CC	mean AC	std. AC	rejected	CPU time
12	0.2145	0.0237	-0.0481	0.0629	1006	45.6
24	0.2160	0.0216	-0.0491	0.0715	1216	40.2
48	0.2170	0.0215	-0.0511	0.0601	1452	37.4
96	0.2154	0.0207	-0.0574	0.0639	1673	37.7

Table 5.7: Numerical results using the r-blocks method for the network of workers in the first period.

s	mean CC	std. CC	mean AC	std. AC	CPU time
20	0.2099	0.0185	-0.0353	0.0635	8.1
40	0.2070	0.0201	-0.0707	0.0700	7.9
80	0.2111	0.0159	-0.1054	0.0714	7.8
160	0.2211	0.0212	-0.0006	0.0474	7.6

Table 5.8: Numerical results using the s-pivots method for the network of workers in the first period.

s	mean CC	std. CC	mean AC	std. AC	CPU-time
20	0.2139	0.0215	-0.0475	0.0701	8.0
40	0.2176	0.0237	-0.0768	0.0717	7.6
80	0.2236	0.0272	-0.0208	0.0598	7.3
160	0.2106	0.0178	-0.0960	0.0141	6.8

Table 5.9: Numerical results using the q-kernel method for the network of workers in the first period.

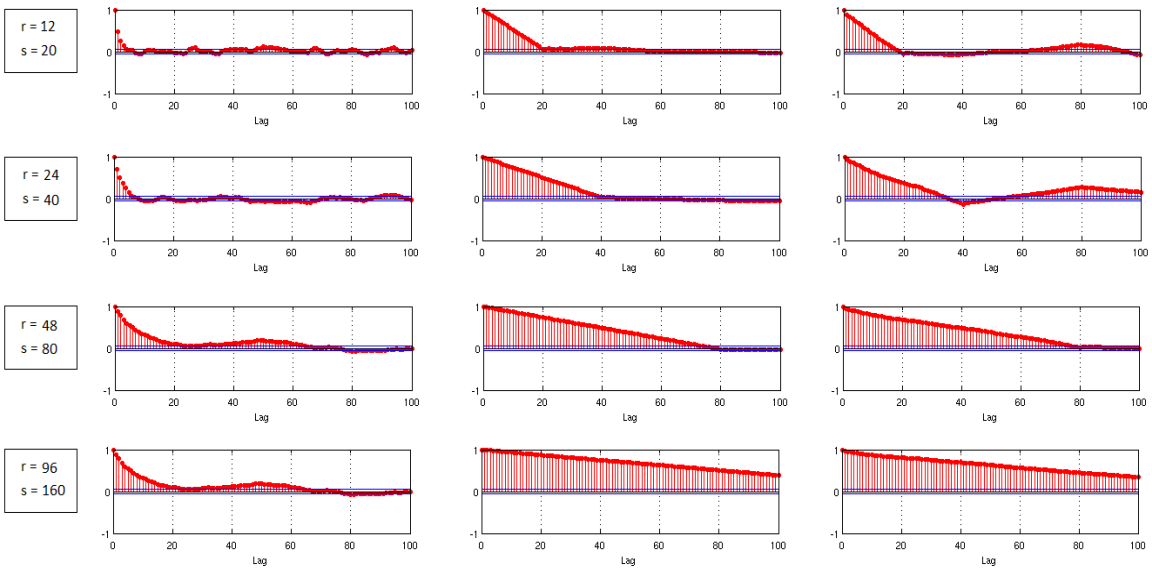


Figure 5.6: Autocorrelation function of CC for the r-blocks (left plots), s-pivots (center plots) and q-kernel (right plots) algorithm corresponding to the samples in tables 5.7, 5.8 and 5.9.

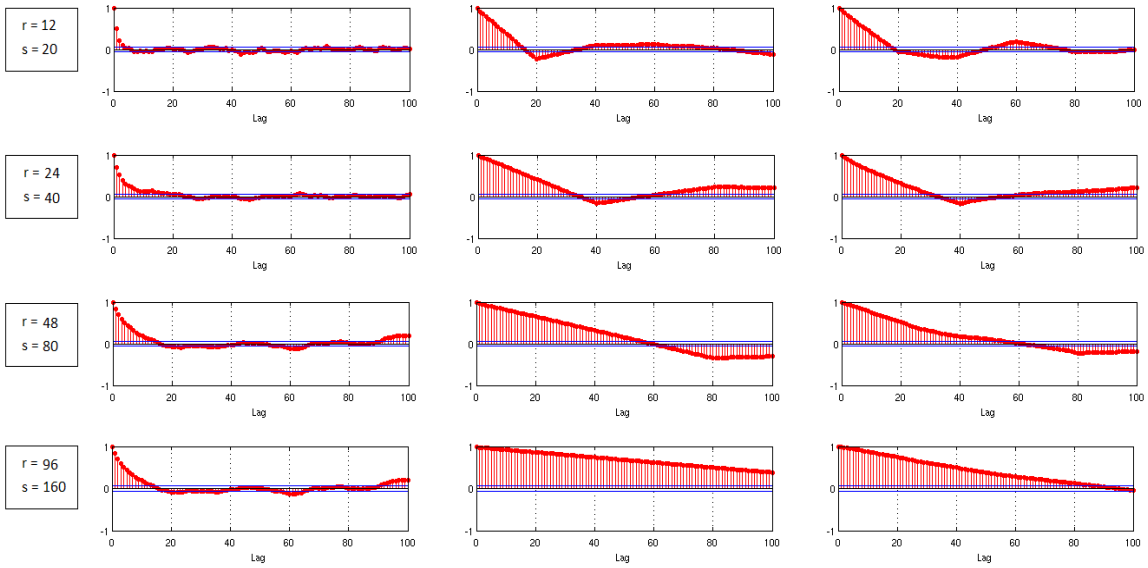


Figure 5.7: Autocorrelation function of AC for the r-blocks (left plots), s-pivots (center plots) and q-kernel (right plots) algorithm corresponding to the samples in tables 5.7, 5.8 and 5.9.

Finally, for the network of workers associated to the second period, the corresponding theoretical values of CC and AC are $2 \cdot 223/39^2 = 0.2932281$ and $-\frac{1}{39/2-1} = -0.05405405$, which are also consistent with the numerical results in tables 5.10, 5.11 and 5.12, corresponding to the plots in figures 5.8 and 5.9.

r	mean CC	std. CC	mean AC	std. AC	rejection	CPU time
12	0.3002	0.0154	-0.0531	0.0540	910	71.4
24	0.3022	0.0168	-0.0527	0.0526	1082	48.4
48	0.3006	0.0164	-0.0441	0.0572	1265	42.1
96	0.2972	0.0134	-0.0334	0.0624	1305	41.5

Table 5.10: Numerical results using the r-blocks method for the network of workers in the second period.

s	mean CC	std. CC	mean AC	std. AC	CPU time
20	0.3001	0.0125	-0.0391	0.0521	8.5
40	0.3055	0.0135	-0.0578	0.0614	8.2
80	0.2993	0.0171	-0.0667	0.0518	8.1
160	0.3093	0.0101	-0.0118	0.0330	8.0

Table 5.11: Numerical results using the s-pivots method for the network of workers in the second period.

s	mean CC	std. CC	mean AC	std. AC	CPU-time
20	0.3001	0.0125	-0.0391	0.0521	8.5
40	0.3055	0.0135	-0.0578	0.0614	8.2
80	0.2993	0.0171	-0.0667	0.0518	8.1
160	0.3093	0.0101	-0.0118	0.0330	8.0

Table 5.12: Numerical results using the q-kernel method for the network of workers in the second period.

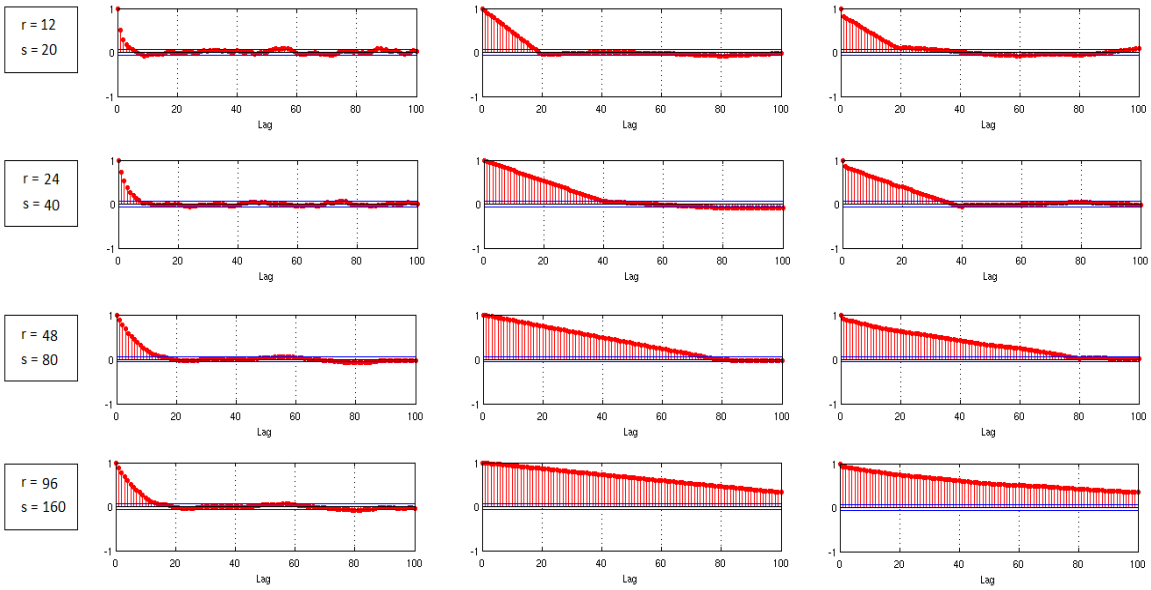


Figure 5.8: Autocorrelation function of CC for the r-blocks (left plots), s-pivots (center plots) and q-kernel (right plots) algorithm corresponding to the samples in tables 5.10, 5.11 and 5.12.

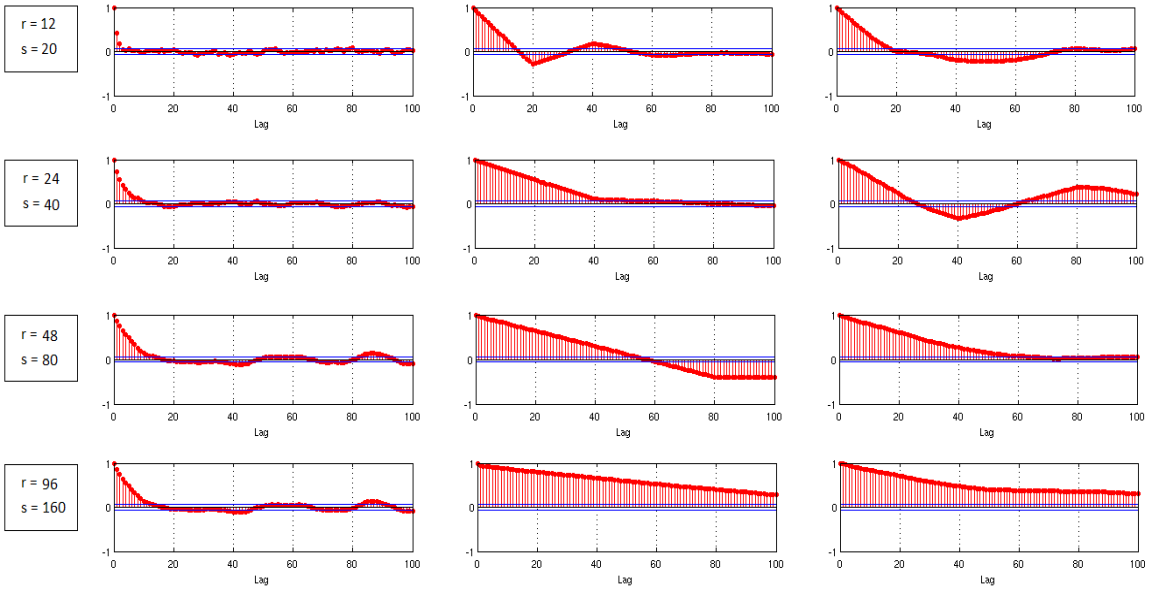


Figure 5.9: Autocorrelation function of AC for the r-blocks (left plots), s-pivots (center plots) and q-kernel (right plots) algorithm corresponding to the samples in tables 5.10, 5.11 and 5.12.

We see from the plots in figures 5.6, 5.7, 5.8 and 5.9 a clear autocorrelated behavior of the s-pivots and the sr-kernel methods, even when $s = 20$. On the other hand, the autocorrelations of the sample obtained with the r-blocks procedure quickly tend to zero for small lags.

The conclusions about this comparison between the three methods might be controversial. On the one hand, the s-pivots and q-kernel methods outperform the r-blocks method in terms of efficiency. On the other hand, the autocorrelations reduces when s and r are small, though the

s-pivots and q-kernel methods maintain a strongly autocorrelated behavior even when $s = 20$ – solving the independent full LP every 20 pivots. Thus, despite the higher efficiency of the s-pivots and q-kernel methods, the generation of less autocorrelated networks by the r-blocks method allows for a much smaller sample size when simulating large networks. Moreover, the availability of a closed-form probability density function (5.13) allows a rigorous evaluation of the probabilistic properties of the r-blocks method, whose numerical correctness supports the theoretical claim of Section 5.3.

The analysis of this simple model allowed us to validate our procedures and result of little utility as a model of network formation. In fact, as already mentioned, the results of Newman [171] provide a clear explanation of the fact that the $G(n, m)$ model poorly fits the most relevant structural properties of real-world networks. Nonetheless, for more complicated models, as the ones analyzed in the next subsection, the described LP-based generation methods might be fruitfully applied to obtains empirical distribution of relevant network features. Based on the simulated samples and using the MSS and VSS, as studied in Chapter 2, the goodness of fit of random network models is analyzed in the next section.

5.4 Goodness of fit of random network models

Most of probabilistic models of network consists of multidimensional random variables with complex pattern of dependency. Their goodness of fit is often assessed in term of high-level network features, such as the variance of the degree distribution, the assortatively coefficient and the variance of the local clustering coefficients. In this section we provide graphical approaches to assess the goodness of fit of probabilistic models of networks based on the MSS and VSS, studied in Chapter 2.

Let's consider the previously described social network of dolphins by Lusseau [152] and a conditionally uniform random model where the degree sequence and the within group densities (associated to the three communities in Figure 3.2) are kept constant, along with the graph connectivity. (No theoretical results is available for this probabilistic model). Since the mathematical programming formulation does not characterize the convex hull of the integer solutions, primal dual solution verifying $x - \text{round}(x) > \varepsilon$ are rejected. The expectation and standard deviation of the CC and AC for a sample of 10.000 networks (primal-dual solutions) generated by the r-blocks method with $r = 1$, are shown in Table 5.13.

	sample mean	sample std.	observed value	one tail p-value	corr CC – AC
CC	0.199145	0.01949775	0.2501595	0.075000	-0.06184976
AC	-0.065980	0.07585131	-0.0436017	0.694120	

Table 5.13: Numerical results for the dolphin' social network, under a sample of 10.000 networks, generated by the sequential r-block method. A burn-in period of 100 networks has been discarded.

From Table 5.13 the conclusion is that the expected CC under the specified model is quite close to the observed one, but the small variability of the empirical distribution result in a particularly small p-value, so that the CC cannot be explained by the within-community-densities and the degree sequence. On the other hand, the AC seems to be likely induced by the fixed structural properties we considered. The associated empirical distribution of CC and AC is shown in the density plot of Figure 5.10.

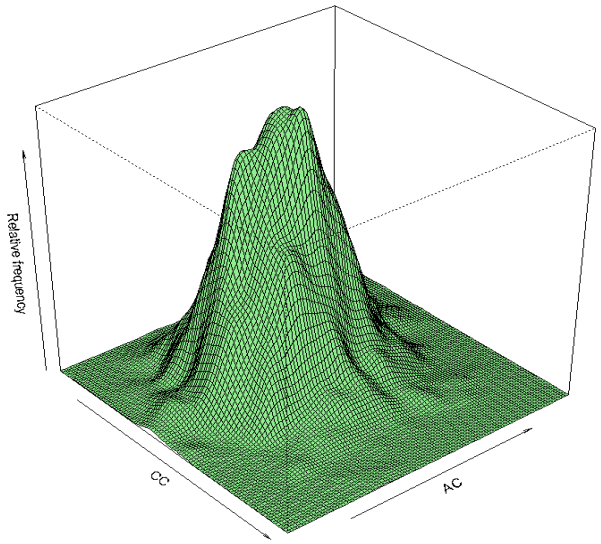


Figure 5.10: Empirical distribution of CC and AC, corresponding to the numerical results in Table 5.13.

The MSS and VSS for each of the simulated networks are calculated, as shown in Figure 5.15, where the means ± standard deviations are denoted by the gray envelope.

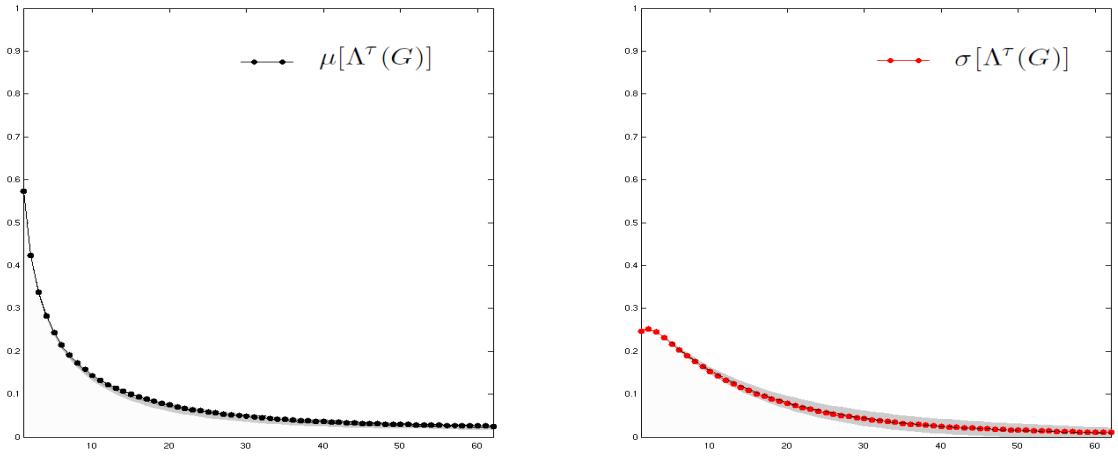


Figure 5.11: The MSS and VSS of 10.000 connected networks with 159 edges and the MSS and VSS of the observed dolphin's social network.

An accurate fit of this model with respect to the MSS and VSS of the observed network is obtained, as shown in Figure 5.10. However the variability of the MSS and VSS increases with τ, outlining a progressively higher differentiation in the pattern of association between structural similarity of order τ and the tie strengths[3]. This fact is somehow reflected in the disagreement

[3]It must be noted that both sequences of structural similarities are strongly affected by the degree sequence and, in the case the probabilistic model preserving the degrees, the overall MSS and VSS might overfit. Overfitting occurs when a statistical model describes random error instead of the underlying complex pattern of dependencies. The conditionally uniform random models with fixed degree sequence and within group densities results in excessively complex modes, such as having too many constraints strongly reducing the dimension of the sample space. It thus has poor predictive performance, as it can amplify slight fluctuations in the data.

between the observed CC and the one expected under simulated model, as shown in Table 5.13.

Consider the network data set representing marriage and business relations among 15 Renaissance Florentine families, as collected by Padgett [183] and two conditionally uniform models:

- uniform undirected valued networks conditioned to the density;
- uniform edge-colored undirected networks conditioned to the within-color densities;

The first model summarize the qualitative nature of a connection in an edge value, defined in accordance with (2.23), as discussed in Chapter 2. In this case an edge is regarded as strong if both a marriage and a business relation occur between its endpoints. On the other hand, the second model takes into account the connection types, i.e. marriage and business relations, and regards the total amount of connections of a given type as a fixed quantity.

For the first model, a chain of 10000 networks associated to the q-kernel method, with $s = 50$, has been simulated, in accordance with (3.16). Correspondingly, the uniform edge-colored undirected network model conditioned to the within-color densities has been simulated by the r-block method, with $r = 1$. The respective results are reported in tables 5.14 and 5.15.

	sample mean	sample std.	observed value	one tail p-value	corr CC – AC
CC	0.3088	0.1146	0.6384	0.0046	-0.0539
AC	-0.1524	0.1548	-0.4786	0.0146	

Table 5.14: Numerical results from the sample obtained with the r-blocks method, uniform undirected valued networks conditioned to the density.

	sample mean	sample std.	observed value	one tail p-value	corr CC – AC
CC	0.3247	0.0702	0.6384	0.0000	-0.1414
AC	-0.1432	0.1311	-0.4786	0.0044	

Table 5.15: Numerical results from the sample obtained with the r-blocks method, uniform edge-colored undirected networks conditioned to the within-color densities.

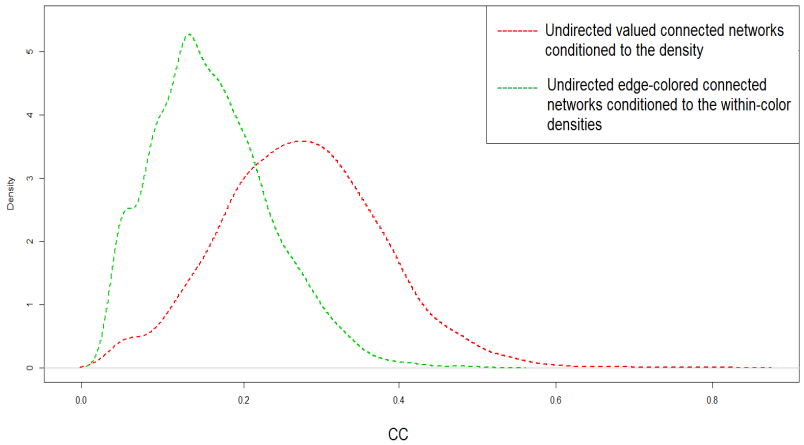


Figure 5.12: Empirical distribution of CC, corresponding to tables 5.14 (red line) and 5.15 (green line).

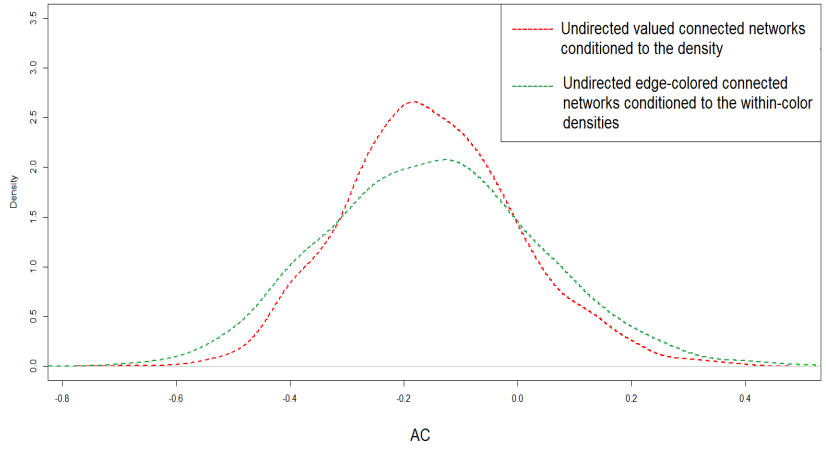


Figure 5.13: Empirical distribution of AC, corresponding to tables 5.14 (red line) and 5.15 (green line).

The two described models, associated to different ways of aggregating the edge information, result in substantial difference of the empirical distribution of CC and AC. This fact brings our discussion back to the problem of edge definition, namely, the problem of establishing *when is an edge an edge*, as addressed by Butts [45] and by Borgatti et al. [37] (see Section 1.1). Any change of the edge set can substantially influence the topology of the resulting family of network, with considerable implications for subsequent analysis, as different aggregation decisions can produce networks with very different structural features.

The numerical results in tables 5.14 and 5.15 show that both models fail to explain the observed clustering coefficient and assortativity coefficient of the network of Renaissance Florentine families and suggest the existence of internal processes between nodes and their relational strategies, which cannot be captured by conditionally uniform distributions.

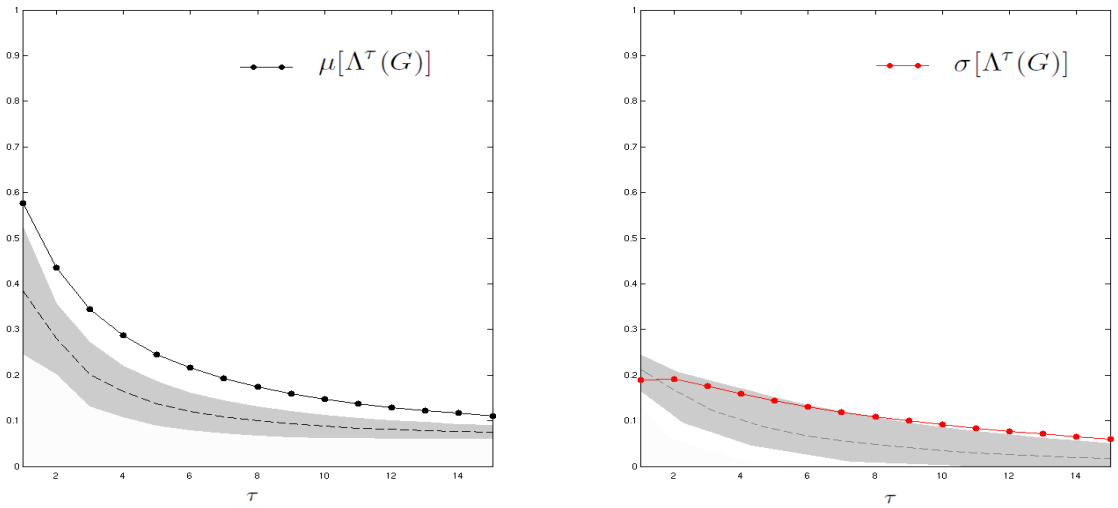


Figure 5.14: The MSS and VSS of 10.000 connected networks from the uniform edge-colored undirected networks conditioned to the within-color densities.

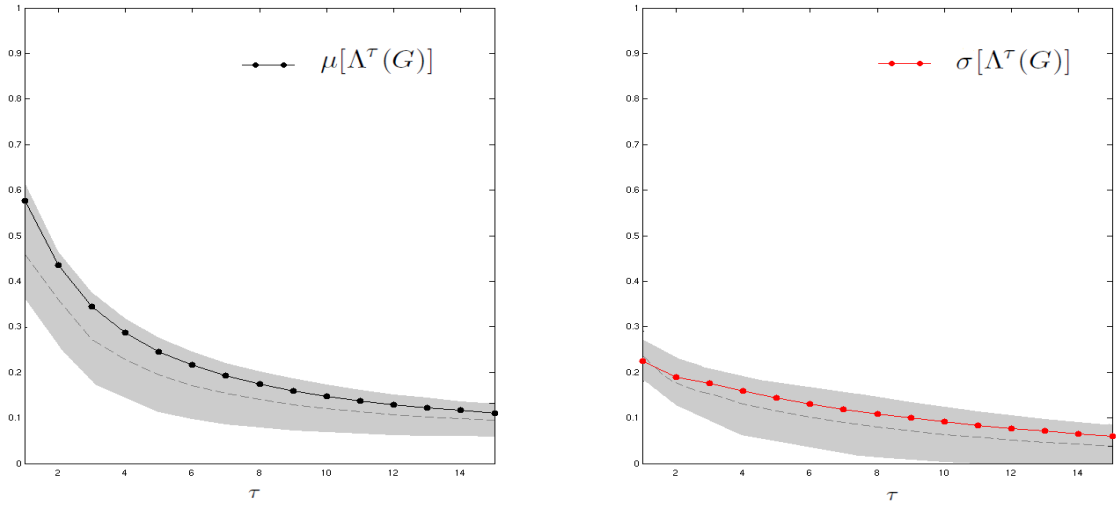


Figure 5.15: The MSS and VSS of 10.000 connected networks uniform undirected valued networks conditioned to the row marginal of the AM.

The following section propose alternative probabilistic modeling for "strange networks", i.e. networks exhibiting structural peculiarities, which are unlikely to appear under general random models, as in the case of many observed social networks.

5.5 Exponential family of random networks

An important limitation of conditionally uniform models is that the rejection of the null hypothesis does not provide a clear insight for the construction of meaningful probabilistic model for the phenomenon being studied. The only conclusion of a conditionally uniform model is that the observed value for the test statistic is unlikely given the conditioning statistics if all else would be random. Because of these limitations non-uniform model have been introduced for the analysis of CNs.

Exponential Random Graph Models (ERGM) has been widely used to capture the fact that the probability of a network should reflect its structural properties. Here we are considering Conditionally Exponential Random Graph Models (CERGOM), which keeps the main probabilistic properties of the classical ERGM into a constrained sample space χ, as shown in the next section. Highly probable networks are regarded as "reliable" in the sense of being consistent with the real-world scenario captured by the specified probabilistic model, so that the problem of designing reliable networks is here translated into the one of maximizing graph probability under conditionally exponential models.

To have a first insight to the ERGM, let χ be the sample space of all networks belonging to a specified family and consider a collection of independent and identically distributed networks $\mathbf{x}_1, \ldots, \mathbf{x}_N \sim p$. Let $S_j(\mathbf{x}_i)$, for $j = 1, \ldots, s$, be a structural feature of network $\mathbf{x}_i$, for $i = 1, \ldots, N$, and $\widehat{\mu}_j = \sum_{i=1}^{N} S_j(\mathbf{x}_i)/N$ the empirical expectation of S_j, for $j = 1, \ldots, s$.

The ERGM arises as an answer to the question *can we recover the* p^* *from* $\widehat{\mu}_1, \ldots, \widehat{\mu}_s$? A reasonable requirement a probability measure p must verify is that

$$\begin{aligned} \mathbb{E}_p[S_j(\mathbf{x})] = \textstyle\int_\chi S_j(\mathbf{x})p(\mathbf{x})dx = \widehat{\mu}_j, \qquad f = 1, \ldots, s. \\ \text{where } \chi = \{\mathbf{x} \in \{0,1\}^N \ : \ A\mathbf{x} = \mathbf{b}\}. \end{aligned} \tag{5.31}$$

(We are using the Lebesgue integral as a generalization of summation for continuous spaces, that is to say, spaces of valued networks, where χ is the set of valued AMs verifying a specified system linear constraints. In the case considered in this section the integral can be replaced by a summation.)

The problem is to find the probability distribution $p(\mathbf{x})$ which verifies $\mathbb{E}_p[S_j(\mathbf{x})] = \widehat{\mu}_j$ and reflects maximal ignorance on $p(\mathbf{x})$. In thermodynamics, S is usually assumed to be the temperature of a closed system and χ is the set of possible configuration that system might take. The probability of energy levels of a closed system is derived by assuming (5.31) and choosing the $p(\mathbf{x})$ which maximizes the entropy,classically interpreted in Information Theory and Thermodynamics as a measure of uncertainty and disorder. In statistics, this approach to deduce an appropriate functional form for the probability distribution from a set of given observations has been studied first introduced by Jaynes [126], who assume certain highly plausible properties of the random variable and then posit maximum entropy with respect to all other properties of the distribution. It leads to pick the distribution $\widehat{p(\mathbf{x})}$ which solves the following problem:

$$\max \quad -\int_{\chi} p(\mathbf{x}) \log p(\mathbf{x}) dx \tag{5.32a}$$

$$\text{subject to} \quad \int_{\chi} S_j(\mathbf{x}) p(\mathbf{x}) dx = \widehat{\mu}_j, \qquad f = 1, \ldots, s \tag{5.32b}$$

$$\int_{\chi} p(\mathbf{x}) dx = 1. \tag{5.32c}$$

The classical method for solving problem (5.32) is to apply the Lagrange multipliers to each of the constraints and maximize the augmented functional with respect to $p(\mathbf{x})$.

$$L(p, \boldsymbol{\theta}, \alpha) = -\int_{\chi} p(\mathbf{x}) \log p(\mathbf{x}) dx + \sum_{j=0}^{s} \theta_j \left(\int_{\chi} S_j(\mathbf{x}) p(\mathbf{x}) dx - \widehat{\mu}_j \right) + \alpha \left(\int_{\chi} p(\mathbf{x}) dx - 1 \right) \tag{5.33}$$

By applying the Euler equation of calculus of variation to the lagrangian function, we get in (5.34) the functional form of a ERGM, where χ becomes the set of all simple graphs.

$$\frac{\partial L}{\partial p(\mathbf{x})} = -\log(p(\mathbf{x})) - 1 + \sum_{j=0}^{s} \theta_j S_j(\mathbf{x}) + \alpha = 0 \tag{5.34a}$$

$$\widehat{p(\mathbf{x})} = \exp\left(\sum_{j=0}^{s} \theta_j S_j(\mathbf{x}) \right) \exp(1 - \alpha) = 0 \tag{5.34b}$$

$$\widehat{p(\mathbf{x})} = \frac{\exp(\boldsymbol{\theta} \mathbf{S}(\mathbf{x}))}{Z(\boldsymbol{\theta}, \boldsymbol{\chi})} \tag{5.34c}$$

The boldface symbols $\boldsymbol{\theta}$ and $\mathbf{S}(\mathbf{x})$ denote the vectors $[\theta_1, \ldots, \theta_s]^T$ and $[S_1(\mathbf{x}), \ldots, S_s(\mathbf{x})]^T$ respectively. The parameter θ_j is the Lagrangian multiplier of the constraints $\mathbb{E}_p[S_j(\mathbf{x})] = \widehat{\mu}_j, j = 1, \ldots, s$. It controls the tendency of networks with parameters $S_j(\mathbf{x})$ to be observed in the data. Hence, the exponential family of distributions is an example of a parameterized set of distributions where the Lagrange multipliers play the role of the parameters of the distribution (another example of this kind is the class of mixture distributions). The scalar quantity $\kappa = \int_{\mathbf{x} \in \boldsymbol{\chi}} e^{\sum_{j=1}^{s} \theta_j S_j(\mathbf{x})} d\mathbf{x}$ is known as the partition function. Many papers from statistical physics call the quantity $\sum_{j=1}^{s} \theta_j S_j(\mathbf{x})$ the *graph Hamiltonian* and denote it as $H(\mathbf{x})$.

The just described CERGM represents a generalization of the classical ERGM, where the sample space χ can be arbitrarily defined. If χ is the set of all simple graphs, then $|\chi| = 2^{\binom{n}{2}}$. If $\boldsymbol{\chi}$ is the set of all undirected graphs with fixed number of edges d, then $|\chi| = \binom{n(n-1)/2}{d}$.

The Erdos and Rnyi model, we introduced in the previous section, is indeed an ERGM with $S(\mathbf{x}) = \sum_{i,j \in V} x_{ij}$. If π is the probability that an edge is observe, then the Erdos and Rnyi model in ERGM form is given in 5.35.

$$p(\mathbf{x}) = \pi^{(\sum \mathbf{x})}(1-\pi)^{(\binom{n}{2} - \sum \mathbf{x})} = \tag{5.35a}$$

$$= \exp\left(\log(\pi/1-\pi)\sum \mathbf{x} + \binom{n}{2}\log(1-\pi)\right) = \tag{5.35b}$$

$$= (1-\pi)^{\binom{n}{2}} \exp\left(\log(\pi/1-\pi)\sum \mathbf{x}\right) = \tag{5.35c}$$

$$= \frac{\exp\left(\boldsymbol{\theta}\mathbf{S}(\mathbf{x})\right)}{Z(\boldsymbol{\theta}, \boldsymbol{\chi})} \tag{5.35d}$$

The partition function and the natural parameter are $1/Z(\boldsymbol{\theta}, \boldsymbol{\chi}) = (1-\pi)^{\binom{|V|}{2}}$ and $\theta = \log(\pi/1-\pi)$, respectively. The ensemble of the Bernoulli model may be extended to allow multiple edge graphs, or it may be changed to directed graphs, and the solubility is not lost.

We say that a ERG model has an exact solution if the partition function $Z(\theta, \chi)$ can be computed without summing over all the ensembles of networks in the family χ. In theory, to use an ERG model, it is always possible to work directly from its defining equations (5.34), but one can easily realize that this is computationally impractical when the summation in $Z(\theta, \chi)$ involves $2^{\binom{n}{2}}$ terms. For instance, the ensemble in a binary graph with 30 the exact evaluation of $Z(\theta, \chi)$ involves more than 10^{130} terms.

One of the convenient ways to approximate these random draws is by Monte Carlo simulation, such as Gibbs sampling and Metropolis-Hastings algorithm [187, 197].

The use of the Metropolis–Hastings algorithm to draw samples from the CERGMs is a straightforward extension of the conditionally uniform case presented in Section 5.3. Consider the probability density function of primal-dual solutions (5.15). Given an initial arbitrary primal-dual solution $\mathbf{t}^0 = (\widehat{\mathbf{x}}^0, \widehat{\mathbf{y}}^0, \widehat{\mathbf{z}}^0)$, the following rule is applied to go from a current state $\mathbf{t}^k = (\widehat{\mathbf{x}}^k, \widehat{\mathbf{y}}^k, \widehat{\mathbf{z}}^k)$ to a new state $\mathbf{t}^{k+1} = (\widehat{\mathbf{x}}^{k+1}, \widehat{\mathbf{y}}^{k+1}, \widehat{\mathbf{z}}^{k+1})$:

1. Initialization: choose an arbitrary point $\mathbf{t}^0$ to be the first sample and let (5.15) be the proposal distribution, which suggests a candidate for the next sample value $\mathbf{t}^{k+1}$, given the previous sample value $\mathbf{t}^k$. (In the case of (5.15) the transaction probabilities are independents.)

2. Candidate state: propose a candidate point $\mathbf{t}^{k+1}$ from a proposal distribution, (5.15) that may depend on the current $\mathbf{t}^k$. (In our case the proposal entails the solution of a LP.)

3. Acceptance criterion: accept the candidate state with probability

$$\alpha(\mathbf{t}^k, \mathbf{t}^{k+1}) = \min\left(1, \frac{f_{\mathcal{I}_D}(\mathbf{t}^k)}{f_{\mathcal{I}_D}(\mathbf{t}^{k+1})} \frac{\exp(\boldsymbol{\theta}\mathbf{S}(\mathbf{t}^{k+1}))}{\exp(\boldsymbol{\theta}\mathbf{S}(\mathbf{t}^k))}\right). \tag{5.36}$$

As in the conditionally uniform case analyzed in Section 5.3, a candidate value is first generated by solving an LP and the acceptance probability is then computed to update the chain. The only differentiation is the definition of the acceptance probability 5.36.

The proof of the correctness of the Metropolis-Hastings chain (its convergence to the desired CERGM) is easily obtained by detailed-balance:

$$P(\mathbf{t}^{k+1} \mid \mathbf{t}^k)\frac{\exp\left(\boldsymbol{\theta}\mathbf{S}(\mathbf{t}^k)\right)}{Z(\boldsymbol{\theta},\boldsymbol{\chi})} = P(\mathbf{t}^k \mid \mathbf{t}^{k+1})\frac{\exp\left(\boldsymbol{\theta}\mathbf{S}(\mathbf{t}^{k+1})\right)}{Z(\boldsymbol{\theta},\boldsymbol{\chi})} \tag{5.37}$$

for all $\mathbf{t}^k, \mathbf{t}^{k+1} \in \mathcal{I}_D$, where $P(\mathbf{t}^{k+1} \mid \mathbf{t}^k)$ is not has been defined in (5.29). When $\mathbf{t}^k = \mathbf{t}^{k+1}$ detailed-balance obviously holds, and for $\mathbf{t}^k \neq \mathbf{t}^{k+1}$ the only way to arrive to $\mathbf{t}^{k+1}$ is by accepting it as a proposed candidate. Thus, (5.37) just says that

$$\alpha(\mathbf{t}^k, \mathbf{t}^{k+1}) f_{\mathcal{I}_D}(\mathbf{t}^{k+1})\frac{\exp\left(\boldsymbol{\theta}\mathbf{S}(\mathbf{t}^k)\right)}{Z(\boldsymbol{\theta},\boldsymbol{\chi})} = \alpha(\mathbf{t}^{k+1}, \mathbf{t}^k) f_{\mathcal{I}_D}(\mathbf{t}^k)\frac{\exp\left(\boldsymbol{\theta}\mathbf{S}(\mathbf{t}^{k+1})\right)}{Z(\boldsymbol{\theta},\boldsymbol{\chi})} \tag{5.38}$$

which is true, since when $\alpha(\mathbf{t}^{k+1}, \mathbf{t}^k)$ on one side is 1, $\alpha(\mathbf{t}^k, \mathbf{t}^{k+1})$ on the other side isn't and the denominator cancels with the term outside the parenthesis producing the equality of both sides.

5.5.1 Maximizing graph probability under conditionally exponential models

Simulating random networks consists in generating networks whose probability of being selected is specified by a well-defined probabilistic model. A different case of network generation is when the selected networks are only the ones which maximize the probability in the specified model. In this case we are not interested in whatever bunch of networks which reflect the probability distribution, but exactly in the probability maximizers. Networks maximizing the probability of being selected are regarded as reliable under the specified probability distribution, in the sense that they reproduce the required topological features, captured by the probabilistic model[4].

In this section we are studying the reliability maximization, based on the CERGM, whose statistical properties has been just described to capture complex topological features of real-world networks. This analysis allows casting a mathematical bridge between probabilistic and optimization based models of network formation, and introduces the microeconomic problems of Chapter 6[5].

Finding a reliable network under the specified CERGM consist in maximizing $\kappa \exp(\boldsymbol{\theta}\mathbf{S}(\mathbf{x}))$, subject to $\mathbf{x} \in \chi$. The highly combinatorial nature of these problems require to translate them into (or reformulate them by) solvable systems of linear constraints.

Consider the set χ of undirected networks with fixed number of edges d and an exponential model on the sample space χ, with graph Hamiltonian $H(\mathbf{x}) = \theta \sum_{i<j<k} x_{ij}x_{jk}x_{ij}$, representing the number of closed triangles in the network[6]. Maximizing the logarithm of the probability

[4]Designing reliable networks generally consists in finding topological structures, which are able to successfully carry out desired processes and operations, captured by a probabilistic model. When this set of activities performed within a network are unknown and the only available information is a probabilistic model reflecting topological network features, highly probable networks are regarded as "reliable", in the sense of being consistent with those probabilistic models.

[5]Optimization based models of network formation allows takeing into account the emerging properties from the point of a global planner, who wish to allocate connections among nodes in such a way as to successfully carry out processes and operations performed within the networks.

[6]For particularly small d (high sparsity) an optimal network consist of a fully connected subgraph and several disconnected nodes. One possibility to overcome this drawback is either to include more network properties in the graph Hamiltonian or to redefine the sample space χ in such a way that the unwanted trivial solutions are discarded, for example by forcing network connectivity. It has been shown in Section 2.4 that a straightforward way to algebraically force connectivity is to require the existence of a flow circulating within the network, from one node to all the others. Thus, we make use of an artificial flow of $n-1$ units, departing from one node $h \in \mathcal{V}$ and arriving to each of the $n-1$ remaining nodes. The existence of such flow is a sufficient and necessary condition for the network to be connected.

inside the specified sample space leads to the non-linear binary problem

$$\begin{aligned} \max \quad & \sum_{i<j<k} x_{ij}x_{jk}x_{ij} \\ \text{s. to} \quad & \sum_{i<j} x_{ij} = d \\ & x_{ij} \in \{0,1\} \ (i,j) \in \mathcal{H}^2. \end{aligned} \tag{5.39}$$

Let w_{ijk} be the binary indicator of the (i,j,k) triad, which is equal to one if a closed triangle between i, j and k exists, and zero otherwise. A system of linear constraints used to characterize the state of w_{ijk} is (2.24b), as described in Section 2.4. By introducing (2.24b) in (5.39) and replacing the objective function $\sum_{i<j<k} x_{ij}x_{jk}x_{ij}$ with $\sum_{(i,j,k)\in\mathcal{H}^3} w_{ijk}$, the maximization of the graph probability under the specified model becomes a linear program in binary variables, which can be solved up to optimality by standard integer programming technics [209, 210].

Consider a CERGM whose specification is given by the following graph Hamiltonian and sample space:

$$\begin{aligned} H(x) = \theta_1 S_1(\mathbf{x}) + \theta_2 S_2(\mathbf{x}) + \theta_3 S_3(\mathbf{x}) \\ \text{where } \mathbf{x} \text{ verifies (2.24}b\text{).} \end{aligned} \tag{5.40}$$

The three network statistics are the following:

- $S_1(\mathbf{x}) = \sum_{(i,j,k)\in\mathcal{H}^3} w_{ijk}$;
- $S_2(\mathbf{x}) = \sum_{(i,j)\in\mathcal{H}^2} x_{ij}$;
- $S_3(\mathbf{x}) = \max_{\mathbf{v},\mathbf{z}} \sum_{i\in\mathcal{V}} u^i(\mathbf{v},\mathbf{z})$

  subject to

$$\begin{aligned} & \sum_{k:(k,u)\in\mathcal{E}} z_j^{ku} - \sum_{h:(u,h)\in\mathcal{E}} z_j^{uh} = v_j^u - q_j^u, && j \in \mathcal{C},\ u \in \mathcal{V} \\ & \sum_{j\in\mathcal{C}} z_j^{uh} \le M x_{uh} && (u,h) \in \mathcal{H}^2 \end{aligned}$$

The network statistics $S_3(\mathbf{x})$ represents the aggregate utility of nodes (agents), associated to the ownership of m types of commodities from a reallocation of their initial endowments $\mathbf{q}$. The set $\mathcal{C}$ is a collection of m types of commodities, Ξ_i a commodity space, representing the feasible bundle of commodity that agent $i \in \mathcal{V}$ may hold. The initial endowments $q_j^i \in \Xi_i$ of agent $i \in \mathcal{V}$ for each commodity $j \in \mathcal{C}$ are fixed quantities and the utility functions are defined on the set of all possible final allocations and flow of reallocations between agents. Since no production is allowed, the aggregated stock of commodities stays constant, as expressed by the balance equations of flow circulating through the network. (See Subsection 1.1.4 for more details about multicommodity network flow problems.)

Thus, the network statistics $S_3(\mathbf{x})$ is defined as the optimal value of an optimization problem on the feasible region of a multicommodity network flow problem with decision variables representing the nodal demands and the circulating flows of commodities (v_j^i and z_j^{hk}, for $j \in \mathcal{C}$, $i \in \mathcal{V}$, $(u,h) \in \mathcal{H}^2$) .

Simulating by the Metropolis-Hastings algorithm a bunch of networks from a CERGM with the graph Hamiltonian and sample space in (5.40) entails the iterative solution of multicommodity network flow problems to evaluate $S_3(\mathbf{x})$. If our interest is in the reliability maximization, a much bigger problem must be solved to optimize (5.40) with respect to $\mathbf{x}$. A detailed investigation of this subject is out of the scope of this thesis and must be regarded as an open problem to be addressed in future researches.

Part III

Strategic models of network formation

Abstract

In the social sciences there is a standing debate over the primacy of structure or agency in shaping human behavior. Agency is the capacity of individuals to act independently and to make their own free choices. Structure is the recurrent patterned arrangements which influence or limit the choices and opportunities available. This part of the thesis studies the arise of macroscopic properties of the social structure from the microeconomical individual interaction. The main purpose is to place the problem of microeconomical interaction on a general and flexible mathematical programming framework, related with the previous analysis of random models and strategic models of network formation. Our discussion in Chapter 1 suggested a general view of closed network structures. We will numerically show by multi-agent simulation the economical reasons behind the emergence of these structures.

Keywords: Complex Networks, Microeconomics, Multi-agent systems, Combinatorial Optimization, Simulation and numerical modeling.

Chapter 6

Mathematical programming approaches for different scenarios of bilateral bartering

6.1 Social capital and the economic effect of the interaction structure

Our discussion in Chapter 1 suggests a general view of the social structure as a form of capital, which facilitates access to goods and services. From this outlook, Coleman [68] spoke of social capital alongside physical capital, suggesting an aggregate of durable commodities which is worth for the acquisition and production of other commodities (capital as a mean of production).

From a sociological point of view, this idea of social capital strongly supports a structuralist understanding of society as prior to individuals, by stressing the one-way effect of the network of interpersonal relationships on the realization of economically valuable processes and tasks and on the emergence of macroscopic social phenomena. The analysis of the effect of the structure of interactions constitute a prominent line of research in social sciences.

As suggested in the Subsection 1.1.4 one of the seminal multi-agent system for this kind of analysis is the Axelrod's model of the dissemination of culture [6], where the discrete time dynamic evolution of cultural features strongly depends on the structure of the topological feature of the interaction structure.

Another interesting study in this direction has been conducted by Masuda and Aihara [157], who assumed that each vertex of a network, which ranges from a regular lattice to a random graph, by changing a re-writing probability π, is occupied by a player of a prisoner's dilemma game. In each round, a player interacts with its immediate neighbors. Each player has two strategies: cooperation (denoted by **C**) or defection (denoted by **D**) in each round. When a player chooses **C**, it receives payoff R (reward) or S (sucker) as the opponent chooses **C** or **D**, respectively. A player that chooses **D** receives T (temptation) or P (punishment) as the opponent chooses **C** or **D**, respectively. Given $T > R > P > S$, a player is always tempted to defect no matter whether the opponent takes **C** or **D**. The combination of **D** and **D**, in which both of the players get unsatisfactory payoff P, is the unique Nash equilibrium in a single game. Each player sums the payoff received by playing a single Prisoner Dilemma with its neighbors and compares the sum with those of the neighbors. Among them, the strategy with the maximal

payoff is copied as the players strategy in the next round. As Masuda noted, for small reward, it is not so tempting for players to defect. Therefore, the proportion of cooperators converges to 1 regardless of the network structure. However, the number of cooperators highly depends on π for $1.3 < T < 2.3$. In this case, the clustering needs to be larger for cooperators to survive. For large T, players are inclined to betray. Even if cooperators happen to form tight clusters, they cannot survive once they face defectors and the cooperators extinguish for whatever network structure. Thus, three different dynamics and dependence on the network structure exist with respect different value of temptation. For intermediate reward T the dependence on the network structure is remarkable. Among all networks, small-world topology, obtained for $0.4 < \pi < 0.6$, is the optimal structure when we take into account the speed at which cooperative behavior propagates, which may explain why the small-world properties are self-organized in real networks.

As already mentioned in the Subsection 1.1.4, the one-way effect of the network of interpersonal relationships on the realization of economically valuable processes has also been studied by Wilhite [229], who considered the price volatility and convergence time of a simple exchange economy where the structure of bilateral trades varies.

The approach adopted in this chapter is based on the opposite consideration: the ways in which individual actions and economic processes support the emergency of network structures and pattern of interaction is analyzed. This idea has been widely discussed among Marxist sociologists by referring to the modes of production and the contrast between basis and superstructure [59]. The basis comprehends the relations of production - employer-employee work conditions, the technical division of labor, and property relations - into which people enter to produce the necessities and amenities of life.

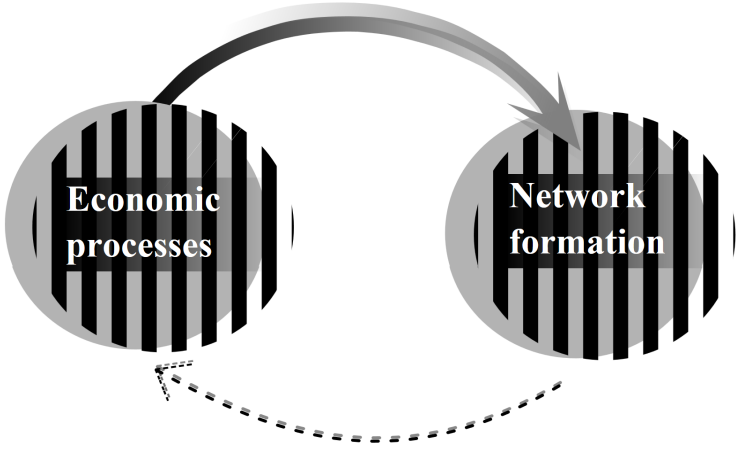


Figure 6.1: From agency to structure.

In Marx' thinking these relations fundamentally determine society's other relationships like common knowledge, mutual influence, norms, constituting the superstructure; thus, the base conditions the superstructure, though their relation is not strictly causal, because the superstructure often influences the base; however the influence of the base predominates. However, from a pure Marxist outlook, this economic base of society is seen as determining everything else in the superstructure, including social, political and intellectual consciousness. Marx postulated the theoretic essentials of the base-superstructure concept in the Preface to *A Contribution to the Critique of Political Economy* [156]:

> *In the social production of their existence, men inevitably enter into definite relations,*

> *which are independent of their will, namely relations of production appropriate to a given stage in the development of their material forces of production. The totality of these relations of production constitutes the economic structure of society, the real foundation, on which arises a legal and political superstructure and to which correspond definite forms of social consciousness. The mode of production of material life conditions the general process of social, political and intellectual life. It is not the consciousness of men that determines their existence, but their social existence that determines their consciousness. At a certain stage of development, the material productive forces of society come into conflict with the existing relations of production or this merely expresses the same thing in legal terms with the property relations within the framework of which they have operated hitherto. From forms of development of the productive forces these relations turn into their fetters.*

Marxist historical materialism conceives epochs as characterized by patterns of social interactions people must enter into in order to satisfy their economic needs and carry out their purposes; such a pattern are called *relations of production*.

Many studies of computer simulation [185] have attempted to reproduce the process of interactions which leads to the emergence of any sort of superstructure like common knowledge, mutual influence, norms, inequalities, etc.

Our approach to mathematically deal with this problem is based on the economic interactions between rational agents aiming at the maximization. To do so several processes of bartering private goods with exogenous prices are investigated in this chapter, both from the local viewpoint of agents and from the perspective of a global planner.

Consider a collection of m types of commodities, call i $\mathcal{C}$, a commodity space Ξ_i (representing the feasible bundle of commodity that agent $i \in \mathcal{V}$ may hold and usually given by a subset of the nonnegative orthant in $\mathbb{R}^m$), the initial endowments $q_j^i \in \Xi_i$ of agent $i \in \mathcal{V}$ for each commodity $j \in \mathcal{C}$ (corresponding a budget constraints), utility functions $u^i : \Xi \longmapsto \mathbb{R}$, representing preference relation $\preceq_i$ on $\mathcal{Y}$, where $\Xi = \Xi_1 \times \ldots \times \Xi_n$.

When agents attempt to simultaneously maximize their respective utilities, conditioned to balance constraints, the resulting problems are $\max u^i(\mathbf{v})$ s.to $\sum_{i\in\mathcal{V}} v_j^i = \sum_{i\in\mathcal{V}} q_j^i$ for $j \in \mathcal{C}$, where v_j^i, is the amount of commodity j demanded by agent i (from now on the superindex shall denote the agent and the subindex shall denote the commodity).

Arrow and Debreu [80] showed that under certain economic conditions (convex preferences, perfect competition and demand independence) there must be a vector of prices $\widehat{P} = (\widehat{p}_1, \widehat{p}_2, \widehat{p}_3, \ldots, \widehat{p}_m)^T$, such that aggregate supplies will equal aggregate demands for every commodity in the economy.

As studied by Dreze [83], when prices are regarded as fixed, markets do not clear and the imbalance between supply and demand is resolved by some kind of quantity rationing. This models have played an important role in maroeconomic models, especially on those models related to wage rigidities and unemployment. Under fixed prices, markets do not clear and the imbalance between supply and demand is resolved by some kind of quantity rationing [83]. In out analysis this quantity rationing is implicit in the process and not explicitly taken into account.

Another specific scenario of market of private goods has been studied by Shapley and Shubik [211], who characterized the equilibria under the assumption that each agent can consume at most one indivisible good. After them, many authors have been studying markets with indivisible goods (see for example, Kaneko [131], Quinzii [193], Scarf [207], and the most recent literature like Danilov et al. [76], Caplin and Leahy [48]). The main focus was to address the question of existence of market clearing prices in the cases of not infinitesimally divisible allocations.

This chapter provides mathematical-programming based approaches for the analysis of markets of private goods, with particular attention to the dynamics of bartering and the effects of the coevolution of agents economical states and social structure. Particularly, our goal is to

provide novel mathematical-programming based approaches to study barter processes, which are commonly used in everyday life by economic agents to solve bargaining problems associated to n-consumer-m-commodity markets of integer commodities and fixed exogenous prices.

6.2 Markets with fixed exogenous prices

As already discussed in Section 1.2, the linear system characterizing the space of possible allocations is (1.26). Here the conservation of commodity (i.e., the overall amount of commodity of each type must be preserved) is generalized to include arbitrary weights in the last m rows of (1.26). Based on this observation consider the following multi-objective integer non-linear optimization problem (MINOP)

$$\max \quad [u^i(\mathbf{v}),\ i = 1, \dots, n] \tag{6.1a}$$

s. to

$$\begin{bmatrix} P & & & \\ & P & & \\ & & \ddots & \\ & & & P \\ d^1 I & d^2 I & \dots & d^n I \end{bmatrix} \mathbf{v} = \begin{bmatrix} b^1 \\ b^2 \\ \vdots \\ b^n \\ \mathbf{b}^0 \end{bmatrix} \tag{6.1b}$$

$$\begin{array}{l} u^i(\mathbf{v}) \geq u^i(\mathbf{q}) \quad i = 1 \dots, n \\ \mathbf{v} \in \mathbb{Z}^{mn} \geq 0, \end{array} \tag{6.1c}$$

where $u^i : \mathbb{R}^{mn} \to \mathbb{R}$, $P \in \mathbb{Q}^{1\times m}$, $d^i \in \mathbb{Q}$, $b^i \in \mathbb{Q}$, $i = 1, \dots, n$, and $\mathbf{b}^0 \in \mathbb{Q}^m$. The conditions $u^i(\mathbf{v}) \geq u^i(\mathbf{q}), \quad i = 1 \dots, n$, guarantee that no agent gets worse under a feasible reallocation, which is known in general bargaining literature as the *disagreement point*. The constraint matrix has a primal block-angular structure with n identical diagonal blocks involving m decision variables. The set of non negative solution of (1.26) coincide with the feasible region of (6.1) for $d_i = 1$, for $i = 1, \dots, n$.

A specialized interior point method for LPs with primal block-angular structure has been described and investigated in Chapter 4. In Appendix C the same factorization principle is used to deal with markets with exogenous prices.

6.2.1 The elementary reallocation problem

In everyday life, barter processes among people tends to achieve the Pareto frontier of problem (6.1) by a sequence of reallocations. We consider a process based on a sequence of two-commodity-two-agent reallocations, denoted as SER. Any step of this sequence requires the solution of a MINOP involving 4 variables and 4 constraints of problem (6.1).

Let $\mathbf{e}$ be a feasible solution of (6.1b) and (6.1c) and suppose we want to produce a feasible change of 4 variables, such that 2 of them belong to the ith and jth position of the diagonal block h and the other belong to the ith and jth position of the diagonal block k.

It can be easily shown that a feasibility condition of any affine change of these 4 variables $e^h_i + \Delta^h_i, e^k_i + \Delta^k_i, e^h_j + \Delta^h_j, e^k_j + \Delta^k_j$ is that $\Delta^h_i, \Delta^k_i, \Delta^h_j, \Delta^k_j$ must be an integer solution of the

following system of equations

$$\begin{bmatrix} p_i & p_j & 0 & 0 \\ 0 & 0 & p_i & p_j \\ d^h & 0 & d^k & 0 \\ 0 & d^h & 0 & d^k \end{bmatrix} \begin{bmatrix} \Delta_i^h \\ \Delta_j^h \\ \Delta_i^k \\ \Delta_j^k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{6.2}$$

The solution set are the integer points in the null space of the matrix of system (6.2), which will be named A. A is a two-agent-two-commodity constraint matrix, and its rank is three (just note that the first column is a linear combination of the other three using coefficients $\alpha_2 = \frac{p_i}{p_j}$, $\alpha_3 = \frac{d^h}{d^k}$ and $\alpha_4 = -\frac{p_i d^h}{p_j d^k}$). Therefore the null space has dimension one, and its integer solutions are found on the line

$$\begin{bmatrix} \Delta_i^h \\ \Delta_j^h \\ \Delta_i^k \\ \Delta_j^k \end{bmatrix} = q \begin{bmatrix} p_j d^k \\ -p_i d^k \\ -p_j d^h \\ p_i d^h \end{bmatrix}, \tag{6.3}$$

for some $q = \alpha F(p_i, p_j, d^k, d^h)$, where $\alpha \in \mathbb{Z}$ and $F : \mathbb{Q}^4 \to \mathbb{Q}$ provides a factor which transforms the null space direction in the nonzero integer null space direction of smallest norm. We note that this factor can be computed as $F(p_i, p_j, d^k, d^h) = G(p_j d^k, p_i d^k, p_j d^h, p_i d^h)$, where

$$G(v_i = \frac{r_i}{q_i}, i = 1, \ldots, l) = \frac{\text{lcm}(q_i, i = 1, \ldots, l)}{\text{gcd}(\text{lcm}(q_i, i = 1, \ldots, l) \cdot v_i, i = 1, \ldots, l)}, \tag{6.4}$$

r_i and q_i being the numerator and denominator of v_i ($q_i = 1$ if v_i is integer), and lcm and gcd being, respectively, the least common multiple and greatest common divisor functions.

Hence, given a feasible point $\mathbf{e}$, one can choose 4 variables, such that 2 of them belong to the ith and jth position of a diagonal block h and the others belong to the ith and jth position of a diagonal block k, in $m(m-1)n(n-1)/4$ ways. Each of them constitutes an ERP, whose Pareto frontier is in $\mathbf{q} + \text{null}(A)$. The SER is a local search, which repeatedly explores a neighborhood and chooses both a locally improving direction among the $m(m-1)n(n-1)/4$ possible ERPs and a feasible step length $q = \alpha F(p_i, p_j, d^k, d^h)$, $\alpha \in \mathbb{Z}$. For problems of the form of (6.1) the SER might be written as follows:

$$\mathbf{v}^{t+1} = \mathbf{v}^t + \alpha F(p_i, p_j, d^k, d^h) \begin{bmatrix} \vdots \\ p_j d^k \\ \vdots \\ -p_i d^k \\ \vdots \\ -p_j d^h \\ \vdots \\ p_i d^h \\ \vdots \end{bmatrix} \begin{matrix} \vdots \\ h, i \\ \vdots \\ h, j \\ \vdots \\ k, i \\ \vdots \\ k, j \\ \vdots \end{matrix} = \mathbf{v}^t + \alpha F(p_i, p_j, d^k, d^h) \Delta_{ij}^{kh}, \tag{6.5}$$

t being the iteration counter. In shorter notation, we write (6.5) as $\mathbf{v}^{t+1} = \mathbf{v}^t + \alpha S_{ij}^{kh}$, where

$$S_{ij}^{kh} = F(p_i, p_j, d^k, d^h) \Delta_{ij}^{kh} \tag{6.6}$$

is a direction of integer components. Since the nonnegativity of $\mathbf{v}$ have to be kept along the iterations, then we have that

$$-\frac{\max\left\{x_i^h/(p_j d^k), x_j^k/(p_i d^h)\right\}}{F(p_i,p_j,d^k,d^h)} \leq \alpha \leq \frac{\min\left\{x_j^h/(p_i d^k), x_i^k/(p_j d^h)\right\}}{F(p_i,p_j,d^k,d^h)}, \tag{6.7}$$

or, equivalently,

$$-\max\left\{x_i^h/(p_j d^k), x_j^k/(p_i d^h)\right\} \leq q \leq \min\left\{x_j^h/(p_i d^k), x_i^k/(p_j d^h)\right\}. \tag{6.8}$$

(The step length is forced to be nonnegative when the direction is both feasible and a descent direction; in our case the direction is only known to be feasible, and then negative step lengths are also considered.)

An important property of an elementary reallocation is that under the assumptions that $\frac{\partial u^k(\mathbf{v})}{\partial_{x_i^k}} : \mathbb{R}^{mn} \to \mathbb{R}$ is (i) non increasing, (ii) nonnegative and (iii) $\frac{\partial u^k(\mathbf{v})}{\partial_{x_i^j}} = 0$ for $j \neq k$ (i.e., u^k only depends on $\mathbf{v}^k$), which are quite reasonable requirements for consumer utilities, then $u^k(\mathbf{v} + \alpha S_{ij}^{kh})$ is a unimodal function with respect to α, as shown by the next proposition.

Proposition 11. *Under the definition of u^k and S_{ij}^{kh}, for every feasible point $\mathbf{v} \in R^{mn}$, $u^k(\mathbf{v} + \alpha S_{ij}^{kh})$ is a unimodal function with respect to α in the interval defined by (6.7).*

Proof. Let us define $g(\alpha) = u^k(\mathbf{v} + \alpha S_{ij}^{kh})$, differentiable with respect to α. It will be shown that for all α in the interval (6.7), and $0 < \tau \in \mathbb{R}$, $g'(\alpha) < 0$ implies $g'(\alpha + \tau) < 0$, which is a sufficient condition for the unimodality of $g(\alpha)$. By the chain rule, and using (6.5) and (6.6), the derivative of $g(\alpha)$ can be written as

$$\begin{aligned} g'(\alpha) &= \nabla_{\mathbf{v}} u^k(\mathbf{v} + \alpha S_{ij}^{kh}) S_{ij}^{kh} \\ &= F(p_i,p_j,d^k,d^h)\left(\frac{\partial u^k(\mathbf{v} + \alpha S_{ij}^{kh})}{\partial_{x_i^k}}(-p_j d^h) + \frac{\partial u^k(\mathbf{v} + \alpha S_{ij}^{kh})}{\partial_{x_j^k}} p_i d^h\right). \end{aligned} \tag{6.9}$$

If $g'(\alpha) < 0$ then, from (6.9) and since $F(p_i,p_j,d^k,d^h) > 0$, we have that

$$\frac{\partial u^k(\mathbf{v} + \alpha S_{ij}^{kh})}{\partial_{x_i^k}} p_j d^h > \frac{\partial u^k(\mathbf{v} + \alpha S_{ij}^{kh})}{\partial_{x_j^k}} p_i d^h. \tag{6.10}$$

Since from (6.5) the component (k,i) of S_{ij}^{kh} is $F(p_i,p_j,d^k,d^h)(-p_j d^h) < 0$, and $\frac{\partial u^k(\mathbf{v})}{\partial_{x_i^k}}$ is non increasing, we have that for $\tau > 0$

$$\frac{\partial u^k(\mathbf{v} + (\alpha+\tau) S_{ij}^{kh})}{\partial_{x_i^k}} \geq \frac{\partial u^k(\mathbf{v} + \alpha S_{ij}^{kh})}{\partial_{x_i^k}}. \tag{6.11}$$

Similarly, since the component (k,j) of S_{ij}^{kh} is $F(p_i,p_j,d^k,d^h)(p_i d^h) > 0$, we have

$$\frac{\partial u^k(\mathbf{v} + \alpha S_{ij}^{kh})}{\partial_{x_j^k}} \geq \frac{\partial u^k(\mathbf{v} + (\alpha+\tau) S_{ij}^{kh})}{\partial_{x_j^k}}. \tag{6.12}$$

Multiplying both sides of (6.11) and (6.12) by, respectively, $p_j d^h$ and $p_i d^h$, and connecting the resulting inequalities with (6.10) we have that

$$\frac{\partial u^k(\mathbf{v} + (\alpha+\tau) S_{ij}^{kh})}{\partial_{x_i^k}} p_j d^h > \frac{\partial u^k(\mathbf{v} + (\alpha+\tau) S_{ij}^{kh})}{\partial_{x_j^k}} p_i d^h,$$

which proofs that $g'(\alpha + \tau) < 0$. □

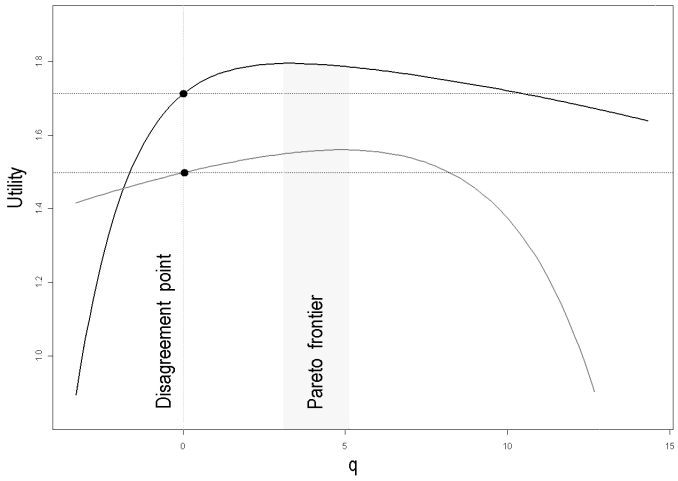


Figure 6.2: Plots of $g^1(\alpha)$ and $g^2(\alpha)$, and interval of α associated to the Pareto frontier. The disagreement point corresponds to $g^1(0)$ and $g^2(0)$, the utilities in the current iterate.

Using Proposition 11 and the characterization of the space of integer solutions of (6.2), we are able to derive a closed expression of the Pareto frontier of the ERP, based on the behavior of $u(\mathbf{v}+\alpha S_{ij}^{kh})$ (see Corollary 1 below), as it is shown in this example:

Example 3. *Consider the following ERP with initial endowments* [40, 188, 142, 66].

$$
\begin{aligned}
\max \quad & [2-e^{-0.051x_1^1}-e^{-0.011x_2^1}, 2-e^{-0.1x_1^2}-e^{-0.031x_2^2}] \\
s.\ to \quad & \\
& 5x_1^1+10x_2^1 = 2080 \\
& 5x_1^2+10x_2^2 = 1370 \\
& \\
& 5x_1^1+6x_1^2 = 1052 \\
& 5x_2^1+6x_2^2 = 1336 \\
& \\
& 2-e^{-0.05x_1^1}-e^{-0.01x_2^1} \geq 1.68 \\
& 2-e^{-0.1x_1^2}-e^{-0.031x_2^2} \geq 1.50 \\
& \\
& x_j^i \geqq 0 \in Z i=1,2; j=1,2;
\end{aligned}
\tag{6.13}
$$

The utility functions $g^1(\alpha)=u^1(\mathbf{v}+\alpha S_{12}^{12})$ *and* $g^2(\alpha)=u^2(\mathbf{v}+\alpha S_{12}^{12})$ *are*

$$
g^1(\alpha)=u^1(\mathbf{v}+\alpha S_{12}^{12}) = u^1\left(\begin{bmatrix}40\\188\\142\\66\end{bmatrix}+\alpha\begin{bmatrix}12\\-6\\-10\\5\end{bmatrix}\right) = 2-e^{-0.051(40+12\alpha)}-e^{-0.011(188-6\alpha)}
$$

$$
g^2(\alpha)=u^2(\mathbf{v}+\alpha S_{12}^{12}) = u^2\left(\begin{bmatrix}40\\188\\142\\66\end{bmatrix}+\alpha\begin{bmatrix}12\\-6\\-10\\5\end{bmatrix}\right) = 2-e^{-0.1(142-10\alpha)}-e^{-0.031(66+5\alpha)},
$$

which are plotted in Figure 6.2. The continuous optimal step lengths for the two respective agents are $\operatorname{argmax} g^1(\alpha)=3.33$ *and* $\operatorname{argmax} g^2(\alpha)=8.94$. *Due to the unimodality of* $u^k(\mathbf{v}+\alpha S_{ij}^{hk})$,

all efficient solutions of (6.13) are given by integer step lengths $\alpha \in [3.33, 8.94]$ *(see Figure 6.2), i.e., for* $\alpha \in \{4, 5, 6, 7, 8\}$ *we have*

$$\begin{array}{lllll} g^1(4) = 1.82412 & g^1(5) = 1.81803 & g^1(6) = 1.80882 & g^1(7) = 1.79752 & g^1(8) = 1.78465, \\ g^2(4) = 1.93043 & g^2(5) = 1.94035 & g^2(6) = 1.94873 & g^2(7) = 1.95558 & g^2(8) = 1.96057. \end{array}$$

Due to the unimodality of both utility functions with respect to α*, no efficient solution exists for an* α *outside the segment* $[3.33, 8.94]$.

The above example illustrates a case where the segment between argmax $u^h(\mathbf{v} + \alpha S_{ij}^{kh})$ and argmax $u^k(\mathbf{v} + S_{ij}^{kh})$ contains five integer points. In this case the efficient solutions of the ERP are the ones associated with these integer step lengths. The following statement gives a constructive characterization of the Pareto frontier of an ERP.

Corollary 1. *Let* $\mathcal{A}$ *be the set of integer points in the interval* $[a^{down}, a^{up}]$*, where* $a^{down} = \min\{\text{argmax}_\alpha u^k(\mathbf{v} + \alpha S_{ij}^{kh}),\ \text{argmax}_\alpha u^h(\mathbf{v} + \alpha S_{ij}^{kh})\}$ *and* $a^{up} = \max\{\text{argmax}_\alpha u^k(\mathbf{v} + \alpha S_{ij}^{kh}),\ \text{argmax}_\alpha u^h(\mathbf{v} + \alpha S_{ij}^{kh})\}$*, and let* $[\alpha^{down}, \alpha^{up}]$ *be the interval of feasible step lengths defined in (6.7). Then, due to Proposition 11, the set* V^* *of Pareto efficient solutions of an ERP can be analytically obtained as follows:*

i. $V^* = \{[u^h(\mathbf{v} + \alpha S_{ij}^{kh}), u^k(\mathbf{v} + \alpha S_{ij}^{kh})] : \alpha \in \mathcal{A}\}$ *if* $\mathcal{A}$ *is not empty and does not contain the zero.*

ii. If $\mathcal{A}$ *is empty and there exists an integer point between* 0 *and* a^{down} *but no integer point between* a^{up} *and* α^{up} *then* V^* *contains the unique point given by* $[u^h(\mathbf{v} + \alpha S_{ij}^{kh}, u^k(\mathbf{v} + \alpha S_{ij}^{kh})]$*such that* α *is the greatest integer between 0 and* a^{down}*.*

iii. If $\mathcal{A}$ *is empty and there exists an integer point between* a^{up} *and* α^{up} *but no integer point between* 0 *and* a^{down} *then* V^* *contains either the unique point given by* $[u^h(\mathbf{v} + \alpha S_{ij}^{kh}, u^k(\mathbf{v} + \alpha S_{ij}^{kh})]$ *such that* α *is the smallest integer between* a^{up} *and* α^{up}*, or* $\alpha = 0$*, or both of them if they do not dominate each other. (In this case the three possibilities must be checked, since if for only one of the utilities —let it be h, for instance—* $u^h(\mathbf{v}) > u^h(\mathbf{v} + \bar{\alpha} S_{ij}^{kh})$*,* $\bar{\alpha}$ *being the smallest integer between* a^{up} *and* α^{up}*, then both values* 0 *and* $\bar{\alpha}$ *are Pareto efficient.)*

iv. If $\mathcal{A}$ *is empty and there are integer points both between* a^{up} *and* α^{up} *and between* 0 *and* a^{down} *then* V^* *contains the points given by* $[u^h(\mathbf{v} + \alpha S_{ij}^{kh}, u^k(\mathbf{v} + \alpha S_{ij}^{kh})]$ *such that* α *is either the smallest integer between* a^{up} *and* α^{up}*, or the greatest integer between 0 and* a^{down}*, or both points if they do not dominate each other.*

v. In the case that $\mathcal{A}$ *contains the zero, then no point dominates the initial endowment* $\mathbf{v}$*, so that the only point in the Pareto frontier is* $\mathbf{v}$*.*

Corollary 2. *Consider the case of an economy where agents have linear utility functions with gradients* $\mathbf{c}^1, \ldots, \mathbf{c}^n$ *and let again* Γ *be the set of integer points in the interval* $[a^{down}, a^{up}]$*, where* $a^{down} = \min\{\text{argmax}_\alpha \alpha \mathbf{c}^k S_{ij}^{kh}, \text{argmax}_\alpha \alpha \mathbf{c}^h S_{ij}^{kh}\}$ *and* $a^{up} = \max\{\text{argmax}_\alpha \alpha \mathbf{c}^k S_{ij}^{kh}, \text{argmax}_\alpha \alpha \mathbf{c}^h S_{ij}^{kh}\}$*, and let* $[\alpha^{down}, \alpha^{up}]$ *be the interval of feasible step lengths defined in (6.7). It might be easily seen that either* $\Gamma = \mathbb{Q}$ *or* $\Gamma = \varnothing$*. The set* $\Gamma = \mathbb{Q}$ *in the case* $(c_i^h p_j d^k - c_j^h p_i d^k)$ *and* $(c_j^k p_i d^h - c_i^k p_j d^h)$ *have opposite sighs, whereas* $\Gamma = \varnothing$ *if* $(c_i^h p_j d^k - c_j^h p_i d^k)$ *and* $(c_j^k p_i d^h - c_i^k p_j d^h)$ *have the same sign. Then, due to Proposition 11, the set* $\mathcal{V}^*$ *of Pareto efficient solutions of an ERP may contain at most one point:*

i. if there is at least one non-null integer between $-\max\{x_i^h/(p_j d^k), x_j^k/(p_i d^h)\}/F(p_i, p_j, d^k, d^h)$ *and* $\min\{x_j^h/(p_i d^k), x_i^k/(p_j d^h)\}/F(p_i, p_j, d^k, d^h)$ *and* $\Gamma = \varnothing$*, then* $\mathcal{V}^*$ *only contains the unique point corresponding to the allocation* $\mathbf{v}^{t+1} = \mathbf{v}^t + \alpha S_{ij}^{kh}$ *for a step-length* α *which is either equal to* $-\max\{x_i^h/(p_j d^k), x_j^k/(p_i d^h)\}/F(p_i, p_j, d^k, d^h)$ *(if* $(c_i^h p_j d^k - c_j^h p_i d^k)$ *and* $(c_j^k p_i d^h - c_i^k p_j d^h)$ *are negative) or for equal to* $\min\{x_j^h/(p_i d^k), x_i^k/(p_j d^h)\}/F(p_i, p_j, d^k, d^h)$ *(if* $(c_i^h p_j d^k - c_j^h p_i d^k)$ *and* $(c_j^k p_i d^h - c_i^k p_j d^h)$ *are positive).*

ii. $\mathcal{V}^*$ *only contains the disagreement point in the opposite case.*

Having a characterization of the Pareto frontier for any ERP in the sequence allows not just a higher efficiency in simulating the process but also the possibility of measuring the number of non dominated endowments of each of the $m(m-1)n(n-1)/4$ ERPs, which might be used as a measure of uncertainty of the process. Indeed, the uncertainty of a barter process of this type might come from different sides: i) how to choose the couple of agents and commodities in each step? ii) which Pareto efficient solution of each ERP to use to update the endowments of the system? In the next subsection we shall study different criteria for answering the first two questions.

Note that the set of non-dominated solutions of the ERP, obtained by the local search movement (6.5) might give rise to imbalances between supply and demand, as described by Dreze [83] for the continuous case. To resolve this imbalance Dreze introduce a quantity rationing, which can by also extended to the ERP.

Consider a rationing scheme for the ERP as a pair of vectors $l \in \mathbb{Z}^m$, $L \in \mathbb{Z}^m$, with $L \geq 0 \geq l$, such that the t^{th} and $(t+1)^{th}$ ER verifies $l_i \leq \mathbf{x}_i^{t+1} - \mathbf{x}_i^t \leq L_i$, for $i = 1, \ldots, n$, where l_i and L_i are the i^{th} components of l and L respectively. Thus, for two given agents h and k and two given commodities i and j we have

$$l_i \leq \alpha F(p_i, p_j, d^k, d^h) \begin{bmatrix} \vdots \\ p_j d^k \\ \vdots \\ -p_j d^h \\ \vdots \end{bmatrix} \leq L_i, \qquad l_j \leq \alpha F(p_i, p_j, d^k, d^h) \begin{bmatrix} \vdots \\ -p_i d^k \\ \vdots \\ p_i d^h \\ \vdots \end{bmatrix} \leq L_j, \tag{6.14}$$

An open problem, which is not investigated in this paper, is the formulation of equilibrium conditions for this rationing scheme. One possibility might be the construction of two intervals for l and L which minimize the overall imbalances, under the conditions that (6.14) is verified in each ERP, as long as l and L are inside the respected intervals. The integrality of the allocation space Ξ forbids a straightforward application of the equilibrium criteria proposed by Dreze [83] to the markets we are considering in this work.

6.2.2 Taking a unique direction of movement

We know that the sequence of elementary reallocations formalized in (6.2) requires the iterative choice of couples of agents (h, k) and couples of commodities (i, j), i.e. directions of movement among the $m(m-1)n(n-1)/4$ in the neighborhood of the current solution. In the case of single objective problems, this choice can be made mainly in two different ways: first improving and best improving directions of movement. In the multi-objective case the choice of the direction is more complicated, as the Pareto efficiency only generates a partial order in $\mathbb{R}^n$.

However, search methods emulating the first and best improving directions of movement can be extended to the multi-objective case.

The best improving direction requires an exhaustive exploration of the neighborhood and in the case of many objective functions a welfare criterion is necessary. Noting that each direction of movement in the current neighborhood constitutes a particular ERP, a welfare criterion might be the uncertainty of each elementary reallocation, measured by the number of points in the Pareto frontier of ERPs, as described in the previous subsection. A usual welfare criterion is a norm of the objective vector (e.g., Euclidean, L_1 or L_∞ norms). Also the average marginal rate of substitution could represent an interesting criterion to select the direction of movement as a high marginal rate of substitution suggests a kind of mismatch between preferences and endowments.

Differently, the first improving direction does not require an exhaustive exploration of the neighborhood but the definition of a total order of the directions. Each of them must be explored (by solving the associated ERP) in accordance with this order, so that the first improving direction can be selected.

If at iteration t an improving direction exists the respective endowments are updated in accordance with the solution of the selected ERP: for each couple of commodities (i, j) and each couple of agents (h, k), agent k gives $\alpha F(p_i, p_j, d^k, d^h)p_j d^k$ units of i to agent h and in return he/she gets $\alpha F(p_i, p_j, d^k, d^h)p_i d^k$ units of j, for some $\alpha \in \mathbb{Z}$. At iteration $t+1$, a second couple of commodities and agents is considered in accordance with the defined criterion. If we use a first improving criterion, the process stops when the endowments keep in *status quo* continuously during $m(m-1)n(n-1)/4$ explorations, i.e. when no improving direction is found in the current neighborhood.

6.2.3 Linear objectives

In microeconomic theory the utility functions are rarely linear, however the case of linear objectives appears particularly suitable from an optimization point of view and allows a remarkable reduction of operations, as the ERPs cannot have more than one Pareto-efficient solution (see Corollary 1).

Consider a given direction of movement S_{ij}^{kh}. We know that a feasible step length α belongs to the interval defined by (6.7). Since in the case of one linear objective the gradient is constant, for any direction of movement (i, j, k, h) the best Pareto improvement (if there exists one) must happen in the endpoints of the feasible range of α (let $\alpha^{down}(i, j, k, h)$ and $\alpha^{up}(i, j, k, h)$ denote the left and right endpoints of the feasible range of α, when the direction of movement is (i, j, k, h)). Therefore, the line search reduces to decide either $\alpha^{down}(i, j, k, h)$, $\alpha^{up}(i, j, k, h)$ or none of them.

Despite the idea behind the SER is a process among self-interested agents, which are by definition local optimizers, this algorithm could also be applied to any integer linear programming problem of the form of (6.1) with one linear objective: $u(\mathbf{v}) = c'\mathbf{v}$. In this case however the branch and cut algorithm is much more efficient even for big instances, as we will show in the next section.

If a first-improve method is applied, an order of commodities and agents is required when exploring the neighborhood and the equilibrium allocation might be highly affected by this order (path-dependence). The pseudocode of Algorithm 7 describes the first improve search of the barter algorithm applied to the case of one linear objective function.

Note that if the nonnegativity constraints are not taken into account, problem (6.1) is unbounded for linear utility functions. This corresponds to the fact that without lower bounds

the linear version of this problem would make people infinitely get into debt. As a consequence, the only possible stopping criterion, when the objective function is linear, is the fulfilment of nonnegativity constraints, i.e. a given point $\mathbf{v}$ is a final endowment (an equilibrium of the barter process) if we have that for any direction of movement and for any given integer α if $c'(\mathbf{v} + \alpha S_{ij}^{kh}) > c'\mathbf{v}$ then $\mathbf{v} + \alpha S_{ij}^{kh}$ has some negative component. In some sense the optimality condition is now only based on feasibility.

Algorithm 7 First-improve SER with linear utility function

1: Initialize the endowments $E = <\mathbf{q}^1, \ldots, \mathbf{q}^n>$ and utilities $U = <u^1, \ldots, u^n>$. Let $t = 0$;
2: Let (h, k, i, j) be the t^{th} direction in the order set of directions;
3: **if** $c'(\overline{x} + \alpha^{down}(i,j,k,h)S_{ij}^{kh}) > c'(\overline{x} + \alpha^{up}(i,j,k,h)S_{ij}^{kh})$ and $c'(\overline{x} + \alpha^{down}(i,j,k,h)S_{ij}^{kh}) > c'(\overline{x})$ **then**
4: Update the incumbent $\overline{x} = \overline{x} + \alpha^{down}(i,j,k,h)S_{ij}^{kh}$ and GOTO 3;
5: **else if** $c'(\overline{x} + \alpha^{up}(i,j,k,h)S_{ij}^{kh}) > c'(\overline{x} + \alpha^{down}(i,j,k,h)S_{ij}^{kh})$ and $c'(\overline{x} + \alpha^{up}(i,j,k,h)S_{ij}^{kh}) > c'(\overline{x})$ **then**
6: Update the incumbent $\overline{x} = \overline{x} + \alpha^{up}(i,j,k,h)S_{ij}^{kh})$ and GOTO 3;
7: **else**
8: $t = t + 1$;
9: **if** $t < m(m-1)n(n-1)$ **then**
10: GOTO 4;
11: **else**
12: RETURN
13: **end if**
14: **end if**

6.2.4 The final allocation and the convergence of the SER

For the case of a continuous commodity space and exogenous prices, pairwise optimality implies global optimality, as long as all agents are initially endowed with some positive amount of a commodity [89]. Unfortunately, the SER described in this paper does not necessarily lead to Pareto efficient endowments. Let $T_{\mathbf{x}}(\alpha) = \mathbf{v} + \sum_{k \neq h} \sum_{i \neq j} \alpha(i,j,k,h) S_{ij}^{kh}$, representing a simultaneous reallocation of m commodities among n agents, with step length α_{ij}^{kh} for each couple of commodities ij and agents hk, starting from $\mathbf{x} \in \Lambda$. Whereas a SER is required to keep feasibility along the process, a simultaneous reallocation $T_{\mathbf{x}}(\alpha)$ of m commodities among n agents does not consider the particular path and any feasibility condition on the paths leading from $\mathbf{v}$ to $T_{\mathbf{v}}(\alpha)$. Hence, remembering that all SERs described in this section stop when no improving elementary reallocation exists in the current neighborhood, we can conclude that the non existence of a feasible improving ER does not entail the non existence of an improving simultaneous reallocation of m commodities among n agents. In this sense a SER provides a lower bound of any sequence of reallocations of more than two commodities and two agents at a time.

Consider the Lyapunov function $U(t) = \sum_{i=1}^{n} u^i(\mathbf{v}(t))$, associating a real value to each point in the allocation space [220]. As $U(t)$ increases monotonically along the SER (6.5) and the allocation space is a finite set, then $\lim_{t \to \infty} U(t) = U^*$.

Some understanding of the evolution of $U(t)$ along the SER iteration can be provided.

Proposition 12. *Consider a SER with m commodities among n agents with linear utility functions, i.e. $u^h = \mathbf{c}^h\mathbf{x}(t)$, where $c_i^h \leq 1$ (the utility functions can be rescaled by a common constant without affecting the SER). The change in the Lyapunov function from iteration $t-1$ to iteration t is bounded from above by*

$$U(t) - U(t-1) \leq \frac{q_{\max}}{p_{\min}} \frac{d^{\max}}{d^{\min}}, \tag{6.15}$$

where $d^{\max}$ *and* $d^{\max}$ *are the minimum and maximum elements of* $d^i \in \mathbb{Q}$*, for* $i = 1 \dots n$*, as defined in (6.1);* $p_{\min}$ *is the minimum price and* $q_{\max} = \max\{\sum_h q_j^h \ : \ j = 1 \dots m\}$.

Proof. Let (k, h, i, j) be the direction of movement selected at iteration t of the SER, $\mathbf{x}(t)$ the corresponding allocation and $\delta_t = U(t) - U(t-1)$ be the change in the Lyapunov function from iteration $t-1$ to iteration t. In the general case we have

$$\delta_t = u^h(\mathbf{v}(t) + \alpha S_{ij}^{kh}) + u^k(\mathbf{v}(t) + \alpha S_{ij}^{kh}) - u^h(\mathbf{v}(t)) - u^k(\mathbf{v}(t)), \tag{6.16}$$

which the case of linear utility functions (i.e. $u^h = \mathbf{c}^h\mathbf{v}(t)$) becomes

$$\delta_t = \alpha F(p_i, p_j, d^k, d^h) \left(\begin{bmatrix} c_i^h \\ c_i^h \end{bmatrix}^T \begin{bmatrix} p_j d^k \\ -p_i d^k \end{bmatrix} + \begin{bmatrix} c_i^h \\ c_i^h \end{bmatrix}^T \begin{bmatrix} -p_j d^h \\ p_i d^h \end{bmatrix} \right), \tag{6.17}$$

in accordance with (6.7). Based on Corollary 2, we have

$$\delta_t = -\max\left\{ \frac{v_i^h(t)}{p_j d^k}, \frac{v_j^k(t)}{p_i d^h} \right\} \left(\begin{bmatrix} c_i^h \\ c_i^h \end{bmatrix}^T \begin{bmatrix} p_j d^k \\ -p_i d^k \end{bmatrix} + \begin{bmatrix} c_i^h \\ c_i^h \end{bmatrix}^T \begin{bmatrix} -p_j d^h \\ p_i d^h \end{bmatrix} \right), \tag{6.18a}$$

if $(c_i^h p_j d^k$ - $c_j^h p_i d^k)$ and $(c_j^k p_i d^h$ - $c_i^k p_j d^h)$ are negative.

$$\delta_t = \min\left\{ \frac{v_j^h(t)}{p_i d^k}, \frac{v_i^k(t)}{p_j d^h} \right\} \left(\begin{bmatrix} c_i^h \\ c_i^h \end{bmatrix}^T \begin{bmatrix} p_j d^k \\ -p_i d^k \end{bmatrix} + \begin{bmatrix} c_i^h \\ c_i^h \end{bmatrix}^T \begin{bmatrix} -p_j d^h \\ p_i d^h \end{bmatrix} \right), \tag{6.18b}$$

if $(c_i^h p_j d^k$ - $c_j^h p_i d^k)$ and $(c_j^k p_i d^h$ - $c_i^k p_j d^h)$ are positive. Without lose of generality, let $p_j \leq 1$ (prices can be rescaled by choosing one commodity as a *numeraire*). Then, in the economically meaningful case of having $d^h = 1$, for $h = 1 \dots n$, we have

$$\delta_t \leq = \frac{q_{\max}}{p_{\min}} \frac{d^{\max}}{d^{\min}} \tag{6.19}$$

since $q_{\max} \geq v_i^h(t)$, for all $h = 1 \dots n$ and $i = 1 \dots n$. □

In the economically meaningful case of $d^h = 1$, for all $h = 1 \dots n$, the immediate economical interpretation of this result is that a high rage of variation of prices might result in big changes of the aggregated utility, from one bilateral exchange to another. The effect of the variability of prices on the computational performance of the SER will be studied in Subsection 6.3.2.

6.2.5 The SER within the graph Hamiltonian of a CERGM

In Subsection 5.5.1 the problem of maximizing graph probability under conditionally exponential models has been studied. We have seen that finding a reliable network under the specified CERGM consist in maximizing $\kappa \exp(\boldsymbol{\theta}\mathbf{S}(\mathbf{x}))$, subject to $\mathbf{x} \in \chi$. A CERGM is quite flexible, in the sense that it might include whatever network statistics in the graph Hamiltonian, as well as whatever *non-network statistics*, such as properties of nodes and properties of edges. Consider a CERGM with the following specification:

$$\begin{aligned} H(x) = \theta_1 S_1(\mathbf{x}) + \theta_2 S_2(\mathbf{x}) + \theta_3 S_3(\mathbf{x}) \\ \text{where } \mathbf{x} \in \chi. \end{aligned} \tag{6.20}$$

The three network statistics are the following:

- $S_1(\mathbf{x})$ = number of triads;

- $S_2(\mathbf{x})$ = number of edges;
- $S_3(\mathbf{x}, \mathbf{q}) = \sum_{i \in \mathcal{V}} u^i(\mathbf{v})$, where $\mathbf{v}$ is the steady state of a SER on the network $\mathbf{x}$ starting from the initial endowment $\mathbf{q}$.

The network statistics $S_3(\mathbf{x}, \mathbf{q})$ represents the aggregate utility of nodes (agents), associated to the ownership of m types of commodities from a reallocation of their initial endowments $\mathbf{q}$, by means of a SER[1]. Thus, the network statistics $S_3(\mathbf{x}, \mathbf{q})$ is defined as the value of a SER on the network$\mathbf{x}$ starting from the initial endowment $\mathbf{q}$.

6.3 Applications in computational economics

The aim of this section is to provide an inclusive application in the field of computational economics of the mathematical programming based models and methods proposed thus far.

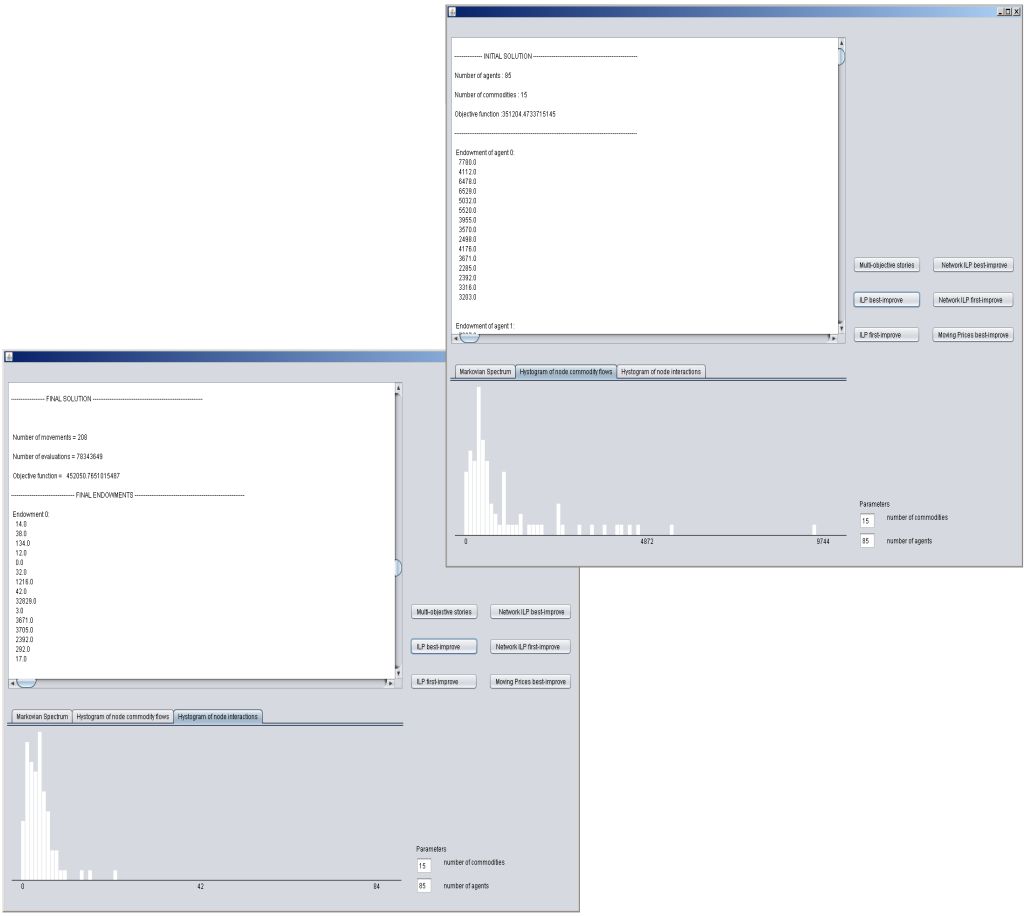


Figure 6.3: Graphical User Interface of the SER processes described in this chapter. The associated Java implementation is freely available in http://www-eio.upc.edu/~nasini/SER/launch.html

We will do so from the point of view of a global planner who wish to investigate optimality conditions of network structures for the realization of economically valuable processes and tasks.

[1]Note that simulating by the Metropolis-Hastings algorithm a bunch of networks from a CERGM with the graph Hamiltonian and sample space in (6.20) entails the iterative solution of a SER to evaluate $S_3(\mathbf{x}, \mathbf{q})$.

All the data sets used to replicates the results illustrated in this section might be downloaded from http://www-eio.upc.edu/~nasini/Thesis_datasets/Bartering_datasets, along with a Java code corresponding to the graphical interface in Figure 6.3. The reader could also modify the codes and independently use the same data to run his modified code and check his hypothesis about social bartering.

6.3.1 Numerical comparison between the simultaneous reallocation and the SERs

We first consider the number of ERs required to equilibrate the system and study their relationship with the size of the problem. In fact a numerical comparison with a global solver, such as the branch and cut, is provided to evaluate the efficiency of a decentralized barter economy in comparison with the action of a centralized global planner.

size	initial welfare	first-improve			best-improve		branch and cut	
		neighborhood	ERPs	solution	ERPs	solution	simplex	solution
10	75.134	0.66	267	353.269	91	365.126	87	394.630
10	147.958	0.84	271	763.188	91	767.371	12	769.861
10	1.205.972	0.77	375	3.925.921	74	3.844.165	70	4.060.685
15	297.713	0.70	1.343	1.455.839	215	1.471.387	49	1.488.149
15	326.996	0.71	1.090	2.544.271	237	2.554.755	63	2.614.435
15	625.800	0.71	806	2.640.317	224	2.644.008	76	2.684.016
20	183.573	0.67	2.759	3.432.832	378	3.425.665	110	3.525.421
20	1.064.023	0.81	1.582	4.197.757	361	4.194.187	94	4.331.940
20	201.377	0.78	2.629	1.017.906	351	1.089.860	80	1.180.977
25	228.365	0.89	4.358	2.221.790	648	2.226.152	237	2.271.552
25	687.492	0.65	2.806	3.416.982	572	3.403.937	113	3.462.043
25	323.495	0.61	4.706	2.262.657	666	2.245.817	50	2.474.429
30	973.955	0.79	6.648	5.428.473	975	5.427.207	101	5.377.843
30	1.811.905	0.82	13.126	8.945.605	1.084	8.953.611	127	9.080.651
30	1.302.404	0.85	12.089	7.583.841	957	7.573.400	132	7.605.525
35	653.739	0.87	13.201	3.456.918	1.310	3.458.570	112	3.474.126
35	564.905	0.80	8.772	3.579.713	1.308	3.585.815	77	3.599.639
35	753.056	0.83	14.199	5.132.226	1.290	5.107.933	67	5.333.123
40	482.570	0.87	16.307	2.429.707	1.608	2.428.731	145	2.446.953
40	430.174	0.68	7.885	5.281.060	1.640	5.229.740	90	5.279.631
40	2.795.862	0.79	14.240	19.175.278	1.578	14.503.963	186	19.276.444
45	3.392.010	0.98	62.398	22.681.229	2.300	22.664.443	162	22.728.195
45	842.645	0.92	12.900	6.606.875	2.137	6.642.397	204	6.755.016
45	1.909.859	0.97	48.688	15.979.841	2.173	15.865.744	180	16.071.407
50	839.559	0.93	20.615	4.822.082	2.105	4.859.830	137	4.895.655
50	718.282	0.97	20.744	3.586.560	2.459	3.588.633	160	3.610.194
50	1.570.652	0.99	58.165	18.872.864	2.530	19.018.519	180	19.069.868
55	351.051	0.98	20.344	2.761.203	2.935	2.748.862	1.242	2.799.187
55	413.656	0.96	26.780	4.566.394	2.922	4.569.975	336	4.585.475
55	551.355	0.99	32.053	5.136.295	3.139	5.135.647	253	5.157.444
60	468.575	0.99	27.208	1.941.409	3.568	1.949.786	271	1.995.930
60	501.366	0.99	34.323	5.051.429	3.521	5.051.836	313	5.067.154
60	575.950	0.98	43.227	4.751.072	3.589	4.747.097	273	4.801.179

Table 6.1: Numerical results of the SER and Branch and Cut for different instances of problem (6.1). The first column shows the number of agents and commodities of the problem. Columns 'ERPs' provide the number of elementary reallocations and column 'neighborhood' shows the proportion of neighborhood which has been explored. Columns 'solution' give the maximum total utility found. Column 'simplex' gives the number of simplex iterations performed by branch and cut.

We have already seen that a SER can also be applied to any integer linear programming problem of the form (6.1), where the individual utilities are aggregated in a single welfare function. If this aggregated welfare is defined as a linear function of the endowments of the form $u(\mathbf{x}) = c^T\mathbf{x}$, the comparison of the SERs with the standard branch and cut algorithm is easily carried out.

Considering the ERP as the basic operation of a SER and the simplex iteration as the basic operation of the branch and cut algorithm, the comparison between the two methods is numerically shown in Table 6.1 for three replications of 11 problems with the same number of agents and commodities, which amounts to 33 instances. The branch and cut implementation of the state-of-the-art optimization solver Cplex was used.

The numerical results in Table 6.1 shows 33 problems where the number of agents and commodities is the same, as reported in the first column. For each of the 11 different sizes 3 replicates are computed.

The second column of Table 6.1 shows the initial levels of social welfare, $c^T\mathbf{e}$. Columns *solution* give the maximum utility found for the three respective methods (first-improve local search, best-improve local search, branch and cut algorithms).

The first-improve local search results in a reduced neighborhood explorations along the sequence of movements, as suggested by the values in the column *neighborhood*, which show the proportion of possible combination of agents and commodities explored before moving to an improving direction (in comparison to the whole $m(m-1)n(n-1)/4$ candidate solutions).

The fourth and fifth columns of Tab. 6.1, named 'ERP', reports the number of movements, i.e. the number of ERPs for which the step-length α (as defined in (6.7)) has been non-null. The first-improve local search gives rise to a higher amount of ERPs, in comparison with the best-improve version. In addition, in most of the cases the best-improve search results in better allocations, as their value appear particularly close to the optimal solution (see the seventh and ninth columns of Tab. 6.1).

On the other hand, when competing with the simultaneous reallocation of all M commodities among the N agents, the sequence of best-improve elementary reallocations fails to reach comparatively good results in terms of number of elementary operations performed and goodness of the achieved final allocation.

The scatter plots in figures 6.4 and 6.5 show the relationship between the problem size (number of agents and commodities) and the elementary operations required for convergence (the ERPs for the best-improve SER and simplex iteration for the branch and cut), with the least square interpolation of algebraical curves and R^2 coefficient of determination.

$$\text{ERPs} = \beta_0 + \beta_1(\text{size})$$

$$\beta_0 = -946.2, \quad \beta_1 = 69.5$$

$$R^2 = 0.270$$

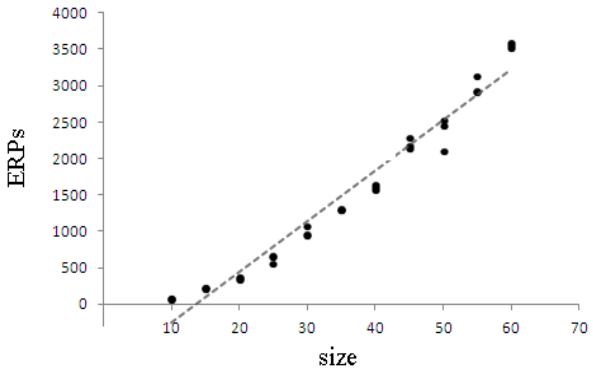


Figure 6.4: Scatter plot and least square approximation of a straight line through the relationship between the problem size and the number of ERPs for the best-improve SER method.

$$\text{simplex} = \beta_0 + \beta_1(\text{size})$$

$$\beta_0 = -63.4, \quad \beta_1 = 6.76$$

$$R^2 = 0.270$$

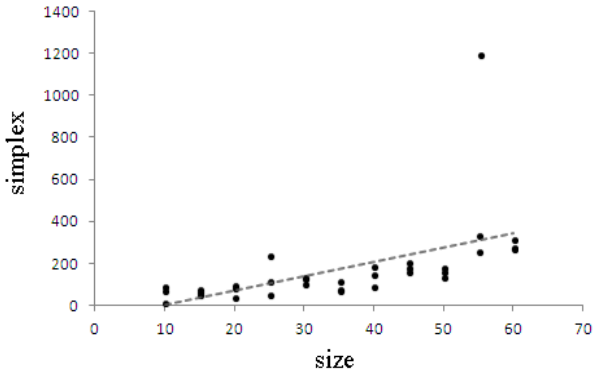


Figure 6.5: Scatter plot and least square approximation of a straight line through the relationship between the problem size and the number of simplex iteration for the branch and cut.

The scatter plots and least square approximation in Fig. 6.6 and 6.7 tries to explain the relationship between the problem size and the number of elementary operations (ERPs for the best-improve SER and the simplex pivots for the branch and cut) by an exponential curve of the form $y = \beta_0 \exp(\beta_1 x)$, with the corresponding R^2 coefficient of determination. The same kind of plots are shown in Fig. 6.8 and 6.9 for the least square interpolation of a polynomial curve of the form $y = \beta_0 x^{\beta_1}$.

$$\text{ERPs} = \beta_0 \exp(\beta_1 \text{size})$$

$$\beta_0 = 85.8, \quad \beta_1 = 0.068$$

$$R^2 = 0.919$$

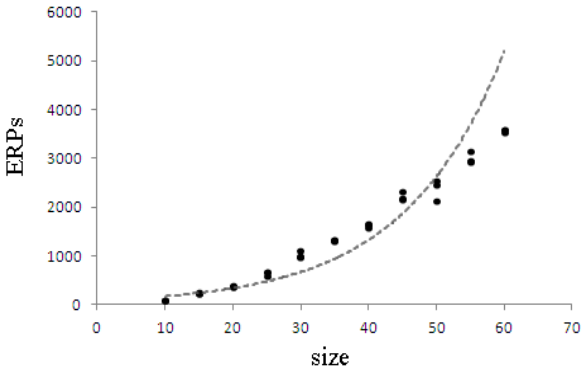


Figure 6.6: Scatter plot and least square approximation of an exponential curve through the relationship between the problem size and the number of ERPs for the best-improve SER method.

$$\text{simplex} = \beta_0 \exp(\beta_1 \text{size})$$

$$\beta_0 = 31.4, \quad \beta_1 = 0.039$$

$$R^2 = 0.602$$

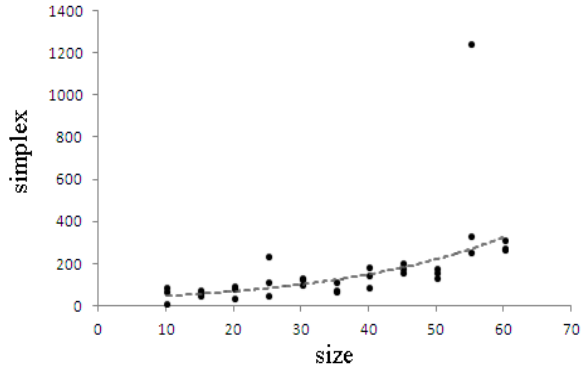


Figure 6.7: Scatter plot and least square approximation of an exponential curve through the relationship between the problem size and the number of simplex iteration for the branch and cut.

$$\text{ERPs} = \beta_0(\text{size})^{\beta_1}$$

$$\beta_0 = 0.79, \quad \beta_1 = 2.07$$

$$R^2 = 0:995$$

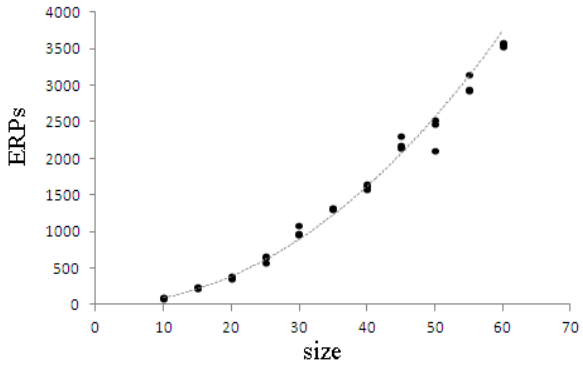


Figure 6.8: Scatter plot and least square approximation of a polynomial curve of the form $y = \beta_0 x^{\beta_1}$ through the relationship between the problem size and the number of ERPs for the best-improve SER method.

$$\text{simplex} = \beta_0(\text{size})^{\beta_1}$$

$$\beta_0 = 2.84, \quad \beta_1 = 1.100$$

$$R^2 = 0.567$$

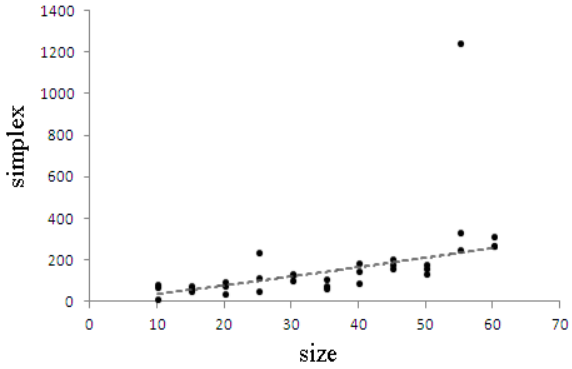


Figure 6.9: Scatter plot and least square approximation of a polynomial curve through the relationship between the problem size and the number of simplex iteration for the branch and cut.

This results quite clearly suggest a quadratic growth of the expected ERPs with respect to the size of the problem, in accordance with the a coefficient of determination of 0.995. Instead the number simplex iteration of the branch and cut algorithm seems not to be well fitted by any of the proposed curves.

From the same computational view, other sequences of reallocation have been studied by Bell [13], who analyzed the performance of the process under a variety of network structures restricting the interactions to be performed only among adjacent agents. She studied a population of Cobb Douglas' agents trading continuous amount of two commodities with local Walrasian prices and focused on the speed of convergence to an equilibrium price and allocation, observing that more centralized networks converge with fewer trades and have less residual price variation than more diverse networks.

Bell relied only on the number of trades as a measure of the speed of convergence, which we regarded as movements in the local search formalizing the process. Instead, ten years ago Wilhite [229] also toke into account the cost imposed by searching and negotiating, which we regarded as the exploration of the neighborhood[2].

[2]Note that in the special case of being interested in an aggregate social welfare, a system of many local optimizers (agents) could be highly inefficient if compared with a global optimizer, who acts for the 'goodness' of the system, as in the case of branch and cut. Also the increase of elementary operation of the barter algorithm is much higher than the one of the branch and cut, particularly when the direction of movement is selected in a best-improve way, as it is shown in Table 6.1. The economical interpretation suggests that if the time taken to reach an equilibrium allocation is too long, it is possible that this equilibrium is eventually never achieved in

6.3.2 The effect of preferences, prices, endowments

The aim of this section is to study how the initial condition of the economy, that is to say, preferences, prices and endowments, are able to affect the computational performance of the barter processes previously defined and the emerging social structure of economical interaction.

A first question when sequences of elementary reallocations are studied might be related to the analysis of which initial condition of the system is more likely to affect the number of non dominated allocations (improving directions), the number of negotiations (neighborhoods explored) and the emerging structure of interaction among agents.

To study the number of non dominated allocations obtained as a result of sequences of elementary reallocations, a method for the enumeration of all possible non-dominated paths from the known initial endowments is described. To do so, the $m(m-1)n(n-1)/4$ directions are explored in each step in such a way that a bundle non dominated reallocations are kept. Let r be the number of non-dominated reallocation in the first iteration; for each $i = 1, \ldots, r$ a collection of $l_i \leq m(m-1)n(n-1)/4$ non-dominated directions are obtained. The bundle of non-dominated solutions are thus updated in each wave by adding and allocation in accordance with this enumerative procedure.

This procedure requires a method to find Pareto-optimal vectors each time $m(m-1)n(n-1)/4$ ERPs are solved. Corley and Moon [73] proposed an algorithm to find the set V^* of Pareto vectors among r given vectors $\mathcal{V} = \{v_1, v_2, \ldots, v_r\}$, where $v_i = (v_{i1}, v_{i2}, \ldots, v_{in}) \in \mathbb{R}^n$, $i = 1, 2, \ldots, r$. Sastry and Mohideen [205] observed that the latter algorithm is incorrect and presented a modified version. In our implementation of the the best-improve barter process, we use the modified Corley and Moon algorithm of [205], shown below.

Step 1. Set $i = 1$, $j = 2$.

Step 2. If $i = r - 1$, go to Step 6. For $k = 1, 2, \ldots, n$, if $v_{jk} \leq v_{ik}$, then go to Step 3; else, if $v_{ik} \leq v_{jk}$, then go to Step 4. Otherwise, go to Step 5.

Step 3. Set $i = i + 1$, $j = i + 1$; go to Step 2.

Step 4. If $j = r$, put $v_i \in v \min V$ and $v_j = \{\infty, \infty, \ldots, \infty\}$; go to Step 3. Otherwise, set $v_{jk} = v_{rk}$, where $k = 1, 2, \ldots, n$; set $r = r - 1$ and go to Step 2.

Step 5. If $j = r$, put $v_i \in v \min V$; go to Step 3. Otherwise, set $j = j + 1$ and go to Step 2.

Step 6. For $k = 1, 2, \ldots, m$, if $v_{jk} \leq v_{ik}$, then put $v_j \in v \min V$ and stop; else, if $v_{ik} \leq v_{jk}$, then put $v_i \in v \min V$ and stop; Otherwise, put $v_i, v_j \in v \min V$ and stop.

The nice property of the modified Corley and Moon algorithm is that it doesn't necessarily compare each of the $r(r-1)/2$ couples of vectors for each of the n components. This is actually what the algorithm do in the worst case, so that the complexity could be written as $\mathcal{O}(nr^2)$, which is linear with respect of the dimension of the vectors and quadratic with respect to the number of vectors.

The pseudo-code to generate all sequences of elementary reallocations for n linear agents, keeping the Pareto-improvement in each interaction, is shown in Algorithm 8.

real social systems, where perturbing events (change in preferences, appetence of new types of commodities, etc.) might take place.

Algorithm 8 Generating paths of all improving directions of movement

Initialize the endowments $E = < \mathbf{q}^1, \ldots, \mathbf{q}^n >$ and utilities $U = < u^1, \ldots, u^n >$.
Initialize the incumbent allocations $\widetilde{E}^t = \{E\}$ and the incumbent utilities $\widetilde{U}^t = \{U\}$.
repeat
  for $\mathbf{v} \in \widetilde{E}^t$ **do**
    Let $< S_{\mathbf{v}}, G_{\mathbf{v}} >$ be the set of movements and utilities $\{(\mathbf{v} + \alpha S_{ij}^{kh},\ c'(\mathbf{v} + \alpha S_{ij}^{kh}))\}$ for each couple of commodities and agents (i, j, k, h) and $\alpha \in \{\alpha^{down}(i, j, k, h), \alpha^{up}(i, j, k, h)\}$
  end for
  Let $< S, G >= \bigcup_{x \in \widetilde{E}} < S_{\mathbf{v}}, U_{\mathbf{v}} >$ and $< S, G >= CorleyMoon(< S, G >)$
  Let $\widetilde{E}^{t+1} = \widetilde{E}^t \cup S$ and $\widetilde{U}^{t+1} = \widetilde{U}^t \cup G$
  Let $t = t + 1$
until $\widetilde{E}^t = \widetilde{E}^{t-1}$

The function $CorleyMoon()$ applies the modified Corley and Moon algorithm to a set of utility vectors and allocation vectors and return the Pareto-efficient utility vectors with the associated allocations.

	allocations			utilities
iteration 0	18 3 3	13 4 55	22 2 2	1422 559 1220
iteration 1	21 3 0	10 4 58	22 2 2	1608 574 1220
	18 3 3	11 4 57	24 2 0	1422 569 1324
iteration 2	24 0 0	7 7 58	22 2 2	1800 571 1220
	19 5 0	10 4 58	24 0 2	1480 574 1326
	21 3 0	8 6 58	24 0 2	1608 572 1326
	21 3 0	8 4 60	24 2 0	1608 584 1324
	21 3 0	10 6 56	22 0 4	1422 567 1430
	21 0 3	8 7 57	24 2 0	1614 566 1324
iteration 3	21 3 0	8 4 60	24 2 0	1608 584 1324
	22 2 0	7 7 58	24 0 2	1672 571 1326
	24 0 0	5 9 58	24 0 2	1800 569 1326
	24 0 0	5 7 60	24 2 0	1800 581 1324
	24 0 0	7 9 56	22 0 4	1614 564 1430
	19 5 0	8 4 60	26 0 0	1480 584 1430
	19 5 0	10 2 60	24 2 0	1480 586 1324
	21 1 2	8 6 58	24 2 0	1608 582 1430
iteration 4	21 3 0	8 4 60	24 2 0	1608 584 1324
	24 0 0	5 7 60	24 2 0	1800 581 1324
	19 5 0	8 4 60	26 0 0	1480 584 1430
	19 5 0	10 2 60	24 2 0	1480 586 1324
	21 1 2	8 6 58	24 2 0	1608 582 1430
	21 3 0	8 6 58	24 0 2	1800 579 1430
	22 0 2	7 9 56	24 0 2	1672 581 1430
	24 0 0	7 7 58	22 2 2	1736 582 1324
	20 2 2	7 7 58	26 0 0	1672 583 1324
iteration 5	21 3 0	8 4 60	24 2 0	1608 584 1324
	24 0 0	5 7 60	24 2 0	1800 581 1324
	19 5 0	8 4 60	26 0 0	1480 584 1430
	19 5 0	10 2 60	24 2 0	1480 586 1324
	21 1 2	8 6 58	24 2 0	1608 582 1430
	21 3 0	8 6 58	24 0 2	1800 579 1430
	22 0 2	7 9 56	24 0 2	1672 581 1430
	24 0 0	7 7 58	22 2 2	1736 582 1324
	20 2 2	7 7 58	26 0 0	1672 583 1324
	21 0 3	8 7 57	24 2 0	1544 583 1430

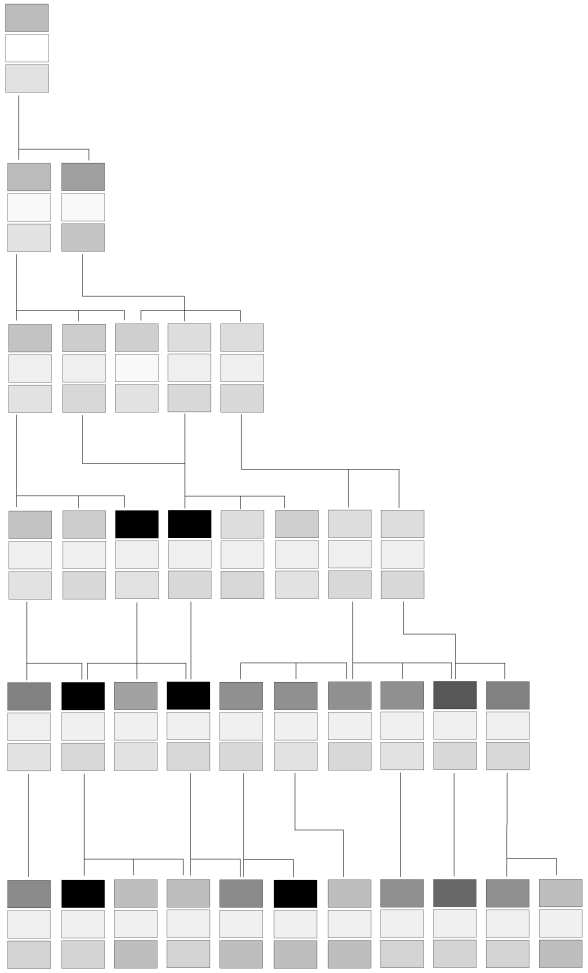


Figure 6.10: Worked example of the generation of all possible SERs, as described in algorithm 8.

Consider a barter process of 3 commodities among 3 agents and let the initial endowments be $\mathbf{q}^1 = (18\ \ 3\ \ 3)$, $\mathbf{q}^2 = (13\ \ 4\ \ 55)$ and $\mathbf{q}^3 = (22\ \ 2\ \ 2)$. The coefficients of the linear objective functions are $\mathbf{c}^1 = (75\ \ 11\ \ 13)$, $\mathbf{c}^2 = (4\ \ 3\ \ 9)$ and $\mathbf{c}^3 = (55\ \ 2\ \ 3)$. Starting from the initial solution, the sequence of two-agent-two-commodity barter leads to the movements of Figure 6.10.

The scale of grey denotes the utility level. Starting from the initial endowments, 28 different stories of elementary reallocations might be generated, although many of them lead to the same stable allocation (local optima). We found 11 stable allocations which might be reached by some sequence of elementary allocation keeping the Pareto-optimality in each ERP.

We consider a theoretical case where 2 agents with linear utility functions have to trade 9 commodities. The following three factors are taken into account:

- $Fact_1$: the variability of prices;

- $Fact_2$: association between $\mathbf{q}^1$ and $\mathbf{c}^1$ and between $\mathbf{q}^2$ and $\mathbf{c}^2$ (*initial stability*);

- $Fact_3$: association between $\mathbf{q}^1$ and $\mathbf{c}^2$ and between $\mathbf{q}^2$ and $\mathbf{c}^1$ (*dissortative matching*).

The aforementioned factors are measured at three levels and 4 randomized replicates have been simulated for each combination of factors. A multivariate analysis of variance (MANOVA) is performed, considering the two following response variables

- $Resp_1$: the number of non dominated allocations related to improving paths of algorithm 8;

- $Resp_2$: the number of neighborhoods explored.

The MANOVA results[3] in Table 6.2 illustrates the effects and the significance of $Fact_3$, corresponding to the association between the initial endowments and the marginal utilities of opposite agents. The correlation between the amounts of the initial endowments and the coefficients of the objective function of the same agent does not appear by itself to have a significant effect on the response variables.

	df	Pillai	approx F	p-value
$Fact_1$	2	0.098426	2.3033	0.06028
$Fact_2$	2	0.034673	0.7851	0.53624
$Fact_3$	2	0.133653	3.1867	0.01474
$Fact_1 \times Fact_2$	4	0.037110	0.4207	0.90758
$Fact_1 \times Fact_3$	4	0.070324	0.8109	0.59384
$Fact_2 \times Fact_3$	4	0.166118	2.0155	0.04701
Residuals	89			

Table 6.2: MANOVA analysis of the paths of all improving directions

[3]The multiple analysis of variance is used to compare multivariate (population) means of several combinations of factors. The third and fourth columns of Table 6.2 report commonly used test statistics which provide a p-value assuming an F distribution under the null hypothesis.

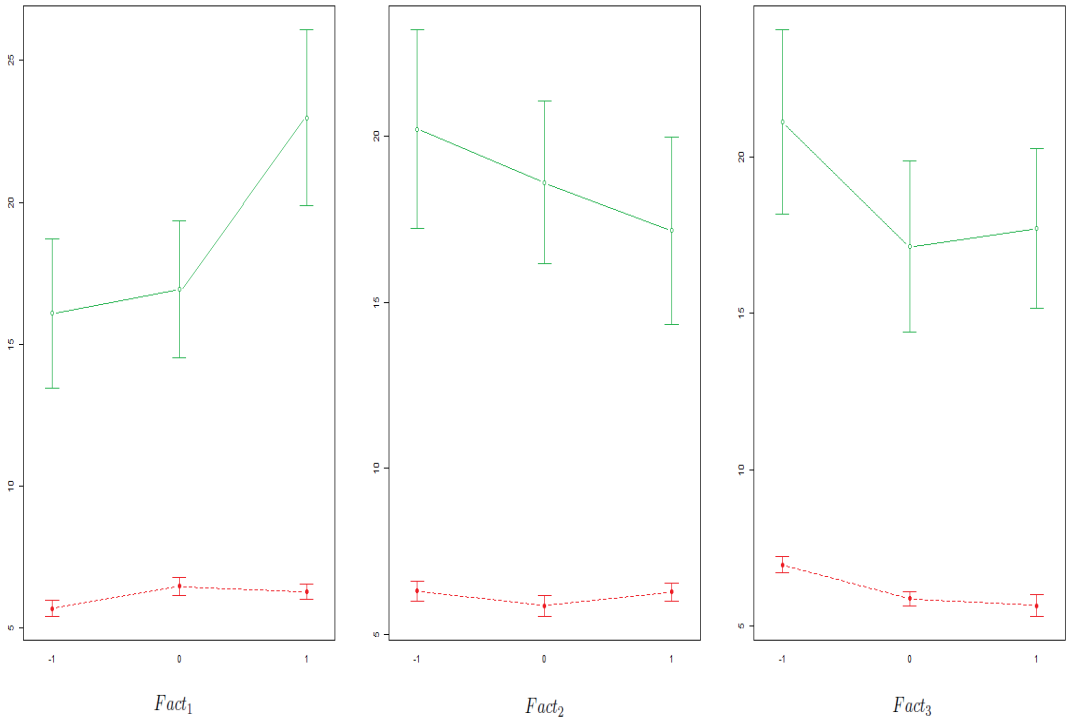


Figure 6.11: The numerical results associated to Table 6.2 are shown. The dotted red lines denote the number of non-dominated allocations, whereas the continuous green lines denote the number of of neighborhoods explored.

The graphical illustration in Figure 6.11 supports the MANOVA results, by showing the values of the two response variables for each level of the factors. The price variability seems to have a non-linear effect to both response variables (left panel). The association between the initial endowment and the marginal utility of the same agent doesn't seem to produce a consistent change in the number of neighborhoods explored (red line in the central panel), though it does have a clear average linear effect on the number of non-dominated allocations. Differently, the correlation between the initial endowment of an agent and the coefficients of the utility function of the other exhibits negative associations with the two response variables.

This experimental result should interpreted as exploratory and aiming to provide clues and suggestions for further analysis about the effect of the initial condition of the system on the outcomes and performance of the SERs. In this respect, the significant effects of dissortative matching advise for the analysis of the dissortative behavior of the economical interaction network.

Any SER intrinsically gives rise to two types of network structures generated by the set of couples of agents interacting along the process:

- the *between–node–interaction network* (whose edge set is represented by the number of exchanges, that is to say, the number of times a ERP is solved per each couple of agents),
- the *between–node–flow network* (whose edge set is represented by amount of exchanged commodities for each couple of agents).

Both networks can be seen as dynamically changing along the process. Such structures might be statistically analyzed in term of their topological properties. We consider three kind of assortativity measures reflecting the preference for an agent to interact with others that are similar or different in some ways:

- $Type_1$: couples of agents with highly different marginal utilities are more often commercial partners: $\rho(\delta(c^h,\, c^k), x_{hk})$;
- $Type_2$: agents who are more sociable (trade more often) interact frequently with agents who are not sociable: $\rho(\delta(f_h,\, f_k), x_{hk})$–;
- $Type_3$: the more two agents are different with respect to their marginal utilities, the more they are different with respect to their commercial interactions: $\rho(\delta(c^h,\, c^k), \delta(f_h,\, f_k))$.

The Greek letter δ denotes the Euclidean distance, ρ is the Pearson correlation, x_{hk} is the valued of the connection between agent h and k and f_h is the total value of connections of agent h, corresponding to the h^{th} row of the AM. The numerical values in Table 6.3 corresponds to the aforementioned assortativities applied to the *interaction network*, corresponding to the instances of Table 6.1.

Size	Between–node–flow			Between–node–interaction		
	$Type_1$	$Type_2$	$Type_3$	$Type_1$	$Type_2$	$Type_3$
10	0.40	0.48	0.63	0.70	0.67	0.74
10	0.46	0.66	0.61	0.85	0.63	0.74
10	0.60	0.48	0.75	0.71	0.70	0.75
15	0.47	0.31	0.62	0.74	0.48	0.56
15	0.33	0.36	0.58	0.58	0.44	0.67
15	0.24	0.48	0.53	0.56	0.74	0.66
20	0.28	0.41	0.61	0.39	0.62	0.54
20	0.23	0.18	0.46	0.54	0.48	0.55
20	0.12	0.06	0.37	0.48	0.45	0.42
25	0.14	0.18	0.39	0.55	0.66	0.53
25	0.36	0.32	0.60	0.65	0.56	0.66
25	0.14	0.17	0.51	0.48	0.70	0.49
30	0.09	0.08	0.40	0.42	0.55	0.53
30	0.24	0.20	0.67	0.56	0.62	0.68
30	0.26	0.33	0.60	0.61	0.63	0.65
35	0.11	0.29	0.40	0.44	0.59	0.43
35	0.14	0.28	0.50	0.46	0.55	0.48
35	0.14	0.26	0.49	0.46	0.58	0.53
40	0.25	0.22	0.53	0.44	0.64	0.58
40	0.28	0.23	0.58	0.68	0.52	0.64
40	0.26	0.18	0.69	0.64	0.64	0.60
45	0.23	0.30	0.55	0.62	0.60	0.54
45	0.29	0.24	0.61	0.57	0.59	0.58
45	0.21	0.21	0.63	0.58	0.57	0.61
50	0.08	0.28	0.36	0.35	0.55	0.32
50	0.16	0.32	0.41	0.45	0.62	0.42
50	0.24	0.17	0.60	0.51	0.50	0.65
55	0.14	0.53	0.17	0.39	0.52	0.48
55	0.17	0.33	0.38	0.29	0.53	0.44
55	0.19	0.37	0.38	0.47	0.56	0.43
60	0.35	0.45	0.60	0.54	0.57	0.62
60	0.20	0.30	0.43	0.34	0.50	0.52
60	0.16	0.38	0.29	0.39	0.51	0.48

Table 6.3: Three types of network assortativity.

The significative effect of $Fact_3$ (the association between the initial endowment and the marginal utility of the other agent) in the MANOVA of Table 6.2 seems coherent with the $Type_1$ assortativity reported in Table 6.3, in the vague sense that the difference in the agents

marginal utilities is likely to result in high exchange opportunities for agents and, conversely, in many possible convenient allocations (in the sense of Pareto).

Surprisingly, as far as the network corresponding to the between–node–flow is concerned, the $Type_3$ assortativity appear comparatively higher then the others. It might be argued that this is due to the fact that nodes with similar marginal utilities have similar abilities in catching the same exchange opportunities existing in the market. An analogous result is observed for the networks corresponding to the between–node–interaction.

Regarding the $Type_2$ dissortativity of the between–node–interaction, the values in Table 6.3 provide a clear connections with the results of Cook et al. [72], who observed that most central nodes (in the sense of eigenvector centrality) were not the most successful in achieving high bargaining power. It can be argued that this achievement relies on his/her connections with poorly connected nodes[4], as noted by Bonacich [32]:

> *in bargaining situations, it is advantageous to be connected to those who have few options; power comes from being connected to those who are powerless. Being connected to powerful others who have many potential trading partners reduces one's bargaining power.*

Note that the goodness of being connected with powerful or powerless neighbors depends on the type of commodity flowing within the network. If the utility of nodes are related to the amount of obtained information, the non rival nature of information suggests a positive association between the power of a node and the power of its neighbors.

The dissortative behavior of the valued networks generated by the barter process can be probabilistically analyzed using the conditionally uniform random network models studied in Chapter 5. For each of the three problems of size 60 in Table 6.1, the results in Table 6.4 show the sample mean and standard deviation of the clustering coefficient and assortativity coefficient of a sample of 20.000 valued networks with fixed density (summation of the AM components) generated by the q-kernel method, as shown in the Subsection 3.3.3.

Network	Property	sample mean	sample std.	observed value	one tail p-value	corr CC – AC
Flow	CC	0.0583	0.0099	0.0107	0.9951	0.1075
	AC	-0.0181	0.0054	-0.0454	0.0000	
	CC	0.0613	0.0114	0.0101	0.9951	-0.0847
	AC	-0.0196	0.0056	-0.0491	0.0000	
	CC	0.0615	0.0096	0.0390	0.9974	0.1387
	AC	-0.0188	0.0058	-0.0316	0.0379	
Interaction	CC	0.0901	0.0092	0.0822	0.7832	-0.1125
	AC	-0.0220	0.0110	-0.0454	0.0220	
	CC	0.1085	0.0050	0.1125	0.0992	0.1344
	AC	-0.0221	0.0115	-0.0491	0.0027	
	CC	0.1125	0.0042	0.1178	0.0576	-0.0250
	AC	-0.0203	0.0128	-0.0326	0.0411	

Table 6.4: Numerical results from the sample obtained with the q-kernel method, for each of the six networks associated to the the three barter processes of size 60 in Table 6.1. The model is based on the conditionally uniform distribution of valued networks with fixed density (summation of the AM components). The sixth column reports the left-tailed p-values.

Similarly, for the same samples of Table 6.4, the results in Table 6.5 show the sample mean and standard deviation of the clustering coefficient and assortativity coefficient of a sample of 10.000 valued networks with fixed row marginal of the AM generated by the q-kernel method.

[4]This results contradict most social psychological literature showing that, in experimentally restricted communication networks, the leadership role typically devolves upon the individual in the most central position [15, 145]

Network	Property	sample mean	sample std.	observed value	one tail p-value	corr CC – AC
Flow	CC	0.0170	0.0031	0.0107	0.9833	0.0462
	AC	-0.0079	0.0053	-0.0454	0.0000	
	CC	0.0177	0.0041	0.0101	1.0000	0.0286
	AC	-0.0064	0.0060	-0.0491	0.0000	
	CC	0.0433	0.0072	0.0390	0.7895	-0.0462
	AC	-0.0144	0.0073	-0.0316	0.0092	
Interaction	CC	0.0515	0.0182	0.0822	0.1179	0.0067
	AC	-0.0251	0.0102	-0.0454	0.0339	
	CC	0.0848	0.0168	0.1125	0.0870	-0.1542
	AC	-0.0173	0.0143	-0.0491	0.0254	
	CC	0.0633	0.0169	0.6384	0.0332	0.0932
	AC	-0.0154	0.0101	-0.4786	0.0433	

Table 6.5: Numerical results from the sample obtained with the q-kernel method, for each of the six networks associated to the the three barter processes of size 60 in Table 6.1. The model is based on the conditionally uniform distribution of valued networks with row marginal density. The sixth column reports the left-tailed p-values.

The results in tables 6.4 and 6.5 are quite confirmatory, as the negative values of the CC and AC between row marginal can not be explained based on the supposed conditional randomness[5].

In a series of computational experiments Kang [130] showed an interesting relationship between the variation at the individual level of a network and its assortative behavior. He found that when actors are connected with similarly central alters, the overall variation at the individual centralities (network centralization) is low. This micromacro linkages from a social network perspective might indeed be reflected in the MSS and VSS, as suggested by the definition of structural similarity (1.2).

For each of the three problems of size 60 in Table 6.1, the plots in figures 6.12, 6.13 and 6.14 show the corresponding MSS and VSS of the two generated networks.

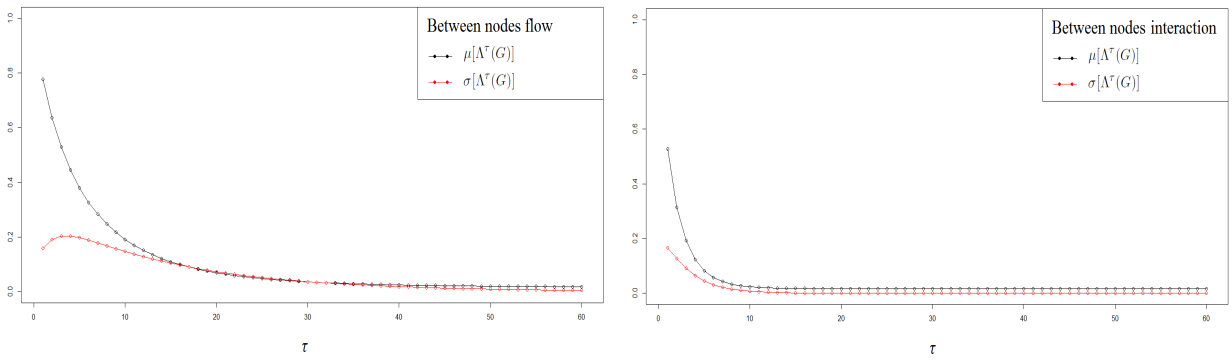


Figure 6.12: MSS and VSS of the two networks generated by the SER in an economy with 100 agents and 4 commodities with linear utility functions.

[5]Note that the CC and AC under consideration refer to their generalized versions for valued networks, as shown in 1.10

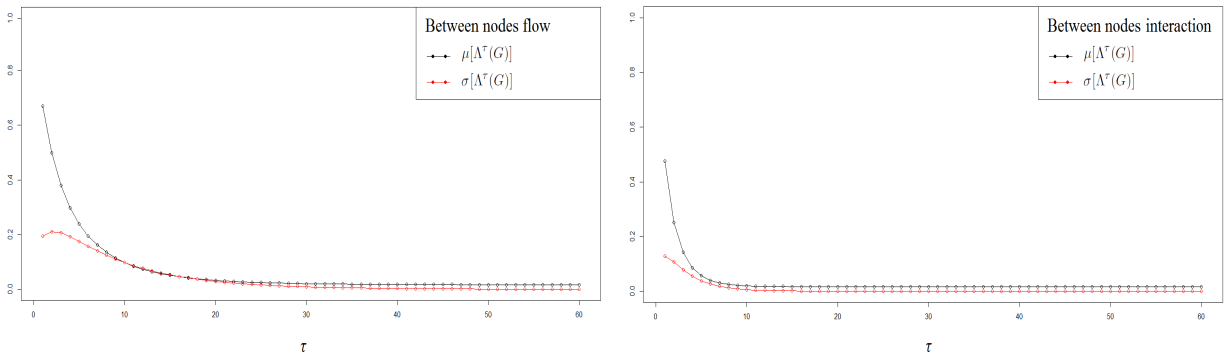


Figure 6.13: MSS and VSS of the two networks generated by the SER in an economy with 100 agents and 4 commodities with linear utility functions.

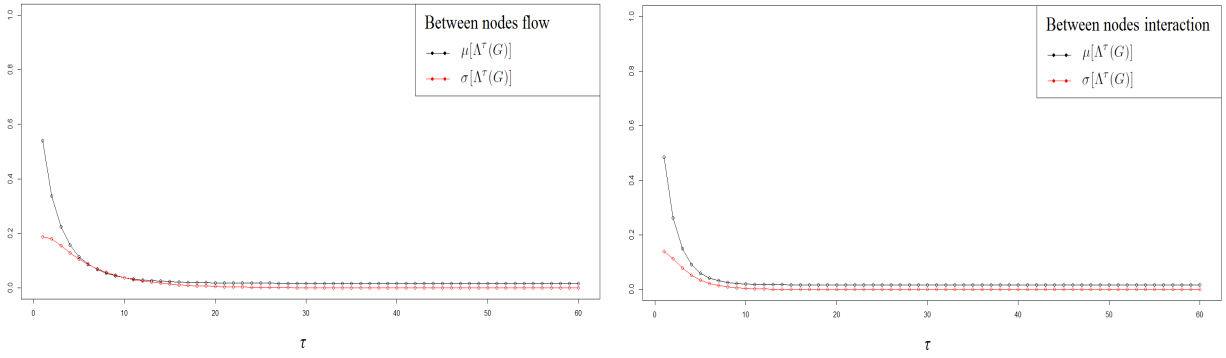


Figure 6.14: MSS and VSS of the two networks generated by the SER in an economy with 100 agents and 4 commodities with linear utility functions.

An interesting question might be whether the inequality in the distribution of the initial endowments and the variability of the marginal utilities affect the MSS and VSS of the two networks generated by the SER. To computationally answer this question a fully crossed design of experiment is performed taking into account the following two factors with two levels:

- $Fact_1$: inequality of the initial endowments;
  - Lev_1: uniform initial endowments $q_j^h = q_i^k$, for $h, k \in \mathcal{V}$, $i, j \in \mathcal{C}$;
  - Lev_2: multinomial distribution of the within agents endowments $(q_1, \ldots, q_m)$;
- $Fact_2$: variability of the marginal utilities;
  - Lev_1: uniform marginal utilities: $c_j^h = c_i^k$, for $h, k \in \mathcal{V}$, $i, j \in \mathcal{C}$;
  - Lev_2: Dirichlet distributed marginal utilities: $(c_1, \ldots, c_m)$.

The computational experiment takes into account 10 randomized replicates for each combination of factors and a MANOVA is performed, considering the six following response variables[6]:

[6]It must be noted that when both factors are at Lev_1 the initial allocation is an equilibrium for the SER and no interaction occurs among agents.

- $Resp_1$: mean similarity spectrum of order 1, that is, $\mu[\Lambda^1(G)]$;
- $Resp_2$: mean similarity spectrum of order 2, that is, $\mu[\Lambda^2(G)]$;
- $Resp_3$: mean similarity spectrum of order 3, that is, $\mu[\Lambda^3(G)]$;
- $Resp_4$: variance similarity spectrum of order 1, that is, $\sigma[\Lambda^1(G)]$;
- $Resp_5$: variance similarity spectrum of order 2, that is, $\sigma[\Lambda^2(G)]$;
- $Resp_6$: variance similarity spectrum of order 3, that is, $\sigma[\Lambda^2(G)]$.

The MANOVA results in tables 6.6 and 6.7 correspond to the network of between–node–interaction and between–nodes–flow respectively. Both factors have a significative effect on the joint distribution of MSS and VSS, for $\tau = 1, 2$ and 3, suggesting that in interaction networks, the average and variance profiles of the association between structural similarities of order 1, 2 and 3 and the tie strengths are a results of nodal attributes, such as preferences and endowments.

	df	Pillai	approx F	p-value
$Fact_1$	1	0.92704	65.6	2.99e-16
$Fact_2$	1	0.99913	5922.0	$<$ 2.2e-16
$Fact_2 \times Fact_3$	1	0.92704	65.6	2.99e-16
Residuals	36			

Table 6.6: MANOVA of the six response variables for the between–node–interaction.

	df	Pillai	approx F	p-value
$Fact_1$	1	0.84631	28.5	2.613e-11
$Fact_2$	1	0.99979	24650.1	$<$ 2.2e-16
$Fact_2 \times Fact_3$	1	0.84631	28.5	2.613e-11
Residuals	36			

Table 6.7: MANOVA of the six response variables for the between–nodes–flow.

For both network structures (the between–node–interaction network and the between–node–flow network) the initial endowments and marginal utilities result to be a fundamental factor, which is able to shape the emerging pattern of interaction, as suggested by the MANOVAs in table 6.6 and 6.7.

The global picture emerging from the observed computational results strongly supports the economical and sociological literature, discussed in Chapter 1. This is particularly true when the dissortative pattern and the network centralization are taken into account [15, 72, 130, 145]. Indeed, this strategic model of network formation is capable of internalizing many and varied assumption on agent behavior, allowing to test hypothesis on the arising of open and closed network structures from the economical interaction.

Chapter 7

Conclusions

This Ph.D. thesis was set out to explore novel mathematical programming based approaches applied to complex network problems and to search for reasonable compromises between accurate models of social structures and computationally solvable ones. It resulted in a fruitful attempt to conjugate novel mathematical and computational methodologies into the analysis of economic and social phenomena, based on a critical revision and reformulation of two classical ways of modeling social structures: i) *random models of network formation* and ii) *strategic models of network formation*. A mathematical linkage between these two approaches has been introduced in the final parts of Chapter 5 and Chapter 6, when analyzing the problem of maximizing graph probability under conditionally exponential models (see subsections 5.5.1 and 6.2.5). Thus, the precise question to be answered is: *what are the real achievements of this work?*

- A low-dimensional representation of binary and valued networks has been proposed and analyzed in Chapter 2, based on spectral graph properties. A direct application of this result has been shown in Section 5.4, when assessing the goodness of fit of random network models.

- In Chapter 3 we have been able to efficiently simulate from families of networks with complex combinatorial properties by means of LP-based methods. The resulted approaches have been shown to be general and applicable to a vast variety of families of networks as long as we are able to define them as a system of linear constraints with TU matrix structures.

- A probability density function of the primal-dual solution has be derived from the uniqueness of the central path of an LP, as shown in Section 5.3.

- A specialized interior point approach to deal with primal-block angular LPs has been studied in Chapter 4, allowing to increase the efficiency of the network generation procedures described in Chapter 3.

- Two ways of preconditioning matrix $D - C^T B^{-1} C$ (associated to the specialized interior point method of Chapter 4) have been studied, resulting in two complementary geometrical properties of blocks and linking constraints. The numerical behavior of the two resulting preconditioners has been shown when solving the LPs associated to a specified family of networks, confirming the theoretical results.

- A strategic model of network formation has been studied in Chapter 6, providing a fruitful mathematical linkage between its optimization-like properties and its multi-agents prop-

erties. In the subsections 5.5.1 and 6.2.5 the analyzed model has been included within the probabilistic framework a CERGMs, opening various possibilities of futures investigations.

- Based on computer simulation, experiment-like approaches have been used in Chapter 6 to analyze how the numerical performance of the SER as a function of input parameters. A similar analysis has been carried out to study the change in the network properties (clustering coefficient and assortativity coefficient), when the boundary conditions of the SER vary.

The researches in the context of this Ph.D. thesis gave rise to the following publications in peer-reviewed journals, scientific conferences and research reports.

Peer-reviewed publications

- Castro J., Nasini S., (2014), On geometrical properties of preconditioners in IPMs for classes of block-angular problems, to be submitted to *Mathematical Programming*.
  - Corresponding to Chapter 4.
- Nasini S., Castro J., Fonseca P., (2015), A Mathematical programming approach for different scenarios of bilateral bartering, accepted to *SORT-Statistics and Operation Research Transactions*.
  - Corresponding to Chapter 6.
- Nasini S., Castro J., Fonseca P., (2013), Novel representation of network structures by spectral theory consideration, under review in *Journal of Social Networks*.
  - Corresponding to Chapter 2.
- Castro J., Nasini S., (2013), Mathematical programming approach for classes of random networks, under review in *European Journal of Operation Research*.
  - Corresponding to chapters 3 and 5.

Scientific conferences

- Nasini S., Castro J., Specialized interior point methods for classes of random network problem, *20th Conference of the International Federation of Operational Research Societies IFORS*, Polytechnic University of Catalonia, Barcelona, Catalonia, July 2014. Invited presentation.
  - Corresponding to chapters 4 and 5.
- Nasini S., Castro J., Preconditioning IPMs for block-angular problems with "almost linearly dependent" constraints, *International Conference on Applied Mathematical Programming and Modelling APMOD*, University of Warwick, Warwick, United Kingdom, April 2014. Invited presentation.
  - Corresponding to Chapter 4.
- Nasini S., Castro J., Generating random networks by linear programming approaches, *Joint International 26th European Conference on Operational Research (EURO 2013)-INFORMS*, Rome, Italy, July 2013. Invited presentation.

  - Corresponding to chapters 3 and 5.

- Nasini S., Castro J., *International Network Optimization Conference 2013*, Tenerife, Spain, May 2013.

  - Corresponding to Chapter 3.

Research reports

- Nasini S., (2014), Maximizing graph probability under conditionally exponential models, *Arxiv-Cornell University Library*, available at http://arxiv.org/abs/1409.5476.

  - Corresponding to sections 2.4 and 5.5.

- Nasini S., Castro J., Fonseca P., (2014), Bartering integer commodities with exogenous prices, accepted by *Arxiv-Cornell University Library*, http://arxiv.org/abs/1401.3145.

  - Corresponding to Chapter 6.

I hope this thesis have been able to provide a clear understanding of the great variety of the possible modeling possibilities, when dealing with social and economic systems.

As already mentioned in the introduction of this thesis, it would have been much easier for me to have followed a more canonical and confirmatory research line, rather than such an explorative journey into the great variety of methodological possibilities that Mathematical Programming can provide to the field of CNs. Nonetheless, this way of acting allowed not only to obtain the discussed results, but also to introduced many new questions, giving rise to new open problems:

- based on the ability of simulating CERGMs (see Chapter 5), the problem of setting a general inferential framework can be studied under the Bayesian framework proposed by Friel and Caimo [46];

- based on the relation between the principal angles of the two subspaces discussed in Chapter 4, the formulation of a combined preconditioner which dynamically combines D and Θ_0 along the IP iterations can be analyzed.

Appendix A

Glossary of Complex Networks

This appendix contains a list of the main concepts used in the field of Complex Networks. The aim is to support readers who are not familiar with these notions to have their basic coordinates and guidelines. The notation used in this glossary refers to the one introduced in Subsection 1.1.2.

Assortativity coefficient (Graph Theory)

The assortativity coefficient is the Pearson correlation coefficient of degree between pairs of connected nodes. If nodes with a high degree tend to be connected to other nodes with a high degree, and nodes with a low degree to other nodes with a low degree, the graph is said to be assortative.

Average path length (Graph Theory)

The average path length is a network feature summarizing the geodesic distances between pairs nodes. It is defined as the average number of steps along the shortest paths for all pairs of nodes: $\frac{1}{n(n-1)} \sum_{i \neq j} \delta_{ij}$, where $\delta(i, j)$ is the shortest path between nodes i and j.

Betweeness (Graph Theory)

The betweenness is a measure of nodes' centrality in a network. For a given node $i \in \mathcal{V}$, it is equal to the number of shortest paths from all vertices to all others that pass through i. Betweenness centrality of a node i is the sum of the fraction of all-pairs shortest paths that pass through i:

$$c_B(i) = \sum_{s,t \in \mathcal{V}} \frac{\sigma(s, t|i)}{\sigma(s, t)}, \tag{A.1}$$

where $\sigma(s, t)$ is the total number of shortest paths from s to t and $\sigma(s, t|i)$ is the number of those paths that pass through i.

Centrality indexes (Graph Theory)

The centrality of a node is a concept capturing the intuitive idea of how important/influential is its location within the graph. Many different criteria to operationalize this concept have been proposed: betweeness, closeness, eigenvector, etc.

Closeness (Graph Theory)

The closeness centrality of a node $i \in \mathcal{V}$ is defined as the inverse of the average length of the shortest paths between i all the other node in $\mathcal{V}$. Nodes who are able to reach other nodes at shorter path lengths have favored positions and centrality.

Conductance (Graph Theory)

The conductance of a network measures how "well-knit" the network is and controls how fast a random walk converges to a uniform distribution. In Markov chains the concept of conductance is also used in a similar way and denotes a scalar measure of the tendency of a Markov chain to move out of a subset of states:

$$\Phi_L = \min_{\mathcal{V}_* \subset \mathcal{V}, 2|\mathcal{V}_*| \leq |\mathcal{V}|} \frac{\sum_{i \in \mathcal{V}_*, j \in \overline{\mathcal{V}_*}} x_{ij}}{\sum_{i \in \mathcal{V}_*, j \in \mathcal{V}} x_{ij}}, \tag{A.2}$$

where x_{ij} is the (i, j) element of the adjacency matrix, $\overline{\mathcal{V}_*}$ is the complement of $\mathcal{V}_*$ with respect to $\mathcal{V}$. The conductance of the overall chain is defined in 2.8, as the minimum conductance among all subsets of nodes $\mathcal{V}_* \subset \mathcal{V}$.

Connectivity and cuts (Graph Theory)

A graph $\mathcal{G}$ is said to be connected if for every pair of vertices there is a path joining them. A node-cut of a connected graph $\mathcal{G}$ is a set of vertices whose removal from $\mathcal{G}$ results in a disconnected graph or in a graph with a single vertex. The vertex connectivity number $\kappa(\mathcal{G})$ is defined as the minimum number of vertices whose removal from $\mathcal{G}$ results in a disconnected graph or in a graph with a single vertex. The edge connectivity number $\alpha(\mathcal{G})$ is defined as the minimum number of edges whose removal from $\mathcal{G}$ results in a disconnected graph. Whitney theorem ensures that $\kappa(\mathcal{G}) \leq \alpha(\mathcal{G})$.

Clustering coefficient (Graph Theory)

The clustering coefficient is a measure which capture how nodes in a graph tend to form densely connected groups. Two proposed measures exist: the global and the local clustering coefficients. The global clustering coefficient is defined as the ration between three times the number of closed triangle and the number of times three nodes are connected. The local clustering coefficient is the average of the local clustering coefficients of each node, which represent the proportion of possible connections existing in a neighborhood of a node. (See (1.10) for the mathematical definition.)

Degree sequence (Graph Theory)

The degree f_i of a node $i \in \mathcal{V}$ is defined as $\sum_{j \in \mathcal{V}} x_{ij}$. The question of whether a given sequence of n integer and positive numbers can be a degree sequence of a simple graphs has been answered by the ErdosGallai theorem, stating that a non-increasing sequence $f_1, \ldots, f_n$ of n integer and positive numbers is the degree sequence of a simple graph if and only if the sum of the sequence is even and $\sum_{i=1}^{k} f_i \leq k(k-1) + \sum_{i=k+1}^{n} \min(f_i, k)$ for $k = 1, \ldots, n$.

Demand function (Microeconomics)

The demand function is the optimal quantity of a commodity (the maximizer of the agent's utility) that an agent might wish to buy, as a function of a given parameter of the problem, such as a budget constraint. Sometimes it can be expressed in term of a closed-form expression, showing the relationship between the quantity of a commodity demanded and the factors affecting the willingness of an agent to buy the commodity.

Density (Graph Theory)

The density of a graph is the ratio of the number of edges and the number of possible edges. In a simple graph the density is $(\sum_{ij} x_{ij})/n(n-1)$.

Eigenvector centrality (Graph Theory)

Eigenvector centrality is a measure of the influence of a node in a network. It assigns relative scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes. Google's PageRank is a variant of the Eigenvector centrality measure.

Embeddedness (Graph Theory)

The embeddedness of an edge in a network is the number of common neighbors the two endpoints have. The embeddedness of a node coincides with its local clustering, i.e. the relative amount of edges in its neighborhood (See Subsection 1.1.4 for a more extensive discussion.)

Equilibrium (Game Theory)

A Nash equilibrium is a solution concept (a condition which identifies the equilibrium) of a game involving two or more players in which no player has anything to gain by changing only his or her own strategy unilaterally. A strategy vector $\mathbf{v} = [v_1 \dots v_n]$, such that $v_1 \in \Xi_1$, ..., $v_n \in \Xi_n$, is said to be a Nash equilibrium if for all players i and each alternate strategy $\mathbf{v}_i' \in \Xi_i$, we have that $u_i(\mathbf{v}_i, \mathbf{v}_{-i}) \geq u_i(\mathbf{v}_i', \mathbf{v}_{-i})$. A dominant strategy solution is a Nash equilibrium. If a solution is strictly dominating (switching to the solution always improves the outcome), it is also the unique Nash equilibrium. Note that a Nash equilibrium is not always an optimal solution.

Erdos-Ranyi model (Random Graph)

The study of random networks begins with the seminal work of P. Erdos and A. Ranyi [86], who considered a fixed set of vertices and an independent and equal probability of observing edges among them. There are two closely related variants of the Erdos-Ranyi model.

1. *The $G(n,p)$ model.* A network is constructed by connecting nodes randomly with independent probability p. Equivalently, all networks with n nodes and M edges have equal probability of $p^M(1-p)^{\binom{n}{2}-M}$.
2. *The $G(n,M)$ model.* A network is chosen uniformly at random from the collection of all graphs with n nodes and M edges. For example, in the $G(n=4, M=3)$ model, each of the three possible graphs on three vertices and two edges are included with probability *1/3*.

Both models possess the considerable advantage of being exactly solvable for many of its average properties.

Exponential random graph model (Random Graph)

Let $S_j(\mathcal{G}_i)$, for $j = 1, \dots, s$, be a structural feature of an observed network $\mathcal{G}_i$, for $i = 1, \dots, N$. The observed networks are seen as N particular realizations out of a large set of possible patterns. If we compute the sample means $\widehat{\mu}_j = \sum_{i=1}^{N} S_j(\mathbf{x}_i)/N$, for $j = 1, \dots, s$, we have that the $\widehat{\mu}_j$ is the empirical expectation of S_j. The ERGM arises as an answer to the question *can we recover* p (the probability measure of the networks) *from* $\widehat{\mu}_1, \dots, \widehat{\mu}_s$? To this question a reasonable requirement a probability measure p must verify is that $\mathbb{E}_p[S_j(\mathbf{x})] = \int_\chi S_j(\mathbf{x})p(\mathbf{x})dx = \widehat{\mu}_j$, for $f = 1 \dots s$, where χ is the considered set of graphs. The functional form of the ERGM is $\widehat{p(\mathbf{x})} = \kappa \exp(\boldsymbol{\theta}\mathbf{S}(\mathbf{x}))$. The boldface symbols $\boldsymbol{\theta}$ and $\mathbf{S}(\mathbf{x})$ denote the vectors $[\theta_1, \dots, \theta_s]^T$ and $[S_1(\mathbf{x}), \dots, S_s(\mathbf{x})]^T$ respectively.

Game, normal form (Game Theory)

Normal form games consist of a set $\mathcal{V}$ of players, with $|\mathcal{V}| = n$, each with a finite set of *actions available*, along with a specification of the *utility* to each player. Let Ξ_i be a variable representing the chosen action of player i, and v_i as a specific value of Ξ_i. A pure strategy profile is a vector of strategies to players, that is an n-tuple $\mathbf{v} = [v_1 \dots v_n]$, such that $v_1 \in \Xi_1, \dots, v_n \in \Xi_n$. The utility function of player i is $u_i : \Xi_1 \times \dots \times \Xi_n \to \mathbb{R}$, so that the value $u_i(\mathbf{v})$ is the payoff to player i resulting from the joint action. A game has a *dominant strategy solution* if each player in the game has a best strategy, independent of the strategies played by the other players. Since, games rarely possess *dominant strategy solutions*, a desirable solution is one in which individual players act in accordance with their incentives, maximizing their own payoff.

Graph (Graph Theory)

A simple graph $\mathcal{G}$ is defined as a finite set of elements, say $V(G)$, conventionally named vertices or nodes ($n = |\mathcal{V}|$), and a set of pairs of them, say $\mathcal{E}(\mathcal{G}) \subseteq \mathcal{V}(\mathcal{G}) \times \mathcal{V}(\mathcal{G})$. A simple graph is normally represented in terms of a $n \times n$ binary matrix $X(\mathcal{G})$, called *adjacency matrix*, whose (i, j)-entry, x_{ij}, are equal to 1 if there is a link between the corresponding row and column elements and 0 otherwise. A simple graph has no loop, so that the diagonal elements of $X(\mathcal{G})$ are null.

Laplacian matrix and graph spectrum (Graph Theory)

Let f_i be the degree of a vertex i, that is $f_i = \sum_{j \in \mathcal{V}(\mathcal{G})} x_{ij}$. Let $D(\mathcal{G}) = \text{diag}(f_1, \dots, f_n)$. The Laplacian matrix of $\mathcal{G}$ is defined as $L(\mathcal{G}) = D(\mathcal{G}) - X(\mathcal{G})$ and the sequence $\lambda_1^{(L}, \dots, \lambda_n^{(L}$, being the multiset of eigenvalues of $L(\mathcal{G})$, is called graph spectrum.

Multi-agent system (Simulation)

A multi-agent system is a programming paradigm based on the idea of multiple interacting entities within an environment. These are the fundamental properties required to the system and its agents, as in stated by Shoham [213]:

i. *Independence.* a single agent should be able to independently accomplish tasks;

ii. *Communication.* agents should be able to communicate through messages;

iii. *Intelligence.* agents have a symbolic representation of knowledge and have a means to apply rules to deduct new knowledge;

Several recent works have introduced graph-theoretic frameworks into multi-agent systems [136, 177], so that each node represents a single agent, and the edges represent pairwise channels of interaction between agents. In the context of Complex Networks, agents i) read the current messages from its neighbors, ii) update its mental state, iii) execute the commitments for the current time.

Neighborhood (Graph Theory)

The neighborhood of a vertex i in a graph $\mathcal{G}$, commonly denoted by $N_G(i)$, is the induced subgraph of $\mathcal{G}$ consisting of all vertices adjacent to i and all edges connecting two such vertices.

Pareto Efficiency (Microeconomics)

Let $\mathcal{U} \subset R^n$ be the space of allocations of an n agents bargaining problem. Points in $\mathcal{U}$ can be compared by saying that $\overline{\mathbf{u}} \in \mathcal{U}$ strictly dominates $\mathbf{u} \in \mathcal{U}$ if each component of $\overline{\mathbf{u}}$ is not less than the corresponding component of $\mathbf{u}$ and at least one component is strictly greater, that is, $u^i \leq \overline{u}^i$ for each element i and $u^j < \overline{u}^j$ for some element j. This is written as $\mathbf{u} \prec \overline{\mathbf{u}}$. Then, the Pareto frontier is the set of points of $\mathcal{U}$ that are not strictly dominated by others.

Pairwise stability (Game Theory)

Pairwise stability is a solution concepts for games of network formation. A binary network $\mathcal{G}$ is pairwise stable if

i. for all $(i,j) \in \mathcal{E}$, $u_i(\mathcal{G}) \geq u_i(\mathcal{G} - (ij))$ and $u_j(\mathbf{x}) \geq u_j(\mathcal{G} - (ij)$ and

ii. for all $(i,j) \notin \mathcal{E}$ if $u_i(\mathcal{G}) < u_i(\mathcal{G} + (ij)$ then $u_j(\mathcal{G}) > u_j((\mathcal{G} + (ij))$.

The notation $\mathcal{G} - (ij)$ refers to the network structure obtained from $\mathcal{G}$ by removing the (i,j) connection. Thus, a network is pairwise stable if no agent wants to sever a tie and no couple of agents simultaneously want to add a tie.

Preference relation and utility function (Microeconomics)

Consider a set $\mathcal{V}$ of players, with $|\mathcal{V}| = n$ and let Ξ_i be the space of actions of player i, and $v_i \in \Xi_i$. Let $\boldsymbol{\Xi} = \Xi_1 \times \ldots \times \Xi_n$. The preference relation $\preceq_i$ of player $i \in \mathcal{V}$ on $\boldsymbol{\Xi}$ is represented by the utility function $u_i : \boldsymbol{\Xi} \to \mathbb{R}$, resulting from the joint action: $u_i(\mathbf{v})$. Arrow and Debreu [80] showed that if the set $\{(\mathbf{v}_a, \mathbf{v}_b) \in \boldsymbol{\Xi} \times \boldsymbol{\Xi} : \mathbf{v}_a \preceq_i \mathbf{v}_b\}$ is closed relative to $\boldsymbol{\Xi} \times \boldsymbol{\Xi}$ the preference relation can be represented by a real-valued function $u^i : \boldsymbol{\Xi} \longmapsto \mathbb{R}$, such that, for each $\mathbf{v}_a$ and $\mathbf{v}_b$ belonging to $\boldsymbol{\Xi}$, $u^i(\mathbf{v}_a) \leqslant u^i(\mathbf{v}_b)$ if and only if $\mathbf{v}_a \preceq_i \mathbf{v}_b$.

Path and cycle (Graph Theory)

A path in a graph is a sequence of vertices such that from each of its vertices there is an edge to the next vertex in the sequence. A path always has a first vertex, called its start vertex, and a last vertex, called its end vertex. The other vertices in the path are internal vertices. A cycle is a path such that the start vertex and end vertex are the same.

Stochastic and doubly stochastic matrices

A doubly stochastic matrix P, is a square matrix, whose component $p_{ij} \in [0,1] \subset \mathbb{R}$ and each of whose rows and columns sum to 1. The class of $n \times n$ doubly stochastic matrices is a convex polytope in $\mathbb{R}^{2n}$, known as the Birkhoff polytope. (See Chapter 2 for a more extensive discussion.)

Watts-Strogatz model (Random Graph)

The WattsStrogatz model is a random graph generator, which allow to simulate graphs with short average path lengths and high clustering coefficient. It is characterized by three parameters: the number of nodes n, the average degree k, and the probability of rewiring p. The mechanism of network formation is carried out in accordance with the following procedure:

1. Arrange the n nodes on a circle and connect each node with $k/2$ neighbor nodes on both sides. (It results in a network with $kN/2$ links)
2. Rewire each of the existing links with probability p under the regulation that there are no self-loops or multiple links. (After rewiring there will be $pkN/2$ rewritten connections on average.)

It has been observed that in a certain range of p, the network presents the so-called small-world property, including short average path lengths and high clustering coefficient.

Appendix B

Primal-dual interior point methods

Consider the problem of optimizing a quadratic function of several variables subject to linear constraints on these variables:

$$\begin{array}{rl} \min & \mathbf{c}^T\mathbf{x} + \frac{1}{2}\mathbf{x}^T Q\mathbf{x} \\ \text{subject to} & A\mathbf{x} = \mathbf{b} \\ & 0 \leq \mathbf{x} \leq \mathbf{u} \end{array} \tag{B.1}$$

where $\mathbf{c}, \mathbf{x}, \mathbf{u} \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$, $Q \in \mathbb{R}^{n\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. Replacing inequalities in (B.1) by a logarithmic barrier with parameter $\mu > 0$ we obtain the logarithmic barrier problem

$$\begin{array}{rl} \min & B(\mathbf{x},\mu) \triangleq \mathbf{c}^T\mathbf{x} + \dfrac{1}{2}\mathbf{x}^T Q\mathbf{x} + \mu\left(-\displaystyle\sum_{i=1}^{n} \ln x_i - \sum_{i=1}^{n} \ln(u_i - x_i)\right) \\ \text{subject to} & A\mathbf{x} = \mathbf{b}. \end{array} \tag{B.2}$$

where x_i and u_i are the i-th components of $\mathbf{x}$ and $\mathbf{u}$ respectively. The KKT conditions of (B.2) are:

$$A\mathbf{x} = \mathbf{b}, \tag{B.3a}$$
$$A^T\mathbf{y} - Q\mathbf{x} + \mathbf{z} - \mathbf{w} = \mathbf{c}, \tag{B.3b}$$
$$XZ\mathbf{e} = \mu\mathbf{e}, \tag{B.3c}$$
$$(U - X)W\mathbf{e} = \mu\mathbf{e}, \tag{B.3d}$$
$$(\mathbf{z}, \mathbf{w}) > 0 \quad \mathbf{u} > \mathbf{x} > 0. \tag{B.3e}$$

Here, $\mathbf{e} \in \mathbb{R}^n$ is a vector of ones; $\mathbf{y} \in \mathbb{R}^m$, $\mathbf{z}, \mathbf{w} \in \mathbb{R}^n$ are the Lagrange multipliers (or dual variables) of $A\mathbf{x} = \mathbf{b}$, $\mathbf{x} \geq 0$ and $\mathbf{x} \leq \mathbf{u}$, respectively; and matrices $X, Z, U, W \in \mathbb{R}^{n\times n}$ are diagonal matrices made up of vectors $\mathbf{x}, \mathbf{z}, \mathbf{u}, \mathbf{w}$. Equations (B.3a)–(B.3b) impose, respectively, primal and dual feasibility, while (B.3c)–(B.3d) impose complementarity. The normal equations for the Newton direction $(\Delta x, \Delta y, \Delta z)$ of (B.3) reduce to

$$(A\Theta A^T)\Delta y = g \tag{B.4}$$
$$\Theta = (Q + (U - X)^{-1}W + X^{-1}Z)^{-1}, \tag{B.5}$$

If we let $A \in \mathbb{R}^{m\times p}$ be the coefficient matrix of (4.1), $b \in \mathbb{R}^m$ its right-hand term and $c \in \mathbb{R}^p$ the gradient of the objective function. The KKT conditions of (4.1) might be formulated in terms of a mapping, namely $G : \mathbb{R}^{3N+M} \longrightarrow \mathbb{R}^{3N+M}$.

$$G(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{w}) = \begin{bmatrix} A\mathbf{x}-\mathbf{b} \\ A^T\mathbf{y}+\mathbf{z}-\mathbf{w}-\mathbf{c} \\ XZ\mathbf{e} \\ (U-X)W\mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mu\mathbf{e} \\ \mu\mathbf{e} \end{bmatrix} \tag{B.6a}$$

$$(\mathbf{z},\mathbf{w}) > 0; \qquad 0 \leq \mathbf{x} \leq 1. \tag{B.6b}$$

The set of primal-dual solutions $\mathfrak{C} = \{(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{w}),\ G(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{w}) = [0,0,\mu\mathbf{e},\mu\mathbf{e}]^T,\ \mu > 0\}$ of (B.6) is known as the central path. Primal-dual path-following interior-point algorithms approximately follow the central path by applying Newton method to the nonlinear function (B.6), reducing the barrier parameter μ at each iteration. When μ approaches the zero these solutions converge to the optimal solution of the original problem. The Newton's direction $(\Delta_x,\Delta_y,\Delta_z,\Delta_w)$ is given by the product between the inverse Javobian of $\mathbf{F}$ and the value of $\mathbf{F}$ in the current iterate: $(\Delta_x,\Delta_y,\Delta_z,\Delta_w) = \nabla G(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{w})^{-1}G(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{w})$. Thus, the solution of the following system is required.

$$\begin{bmatrix} A & & & \\ & A^T & I & -I \\ Z & & X & \\ -W & & & U-X \end{bmatrix} \begin{bmatrix} \Delta_x \\ \Delta_y \\ \Delta_z \\ \Delta_w \end{bmatrix} = \begin{bmatrix} A\mathbf{x}-\mathbf{b} \\ A^T\mathbf{y}+\mathbf{z}-\mathbf{w}-\mathbf{c} \\ XZ\mathbf{e} \\ (U-X)W\mathbf{e} \end{bmatrix} \tag{B.7}$$

In practice, variables Δ_z and Δ_w are eliminated multiplying by X^{-1} the third block of equations and by $(I-X)^{-1}$ the fourth one and subtracting to the second blocks of equations. (See Wright [231] and Castro [53] for more details.) The system reduces to the indefinite augmented system form:

$$\begin{bmatrix} A & \\ -\Theta^{-1} & A^T \end{bmatrix} \begin{bmatrix} \Delta_x \\ \Delta_y \end{bmatrix} = \begin{bmatrix} A\mathbf{x}-\mathbf{b} \\ A^T\mathbf{y}-\mu(U-X)^{-1}\mathbf{e} \end{bmatrix}, \tag{B.8}$$

where $\Theta = ((I-X)^{-1}W + X^{-1}Z)^{-1}$ is a diagonal matrices. Multiplying the last block of equations by $A\Theta$ and summing to the first, we obtain that the direction of movement in the dual space is obtained by solving $A\Theta A^T\Delta_y = A[\mathbf{x}-\mathbf{b}+\Theta A^T\mathbf{y}-\mu\Theta(U-X)^{-1}\mathbf{e}]$, with respect to Δ_y. This is known as the system of normal equation and represents indeed the most time consuming step of the primal-dual interior point method.

When the problem has a primal block-angular structure as in (4.1), the specialized interior-point algorithm [50, 51] described in Chapter 4, the normal equations for the Newton direction of Δ_y becomes

$$\begin{aligned} A\Theta A^T\Delta_{\mathbf{y}} &= \left[\begin{array}{ccc|c} N_1\Theta_1N_1^T & & & N_1\Theta_1L_1^T \\ & \ddots & & \vdots \\ & & N_k\Theta_kN_k^T & N_k\Theta_kL_k^T \\ \hline L_1\Theta_1N_1^T & \dots & L_k\Theta_kN_k^T & \Theta_0+\sum_{i=1}^k L_i\Theta_iL_i^T \end{array}\right]\Delta_y \\ &= A[\mathbf{x}-\mathbf{b}+\Theta A^T\mathbf{y}-\mu\Theta(I-X)^{-1}\mathbf{e}], \end{aligned} \tag{B.9}$$

As we said, the primal-dual path-following algorithms restrict iterates to a neighborhood of the central path, by applying the Newton method to (B.6) and reducing the barrier parameter μ at each iteration. This class of algorithms differentiate in accordance with the ways in which μ is reduced along the iterative process: short-step path-following methods, long-step path-following methods; predictor-corrector path-following methods. Full details can be found in Wright [231].

Appendix C

A specialized interior point method for markets with exogenous prices

In this appendix we consider how the inspiring idea behind the specialized interior point algorithm for primal block-angular problems, studied in Chapter 4, can be fruitfully applied to the continuous relaxation of the concave maximization problem (6.1), with aggregated utility: $\sum_{h=1}^{n} \alpha_h u^h(\mathbf{x})$, where $\alpha_1, \ldots, \alpha_n$ are positive weights.

In order to exploit the underlying idea of the specialized point method for block-angular linear program [50, 53], we consider a modified version of problem (6.1), in which the linking constraints are relaxed in the form of inequalities: $[I \quad I \ldots \quad I]\mathbf{v} + \mathbf{v}_0 = \mathbf{b}^0$, where $0 \leq \mathbf{v} \leq u_v$ and $0 \leq \mathbf{v}_0 \leq u_s$. Similarly, in this modified version we also reply the inequalities associated to the disagreement point (agents rationality) with equality constraints, by adding slack variables: $u^h(\mathbf{v}) - u^h(\mathbf{q}) - s^h = 0$, for $h = 1 \ldots, n$, with the condition $0 \leq \mathbf{s} \leq u_s$, where $\mathbf{s}^T = [s^1 \ldots s^n]$. If we let $A \in \mathbb{Q}^{n+m \times mn+m}$ be the coefficient matrix associated to (MCAPFP), the resulting μ-KKT conditions are:

$$
\begin{aligned}
A\mathbf{v} &= \mathbf{b}, && \text{(C.1)}\\
u^h(\mathbf{v}) - u^h(\mathbf{q}) - s^h &= 0 \qquad h = 1 \ldots, n, && \text{(C.2)}\\
A^T\mathbf{y} + \mathbf{z}_v - \mathbf{w}_v + \sum_{h=1}^{n} t^h \left[\frac{\nabla u^h(\mathbf{v})}{\mathbf{0}}\right] &= \sum_{h=1}^{n} \alpha^h \left[\frac{\nabla u^h(\mathbf{v})}{\mathbf{0}}\right] && \text{(C.3)}\\
T\mathbf{e}_s + \mathbf{z}_s - \mathbf{w}_s &= 0 && \text{(C.4)}\\
XZ_v\mathbf{e}_v &= \mu\mathbf{e}_v, && \text{(C.5)}\\
(U_v - X)W_v\mathbf{e}_v &= \mu\mathbf{e}_v, && \text{(C.6)}\\
SZ_s\mathbf{e}_s &= \mu\mathbf{e}_s, && \text{(C.7)}\\
(U_s - S)W_s\mathbf{e}_s &= \mu\mathbf{e}_s, && \text{(C.8)}
\end{aligned}
$$

where $\mathbf{e}_v \in \mathbb{R}^{nm+m}$ and $\mathbf{e}_s \in \mathbb{R}^n$ are a vectors of ones; $\mathbf{y} \in \mathbb{R}^{m+n}$ and $\mathbf{z}_v, \mathbf{w}_v \in \mathbb{R}^{nm+m}_{+\cup\{0\}}$ are the Lagrange multipliers (or dual variables) of $A\mathbf{v} = \mathbf{b}$ and $\mathbf{v} \geq 0$, $\mathbf{v} \leq u_v$ respectively; similarly, $\mathbf{t} = [t_1 \ldots t_n]^T \in \mathbb{R}^n$ is the vector of Lagrangian multipliers of $u^h(\mathbf{v}) - u^h(\mathbf{q}) + s_h = 0$, for $h = 1 \ldots, n$ and $\mathbf{z}_s, \mathbf{w}_s \in \mathbb{R}^{2n}_{+\cup\{0\}}$ are the Lagrange multipliers of $\mathbf{s} \geq 0$, $\mathbf{s} \leq u$ respectively. Primal variables must be inside the intervals $0 < \mathbf{v} < u_v$, $0 < \mathbf{s} < u_s$, $0 < \mathbf{v}_0 < u_v$. Matrices $X, Z_v, U_v, W_v \in \mathbb{R}^{(nm+m)\times(nm+m)}$ are diagonal matrices made up of vectors $\mathbf{v}, \mathbf{z}_v, \mathbf{u}_v, \mathbf{w}_v$; matrices $S, T, Z_s, U_s, W_s \in \mathbb{R}^{n\times n}$ are diagonal matrices made up of vectors $\mathbf{s}, \mathbf{t}, \mathbf{z}_s, \mathbf{u}_s, \mathbf{w}_s$. Matrix $T \in \mathbb{R}^{n\times n}$ is diagonal with components $t_1, \ldots, t_n$.

Applying Newton method to (C.1) and reducing the barrier parameter μ at each iteration, we have that the $\mathbf{v}$ solution of (C.1) converge to the optimal commodity allocation, when μ approaches the zero. The Newton's direction $(\Delta_x, \Delta_s, \Delta_y, \Delta_t, \Delta_{z_v}, \Delta_{z_s}, \Delta_{w_v}, \Delta_{w_s})$ is obtained by solving the following system in each iteration.

$$\left[\begin{array}{cc|cc|c|cc} A & & & & & & \\ V & -I & & & & & \\ \hline Q & & A^T & V^T & & & \\ & & & -I & \text{I} & \text{-I} & \\ \hline Z_v & & & & X & & \\ & Z_s & & & & S & \\ \hline -W_v & & & & & U_v - X & \\ & -W_s & & & & & U_s - S \end{array}\right] \left[\begin{array}{c} \Delta_x \\ \Delta_s \\ \hline \Delta_y \\ \Delta_t \\ \hline \Delta_{z_v} \\ \Delta_{z_s} \\ \hline \Delta_{w_v} \\ \Delta_{w_s} \end{array}\right] = \left[\begin{array}{c} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \hline \mathbf{r}_3 \\ \mathbf{r}_4 \\ \hline \mathbf{r}_5 \\ \mathbf{r}_6 \\ \hline \mathbf{r}_7 \\ \mathbf{r}_8 \end{array}\right] \tag{C.9}$$

where the right-hand term is defined as

$$\begin{aligned} \mathbf{r}_1 &= A\mathbf{v} - \mathbf{b} \\ \mathbf{r}_2 &= \begin{bmatrix} u^1(\mathbf{v}) - u^1(\mathbf{q}) - s^1 \\ \vdots \\ u^n(\mathbf{v}) - u^n(\mathbf{q}) - s^n \end{bmatrix} \\ \mathbf{r}_3 &= A^T\mathbf{y} + \mathbf{z} - \mathbf{w} - \sum_{h=1}^{n}(t^h - \alpha^h)\begin{bmatrix} \nabla u^1(\mathbf{v}) \\ \hline \mathbf{0} \end{bmatrix} . \\ \mathbf{r}_4 &= T\mathbf{e}_s + \mathbf{z}_s - \mathbf{w}_s \\ \mathbf{r}_5 &= XZ_v\mathbf{e}_v - \mu\mathbf{e}_v \\ \mathbf{r}_6 &= XZ_s\mathbf{e}_s - \mu\mathbf{e}_s \\ \mathbf{r}_7 &= (U_v - X)W_v\mathbf{e}_v - \mu\mathbf{e}_v \\ \mathbf{r}_8 &= (U_s - X)W_s\mathbf{e}_s - \mu\mathbf{e}_s \end{aligned} \tag{C.10}$$

Considering that $\frac{\partial u^h(\mathbf{v})}{\partial_{x_j^k}} = 0$ (for $k \neq h$), as we assumed in Chapter 5, matrix $Q(\mathbf{v})$ results to be block-diagonal:

$$Q(\mathbf{v}) = \left[\begin{array}{ccc|c} Q^1(\mathbf{v}) & & & \\ & Q^2(\mathbf{v}) & & \\ & & \ddots & \\ & & Q^n(\mathbf{v}) & \\ \hline & & & \mathbf{0} \end{array}\right], \tag{C.11}$$

where, for each agent $h = 1, \ldots, n$ and each couple of commodities $i, j = 1, \ldots, m$, we have $Q_h(\mathbf{v}) \in \mathbb{R}^{m \times m}$ to be defined as:

$$Q_{ij}^h(\mathbf{v}) = (t^h - \alpha^h)\frac{\partial^2 u^h(\mathbf{v})}{\partial_{x_j^h}\partial_{x_i^h}}. \tag{C.12}$$

Matrix $V \in \mathbb{R}^{mn+m \times n}$ is also block-diagonal:

$$V = \left[\begin{array}{ccc|c} \nabla u^1(\mathbf{v})^T & & & \\ & \nabla u^2(\mathbf{v})^T & & \\ & & \ddots & \\ & & \nabla u^n(\mathbf{v})^T & \\ \hline & & & 0 \end{array}\right]. \tag{C.13}$$

By collecting variables $\Delta_{\widehat{x}} = [\Delta_x \mid \Delta_s]$, $\Delta_{\widehat{y}} = [\Delta_y \mid \Delta_t]$, $\Delta_{\widehat{z}} = [\Delta_{z_v} \mid \Delta_{z_s}]$ and $\Delta_{\widehat{w}} = [\Delta_{w_v} \mid \Delta_{w_s}]$ and performing elementary row operations, system (C.9) might be reduced to

$$\begin{bmatrix} \widehat{A} & & & \\ \widehat{\Theta} & \widehat{A}^T & & \\ \widehat{X}^{-1}\widehat{Z} & & I & \\ -(\widehat{U}-\widehat{X})^{-1}\widehat{W} & & & I \end{bmatrix} \begin{bmatrix} \Delta_{\widehat{x}} \\ \Delta_{\widehat{y}} \\ \Delta_{\widehat{z}} \\ \Delta_{\widehat{w}} \end{bmatrix} = \begin{bmatrix} \widehat{\mathbf{r}}_1 \\ \widehat{\mathbf{r}}_2 \\ \widehat{\mathbf{r}}_3 \\ \widehat{\mathbf{r}}_4 \end{bmatrix}, \tag{C.14}$$

where $\widehat{A} \in \mathbb{R}^{2(n+m)\times(nm+n+m)}$ is defined as:

$$\widehat{A} = \begin{bmatrix} A & \\ V & -I \end{bmatrix}. \tag{C.15}$$

and $\widehat{Z} \in \mathbb{R}^{(nm+n+m)\times(nm+n+m)}$, $\widehat{X} \in \mathbb{R}^{(nm+n+m)\times(nm+n+m)}$, $\widehat{W} \in \mathbb{R}^{(nm+n+m)\times(nm+n+m)}$, $\widehat{U} \in \mathbb{R}^{(nm+n+m)\times(nm+n+m)}$ are also defined by concatenating the corresponding diagonal matrices in (C.9), as well as the right-hand term: $\widehat{\mathbf{r}}_4 = -(\widehat{U}-\widehat{X})^{-1}[\mathbf{r}_7 \mid \mathbf{r}_8]$, $\widehat{\mathbf{r}}_3 = \widehat{X}^{-1}[\mathbf{r}_5 \mid \mathbf{r}_6]$, $\widehat{\mathbf{r}}_2 = [\mathbf{r}_3 \mid \mathbf{r}_4] - \widehat{\mathbf{r}}_3 + \widehat{\mathbf{r}}_4$, $\widehat{\mathbf{r}}_1 = [\mathbf{r}_1 \mid \mathbf{r}_2]$. Thus, variables $\Delta_{z_v}, \Delta_{z_s}, \Delta_{w_v}$ and Δ_{w_s} might be eliminated after solving the indefinite augmented form:

$$\begin{bmatrix} \widehat{A} & \\ \widehat{\Theta} & \widehat{A}^T \end{bmatrix} \begin{bmatrix} \Delta_{\widehat{x}} \\ \Delta_{\widehat{y}} \end{bmatrix} = \begin{bmatrix} \widehat{\mathbf{r}}_1 \\ \widehat{\mathbf{r}}_2 \end{bmatrix}, \tag{C.16}$$

Matrix $\widehat{\Theta} \in \mathbb{R}^{(nm+m+n)\times(nm+m+n)}$ is

$$\widehat{\Theta} = \begin{bmatrix} \Theta^x & \\ & \Theta^s \end{bmatrix} = \begin{bmatrix} Q - X^{-1}Z_v + (U_v - X)W_v & \\ & -S^{-1}Z_s + (U_s - S)W_s \end{bmatrix}, \tag{C.17}$$

where $\Theta^x \in \mathbb{R}^{(nm+m)\times(nm+m)}$ and $\Theta^s \in \mathbb{R}^{n\times n}$. Multiplying by $-A\Theta^{-1}$ the last block of equations and summing it to the first one, we obtain that the coefficient matrix of the system to be solved to compute Δ_y is

$$\begin{aligned} \widehat{A}\widehat{\Theta}\widehat{A}^T &= \begin{bmatrix} A\Theta^x A^T & A\Theta^x V^T \\ V\Theta^x A^T & V\Theta^x V^T - \Theta^s \end{bmatrix} \\ &= \left[\begin{array}{ccc|c|ccc} P\Theta_1^x P^T & & & d_1 P\Theta_1^x & P\Theta_1^x \nabla_{u^1} & & \\ & \ddots & & \vdots & & \ddots & \\ & & P\Theta_n P^T & d_n P\Theta_n^x & & & P\Theta_n^x \nabla_{u^1}^T \\ \hline d_1\Theta_1^x P^T & \dots & d_n\Theta_n^x P^T & \Theta_0 + \sum_{h=1}^{n} d_h^2 \Theta_h^x & d_1\Theta_1^x\nabla_{u^1} & \dots & d_n\Theta_n^x\nabla_{u^n} \\ \hline \nabla_{u^1}^T\Theta_1^x P^T & & & d_1\nabla_{u^1}^T\Theta_1^x & \nabla_{u^1}^T\Theta_1^x\nabla_{u^1} - \Theta_1^s & & \\ & \ddots & & \vdots & & \ddots & \\ & & \nabla_{u^n}^T\Theta_n^x P^T & d_n\nabla_{u^n}^T\Theta_n^x & & & \nabla_{u^n}^T\Theta_n^x\nabla_{u^n} - \Theta_n^s \end{array}\right] \\ &= \left[\begin{array}{c|c|c} B & C_0 & C_1 \\ \hline C_0^T & D_0 & D_\nabla^T \\ \hline C_1^T & D_\nabla & D_1 \end{array}\right], \end{aligned} \tag{C.18}$$

where

$$\nabla_{u^h} = \begin{bmatrix} \dfrac{\partial u^h(\mathbf{v})}{\partial_{x_m^h}} \\ \vdots \\ \dfrac{\partial u^h(\mathbf{v})}{\partial_{x_m^h}} \end{bmatrix} \qquad h = 1, \ldots, n. \tag{C.19}$$

Thus, by noting that the first n components of the Newton direction Δ_y are associated to the block-angular constraints $\sum_{i \in C} p_i x_i^h = \sum_{i \in C} p_i v_i^h$, for $h = 1, \ldots, n$, whereas the second m components of Δ_y are associated to the linking constraints $\sum_{h \in A} x_i^h = \sum_{h \in A} v_i^h$, for $i = 1, \ldots, m$, we define $\Delta_y = [\Delta_{y_1} \ \Delta_{y_2}]$ and see that the system to be solved to compute Δ_y is

$$\begin{aligned} \widehat{A}\widehat{\Theta}\widehat{A}^T\Delta_{\widehat{y}} &= \left[\begin{array}{cc|c} B & C_0 & C_1 \\ C_0^T & D_0 & D_\nabla^T \\ \hline C_1^T & D_\nabla & D_1 \end{array}\right] \left[\begin{array}{c} \Delta_{y_1} \\ \Delta_{y_2} \\ \hline \Delta_t \end{array}\right] = \left[\begin{array}{c} g_1 \\ g_2 \\ \hline g_3 \end{array}\right] \\ &= \widehat{\mathbf{r}}_1 - \widehat{A}\widehat{\Theta}^{-1}\widehat{\mathbf{r}}_2, \end{aligned} \tag{C.20}$$

so that we can sequentially solve the following two systems

$$\begin{bmatrix} D_0 - C_0^T B^{-1} C_0 & D_\nabla^T - C_0^T B^{-1} C_1 \\ D_\nabla - C_1^T B^{-1} C_0 & D_1 - C_1^T B^{-1} C_1 \end{bmatrix} \begin{bmatrix} \Delta_{y_2} \\ \Delta_t \end{bmatrix} = \left(\begin{bmatrix} g_2 \\ g_3 \end{bmatrix} - \begin{bmatrix} C_0^T \\ C_1^T \end{bmatrix} B^{-1} g_1 \right), \tag{C.21}$$

$$B\Delta_{y_1} = \left(g_1 - \begin{bmatrix} C_0 & C_1 \end{bmatrix} \begin{bmatrix} \Delta_{y_2} \\ \Delta_t \end{bmatrix} \right) \tag{C.22}$$

System (C.22) is directly solvable, as $B \in \mathbb{R}^{n \times n}$ is diagonal, so that the main computational effort is to solve (C.21). However, the structure of (C.21) might also been exploited, by noting that $D_1 - C_1^T B^{-1} C_1 \in \mathbb{R}^{n \times n}$ is a diagonal matrix and rewriting (C.21) in the form

$$\begin{aligned} &\left[\begin{array}{c|ccc} \Theta_0 + \sum_{h=1}^n d_h^2 \Upsilon_h & d_1^2 \Upsilon_1 \nabla_{u^1} & \ldots & d_n^2 \Upsilon_n \nabla_{u^n} \\ \hline d_1^2 \nabla_{u^1}^T \Upsilon_1 & \nabla_{u^1}^T \Upsilon_1 \nabla_{u^1} - \Theta_1^s & & \\ \vdots & & \ddots & \\ d_n^2 \nabla_{u^n}^T \Upsilon_n & & & \nabla_{u^n}^T \Upsilon_n \nabla_{u^n} - \Theta_n^s \end{array}\right] \begin{bmatrix} \Delta_{y_2} \\ \Delta_t \end{bmatrix} \\ &= \begin{bmatrix} D_\Upsilon & C_\Upsilon \\ C_\Upsilon^T & B_\Upsilon \end{bmatrix} \begin{bmatrix} \Delta_{y_2} \\ \Delta_t \end{bmatrix} = \begin{bmatrix} g_2 - C_0^T B^{-1} g_1 \\ g_3 - C_1^T B^{-1} g_1 \end{bmatrix}, \end{aligned} \tag{C.23}$$

where

$$\Upsilon_h = \Theta_h^x - \frac{\Theta_h^x P^T P \Theta_h^x}{P \Theta_h^x P^T}, \qquad h = 1, \ldots, n. \tag{C.24}$$

By eliminating Δ_t from the first group of equations in (C.23), we obtain

$$(D_\Upsilon - C_\Upsilon^T B_\Upsilon^{-1} C_\Upsilon)\Delta_{y_2} = g_{\Upsilon_1} \tag{C.25a}$$

$$B_\Upsilon \Delta_t = g_{\Upsilon_2} \tag{C.25b}$$

where $g_{\Upsilon_1} = g_2 - C_0^T B^{-1} g_1 - g_2 - C_\Upsilon^T B_\Upsilon^{-1}(g_3 - C_1^T B^{-1} g_1)$ and $g_{\Upsilon_2} = g_3 - C_1^T B^{-1} g_1 - C_\Upsilon \Delta_{y_2}$. Since B_Υ is diagonal, Δ_t can be directly obtained, so that solving (C.9) – a system of size $2(n+m)$ – reduced to the much smaller problem (C.25a) – a system of size m –.

It might be possible to solve (C.25a) by a preconditioned conjugate gradient, in accordance with the specialized interior point algorithm for primal block-angular problems, introduced by Castro [50, 53].

Appendix D

Random pivoting

In Section 3.3.2 the s-pivots method to generate random networks have been described. We describe in this Appendix some technical details about the computer implementation of the s-pivots methods.

Consider again a polytope $CR(\chi) = \{\mathbf{x} \in [0,1]^{n'} : A\mathbf{x} = \mathbf{b}\}$, where $A \in \mathbb{R}^{m' \times n'}$, $m' < n'$. Based on the equivalence between extreme points and basic solutions, we introduced in Section 3.3.2 the affine transformation

$$\Delta_k(q) = \begin{bmatrix} -B_k^{-1} N_k e_q \\ u_q \end{bmatrix}. \tag{D.1}$$

where u_q is the q-th column vector of the identity matrix and B_k and N_k are the basic and nonbasic submatrices of A. Thus, given a basic solution $\mathbf{x}^k$ we can obtain another one by moving along the simplex-like direction. If the extreme points of $CR(\chi)$ are all integer, it turns out that all basic variables must be at their limits (either at 0 or 1), as well as the non-basic variables. The basic solutions are thus fully degenerate.

From now on we shall use $\delta_B(q)$ to denote in a more compact notation $[\, B^{-1} u_q \mathcal{N}_B^T \mid -u_q \,]^T$, as a function of q.

We say that the set of extreme points which can be achieved from $\mathbf{x}_k$ by applying $\mathbf{x}_k \pm \delta_{\mathcal{B}}(q)$, for some $q = 1, \ldots, n' - m'$, is the set of neighbors of $\mathbf{x}_k$: $\partial(\mathbf{x}_k) = \{y \in \varphi : y = \mathbf{x}_k \pm \delta_{\mathcal{B}}(q);\ q = 1, \ldots, n' - m'\}$.

Let $\mathbf{x}^T = [\mathbf{x}_B^T, \mathbf{x}_N^T]$ for a suitable permutation of the variables. After selecting $q \in \{1, \ldots, n' - m'\}$, $\lambda \in \mathbb{Z}$ must be chosen so that $0 \leq \mathbf{x}_k + \lambda \delta_{\mathcal{B}}(q) \leq 1$, resulting in the minimum ratio equation:

$$\lambda = \begin{cases} \min\left[1, \left\{\dfrac{1 - \mathbf{x}_{B_p}}{-y_p} \colon y(p) < 0\right\}, \left\{\dfrac{\mathbf{x}_{B_p}}{y_p} \colon y_p > 0\right\}\right] & \text{if } \mathbf{x}_{N_p} = 0 \\ \min\left[1, \left\{\dfrac{1 - \mathbf{x}_{B_p}}{y_p} \colon y(p) > 0\right\}, \left\{\dfrac{\mathbf{x}_{B_p}}{-y_p} \colon y_p < 0\right\}\right] & \text{if } \mathbf{x}_{N_p} = 1 \end{cases}, \tag{D.2}$$

where $\mathbf{x}_{N_p}$ is the present value of the component of vector $\mathbf{x}_N$ representing the entering variable, i.e. the current value of the non-basic variable associated with the q^{th} position in the ordering of the non-basic columns, y_p is the p^{th} component of vector $[B^{-1} N u_q^T]$ and $\mathbf{x}_{B_p}$ is the present value of the p^{th} basic variable.

It turns out that, if we let q be a random variable in $\{1, \ldots, n' - m'\}$, this procedure gives rise to a Markov chain over the extreme points of $CR(\chi)$. If the probability distribution of q is properly chosen, that Markov chain is irreducible, as every extreme point might be achieved starting from whatever other.

To draw a sample of size $\overline{k}$ from this chain, a basic inverse B_k^{-1} for every iteration $k = 1 \dots \overline{k}$ is required. The availability of the observed network simplifies the construction of an initial basic matrix. As all components are at their limits, (either 0 or 1), all columns of A might be potentially in the orthogonal basis for its column space, so that the only task reduces to the search for m' linearly independent columns of A.

The availability of the observed network simplifies the construction of an initial basic matrix. As all components are at their limits, (either 0 or 1), all columns of A might be potentially in the orthogonal basis for its column space, so that the only task reduces to the search for m' linearly independent columns of A. Applying a QR-decomposition, matrix A is factorized into a product of an orthogonal matrix $Q \in \mathbb{R}^{m' \times n'}$ and an upper triangular matrix $R \in \mathbb{R}^{m' \times n'}$. The first m' columns of Q form an orthogonal basis for the column space of A. There are several methods for actually computing the QR-decomposition, and most of them terminates in $\mathcal{O}(m'^3)$ operations.

Thus, having a basic feasible solution of a system, $A\mathbf{x} = \mathbf{b}, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}$, the operations which must be iteratively performed are computationally easy: i) finding an entering column $\mathcal{N}u_q$, which allows carrying out a simplex pivot with a non-zero step-length λ; ii) updating the inverse basis B^{-1}, every time a column of the orthogonal basis is replace by a new one.

We built the so-called Sherman-Morrison formula [212] to yield an efficient update of B^{-1} when a single column of B changes. If we wish to replace the j^{th} column of B, call it B_j, with the column $\mathcal{N}_B u_q$, the Sherman-Morrison formula allows obtaining the new inverse basis by means of a simple matrix summation. Let B^{-1} and $\widetilde{B}^{-1}$ be the inverse matrices of B and $\widetilde{B}$ respectively, such that $\widetilde{B}$ differs from B only for its j^{th} column.

Note that we can write $\widetilde{B} = B + (y - B_j)u_j^T$. As $(y - B_j)u_j^T$ is a rank-1 matrix, then $(B+(y-B_j)u_j^T)^{-1}$ is $Q = (I-aB^{-1}(y-B_j)u_j^T)B^{-1}$, where the scalar $a = (1+u_j^T B^{-1}(y-B_j))^{-1}$. Equations (D.3) shows that $(B + (y - B_j)u_j^T)Q = I$, so that Q is the updated inverse.

$$(B + (y - B_j)u_j^T)Q = \tag{D.3a}$$

$$= (B + (y - B_j)u_j^T)\left(I - \frac{B^{-1}(y - B_j)u_j^T}{1 + u_j^T B^{-1}(y - B_j)}\right)B^{-1} = \tag{D.3b}$$

$$= (B + (y - B_j)u_j^T)\left(B^{-1} - \frac{B^{-1}(y - B_j)u_j^T B^{-1}}{1 + u_j^T B^{-1}(y - B_j)}\right) = \tag{D.3c}$$

$$= I + (y - B_j)u_j^T B^{-1} - \frac{(y - B_j)u_j^T B^{-1}(y - B_j)u_j^T B^{-1} + (y - B_j)u_j^T B^{-1}}{1 + u_j^T B^{-1}(y - B_j)} = \tag{D.3d}$$

$$= I + (y - B_j)u_j^T B^{-1} - \frac{(y - B_j)(1 + u_j^T B^{-1}(y - B_j))u_j^T B^{-1}}{1 + u_j^T B^{-1}(y - B_j)} = \tag{D.3e}$$

Note that $(1 + u_j^T B^{-1}(y - B_j))$ is a scalar, so that the resulting expression is

$$(B + (y - B_j)u_j^T)Q = I + (y - B_j)u_j^T B^{-1} - (y - B_j)u_j^T B^{-1} = I. \tag{D.3f}$$

If we write $E = (I - aB^{-1}(y - B_j)u_j^T)$, the update of the inverse basis is obtained by pre-multiplying EB^{-1}. The form of E is the one of an identity matrix of size M, whose p^{th} column is replaced by the vector $[-y(1)/y(p), -y(2)/y(p), \dots, 1/y(p), \dots, -y(M-1)/y(p), -y(M)/y(p)]^T$.

This product form of the inverse is used by almost all commercial linear-programming codes, as only requires $2M^2$ scalar multiplications for each inverse basis update.

Appendix E

MCMC methods to sample from an arbitrary random vector

MetropolisHastings algorithm is a statistical simulation method, design to obtain a sequence of random samples from a probability distribution for which direct sampling is hard to be carried out. The idea behind MCMC is to sample from an arbitrary random vector $\mathbf{x} \in \mathbb{R}^N$ with probability density function f by simulating a Markov chain whose equilibrium distribution is f. Often we only know $f(\mathbf{x})$ up to a proportionality constant independent of $\mathbf{x}$, as in the case described in Section 5.3.

For the purpose of illustration, let us consider the plain vanilla Metropolis algorithm, a special case of the MetropolisHastings algorithm where the proposal distribution is symmetric. The plain vanilla Metropolis algorithm generates a stochastic sequence of states, from an arbitrary starting vector $\mathbf{x}^0$, and the following rule to go from a current state $\mathbf{x}^k$ to a new state $\mathbf{x}^{k+1}$:

1. Propose a candidate state $\mathbf{x}^{k+1}$ from a proposal distribution, $q(y|\mathbf{x}^k)$ that may depend on the current $\mathbf{x}^k$. This proposal distribution is assumed to be symmetric, i.e., $q(\mathbf{x}^{k+1}|\mathbf{x}^k) = q(\mathbf{x}^k|\mathbf{x}^{k+1})$ this can be easily implemented.

2. If the new proposed state $\mathbf{x}^{k+1}$ uses less energy than the current state $\mathbf{x}^k$ then go there with probability one. If the new state is more expensive, in terms of energy, than the one we are currently on, then test your luck and go there with a probability exponentially decreasing in the difference of energy. More formally the acceptance probability is,

$$\alpha(\mathbf{x}^k, \mathbf{x}^{k+1}) = \min\left\{1, \frac{f(\mathbf{x}^{k+1})}{f(\mathbf{x}^k)}\right\} \tag{E.1}$$

A simple modification to the acceptance probabilities used in the plain vanilla algorithm, allows to use non-symmetrical proposal distributions and still have detailed balance, as in the case described in Section 5.3. Change the previous formula for α to,

$$\alpha(\mathbf{x}^k, \mathbf{x}^{k+1}) = \min\left\{1, \frac{q(\mathbf{x}^k|\mathbf{x}^{k+1})f(\mathbf{x}^{k+1})}{q(\mathbf{x}^{k+1}|\mathbf{x}^k)f(\mathbf{x}^k)}\right\} \tag{E.2}$$

All we need to do is to check detailed balance, i.e.,

$$p(y|\mathbf{x}^k)f(\mathbf{x}^k) = p(\mathbf{x}^k|\mathbf{x}^{k+1})f(\mathbf{x}^{k+1}) \text{ for all } \mathbf{x}^k,\ \mathbf{x}^{k+1} \tag{E.3}$$

But this is straight forward to check. When $\mathbf{x}^k = \mathbf{x}^{k+1}$ it is obviously true, and for $\mathbf{x}^k \neq \mathbf{x}^{k+1}$ the only way to arrive to y is by accepting it as a proposed candidate. Thus, the above equation just says that,

$$\alpha(\mathbf{x}^k, \mathbf{x}^{k+1}) q(\mathbf{x}^{k+1}|\mathbf{x}^k) f(\mathbf{x}^k) = \alpha(\mathbf{x}^{k+1}, \mathbf{x}^k) q(\mathbf{x}^k|\mathbf{x}^{k+1}) f(\mathbf{x}^{k+1}) \tag{E.4}$$

which is the same as,

$$\min\left\{1, \frac{q(\mathbf{x}^k|y) f(\mathbf{x}^{k+1})}{q(\mathbf{x}^{k+1}|\mathbf{x}^k) f(\mathbf{x}^k)}\right\} q(y|\mathbf{x}^k) f(\mathbf{x}^k) = \min\left\{1, \frac{q(y|\mathbf{x}^k) f(\mathbf{x}^k)}{q(\mathbf{x}^k|\mathbf{x}^{k+1}) f(\mathbf{x}^{k+1})}\right\} q(\mathbf{x}^k|\mathbf{x}^{k+1}) f(\mathbf{x}^{k+1}) \tag{E.5}$$

which is true, since when the min. on one side is 1, the min. on the other side isn't and the denominator cancels with the term outside the parenthesis producing the equality of both sides. Detailed balance is enough to assure that f is the stationary distribution of the described chain.

Clearly, the Metropolis-Hastings algorithm works best if the proposal distribution matches the shape of the target distribution, allowing a low amount of rejections.

Bibliography

[1] Ahuja, R.K., Magnanti, T.L., Orlin, J.B., (1991). *Network Flows: Theory, Algorithms, and Applications*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

[2] Arrow, K. J., Debreu, G., (1954). Existence of an equilibrium for a competitive economy, *Econometrica*, 22, 265-290.

[3] Arrow, K. J., Debreu, G., (1983). Existence of an equilibrium for a competitive economy, *Econometrica*, 22, 265–290.

[4] Auman R., Dreze, J., (1986). Values of Markets with Satiation or fixed prices, *Econometrica*, 54, 1271–1318.

[5] Axelrod, R., Hamilton, W. D. (1981). The Evolution of Cooperation, *Science*, 211, 1390-96.

[6] Axelrod, R., (1997). *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*, Princeton University Press.

[7] Axelrod, R. (1998). The Dissemination of Culture: A Model with Local Convergence and Global Polarization, *Journal of Conflict Resolution.*

[8] Axtell, R., (1988). The complexity of exchange, *In Working Notes: Artificial Societies and Computational Markets. Autonomous Agents 98 Workshop, Minneapolis/St. Paul (May).*

[9] Aumann, R., Myerson, R., (1988). *Endogenous Formation of Links Between Players and Coalitions: An Application of the Shapley Value*, In: Roth, A.(ed.), The Shapley Value, Cambridge University Press: 175-191.

[10] Babonneau, F., Vial, J.-P., (2009). ACCPM with a nonlinear constraint and an active set strategy to solve nonlinear multicommodity flow problems, *Math. Prog.*, 120, 179–210.

[11] Babonneau, F., Merle, O., Vial, J.-P., (2006). Solving large-scale linear multicommodity flow problems with an active set strategy and proximal-ACCPM, *Oper. Res.*, 54, 184–197.

[12] Barabasi, A.L., R. Albert, (1999). Emergence of scaling in random networks, *Science*, 286: 509512.

[13] Bell, A.M., (1998). Bilateral Trading on Network: A Simulation Study, *In Working Notes: Artificial Societies and Computational Markets. Autonomous Agents 98 Workshop, Minneapolis/St. Paul.*

[14] Berghammer, R., Rusinowska A., de Swart, H., (2010), Applying relation algebra and RelView to measures in a social network, *European Journal of Operational Research*, 202, 182–195.

[15] Berkowitz, L, (1956). Personality and Position, *Sociometry*, 19, 210–22.

[16] Bernard, H.R., Killworth, P.D., (1977). Informant accuracy in social network data II, *Human Comunication Research*, 4, 3–18.

[17] Bernard, H.R., Kdlworth, P.D., Sailer, L., (1980). Informant accuracy in social-network data V. An expernnental attempt to predict actual communication from recall data, *Social Sciences Research* 11: 30-66.

[18] Bernard, H.R., Kdlworth, P.D., Sailer, L., (1981). A note on inferences regarding networks subgroups: Response to Burt and Bittner. *Social Networks*, 3, 89–92.

[19] Barrat, A., Barthelemy, M., Pastor-Satorras, R., Vespignani, A., (2004). The architecture of complex weighted networks, *PNAS* 101, 3747–3752.

[20] Barrat, A., Barthelemy, M., Vespignani, A., (2004). Modeling the evolution of weighted networks, *cond-mat*, 0406238, 3747-3752.

[21] Barrat, A., Barthelemy, M., Vespignani, A., (2004). Weighted evolving networks: coupling topology and weights dynamics, *Phys. Rev. Lett.*, 92, 228–701.

[22] Barry, R.P., Pace, R. K., (1999). Monte Carlo estimates of the log determinant of large sparse matrices, *Linear Algebra and its Applications*, 289, 41–54.

[23] Bertalanffy, L., (1988), *General system theory*, G. Braziller New York.

[24] Biggs, N., Lloyd, E., Wilson, R., (1986). Graph Theory, 1736-1936, *Oxford University Press.*

[25] Bjorck, A., Golub, G.H., (1977), Numerical methods for computing angles between linear subspaces, *Math. Comp.*, 27, 579–594.

[26] Besag, J., (1975). Statistical Analysis of Non-Lattice Data, *The Statistician.* 24, 179-195

[27] Bienstock, D., (2002). *Potential Function Methods for Approximately Solving Linear Programming Problems.* Theory and Practice, Kluwer, Boston.

[28] Bocanegra, S., Castro, J., Oliveira, A.R.L., (2013). Improving an interior-point approach for large block-angular problems by hybrid preconditioners, *European Journal of Operational Research*, 231, 263–273

[29] Bollobas, B., (1985). *Random Graphs.* Cambridge University Press.

[30] Bollobs, B., (1998). *Modern Graph Theory.* Springer, New York.

[31] Bonacich, P., (1972), Factoring and weighting approaches to clique identification, *Journal of Mathematical Sociology*, 2, 113-120.

[32] Bonacich, P., (1987), Power and Centrality: A Family of Measures, *Journal of Mathematical Sociology*, 92, 1170–1182.

[33] Bonacich, P., (2007), Some unique properties of eigenvector centrality, *Journal of Social Networks* 29, 555-564.

[34] Borgatti, S.P., Everett, M.G. and Freeman, L.C, (2002), Ucinet for Windows: Software for Social Network Analysis, *Harvard, MA: Analytic Technologies.*

[35] Boorman, S.A., (1975). A Combinatiorial Optimization Model for Transmission of Job Information Through Contact Networks, *The Bell Journal of Economics*, 6, 216–249.

[36] Bourdieu, P., (1986). The forms of capital, *in J.E. Richardson, Handbook of Theory of Research for the Sociology of Education, Greenwood Press*, 241-258.

[37] Borgatti, S.P., Mehra, A., Brass, D.J., Labianca, G., (2009). Network Analysis in the Social Sciences, *Science*, 323, 892–895.

[38] Box, G., (1976). Science and Statistics, *Journal of the American Statistical Association*, 356, 91–799

[39] Breiger, R.L., Boorman, S. A., Arabie, P., (1975). An Algorithm for Clustering Relation Data, with Application to Social Network Analysis and comparison with Multidimensional Scaling, *Journal of Mathematical Psycology*, 12, 328–383.

[40] Breiger, R., Pattison, P., (1986). Cumulated social roles: The duality of persons and their algebras. *Social Networks*, 8, 215–256.

[41] Brin, S., Page, L, (1998). The anatomy of a large-scale hypertextual Web search engine, *Computer Networks and ISDN Systems*, 30, 107-117.

[42] Burt, R.S., (1992). *Structural Holes: The Social Structure of Competition*, Harvard University Press.

[43] Burt, R.S., (2000). The network structure of social capital, *Research in Organizational Studies*, 22, 345-423.

[44] Buskens, V., Rijt, A., (2009). Dynamics of networks if everyone strives for structural holes, *American Journal of Sociology*, 114, 371-407.

[45] Butts, C.T., (2009). Revisiting the Foundations of Network Analysis, *Science*, 325, 414–416.

[46] Caimo, A., Friel, N., (2011). Bayesian inference for exponential random graph models, *Social Networks*, 33, 41–55.

[47] Callaway, DS., Newman, MEJ., Strogatz, SH., Watts, DJ., (2000). Network Robustness and Fragility: Percolation on Random Graphs *Phys Rev Lett*, 85 5468-5471.

[48] Caplin A., Leahy, J., (2010). A Graph Theoretic Approach to Markets for Indivisible Goods, *Mimeo, New York University*, NBER Working Paper 16284.

[49] Carrington, P.J., Scott, J., Wasserman, S., (2005). *Models and methods in social network analysis*, Cambridge University Press.

[50] Castro, J., (2000). A specialized interior-point algorithm for multicommodity network flows, *SIAM J. Optim.*, 10, 852–877.

[51] Castro, J., (2003). Solving difficult multicommodity problems through a specialized interior-point algorithm, *Ann. Oper. Res.*, 124, 35–48.

[52] Castro, J., (2003). Solving quadratic multicommodity problems through an interior-point algorithm, *System Modelling and Optimization*, 199–212.

[53] Castro, J., (2007). An interior-point approach for primal block-angular problems, *Comput. Optim. Appl.*, 36, 195–219.

[54] Castro, J. (2014). Interior-point solver for convex separable block-angular problems, *Research Report DR 2014/03. Dept. of Statistics and Operations Research, Universitat Politcnica de Catalunya.*

[55] Castro J., Cuesta J., (2011). Quadratic regularizations in an interior-point method for primal block-angular problems. *Mathematical Programming*, 130, 415–445

[56] Castro, J., Cuesta, J., (2010). Existence, uniqueness and convergence of the regularized primal-dual central path, *Oper. Res. Letters*, 38, 366–371.

[57] Castro, J., Nabona, N., (1996). An implementation of linear and nonlinear multicommodity network flows, *Eur. J. Oper. Res.*, 92, 37–53.

[58] Centola D., Gonzalez-Avella, J., Eguluz, V.E., San Miguel, M., (2007). Homophily, Cultural Drift, and the Co-Evolution of Cultural Groups, *Journal of Conflict Resolution*, 51, 905–929

[59] Chandler, D. (2000). Marxist Media Theory. *Aberystwyth University.*

[60] Chandler, D. (1987). Introduction to Modern Statistical Mechanics, *Oxford University Press.*

[61] Chardaire, P., Lisser, A., (2002). Simplex and interior point specialized algorithms for solving nonoriented multicommodity flow problems, *Oper. Res.*, 50, 260–276.

[62] Charon, I., Germa, A., Hudry, O., (1996), Random generation of tournaments and asymmetric graphs with given out-degrees, *European Journal of Operational Research*, 95, 411–419.

[63] Chen, Y., Diaconis, P., Holmes, S., Liu, J., (2005), Sequential Monte Carlo methods for statistical analysis of tables. *Journal of the American Statistical Association*, 100, 109-120.

[64] Chen, Y., (2006), Simple existence conditions for zero-one matrices with at most one structural zero in each row and column. *Discrete Mathematics*, 306, 2870-2877.

[65] Chierichetti, F., Lattanzi, S., Panconesi, A., (2009). Rumour spreading and graph conductance *SODA 10 Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, 1657–1663.

[66] Chung, Fan R. K. (1997), *Spectral Graph Theory*, American Mathematical Society.

[67] Coleman, J.S., (1961). *The Adolescent Society.* Free Press.

[68] Coleman, J.S., (1988). Social capital in the creation of human capital, *American Journal of Sociology*, 94, 95-120.

[69] Coleman, J.S., (1990). Foundations of Social Theory, *Harvard University Press.*

[70] Conejo, A.J., Castillo, E., Minguez, R., Garcia-Bertrand, R., (2006). *Decomposition Techniques in Mathematical Programming. Engineering and Science Applications*, SpringerVerlag, Berlin.

[71] Connor R.C., Heithaus M.R., Barre L.M., (1999). Superalliance of bottlenose dolphins, *Nature* 397: 571-572.

[72] Cook, K. S., R. M. Emerson, M. R. Gilmore, and T. Yamagishi, (1983). The Distri- bution of Power in Exchange Networks: Theory and Experimental Results, *American Journal of Sociology* , 89:275-305.

[73] Corley, H.W., Moon, I.D., (1985). Shoertest Path in Network with Vector Weights, *Journal of Optimization Theory and Applications* 46: 79–86.

[74] Cover, T.M., Thomas, J.A., (1991), *Elements of Information Theory*, Wiley.

[75] Cvetkovic, D.M., Doob, M., Sachs, H., (1980), *Spectra of graphs: theory and application*, Academic Press.

[76] Danilov, V., Koshevoy, G., Murota, K., (2001). Discrete convexity and equilibria in economies with indivisible goods and money, *Journal of Mathematical Social Sciences*, 41, 3, 251–273.

[77] Davidsen, J., Ebel, H., Bornholdt, S., (2002). Emergence of a Small World from Local Interactions: Modeling Acquaintance Networks, *Physical review letters*, 88()12, 128701.

[78] Davis, G.F., Greve, H.R., Sociol, A.J., (1997). Elite Networks and Governance Changes in the 1980s, *American Journal of Sociology*, Vol. 103, 1–37.

[79] Deb, K., (2001). Multi-objective optimization using evolutionary algorithms, *John Wiley and Sons*.

[80] Debreu, G., (1983), Economic theory in the mathematical mode, *Nobel Memorial lecture* 12.

[81] Doreian, P., Batagelj, V., Ferligoj, A., (2005), *Generalized Blockmodeling (Structural Analysis in the Social Sciences)*, Cambridge University Press.

[82] Dorogovtsev, S.N., Mendes, J.F.F., (2004). Minimal models of weighted scale-free networks, *arXiv:cond-mat*/0408343, Preprint. (http://arxiv.org/abs/cond-mat/0408343)

[83] Dreze J. (1975), Existence of an exchange equilibrium under price rigidities. *International Economic Review*, 16, 2, 301–320.

[84] Easley, D., Kleinberg, J., (2010). Networks, Crowds, and Markets: Reasoning About a Highly Connected World, *Cambridge University Press*.

[85] Edgeworth, F.Y., (1881). Mathematical Psychics, An Essay on the Application of Mathematics to the Moral Sciences, *in (1967) L.S.E. Series of Reprints of Scarce Tracts in Economics and Political Sciences* 10.

[86] Erdos, P., Rainyi, A., (1959). On random graphs, *Publicationes Mathematicae* 6: 290-297.

[87] Euler, L., (1736). Solutio problematis ad geometriam situs pertinentis, *Commentarii Academiae Scientiarum Imperialis Petropolitanae* 8, 128–140.

[88] Feld, S., (1991). Why your friends have more friends than you do, *Am. J. Sociol* 96: 1464–1477.

[89] Feldman, A., (1973). Bilateral Trading Processes, Pairwise Optimality, and Pareto Optimality, *Review of Economic Studies*, XL (4), 463–473.

[90] Frangioni, A., Gallo, G., (1999). A bundle type dual-ascent approach to linear multicommodity min cost flow problems, *INFORMS J. Comp.* 11, 370–393.

[91] Frangioni, A., Gentile, C., (2004). New preconditioners for KKT systems of network flow problems, *SIAM J. Optim.* 14, 894–913.

[92] Frank, O. (1971) *Statistical Inference in Graphs*, PhD Thesis, Stockholm University.

[93] Frank, O., Strauss, D., (1986). Markov graphs, *Journal of the American Statistical Association* 81: 832-842.

[94] Franklin, N., (2002). Methods of Mathematical Economics: Linear and Nonlinear Programming, Fixed-Point Theorems, SIAM, *Business and Economics.*

[95] Friedkin, N., (1980). A Test of Structural features of Granovette's strength of weak ties theory, *Social Networks*, 2, 411–422.

[96] Freeman, L. (1979), Centrality in social networks: Conceptual clarification, *Journal of Social Networks*, 1(3): 215-239.

[97] Fremlin, D.H. (2010), *Measure Theory, Volume 2*, Torres Fremlin.

[98] Gale, D., (1957). A theorem on flows in networks, *Pacific Journal of Mathematics*, 7: 1073-1082.

[99] Galeotti, A., Goyal, S., Jackson, M.O., Vega-Redondo, F., Yariv. L. (2010). Network Games. *The Review of Economic Studies.* 77(1), 218–244

[100] Goffin, J. L., Gondzio, J., Sarkissian, R., Vial, J.P., (1996). Solving nonlinear multicommodity flow problems by the analytic center cutting plane method, *Math. Prog.* 76, 131–154.

[101] Golub, G.H., Van Loan, C.F., (1996). *Matrix Computations*, Third Ed., Johns Hopkins Univ. Press, Baltimore.

[102] Gondzio J (1996) Multiple centrality corrections in a primal dual method for linear programming. Comput Opt App. 6:137–156

[103] Gondzio J, Sarkissian R (2003) Parallel interior-point solver for structured linear programs. Math. Prog. 96: 561–584

[104] Hummon, N.P., Doreian, P., (1989). Connectivity in a citation network: the development of a DNA theory, *Social Networks*, Vol. 11, 1989, 39–63.

[105] Golub, B., Jackson, M.O., (2012). Does Homophily Predict Consensus Times? Testing a Model of Network Structure via a Dynamic Process, *Review of Network Economics*, 11, 1–28.

[106] Golender, V. E. , Drboglav, V.V., Rosenblit, A.B., (1981). Graph Potentials Method and Its Application for Chemical Information Processing, *J. Chem. If. Comput. Sci*, 21, 196–204.

[107] Granovetter, M., (1970). Changing Jobs: Channels of Mobility Information in a Suburban Community. *Doctoral dissertation, Harvard University.*

[108] Granovetter, M., (1973). The strength of weak ties, *American Journal of Sociology*, 78, 1360-1380.

[109] Granovetter, M., (1985). Economic action and social structure: The problem of embeddedness, *American Journal of Sociology*, 91, 481-510.

[110] Granovetter, M., (1992), Problems of explanation in economic sociology, *in Nohria, N., Eccles, R.(eds), Networks and Organizations: Structure, Form and Action, Harvard Business School Press.*

[111] Haimes, Y.Y., Lasdon, L.S., Wismer, D.A., (1971). On a bicriterion formulation of the problems of integrated system identification and system optimization, *IEEE Transactions on Systems, Man, and Cybernetics*, 1(3), 296–297.

[112] Handcock, M., (3003), Assessing degeneracy in statistical models of social networks, *working paper n. 39*

[113] Harary, F., (1953). On the notion of balance of a signed graph. *Michigan Math. Journal* 2, 143–146.

[114] Harrison G. W., HirshleiferSource, J., (1989). An Experimental Evaluation of Weakest Link/Best Shot Models of Public Goods Author(s), *The Journal of Political Economy*, 97, 201–225.

[115] Heller, I., Tompkins, C.B.Gh, (1956). 'An Extension of a Theorem of Dantzig's', in K. Kuhn, A. Tucker, *Linear Inequalities and Related Systems*, Princeton University Press, Princeton, 247–254.

[116] Heider, F. (1958). *The Psychology of Interpersonal Relations*, John Wiley and Sons.

[117] Hirshleifer, J. (1983) "From Weakest-Link to Best-Shot: The Voluntary Provision of Public Goods". Public Choice, 41, 371–386.

[118] Holland, P.W., Leinhardt, S., (1976). Local structure in social networks. *In D. Heise (Ed.), Sociological Methodology*, 1: 1–45.

[119] Horn, R.A., Johnson, C.R., (1991). *Topics in Matrix Analysis*, Cambridge University Press.

[120] Hummon, N.P., Doreian, P., (1989). Connectivity in a citation network: the development of a DNA theory, *Social Networks*, 11, 39–63.

[121] Jackson, M.O., (2008). *Social and Economic Networks*, Princeton University Press.

[122] Jackson, M.O., Rogers, B.W, (2004). The Economics of Small Worlds, *Journal of the European Economic Association* 3, 617–627.

[123] Jackson, M.O., Lopez-Pintado, D., (2011). Diffusion and Contagion in Networks with Heterogeneous Agents and Homophily, *Available at SSRN: ssrn.com/abstract=1950476 or http://dx.doi.org/10.2139/ssrn.1950476.*

[124] Jackson, M.O., Watts, A., (2001). The existence of pairwise stable networks, *Seoul Journal of Economics*, 14, 299–321.

[125] Jackson, M.O., Wolinsky, A., (1996). A Strategic Model of Social and Economic Networks, *Journal of Economic Theory*, 71, 44-74.

[126] Jaynes, E.T., (1957), Information theory and statistical mechanics, *Phys. Rev*, 106, 620.

[127] Jevons, W. S., (1871). The Theory of Political economy, *Macmillan & Co.*

[128] Jerrum, M., Sinclair, A., (1988). Approximating the permanent, *SIAM Journal on Computing*, 18(6), 1149-1178.

[129] Kalna, G., Higham, D.J., (2007). A Clustering Coefficient for Weighted Networks, with Application to Gene Expression Data, *AI Communications - Network Analysis in Natural Sciences and Engineering archive* 20: 263-271.

[130] Kang, S.M., (2007). Equicentrality and network centralization: A micromacro linkage, *Social Networks*, 29, 585-601.

[131] Kaneko, M., (1982). The Central Assignment Game and the Assignment Markets, *Journal of Mathematical Economics*, 10, 205–232.

[132] Kapferer, B., (1969). *Norms and the manipulation of relationships in a work context*, in *Social Networks in Urban Situations*, edited by J. C. Mitchell, Manchester University Press.

[133] Kapferer, B., (1972). Strategy and transaction in an African factory, *Manchester University Press.*

[134] Karlberg, M., (1999). Testing transitivity in digraphs. *Sociological Methodology*, 29, 225-251

[135] Katz, L., (1953). A New Status Index Derived from Sociometric Index. *Psychometrika*, 18, 39–43.

[136] Kearns, M., Littman, M., Singh., S., (2001). *Graphical models for game theory.* In Proceedings of the Conference on Uncertainty in Artificial Intelligence.

[137] Kimura, M., (1953). Stepping-stone model of population, *Annual Report of the National Institute of Genetics*, 3, 62-63.

[138] Kleinberg, J., Suri, S., Tardos, E., Wexler, T., (2008). Strategic network formation with structural holes, *In Proc. 9th ACM Conference on Electronic Commerce.*

[139] Klemm, K., Eguiluz, V. M., Toral, R., San Miguel, M., (2003). Global culture: A noise-induced transition in finite systems, *Phys. Rev. E*, 67, 045101(R).

[140] Klemm, K., Serrano, M.A., Eguiluz, V.M., Miguel, V.M., (2012). A measure of individual role in collective dynamics, *arXiv:1002.4042.*

[141] Kochen, M., Pool, I., (1978). Contacts and Influence, *Social networks*, 1, 5–51.

[142] Korte, B., Jens, V., (2008). *Combinatorial Optimization: Theory and Algorithms.* Berlin: Springer-Verlag.

[143] Konig, D. (1936). *Theorie der endlichen und unendlichen Graphen, Leipzig: Akademische Verlagsgesellschaft.* Translated from German by McCoart, R., *Theory of finite and infinite graphs*, Birkhuser, 1990.

[144] Lawler, E. (1976). *Combinatorial Optimization: Networks and Matroids.* Holt, Rinehart and Winston.

[145] Leavitt, H. J, (1951). Some Effects of Certain Communication Patterns on Group Performance, *Journal of Abnormal and Social Psychology*, 46:38-50.

[146] Lehmann, E.L., Romano, J.P., (2005, 3rd edition). *Testing statistical hypotheses*, Springer, New York.

[147] Lemaréchal, C., Ouorou, A., Petrou, G., (2009). A bundle-type algorithm for routing in telecommunication data networks, *Comput. Optim. Appl.* 44, 385–409.

[148] Leskovec, J.A. Damic, L., Huberman, B., (2006). The dynamics of viral marketing, *ACM Transactions on the Web 1*, 1, 5.

[149] Lin N., (2002). Social Capital: A Theory of Social Structure and Action, *Cambridge University Press.*

[150] Lopez, L., Robles, G., Gonzalez, J., (2004). *Applying social network analysis to the information in CVS repositories*, in: Proc. of the 1st Intl. Workshop on Mining Software Repositories, 101-105.

[151] Lorrain, F.,White, H. C., (1971). Structural Equivalence of Infividuals in Social Networks, *Journal of Mathematical Sociology*, 1, 49–80.

[152] Lusseau D., (2003). The emergent properties of a dolphin social network, *Proceedings of the Royal Society of London Series B-Biological Sciences*, 270, 186-188.

[153] Lusseau D., (2004). Identifying the role that individual animals play in their social network, *Proceedings of the Royal Society of London Series B-Biological Sciences*, 271, S477-S481.

[154] Lusseau, D., Schneider, K., Boisseau, O.J., Haase, P., Slooten, E., Dawson, S.M., (2003). The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations, *Behavioral Ecology and Sociobiology*, 54, 396-405.

[155] Manca, A., Sechi, G., Zaddas, P., (2010). Water Supply Network Optimisation Using Equal Flow Algorithms *Water Resour Manage*, 24, 3665-3678.

[156] Marx, K., (1859). *A Contribution to the Critique of Political Economy*, Progress Publishers.

[157] Masuda, N. Aihara, K., (2003). Spatial prisoners dilemma optimally played in small-world networks, *Physics Letters A*, 313, 55–61.

[158] Mathias N., Gopa, V., (2001). Small worlds: how and why, *Physical Review E*, 63, 021117.

[159] McBride, R.D., (1998). Progress made in solving the multicommodity flow problem, *SIAM J. Optim.* 8, 947–955.

[160] McDonald, J.W., Smith, P.W., Forster, J.J., (2007). Markov chain monte carlo exact inference for social networks, *Social Networks* 29, 127–136.

[161] Merris, R., (1997). Doubly stochastic graph matrices, *Univ. Beograd. Publ. Elektrotehn. Fak. Ser Mat*, 8, 6471.

[162] Meyer, C. (2000), Matrix analysis and applied linear algebra, SIAM.

[163] Mintz, B., Schwartz, M., (1985). *The Power Structure of American Business*, Chicago: University of Chicago Press.

[164] Montgomery, J.D. (1992). Job Search and Network Composition: Implications of the Strength-of-Weak-Ties Hypothesis, *American Sociological Review*, 57, 586-96.

[165] Morris, M., Handcock, M.S., Hunter, D.R., (2008). Specification of Exponential-Family Random Graph Models, *Journal of Statistical Software* 24(4).

[166] Moreno, J.L., (1934). *Who Shall Survive?* Washington, DC: Nervous and Mental Disease Publishing Company. *Available at http://www.asgpp.org/docs/WSS/WSS.html.*

[167] Myerson, R., (1991). *Game Theory: Analysis of Conflict*, Harvard University Press.

[168] Nash, J.F, (1951). The bargaining problem, *Econometrica*, 18, 155–162.

[169] Newman, M.E.J., (2001). Scientific collaboration networks: II. Shortest paths, weighted networks, and centrality, Phys. *Rev. E 64*, 016132.

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

[171] Newman, M.E.J., (2002). Random graphs as models of networks, *Working Papers 02-04-020, Santa Fe Institute.*

[172] Newman, M.E.J., (2002). Assortative mixing in networks. *Phys. Rev. Lett.* 89, 208701.

[173] Newman, M.E.J., Park, J., (2003). Why social networks are different from other types of networks, *Phys. Rev. E*, 68, 036122.

[174] Newman, M.E.J., (2004), Detecting community structure in networks, *The European Physical Journal B*, 38, 321–330.

[175] Newman, M.E.J., (2005). Power laws, Pareto distributions and Zipf's law, *Phys. Rev. Lett*, 46: 323-351.

[176] Ng, E., Peyton, BW., (1993). Block sparse Cholesky algorithms on advanced uniprocessor computers. *SIAM J Sci Comput*, 14, 1034–1056.

[177] Nisan, N. Roughgarden, T., Tardos, E., Vazirani, V., (2007). *Algorithmic Game Theory*, Cambridge University Press.

[178] Ouorou, A., (2007), Implementing a proximal point algorithm to some nonlinear multicommodity flow problems, *Networks*, 18, 18–27.

[179] Ouorou, A., Mahey, P., Vial, J.-P., (2000). A survey of algorithms for convex multicommodity flow problems, *Manag. Sci.* 46, 126–147.

[180] Ozlen, M., Azizoglu, M., (2009). Multi-objective integer programming: a general approach for generating all non-dominating solutions, *European Journal of Operational Research*, 199(1), 25–35.

[181] Ozlen, M., Azizoglu, M., Burton, B. A., (Accepted 2012). Optimising a nonlinear utility function inmulti-objective integer programming, *Journal of Global Optimization*, to appear.

[182] Padberg, M., (1999). *Linear optimization and extensions.* 2nd revised and expanded edition, Springer-Verlag, New York.

[183] Padgett, J.F., (1994). Marriage and Elite Structure in Renaissance Florence, 1282–1500, *Conference paper for Social Science History Association annual meeting*, St. Louis, Mo.

[184] Page, L., Brin, S., Motwani, R., Winograd, T. (1999). The PageRank citation ranking: Bringing order to the Web. *published as a technical report*: http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf.

[185] Panait, L., Luke, S., (2005), Cooperative Multi-Agent Learning: The State of the Art, *Autonomous Agents and Multi-Agent Systems*, 11, 387-434.

[186] Pareto, V., (1932). Di un errore del Cournot nel trattare leconomia politica colla matematica, *Giornale degli Economisti*, 4, 1-14.

[187] Park, J., Newman, M.E.J., (2004). Statistical Mecanics of Networks, *Phisical Review E*, 70, 066117.

[188] Penrose, M. (2003). *Random Geometric Graphs*, Oxford University Press.

[189] Pevzner, A., (1995). DNA physical mapping and alternating Eulerian cycles in colored graphs, *Algorithmica*, 13, 77–105.

[190] Portes, A., (1998). Social capital: Its origins and applications in modern sociology, *Annual Review of Sociology*, 24, 1-24.

[191] Porter, M. A., Onnela, J.P., Mucha, J.P., (2009). Communities in Networks, *Not. Amer. Math. Soc.*, 56, 10821097.

[192] Putnam, R. D., (1995). Bowling Alone: America's Declining Social Capital. *Journal of Democracy*, 6, 65-78.

[193] Quinzii, M., (1984). Core and Competitive Equilibria with Indivisibilities, *International Journal of Game Theory*, 13, 41–60.

[194] Rao, A., Jana, A., Bandyopadhyay, S., (1996). A markov chain monte carlo method for generating random (0,1) matrices with given marginals. *Sankhya, ser.*, A 58, 225–242.

[195] Rashevsky, N., (1938). Mathematical Biophysics: Physico-Mathematical Foundations of Biology, *University of Chicago Press : Chicago Press*.

[196] Rashevsky, N., (1947). *Mathematical Theory of Human Relations: An Approach to Mathematical Biology of Social Phenomena*, Bloomington, ID: Principia Press.

[197] Robert, C., Casella, G., (2004), *Monte Carlo Statistical Methods*, Springer.

[198] Roberts, J.M. (2000). Simple methods for simulating sociomatrices with given marginal totals, *Social Networks* 22, 273–283.

[199] Roberts, G.O., Tweedie, R.L., (1996). Geometric convergence and central limit theorems for multidimensional Hastings and Metropolis algorithms, *Biometrika*, 83, 95–110.

[203] Roberts, J.M. (2000). Simple methods for simulating sociomatrices with given marginal totals, *Social Networks*, 22, 273–283.

[201] Rubinstein, A., (1983). Perfect Equilibrium in a bargaining model, *Econometrica*, 50, 97-109.

[202] Ryser, H.J., (1957). Combinatorial properties of matrices of zeros and ones, *Canadian Journal of Mathematics*, 9, 371–377.

[203] Roberts, Jr.J., (2000). Simple methods for simulating sociomatrices with given marginal totals. *Social Networks* 22, 273–283.

[204] Sampson, S. (1969). Crisis in a cloister. *Unpublished doctoral dissertation*, Cornell Universit.

[205] Sastry, V.N., Ismael Mohideen, S., (1999). Modified Algorithm to Compute Pareto-Optimal Vectors, *Journal of Optimization Theory and Applications*, 103, 241–244.

[206] Sbalzariniy, I.F. , Mullery, S., Koumoutsakosyz, P., (2000). *Multi-objective optimization using evolutionary algorithms*, Center for Turbulence Research, Proceedings of the Summer Program.

[207] Scarf, H., (1994). The Allocation of Resources in the Presence of Indivisibilitie, *Journal of Economic Perspectives*, 8, 111–128.

[208] Schaeffer, S., (2007). Graph clustering, *Computer Science Review*, 1, 27-64.

[209] Schrijver, A., (1998). *Theory of linear and integer programming*, John Wiley and Sons.

[210] Schrijver, A., (2003). Combinatorial Optimization. *Springer-Verlag.*

[211] Shapley, L., Shubik, M., (1972). The Assignment Game I: The Core, *International Journal of Game Theory*, 1, 111–130.

[212] Sherman, J., Morrison, W.J., (1949). Adjustment of an Inverse Matrix Corresponding to Changes in the Elements of a Given Column or a Given Row of the Original Matrix, *Annals of Mathematical Statistics*, 21, 124-127.

[213] Shoham, Y., (1993). Agent-oriented programming, *Artificial Intelligence*, 60, 51–92.

[214] Shu, P., Tang, M., Gong,K., Liu, Y., (2012). Effects of Weak Ties on Epidemic Predictability in Community Networks, *Physics.soc-ph, arXiv:1207.0931.*

[215] Skvoretz, J., Agnenssens, F., (2007). Reciprocity, Multiplexity, and Exchange: Measures, *Quality and Quantity*, 41, 341-357.

[216] Snijders, T.A.B., (1991). Enumeration and simulation methods for 0-1 matrices with given marginals, *Psychometrika*, 56, 397–417.

[217] Sinclair, A., Jerrum, M., (1989), Approximate counting, uniform generation and rapidly mixing Markov chains. *Information and Computation*, 82(1), 93-133.

[218] Takaguchi, T., Miyazaki, S., (2009) Spectra Statistics of the Transition Matrices on Watts-Strogatz Model, *Forma*, 24, 37-40.

[219] Tarjan, R.E., (1974). A note on finding the bridges of a graph, *Information Processing Letters*, 160-161.

[220] Uzawa, H., (1962). On the Stability of Edgeworth's Barter Process, *International Economic Review*, 3, 218-232.

[221] Vazirani, V., Nisan, N., Roughgarden, T., Tardos, E., (2007), *Algorithmic Game Theory*, (1st ed.). Cambridge, Cambridge University Press.

[222] Verhelst, N.D., (2008). An effcient MCMC algorithm to sample binary matrices with fixed marginals. *Psychometrika*, 73(4), 705-728.

[223] Vickrey, D., Koller, D., (2002). *Multi-agent algorithms for solving graphical games.* In Proceedings of the National Conference on Artificial Intelligence.

[224] Wasserman, S., Faust, K., (1994). Social Network Aanalyris, *Carnbridge: University Press.*

[225] Wasserman, S., Pattison, P. E., (1996). Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and p*, *Psychometrika*, 61, 401–425.

[226] Watts, J.D., Strogatz, S.H., (1998). Collective dynamics of 'small-world' networks. *Nature*, 393, 440-442

[227] Watts. D.J., (1999). *Small Worlds: The Dynamics of Networks Between Order and Randomness*, Princeton University Press.

[228] White, H. C., (1963). The anatomy of kinship: Mathematical models for structures of cumulated social roles, *Englewood Cliffs, NJ: Prentice-Hall.*

[229] Wilhite, A., (2001). Bilateral Trade and Small-Worl Networks,*Computational Economics*, 18, 49-44.

[230] Wooldridge, M., (2002). *An Introduction to MultiAgent Systems*, (1st ed.). Krst Sussex, John Wiley and Sons Ltd.

[231] Wright, S.J., (1996). Primal-Dual Interior-Point Methods, SIAM, Philadelphia

[232] Xu, R., Wunsch, D., (2005). Survey of Clustering Algorithms, *Transactions on Neural networks*, 16, 645–678.

[233] Ye, Y. (1997). *Interior Point Algorithms. Theory and Analysis*, Wiley.

[234] Zhang, Y., (1998). Solving large-scale linear programs by interior-point methods under the MATLAB environment, *Opt Methods Soft.*, 10, 1–31.