



**HAL**  
open science

# Random Digraphs with Given Expected Degree Sequences: A Model for Economic Networks

Mauro Gallegati

► **To cite this version:**

Mauro Gallegati. Random Digraphs with Given Expected Degree Sequences: A Model for Economic Networks. *Journal of Economic Behavior and Organization*, 2011, 10.1016/j.jebo.2011.01.022 . hal-00989518

**HAL Id: hal-00989518**

**<https://hal.science/hal-00989518>**

Submitted on 12 May 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Accepted Manuscript

Title: Random Digraphs with Given Expected Degree Sequences: A Model for Economic Networks

Author: Mauro Gallegati

PII: S0167-2681(11)00052-7  
DOI: doi:10.1016/j.jebo.2011.01.022  
Reference: JEBO 2690

To appear in: *Journal of Economic Behavior & Organization*

Received date: 25-11-2010  
Revised date: 19-1-2011  
Accepted date: 24-1-2011

Please cite this article as: Gallegati, M., Random Digraphs with Given Expected Degree Sequences: A Model for Economic Networks, *Journal of Economic Behavior and Organization* (2010), doi:10.1016/j.jebo.2011.01.022

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Elsevier Editorial System(tm) for Journal of Economic Behavior & Organization  
Manuscript Draft

Manuscript Number: JEBO-D-10-00619R1

Title: Random Digraphs with Given Expected Degree Sequences: A Model for Economic Networks

Article Type: Full Length Article

Keywords: diversification; complex networks; credit networks; expected degree model; modularity; community structure of networks

Corresponding Author: Dr. Mauro Gallegati,

Corresponding Author's Institution: Universita' Politecnica delle Marche

First Author: Mauro Gallegati, Full Professor

Order of Authors: Mauro Gallegati, Full Professor; leonardo bargigli, Ph.D.

Abstract: Building upon the growing interest for complex network theory, the main ambition of this paper is to contribute to a more effective application of network theory to economic phenomena, and particularly to financial networks. We depart from recent contributions on credit networks in two respects. In the first place, we constrain diversification with the out- and in-degree distribution of nodes, by adopting a suitable extension of the expected degree model for random weighted digraphs. In the second place, we focus on real networks by using this extension as null model for statistical analysis. With the help of statistical inference, we provide a rigorous answer to the following question: do real networks obey our null model? Further, we show that this answer is tightly connected to the existence of clusters or modules, as defined by Newman and Girvan, within networks.

\title{Random Digraphs with Given Expected Degree Sequences: A Model for Economic Networks}  
\author{Leonardo Bargigli<sup>a</sup> \ Mauro Gallegati<sup>a</sup> \ \begin{normalsize}<sup>a</sup>  
Universit<sup>a</sup> Politecnica delle Marche \end{normalsize}  
\ \begin{normalsize}Department of Economics  
\end{normalsize} \ \footnotesize Correspondence to: \ \footnotesize mauro.gallegati@univpm.it \ \footnotesize  
\footnotesize Phone: +39 071 220 7088 \ \footnotesize Fax: +39 071 220 7102  
\footnote{Authors acknowledge the financial support from the European Community Seventh  
Framework Programme (FP7/2007-2013) under Socio-economic Sciences and  
Humanities, grant agreement no. 225408 (POLHIA). We are especially thankful to Hideaki Aoyama  
and Yoshi Fujiwara for sharing Japanese network data.}}

Accepted Manuscript

Dear Barkley,  
we have revised citations and included references into the main text as required. We have also made an extensive language revision of the text, which we hope has solved the language problems of the first version.

Thank you very much for your help and suggestions.

Sincerely,  
Mauro

Accepted Manuscript

# Random Digraphs with Given Expected Degree Sequences: A Model for Economic Networks

January 19, 2011

## Abstract

Building upon the growing interest for complex network theory, the main ambition of this paper is to contribute to a more effective application of network theory to economic phenomena, and particularly to financial networks. We depart from recent contributions on credit networks in two respects. In the first place, we constrain diversification with the out- and in-degree distribution of nodes, by adopting a suitable extension of the expected degree model for random weighted digraphs. In the second place, we focus on real networks by using this extension as null model for statistical analysis. With the help of statistical inference, we provide a rigorous answer to the following question: do real networks obey our null model? Further, we show that this answer is tightly connected to the existence of clusters or modules, as defined by Newman and Girvan (2004), within networks.

## 1 Introduction

For many years an increasing interdependence has been the intended and welcome outcome of the liberalization of real and financial economic flows across the globe. With the start of the global crisis, this trend has come to a sudden reversal, as the unintended, and unwelcome, consequences of interdependence have finally come to the fore for both private investors and policy makers. Today, the need to deal with the unpleasant side of interdependency is highlighted by the new importance attached to the concept of *systemic risk*. By the latter expression, policy makers, and especially central banks, have begun to designate a new way of looking at economic risk, which aims to take into consideration the potentially negative impact, on the wealth of an individual institution, of the behavior of other institutions, directly or indirectly connected with it.

Systemic risk is implicitly defined by the transmission channels we introduce to explain events which cannot be accounted for by looking only at individual institutions. Many such transmission channels have been listed by the literature (see e.g Brunnermeier, 2009), which for our purposes may be categorized in two broad classes. The first class includes all sorts of market-related contagion effects, which are determined by the indirect interaction of agents through asset markets. The most prominent example in this class is the asset price contagion effect, i. e. the negative feedback loop running from asset price declines to forced asset sales and back, via balance sheet constraints. The second class

1  
2  
3  
4  
5  
6  
7  
8  
9 consists of credit-related contagion effects, which are determined by the recip-  
10 rocal claims between institutions on the credit markets. Obviously, the two  
11 classes are themselves tightly connected, since most investors on asset markets  
12 rely on credit to fund their positions, a circumstance which explains why credit  
13 deleveraging may hurt asset prices so badly.

14 Recently, complex network theory has become a prominent tool in this field,  
15 since it provides a rigorous representation of complex multi-part interactions,  
16 as well as the possibility to predict their static and dynamic behavior in many  
17 cases<sup>1</sup>. Regarding the case in point, it's natural to conceive credit markets  
18 as credit networks in which nodes represent agents and links represent credit  
19 claims. This idea, which has been circulating for some years (Greenwald and  
20 Stiglitz, 2003), has received new attention after the outbreak of the global crisis,  
21 which started with the so called "credit crunch", an almost complete breakdown  
22 of most segments of the global credit markets. Additionally, network theory is  
23 suited to represent market interactions too. For example, any asset market may  
24 be conceived of as a bipartite network, i. e. one with two separate categories  
25 of nodes, the first one being the investors, the second one being the assets  
26 exchanged on the market, and the links being weighted by the holdings of each  
27 asset by each investor. In such a framework, it's quite easy to build a projected  
28 network for investors, i. e. one in which investors exposed to the same asset  
29 are connected, but it may be more useful, for some purposes, to work with the  
30 original network. To deal with such bipartite representation, as well as with  
31 any multipartite interaction which is typical of the economic domain (just think  
32 about the loan market, where firms and banks act on the opposite sides of the  
33 market), very flexible tools are needed. The main ambition of this paper is  
34 to provide a set of such tools, thus laying the ground for a more general and  
35 effective application of network theory to the economic domain.

36  
37 Most works on systemic risk focus on credit markets, and specifically on the  
38 interbank market, which is particularly relevant for financial stability and, at  
39 the same time, well suited for a representation with basic network theory. The  
40 early contribution of Allen and Gale (2000) stressed the benefits of increasing  
41 diversification, suggesting that the more connections, the better for financial  
42 stability, even though, to be honest, the authors didn't take into account the  
43 kind of global shock experienced with the crisis. Since the latter triggered a full  
44 rethink of this matter, a growing number of works have begun to challenge this  
45 view, showing that diversification is not always beneficial for stability.

46  
47 Building upon the work of Eisenberg and Noe (2001), Shin (2008) has pro-  
48 vided a model in which balance sheet effects acting across a number of in-  
49 terconnected banks may produce credit booms and busts. The recent model  
50 of Battiston et al. (2009) shows, using the tools of complexity theory, that,  
51 if market-related effects are considered along with credit-related effects by in-  
52 troducing a financial accelerator mechanism, then a potential trade-off between  
53 individual risk and systemic risk may exist for increasing connectivity of the net-  
54 work. Similar results are provided by Gai and Kapadia (2010), who show, with  
55 the help of complex networks theory, that financial systems exhibit a robust-yet-  
56 fragile tendency: while the probability of contagion may be low, once a default  
57 cascade is started its spread may be quite large. This effect is non monotonic in  
58 connectivity: for a given range of values, connectivity increases the chances that

59  
60  
61  
62  
63  
64  
65  

---

<sup>1</sup>For a general introduction see Vega-Redondo (2007).

1  
2  
3  
4  
5  
6  
7  
8  
9 institutions surviving the effects of the initial default will be exposed to more  
10 than one defaulting counterpart after the first round of contagion, thus making  
11 them more vulnerable to a second-round default. Finally, it's noteworthy that  
12 central banks are getting more and more interested in this topic, supporting  
13 network-related research and dissemination, although most applied work in this  
14 direction still looks merely descriptive (Castrén and Kavonius, 2009; ECB, 2010)  
15 <sup>2</sup>.

16 While pursuing the same general goal of a better understanding of complex  
17 interactions within the economic domain, our approach is different in two re-  
18 spects from the existing works on credit networks. In the first place, we adopt  
19 a different view on diversification. While most models use either very simpli-  
20 fied network settings (e.g. Allen et al., 2010), involving few nodes, or regular  
21 topologies, and in both cases agents are assumed to have equally weighted re-  
22 lative exposure, we suppose that agents' diversification is constrained by the  
23 respective weight of their participation to the network. Technically, this means  
24 assuming networks with arbitrary weighted degree distributions. This approach,  
25 which will be explained fully in the next section, brings our model much closer  
26 to empirics, since regular topologies are hardly observed in real world economic  
27 networks, while also carrying some theoretical advantages. To explain the lat-  
28 ter point, we may think of a network representing a market; then our approach  
29 constrains diversification to a given expected supply and demand on part of the  
30 agents. This is convenient, since it allows to sharply separate the problem of  
31 market interaction from the problem of supply and demand determination.

32 In the second place, we turn to empirics by adopting our "constrained di-  
33 versification" model as null model for a statistical analysis of real networks. We  
34 choose this path since we believe that a stronger theory and a stronger empirical  
35 knowledge of real networks go hand in hand, and both are needed to improve our  
36 understanding of this complex domain. In particular, in this paper we wish to  
37 answer the following general question: do real networks obey to our null model?  
38 As it turns out, we can provide a rigorous answer to this question with the help  
39 of statistical inference. Further, we show that this answer is tightly linked to  
40 the existence of clusters or modules in the networks, to be defined, according to  
41 Newman and Girvan (2004), as subset of nodes whose connections exceed the  
42 expectation under an appropriately chosen null model. Thus, a perfect (con-  
43 strained) diversification is proven to be equivalent to the absence of modules in a  
44 network, a circumstance which clears the way for a rigorous meso-level analysis  
45 of economic networks, and of markets in particular.

46 The rest of the paper is organized as follows. The first step is to introduce  
47 our null model (section 2) along with some basic concepts and facts needed  
48 to derive the results presented later (section 3). Then, with the help of two  
49 bounds provided in section 4, we derive a set of statistical tools for the analysis  
50 of economic networks (section 5), which are subsequently tested on a credit  
51 network in section 6. We must remark that these tools can be applied readily to  
52 any interaction occurring among economic agents, provided that this interaction  
53 can be represented by a nonnegative matrix of arbitrary shape. Finally, the  
54 conclusive section summarizes the main results of the paper and suggests future  
55 research paths.

56  
57  
58  
59 <sup>2</sup>A notable exception is provided by Alessandri et al. (2009) who incorporate default cas-  
60 cades in a more general econometric model for estimating the assets distribution of UK banks.

## 2 The Expected Degree Model

Our expected degree (ED) model is a generalization of the one proposed by F. Chung and co-authors in several papers (Chung, 2002; Chung and Lu, 2002; Chung et al., 2003a,b). The latter, on its part, represents an extension of the classical random graph model to accommodate for any arbitrary nodes' degree sequence  $\bar{d} = (\bar{d}_1, \bar{d}_2, \dots, \bar{d}_n)$ , contrary to the basic Erdos-Rényi model whose degrees follow invariably the binomial distribution. In particular, under Chung's formulation, each pair  $(i, j)$  of vertices in  $G(\bar{d})$  is joined by an edge with probability  $\bar{p}_{ij} = \frac{\bar{d}_i \bar{d}_j}{\bar{v}}$  where  $\bar{v} = v(\bar{d}) = \sum_{i=1}^n \bar{d}_i$  is the volume of  $G$ . A key assumption of the model is that  $\max_i \bar{d}_i^2 < \bar{v}$  so that  $\bar{p}_{ij} \leq 1$  for all  $i$  and  $j$ , and the sequence is graphical, in the sense that the necessary and sufficient conditions for a sequence to be realized by a graph are satisfied. In particular, if  $a_{ij}$  is an element of the adjacency matrix  $A$  of  $G$ , then  $a_{ij}$  becomes a random variable and we have  $a_{ij} \in \{1, 0\}$  and thus  $E[a_{ij}] = \bar{p}_{ij}$ . It is straightforward to verify that  $E[d_i] = E[\sum_{j=1}^n a_{ij}] = \sum_{j=1}^n E[a_{ij}] = \sum_{j=1}^n \bar{p}_{ij} = \bar{d}_i$ . Unfortunately, in its original formulation this model is not very interesting for economic applications due to two major shortcomings:

1. economic interactions are typically not symmetric (think of market operations where buyers and sellers find themselves in opposite roles, or to credit operations, where the same occurs to borrowers and lenders) - thus a graph model, which depicts only symmetric relations, doesn't fit very well into this domain
2. the content of economic interactions is usually best described by monetary values or some equivalent, which are naturally conceived as continuous magnitudes, as opposed to the discrete binary representation typical of graph theory

To begin with, we want to show that it is possible to recast the model in such a way as to overcome these limitations. In doing so we take advantage of the extensions of the ED model to weighted and directed networks, proposed by authors studying the problem of community detection in complex networks, like Newman and Girvan (2004) or Arenas et al. (2007). In the first place, given the discussion above, it looks natural to conceive the out- and in-degree values as continuous variables. In the second place, we must take into account the eventual lack of symmetry by letting  $G(\bar{d}_{out}, \bar{d}_{in})$  depend on two distinct sequences  $\bar{d}_{out} = (\bar{d}_{out_1}, \bar{d}_{out_2}, \dots, \bar{d}_{out_n})$  and  $\bar{d}_{in} = (\bar{d}_{in_1}, \bar{d}_{in_2}, \dots, \bar{d}_{in_m})$ . We depart from the extensions mentioned above by allowing possibly  $n \neq m$ , provided that  $v(\bar{d}_{out}) = \sum_{i=1}^n \bar{d}_{out_i} = \bar{v} = \sum_{j=1}^m \bar{d}_{in_j} = v(\bar{d}_{in})$  still holds. The representation obtained in this way is indeed very general, since we only require that  $n, m$  are less or equal to the number of vertices or economic agents involved in the system or market under consideration, which means that any agent could possibly (but not necessarily) operate on both sides of any interaction, and eventually even interact with herself if this makes sense in the specific domain. Later on we will introduce the condition that  $\bar{d}_{out_i}, \bar{d}_{in_j} > 0$  for all  $i, j$  with  $1 \leq i \leq n$  and  $1 \leq j \leq m$ , since we wish to work with a properly degree normalized adjacency matrix (see next section). Although none of the arguments presented in this section depends on this assumption, we must remark that requiring strictly positive in- and out-degree provides more parsimony of representation while not

imposing any significant restriction, since, according to our former assumption, we don't necessarily need to represent agents who are not expected to take part to one of the two poles of the interaction.

If we define  $\Omega_{ij} = \{\omega_1, \omega_2\}$  where  $\omega_1 = \{\varepsilon \in (ij)\}$  and  $\omega_2 = \{\varepsilon \notin (ij)\}$ , for  $\varepsilon > 0$  and  $(ij)$  representing a pair of economic agents, then  $(\Omega_{ij}, \mathcal{F}, P_{ij})$  is a discrete probabilistic model. In fact, if we set  $P_{ij}(\omega_1) = \bar{p}_{ij} = \frac{\bar{d}_{out_i} \bar{d}_{in_j}}{\bar{v}^2}$  and  $P_{ij}(\omega_2) = 1 - P_{ij}(\omega_1)$ , it is clear that  $P_{ij}(\omega_1), P_{ij}(\omega_2) \geq 0$ ,  $P(\Omega_{ij}) = 1$  and  $P_{ij}(\omega_1) + P_{ij}(\omega_2) = 1$ , which means that  $P_{ij}$  satisfies all the axioms of a probability measure for  $(\Omega_{ij}, \mathcal{F})$ . In graph theoretic terms, a succession  $\mathcal{W}_{ij} = \{\omega_n\}$  of events from  $(\Omega_{ij}, \mathcal{F}, P_{ij})$  may describe the  $(i, j)$  nonnegative entry of a random adjacency matrix  $A \in \mathbb{R}^{n \times m}$  representing a weighted directed graph, possibly with self-loops, whenever the weights are of the form  $w = \varepsilon N$  for some integer  $N$ . In fact, from  $\mathcal{W}_{ij}$  we can build a succession of random values  $\{\varepsilon_i\}$  with  $\varepsilon_i = \{0, \varepsilon\}$  such that  $a_{ij} = \sum_{i=0}^{K-1} \varepsilon_i$  for each entry of  $A$  and some integer  $K > 0$ . We must prove that the expected degree distribution derived from  $A$  corresponds to the desired one. The key hypothesis to obtain this result is that  $|\mathcal{W}_{ij}| = M$  for all  $i$  and  $j$ , where we have that  $\bar{v} = \varepsilon M$ . In fact,  $E[a_{ij}] = \sum_{k=0}^{M-1} E[\varepsilon_k] = \sum_{k=0}^{M-1} \bar{p}_{ij} \varepsilon = \frac{\bar{d}_{out_i} \bar{d}_{in_j}}{\bar{v}} = \bar{a}_{ij}$ . Then  $d_{out_i}$  and  $d_{in_j}$  are themselves sums of independent random variables, whose expectation is  $E[d_{out_i}] = \sum_{j=1}^m E[a_{ij}] = \bar{d}_{out_i}$  and  $E[d_{in_j}] = \sum_{i=1}^n E[a_{ij}] = \bar{d}_{in_j}$ , as required by the main assumption of the model.

We can justify the choice made for  $P_{ij}(\omega_1)$  by showing that it is possible to derive the expression for  $\bar{p}_{ij}$  as the solution of an entropy maximization problem over  $P_{ij} \equiv P_{ij}(\omega_1)$ . This solution is commonly used as a statistical tool to estimate the entries of a matrix from its row and column sums, and has found wide application in the economic domain, e. g. to estimate the bilateral exposure of financial entities or goods' flows across sectors in the input-output tables (Castrén and Kavonius, 2009). In particular, following the formulation of Wells (2004), we can state the problem as follows:

$$\min_P g(P) = \sum_{i=1}^n \sum_{j=1}^m P_{ij} \ln P_{ij} \quad (2.1)$$

subject to the following constraints:

$$\begin{aligned} \sum_{j=1}^m P_{ij} &= \bar{r}_i \\ \sum_{i=1}^n P_{ij} &= \bar{c}_j \\ P_{ij} &\geq 0 \end{aligned}$$

where  $\bar{r}_i = \bar{d}_{out_i} / \bar{v} < 1$ ,  $\bar{c}_j = \bar{d}_{in_j} / \bar{v} < 1$  and of course  $\sum_{i=1}^n \bar{r}_i = \sum_{j=1}^m \bar{c}_j = 1$ . The first-order conditions give the solution of the problem:

$$P_{ij} = \exp(\lambda_i + \mu_j - 1) \quad (2.2)$$

where  $\lambda_i$  and  $\mu_j$  are Lagrange multipliers respectively of the first and second constraint. By substituting the solution into the constraints and the normalization condition we obtain

$$\begin{aligned} \sum_{j=1}^m \exp(\lambda_i + \mu_j - 1) &= \exp(\lambda_i - 1/2) \sum_{j=1}^m \exp(\mu_j - 1/2) = \bar{r}_i \\ \sum_{i=1}^n \exp(\lambda_i + \mu_j - 1) &= \exp(\mu_j - 1/2) \sum_{i=1}^n \exp(\lambda_i - 1/2) = \bar{c}_j \\ \sum_{i=1}^n \sum_{j=1}^m \exp(\lambda_i + \mu_j - 1) &= \sum_{j=1}^m \exp(\mu_j - 1/2) \sum_{i=1}^n \exp(\lambda_i - 1/2) = 1 \end{aligned}$$

Combining the latter expression with the former two we get

$$\begin{aligned} \exp(\lambda_i - 1/2) &= \sum_{i=1}^n \exp(\lambda_i - 1/2) \bar{r}_i \\ \exp(\mu_j - 1/2) &= \sum_{j=1}^m \exp(\mu_j - 1/2) \bar{c}_j \end{aligned}$$

If we now substitute the latter expressions into eq. (2.2), we get the desired result

$$P_{ij} = \bar{r}_i \bar{c}_j \sum_{i=1}^n \exp(\lambda_i - 1/2) \sum_{j=1}^m \exp(\mu_j - 1/2) = \bar{r}_i \bar{c}_j = \frac{\bar{d}_{out_i} \bar{d}_{in_j}}{\bar{v}^2} \quad (2.3)$$

Given the properties of entropy, this result means that the probability expression we have chosen for  $P_{ij}(\omega_1)$  is not an arbitrary one. In fact, it corresponds to a set of probability values obtained using as information input only the relative in- and out-degree of nodes in the network. To understand why this solution works, we should ask ourselves which relationship holds between the constraints of problem (2.1) and  $G(\bar{d}_{out}, \bar{d}_{in})$ . In order to do so, the first step is to write down  $P(a_{ij} = w)$  explicitly. If we set  $w = \varepsilon K \geq 0$  and  $x = \varepsilon N > 0$ , then we may write

$$\begin{aligned} P(a_{ij} = w)(x) &= P(K) \\ &= \binom{N}{K} P_{ij}^K (1 - P_{ij})^{N-K} \end{aligned} \quad (2.4)$$

Since each  $a_{ij}$  follows an independent binomial distribution, the likelihood of observing a given  $w$  is maximized by the mode of the Beta distribution with parameters  $(K + 1, N - K + 1)$ . Next we observe that, if we wish to equalize the expected volume of  $G(\bar{d}_{out}, \bar{d}_{in})$  and the actual volume of  $G$ , the following must hold:

$$\bar{v} = E[v] = \sum_{ij} E[a_{ij}] = \sum_{ij} P_{ij}x \quad (2.5)$$

A natural candidate solution of equation (2.5) is  $x = \bar{v}$ , which implies  $\sum_{ij} P_{ij} = 1$ . Let's suppose that we start with an unconstrained entropy maximizing assignment  $P_{ij} = p$  for all  $i, j$ . Next suppose we wish to maximize  $P(d_i = \bar{d})(\bar{v})$  with respect to  $p$ , now considered as a variable, for a given observed degree value  $\bar{d} = \varepsilon K$ . The solution is again provided by the mode of the Beta distribution with parameters  $(K + 1, N - K + 1)$  since  $d_i$  is a sum of binomially distributed variables with equal probability parameter  $p$ :

$$\frac{\alpha - 1}{\alpha + \beta - 2} = \frac{K}{N} = \frac{\bar{d}}{\bar{v}} \quad (2.6)$$

Clearly the set of maximum likelihood solutions for each  $d_i$  provides us with the constraints used for problem (2.1). Then it's clear that  $x = \bar{v}$  is the only value which makes the latter constraints, derived for each degree sequence, consistent with condition (2.5), a conclusion which justifies our previous seemingly arbitrary assumption.

The problem (2.1) is also tightly related to a different class of models (Garlaschelli and Loffredo, 2008, 2009), which apply the maximum likelihood method to estimate, under a set of local constraints, the probability of observing a given  $G$  within the ensemble of graphs which may be represented by a  $n \times m$  adjacency matrix. Since each  $G$  is equivalent to a  $nm$  long ordered sequence of weights, each possible nonnegative sequence  $w = (w_0 \dots w_{nm-1})$  with  $w_i \leq x$  and  $w_i = \varepsilon K_i$  for some integer  $K_i$ , corresponds to a Graph in the ensemble and *vice versa*. The probability of observing a graph  $G$  in the ensemble is

$$P(G) = \prod_{i,j} P_{ij}(w) \quad (2.7)$$

where we leave  $P_{ij}(w)$  undefined for the moment. By maximizing  $-P(G) \ln P(G)$  it's possible to show that  $P(G)$  takes the following form (Garlaschelli and Loffredo, 2009):

$$P(G) = \frac{e^{-H(G)}}{Z} \quad (2.8)$$

where  $H(G)$  is an Hamiltonian for the vector of constraints and  $Z$  is a partition function which serves as a normalizing factor. In the case of  $G(\bar{d}_{out}, \bar{d}_{in})$  the Hamiltonian reads  $H(G) = \sum_{ij} (\theta_i^{out} + \theta_j^{in}) a_{ij}$ . Comparing equations (2.7) and (2.8) we derive

$$P_{ij}(w) = \frac{e^{-(\theta_i + \theta_j)w}}{\sum_{k=0}^N e^{-(\theta_i + \theta_j)k\varepsilon}} \quad (2.9)$$

To see the equivalence between (2.9) and (2.4) we may take logs of both expressions:

$$\begin{aligned}
\log P(a_{ij} = w)(x) &= \log \binom{N}{K} + K \log(P_{ij}) + (N - K) \log(1 - P_{ij}) \\
&= [\log(P_{ij}) - \log(1 - P_{ij})] K + \log \binom{N}{K} + N \log(1 - P_{ij}) \\
\log P_{ij}(w) &= -(\theta_i + \theta_j)w - \log \left( \sum_{k=0}^N e^{-(\theta_i + \theta_j)k\varepsilon} \right)
\end{aligned}$$

A comparison of the two expressions leads us to the following identification:

$$\log(P_{ij}) - \log(1 - P_{ij}) \equiv -(\theta_i + \theta_j)\varepsilon$$

Rearranging we obtain

$$\begin{aligned}
P_{ij} &= \frac{e^{-(\theta_i + \theta_j)\varepsilon}}{1 + e^{-(\theta_i + \theta_j)\varepsilon}} \\
1 - P_{ij} &= \frac{1}{1 + e^{-(\theta_i + \theta_j)\varepsilon}} \\
&\text{and} \\
(1 - P_{ij})^{N-K} &= (1 + e^{-(\theta_i + \theta_j)\varepsilon})^{K-N}
\end{aligned}$$

Thus

$$\begin{aligned}
P_{ij}(w) &= \frac{e^{-(\theta_i + \theta_j)w}}{\sum_{k=0}^N e^{-(\theta_i + \theta_j)k\varepsilon}} \left( \frac{1 + e^{-(\theta_i + \theta_j)\varepsilon}}{1 + e^{-(\theta_i + \theta_j)\varepsilon}} \right)^K \left( \frac{1 + e^{-(\theta_i + \theta_j)\varepsilon}}{1 + e^{-(\theta_i + \theta_j)\varepsilon}} \right)^N \\
&= \binom{N}{K} P_{ij}^K (1 - P_{ij})^{N-K}
\end{aligned}$$

where  $\frac{(1 + e^{-(\theta_i + \theta_j)\varepsilon})^N}{\sum_{k=0}^N e^{-(\theta_i + \theta_j)k\varepsilon}} \equiv \binom{N}{K}$ . This shows the equivalence of (2.9) and (2.4), which implies that the solution of problem (2.1) delivers also the maximum likelihood solution for  $P(G)$ . In addition, our formulation explains why we should set  $x = \bar{v}$  to obtain the desired result.

To underline the theoretical convenience of the choice made for  $P$ , we can remark that this expression has a very natural interpretation in economic terms. In fact, if we think for example in terms of market exchanges, the expression says that, for a given market price, the probability that a small quantity of good is sold from a particular seller to a particular buyer is determined by the joint probability that the former comes to the market to sell (which is given by  $\frac{\bar{d}_{out_i}}{\bar{v}}$ ) and the latter comes to the market to buy (which is given by  $\frac{\bar{d}_{in_j}}{\bar{v}}$ ). This expression describes the expected behavior on the market, once the expected individual supply  $\bar{d}_{out_i}$  and demand  $\bar{d}_{in_j}$  of each agent are known, and the

following two conditions hold: i) each market transaction for  $dv$  is random and independent; ii) no buyer/seller displays a particular preference or dislike for any seller/buyer.

To conclude, one final remark is in place: if self-interactions are unsuitable for the domain under analysis, the model must incorporate the additional constraint  $p_{ij} = 0$  whenever  $i$  and  $j$  refer to the same entity. This can be accomplished with the help of a specific procedure (see e.g. Censor and Zenios, 1997). Given the fact that the credit networks we are going to consider below are bipartite, i. e. networks in which nodes at the two poles of interaction never overlap, we are not going to tackle this problem further in this paper.

### 3 Basic concepts and facts

In this section we are going to present shortly some basic concepts and facts which are necessary to state our main results. For a general treatment the reader can refer to Chung (1997, 2006). The first notion we need to introduce is the normalized adjacency matrix of  $G(\bar{d}_{out}, \bar{d}_{in})$ :

$$K = D_{out}^{-\frac{1}{2}} A D_{in}^{-\frac{1}{2}}$$

where  $D_{out}^{-\frac{1}{2}}, D_{in}^{-\frac{1}{2}}$  are diagonal matrices with entries respectively equal to the reciprocal of the square root of out- and in-degree values of  $A$ .  $K$ , like any other real or complex matrix, has a unique singular value (SV) decomposition  $K = U \Sigma V^T$ , where  $\Sigma = \text{diag}(\sigma_0, \dots, \sigma_{p-1}) \in \mathbb{R}^{p \times p}$  with  $\sigma_0 \geq \sigma_1 \geq \dots \geq \sigma_{p-1} \geq 0$  and  $p = \min\{n, m\}$ . Further, since  $K$  has real entries,  $U \in \mathbb{R}^{n \times p}$  and  $V \in \mathbb{R}^{m \times p}$ , which means that the singular vectors  $u_i, v_i$  of  $K$  have real entries too.

The normalization of  $K$  is realized with respect to the matrix 2-norm, i. e.  $\|K\|_2 \equiv 1$ . To confirm this claim, we need to show that  $\sigma_0 \equiv 1$ , where  $\sigma_0$  stands for the largest SV of  $K$ , since by definition  $\|K\|_2 \equiv \sigma_0$ . The task can be divided into two steps: firstly we show that 1 is always a SV of  $K$ , secondly we show that such value is the largest one. The first step consists of finding two real vectors  $u$  and  $v$  such that  $Kv \equiv u$ . The conclusion readily follows if we choose  $u^T = (\sqrt{d_{out_1}/v}, \dots, \sqrt{d_{out_n}/v})$  and  $v^T = (\sqrt{d_{in_1}/v}, \dots, \sqrt{d_{in_m}/v})$ , because we get the following identity

$$\sqrt{\frac{d_{out_i}}{v}} = u_i = \sum_{j=1}^m k_{ij} v_j = \sum_{j=1}^m \frac{a_{ij}}{\sqrt{d_{out_i}} \sqrt{d_{in_j}}} \sqrt{\frac{d_{in_j}}{v}} = \sqrt{\frac{d_{out_i}}{v}}$$

Since we can compute such vectors for any normalized adjacency matrix, 1 is always a SV of  $K$ .

We can show that 1 is effectively the largest SV of  $K$  through the Perron-Frobenius theorem (PF). In fact, we know that  $\sigma_i(K)^2 = \lambda_i(KK^T) = \lambda_i(K^T K)$  for all  $\lambda_i, \sigma_i > 0$  (Hogben, 2007, pp.5-11), where  $\lambda_i$  stands for a real eigenvalue of the symmetric matrices  $KK^T$  and  $K^T K$ , and the  $\lambda_i$ 's are taken in decreasing order. In addition, the singular vectors  $u_i, v_i$  associated with  $\sigma_i$  are equal to the eigenvectors of  $KK^T$  and  $K^T K$  respectively associated with  $\lambda_i$ . Since 1 is

a SV of  $K$  whose associated singular vectors are positive, then 1 is an eigenvalue of both  $KK^T$  and  $K^TK$  with positive associated eigenvector. From PF we know that the spectral radius  $\rho$  of an irreducible matrix is a simple and positive eigenvalue, and that it is the only eigenvalue of that matrix with a positive eigenvector. But a nonnegative symmetric square matrix  $X \in \mathbb{R}^{n \times n}$  being irreducible is equivalent to the following condition: the graph  $G(X)$ , with vertices  $V(X) = \langle n \rangle$  and edges  $E(X) = \{(ij) : i, j \in \langle n \rangle \wedge x_{ij} = x_{ji} \neq 0\}$ , is connected. Thus, whenever this condition is satisfied for  $G(KK^T)$  and  $G(K^TK)$ , the conclusion follows.

We must underline that there is no loss of generality with this assumption, since we know that the spectrum of a Graph with  $k$  non trivial connected components is the union of the spectra of the disjoint connected Graphs corresponding to each of its components (Chung, 1997, par 1.2). Thus the conclusion  $\sigma_0 \equiv 1$  follows, together with  $\sigma_0 = \dots \sigma_{k-1} = 1$  where  $k$  is the number of non trivial connected components in  $G(KK^T)$  and  $G(K^TK)$ . Further, it's easy to see that  $G(KK^T)$  and  $G(K^TK)$  are connected if and only if  $G(\bar{d}_{out}, \bar{d}_{in})$  is (weakly) connected. In fact, the latter holds only if  $G(K)$ , with vertices  $V(X) = \langle k = \max\{n, m\} \rangle$  and edges  $E(X) = \{(ij) : i, j \in \langle k \rangle \wedge (x_{ij} \neq 0 \vee x_{ji} \neq 0)\}$ , is connected too. Then, if we consider in particular the necessity condition, it's clear that  $G(K)$  being disconnected means that we have a permutation for  $K$  such that

$$K = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

where only  $A$  is required to be square. Thus  $KK^T$  will be of the following form

$$KK^T = \begin{pmatrix} A^2 & 0 \\ 0 & BB^T \end{pmatrix}$$

which implies that  $G(KK^T)$  is disconnected. The same conclusion obviously holds also for  $K^TK$  since  $K^TK = K^T(K^T)^T = K^TK^T$ .

It is possible to say more about the SVs of  $K$  with the help of  $G(KK^T)$  and  $G(K^TK)$ . In particular we can state the following theorem:

**Theorem 3.1.** *If  $K$  is the normalized adjacency matrix of  $G(\bar{d}_{out}, \bar{d}_{in})$ , then  $G(KK^T)$  and  $G(K^TK)$  are not bipartite.*

*Proof.* Suppose  $G(KK^T)$  is bipartite. Then there is a permutation of  $KK^T$  such that

$$KK^T = \begin{pmatrix} O & B \\ B^T & O \end{pmatrix}$$

where  $B \neq 0$  and  $O$  is an all-zero matrix. But this is impossible, since  $K$  has positive row and column sums, which means that the diagonal elements of  $KK^T$  are all positive

$$(KK^T)_{ii} = \sum_{j=1}^m k_{ij}^2 > 0$$

The same conclusion is immediately extended to  $K^T K$ . □

From PF we know that  $|\lambda_i(X)| < \rho(X)$  for all  $\lambda_i(X) \neq \rho(X)$  if the nonnegative matrix  $X$  is aperiodic. For our purposes, this condition can be translated into the following statement: if  $G(KK^T)$  and  $G(K^T K)$  are non bipartite, then  $|\lambda_i(KK^T)| = |\lambda_i(K^T K)| < 1$  for all  $\lambda_i(KK^T) = \lambda_i(K^T K) \neq 1$ . Then, using theorem 3.1, we derive two straightforward conclusions:

1. if  $G(\bar{d}_{out}, \bar{d}_{in})$  is (weakly) connected, then  $\sigma_0(K) > \sigma_1(K)$
2. if  $G(\bar{d}_{out}, \bar{d}_{in})$  has  $k$  (weakly) connected components,  $1 \equiv \sigma_0(K) = \dots = \sigma_{k-1}(K) > \sigma_k(K)$

To summarize, the multi-set of the SVs of  $K$  is equal to the union of the squared spectra of the connected components of  $KK^T$  and  $K^T K$ . This statement allows us to assume throughout this paper, without loss of generality, that  $G(\bar{d}_{out}, \bar{d}_{in})$  is (weakly) connected.

In the next paragraph we will analyze networks' properties by means of the SVs of their normalized adjacency matrix  $K$ . This choice must be justified, since it is more common to find in the literature a characterization of the spectrum of  $G$  by means of the so-called Laplacian. In particular Chung (1997) has introduced into algebraic graph theory the important notion of the normalized Laplacian of a Graph  $\mathcal{L}(G)$ , to be defined (ibidem, par 1.2) as follows

$$\mathcal{L}(G) \equiv_{def} I - D_{-1/2} A D_{-1/2} = I - K \quad (3.1)$$

where  $A$  is the adjacency matrix of  $G$  and  $D$  is a diagonal matrix with non zero entries equal to the degrees of the nodes. From the definition above, it follows immediately that  $\lambda_i(\mathcal{L}) = 1 - \lambda_i(K)$  where  $\lambda_i(\mathcal{L})$  and  $\lambda_i(K)$  are taken in increasing and decreasing order respectively. The emphasis on  $\mathcal{L}(G)$  is motivated by the connection between its spectral properties and Markov processes defined over graphs and digraphs, a field of application which falls beyond the limits of this paper. In addition,  $\mathcal{L}(G)$  cannot be defined whenever  $K(G)$  is rectangular, while at the same time  $K(G)$  retains the most important properties of  $\mathcal{L}(G)$ , for our purposes, when this happens. Then, since  $K$  turns out to be the most flexible option in our case, in this paper we will refer to the SVs of  $K$  as the SVs of  $G(\bar{d}_{out}, \bar{d}_{in})$ .

## 4 Upper and lower bounds for the subdominant singular value of $G(\bar{d}_{out}, \bar{d}_{in})$

In this section we are going to derive an upper bound for the so-called subdominant singular value of  $G(\bar{d}_{out}, \bar{d}_{in})$ , namely  $\sigma_1(K)$ , which is based on the Frobenius norm of a peculiar matrix  $K^*$ , which will be defined below. Since we are assuming  $G(\bar{d}_{out}, \bar{d}_{in})$  to be weakly connected, we know in advance that  $0 \leq \sigma_1 < 1$ .

The interest for non unit eigenvalues and SVs of normalized matrices derives from the theory of Markov processes. In fact, a Markov chain is ergodic if the

1  
2  
3  
4  
5  
6  
7  
8  
9 powers of a stochastic matrix  $S$  converge to a rank one matrix. If  $\lambda_1 < 1$ ,  
10 where  $\lambda_1$  is the second largest eigenvalue of  $S$ , then  $|\lambda_1(S^k)| = |\lambda_1(S)|^k \rightarrow 0$  as  
11  $k \rightarrow \infty$ , i. e. the powers of  $S$  converge to a stochastic matrix of rank one. The  
12 magnitude of  $\lambda_1$  estimates the asymptotic rate of convergence, which means  
13 that  $\lambda_1$  can be used as a coefficient of ergodicity (Rothblum and Tan, 1985). If  
14 alternatively the matrices under multiplication are different, we need a different  
15 coefficient, defined over a specific vector or matrix norm, like for example  
16

$$17 \quad \tau_1(S) = \max_{\|z\|_1, z^T \mathbf{1} = 0} \|S^T z\|_1$$

18  
19 The expression above provides the one-norm coefficient of ergodicity for the  
20 stochastic matrix  $S$ , which bounds from above the non unit eigenvalues of  $S$ , i.  
21 e.  $\lambda_1(S) \leq \tau_1(S)$ , and is submultiplicative. Given the relationship between  $\sigma_1$   
22 and  $\lambda_1$ , ergodic coefficients could also provide, under appropriate conditions, an  
23 upper bound for  $\sigma_1(S)$ , a circumstance which explains our interest for this no-  
24 tion. To start with, we can remark that  $\tau_1(S)$  is simply the norm of  $S$  restricted  
25 to the subspace that is orthogonal to  $\mathbf{1}$ , which is a dominant eigenvector of  $S$ .  
26 This fact has suggested the opportunity to extend ergodic coefficients to non-  
27 negative real and complex matrices of arbitrary shape (Selee and Ipsen, 2010).  
28  
29

30 Before we get into the matter more deeply, it is worth explaining the rele-  
31 vance of  $\sigma_1(K)$  for our purposes. In order to do so, it is convenient to intro-  
32 duce the notion of *discrepancy* (Butler, 2006) between the typical Digraph in  
33 the ensemble  $\mathcal{G}(\bar{d}_{out}, \bar{d}_{in})$  of Digraphs generated using  $\bar{p}_{ij}$  as defined by (2.3),  
34 and a given Digraph  $G$  supposed to belong to  $\mathcal{G}$ . For each  $X \subset V_{out}(G)$  and  
35  $Y \subset V_{in}(G)$  where  $V_{out}$  and  $V_{in}$  are respectively the sets of nodes with positive  
36 out- and in-degree, we can always compute a value  $\beta$  such that  
37

$$38 \quad \left| e(X \rightarrow Y) - \frac{vol_{out}(X)vol_{in}(Y)}{\bar{v}} \right| = \beta \sqrt{vol_{out}(X)vol_{in}(Y)} \quad (4.1)$$

39 Here  $e(X \rightarrow Y)$  stands for the total weight of edges going from nodes in  $X$   
40 to nodes in  $Y$ , and  $vol_{out}(X), vol_{in}(Y)$  are the volumes of  $X$  and  $Y$  respectively.  
41 The connection of (4.1) with  $\mathcal{G}(\bar{d}_{out}, \bar{d}_{in})$  becomes clear if we look at the second  
42 term on the l.h.s. of the expression, which in fact gives the expected weight of  
43 edges between  $X$  and  $Y$  according to  $p_{X,Y} \equiv \sum_{i \in X} \sum_{j \in Y} p_{ij}$ . The relationship  
44 between  $\sigma_1$  and  $\beta$  for all  $X$  and  $Y$  is given by the following  
45  
46  
47

48 **Theorem 4.1.** *For all  $G(\bar{d}_{out}, \bar{d}_{in})$ ,  $X \subset V_{out}(G)$  and  $Y \subset V_{in}(G)$ , we have*  
49 *that  $disc(G) \leq \sigma_1(K)$ , where  $K$  is the normalized adjacency matrix of  $G$  and*  
50  *$disc(G) \equiv_{def} \max_{(X,Y)} \beta(X,Y)$ , with  $\beta(X,Y)$  taken to be the value such that*  
51 *4.1 holds for given  $X$  and  $Y$ .*  
52

53 *Proof.* (Butler, 2006, see) The proof depends on the inequality  $|\langle a, Mb \rangle| \leq$   
54  $\sigma_0(M) \|a\|_2 \|b\|_2$ , where  $\langle \cdot, \cdot \rangle$  denotes the usual scalar product, which holds for  
55 any matrix  $M$  and vectors  $a$  and  $b$ . Further, we know that  $A = \sum_{i=0}^r \sigma_i u_i v_i^T$  for  
56 all  $A \in \mathbb{R}^{n \times m}$ , where the  $\sigma_i$  are the non zero SVs of  $A$  with left and right singular  
57 vectors  $u_i$  and  $v_i$ , and  $r = \min \{n - 1, m - 1\}$ . Then we have as a consequence  
58 that, if  $u$  and  $v$  are the dominant singular vectors of  $K$ ,  $\sigma_0(K - uv^T) = \sigma_1(K)$ .  
59 Finally,  $\frac{vol_{out}(X)vol_{in}(Y)}{\bar{v}} = \langle \psi_X, D_{out}^{\frac{1}{2}} uv^T D_{in}^{\frac{1}{2}} \psi_Y \rangle$ , where  $\psi_X$  and  $\psi_Y$  are the  
60  
61  
62  
63  
64  
65

characteristic vectors of  $X$  and  $Y$ , and  $D_{out}, D_{in}$  have been already defined. Thus we get the following

$$\begin{aligned} \left| e(X \rightarrow Y) - \frac{vol_{out}(X)vol_{in}(Y)}{\bar{v}} \right| &= \left| \left\langle \psi_X, (A - D_{out}^{\frac{1}{2}} w v^T D_{in}^{\frac{1}{2}}) \psi_Y \right\rangle \right| \\ &= \left| \left\langle D_{out}^{\frac{1}{2}} \psi_X, (K - w v^T) D_{in}^{\frac{1}{2}} \psi_Y \right\rangle \right| \\ &\leq \sigma_0(K - w v^T) \left\| D_{out}^{\frac{1}{2}} \psi_X \right\|_2 \left\| D_{in}^{\frac{1}{2}} \psi_Y \right\|_2 \\ &= \sigma_1(K) \sqrt{vol_{out}(X)vol_{in}(Y)} \end{aligned}$$

Since the inequality holds for all  $\beta(X, Y)$ , then it holds also for  $disc(G)$ .  $\square$

Theorem 4.1 provides a powerful insight into the spectral properties of  $G$ , suggesting that if  $G \in \mathcal{G}(\bar{d}_{out}, \bar{d}_{in})$  w. h. p., which implies that discrepancy is small, then the subdominant SV of  $K(G)$  will be smaller than in the opposite case, and possibly close to zero. Here two observations are in order. In the first place, for any digraph it is clearly interesting to test whether it belongs to  $\mathcal{G}$  or not, since in the former case we have a fully diversified network, as clarified above (section 2), while in the latter case the network displays communities as defined by Newman (2006), a circumstance which may have probably relevant consequences for the dynamics of networks in case of sudden negative shocks<sup>3</sup>. On the other hand, unfortunately, the theorem doesn't provide any clue as to determine how small  $\sigma_1$  must be in order to decide whether, w.h.p.,  $G \in \mathcal{G}$ . In order to do so, we need a statistical test to provide us with a reliable estimate of the probability of finding a Digraph  $G$  in  $\mathcal{G}$ , given the magnitude of  $\sigma_1(G)$ .

While finding such a test will be the task of the next paragraph, we need to state some preliminary results here which will be useful for this purpose. First of all, we must provide a definition of the ergodic coefficient for rectangular matrices  $A \in \mathbb{R}^{n \times m}$  and arbitrary vectors  $w \in \mathbb{R}^n$  (Seneta, 1984):

$$\tau_p(w, A) = \max_{\substack{\|z\|_p=1 \\ z^T w=0}} \|A^T z\|_p$$

Then it is possible to demonstrate the following (Rothblum and Tan, 1985, theorem 5.5)

**Theorem 4.2.** *If  $A \in \mathbb{R}^{n \times m}$  and  $w \in \mathbb{R}^n$  with  $w \neq 0$ , then for all vectors  $x \in \mathbb{R}^m$*

$$\tau_p(w, A) \leq \|(A - w x^T)^T\|_p$$

*In particular if  $p = 2$*

$$\tau_2(w, A) \leq \|(A - w x^T)\|_2 \leq \left\| \left( I - \frac{w w^T}{\|w\|_2^2} \right) A \right\|_F$$

*Proof.* see Selee and Ipsen (2010).  $\square$

<sup>3</sup>This topic will be treated more extensively in the conclusive section of the paper.

The rightmost expression provides the required upper bound for  $\sigma_1$ . In fact, by combining theorem 7.15 and corollary 7.16 in Selee and Ipsen (2010), we can state the following

**Theorem 4.3.** *Let  $A \in \mathbb{R}^{n \times m}$  and  $w \in \mathbb{R}^n$  with  $w \neq 0$ . Then*

$$\tau_2(w, A) = \left\| \left( I - \frac{ww^T}{\|w\|_2^2} \right) A \right\|_2$$

*In particular, if  $u$  and  $v$  are the dominant left and right singular vectors of  $A$ , then*

$$\tau_2(u, A) = \tau_2(u, A^T) = \sigma_1(A)$$

*Proof.* For the general proof we refer to Selee and Ipsen (2010). When considering  $K$ , or any other normalized matrix, the derivation of the second statement from the first becomes straightforward:

$$\tau_2(u, K) = \|(I - uu^T)K\|_2 = \|(K - uv^T)\|_2 = \sigma_1(K) \quad (4.2)$$

where we have used  $u^TK = v^T$  and the fact that the matrix 2-norm is equal to the largest SV of a matrix.  $\square$

We can remark that theorem 4.3 provides an alternative proof for theorem 4.2, since we know that  $\|\cdot\|_2 \leq \|\cdot\|_F$  for all matrices (Hogben, 2007, 37.3). Using this fact and theorem 4.3 we get the desired upper bound for  $\sigma_1(A)$

$$\sigma_1(A) \leq \|(I - uu^T)A\|_F$$

In particular, for  $K$  like for any other normalized matrix

$$\sigma_1(K) \leq \|K - uv^T\|_F \quad (4.3)$$

Finally, by posing  $K^* \equiv K - uv^T$ , we define the matrix mentioned at the beginning of the section. As it turns out, the same matrix appeared in the proof of Theorem 4.1, a fact which shows the intimate connection between the latter and Theorem 4.3. This connection gives us the possibility to combine the two in order to derive a lower bound for  $\sigma_1(K)$  in terms of the volume of  $K^*$ :

$$\frac{|vol(K^*)|}{\sqrt{nm}} = \frac{|\langle \mathbf{1}, K^* \mathbf{1} \rangle|}{\sqrt{nm}} = \beta(\mathbf{1}, \mathbf{1}) \leq \sigma_1(K) \quad (4.4)$$

Here  $\beta$  is defined directly with respect to the characteristic vectors of the subsets  $X$  and  $Y$ . Although the bound provided by (4.4) will turn out to be theoretically useful in the next section, it is of little use in applications. In fact, it's clear that if  $K$  and  $E[K]$  have the same degree sequences, as usually required from applications, they will have the same volume too, which means  $|vol(K^*)| = 0$ . On the other hand, even if the latter is true, as long as  $K \neq E[K]$  there are characteristic vectors  $\psi_X \neq \mathbf{1}$  and  $\psi_Y \neq \mathbf{1}$  such that  $\beta(\psi_X, \psi_Y) > 0$ , which provide tighter lower bounds for  $\sigma_1(K)$ . To extract two such characteristic vectors it's natural to resort to spectral partitioning methods (Newman, 2006). This circumstance highlights the connection between our results and the problem of module detection in networks. Although a full clarification of this

connection is left for future work, we will provide some preliminary considerations on this matter in the conclusive section.

## 5 Statistical tests for $\mathcal{G}$

As it is well known, bounds based on entry-wise matrix norms are not very tight, since these norms are obviously very sensitive to the size of matrices. On the other hand, alternative norms involve invariably some variational characterization (Selee and Ipsen, 2010), which becomes difficult to handle if one wishes, like in this case, to work with expectations. In this section we will show that, under specific assumptions, entry-wise norms for  $\sigma_1(K)$  provide a useful bound to test if Digraphs based on real data belong w.h.p. to  $\mathcal{G}$ .

First of all, we need to introduce a result which is essential for our argumentation. To explain its relevance in this context, we should note that, according to the model presented in section 2, if  $G \in \mathcal{G}(\bar{d}_{out}, \bar{d}_{in})$ , then each entry  $a_{ij}$  of  $A(G)$  becomes a sum of  $N$  independent random variables  $\varepsilon_k$  (remember  $\bar{v} = \varepsilon N$ ).

**Theorem 5.1.** *Let  $X_1, \dots, X_n$  be independent random variables satisfying  $X_i \leq \frac{\lambda}{|M|}$  for all  $1 \leq i \leq n$  and  $M \neq 0$ . Let  $X = \sum_{i=1}^n X_i$  and  $\|X\| = \sqrt{X} = \sqrt{\sum_{i=1}^n E(X_i^2)}$ . Then*

$$P(X \leq E[X] - \lambda) = P(X \geq E[X] + \lambda) \leq e^{-\frac{\lambda^2}{2(\|X\|^2 + |M|\lambda/3)}} \quad (5.1)$$

If the  $X_i$  are nonnegative, the lower bound can be strengthened to

$$P(X \leq E[X] - \lambda) \leq e^{-\frac{\lambda^2}{2\|X\|^2}} \quad (5.2)$$

*Proof.* See Chung and Lu (2006).  $\square$

If  $G \in \mathcal{G}(\bar{d}_{out}, \bar{d}_{in})$ , we may take  $X = \sum_{k=1}^N \varepsilon_k = a_{ij}$  and  $E[X] = \bar{a}_{ij} \geq \|X\|^2$  if we choose  $\varepsilon < 1$ . Then the following bounds hold:

$$P(a_{ij} \geq \bar{a}_{ij} + \lambda) \leq e^{-\frac{\lambda^2}{2(\bar{a}_{ij} + \varepsilon\lambda/3)}} \quad (5.3)$$

$$P(a_{ij} \leq \bar{a}_{ij} - \lambda) \leq e^{-\frac{\lambda^2}{2\bar{a}_{ij}}} \quad (5.4)$$

These inequalities, when applied to single entries, are of little use as a test since, generally speaking, it is likely that, as  $A$  gets dimensionally larger, there will be some  $a_{ij}$  close to its expectation even if  $G \notin \mathcal{G}(\bar{d}_{out}, \bar{d}_{in})$ , and likewise some  $a_{ij}$  lying far away from its expectation even if  $G \in \mathcal{G}(\bar{d}_{out}, \bar{d}_{in})$ . This observation provides us with the fundamental justification for a test based on a global property of  $G$ , like the magnitude of  $\sigma_1(G)$ .

Before we tackle this problem, we must make a fundamental remark. The validity of inequalities (5.3) and (5.4) is immediately extended to any sum of entries of  $A$ . In particular, they hold for each  $d_{out_i}$ ,  $d_{in_j}$  and for  $v$ . Now, since these values are set for any  $G \in \mathcal{G}(\bar{d}_{out}, \bar{d}_{in})$  to be equal to their expectation, an obvious consequence of (5.3) and (5.4) is the following:

$$P(d_{out_i} = \bar{d}_{out_i}) \leq 1 \quad P(d_{in_j} = \bar{d}_{in_j}) \leq 1 \quad \forall i, j \quad P(v = \bar{v}) \leq 1 \quad (5.5)$$

Up to now we have made the implicit supposition that we know exactly the out- and in-degree distribution of any  $G$  we use to build  $\mathcal{G}$ . In fact, the latter ensemble is built using  $(\bar{d}_{out}, \bar{d}_{in})$  as a deterministic input. But then we can strengthen conditions (5.5) to equality, something which, on its part, implies that  $G \in \mathcal{G}(\bar{d}_{out}, \bar{d}_{in})$  if and only if  $G \in \mathcal{G}^* \subseteq \mathcal{G}(\bar{d}_{out}, \bar{d}_{in})$  where  $\mathcal{G}^*$  is populated by the Digraphs in the ensemble which have exactly the required out- and in-degree distribution. Using this fact,  $E[K]$  is easily determined. In fact, we get

$$E[K] = \bar{D}_{out}^{-\frac{1}{2}} \bar{A} \bar{D}_{in}^{-\frac{1}{2}} \equiv \bar{K} \quad (5.6)$$

which follows since  $d_{out_i} \equiv \bar{d}_{out_i}$  and  $d_{in_j} \equiv \bar{d}_{in_j}$  in  $\mathcal{G}^*$ . We are going to show that the restriction to  $\mathcal{G}^*$  allows some very stark conclusions regarding  $G \in \mathcal{G}^*$ . As a first step, the following result provides an application of theorem 5.1 to  $K$ :

**Theorem 5.2.** *Suppose  $G \in \mathcal{G}^*$ , with  $K$  as its normalized adjacency matrix, and  $\bar{K}$  being as defined above. Then, for any sum  $X = \sum k_{ij}$  of elements of  $K$  we have the following*

$$P(X \geq E[X] + \lambda) \leq e^{-\frac{\lambda^2}{2(\sum s_{min_{ij}} + \lambda/3)}} \quad (5.7)$$

and, since the  $k_{ij}$  are nonnegative

$$P(X \leq E[X] - \lambda) \leq e^{-\frac{\lambda^2}{2\sum s_{min_{ij}}}} \quad (5.8)$$

where  $s_{min_{ij}} = \min \{s_{row_{ij}}, s_{col_{ij}}\}$ , and the arguments of  $s_{min_{ij}}$  are defined respectively as  $s_{row_{ij}} = a_{ij}/d_{out_i}$  and  $s_{col_{ij}} = a_{ij}/d_{in_j}$ .

*Proof.* The statement follows from theorem 5.1, by observing that, from the application of Holder inequality, the following holds

$$\begin{aligned} E[k_{ij}^2] &= E \left[ \frac{a_{ij}}{d_{out_i}} \frac{a_{ij}}{d_{in_j}} \right] \leq \min \left\{ E \left[ \frac{a_{ij}}{d_{out_i}} \right], E \left[ \frac{a_{ij}}{d_{in_j}} \right] \right\} \\ &= \min \left\{ \frac{\bar{a}_{ij}}{d_{out_i}}, \frac{\bar{a}_{ij}}{d_{in_j}} \right\} \end{aligned}$$

where the latter equality follows from the main hypothesis.  $\square$

The applicability of theorem 5.2, as well as of expressions (5.3) and (5.4), relies on the knowledge we have of  $\bar{K}$ . Unfortunately, the following results are going to show that we cannot take this knowledge for granted. This is because the way  $\mathcal{G}(\bar{d}_{out}, \bar{d}_{in})$  is built is not fully consistent.

**Lemma 5.1.** *Suppose  $G \in \mathcal{G}^*$ , with  $K$  as its normalized adjacency matrix,  $K^*$  and  $\bar{K}$  defined as above. Then  $E[K^*] = 0$ .*

*Proof.* The proof is straightforward:

$$\begin{aligned}
 E[K^*] &= E[K - uv^T] \\
 &= E[K] - uv^T \\
 &= \bar{K} - uv^T \\
 &= 0
 \end{aligned}$$

where  $E[uv^T] = uv^T$  since  $u$  and  $v$  are constant in  $\mathcal{G}^*$  and

$$\bar{k}_{ij} = \frac{\bar{a}_{ij}}{\sqrt{\bar{d}_{out_i} \bar{d}_{in_j}}} = \frac{\sqrt{\bar{d}_{out_i} \bar{d}_{in_j}}}{\bar{v}} = u_i v_j$$

□

In order to extend lemma 5.1 to  $|K^*|$ , we define two distinct matrices  $K_+^*$  and  $K_-^*$  whose elements are again random variables  $k_{+ij}^* = I(\omega_+)k_{ij}^*$  and  $k_{-ij}^* = -I(\omega_-)k_{ij}^*$ , with  $\omega_+ = \{k_{ij}^* \geq 0\}$  and  $\omega_- = \{k_{ij}^* < 0\}$ . Then it is clear that  $|K^*| \equiv K_+^* + K_-^*$ . Further, with the help of Holder inequality we get

$$0 \leq E[k_{+ij}^*] \leq E[k_{ij}^*] = 0 = -E[k_{ij}^*] \geq E[k_{-ij}^*] \geq 0$$

i. e.  $E[k_{+ij}^*] = E[k_{-ij}^*] = 0$  which, taken together, imply  $E[|K^*|] = 0$ . With the help of the latter result we are ready to prove the following

**Theorem 5.3.** *Suppose the hypotheses in lemma 5.1 are true. Then  $K = \bar{K}$ .*

*Proof.* The statement follows by observing that, as a consequence of the extension of lemma 5.1 to  $|K^*|$ , we have that  $\sigma^2(K^*) = 0$ . This means that, for all  $G \in \mathcal{G}^*$ , it must hold that  $K^* = E[K^*] = 0$ , which on its turn implies  $K = \bar{K}$ . □

The paradoxical consequence of theorem 5.3 is that  $G \notin \mathcal{G}$  holds whenever  $G \neq \bar{G}$ , where the latter is defined to be the Graph corresponding to  $\bar{K}$ . This is a serious inconvenient, since  $G$  belongs to the ensemble only if a very restrictive condition holds. The paradox is solved if we drop the hypothesis of an exact knowledge of  $K$ . In general, the opposite supposition seems more convenient since, if  $G$  is the outcome of a random process, then it is a single instance of this process, and its degree sequences cannot be warranted to correspond to the "true" ones. On the other hand, an uncertainty over the "true" degree sequences implies that, on the ground of this instance alone, we have no hints as to which ensemble  $G$  belongs. Let's take for such a "true" value the expectation  $E[K] = \langle K \rangle$  over an unspecified probability measure and denote with  $\langle G \rangle$  the corresponding Graph. Then theorem 5.3 turns out to be extremely useful since it tells us that  $K^* = 0$  must hold for the ensemble  $\mathcal{G}$  obtained from  $\langle K \rangle$ . Using this fact we state the next theorem

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

**Theorem 5.4.** *Suppose  $\langle G \rangle \in \mathcal{G}^*$ . Then  $E[K^*] = 0$  for each  $G$  with a n.a.m.  $K$  such that  $E[K] = \langle K \rangle$ .*

*Proof.* In the first place we remark that, since  $E[K] = \langle K \rangle$ , then  $E[uv^T] = \langle u \rangle \langle v \rangle^T$ , where the two terms of the product stand for the dominant singular vectors of  $\langle K \rangle$ . This is because  $E[uv^T]$  is functionally determined by the row and column sums of  $E[K]$ . Thanks to the main hypothesis, we can use theorem 5.3 to derive  $\langle K \rangle = \bar{K}$ . From lemma 5.1, then, we know that  $\langle K \rangle - \langle u \rangle \langle v \rangle^T = 0$ . The conclusion follows from  $E[K^*] = \langle K \rangle - \langle u \rangle \langle v \rangle^T$ .  $\square$

**Theorem 5.5.** *Suppose the hypotheses in theorem 5.4 are true. Then, for any sum  $X = \sum |k_{ij}^*|$  of elements of  $|K^*|$  we have the following*

$$P(X \geq \lambda) \leq e^{-\frac{3}{2}\lambda} \quad (5.9)$$

*Proof.* From theorem 5.4 we derive  $E[|K^*|] = 0$  which implies  $E[X] = 0$ , while  $X$  is necessarily nonnegative. Then the proof follows from (5.1) by observing that  $\|X\| \leq E[X] = 0$ .  $\square$

Theorem 5.5 provides an upper bound for the probability of observing the volume of  $|K^*|$ . This probability on its part gives an indirect estimation of the difference between  $K$  and  $\langle K \rangle$ , since the volume of  $K^*$  is expected to be equal to zero for  $\langle G \rangle \in \mathcal{G}$ . In the next section this estimation will prove useful for the analysis of the evolution of networks over time.

Under the hypotheses of theorem 5.4, we have that  $E[\beta(\mathbf{1}, \mathbf{1})] = 0$  and  $E[\|K^*\|_F] \leq E[\sqrt{\text{vol}(K^*)}] \leq \sqrt{E[\text{vol}(K^*)]} = 0$  which, taken together and remembering (4.3) and (4.4), imply  $E[\sigma_1(K)] = 0$ . This fact reduces the problem of verifying  $G \in \mathcal{G}$  w.h.p., for some unspecified degree sequence, to that of verifying whether w.h.p.  $\sigma_1(G) = 0$ . This task is drastically simplified by the following

**Theorem 5.6.** *Suppose the hypotheses of theorem 5.4 are true. Then for any  $\epsilon \in \mathbb{R}^+$ , we have that  $P(\sigma_1(G) \geq \lambda) \gg 0$  only if  $\lambda$  can be chosen to be infinitesimal of the same order of  $\epsilon$ .*

*Proof.* Using successively (4.2), the equivalence of matrix norms, Markov inequality and Jensen inequality we get

$$\begin{aligned} P(\sigma_1(G) \geq \lambda) &= P(\|K^*\|_2 \geq \lambda) \\ &\leq P(\|K^*\|_F \geq \lambda) \\ &\leq P(\sqrt{\text{vol}(|K^*|)} + \epsilon \geq \lambda) \\ &\leq \frac{E[\sqrt{\text{vol}(|K^*|)} + \epsilon]}{\lambda} \\ &\leq \frac{\sqrt{E[\text{vol}(|K^*|)]} + \epsilon}{\lambda} \\ &\leq \frac{\epsilon}{\lambda} \end{aligned}$$

$\square$

In particular, with the help of theorem 5.6 we can state the following

**Corollary 5.1.** *Suppose  $\lambda > 0$  is a numerically determined value for  $\sigma_1(G)$ , while  $\epsilon > 0$  is a numerically determined value for  $\sigma_1(\bar{G})$ , where  $\bar{G}$  is now computed from  $G$  itself. Then we have that  $H_0 : \sigma_1(G) = 0$  can be rejected with confidence level  $\alpha \leq 1 - \frac{\epsilon}{\lambda}$ .*

*Proof.* The statement follows from theorem 5.6 by observing that  $1 - \frac{\epsilon}{\lambda} \leq 1 - P(\sigma_1(G) \geq \lambda)$ .  $\square$

## 6 Application to credit networks

In this section we are going to apply the tools developed in section 5 to a dataset of credit relationship data between Japanese firms and banks, which has been analyzed under different perspectives in previous works (e.g De Masi et al., 2009; Fujiwara and Aoyama, 2008; Fujiwara, 2009). More in detail, the dataset refers to firms listed in the Japanese stock-exchange markets. Data are compiled from firms' financial statements, integrated by a survey of *Nikkei Media Marketing, Inc.* in Tokyo. They include the indication of the amount of borrowing obtained from each financial institution, subdivided in short-term debt and long-term debt. Financial institutions, which we refer to as "banks" in the remaining of this paper, consist of long-term credit banks, city banks, regional banks (primary and secondary), trust banks and insurance companies, all of which basically represent the universe of financial institutions in Japan. Although the dataset covers a long span of years, ranging from 1980 to 2005, we are going to focus on the most recent years, namely the period 2000-2005. For an extensive description of this network the reader can refer to De Masi et al. (2009).

Table 1: Descriptive statistics of the Dataset

Year	Firms	Banks	Links
2000	2,629	211	27,389
2001	2,714	204	26,597
2002	2,739	202	24,555
2003	2,700	192	22,585
2004	2,700	190	21,919
2005	2,674	182	21,811

In Graph theoretic terms, bank-firm relationships are represented by a weighted bipartite Graph  $G$ . In fact, the weights of  $G$  are represented by the amounts of credit/debit and its two bipartite sets are represented by firms  $F$  and banks  $B$  respectively. Thanks to our previous theoretical treatment, we can work directly with  $A$  taken as the bi-adjacency matrix of  $G$ , which has shape  $|F| \times |B|$  and entries equal to the amounts of credit/debit between each bank and firm in  $G$ . Further, since we have two different kinds of credit/debit relationship, we may denote the graphs where only short term and long term credit/debit amounts

1  
2  
3  
4  
5  
6  
7  
8  
9 are included with  $G_s$  and  $G_l$  respectively. Then we have that  $G = G_s \oplus G_l$ . It's  
10 clear that each of the three networks will have distinct SVs, which means our  
11 analysis will refer to 18 matrices in total. It is worth remarking that the matrices  
12  $AA^T$  and  $KK^T$  on the one hand, and  $A^T A$  and  $K^T K$  on the other, have in  
13 this case a natural interpretation as the *projected networks* between firms and  
14 banks respectively (De Masi et al., 2009).  
15

16 The goal of our analysis is to answer the following two questions: how likely  
17 is it that the considered networks belong to  $\mathcal{G}$ ? And, in that case, how far is  
18 the network from its expectation? Both questions are indeed relevant for eco-  
19 nomics. By answering the first one, we verify how likely it is that real data are  
20 generated by a random process of the kind described in section 2. On its part,  
21 this recognition leads to opposite but, in both cases, relevant conclusions: on  
22 the one hand, it provides a validation of the theoretical model in the domain under  
23 consideration; on the other, as already hinted at, it shows that the network  
24 in case is less diversified than what would be optimal according to the model.  
25 Once diversification is unequivocally recognized as beneficial for stability, the  
26 latter conclusion clears the way for a policy intervention seeking to prevent the  
27 build-up of systemic risk within the network. By answering the second question,  
28 on the other hand, we look at the evolution of the network over time, since the  
29 difference between real data and their expected value accounts for the level of  
30 diversification of the network in a given period. As it turns out, diversification  
31 is intimately connected with the role, within the network, of modules or  
32 communities as defined by Newman (2006), whose existence, we will argue, is  
33 intimately connected with the SVs of  $K$ .  
34  
35

36 As anticipated in section 5, the examination of single cell values cannot answer  
37 the global question of whether, and with which probability, our networks  
38 belong to  $\mathcal{G}$ . Then, in order to answer our first question, we must resort to  
39 corollary 5.1 and take into consideration the sub-dominant SV of  $K$ .<sup>4</sup> In the  
40 first place, we can observe that the difference between the SVs of  $K$  and those  
41 of its expectation  $\bar{K}$  is striking, as depicted by figures 1, 2 and 3, in each year  
42 and for each kind of credit. In particular, many SVs of  $K$  are close to unity,  
43 a circumstance which finds an intuitive explanation if we recall the discussion  
44 made in section 3. In fact, whenever  $K$  displays a unitary SV, we know that in  
45  $K$  there is an additional (weak) component. Thus, by extension, if a SV is close  
46 to unity, this means that  $K$  has an additional quasi-component, i.e. a region  
47 which is weakly connected to the rest of the network. The peculiar feature of  
48  $\langle K \rangle \in \mathcal{G}$  is then evident, since the fact that its sub-dominant SV is zero tells us  
49 that in this matrix there are no quasi-components, or communities, like we may  
50 prefer to call them. Finally, this shows, as claimed above, that the difference  
51 between  $K$  and  $\langle K \rangle$ , if it exists, lies exactly in the existence of communities in  
52 the former, while, on the other hand, this existence always implies, thanks to  
53 the theoretical properties of  $\langle K \rangle$ , a lack of diversification within  $K$ . If such a  
54  
55

56 <sup>4</sup>A preliminary analysis showed that some networks in the dataset were not connected, due  
57 to the existence of a second very small component, including very few firms and banks. This  
58 is true in particular for the following years: 2000 (long term data and total data), 2004 (short  
59 term data), 2005 (short term data). In accordance with the hypothesis made at the beginning  
60 of section 5, we have removed these components from the networks. As expected, then, table  
61 2 shows that the subdominant SVs are always smaller than unity.  
62  
63

1  
2  
3  
4  
5  
6  
7  
8  
9 difference exists in the first place, then a subsequent deeper analysis of the net-  
10 work is required to identify the communities within it, a point we will examine  
11 again in the conclusive section.  
12

13  
14 Table 2: Values of  $\sigma_1(K)$  and p-values of  $\sigma_1 \geq \lambda$  if  $G \in \mathcal{G}$ .  
15

Year	Short Term	Long Term	Total
2000	0.8203 (1.35e-08)	0.9828 (1.29e-08)	0.9753 (1.07e-08)
2001	0.9634 (1.20e-08)	0.7855 (1.52e-08)	0.7671 (1.40e-08)
2002	0.8822 (1.58e-08)	0.6835 (1.66e-08)	0.6610 (1.72e-08)
2003	0.8432 (1.63e-08)	0.7075 (1.87e-08)	0.6535 (1.93e-08)
2004	0.8564 (1.51e-08)	0.7126 (1.91e-08)	0.6627 (2.00e-08)
2005	0.7826 (1.77e-08)	0.9678 (1.25e-08)	0.7694 (1.58e-08)

16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33 Table 2 finally provides us with the answer to our first question. In fact, the  
34 probabilities of observing the values reported in the table are very small, which  
35 implies that the hypothesis  $G \in \mathcal{G}$  can be rejected at a 99 % confidence level  
36 in all cases. On the other hand, the interpretation advanced for these values  
37 suggests that their variations over time are to be interpreted as a measure of  
38 how close the two main communities within  $K$  are to be disconnected. In other  
39 terms, the values in Table 2, while providing a global test for  $K$ , convey by them-  
40 selves a local, and not a global, information on the networks. To find a proper  
41 measure for the overall distance between  $K$  and  $\bar{K}$  we must resort to theorem  
42 5.5 and apply it to  $D = vol(|K^*|)$ . In table 3 we reproduce the resulting values  
43 for each of our networks, together with the probabilities computed according  
44 to (5.9). Since  $D$  is of course correlated with the dimensions of  $K^*$ , which are  
45 progressively shrinking according to table 1, we add to the table  $D^*$ , an aver-  
46 age value per cell, in order to make the results comparable across matrices of  
47 different size. According to our results, then, there is evidence of an increasing  
48 deviation of  $K$  from  $\langle K \rangle$ , which means that communities have become progres-  
49 sively stronger in the period under analysis or, alternatively, that the network  
50 has become less diversified. The question of whether such differences are indeed  
51 statistically significant is left for future work.  
52  
53

## 54 7 Conclusions

55  
56  
57 In these conclusive remarks we wish to outline some of the general implications  
58 of the results obtained in the previous sections. In our opinion, the most impor-  
59 tant theoretical achievement of our work is the identification of a model which,  
60 while being easily applicable for the interpretation of real data, has also some  
61  
62  
63  
64  
65

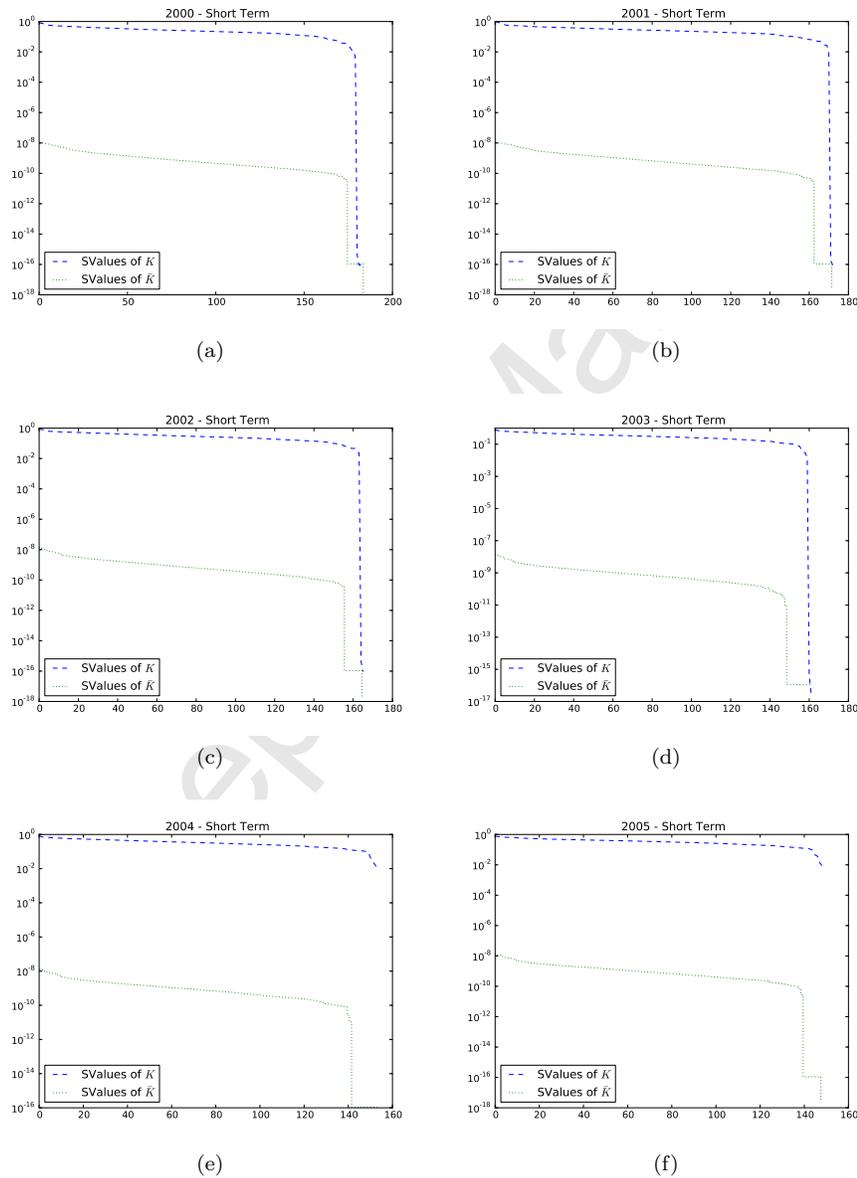
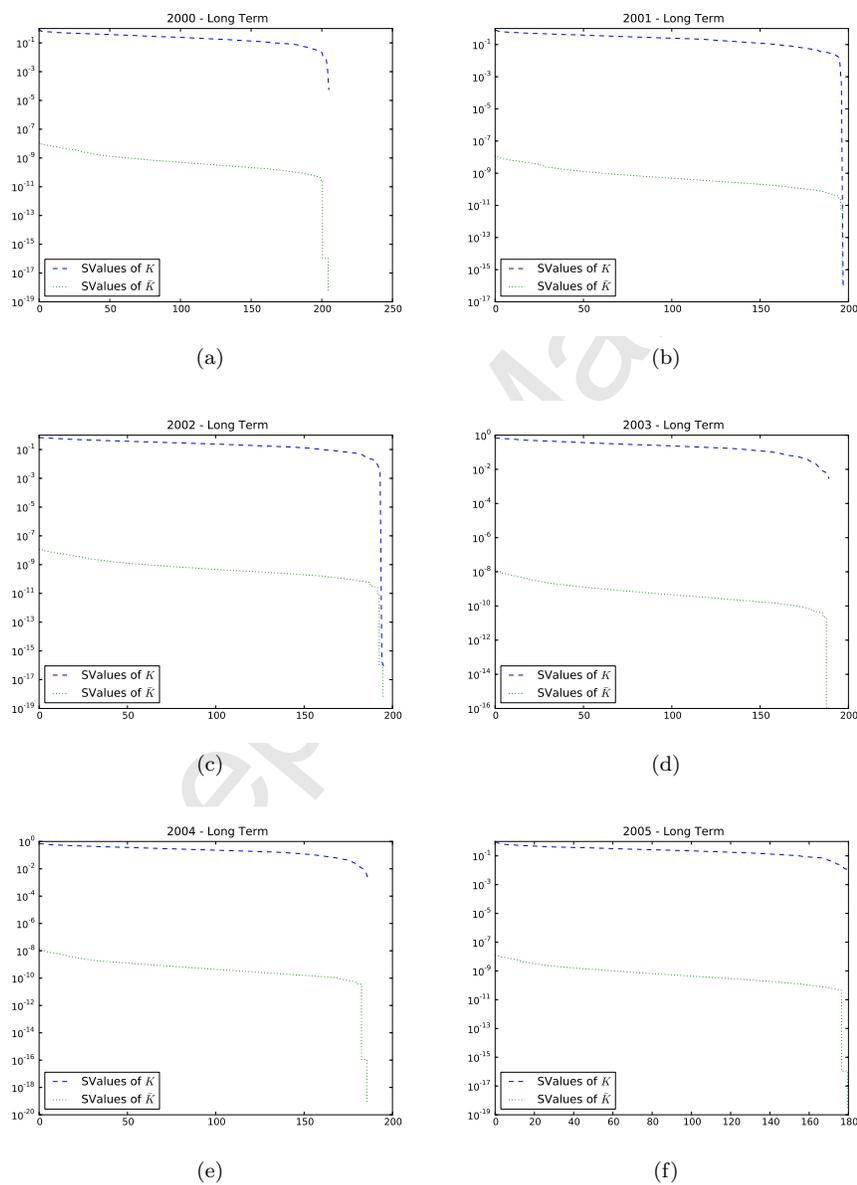


Figure 1: Singular Values of  $K_{st}$  and  $\bar{K}_{st}$ .

Figure 2: Singular Values of  $K_{lt}$  and  $\bar{K}_{lt}$

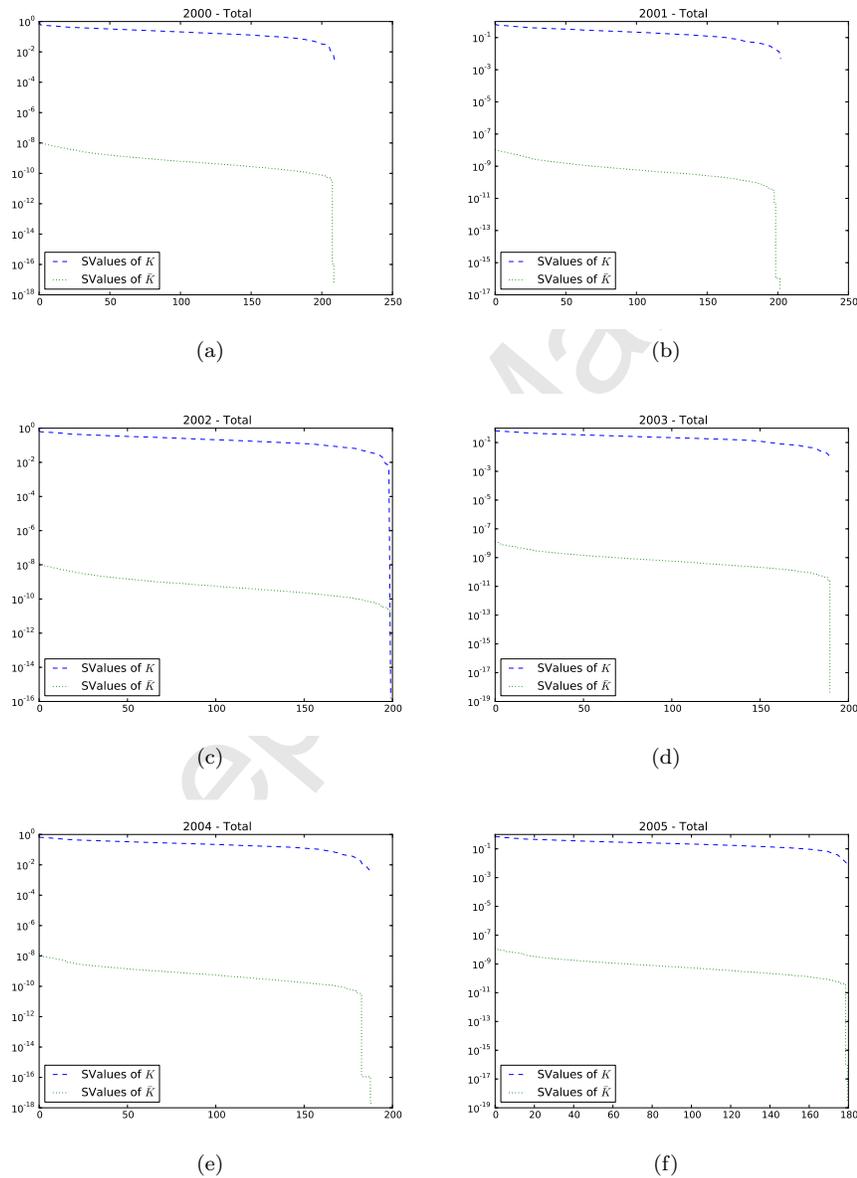
Figure 3: Singular Values of  $K$  and  $\bar{K}$ .

Table 3: Values of  $D = vol(|K^*|)$  and of  $D^* = \frac{vol(|K^*|)}{nm}10^3$ . P-values of  $vol(|K^*|) \geq \lambda$  if  $G \in \mathcal{G}$  in parentheses.

Year	$D$			$D^*$		
	Short Term	Long Term	Total	Short Term	Long Term	Total
2000	421.77 (1.75e-275)	403.76 (9.46e-264)	473.38 (4.18e-309)	0.9528	0.8989	0.8534
2001	419.89 (2.93e-274)	403.03 (2.81e-263)	470.17 (5.17e-307)	0.9959	0.8909	0.8492
2002	405.10 (1.27e-264)	403.96 (6.96e-264)	462.24 (7.50e-302)	1.0011	0.8892	0.8399
2003	407.91 (1.86e-266)	394.93 (5.36e-258)	454.14 (1.43e-296)	1.0672	0.9061	0.8760
2004	405.99 (3.33e-265)	386.59 (1.44e-252)	444.46 (2.88e-290)	1.1761	0.8948	0.8664
2005	413.95 (2.18e-270)	381.55 (2.78e-249)	442.83 (3.31e-289)	1.2647	0.9227	0.9099

very attractive theoretical properties, which depend on the probabilistic model we have adopted. The latter, in fact, is derived from simple, and fairly standard, assumptions regarding the behavior of agents, which have been already highlighted in section 2. The application of the model to real data then conveys a clear cut indication regarding the data generating process of real networks. In particular, we have argued that, whenever real networks don't behave consistently with the model, two specular options are open: either modify the model, e. g. by introducing additional hypotheses, or accept the model for its theoretical benefits, and then act, with the help of policy instruments, to modify real networks in order to bring them closer to the model itself.

In the final part of the paper we have applied the tests devised in section 5 to real credit networks data, showing that such networks don't behave consistently with the ED model of section 2. Consequently, real credit networks are characterized by the existence of modules as defined by Newman (2006), since the latter are expected not to exist within the ensemble  $\mathcal{G}$ . This circumstance may have either a theoretical or a practical bearing, depending on the stance of the reader. From the theoretical point of view, it's clear that we need an alternative model for modularized networks, encompassing factors, such as asymmetric information, institutional specificities or geographical proximity, which could explain the existence of modules. Instead, in order to further develop the practical implications of our work, the first step we need to take is to rigorously identify modules within networks, since these correspond to the regions which may most likely be the epicenter of negative events in case of a unexpected build-up of systemic risk.

Given the latter consideration, our work turns out to be intimately connected with the growing stream of literature tackling the problem of community or module detection in networks (for a recent review see Fortunato, 2010), and particularly with spectral partitioning methods (e.g. Newman, 2006; Donetti and Muñoz, 2005). In a nutshell, the connection with our analysis lies in the

1  
2  
3  
4  
5  
6  
7  
8  
9 parallel, proposed in section 6, between communities and components. Spectral  
10 partitioning methods show that members of the same community have similar  
11 values in the eigenvectors (excluding the first one) associated with the  $k$  largest  
12 positive eigenvalues of  $A$  or, alternatively, with the  $k$  smallest eigenvalues of the  
13 Laplacian (a recent analysis of the spectral properties of modular networks is  
14 provided by Mitrovi and Tadi, 2009). The representation of the nodes as vectors  
15 in a  $k$ -dimensional space then provides us with the information needed to obtain  
16 a partition of the network in its  $k$  most relevant communities. Following our  
17 analysis, this possibility comes as no surprise, but the ideas underlying spectral  
18 methods need to be developed further, since we wish to work with nonnegative  
19 matrices of arbitrary shape for which SVD is required, and moreover we have  
20 to address the numerous open questions highlighted by the literature mentioned  
21 above. On the other hand, a SVD-based algorithm for community extraction  
22 may leverage on the results presented in this paper, which may provide a sound  
23 probabilistic foundation for its outcomes.  
24

25 Finally, our work has some interesting implications for the understanding  
26 of the geometrical properties of any normalized nonnegative matrix  $K$  and, by  
27 extension, of its non normalized counterpart  $A$ . In fact, while the largest SV  
28 of the latter matrix conveys only, so to speak, a scalar information, since it is  
29 invariant over all normalized matrices, the other SVs depend on the rank of  $K$   
30 and  $A$ . Thus, our results provide a probabilistic model, with respect to which  
31 we can evaluate the likelihood of observing the rank of any given nonnegative  
32 matrix, while at the same time giving a clear interpretation of this property in  
33 terms of deviations from such model. In fact, since  $\bar{K}$  is a rank-1 matrix, the  
34 probability of observing  $K$  under the ED model is inversely related to the rank  
35 of  $K$  itself.  
36  
37

## 38 References

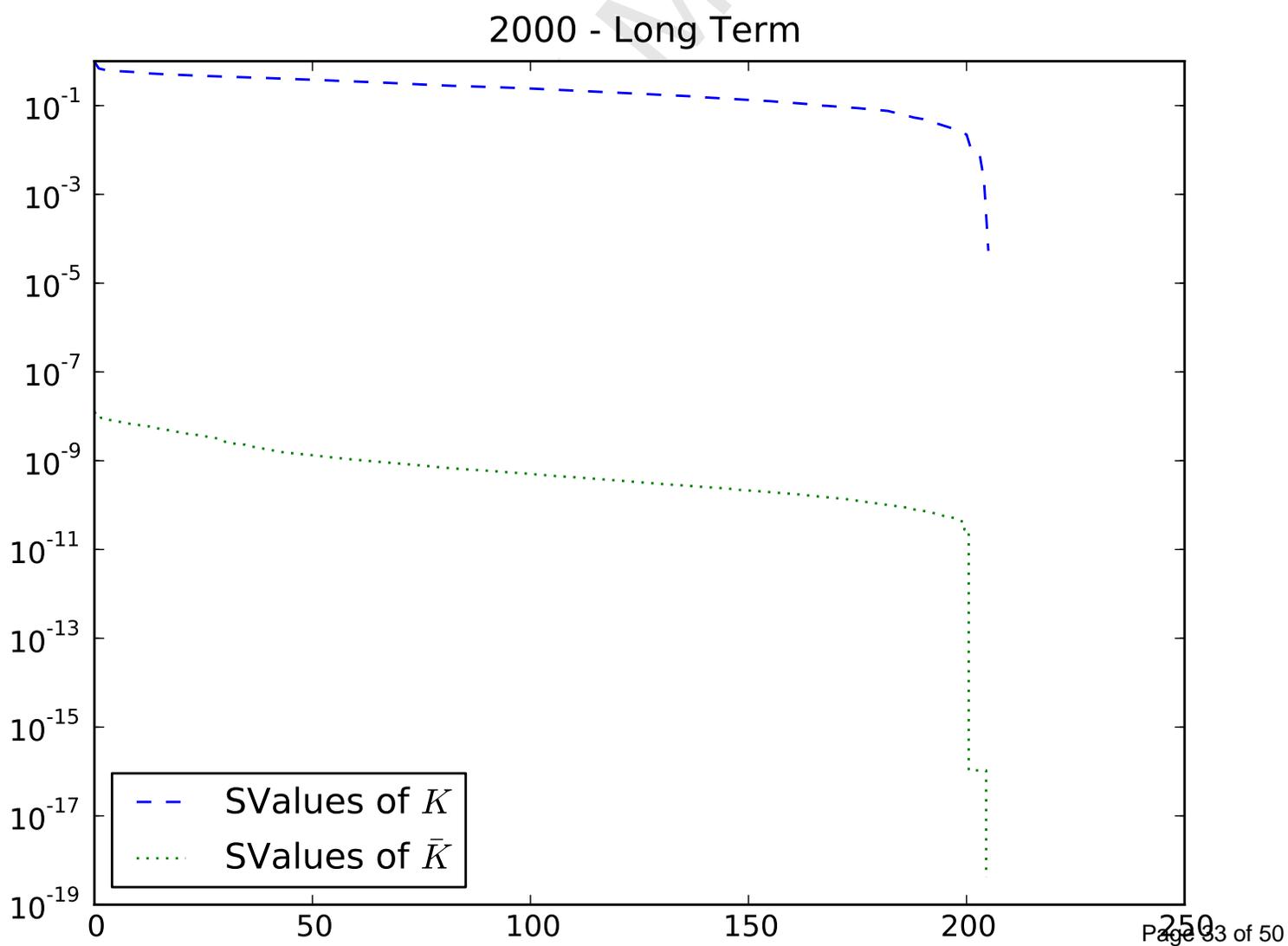
- 39  
40 Kapadia, S., Mora, N., Pühr, C., Alessandri, P., Gai, P., 2009. Towards a  
41 framework of quantifying systemic stability. *International Journal of Central*  
42 *Banking*, 5(3): 47–81, September.  
43  
44 Allen, F., Babus, A., Carletti, E., 2010. Financial connections and systemic  
45 risk. NBER Working Papers 16177, July.  
46  
47 Allen, F., Gale, D., 2000. Financial contagion. *Journal of Political Economy*,  
48 108(1): 1–33.  
49  
50 Arenas, A., Duch, J., Fernández, A., Gómez, S., 2007. Size reduction of complex  
51 networks preserving modularity. *New Journal of Physics*, 9(6): 176.  
52  
53 Battiston, S., Delli Gatti, D., Gallegati, M., Greenwald, B.C., Stiglitz, J.E.,  
54 2009. Liaisons dangereuses: Increasing connectivity, risk sharing, and sys-  
55 temic risk. Working Paper 15611, National Bureau of Economic Research,  
56 January.  
57  
58 Brunnermeier, M.K., 2009. Deciphering the liquidity and credit crunch 2007-  
59 2008. *Journal of Economic Perspectives*, 23(1): 77–100, Winter.  
60  
61  
62  
63  
64  
65

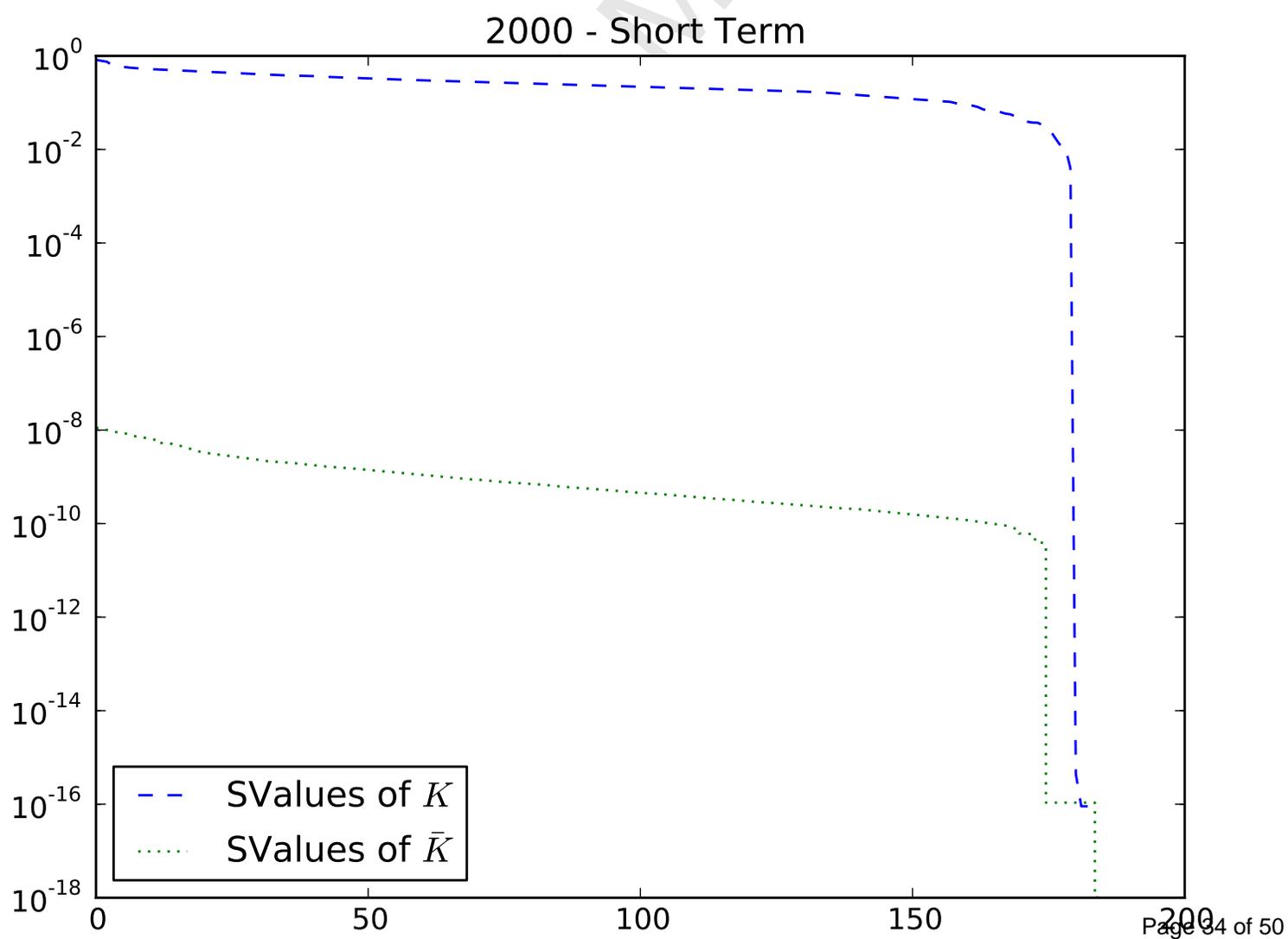
- 1  
2  
3  
4  
5  
6  
7  
8  
9 Butler, S., 2006. Relating singular values and discrepancy of weighted directed  
10 graphs. *SODA '06: Proceedings of the seventeenth annual ACM-SIAM sym-*  
11 *posium on Discrete algorithm*, 1112–1116.
- 12  
13 Castrén, O., Kavonius, I.K., 2009. Balance sheet interlinkages and macro-  
14 financial risk analysis in the euro area. Working Paper Series 1124, European  
15 Central Bank, December.
- 16  
17 Censor, Y., Zenios, S.A., 1997. *Parallel Optimization: Theory, Algorithms, and*  
18 *Applications*. New York: Oxford University Press.
- 19  
20 Chung, F., 2006. *Complex Graphs and Networks*. Conference Board of the  
21 *Mathematical Sciences. Regional Conference Series in Mathematics*, 107.
- 22  
23 Chung, F., Lu, L., 2002a. The average distances in random graphs with given  
24 expected degrees. *Proceedings of the National Academy of Science of the*  
25 *U.S.A.*, 99(25): 15879–15882, December.
- 26  
27 Chung, F., Lu, L., 2002b. Connected components in random graphs with given  
28 expected degree sequences. *Annals of Combinatorics*, 6: 125–145.
- 29  
30 Chung, F., Lu, L., 2006. Concentration inequalities and martingale inequalities:  
31 a survey. *Internet Mathematics*, 3(1): 79–127.
- 32  
33 Chung, F., Lu, L., Vu, V., 2003a. Eigenvalues of random power law graphs.  
34 *Annals of Combinatorics*, 7: 21–33.
- 35  
36 Chung, F., Lu, L., Vu, V., 2003b. Spectra of random graphs with given ex-  
37 pected degrees. *Proceedings of the National Academy of Science of the U.S.A.*,  
38 100(11): 6313–6318, May.
- 39  
40 Chung, F., 1997. *Spectral Graph Theory*. Conference Board of the Mathematical  
41 *Sciences. Regional Conference Series in Mathematics*, 92.
- 42  
43 De Masi, G., Fujiwara, Y., Gallegati, M., Greenwald, B.C., Stiglitz, J.E., 2009.  
44 An analysis of the Japanese credit network. *Quantitative Finance Papers*  
45 0901.2384, arXiv.org.
- 46  
47 Donetti, L., Muñoz, M.A., 2005. Improved spectral algorithm for the detection  
48 of network communities. *AIP Conference Proceedings*, 779: 104-107, July.
- 49  
50 Eisenberg, L., Noe, T.H., 2001. Systemic risk in financial systems. *Management*  
51 *Science*, 47(2): 236-249.
- 52  
53 European Central Bank, 2010. *Recent Advances in Modeling Systemic Risk*  
54 *Using Network Analysis*, Frankfurt am Main, January.
- 55  
56 Fortunato, S., 2010. Community detection in graphs. *Physics Reports*, 486(3-5):  
57 75-174.
- 58  
59 Fujiwara, Y., Aoyama, H., 2008. Large-scale structure of a nation-wide produc-  
60 tion network. *The European Physical Journal B*, 77(4): 565-580.
- 61  
62 Fujiwara, Y., Aoyama, H., Ikeda, Y., Iyetomi, H., Wataru, S., 2009. Structure  
63 and temporal change of the credit network between banks and large firms in  
64 Japan. *Economics e-journal*, 3(7).
- 65

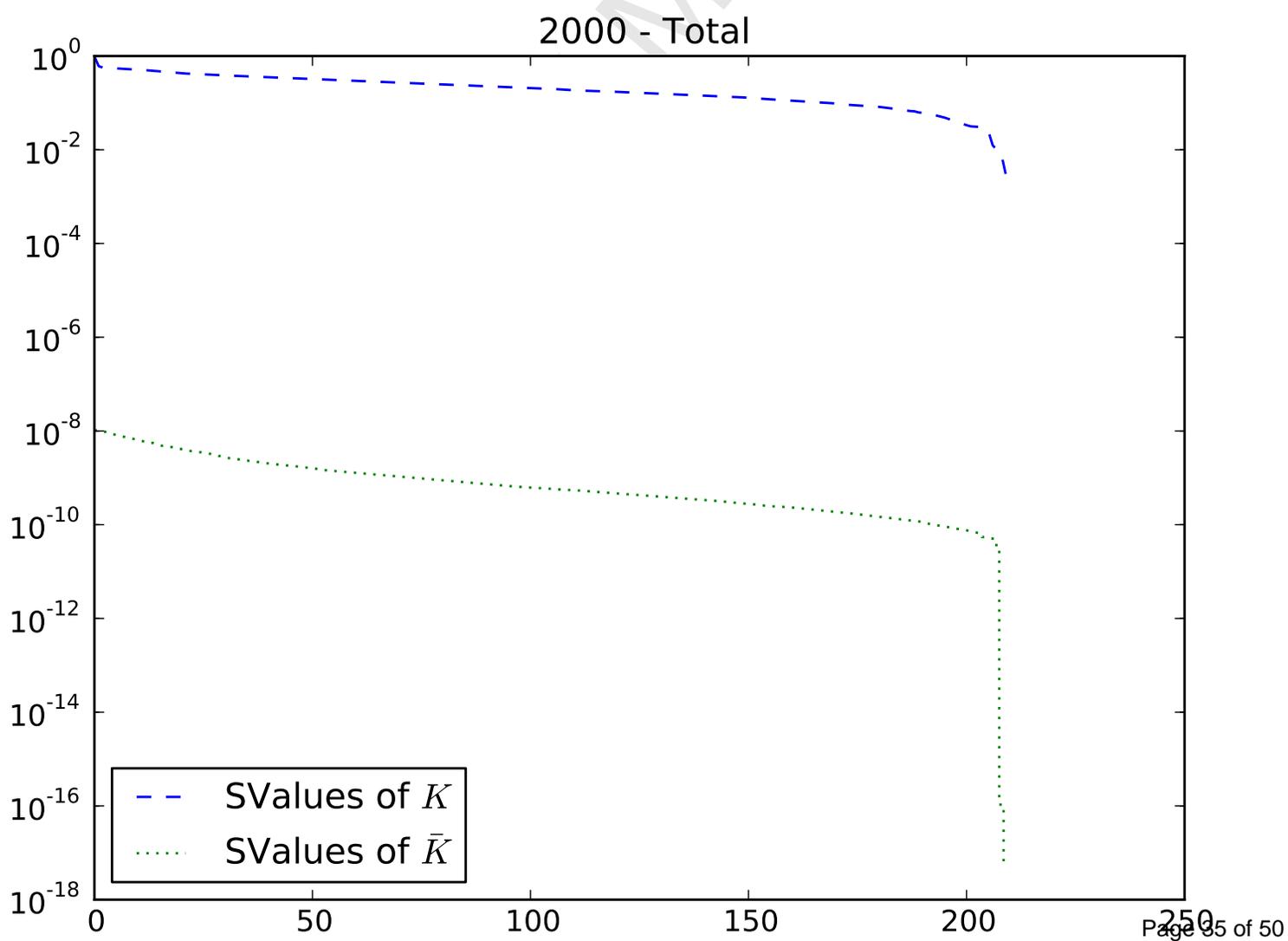
- 1  
2  
3  
4  
5  
6  
7  
8  
9 Gai, P., Kapadia, S., 2010 Contagion in financial networks. Proceedings of the  
10 Royal Society A: Mathematical, Physical and Engineering Science, 466(2120):  
11 2401–2423.
- 12  
13 Garlaschelli, D., Loffredo, M.I., 2008. Maximum likelihood: Extracting unbiased  
14 information from complex networks. Physical Review E, 78(1): 015101, July.
- 15  
16 Garlaschelli, D., Loffredo, M.I., 2009. Generalized bose-fermi statistics and  
17 structural correlations in weighted networks. Physical Review Letters, 102(3):  
18 038701, January.
- 19  
20 Hogben, L. (editor), 2007. Handbook of Linear Algebra. Boca Raton: Chapman  
21 & Hall/CRC.
- 22  
23 Selee, T.M., Ipsen, I.C.F., 2010. Ergodicity coefficients defined by vector norms.  
24 *Mimeo.*
- 25  
26 Mitrović, M., Tadić, B., 2009. Spectral and dynamical properties in classes of  
27 sparse networks with mesoscopic inhomogeneities. Physical Review E, 80(2):  
28 026123, August.
- 29  
30 Newman, M.E.J., 2006. Finding community structure in networks using the  
31 eigenvectors of matrices. Physical Review E, 74:036104.
- 32  
33 Newman, M.E.J., Girvan, M., 2004. Finding and evaluating community struc-  
34 ture in networks. Physical Review E, 69(2): 026113, February.
- 35  
36 Rothblum, U.G., Tan, C.P., 1985. Upper bounds on the maximum modulus  
37 of subdominant eigenvalues of nonnegative matrices. Linear Algebra and its  
38 Applications, 66: 45–86.
- 39  
40 Seneta, E., 1984. Explicit forms for ergodicity coefficients and spectrum local-  
41 ization. Linear Algebra and its Applications, 60: 187–197.
- 42  
43 Shin, H.S., 2008. Risk and liquidity in a system context. Journal of Financial  
44 Intermediation, 17(3): 315–329.
- 45  
46 Greenwald, B.C., Stiglitz, J.E., 2003. Towards a New Paradigm in Monetary  
47 Economics. Cambridge: Cambridge University Press.
- 48  
49 Vega-Redondo, F., 2007. Complex Social Networks. Cambridge: Cambridge  
50 University Press.
- 51  
52 Wells, S., 2004. Financial inter-linkages in the united kingdom’s interbank mar-  
53 ket and the risk of contagion. Bank of England working papers 230, Bank of  
54 England, September.
- 55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

- > We generalize the expected degree model in order to apply it to economic networks
- > We derive new statistical tests for economic networks, based on this theoretical model
- > Applying our tests to credit network data, we show these networks don't follow the model
- > These networks are poorly diversified and contain densely connected communities of nodes
- > Our work suggests ways to improve community detection algorithms for economic networks

Accepted Manuscript

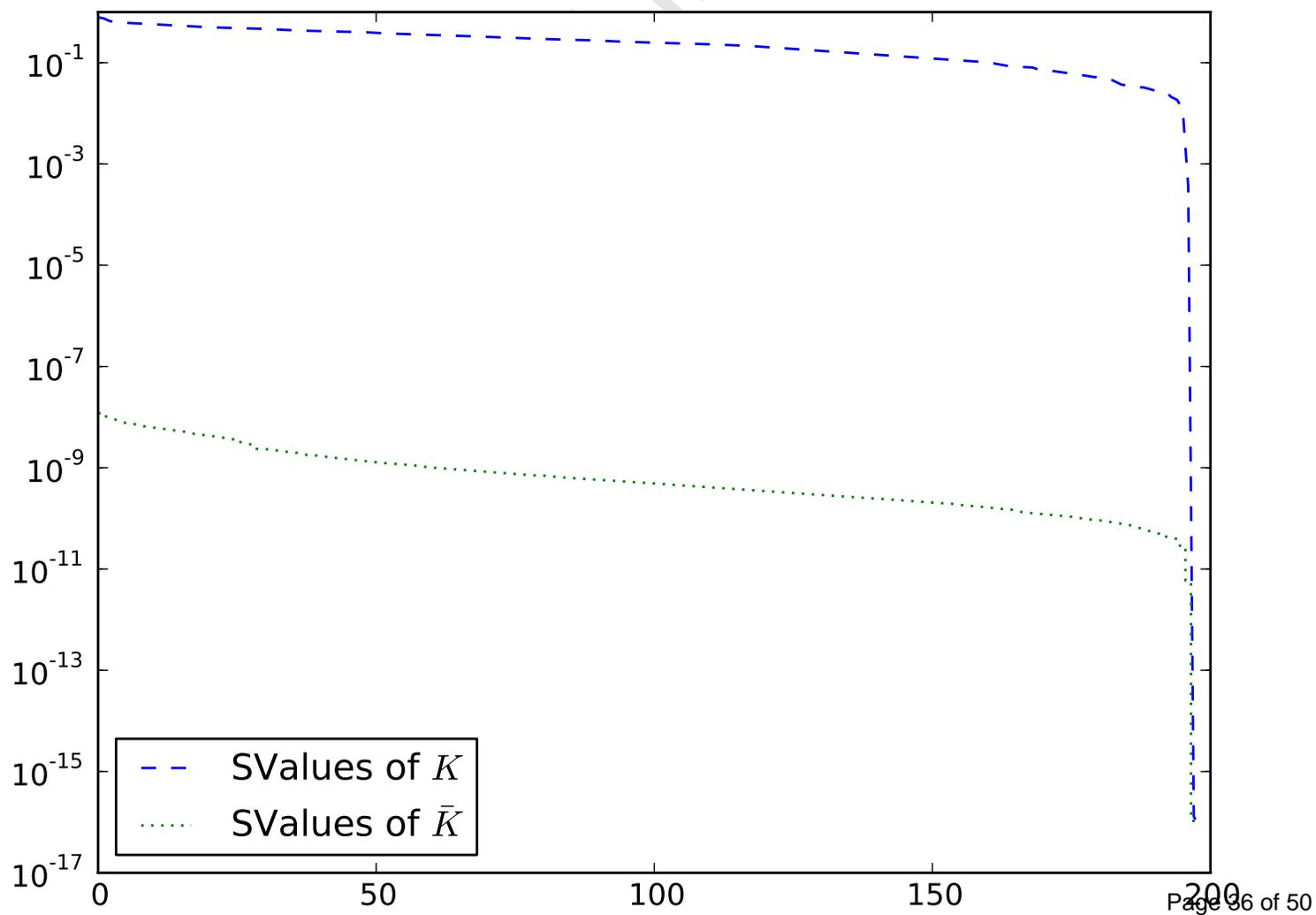


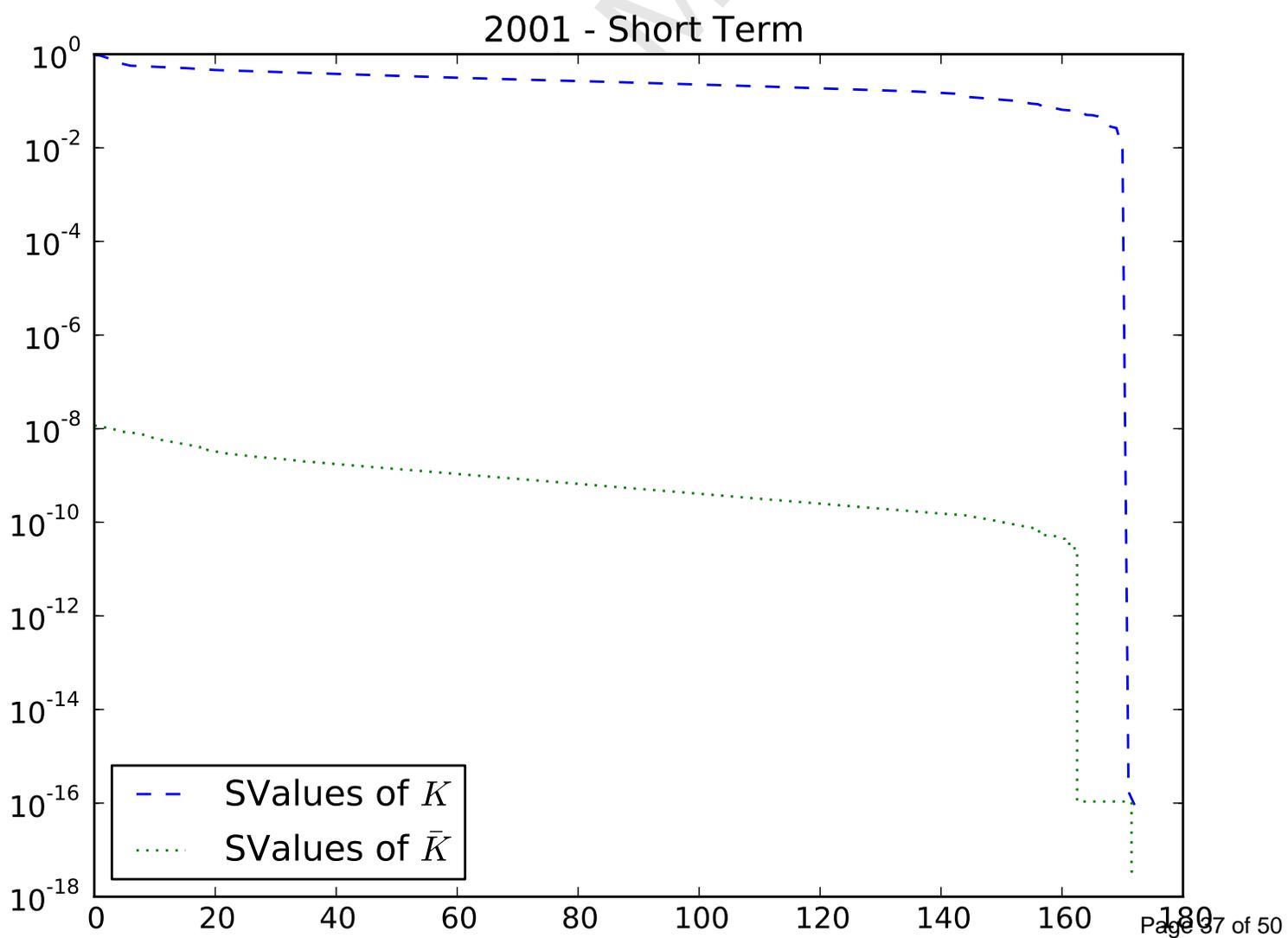


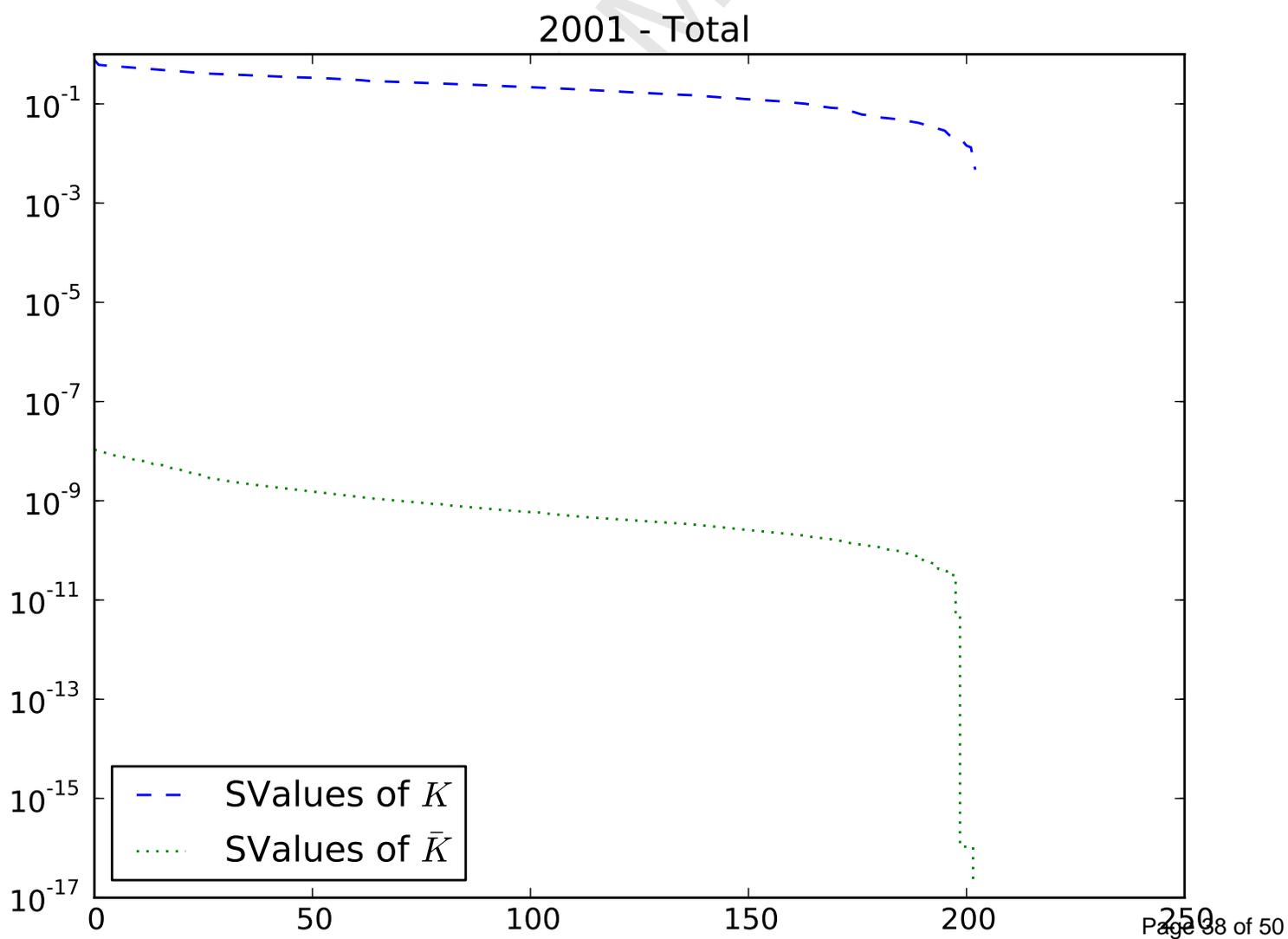


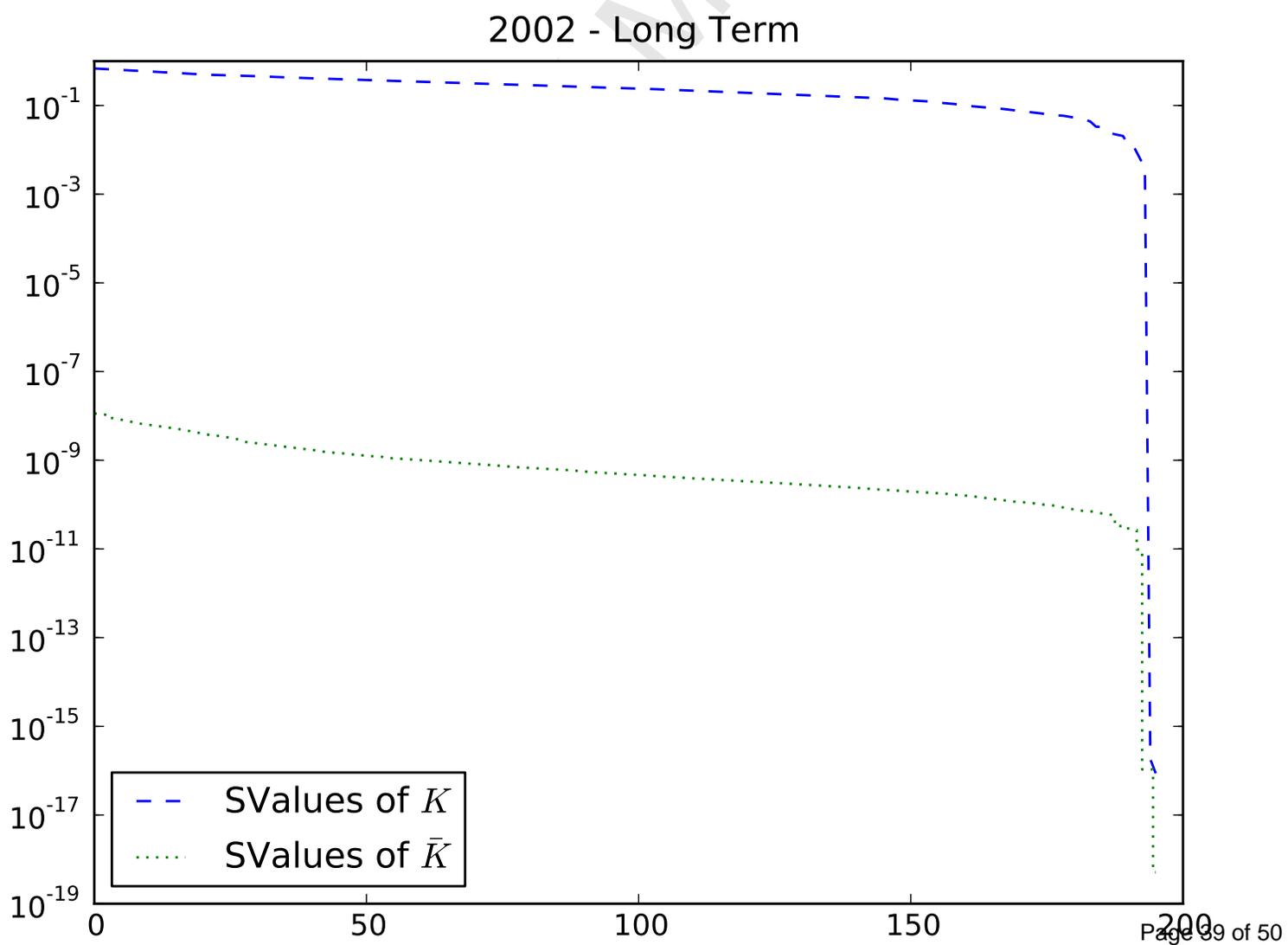
Manuscript

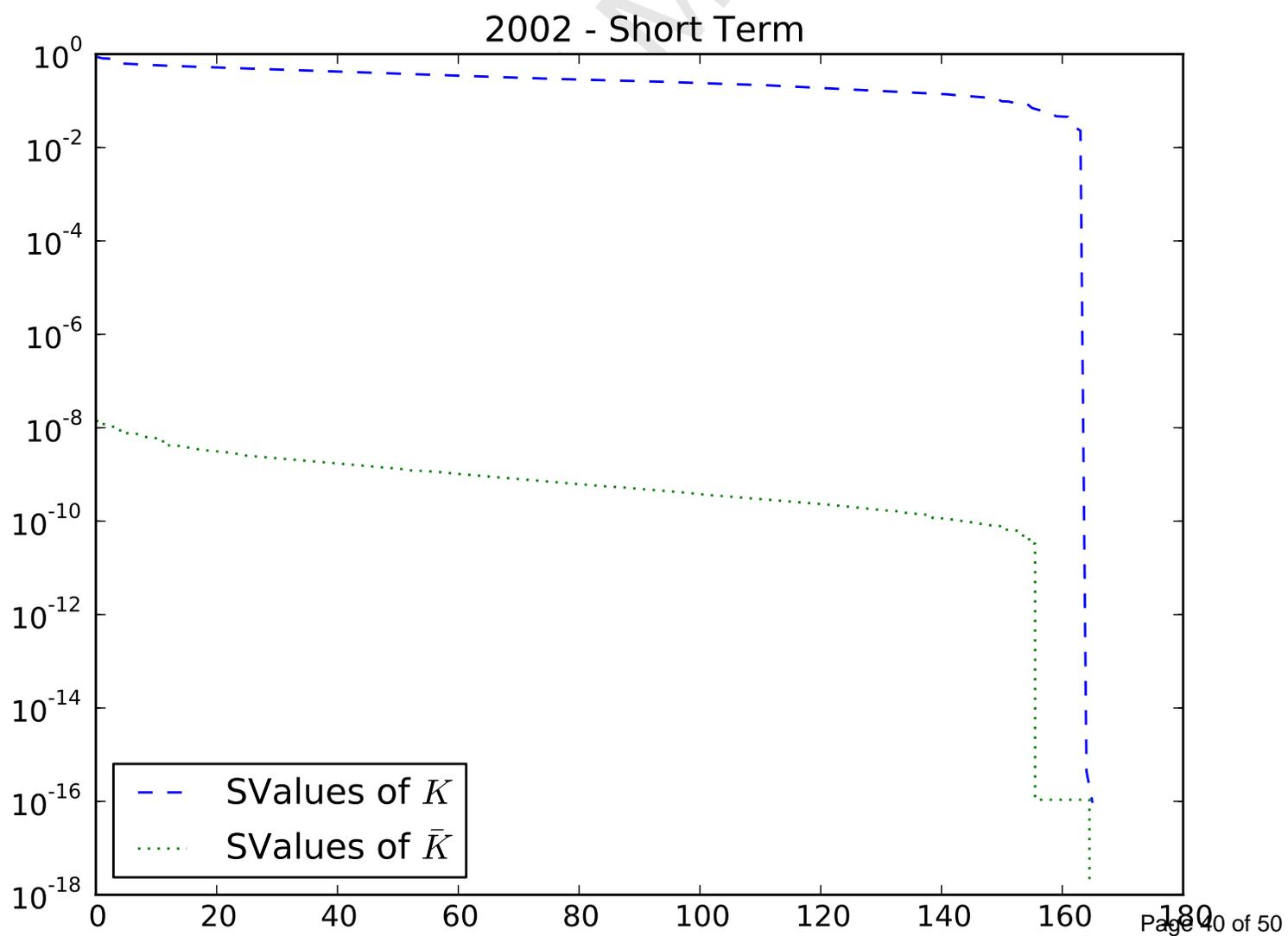
2001 - Long Term

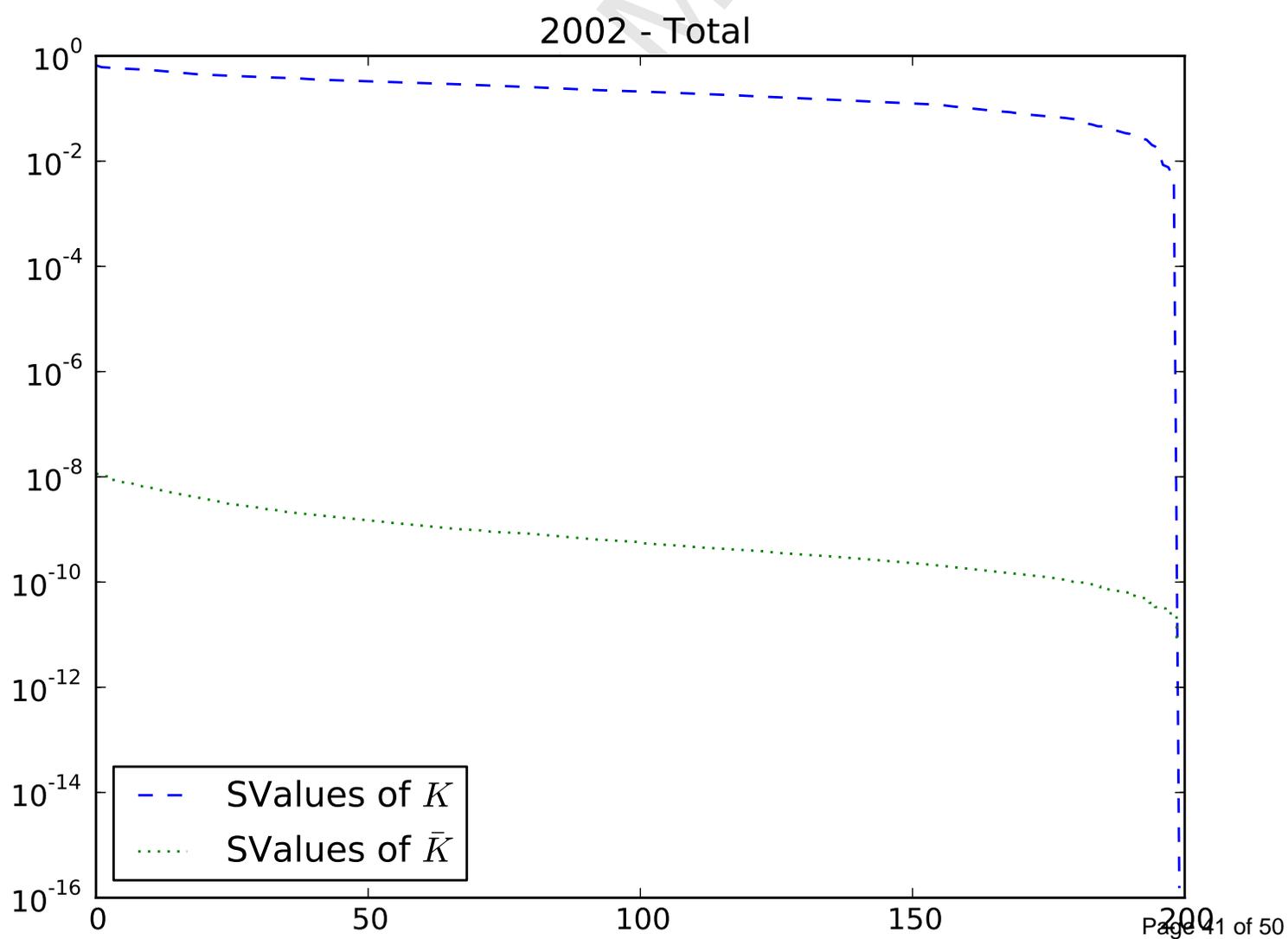




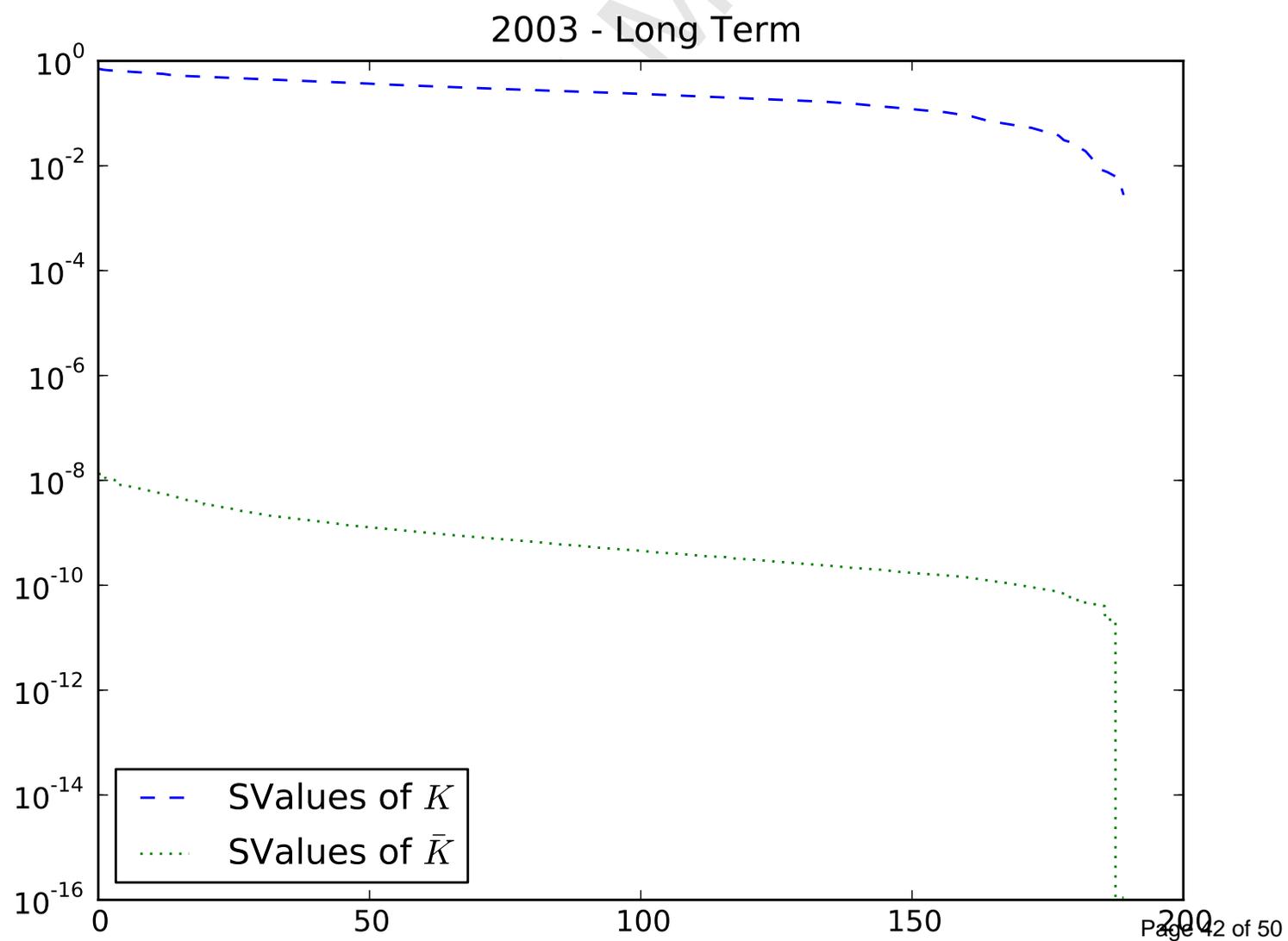


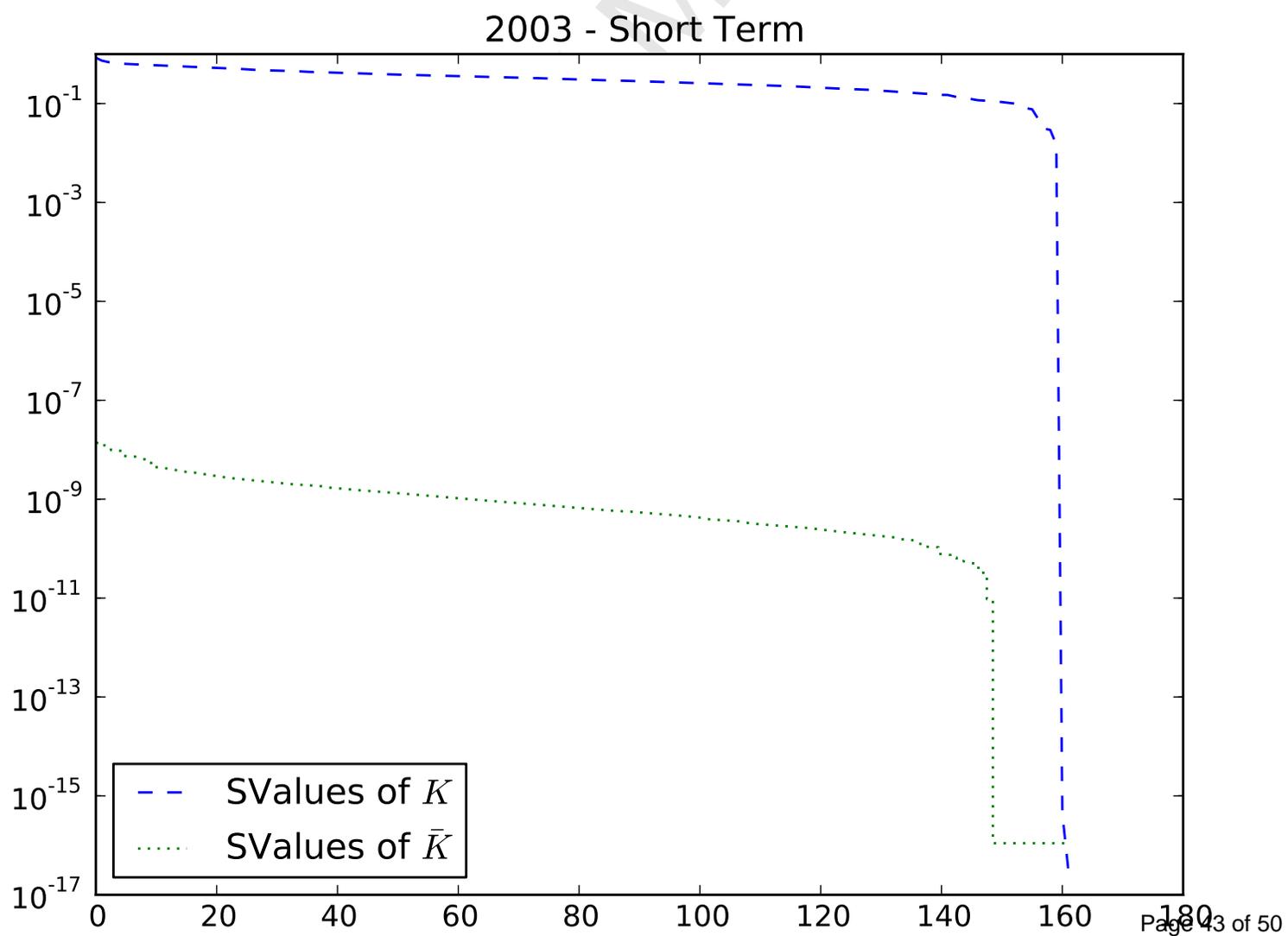




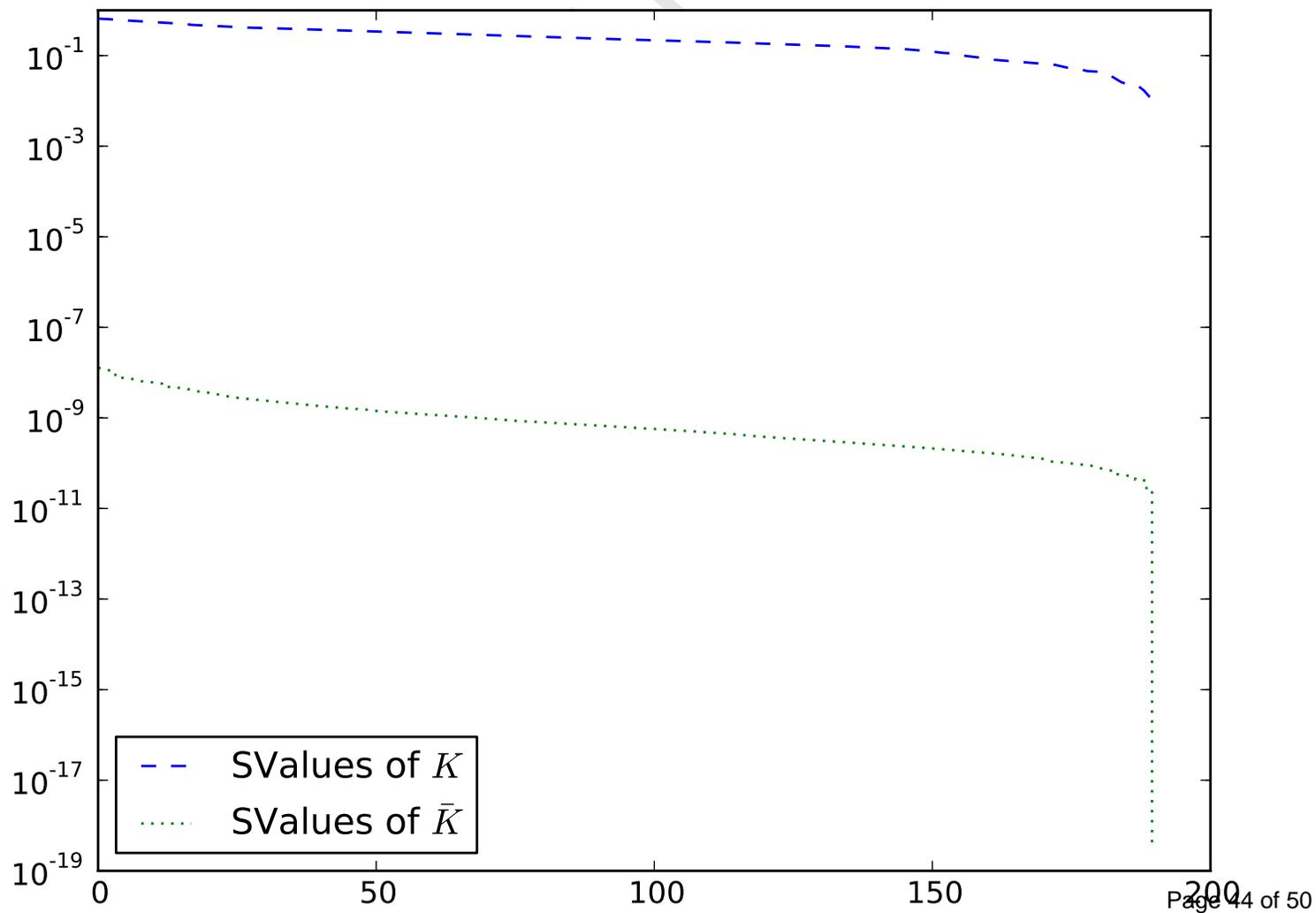


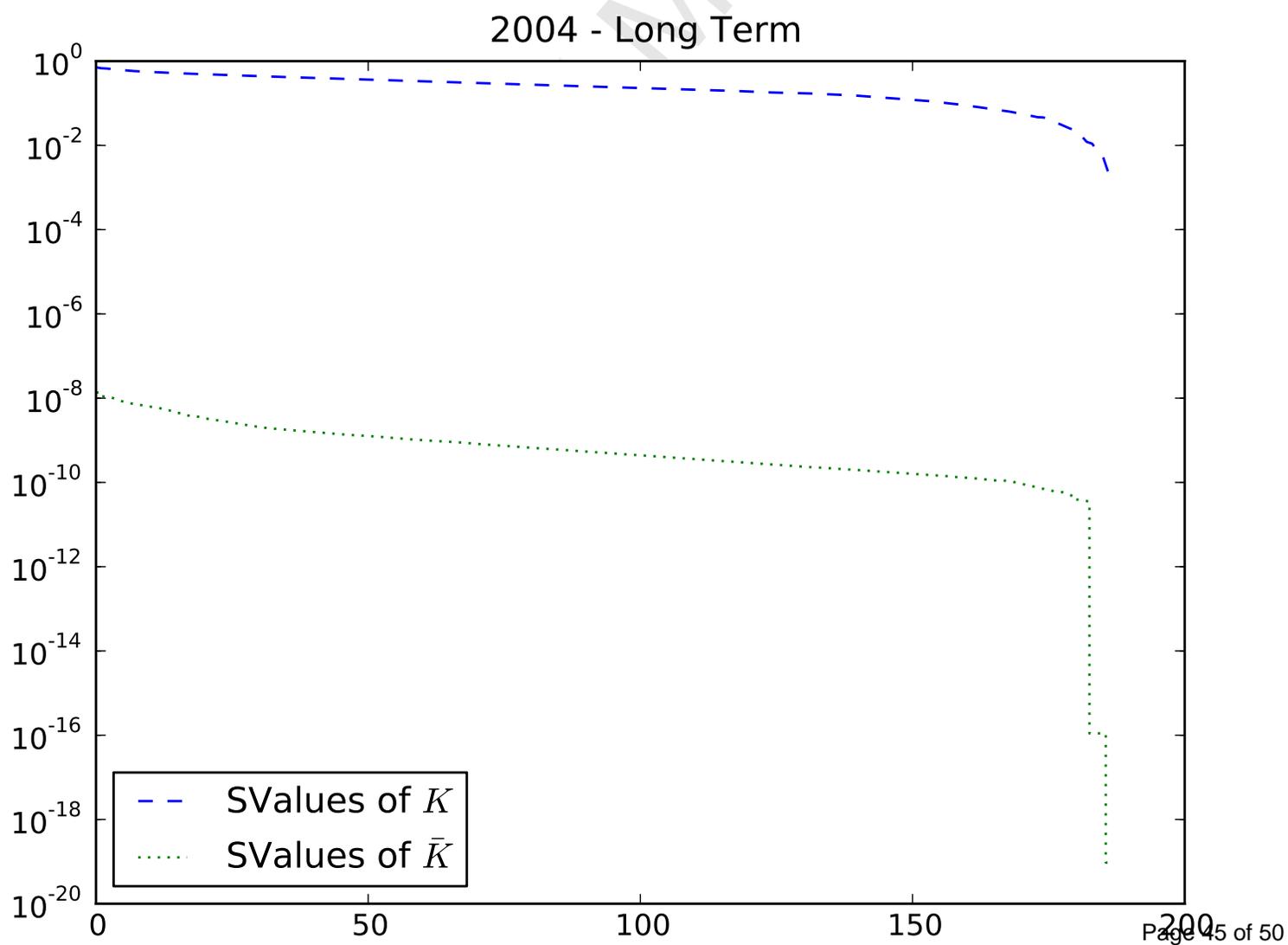
Manuscript

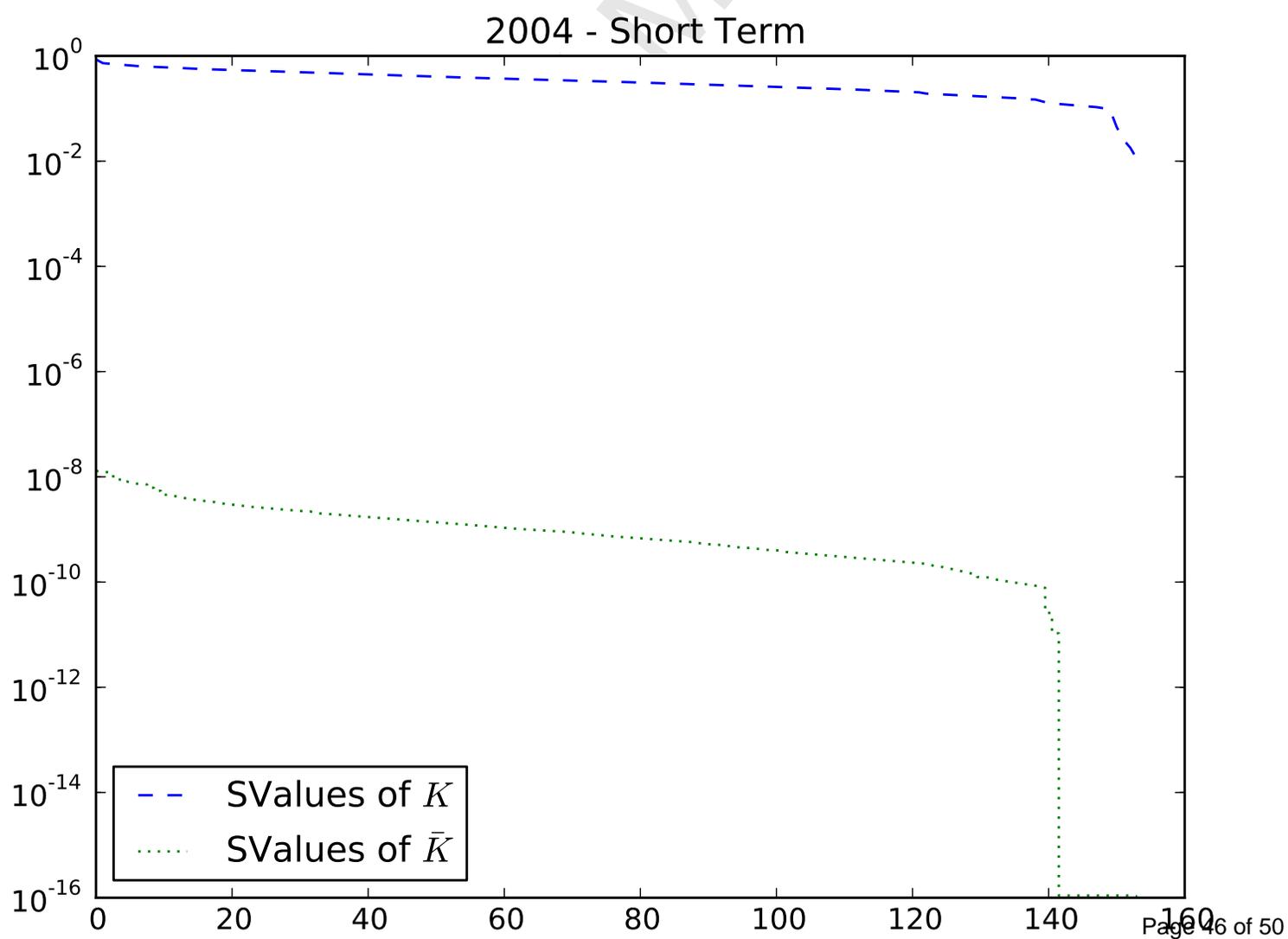


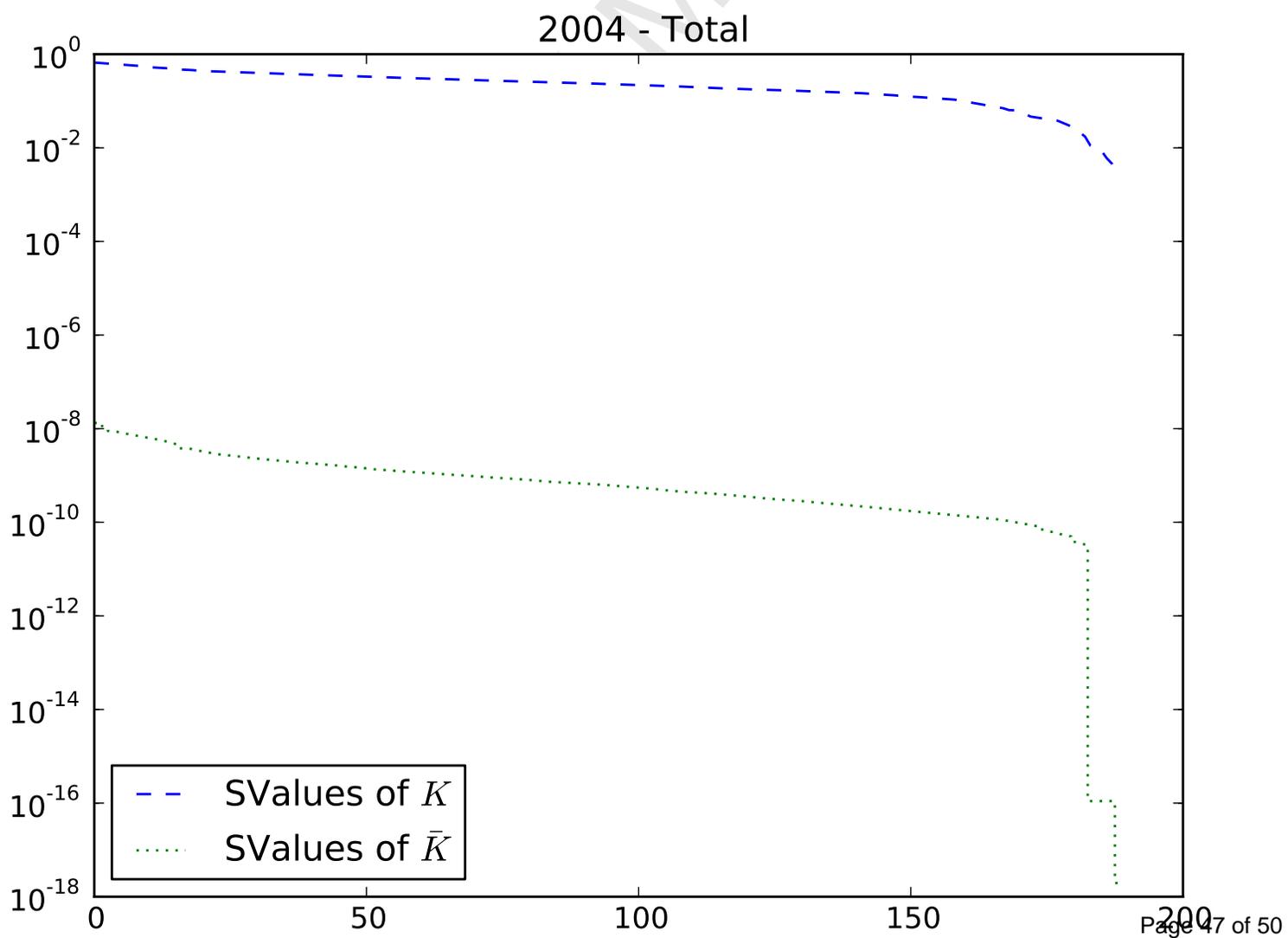


2003 - Total



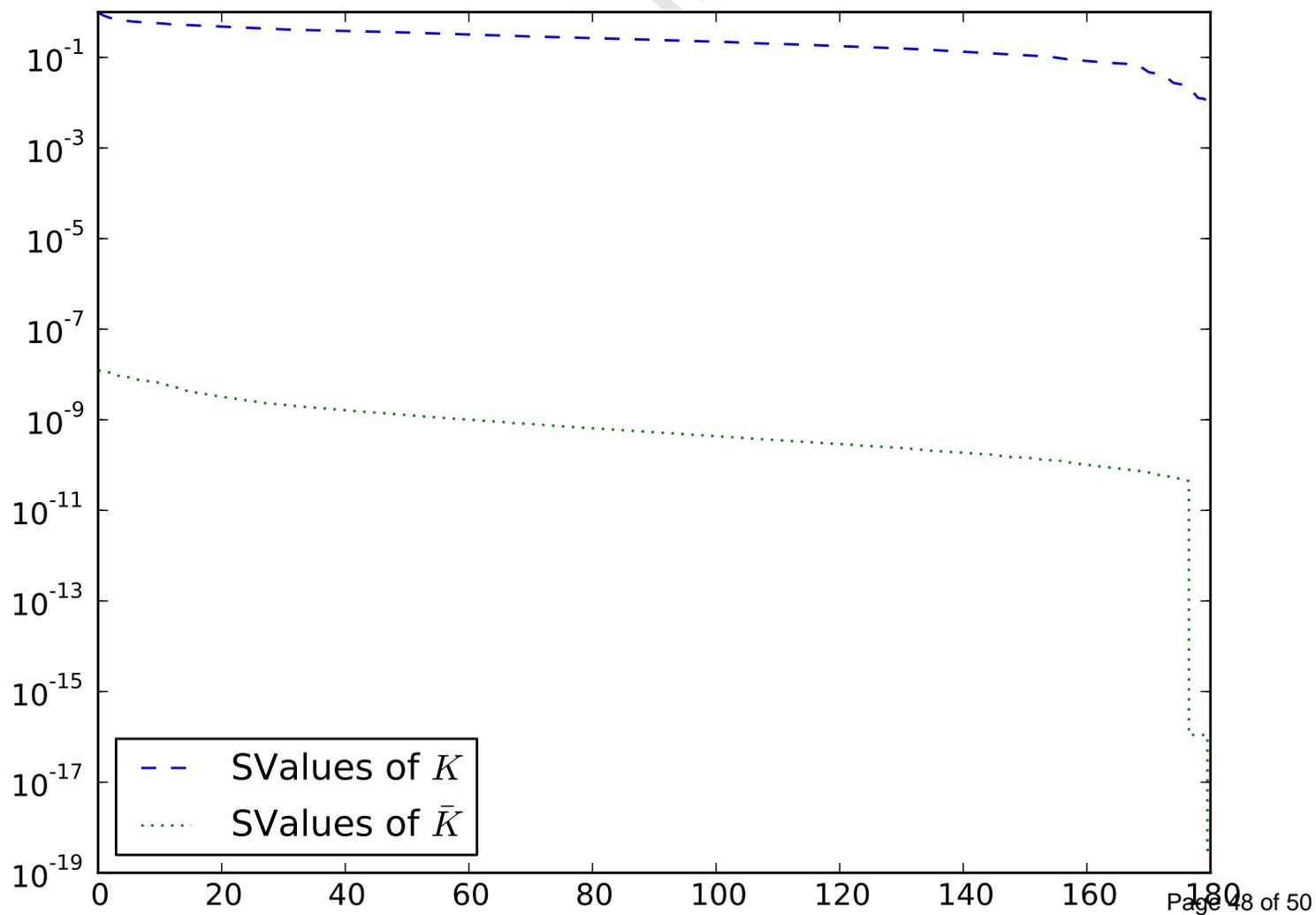


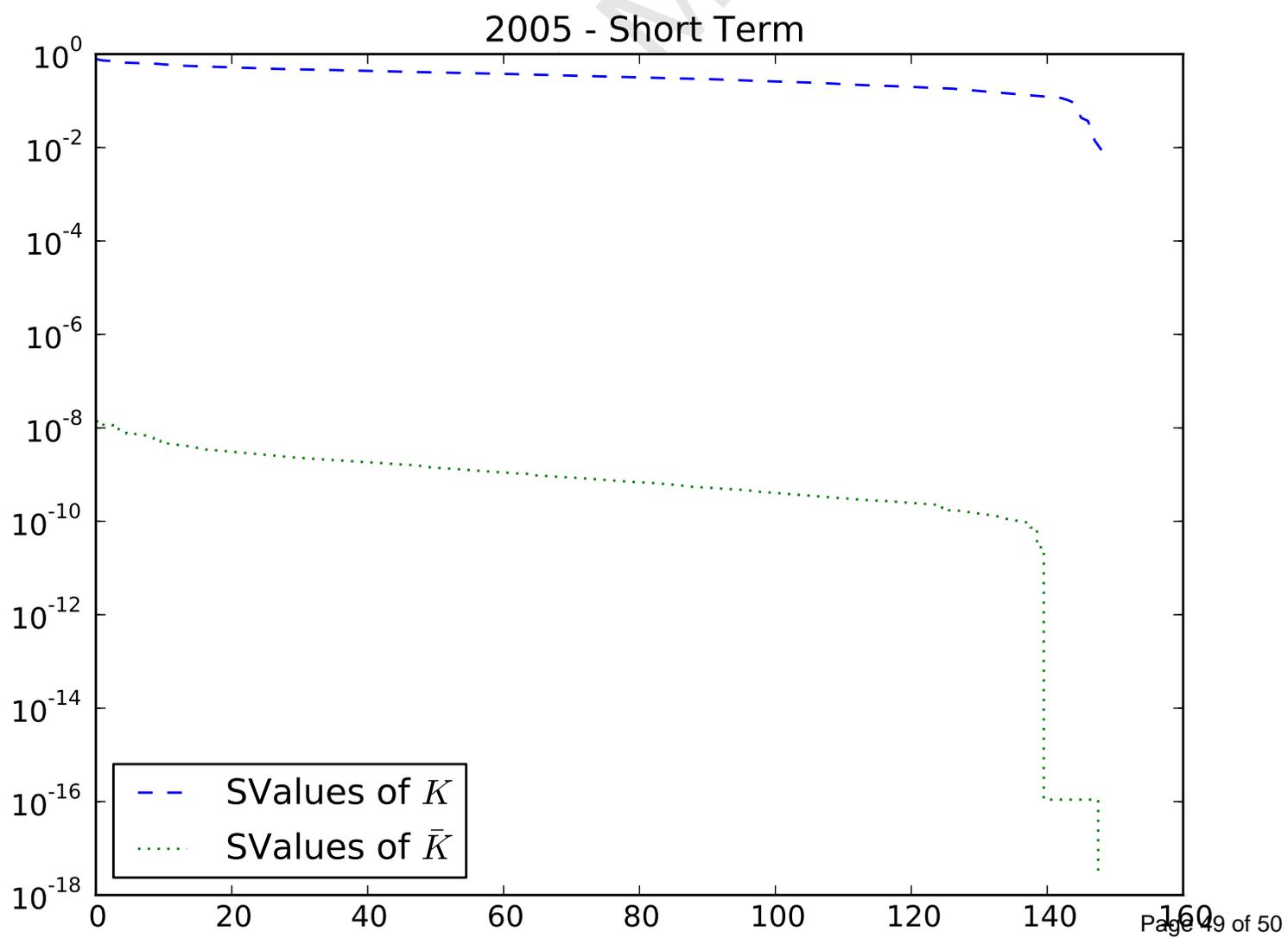




Manuscript

## 2005 - Long Term





2005 - Total

