Contagion Dynamics on Financial Networks*

Monica Billio† Roberto Casarin‡ Michele Costola§
Lorenzo Frattarolo¶

May 28, 2018

Abstract

We provide a graph theoretic background for the analysis of financial networks and review some technique recently proposed for the extraction of financial networks. We develop new measures of network connectivity, that are Von Neumann entropies and disagreement persistence index, using the spectrum of normalized Laplacian and Diplacian. We show that the new measures account for global connectivity patterns given by paths and walks of the network. We apply the new measures to a sequence of inferred pairwise-Granger networks. In the application, we employ the proposed measures for the system immunization and early warning for banking crises.
Keywords: Financial contagion; Financial Network; Financial Crises; Network Entropy
JEL Classification: G12; G29; C51

1 Introduction

Given the relevance of the latest financial and sovereign crises, systemic events are now deeply analysed by scholars and policy makers. As a matter of fact, the studies on the

*We thank the conference participants at the International Conference on Econometrics and Statistics (EcoSta 2017) in Hong Kong, the 10th CSDA International Conference on Computational and Financial Econometrics (CFE 2016) in Seville (Spain) and the Mathematical and Statistical Methods for Actuarial Sciences and Finance (MAF 2016) in Paris (France) for their helpful comments. Michele Costola acknowledges financial support from the Marie Skłodowska-Curie Actions, European Union, Seventh Framework Program HORIZON 2020 under REA grant agreement n. 707070. He also gratefully acknowledges research support from the Research Center SAFE, funded by the State of Hessen initiative for research LOEWE. The computation has been performed using Matlab through the SCSCF (Sistema di Calcolo Scientifico Ca' Foscari) cluster.
†Department of Economics, Ca' Foscari University of Venice, Dorsoduro 3246, 30123 Venice (Italy), billio@unive.it (e-mail). Corresponding author.
‡Department of Economics, Ca' Foscari University of Venice, Dorsoduro 3246, 30123 Venice (Italy), r.casarin@unive.it (e-mail).
§SAFE, House of Finance, Goethe University Frankfurt, Theodor-W.-Adorno-Platz 3, 60323 Frankfurt am Main (Germany), costola@safe.uni-frankfurt.de (e-mail).
¶Department of Economics, Ca' Foscari University of Venice, Dorsoduro 3246, 30123 Venice (Italy), lorenzo.frattarolo@unive.it (e-mail).

consequences of systemic risk are relevant both for the stability of the financial and banking system and in terms of diversification in an investor perspective [**?**]. As in **?**, we define systemic risk "any set of circumstances that threatens the stability of or public confidence in the financial system" where interconnectedness among financial institutions and markets represents a potential channel in propagation of shocks to the system [**??**]. Pairwise Granger causality tests have been used to extract the network of significant linkages among financial institutions and to find which ones are systemically important [**?**]. The question here is to compare the network defined over time (using pairwise Granger causality tests), to identify possible distortions, sources of systemic risk. Recently, entropy measures have been involved in systemic risk measurement for propagation of financial contagion [**?**] and as early warning indicator for banking crises [**?**]. **?** use the Shannon, Tsallis and Rényi entropies of the degree distribution of financial networks. While these measures revealed effective in predicting banking crisis, they take into account only the dispersion of the degree distribution. In this paper we propose an alternative entropy measure, the Von Neumann entropy, which is specifically designed for networks. The Von Neumann entropy has been widely used for the analysis of complex systems and it naturally arises from the relationship between network structure and density matrix of a state of a quantum system [for definition and references, see **?**]. Unfortunately, most of this literature has focused on undirected networks and has used both the adjacency matrix and the Laplacian matrix to obtain the association with a state of a quantum system. In any case, the analogy with a quantum system is not in any way essential to understand the remaining of the paper. Within the few works providing extensions of the Von Neumann entropy for undirected graphs to the case of directed graphs, **?** certainly represents a relevant reference. The peculiarity of directed networks and in particular their asymmetry requires some care and redefinition of the Laplacian [**?**]. In particular, the combinatorial Laplacian for directed networks is formulated in terms of nodes out degree, the Perron vector of the transition matrix, that is related to eigenvector centrality, and can be soundly computed only for strongly connected components, and related to the circulation on the graph [**?**]. In this paper, we build on **?** and propose an alternative definition of Von Neumann entropy for directed graphs which accounts for positive Laplacian matrix. We show that the new measures can be written as weighted sums of walks of the network and relate them to consensus dynamics.

This chapter is organized as follows. Section **??** introduces a notion of financial network

and some background in graph theory. Section **??** reviews some methods of network extraction. Section **??** presents classical network measures and Section **??** discusses our new measures based on the notion of Von Neumann entropy while in Section **??**, we discuss the role of the Diplacian in consensus dynamics and design an associated measure. Section **??** provides an empirical application.

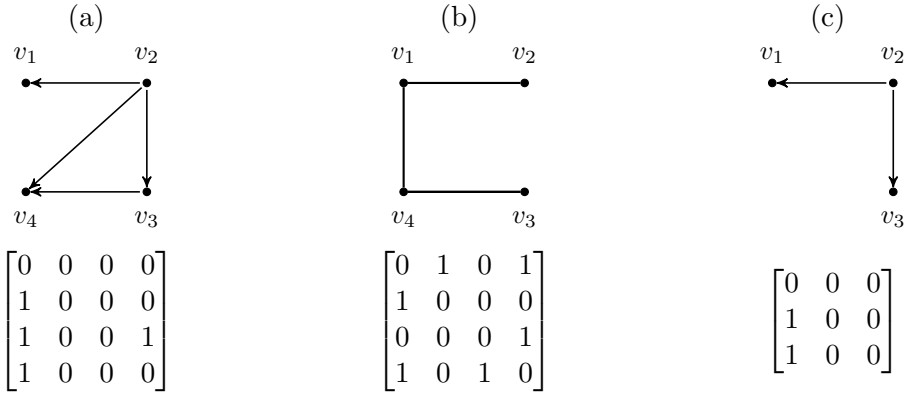
2 Financial Networks

A network can be defined as a set of vertices (or nodes) and arcs (or edges) between vertices. In financial networks, a node represents a financial institution (e.g., a bank, an insurance company, a financial agglomeration) and an edge has the interpretation of financial linkage between two institutions. In mathematical terms a network can be represented through the notion of graph and its properties. In the following sections we provide some background in graph theory useful for a better comprehension of the new indicators developed in this paper and of the analysis of financial networks. For further material on graph theory and random graph we refer the interested reader to **?** and **?**. See **?** for an introduction to network theory in social sciences.

2.1 Graph Theoretic Foundation

A graph is defined as the ordered pair of sets $G = (V, E)$ where $V = \{1, \ldots, n\}$ is the set of vertices (or nodes) and $E \subset V \times V$ the set of edges (or arcs). The order of a graph is the number of vertices in V, that is the cardinality of V denoted with $|V|$. An (directed) edge between two nodes exists if there is a relationship between them and it can be identified as the (ordered) pair $\{u, v\}$ with $u, v \in V$. If there is no direction in the connection between nodes then an edge $\{u, v\}$ is an unordered pair of nodes and the graph G is said to be undirected, whereas if a direction exists, then each edge $\{u, v\}$ is defined as an ordered pair of nodes and the graph G is said to be directed graph (or digraph).

Assume for simplicity the graph $G = (V, E)$ is undirected. If $\{u, v\} \in E$ then u and v are adjacent vertices and they are incident with the edge $\{u, v\}$. For each node u, it possible to define its neighborhood as the set of nodes adjacent to u, that is $N_u = \{v \in V; \{u, v\} \in E\}$. The vertex adjacency structure of a n-order graph $G = (V, E)$ can be represented through a n-dimensional matrix A called adjacency matrix. Each element a_{uv} of the adjacency matrix is equal to 1 if there is an edge from institution u to institution v



$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Figure 1: Panel (a): directed graph $G = (V, E)$ (top) with vertex set $V = \{v_1, v_2, v_3, v_4\}$ and edge set $E = \{e_1, e_2, e_3, e_4\}$, where $e_1 = \{v_2, v_1\}$, $e_2 = \{v_2, v_3\}$, $e_3 = \{v_2, v_4\}$, $e_4 = \{v_3, v_4\}$ and its adjacency matrix (bottom). Panel (b): undirected graph $G = (V, E)$ (top) with vertex set $V = \{v_1, v_2, v_3, v_4\}$ and edge set $E = \{e_1, e_2, e_3\}$, where $e_1 = \{v_1, v_2\}$, $e_2 = \{v_1, v_4\}$, $e_3 = \{v_3, v_4\}$ and its adjacency matrix (bottom). Panel (c): subgraph of the graph given in panel (a).

with $u, v \in V$, and 0 otherwise, where $u \neq v$, since self-loops are not allowed. If the graph is undirected than $a_{uv} = a_{vu}$, that is the adjacency matrix is symmetric.

As an example, Figure **??** includes two graphs, one directed (panel (a)) and the other undirected (panel (b)). The edges of the directed graph are $e_1 = \{v_2, v_1\}$, $e_2 = \{v_2, v_4\}$, $e_3 = \{v_2, v_3\}$ and $e_4 = \{v_3, v_4\}$ and its adjacency matrix is given in the second line of the same panel. The edges of the undirected graph are $e_1 = \{v_1, v_2\}$, $e_2 = \{v_1, v_4\}$, $e_3 = \{v_3, v_4\}$ and its adjacency matrix is given in the second line of the same panel.

In some applications it is useful to focus the analysis on a part of the graph. We say that $G' = (v', E')$ is a subgraph of G if $V' \subset V$ and $E' \subset E$. The subgraph can be induced by a subset of edges or by a subset of nodes. Panel (c) of Figure **??** shows, as an example, the subgraph of the directed graph reported in Panel (a). Given two subgraphs of G, $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, the graph union $G_3 = G_1 \cup G_2$ is defined as the graph $G_3 = (V_3, E_3)$ such that $V_3 = V_1 \cup V_2$ and $E_3 = \{\{u, v\} \in E; u, v \in V_3\}$. Note that $E_1 \cup E_2 \subset E_3$. The graph difference $G_3 = G_2 \backslash G_1$ with $V_1 \subset V_2$, is defined as the graph $G_3 = (V_3, E_3)$ such that $V_3 = V_2 \backslash V_1$ and $E_3 = \{\{u, v\} \in E; u, v \in V_3\}$.

2.2 Graph Connectivity

The two extreme configurations of the connectivity structure of a n-order graph G are given by the graph with empty edge set, i.e. $|E| = 0$, which is called empty graph and denoted with E_n and the complete graph where each node is adjacent to all other nodes in

the graph. In this case, the cardinality of the edge set is maximal, i.e. $|E| = n(n-1)/2$, and graph is denoted with K_n. Panels (a) and (b) of Figure **??** show an example of complete, K_4, and empty, E_4, graphs.

In the connectivity structure of a graph and in spreading of contagion in a network the cohesiveness and the indirect connections between nodes play a crucial role. The cohesiveness can be represented through the number and size of cliques or of communities. The notion of indirect connection can be made more precise through the definitions of walk, trail, path, circuit and cycle.

A clique $C \subset G$, is defined as an ordered pair of sets $C = (V_C, E_C)$ with $V_C \subset V$ and $E_C = \{\{u, v\} \in E; u, v \in V_C\}$ such that $m = |V_C| > 2$, $E_C = K_m$ and $C \cup \{w\}$ with $w \in G \backslash C$ is not complete.

A walk $W_{uv} = (v_0, e_1, \ldots, e_l, v_l)$ between two vertices u and v of G, called endvertices, is identified by an alternating sequence of (not necessary distinct) vertices $V(W_{uv}) = \{v_0, v_1, \ldots, v_l\}$ and edges $E(W_{uv}) = \{e_1, \ldots, e_l\} \subset E$, with $e_1 = (v_0, v_1)$, $e_l = \{v_{l-1}, v_l\}$, and $v_0 = u$ and $v_l = v$. The number of edges $|E(W_{uv})| = l$ in a walk is called "walk length". A walk of length l is called l-walk and denoted with W_l. It is easy to show that the number of l-walks from node u to node v is equal to the (u, v)-th element of A^l that is equal to

$$\sum_{v_1=1}^{n} \sum_{v_2=1}^{n} \cdots \sum_{v_{l-1}=1}^{n} a_{uv_1} a_{v_2 v_3} \ldots a_{v_{l-1} v}. \tag{1}$$

If all edges are distinct then the walk is called a trail. A trail with coincident endvertices is called a circuit (or closed trail). A walk W_l with $l \geq 3$ with $v_0 = v_l$ and vertices v_j, $0 < j < l$ distinct from each other and from v_0, is called cycle and denoted with C_l. An example of cycle C_4 is given in panel (d) of Figure **??**.

A path P_{uv} between vertices u and v of G is a walk with distinct elements in its vertex set. A generic path of length l is denoted with l. The shortest-path P^*_{uv} between two vertices u and v is $\min_l \{P_{uv} = (v_0, e_1, \ldots, e_l, v_l), l \geq 1\}$ that is the path with the minimum length. An example of path P_2 is given in panel (c) of Figure **??**. The notion of shortest path is relevant in spreading of contagion in financial networks. The average shortest path over pairs of nodes reflects the time a shock takes to spread out in the network. The lower the average shortest path is the higher will be speed of shocks transmission. Moreover, if a function of the losses is assigned as weight to the edges between nodes, then the shortest path can be used to provide a measures of the minimum loss following a transmission of

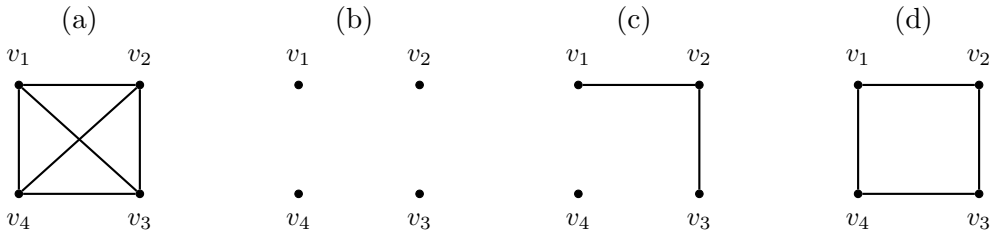


Figure 2: Example of complete graph K_4 (a), empty graph E_4 (b), path P_2 (c) and cycle C_4 (d).

shocks in the financial networks.

The notion of path allows us to introduce the definition of connected graph and some other basic graph structures. A graph is connected if for every pair of distinct vertices u and v there is a path from u to v. A maximal connected subgraph is a component of the graph. A cutvertex is a vertex whose deletion increases the number of components. An edge is a bridge if its deletion increases the number of components. A graph without cycles is a forest or an acyclic graph. A tree is a connected forest.

Also, the notion of path can be used to define the distance between two nodes u and v, $d(u, v)$ as the length of the shortest path or geodesic between u and v. The notion of distance allows us to define the diameter of G, $\text{diam}(G)$ as the $\max_{u,v \in V} d(u, v)$ and the radius of G, $\text{rad}(G)$, as the $\min_u \max_v d(u, v)$. If the graph G is connected then there exists and integer $l > 0$ for which the (u, v)-th element of A^l is not equal to 0 for each pair (u, v) and the lowest integer l^* such that A^{l^*} is not equal to 0 for each pair of nodes is $\text{diam}(G)$. Thus the diameter is equal to the length of the longest shortest path in G.

Finally, note that it is possible to define graph measures using cycles rather than paths. The girth of a graph G, $\text{gir}(G)$, is the length of the shortest cycle in a network (set to infinity if there are no cycles) and the circumference, $\text{circ}(G)$, is the length of the largest cycle.

3 Network Extraction

Financial linkages among all the institutions in the system are commonly unobservable, but they can be inferred from data by applying suitable econometric tools. In the following, we review some techniques used to extract financial networks from a panel of observations which contains, for each institution i, $i = 1, \ldots, n$, the time series of a variable of interest, y_{it}, $t = 1, \ldots, T$, such as financial returns, or realized volatility.

3.1 Pairwise and Conditional Granger Networks

? proposed pairwise Granger causality between returns to extract the network of financial institutions. The adjacency matrix A is estimated using a Granger causality test on pairs of time series to detect the direction and propagation of the shocks among the two institutions considered. In the pairwise-Granger approach to network extraction, the following bivariate vector autoregressive model (VAR) of the order one is estimated

$$\begin{cases} y_{it} = \varphi_{10} + \varphi_{11} y_{it-1} + \varphi_{12} y_{jt-1} + \varepsilon_{it} \\ y_{jt} = \varphi_{20} + \varphi_{21} y_{it-1} + \varphi_{22} y_{jt-1} + \varepsilon_{jt} \end{cases} \tag{2}$$

$\forall i, j = 1, \ldots, n$, $i \neq j$, where ε_{it} and ε_{jt} are uncorrelated white noise processes. Then, a test for the existence of Granger causality is applied. The definition of causality implies, for $t = 1, \ldots, T$,

- if $\varphi_{12} \neq 0$ and $\varphi_{21} = 0$, y_{jt} causes y_{it} and $a_{ji} = 1$;
- if $\varphi_{12} = 0$ and $\varphi_{21} \neq 0$, y_{it} causes y_{jt} and $a_{ij} = 1$;
- if $\varphi_{12} \neq 0$ and $\varphi_{21} \neq 0$, there is a feedback relationship among y_{it} and y_{jt} and $a_{ij} = a_{ji} = 1$, where $a_{i,j}$ is the element in $i - th$ row and $j - th$ column of the adjacency matrix A.

The standard pairwise Granger causality approach deals only with bivariate time series and does not consider the conditioning on relevant covariates. In order to account for spurious causality effects, the model given above can be extended to include exogenous variables and endogenous variables at higher order lags. The maximum lag can be selected according to some criteria, such as AIC or BIC. The conditional Granger approach considers the conditioning on relevant covariates, however, with a higher number of variables relative to the number of data point, it encounters problems of over-parameterization, that leads to a loss of degrees of freedom and to inefficiency in correctly gauging the causal relationships.

3.2 Granger Networks and Graphical Models

?? proposed an alternative approach for network estimation based on graphical models.

In their approach, the following structural VAR (SVAR) model is estimated,

$$\begin{cases} y_{1t} - \sum_{j=2}^{n} \gamma_{1j} y_{jt} = \sum_{j=1}^{n} \varphi_{1j} y_{jt-1} + \varepsilon_{1t} \\ y_{2t} - \sum_{\substack{j=1 \\ j\neq 2}}^{n} \gamma_{2j} y_{jt} = \sum_{j=1}^{n} \varphi_{2j} y_{jt-1} + \varepsilon_{2t} \\ \vdots \\ y_{nt} - \sum_{j=1}^{n-1} \gamma_{nj} y_{jt} = \sum_{j=1}^{n} \varphi_{nj} y_{jt-1} + \varepsilon_{nt} \end{cases} \tag{3}$$

where ε_{it}, $i = 1, \ldots, n$, are uncorrelated white noise processes, φ_{ij} are the autoregressive coefficients of the lagged dependence structure and γ_{ij} the structural coefficients of the contemporaneous dependence structure. In this approach two kind of networks are extracted: the undirected acyclic graph G_0 for the contemporaneous dependence and the directed graph G for the lagged dependence structure. To this aim, the coefficients are re-parametrized as $\gamma_{ij} = a_{0,ij}\gamma^*_{ij}$ where $a_{0,ij} \in \{0, 1\}$ is the element of the adjacency matrix of G_0 and $\varphi_{ij} = a_{ij}\varphi^*_{ij}$ where a_{ij} is an element of the adjacency matrix of G. Then the binary connectivity variables are estimated by applying a Markov-chain Monte Carlo search algorithm. For the contemporaneous dependence graph G_0 an acyclic constraint is used to identify the causal directions in the system and to produce an identifiable SVAR model.

3.3 Quantile Networks

In the quantile approach the network is extracted by using a pairwise quantile regression [e.g., see **?**] at a given quantile level $q \in (0, 1)$,

$$\begin{cases} y_{it} = \varphi_{01} + \varphi_{11} y_{jt-1} + \varepsilon_{1t}, \\ y_{jt} = \varphi^*_{02} + \gamma^*_{21} y_{it} + \varphi^*_{21} y_{jt-1} + \varepsilon^*_{2t} \end{cases} \tag{4}$$

and

$$\begin{cases} y_{it} = \varphi^*_{01} + \gamma^*_{12} y_{jt} + \varphi^*_{11} y_{jt-1} + \varepsilon^*_{1t}, \\ y_{jt} = \varphi_{02} + \varphi_{21} y_{jt-1} + \varepsilon_{2t} \end{cases} \tag{5}$$

$\forall i, j = 1, \ldots, n$, $i \neq j$, where ε^*_{1t} and ε^*_{2t} are uncorrelated white noise processes.

- if $\gamma^*_{21} y_{it} \neq 0$ and $\gamma^*_{12} y_{jt} = 0$, y_{jt} is tail dependent from y_{it} and $a_{ij} = 1$, but not vice

versa, $a_{ji} = 0$;

- if $\gamma^*_{21} y_{it} = 0$ and $\gamma^*_{12} y_{jt} \neq 0$, y_{it} is tail dependent from y_{jt} and $a_{ji} = 1$, but not vice versa, $a_{ij} = 0$;

- if $\gamma^*_{21} y_{it} \neq 0$ and $\gamma^*_{12} y_{jt} \neq 0$, y_{it} is a tail mutual independence among y_{it} and y_{jt} and $a_{ji} = a_{ij} = 0$;

The relationship between the financial institutions using the quantile regression is asymmetric which implies the network extracted can be represented as a directed graph with an asymmetric adjacency matrix. The approach given above can be extended to include more lags and further covariates to control for spurious linkages.

4 Classical Network Measures

In this section, we present the commonly used measures in network analysis. See also **?** for a review. The structure and connectivity features of a network can be characterized by means of some measures. Node-specific measures are evaluated at the node level and reveal the role of a node in the connectivity structure of the network and its relationship with the other nodes. Local measures can be used to identify systemically important financial institutions. Global measures aim to describe the connectivity structure or topological features of the network, and therefore can be used to analyse the stability and fragility of the financial system.

4.1 Node-specific Measures (i.e local measures)

In the undirected network, the degree indicates the number of adjacent nodes, that is the number of nodes to which the node is connected. If a_{uv} is the u-th row and v-th column element of the adjacency matrix A, then the degree is equal to

$$d_u = \sum_{v=1}^{n} a_{uv} \tag{6}$$

In directed graphs, for a given node i it is useful to define the number of edges directed from othr nodes to node i (in-degree) and from node i to other nodes (out-degree), and

the total number of incident edges (total degree), that are

$$d_u^{out} = \sum_{v=1}^{n} a_{vu} \tag{7}$$

$$d_u^{in} = \sum_{v=1}^{n} a_{uv} \tag{8}$$

$$d_u^{tot} = d_u^{in} + d_u^{in} \tag{9}$$

The measures d_u^{out} and d_u^{in} are also known as in-degree and out-degree centrality measures and assess the centrality of a node in the network. They can be used to identify which are the nodes in the network spreading risk (spreaders) and which absorbing it (receivers).

While the degree measures how connected a node is, closeness account for connectivity patterns (such as paths and cycles) and indicates how easily a node can reach other nodes. The closeness centrality of a node u is defined as

$$\mathrm{Clo}_u(G) = \frac{n-1}{\sum_{v \in V, v \neq u} l(u,v)} \tag{10}$$

where $l(u,v)$ is the length of the shortest path between u and v.

A measure related to the closeness is the betweenness centrality which indicates how relevant is a node in terms of connecting other nodes in the graph. Let $n(u,v)$ be the number of shortest paths P_{uv}^* from u to v, and $n_w(u,v) = |\{P_{uv}^*; w \in P_{uv}^*\}|$, i.e. the number of shortest paths from u to v going through the node w, then the betweenness centrality

$$\mathrm{Bet}_w(G) = \sum_{u \neq v, w \notin \{u,v\}} \frac{n_w(u,v)/n(u,v)}{(n-1)(n-2)} \tag{11}$$

? introduced a measure of centrality for a undirected graph[1], called eigenvector centrality, which accounts for the centrality of the neighbourhood of a given node. This measure describe the influence of a node in a network and is defined as

$$\lambda x_u = \sum_{v \in N_u} a_{uv} x_v, \tag{12}$$

where the score x_u is related by the score of its neighbourhood $N_u = \{v \in V; a_{uv} = 1\}$. It is easy to show that the score vector $\mathbf{x} = (x_1, \ldots, x_n)'$ satisfies the equation $A\mathbf{x} = \lambda\mathbf{x}$,

[1]The concept of the eigenvector centrality could be applied to the directed network, but since the adjacency matrix is not symmetric the right and the left eigenvector are different.

where λ is an eigenvalue of the matrix A. Eigenvector centrality explains the propagation of economic shocks better than other measure as closeness and betweenness centrality since it accounts not only for the number of connections of each node with the adjacent nodes, but also for its weight and for the weights of the paths connecting the node to the other nodes of the graph.

? introduced a related centrality measure, the $c(\beta)$ centrality, defined as

$$c(\beta) = \sum_{j=1}^{\infty} \beta^{j-1} A^j \boldsymbol{\iota}$$

with $|\beta| < 1/\lambda_1$ where longer paths are weighted less through higher powers of the discount parameter β. $\boldsymbol{\iota}$ is the n-dimensional unit vector and λ_1 is the largest eigenvalue. This measure is a weighted sum over all possible paths connecting other vertices to each position. $c(\beta)$ centrality has the eigenvector centrality as limiting case for $\beta \to 1/\lambda_1$.

The $c(\beta)$ centrality is strictly related to another widely used measure, that is the Katz centrality. **?** proposed a node centrality measure which is a weighted sum of the walks of a given node neighbours, with weights driven by an attenuation parameter. The Katz centrality of the node u is defined as

$$x_u = \beta \sum_{v=1}^{n} a_{uv} x_v + \alpha \tag{13}$$

where $0 < \beta < 1$ is the attenuation parameter and α is an arbitrary term that avoids to consider in the centrality score all vertices with null degree. The parameter α is usually set to 1. Equation **??** can be written in matrix form: $\mathbf{x} = (I_n - \beta A)^{-1} \alpha \boldsymbol{\iota} = (I_n + \beta A + \beta^2 A^2 + \ldots + \beta^k A^k + \ldots) \alpha \boldsymbol{\iota}$ which shows how neighbours, and neighbours of neighbours affect the nodes centrality. Finally, note that the Katz centrality has as special cases the $c(\beta)$ centrality for $\alpha = 1$ and the Bonacich's centrality for $\alpha = 0$ and $\beta = 1/\lambda_1$.

4.2 Global Measures

The density of a n-order graph is given by the ratio between the number of edges in the edge set E, denoted with $e(G)$, and the number of edges of the complete graph K_n. If the K_n graph is undirected the cardinality of the edge set is $e(K_n) = n(n-1)/2$, if it is

directed the cardinality is $e(K_n) = n(n-1)$. Thus the graph density is

$$\mathrm{De} = \frac{e(G)}{K_n} \tag{14}$$

The density is null if G is the empty graph E_n and it is equal to one if G is the complete graph K_n. Density is a good indicator of the level of interconnectedness in a financial network. Nevertheless this measure relies on the adjacency of the nodes and does not consider indirect connectivity patterns such as paths and cycles.

A way to account for connectivity patterns is to analyse the cliquishness of a graph. Unfortunately, the clique structure can be very sensitive to slight changes in the graph and thus, a general procedure for extracting cliques can fail in finding the clique structure. A common way to measure cliquishness is to employ the average clustering coefficient of a n-order graph $G = (V, E)$, $\mathrm{Cl}(G)$, which counts the fraction of fully connected triples of nodes out of the potential triples where at least two links are presents. In formulas, we have

$$\mathrm{Cl}(G) = \frac{1}{n}\sum_{u=1}^{n} Cl_u(G) \tag{15}$$

where

$$\mathrm{Cl}_u(G) = \frac{|\{(u', v') \in E; u' \neq v', u' \in N_u(G), v' \in N_u(G)\}|}{|\{(u', v') \in K_n; u' \neq v', u' \in N_u(K_n), v' \in N_u(K_n)\}|}$$

where $N_u(G)$ indicates the neighbourhood of a node u in G, and so if G is the complete graph K_n, $N_u(K_n)$ is of the set all the nodes different from u.

Assortativity can be defined as the difference between the number of edges among vertexes having the same characteristics and therefore belonging to the same class, and the expected number of edges among these vertexes if the attachment were purely random [**??**]. Let m_i be the class of the vertex i, and n_m the number of classes in the network. In a directed network, the number of edges among the vertexes of the same class is

$$\sum_{u,v \in V} a_{uv}\delta_{m_u m_v}, \tag{16}$$

where δ_{xy} is the Kronecker delta.[2] Assuming a random graph, the expected number of

[2]The Kronecker delta δ_{xy} is a function of two variables which is equal to 1 if they are equal and 0 otherwise.

edges among vertices of the same class is equal to

$$\sum_{u,v \in V} \frac{d_u^{out} d_v^{in}}{e(G)} \delta_{m_u m_v}, \tag{17}$$

where d_u^{out} and d_v^{in} denote the nodes out- and in-degree. Thus, the assortativity measure of the graph $G = (V, E)$ can be defined as

$$Q = \frac{1}{e(G)} \left(e(G) - \frac{d_u^{out} d_v^{in}}{e(G)} \delta_{m_u m_v} \right). \tag{18}$$

The maximum assortativity value $Q_{\max}$ is attained when all edges of E are adjacent to all vertices of the same category, so

$$Q_{\max} = \frac{1}{e(G)} \left(e(G) - \sum_{u,v \in V} \frac{d_u^{out} d_v^{in}}{e(G)} \delta_{m_u m_v} \right). \tag{19}$$

If $0 < Q \leq Q_{\max}$ the nodes have an homophily behaviour and if $Q/Q_{\max}$ is close to zero there is any preferential attachment of the nodes and the network is a random graph. If $-Q_{\max} \leq Q < 0$ the nodes exhibit disassortative patterns, in sense that it is likely that a given node is connected to nodes in a different class.

The assortativity measure is very important in systemic risk analysis because it allows to detect clusters in the network that can be useful to block the shocks propagation in some circumstances as a firewall. The assortativity can be also applied to the vertex degree in order to capture the tendency of each node to connect with vertices having similar or different degree. This measure is able to detect a core-periphery structure of the graph when the assortativity coefficient is high. Let e_{jk} be the edges fraction connecting vertices of degree j to vertexes of degree k, q_j^{out} and q_j^{in} the probabilities to have an excess in-degree, out-degree,[3] respectively, equal to j and σ_q^{out} and σ_q^{in} standard deviation of the degree distributions of q_j^{out} and q_j^{in}, respectively. Then, following **?**, the assortativity by degree for directed graphs is defined as

$$r = \frac{\sum_{jk} jk(e_{jk} - q_j^{in} q_k^{out})}{\sigma_{in} \sigma_{out}}. \tag{20}$$

Assortativity can be similarliy defined for undirected graph [see **?**].

[3] As calculated in Equation **??** above.

5 Entropy Measures

5.1 Von Neumann Entropy for Directed Graph

Following **?**, we introduce a Markov chain process on the graph with transition matrix P with entries

$$p_{uv} = \begin{cases} \frac{1}{d_u^{out}} a_{uv} & \text{if } \{u,v\} \in E \\ 0 & \text{otherwise.} \end{cases}$$

Let $D = \text{diag}\{(d_1^{out}, d_2^{out}, \ldots, d_n^{out})\}$ then P can be in written in the matrix form: $P = D^{-1}A$. Additionally, given the vector $\boldsymbol{\varphi}$ of the ergodic probability of the Markov chain associated to P,[4] we define the diagonal matrix $\varphi = \text{diag}\{\boldsymbol{\varphi}\}$ and the Laplacian [**?**] for undirected graphs:

$$L_1 = I_n - \frac{\varphi^{1/2} P \varphi^{-1/2} + \varphi^{-1/2} P' \varphi^{1/2}}{2} \tag{21}$$

$$L_2 = \varphi - \frac{\varphi P + P' \varphi}{2} \tag{22}$$

where I_n is the identity matrix. L_1 can be related to a random walk onto the graph starting at a given node chosen with uniform probability while L_2 to a random walk starting at a node chosen according to the ergodic probability vector φ. In fact, those Laplacians can be interpreted as the Laplacians for an equivalent weighted undirected graph obtained by changing the weights of the starting graph but not its connectivity [**?**]. The difference among the two is in the equivalent ergodic probabilities that are uniform in the L_1 and again equal to φ for L_2. In this line, L_1 is largely driven by long-run effects with respect to L_2.

The induced L_i by the directed graph is symmetric but not positive definite and consequently, it is not suitable to be used as a proper density matrix [**?**]. Using the results in **?** and **?**, we obtain a density matrix ρ, based on the Laplacian but corrected for directed graphs. A proper density matrix, ρ, is a symmetric positive definite matrix with unitary trace. The functional form of the density matrix describes the correspondence of the graph to a given quantum system. In **?** for undirected graphs, the density matrix is defined as the Laplacian normalized by its trace while **?** generalizes the same construction

[4] $\boldsymbol{\varphi}$ satisfies the equation: $\boldsymbol{\varphi} P = \lambda \boldsymbol{\varphi}$, where λ is the associated eigenvalue and $\sum_{i=1}^{n} \varphi_i = 1$.

for directed graphs,

$$\rho_{Li} = \frac{L_i}{\text{trace}\left(L_i\right)}. \tag{23}$$

Formally, any positivity preserving transformation of the Laplacian could provide a proper density matrix ρ. In this regard, we consider the exponential function as an alternative transformation,

$$\rho_{Ei} = \frac{\exp\left(L_i\right)}{\text{trace}\left(\exp\left(L_i\right)\right)}. \tag{24}$$

The linear transformation takes into consideration only the 1-step walk probability which can be viewed as the short run effect of propagation in the network. Differently, the exponential considers all the possible walk probabilities which are weighted by the inverse of the factorial of the walk length. Thus, it can be interpreted as the long run effect of propagation. According to the random walk interpretation of the two Laplacians, we do not expect large differences among $\rho(L_1)$ and $\rho(E_1)$. Given the density matrix ρ, we can measure the complexity of the network using the Von Neumann entropy,

$$S\left(\rho\right) = -\text{trace}\left(\rho\log\left(\rho\right)\right), \tag{25}$$

which is equivalent to the Shannon entropy of the eigenvalues of ρ,

$$S\left(\rho\right) = -\sum_{i=1}^{n}\lambda_i^{\rho}\log\left(\lambda_i^{\rho}\right). \tag{26}$$

and is bounded [**?**],

$$S\left(\rho\right) \leq -\log\left(n\right). \tag{27}$$

For undirected graphs and ρ_{Li} the maximum entropy is associated with the complete graph [**?**]. For directed graphs, linear density matrix and the Laplacians L_1, **?** show that the maximum value of entropy is associated with the star graph according to the quadratic approximation.

5.2 A Von Neumann Entropy Decomposition

We present some preliminary results which will be used to state some properties of the Von Neumann entropy proposed in this paper. Let L be one the two Laplacian matrices

L_i, $i = 1, 2$, then the following results hold.

Theorem 1 *The following property holds for the Laplacian L*

- *For* $L = L_1$

$$tr(L^m) = n + \sum_{k=0}^{m-1} \sum_{l=0}^{m-k} \omega_{lk} tr(R^{(lk)}) \tag{28}$$

$m = 0, 1, 2, \ldots,$ *where*

$$\omega_{lk} = \binom{m}{k} (-2)^{m-k} \binom{m-k}{l}$$

are weights and $R^{(lk)}$ *is a matrix with the* (i, j)*-th element*

$$R_{ij}^{(lk)} = \sum_{h=1}^{n} W_{ih}^{(l)} W_{jh}^{(m-k-l)} \frac{\varphi_i}{\varphi_h}$$

with

$$W_{ij}^{(q)} = \sum_{i_2 \cdots i_q} \prod_{r=1}^{q} a_{i_r\, i_{r+1}} \frac{1}{d_{i_r}^{in}}$$

where $i_1 = i$ *and* $i_{q+1} = j$, $q = 0, 1, \ldots, n-1$ *the length of the walk and* $d_{i_r}^{in}$ *is defined in Equation* **??**.

- *For* $L = L_2$

$$tr(L^m) = \sum_{i=1}^{n} \varphi_i^m + \sum_{k=0}^{m-1} \omega_k tr(R^{(k)}) \tag{29}$$

$m = 0, 1, 2, \ldots,$ *where*

$$\omega_k = \binom{m}{k} (-2)^{-(m-k)}$$

are weights, $R^{(k)} = \varphi^k W^{(m-k)}$, *and* $W^{(q)}$ *is a matrix with the* (i, j)*-th element*

$$W_{ij}^{(q)} = \sum_{i_2 \cdots i_q} \prod_{r=1}^{q} \left(\frac{\varphi_{i_r}}{d_{i_r}^{-}} a_{i_r i_{r+1}} + a_{i_{r+1} i_r} \frac{1}{\varphi_{i_{r+1}} d_{i_{r+1}}^{-}} \right)$$

with $i_1 = i$ *and* $i_{q+1} = j$.

Proof. See Appendix **??**.

In the theorem given above the matrices $W^{(q)}$ have an interpretation in terms of path of random walk on a network. More specifically, for $L = L_1$ $W_{ij}^{(q)}$ is the transition q steps forward in time starting in i at time 1 and arriving in j at time $(q + 1)$. Whereas for $L = L_2$ $W_{ij}^{(q)}$ is the transition q steps forward and q steps backward in time, starting in i at time 0 and arriving in j at time $2q$.

We show that the Von Neumann entropy based on the transformed Laplacian accounts for various features of the associated graph. We focus on the numerator of Eq. **??**

$$\exp\{L_i\} = \sum_{m=0}^{\infty} \frac{1}{m!} L_i^m \tag{30}$$

where $m!$ is factorial. It can be approximated by

$$\tilde{L}_i^{(M)} = \sum_{m=0}^{M} \frac{1}{m!} L_i^m \tag{31}$$

with $M < \infty$. Then, the Von Neumann entropy S given in equation **??** can be approximated as

$$S_i^{(M)}(\boldsymbol{\lambda}) = \frac{1}{n}\mathrm{tr}(L_i^{(M)}) - \frac{1}{n^2}\mathrm{tr}((L_i^{(M)})^2) \tag{32}$$

Proof. See Appendix **??**.

Theorem 2 *The approximated quadratic Von Neumann entropy for various degrees of approximation can be written as*

- *For* $M = 1$

$$\begin{aligned}
&S^{(M)}(\boldsymbol{\lambda}_1) = 2 - \frac{3}{n} - \frac{1}{4n^2} tr((P')^2) - \frac{2}{n^2} tr(\varphi P \varphi^{-1} P') - \frac{1}{n^2} tr(P^2)) \\
&S^{(M)}(\boldsymbol{\lambda}_2) = 1 - \frac{2}{n^2} + \left(\frac{1}{n} - \frac{2}{n^2}\right) \left(tr(P(\varphi + \varphi^{-1})/2)\right) \\
&-\frac{1}{n^2}\left(\sum_{i=1}^{n} \varphi_i^2 + \frac{1}{4}(tr((\varphi P)^2) + tr((P'\varphi^{-1})^2) + tr(P'P) + tr(PP')) - tr(\varphi^2 P)\right)
\end{aligned}$$

- *For* $M = 2$

$$\begin{aligned}
S^{(M)}(\boldsymbol{\lambda}_1) &= 2 - \frac{25}{4n} - \frac{1}{n^2}\left(\sum_{k=0}^{3}\sum_{l=0}^{4-k} \tilde{\omega}_{lk} tr(R^{(lk)})\right) \\
S^{(M)}(\boldsymbol{\lambda}_2) &= 2 - \frac{3}{n} - \frac{1}{n^2}\sum_{k=0}^{3} \tilde{\omega}_k tr(R^{(k)})
\end{aligned}$$

where $R^{(lk)}$ *and* $R^{(k)}$ *have been defined in Theorem* **??**.

Note that our quadratic entropy for the case $M = 1$ has an analytical relationship with the quadratic entropy $S_Q = n^{-1}\mathrm{tr}(L_1) - n^{-2}\mathrm{tr}(L_1)$ of **?** as stated in the following.

Corollary 1

$$S_1 - S_Q = 1 - \frac{3}{n} \tag{33}$$

Proof. See Appendix **??**.

6 Diplacian and convergence rate to distributed consensus

In the computation of the Von Neumman entropy we considered a symmetrized version of the Laplacian matrix for directed graphs. This choice, as pointed out in two recent papers, could hamper the comprehension of several important aspect related to diffusion on the graph [**??**]. In particular, we investigate the relationship between the eigenvalues of the diplacian introduced in **?**.
The Diplacian is defined as

$$\Gamma = \varphi^{1/2} (I - P) \varphi^{-1/2}$$

and the rate of convergence of autonomous agents on the network to a consensus [**?**]. The application of those techniques to financial networks could be understood as measuring the structural speed of coordination so that a low rate implies persistence of disagreement in the market that in line with **?** is "magnified when major events occur in financial markets". Our approach considers a limited communication network [see **?**] among agents, proxied by the causality relationship between those stocks returns. The convergence to a final group decision is then a convergence to the consensus of investors, trading in different stocks, on common Arrow-Debreu securities prices. Consequently, we propose the persistence of disagreement as a general proxy for the presence of market frictions. Consider a graph with adjacency matrix A and elements a_{ij} and with out-degree diagonal matrix D with non zero elements d_i^{out}. We investigate the following discrete time multi-agent dynamical system on the network:

$$x_{it} = x_{it} + \frac{1}{2d_i^{out}} \sum_{j=1}^{n} a_{ij} (x_{jt} - x_{it}) \tag{34}$$

Similar systems, pioneered by **?**, are considered in models of belief evolution of bounded rational agents, with the bounded rationality motivated by a persuasion bias [**??**]. The closer analogue to our approach is the one considered in the Theorem 2 of **?** and can be

rewritten in vectorial form as

$$\mathbf{x}_t = \frac{1}{2}I_n + \frac{1}{2}\left(I_n - D^{-1}\left(D - A\right)\right)\mathbf{x}_{t-1} = P_L X_{t-1}$$

Where $\mathbf{x}_t = (x_{it}, \ldots, x_{nt})$ is the state vector of the agents $P = D^{-1}A$ is the transition probability matrix of the Markov chain associated with random walks on G, where at each vertex i, a random walk has probability $p_{ij} = a_{ij}/d_i^{out}$ of transiting from vertex i to vertex j and $P_L = (I_n - P)$ corresponds to the transition matrix of the lazy random walk introduced in **?** and further studied in **?**. Using their results, if the graph is strongly connected, P_L is irreducible and aperiodic, so that Perron -Frobenious theorem applies and we can easily adapt Theorem 2 of **?**, being sure that the system converge to a consensus with group decision value $\boldsymbol{\varphi}'\mathbf{x}_0$. The group decision is a conserved quantity of the dynamics:

$$\boldsymbol{\varphi}'\mathbf{x}_t = \boldsymbol{\varphi}' P_L \mathbf{x}_{t-1} \boldsymbol{\varphi}'\mathbf{x}_{t-1} = \alpha,$$
$$\alpha \in \mathbb{R}.$$

Consequently, we can define a disagreement vector and its dynamics

$$\boldsymbol{\xi}_t = \mathbf{x}_t - \alpha\mathbf{1}$$
$$\boldsymbol{\xi}_t = P_L \boldsymbol{\xi}_{t-1}$$

The disagreement dynamics allows us to study speed of convergence to this decision value. We exploit the theoretical results on lazy random walks on strongly connected directed graphs due to **?** and **?**. In particular in **?** the decomposition of the Diplacian Γ in its symmetric and asymmetric part is introduced

$$\Gamma = L + \Delta,$$
$$L = \frac{\Gamma + \Gamma'}{2} \tag{35}$$
$$\Delta = \frac{\Gamma - \Gamma'}{2} \tag{36}$$

In the following theorem the convergence rate is expressed in terms of λ_2 the second smallest eigenvalue of L and of the second largest singular value $\sigma_{n-1}\left(I_n - L\right)$ of $I_n - L$ and the largest singular value $\sigma_n\left(\Delta\right)$ of the skew-symmetric part of the diplacian Δ.

Theorem 3 *Consider the discrete-time system introduced in (**??**) on a a strongly con-*

nected directed network. A consensus is globally exponentially reached according to

$$\begin{aligned} \|\boldsymbol{\xi_t}\| &\leq \exp\left\{\frac{1}{2}\left[\log\left(\frac{\max(\boldsymbol{\varphi})}{\min(\boldsymbol{\varphi})}\right)+\log(\mu)\,t\right]\right\}\|\boldsymbol{\xi_0}\| \\ \mu &= \frac{3}{4}-\frac{\lambda_2}{2}+\frac{\left(\sigma_{n-1}(I_n-L)+\sigma_n(\Delta)\right)^2}{4}. \end{aligned}$$

where μ is the disagreement persistence index, measuring the convergence rate to consensus.

Proof. See Appendix **??**.

? show that for symmetric adjacency matrices, i.e. for undirected graphs,$\sigma_n(\Delta)=0$ and they propose it as a measure of asymmetry (directedness). Moreover they underline that in this case

$$\mu = 1-\frac{\lambda_2}{2}$$

a bound previously derived in **?**.

The expression in Theorem **??** implies a slower convergence if the graph is directed and shows an initial magnifing effect of the heterogeneity of importance of the nodes in the group decision. The latter effect is not present for directed balanced graphs i.e. for directed graphs that have row and column sum equal, because in that case the group decision is the average of the initial state vector. Going further,if we assume an initial disagreement vector of unitary norm we can evaluate the time needed to reach consensus. In fact, for each $\epsilon>0$, we have $\|\boldsymbol{\xi_t}\|\leq\epsilon$ in a time

$$t \leq \frac{2\log(\epsilon)-\log\left(\frac{\max(\boldsymbol{\varphi})}{\min(\boldsymbol{\varphi})}\right)}{\log\mu}$$

In the empirical result section we will evaluate μ on the giant strongly connected component and multiply it by a weight proportional to the size of the component in order to take into consideration the impact of the number of coordinated agents.

7 Empirical Analysis

7.1 Data Description

The dataset is composed by the daily closing price series for the European financial institutions (active and dead) from 29^{th} December 1999 to 16^{th} January 2013. We analyse

a total of 437 European financial institutions according to the Industrial Classification Benchmark (ICB). We select the MSCI Europe index as the proxy for the market index, which covers the 15 European countries where the financial institutions are based.
To estimate dynamic Granger networks, we use a rolling window approach [e.g., see **???**] with a window size of 252 daily observations, that is approximatively one year.[5]

The sequence of adjacency matrices of the directed graph extracted with the pairwise Granger approach is represented in Fig. **??** (a weekly sampling frequency has been used for expository purposes). Red boxes highlight the adjacency matrix at a given date. In each box blue dots represent directed edges between nodes.

The question here is to compare the network defined over time, to identify possible distortions and sources of systemic risk. Figure **??** shows the graphs of four financial networks extracted in January 2001 (Panel a), August 2004 (Panel b), October 2008 (Panel c), and March 2012 (Panel d). The size of the nodes is proportional to the node degree. For expository purposes nodes with a degree lower then a given threshold have been removed (see caption of Figure **??**). From a visual inspection, the graphs in panel (a) and (b) exhibit a lower number of edges then the graphs in panel (b) and (d), which reflects a higher level of financial interconnectedness. The node color from black to red reflects the node eigenvector centrality. Red indicates a large centrality level, black a low centrality. A more precise description of the four networks can be achieved by applying the network measures discussed in the previous section.

Table **??** shows the network measures for the four networks depicted in Figure **??**. Panels (a) and (b) refer to two tranquil periods whereas (c) and (d) to instability periods due to the global financial and European sovereign debt crises

A comparison of statistics in panels (a) and (b) with the one in panels (c) and (d) reveals that the level of connectivity changes during periods of instability. More specifically, the number of direct connections increases, since the average degree and density for (a) and (b) are lower than (c) and (d). Also the number of indirect connections increases, since the network diameter, that is the longer shortest path, is 4 in (a) and (b) and 3 in (c) and (d). The decrease in the diameter indicates that a financial shock propagates faster in (c) and (d) than in (a) and (b). This is supported also by the average path length (APL)

[5]The rolling-window estimations have been parallelized and implemented in Matlab. It takes approximately 72 hours on the SCSCF (Sistema di Calcolo Scientifico Ca' Foscari) cluster multiprocessor system which consists of 4 nodes; each comprises four Xeon E5-4610 v2 2.3GHz CPUs, with 8 cores, 256GB ECC PC3-12800R RAM, Ethernet 10Gbit, 20TB hard disk system with Linux.

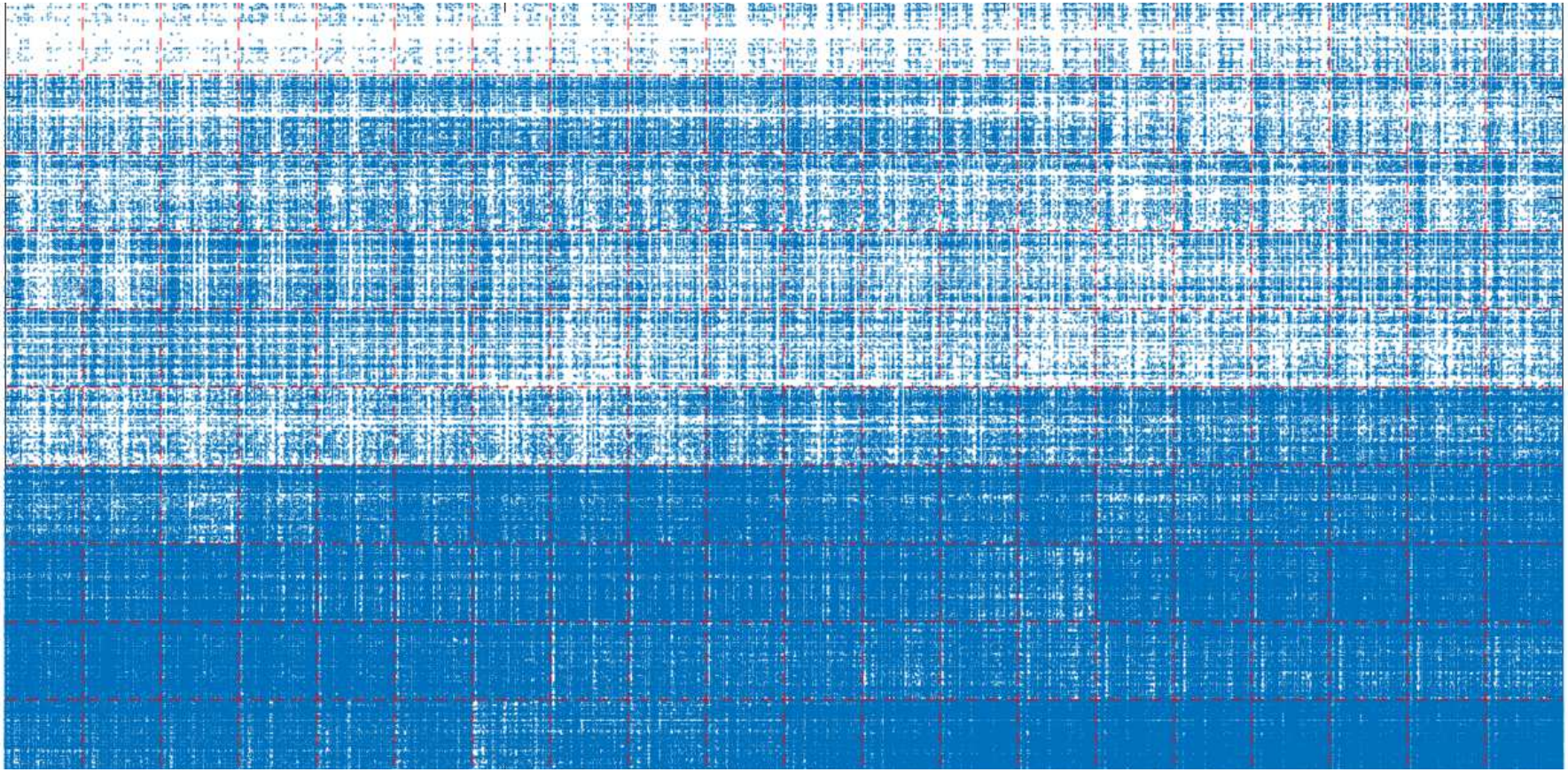


Figure 3: Sequence of adjacency matrices from 29^{th} December 1999 to 16^{th} January 2013 (re-sampled at weekly frequency for expository purposes). Red boxes highlight the adjacency matrix at a given week. In each box blue dots represent directed edges between nodes.

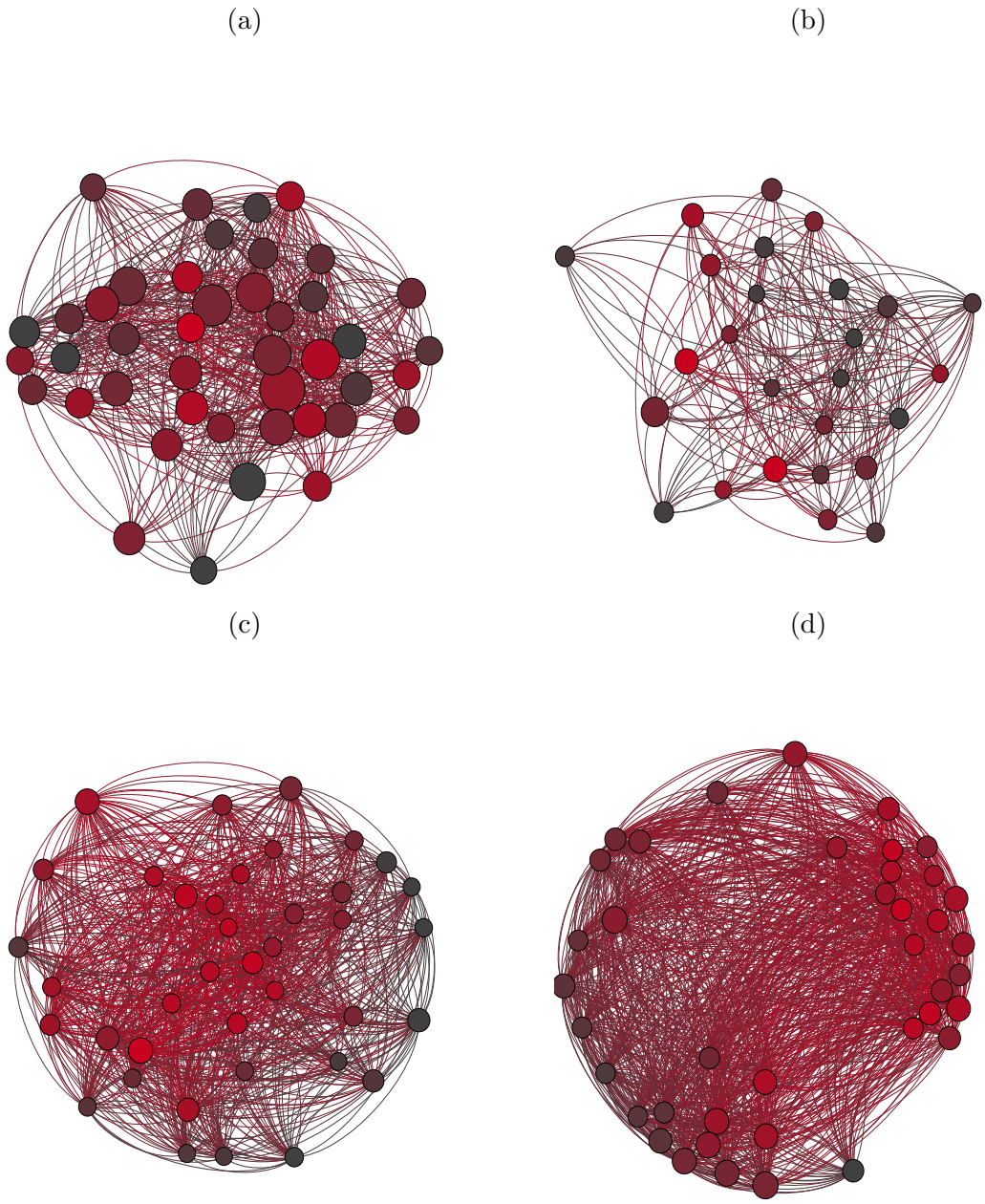


Figure 4: Financial networks extracted in January 2001 (a), August 2004 (b), October 2008 (c), and March 2012 (d). The size of the nodes are proportional to the node degree. For expository purposes, nodes with a degree lower than 70 (a), 90 (b), 300 (c) and 350 (d) have been removed. The difference in threshold is due to different network density. The node color from black to red reflects the node eigenvector centrality. Red indicates a large centrality level, black a low centrality.

	AD	AID	De	Di	Ra	APL
Panel (a)	55.065	27.53	0.149	4	3	2.001
Panel (b)	47.517	23.758	0.161	4	3	2.008
Panel (c)	213.935	106.967	0.291	3	2	1.731
Panel (d)	257.636	128.812	0.356	3	2	1.652

	WCC	SCC	ACluCoe	ACloCen	ABetCen	AEigCen
Panel (a)	1	1	0.181	0.503	185.134	0.332
Panel (b)	1	1	0.232	0.505	149.196	0.236
Panel (c)	1	2	0.377	0.586	267.54	0.272
Panel (d)	1	1	0.417	0.6145	236	0.3077

	$S(\rho_{L_1})$	$S(\rho_{L_2})$	$S(\rho_{E_1})$	$S(\rho_{E_2})$	μ
Panel (a)	-50.4660	-720.6452	-10.8064	-0.0065	0.3341
Panel (b)	-49.9492	-1238.3541	-10.3041	-0.0223	0.2450
Panel (c)	-16.0168	-1421.1440	-3.2444	-0.0041	0.5296
Panel (d)	-11.6902	-673.3415	-2.1504	-0.0014	0.4500

Table 1: Network statistics for the four financial networks in Figure **??** extracted in January 2001 (Panel a), August 2004 (Panel b), October 2008 (Panel c), and March 2012 (Panel d). Average degree (AD), average in-degree (AID), density (De), diameter (Di), radius (Ra), average path length (APL), number of weakly and strongly connected components (WCC and SCC), average clustering coefficient (ACluCoe), average closeness centrality (ACloCen), average betweenness centrality (ABetCen), average eigenvector centrality (AEigCen), the Von Neumann entropies for the linear, i.e. $S(\rho_{L_1})$ and $S(\rho_{L_2})$, and exponential, i.e. $S(\rho_{E_1})$ and $S(\rho_{E_2})$, functional forms using the two Laplacians L_1 and L_2 and the disagreement persistence index (μ).

which is larger in (a) and (b) with respect to (c) and (d) (see Table **??**. The same applies to the average cluster coefficient (ACluCoe), average closeness centrality (ACloCen) and average betweenness centrality (ABetCen). Finally, the increase of the (log) Von Neumann entropies and the disagreement persistence index (μ) in (c) and (d) with respect than to (a) and (b) indicates a more complex topology of the network.

Figure **??** reports the Von Neumann entropies evaluated over the sequence of inferred networks of Figure **??**. In the different panels, we have the Von Neumann entropies for the linear, i.e. $S(\rho_{L_1})$ and $S(\rho_{L_2})$, and exponential, i.e. $S(\rho_{E_1})$ and $S(\rho_{E_2})$, functional forms using the two Laplacians L_1 and L_2. Given the first Laplacian, L_1, the dynamic of the linear $S(\rho_{L_1})$ (top-left panel) and the exponential $S(\rho_{E_1})$ (bottom-left panel) are almost identical. The second Laplacian provides different dynamics in the linear $S(\rho_{L_2})$ entropy and the exponential one $S(\rho_{E_2})$. While the Von Neumann entropies depict a structural change in the European financial system starting from 2006, the disagreement persistence index (μ) reported Figure **??** seems to capture some well-known events which provoked turbulences and stress in the financial system. In fact, the indicator signals a first level of stress in the mid 2006 when the U.S market price reaches its peak and there are some tensions in the financial market due to rising of the interest rates. The second signal is located at the beginning of the Global financial crisis in August 2007 when BPN Paribas suspends three funds due to the subprime mortgage sector large decline and the inability by the fund managers to determine their market value. The level of the indicator remains at its higher levels during the crisis and slowly declines after the peak in the correspondence of the Bankruptcy of Lehman Brother on the 15th September 2008. The beginning of the European sovereign debt crisis is captured by the a new rise of the disagreement persistence index occurred in April 2010 with the Greece government debt crisis and the downgrade of Greek bonds to junk status. The persistence of the high level of the indicator culminated with the peak in August 2011 in the correspondence of the Euro market fall due the contagion to the peripheral countries such as Italy and Spain. Then, the index decreases with some short spikes before and immediately after the ECB's asset purchase programm.

7.2 Early Warning Indicators

As in **?**, we observe a persistence in the dynamics of entropy and thus we test the ability of the Von Neumann entropy and the disagreement persistence index as early warning

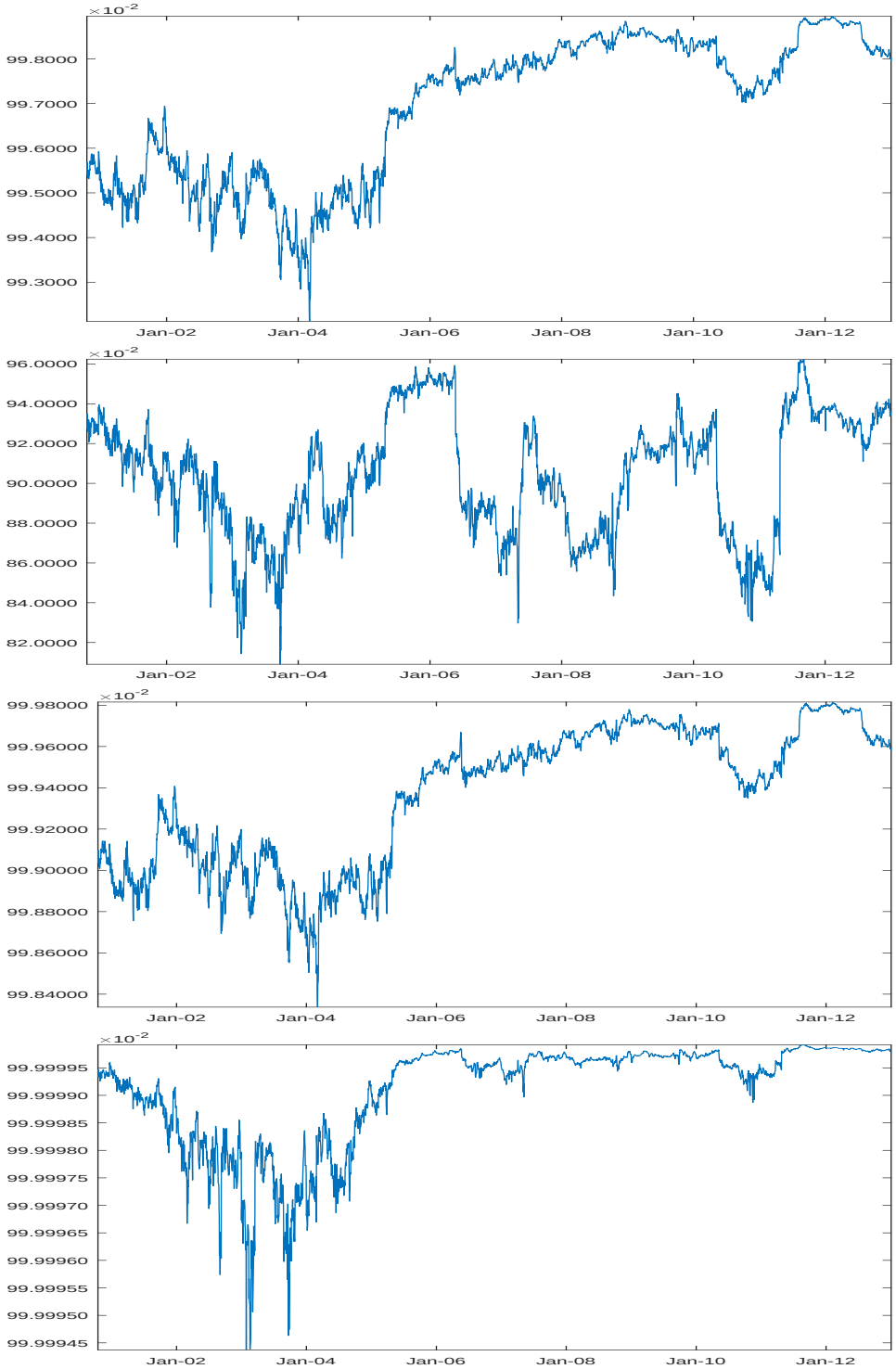


Figure 5: Von Neumann entropies of the Granger network of European financial institutions during 2000-2014. First panel: $S(\rho_{L_1})$ indicates the Von Neumann entropy with the linear functional form and Laplacian L_1. Second panel: $S(\rho_{L_2})$ indicates the Von Neumann entropy with the linear functional form and Laplacian L_2. Third panel: $S(\rho_{E_1})$ indicates the Von Neumann entropy with the exponential functional form and Laplacian L_1. Fourth panel: $S(\rho_{E_2})$ indicates the Von Neumann entropy with the exponential functional form and Laplacian L_2.

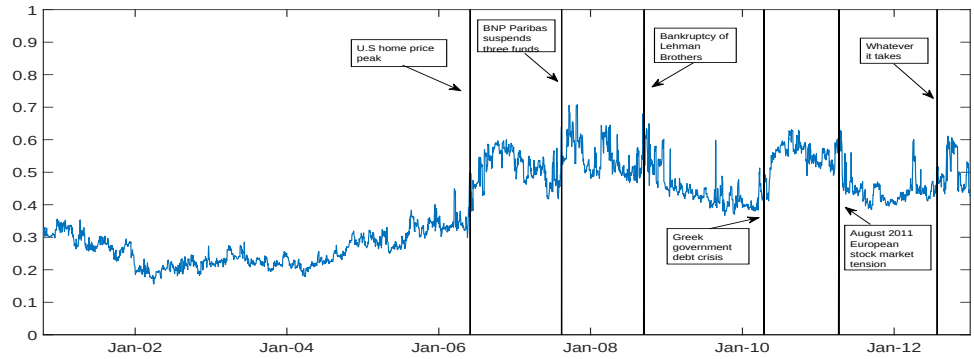


Figure 6: Disagreement persistence index (μ) of the Granger Network of European financial institutions during 2000-2014.

indicators in nowcasting banking crises. An early warning system issues a signal in case the likelihood of a crisis crosses a specified threshold [**?**]. In this regard, we select as the European banking crisis indicator the one presented in **?** and **?** which represents one of the target variables monitored by European Systemic Risk Board (ESRB). The indicator identifies significant signs of financial distress in the banking system as evidenced bank runs in relevant institutions or losses in the banking system (nonperforming loans above 20% or bank closures of at least 20% of banking system assets), or significant public intervention in response to or to avoid the realisation of losses in the banking system. As stated in **?**, we define an indicator on European basis to be used in the early warning system since the crisis indicator is given on a per-country basis,

$$C_t = \begin{cases} 1 & \text{if more than one country is in crisis at time } t \\ 0 & \text{otherwise.} \end{cases} \tag{37}$$

The banking crisis indicator in **?** has its last record in December 2012 and thus we consider the period from is from October 2000 to December 2012. Since the crisis indicator is at a quarterly frequency and networks are extracted from daily returns, we assume that the daily crisis indicator is equal to 1 for all days in a given quarter, if the quarterly indicator equals 1 for that quarter [see **?**]. Consequently, we make use of the logistic model with Von Neumann entropies and the disagreement persistence index as covariates. As comparison, we include the DCI [**?**], that is the density of the network (De) defined in Equation **??**, and the Shannon entropy (H) of the degree distribution as proposed in **?**.

Hence, we set the following logistic regression model,

$$\mathbb{P}(C_t = 1|S_t) = \Phi(\beta_0 + \beta_1 S_t), \tag{38}$$

$t = 1, \ldots, T$, where $\Phi(x) = 1/(1 + \exp(-x))$ is the logistic function and S_t is an early warning indicator. The estimation results from the logit specification are presented in Table **??**. All indicators are significant at 1% confidence-level. All the Von Neumann entropies (except for the one with the linear functional form and the second Laplacian) and the disagreement persistence index provide better estimates in terms of Adjusted-R-squared, AIC and BIC criteria with respect to the DCI and Shannon Entropy. In particular, the Von Neumann entropy with the exponential functional form and Laplacian L_1 shows the best fitting with an adjusted-R-squared of 0.34. The corresponding estimated response variables are reported in Figure **??**. Von Neumann entropies with exponential functional form together Laplacian L_1 (solid line) and with linear functional form together with Laplacian L_2 (dashed line), $S(\rho_{E1})$ and $S(\rho_{L2})$, respectively, provide the best fit with the actual crises indicator (stepwise red line) defined in Equation **??**. It is worth noting that the estimated response variable in all the logit model specifications returns a high probability level of crisis in correspondence of the European sovereign debt crisis after a recovery from the Global financial crisis in 2007 and 2008.

	DCI	H	$S(\rho_{L1})$	$S(\rho_{L2})$	$S(\rho_{E1})$	$S(\rho_{E2})$	μ
β_0	0.54^{***}	-22.35^{***}	-1009^{***}	-3.11^{***}	-4717^{***}	-2838741^{***}	4.25^{***}
	(0.13)	(1.29)	(36.78)	(1.07)	(173.74)	(134373)	(0.16)
β_1	-4.10^{***}	25.16^{***}	1012^{***}	3.22^{***}	4719^{***}	2838742^{***}	10.25^{***}
	(0.70)	(1.47)	(36.87)	(1.19)	(173.83)	(134373)	(0.37)
Adj-R^2	0.01	0.10	0.33	0.00	0.32	0.34	0.28
AIC	4356	4054	3149	4383	3208	3068	3360
BIC	4368	4066	3162	4395	3220	3080	3372

Table 2: Logit specification where the dependent variable is the banking crisis indicator from **?** and the explanatory variables are from the European Granger Network (columns) : 1) Dynamic causality index (DCI) as proposed in **?**; 2) Shannon entropy on In-Out degree (H) as in **?**; 3) $S(\rho_{L_1})$ indicates the Von Neumann entropy with the linear functional form and Laplacian L_1; 4) $S(\rho_{L_2})$ indicates the Von Neumann entropy with the linear functional form and Laplacian L_2; 5) $S(\rho_{E_1})$ indicates the Von Neumann entropy with the exponential functional form and Laplacian L_1; 6) $S(\rho_{E_2})$ indicates the Von Neumann entropy with the exponential functional form and Laplacian L_2 and 7) μ indicates the disagreement persistence. Significance level: 1% (***). Standard errors in parentheses. The number of observations are 3187.

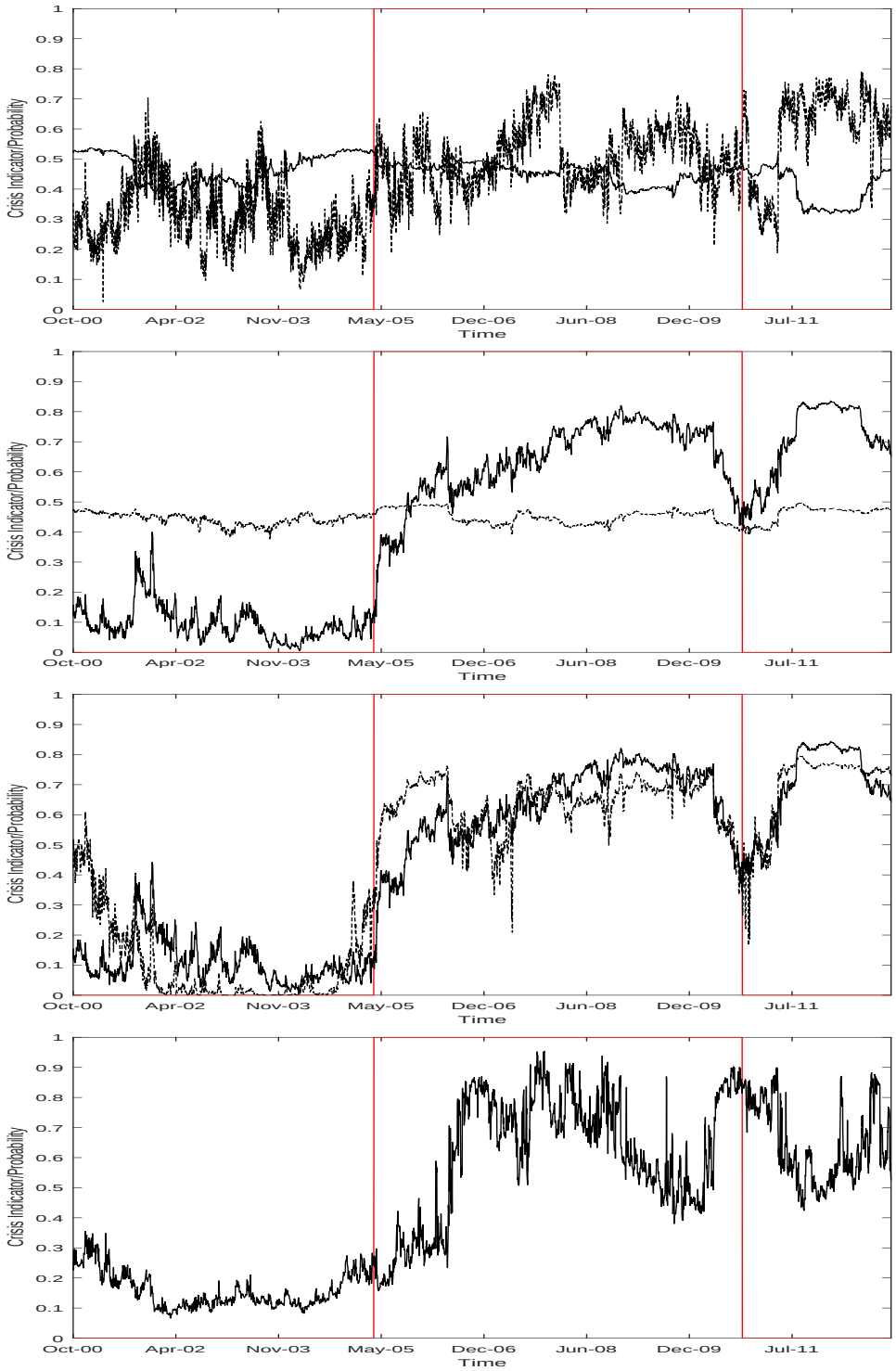


Figure 7: Overview of the actual (stepwise red lines) defined in Equation **??** and the estimated response variable over time: *(first panel)* Dynamic causality index (DCI, solid line) as proposed in **?** and Shannon entropy on In-Out degree (dashed line) as in **?**; *(second panel)* Von Neumann entropies with linear functional form and Laplacian L_1 (solid line) and Laplacian L_2 (dashed line); *(third panel)* Von Neumann entropies with exponential functional form and Laplacian L_1 (solid line) and Laplacian L_2 (dashed line) and *(fourth panel)* μ indicates the Disagreement persistence (solid line).

7.3 EWI evaluation and loss function

In the previous section we compare the models by applying some goodness of fit statistical measures. In practice an economic comparison is needed, especially when policymakers are interested using early warning signals to detect vulnerabilities in the financial system according to the indicators of interest (i.e. banking crises). For detecting crisis events using information from indicator C_t, we use $\hat{\Phi}_t = \Phi(\hat{\beta}_0 + \hat{\beta}_1 S_t)$, the predicted probability of crisis returned by the logit model. Then, the predicted probability is turned into a binary prediction, which takes the value of 1 if $\hat{\Phi}_t$ exceeds a specified threshold c and 0 otherwise, i.e.

$$\hat{C}_t = \begin{cases} 1 & \text{if } \hat{\Phi}_t \geq c, \\ 0 & \text{otherwise.} \end{cases} \tag{39}$$

We set the threshold $c = 0.50$.[6]. The key aspect of these models is the forecast evaluation and therefore the quality of the issued signals. In fact, policymakers aim to distinguish for two type of potential errors: the missed crisis (type 1 error) and the false signal of crisis (type 2 error). Therefore, the pairs of values of the actual and predicted crisis $(C_t, \hat{C}_t)$ can form four possible combinations: both equal to 1 or 0, or different. We can represent those values into a contingency matrix [see **????**] which describes this relationship as reported in Figure **??**.

The type 1 error (T_1) represents the share of missed crises on the total crises,

$$\frac{FN}{TP + FN}, \tag{40}$$

while the type 2 error (T_2) represents the share of issued false alarms on the total tranquil periods,

$$\frac{FP}{TN + FP}. \tag{41}$$

The percentage of correctly predicted indicators is the correct signal (TP) and correct no signal (TN) over the total realizations. We evaluate the impact of the two type of errors using a loss function. In this regard, we adopt the one proposed by **?** L which is simple to implement and robust to small perturbations,

$$L(\theta) = \theta T_1 + (1 - \theta) T_2, \tag{42}$$

[6] As defined in Equation 17 in **?**

	Crisis ($C_t = 1$)	Not crisis ($C_t = 0$)
Signal issued ($\hat{C}_t = 1$)	Correct signal *(true positive, TP)*	False alarm *(false positive, FP)*
No signal issued ($\hat{C}_t = 0$)	Missed crisis *(false negative, FN)*	Correct no signal *(true negative, TN)*

Figure 8: The contingency matrix for the crises indicator C_t and the binary predictor $\hat{C}_t$. The type 1 error (T_1) represents the share of missed crises on the total crises, $FN/(TP + FN)$, while the type 2 error (T_2) represents the share of issued false alarms on the total tranquil periods, $FP/(TN + FP)$. The percentage of correctly predicted indicators is the correct signal (TP) and no signal (TN) over the possible situations.

where θ is the relative risk aversion parameter of the decision maker among type 1 and type 2 errors. If $\theta > 0.5$, the aversion is greater for missing a crisis (false negative) than a false alarm (false positive). As suggested by **?**, we set θ equal to 0.5 since it is uncommon to have preferences for values above this value among the financial stability community. If the binary predictor performs perfectly the value of the loss function is 0. Conversely, the worst it performs the more is close to 1. Other robust methods such as the ROC curve can be used to evaluate the TP and TN by varying the threshold [e.g, see **?**]. Table **??** reports the results of the predicted crises specified according to the contingency matrix. Results show the goodness of the Von Neumann entropies and the disagreement persistence index with respect to the DCI and the Shannon entropy on the In-Out degree. In particular, the Von Neumann entropy with the exponential functional form and Laplacian L_2, $S(\rho_{E_2})$, shows the lowest value in the loss function (0.19) with the highest true positive (TP) and the lowest false negative (FN) values. Moreover, the Von Neumann entropy with the exponential functional form and the Laplacian L_2 shows the lowest type 2 errors (T_2) while the exponential one with Laplacian L_1 shows the lowest type 1 errors (T_1). Since $S(\rho_{E_2})$ has a higher T_1 but a lower T_2 with respect to the $S(\rho_{E_1})$, if we set a lower value for θ,

	DCI	H	$S(\rho_{L_1})$	$S(\rho_{L_2})$	$S(\rho_{E_1})$	$S(\rho_{E_2})$	μ
correctly predicted	0.59	0.67	0.78	0.48	0.79	0.80	0.71
TP	0.31	0.31	0.41	0.23	0.41	0.43	0.34
TN	0.29	0.36	0.37	0.25	0.38	0.37	0.37
FP	0.26	0.19	0.18	0.30	0.17	0.18	0.18
FN	0.14	0.14	0.04	0.22	0.04	0.02	0.11
T_1	0.32	0.31	0.09	0.49	0.10	0.05	0.24
T_2	0.48	0.34	0.33	0.54	0.30	0.33	0.33
$L(\theta = 0.5)$	0.40	0.33	0.21	0.51	0.20	0.19	0.28
$L(\theta = 0.4)$	0.42	0.33	0.23	0.52	0.22	0.22	0.29
$L(\theta = 0.3)$	0.43	0.33	0.26	0.53	0.24	0.25	0.30

Table 3: Percentage of correctly predicted banking crises with the logit models using: a) Dynamic causality index (DCI) as proposed in **?**; b) Shannon entropy on In-Out degree as in **?**; c) $S(\rho_{L_1})$ indicates the Von Neumann entropy with the linear functional form and Laplacian L_1; d) $S(\rho_{L_2})$ indicates the Von Neumann entropy with the linear functional form and Laplacian L_2; e) $S(\rho_{E_1})$ indicates the Von Neumann entropy with the exponential functional form and Laplacian L_1; f) $S(\rho_{E_2})$ indicates the Von Neumann entropy with the exponential functional form and Laplacian E_2 and g) μ indicates the disagreement persistence index. Banking crisis indicators are defined as more than one countries on crisis. The data are from **?**. T_1 and T_2 represent the share of missed crises on the total crises ($FN/(TP + FN)$) and the share of issued false alarms on the total tranquil periods ($FP/(TN + FP)$), respectively. The percentage of correctly predicted indicators is the correct signal (TP) and no signal (TN) over the possible situations. $L(\theta)$ is the loss function and θ represents the relative risk aversion parameter of the decision maker among type 1 and type 2 errors.

the loss function penalizes more T_2 errors and consequently $S(\rho_{E_1})$. By selecting $\theta = 0.4$, both the Von Neumann entropies with the exponential functional form perform equally well, while for values below or equal to 0.3 the entropy with the Laplacian L_1 performs better.

8 Conclusion

In this paper, we provide a graph theoretic background for the analysis of financial networks and review some technique recently proposed for the extraction of financial networks. We develop new measures of network connectivity based on the notion of Von Neumann entropy and show that they account for global connectivity patterns given by paths and walks of the network. We apply the new measures to a sequence of inferred pairwise-Granger networks. In the application, we show how to use the proposed measures to achieve effective immunization of the financial system from the spread of contagion. Finally, we show that entropy measures can be successfully employed to generate early warning signals for

banking crises.

A Proofs of the results in the paper

A.1 Proof of Theorem ??

- Let $L = L_1$, $B = \varphi^{1/2} P \varphi^{-1/2}$ and $C = \varphi^{-1/2} P' \varphi^{1/2}$. Note that

$$B^l C^{r-l} = (\underbrace{\varphi^{1/2} P \varphi^{-1/2} \dots \varphi^{1/2} P \varphi^{-1/2}}_{\text{l times}})(\underbrace{\varphi^{-1/2} P' \varphi^{1/2} \dots \varphi^{-1/2} P' \varphi^{1/2}}_{\text{r-l times}})$$
$$= \varphi^{1/2} P^l \varphi^{-1/2} \varphi^{-1/2} (P')^{r-l} \varphi^{1/2} = \varphi^{1/2} P^l \varphi^{-1} (P')^{r-l} \varphi^{1/2}$$

then

$$\begin{aligned} \mathrm{tr}(L^m) &= \mathrm{tr}(I^m) + \sum_{k=0}^{m-1} \binom{m}{k} (-2)^{-(m-k)} \mathrm{tr}((B+C)^{m-k}) \\ &= n + \sum_{k=0}^{m-1} \binom{m}{k} (-2)^{-(m-k)} \sum_{l=0}^{m-k} \binom{m-k}{l} \mathrm{tr}(P^l \varphi^{-1} (P')^{m-k-l} \varphi) \end{aligned} \tag{43}$$

The results follows by considering that the (i,j)-th element of P^q is equal to the (j,i)-th element of $(P')^q$ which is

$$W_{ij}^{(q)} = \sum_{i_2} \dots \sum_{i_q} \prod_{r=1}^{q} P_{i_r i_{r+1}} = \sum_{i_2 \cdots i_q} \prod_{r=1}^{q} a_{i_r i_{r+1}} \frac{1}{d_{i_r}^-}$$

where $i_1 = i$ and $i_{q+1} = j$, and by applying the definition of trace

$$\mathrm{tr}(P^l \varphi^{-1} (P')^{r-l} \varphi) = \sum_{i=1}^{n} \sum_{h=1}^{n} W_{ih}^{(l)} W_{ih}^{(r-l)} \varphi_i \varphi_h^{-1}$$

- Let $L = L_2$ and $B = \varphi P + P' \varphi^{-1}$, then

$$\mathrm{tr}(L^m) = \mathrm{tr}(\varphi^m) + \sum_{k=0}^{m-1} \binom{m}{k} (-2)^{-(m-k)} \mathrm{tr}(\varphi^k B^{m-k}) \tag{44}$$

Let $R^{(k)} = \varphi^k B^{m-k}$ with (i,j)-th element

$$(\varphi^k B^{m-k})_{ij} = \varphi_i^k W_{ij}^{(m-k)}$$

where

$$W_{ij}^{(q)} = \sum_{i_2=1}^{n} \dots \sum_{i_q=1}^{n} \prod_{r=1}^{q} \left(\frac{\varphi_{i_r}}{d_{i_r}^-} a_{i_r i_{r+1}} + a_{i_{r+1} i_r} \frac{1}{\varphi_{i_{r+1}} d_{i_{r+1}}^-} \right)$$

is the (i,j)-th element of $W^{(q)} = B^q$ one obtains

$$\mathrm{tr}(L^m) = \sum_{i=1}^{n} \varphi_i^m + \sum_{k=0}^{m-1} \binom{m}{k} (-2)^{-(m-k)} \mathrm{tr}(R^{(k)}) \tag{45}$$

A.2 Proof of Theorem ??

Note that the Von Neumann entropy can be rewritten in terms of trace of the Laplacian, that is

$$S^{(M)}(\boldsymbol{\lambda}_i) = \frac{1}{n}\mathrm{tr}(\tilde{L}_i^{(M)}) - \frac{1}{n^2}\mathrm{tr}((\tilde{L}_i^{(M)})^2) \tag{46}$$

where $\boldsymbol{\lambda}_i$ is the vector of eigenvalues of $\tilde{L}_i^{(M)}$.

By replacing $L_1^{(M)}$ by its definition with $M = 1$ and by applying the results in Theorem **??** one obtains

$$\begin{aligned}
&S^{(M)}(\boldsymbol{\lambda}_1) = \frac{1}{n}\mathrm{tr}(I + L_1) - \frac{1}{n^2}\mathrm{tr}((I + L_1)^2) \\
&= 2 - \frac{3}{n} - \frac{1}{n^2}\mathrm{tr}(L_1^2) \\
&= 2 - \frac{3}{n} - \frac{1}{n^2}\left(\frac{1}{4}(\mathrm{tr}(P') + 2\mathrm{tr}(\varphi P \varphi^{-1} P') + \mathrm{tr}(P^2)) - 2\frac{1}{2}(\mathrm{tr}(P') + \mathrm{tr}(P))\right) \quad (47) \\
&= 2 - \frac{3}{n} - \frac{1}{4n^2}\left(2\mathrm{tr}(P^2) + 2\mathrm{tr}(\varphi P \varphi^{-1} P')\right) \quad (48)
\end{aligned}$$

by using the fact that $\mathrm{tr}(P') = \mathrm{tr}(P) = 0$ since the transition matrix P has zero elements on the main diagonal. This entropy can be decomposed further using out- and in-degree of the node following **?**.

For $L_2^{(M)}$ we have instead

$$\begin{aligned}
&S^{(M)}(\boldsymbol{\lambda}_2) = \frac{1}{n}\mathrm{tr}(I + L_1) - \frac{1}{n^2}\mathrm{tr}((I + L_1)^2) \\
&= 1 - \frac{2}{n^2} + \left(\frac{1}{n} - \frac{2}{n^2}\right)\left(\mathrm{tr}(P(\varphi + \varphi^{-1})/2)\right) \\
&-\frac{1}{n^2}\left(\sum_{i=1}^{n}\varphi_i^2 + \frac{1}{4}(\mathrm{tr}((\varphi P)^2) + \mathrm{tr}((P'\varphi^{-1})^2) + \mathrm{tr}(P'P) + \mathrm{tr}(PP')) - \mathrm{tr}(\varphi^2 P)\right) \quad (49)
\end{aligned}$$

By replacing $L_1^{(M)}$ by its definition with $M = 2$ and by applying the results in Theorem

?? one obtains

$$
\begin{aligned}
&S^{(M)}(\boldsymbol{\lambda}_1) = \frac{1}{n}\mathrm{tr}(I + L_1 + L_1^2) - \frac{1}{n^2}\mathrm{tr}(I + 2L_1 + L_1^2 + \frac{3}{2}L_1^3 + \frac{1}{4}L_1^4) \\
&= 2 - \frac{3}{n} - \frac{2}{n^2}\mathrm{tr}(L_1^2) - \frac{1}{n^2}\mathrm{tr}(L_1^3) - \frac{1}{4n^2}\mathrm{tr}(L_1^4) \\
&= 2 - \frac{3}{n} - \frac{2}{n^2}\sum_{k=0}^{1}\sum_{l=0}^{2-k}\omega_{lk}\mathrm{tr}(R^{(lk)}) - \frac{1}{n^2}\sum_{k=0}^{2}\sum_{l=0}^{3-k}\omega_{lk}\mathrm{tr}(R^{(lk)}) - \frac{1}{4n^2}\mathrm{tr}(\sum_{k=0}^{3}\sum_{l=0}^{4-k}\omega_{lk}\mathrm{tr}(R^{(lk)})) \\
&= 2 - \frac{25}{4n} - \frac{1}{n^2}\left(\sum_{k=0}^{3}\sum_{l=0}^{4-k}\tilde{\omega}_{lk}\mathrm{tr}(R^{(lk)})\right) \qquad (50)
\end{aligned}
$$

with

$$
\begin{aligned}
&\tilde{\omega}_{00} = \frac{13}{4}\omega_{00},\ \tilde{\omega}_{10} = \frac{13}{4}\omega_{10},\ \tilde{\omega}_{20} = \frac{13}{4}\omega_{20},\ \tilde{\omega}_{30} = \frac{5}{4}\omega_{30},\ \tilde{\omega}_{40} = \frac{1}{4}\omega_{40} \\
&\tilde{\omega}_{01} = \frac{13}{4}\omega_{01},\ \tilde{\omega}_{11} = \frac{13}{4}\omega_{11},\ \tilde{\omega}_{21} = \frac{5}{4}\omega_{21},\ \tilde{\omega}_{31} = \frac{1}{4}\omega_{31} \\
&\tilde{\omega}_{02} = \frac{5}{4}\omega_{02},\ \tilde{\omega}_{12} = \frac{5}{4}\omega_{12},\ \tilde{\omega}_{22} = \frac{1}{4}\omega_{22} \\
&\tilde{\omega}_{03} = \frac{1}{4}\omega_{03},\ \tilde{\omega}_{13} = \frac{1}{4}\omega_{13}
\end{aligned}
$$

where we used $\mathrm{tr}(L_1) = n$.

By replacing $L_2^{(M)}$ by its definition with $M = 2$ and by applying the results in Theorem **??** one obtains

$$
\begin{aligned}
&S^{(M)}(\boldsymbol{\lambda}_2) = 2 - \frac{3}{n} - \frac{2}{n^2}\mathrm{tr}(L_2^2) - \frac{1}{n^2}\mathrm{tr}(L_2^3) - \frac{1}{4n^2}\mathrm{tr}(L_2^4) \\
&= 2 - \frac{3}{n} - \frac{2}{n^2}\left(\sum_{i=1}^{n}\varphi_i^2 + \sum_{k=0}^{1}\omega_k\mathrm{tr}(R^{(k)})\right) - \frac{1}{n^2}\left(\sum_{i=1}^{n}\varphi_i^3 + \sum_{k=0}^{2}\omega_k\mathrm{tr}(R^{(k)})\right) \\
&-\frac{1}{4n^2}\left(\sum_{i=1}^{n}\varphi_i^4 + \sum_{k=0}^{3}\omega_k\mathrm{tr}(R^{(k)})\right) \\
&= 2 - \frac{3}{n} - \frac{1}{n^2}\sum_{k=0}^{3}\tilde{\omega}_k\mathrm{tr}(R^{(k)}) \qquad (51)
\end{aligned}
$$

with

$$
\tilde{\omega}_0 = \frac{13}{4}\omega_0,\ \tilde{\omega}_1 = \frac{13}{4}\omega_1,\ \tilde{\omega}_2 = \frac{5}{4}\omega_2,\ \tilde{\omega}_3 = \frac{1}{4}\omega_3
$$

A.3 Proof of Corollary ??

Consider $S^{(M)}(\boldsymbol{\lambda}_1)$ given in Equation **??** with $M = 1$, and note that $\mathrm{tr}(L_1) = n$, then

$$S^{(1)}(\boldsymbol{\lambda}_1) - S_Q = 2 - \frac{3}{n} - \frac{1}{n^2}\mathrm{tr}(L_1^2) - (\frac{1}{n}\mathrm{tr}(L_1) - \frac{1}{n^2}\mathrm{tr}(L_1^2)) = 1 - \frac{3}{n} \tag{52}$$

A.4 Proof of Theorem ??

In order to prove Theorem **??** we need the following lemma

Lemma 1 *Given the matrix $M = \varphi^{1/2} P_L \varphi^{-1/2}$ we have*

$$\max_{\boldsymbol{\varphi}'\mathbf{v}=0} \frac{\mathbf{v}' P_L' \varphi P_L \mathbf{v}}{\mathbf{v}'\varphi\mathbf{v}} = \max_{\boldsymbol{\varphi}^{1/2\prime}\mathbf{u}=0} \frac{\mathbf{u}' M' M \mathbf{u}}{\mathbf{u}'\mathbf{u}} \leq \mu$$

Proof.

$$\begin{aligned} & \max_{\boldsymbol{\varphi}'\mathbf{v}=0} \frac{\mathbf{v}' P_L' \varphi P_L \mathbf{v}}{\mathbf{v}'\varphi\mathbf{v}} \\ = & \max_{\boldsymbol{\varphi}'\varphi^{-1/2}\varphi^{1/2}\mathbf{v}=0} \frac{\mathbf{v}'\varphi^{1/2}\varphi^{-1/2} P_L' \varphi^{1/2}\varphi^{1/2} P_L \varphi^{-1/2}\varphi^{1/2}\varphi^{1/2}\varphi^{1/2}\mathbf{v}}{\mathbf{v}'\varphi^{1/2}\varphi^{1/2}\mathbf{v}} \\ = & \max_{\boldsymbol{\varphi}^{1/2\prime}\mathbf{u}=0} \frac{\mathbf{u}' M' M \mathbf{u}}{\mathbf{u}'\mathbf{u}} \end{aligned}$$

According to **?** we can bound the last term using the second eigenvalue λ_2 of the symmetrized laplacian L

$$\max_{\boldsymbol{\varphi}^{1/2\prime}\mathbf{u}=0} \frac{\mathbf{u}' M' M \mathbf{u}}{\mathbf{u}'\mathbf{u}} \leq 1 - \frac{\lambda_2}{2}$$

? give a tighter bound :

$$\max_{\boldsymbol{\varphi}^{1/2\prime}\mathbf{u}=0} \frac{\mathbf{u}' M' M \mathbf{u}}{\mathbf{u}'\mathbf{u}} \leq \frac{3}{4} - \frac{\lambda_2}{2} + \frac{\left(\sigma_{n-1}\left(I_n - L\right) + \sigma_n\left(\Delta\right)\right)^2}{4}$$

In addition, using the primitivity of P and the results in theorem 4.3 in **?**, $\mu < 1$.

□

To show the results in Theorem **??**, we note that

$$\boldsymbol{\varphi}'\xi_t \;=\; \boldsymbol{\varphi}'\mathbf{x}_t - \alpha\boldsymbol{\varphi}'\mathbf{1} = 0 \tag{53}$$

So the disagreement vector satisfies, at any time, the constraint imposed on the maximiza-

tion of the quotient in Lemma **??**.

Let $V(t) = \boldsymbol{\xi}_t' \varphi \boldsymbol{\xi}_t$ a candidate Lyapunov function for the system (**??**). It is a valid candidate because it is well known that the Perron vector $\boldsymbol{\varphi}$ is strictly positive and so $V(t) = 0$ if and only if $\boldsymbol{\xi}_t$ is equal to the zero vector. Then a bound on $V(t+1)$ is easily obtained using Lemma **??**:

$$\begin{aligned} V(t+1) &= \boldsymbol{\xi}_{t+1}' \varphi \boldsymbol{\xi}_{t+1} = \boldsymbol{\xi}_t' P_L' \varphi P_L \boldsymbol{\xi}_t \\ &\leq \mu \boldsymbol{\xi}_t' \varphi \boldsymbol{\xi}_t \\ &\leq \mu V(t) \end{aligned}$$

since $\mu < 1$ the system is asymptotically stable. Accordingly we have

$$\min(\boldsymbol{\varphi}) \boldsymbol{\xi}_t' \boldsymbol{\xi}_t \leq \boldsymbol{\xi}_t' \varphi \boldsymbol{\xi}_t \leq \mu^t \boldsymbol{\xi}_0' \varphi \boldsymbol{\xi}_0 \leq \max(\boldsymbol{\varphi}) \mu^t \boldsymbol{\xi}_0' \boldsymbol{\xi}_0$$

from which the result follow.