Emergence of Complexity in Financial Networks

Guido Caldarelli[1], Stefano Battiston[2], Diego Garlaschelli[3] and Michele Catanzaro[1]

[1] INFM UdR Roma1 Dipartimento di Fisica Università La Sapienza P.le Moro 5 00185 Roma, Italy. gcalda@pil.phys.uniroma1.it, michele@pil.phys.uniroma1.it
[2] Laboratoire de Physique Statistique ENS, 24 rue Lhomond, 75005 Paris France. battiston@ens.fr
[3] INFM UdR Siena and Dipartimento di Fisica Università di Siena, Via Roma 56, 53100 Siena Italy. diego@pil.phys.uniroma1.it

Abstract. We present here a brief summary of the various possible applications of network theory in the field of finance. Since we want to characterize different systems by means of simple and universal features, graph theory could represent a rather powerful methodology. In the following we report our activity in three different subfields, namely the board and director networks, the networks formed by prices correlations and the stock ownership networks. In most of the cases these three kind of networks display some scale-free properties making them interesting in their own. Nevertheless, we want to stress here that the main utility of this methodology is to provide new measures of the real data sets in order to validate the different models.

1 Introduction

The study of topological properties of networks has recently received great attention [1]. In particular it has been shown that many natural systems display an unexpected amount of correlation[2] with respect to traditional models[3]. Graphs are mathematical objects formed by vertices connected by arcs. An important characterization of a graph is given by the degree of vertices, that is the number of arcs per vertex. In an Erdős-Rényi random graph the degree distribution has a poissonian form, whereas in many cases of interest ranging from the WWW[4,5] to the Internet[6,7] to social networks[8] the degree is power law distributed. The scale-free behavior can be reproduced by two classes of models. The growth models where new sites enter and choose a site to be linked with through the "rich gets richer" rule of preferential attachment[9]. Interestingly, there are also other ways to reproduce such scale invariance by means of static models where "good gets richer"[10]). Choice between different models therefore should be driven by observation of the real data. One has to consider if a growth of the network is present or not during the time evolution. While growth is certainly present in technological networks like Internet and WWW, this could not be the case for the systems presented here.

Besides the distribution of the connectivity degree of vertices, assortativity and clustering are among the most interesting quantities to consider for classifying and describing these complex networks. Assortativity can be defined introducing the quantity $Knn(k)$, giving the average degree of the site neighbors of one site whose degree is k. Knn increases if nodes are correlated by degree (assortative networks). It decreases if they are anti-correlated (dis-assortative networks). The tendency for nodes in a social network to form connections preferentially to others similar to them[11] has been proposed as the key ingredient for the formation of communities in networks[12,13]. It is possible to distinguish the technological networks, where instead, the behavior is rather degree-disassortative, so that vertices tend to be linked to others different from them. Despite the relative simplicity of such behavior few models[14–16] of network growth are able to reproduce the formation of communities and no one explains the difference between social and technological networks.

Clustering coefficient c_i for every site i gives the probability that two nearest neighbors of vertex i are also neighbors each other. $cc(k)$, is the average clustering coefficient for sites whose degree is k, and it measures the tendency to form cliques where each nearest neighbor of a node (with degree k) is connected to each other. In real networks this usually decreases with a power-law $cc(k) \propto k^{\psi}$ because hubs tend to play the role of connections between separate clusters in the graph, i.e. clusters that have few other interconnections than the ones passing through the hub. Then the high degree node tends to have low clustering coefficient.

We focus in this paper on three different kinds of complex networks with relevance in finance: the network of boards and directors of the largest corporations, the network of stock price correlation and the network of shareholders in the stock market. For each network we report the more relevant topological properties and we present models of network formation accounting for some of the observed properties.

2 The Board and Director Networks

We start with the network of boards and directors, a complex network in finance which is also a social network.

This can be represented as a bipartite graph where two classes of nodes are present (boards and directors) and an edge is always disposed between nodes belonging to different classes. In particular, a director is linked with a society if he sits in its board. Such a graph can be projected into the board network and the director network, where two boards (directors) are connected with an edge weighted proportionally to the numbers of directors (boards) they have in common. In fact, very often two boards share some directors, in which case they are said to be *interlocked*.

An example of board network is shown in figure 1: nodes represent boards of directors, two boards are connected by an edge if they are interlocked. The network represents the boards 1 degree of separation (in term of edges) away from the board of Chase Manhattan Bank.

Boards of directors of corporations make decisions about the long-term strategy, such decisions having considerable impact on the economic performance of the corporation and collectively on the economy of a country.

Because large corporations' boards are organized in a networks leading the economy of a country, some issues are particularly relevant about these networks: what the topological properties are and if they are similar in different countries. What these topological properties mean about the corporate directorate elite as a leading class. What mechanism of network formation can explain the observed features. What effect the network structure has on the process of decision making.

2.1 Topological properties of board and director networks

A number of recent works has been devoted to the study of the topological properties of the board and director networks. Davis et al. [17] have studied the network of the boards of Fortune 1000 in 1999 and have shown that both the director network and the board network have Small World properties.

Newman et al. [18] have applied on the same data set a random graph model showing that using the generating function method, it is possible to reproduce very accurately the degree distribution of the director network. On the contrary, their model fails in predicting the degree distribution of the board network. In fact the director network turns out to be *assortative* as observed commonly in social networks, meaning that directors with high (low) degree tend to be connected to directors with high (low) degree. As a consequence even if the random graph model predicts the right degree distribution for the director network it underestimate the number of boards with high number of interlocks and with small number of interlocks.

As a general empirical finding, social networks are characterized by assortativity and high *clustering coefficient cc* (the latter measuring the average fraction of connection between the first neighbors of a node out of all the possible connections among them). We report in table 1 the values of assortativity coefficient r (see [19]) and clustering coefficient cc for two novel data sets we collected and analysed, namely the networks of boards of the companies quoted on the Italian stock Market in 1986 and 2002. In figure 2 we report the average nearest neighbors degree Knn and the clustering coefficient cc as a function of the degree of the nodes. As a general trend, nodes with high degree tend to be connected to nodes with high degree, nodes with high degree tend to have low values of cc.

Newman et al. [20] have recently argued that the presence of groups or communities in a social network is able to produce alone both assortativity and clustering. They develop a model in which nodes belong to one or more groups and have probability p to be connected to another node of the same group. Instead they are never connected to nodes of groups they do not belong to. If groups have heterogenous size, than nodes who belong to a small group tend to have low degree and are connected to others in the same group, who also have low degree.

This model explains about 40 p.c. of the observed assortativity in the Fortune 1000 network. This means that some additional sociological mechanism is at work, probably the fact that new board members are more likely to be recruited among those who are already connected to some of the current board member.

Some of us have recently proposed a new model [21] to reproduce assortativity in social networks. This is a network growth model and is a generalization of the Barabasi-Albert preferential attachment model [9], which is known to produce a scale free network with no clustering and no assortativity. The generalization of that model proposed by Catanzaro et al. allows for growth by addition of new links between old nodes. In details, at every step of growth:

1. with probability p a new node is wired to an existing one with the Barabási-Albert preferential attachment rule ('rich gets richer').
$$p\frac{k_i}{\sum_{j=1,N} k_j}. \tag{1}$$
2. with probability $(1-p)$ a new edge is added (if absent) between two existing nodes. These are chosen on the basis of their degree. In other words, the probability of adding an edge between node 1 and node 2 is a $\tilde{P}(k_1, k_2)$. This can be written as $P_1(k_1)P_2(k_2|k_1)$, being the second factor a conditioned probability. $P_1(k_1)$ is the rule for choosing the first of the two nodes, and again it is determined by the preferential attachment. The functional form of $P_2(k_2|k_1)$ can be chosen so as to favor links between similar or different degree. In this way, the probability of adding a new edge and connecting two old non-linked nodes is
$$(1-p)\frac{k_i}{\sum_{j=1,N} k_j}P_2(k_2|k_1) \tag{2}$$

In the limit of $p = 1$ the model reduces to a traditional BA tree. By tuning the parameter p, it is possible to weight the role of growing (addition of new nodes) and mixing (addition of new edges) in the microscopical behavior of the network.

The authors explore two different functional forms for $P_2(k_2|k_1)$: an inverse dependence
$$P_2(k_2|k_1) \propto \frac{1}{|k_1 - k_2| + 1} \tag{3}$$

and an exponential dependence

$$P_2(k_2|k_1) \propto e^{-|k_1-k_2|}. \quad (4)$$

In the first case the model produces a scale free network: the connectivity degree is power law distributed with exponent monotonically increasing with p.

In the second case for $p < 0.5$ a peak at high degree appears in the degree distribution.

In both cases the resulting network presents a core-periphery structure, where hubs (highly connected nodes) connect with other hubs. This structure is emphasized in the exponential case, where the assortativity becomes so large to induce a phase transition from a scale-free graph to a network with a characteristic scale for high degrees.

One one hand, the model reproduces assortativity, reducing its emergence to the role of mixing in growth. On the other hand, it fails in reproducing the fact that the clustering coefficient decreases with the degree k in real social networks. In real networks hubs tend to play the role of connections between separate clusters in the graph, with few links between each other (apart from the ones attached to the hub). Therefore this nodes tend to have low clustering coefficient. In this model, on the other hand, all the hubs are aggregated together. Thus, even producing an assortative network it cannot reproduce a network with $cc(k)$ decreasing with k.

See [21] for more details.

We here suggest that the choice of $P1(K1)$ might be also relevant. Choosing the first node of a new edge at random instead that with a BA rule, would imply that nodes with any degree will tend to be connected with nodes of similar degree, without favoring a cluster of hubs.

A simple rationale for this rule in the director network follows from the assumption that degree is related to prestige **??**. It is quite reasonable to assume that individuals would like to be connected to others with similar or higher prestige and to loose connection with individuals with lower prestige. As a result individuals tend to be connected to others with similar prestige. Now, this is reflected in the recruiting process of new members of a board. Directors try to recruit directors with more prestige, but they manage to recruit directors with similar prestige.

As a conclusion, a general model for network formation reproducing all the statistical features observed in social networks is still object of search. We reported some common features of boards and directors networks for US and Italy and we discussed some recent models producing networks with assortativity and high clustering coefficient.

2.2 The impact of network structure on the decision making process

Topological properties reveal interesting features of the boards-directors economical system. For example, it should be noted that both the projections of the bipartite graph display a giant connected component containing around 90 p.c. of the nodes. Therefore, the interlock is strong enough to bring the network over the percolation threshold. Moreover, while positive assortativity suggests prestige related dynamics (see above), high clustering coefficient and small world property describe a social system where common "friendships" and close proximity to each other are typical. As well, it is interesting to observe that the Italian network, for which we have two time snapshots, seems to be stationary, as suggested by the substantial invariance of the shape of the degree distribution in the years. A deeper investigation is required to explore the differences (if any) between Italian and US market.

Some recent works have focussed on the influence of the structure of the interlock network on the decisions made by boards. There are essentially two kinds of decisions a board is faced to. *Local* decisions regard topics specific to the board, such as the appointment of a vice president, for which boards can be assumed not to influence each other. Battiston et al. [22] have shown the role of subsets of well connected directors on decisions of this type.

By contrast, *global* decisions concern topics of general interest to the economy such as whether to increase or decrease investments in development or in advertisement, which depend on the belief in economical growth or recession. In these cases, decisions previously made in some boards might influence other boards, through the presence of shared directors.

In a recent model, Battiston et al. [23] investigate the conditions under which a large majority of boards making a same decision can emerge in the network. In their model board directors are engaged in a decision making dynamics based on "herd behavior" and boards influence each other through shared directors.

They find that imitation of colleagues and opinion bias due to the interlock do not trigger an avalanche of identical decisions over the board network, whereas the information about interlocked boards' decisions does. There is no need to invoke global public information, nor external driving forces. This model provides a simple endogenous mechanism to explain the fact that boards of the largest corporations of a country can, in the span of a few months, take the same decisions about general topics, despite the a priori uncertainty of the economic trend.

3 Network of price Correlations

The case of study is given by a network whose vertices are a fixed number of stocks continuously traded at the New York Stock Exchange (NYSE) and the arcs are obtained by considering the return cross-correlations. The network is a Minimal Spanning Tree (MST) connecting all the stocks. Spanning trees are particular types of graphs. They connect all the vertices in a graph without forming any loop.

It is interesting to note that when the stock correlations are described through this method it is very easy to validate the models of portfolio dynamics. Specifically we show that a simple model of uncorrelated Gaussian return time series and the widespread one-factor model do not reproduce the topological quantities of interest. This last model is the starting point of the Capital Asset Pricing Model[24] that is one of the most widely known models.

3.1 The MST formation

The topological characterization of the correlation based MST of real data has been already studied in Ref. [25]. Differently from this approach, here we use a number of stocks N smaller but we use a number of time records T larger than the number of stocks. Our choice is motivated by the request that the correlation matrix be positive definite. When the number of variables is larger than the number of time records the covariance matrix is only positive semi-definite[26]. Moreover, the application of the random matrix theory to the spectral properties of the correlation matrix can be applied only when $T/N > 1$.

We consider the daily price return $r_i(t)$ of asset i on day t. Given a portfolio composed of N assets traded simultaneously in a time period of T trading days, we extract the $N \times N$ correlation matrix. ¿From each correlation coefficient $\rho_{i,j}$ we computed a metric distance $d_{i,j} = \sqrt{2(1-\rho_{i,j})}$ between asset i and j through the relation [27,28]. The distance matrix is then used to determine the MST connecting all the assets. We start with the two nearest sites and we draw them, then we consider the second shorter distance. This can link two different vertices from the previous ones or may link one of the older with a new one. We draw also this link, now for the third distance we can have a link forming a loop. If this is the case we do not draw such distance and we pass on the next couple of sites. This procedure iteratively form the Minimum Spanning Tree. The method of constructing the MST linking N objects is known in multivariate analysis as the nearest neighbor single linkage cluster algorithm[26].

3.2 The Data and the One Factor Model

The data set used here consists of daily closure prices for 1071 stocks traded at the NYSE and continuously present in the 12-year period 1987-1998 (3030 trading days). It is worth noting that the ratio $T/N \simeq 2.83$ is significantly larger than one. With our choice the correlation matrix is positive definite and the theoretical results of the random matrix theory are valid. **Figure 3** shows the MST of the real data. The color code is chosen by using the main industry sector of each firm according to the Standard Industrial Classification system[30] for the main industry sector of each firm and the correspondence is reported in the figure caption. Regions corresponding to different sectors are clearly seen. Examples are clusters of stocks belonging to the financial sector (purple), to the transportation, communications, electric gas and sanitary services sector (green) and to the mining sector (red). The mining sector stocks are observed to belong to two subsectors one containing oil companies (located on the right side of the figure) and one containing gold companies (left side of the figure).

The empirical MST of real data can be compared with the results obtained from simple models of the simultaneous dynamics of a portfolio of assets. The simplest model assumes that the return time series are uncorrelated Gaussian time series, i.e. $r_i(t) = \epsilon_i(t)$, where $\epsilon_i(t)$ are Gaussian random variables with zero mean and unit variance. This type of model has been considered in Ref. [31,32] as a null hypothesis in the study of the spectral properties of the correlation matrix. In the cited references it has been shown that the spectrum of the real correlation matrix has a very large eigenvalue corresponding to the collective motion of the assets. A random model does not explain this empirical observation and therefore this fact clarifies why a better modeling of the portfolio dynamics is obtained by using the one-factor model. The one-factor model assumes that the return of assets is controlled by a single factor (or index). Specifically for any asset i we have

$$r_i(t) = \alpha_i + \beta_i r_M(t) + \epsilon_i(t), \tag{5}$$

where $r_i(t)$ and $r_M(t)$ are the return of the asset i and of the market factor at day t respectively, α_i and β_i are two real parameters and $\epsilon_i(t)$ is a zero mean noise term characterized by a variance equal to $\sigma^2_{\epsilon_i}$. Our choice for the market factor is the Standard & Poor's 500 index and we assume that $\epsilon_i = \sigma_{\epsilon_i} w$, where w is a random variable distributed according to a Gaussian distribution.

We estimate the model parameters for each asset from real time series with ordinary least squares method[24] and we use the estimated parameters to generate an artificial market according to Eq. (5). A consequence of this equation is that the variance (the squared volatility) of asset i can be written as the sum of a term depending on the market factor and an idiosyncratic term. The fraction of variance explained by the factor r_M is approximately described by an exponential distribution with a characteristic scale of about 0.16. The random model can be considered as the limit of the one factor model when the fraction of variance explained by the factor goes to zero.

3.3 Results of the models

In the MST obtained with the random model few nodes have a degree larger than few units. This implies that the MST is composed by long files of nodes. These files join at nodes of connectivity equal to few units. The MST obtained with the one-factor model is very different from the one obtained with the random model. In Figure 2 we show the MST obtained in a typical realization of the one-factor model performed with the control parameters obtained as described above. It is evident that the structure of sectors of Fig.3 is not present in Fig.4. In fact the MST of the one-factor model has a star-like structure with a central node. The largest fraction of node links directly to the central node and a smaller fraction is composed by the next-nearest neighbors. Very few nodes are found at a distance of three links from the central node. The central node corresponds to General Electric and the second most connected node is Coca Cola. It is worth noting that these two stocks are the two most highly connected nodes in the real MST also.

In order to characterize quantitatively the structure of the MST we make use of two topological quantities. The first one is the distribution of the degree k. In random graph this quantity is distributed according to a binomial distribution which for large networks tends to a Poisson distribution.

The second topological quantity is frequently used for oriented graphs. For any vertex i in the tree we count the total number of vertices a in the uphill subtree whose root is i. This quantity is called drainage basin area in oriented graphs of river networks[33], whereas it is usually referred as the in-degree component in graph theory. To calculate the in-degree component in a correlation based MST, we orient the MST according to the number of steps each node is far from the most connected node (sink). When more than one sink is present in the MST a preferential one is randomly chosen among them.

We report in **Figure 5** the frequency distribution for the degree k for the real data and for the average over 100 realizations of the random model and of the one factor model. The degree distribution for the MST of the real data shows a power law behavior with exponent -2.6 for one decade followed by a set of isolated points with high degree. A power law behavior with a similar exponent has been observed in Ref.[25] and in another recent study[34]. The highest degree $k_{max} = 115$ is observed for the General Electric, one of the most capitalized company in the NYSE. As pointed out in a previous work[29], some important companies clearly emerge for its high degree value indicating that they act as a reference for other companies. The random model displays an approximately exponential decay of the degree distribution. The value of the maximum degree is small, $k_{max} = 7.34 \pm 0.92$, showing that no asset plays a central role in the MST. The correlation based MST of the random model can be considered as the MST of a set of N points randomly distributed in an Euclidean space with $d = T$ dimension [**?**]. The N points have independent identically Gaussian distributed coordinates $\mathbf{r_i} = (r_i(1), r_i(2), ..., r_i(T))$ with $i = 1, 2, ..., N$. It has been shown that the distribution of degree of the random MST in Euclidean space converges to a specific distribution in the mean field limit $d \to \infty$ [35]. The numerical values of the degree frequency obtained from this mean field limit are shown as a star in Figure 3 for $k = 1, .., 7$. The agreement of theoretical values with the numerical simulations is very good showing that the mean field limit is already a good approximation for our T parameter.

The MST obtained from the one factor model is characterized by a rapidly power-law decaying degree distribution and by an asset with a very high value of the degree, which is indicated by an arrow in Fig. 5. The value of the maximum degree is $k_{max} = 718 \pm 29$. The corresponding asset is the center of the star like structure of Figure 2. The region with highest value of the degree contains information about the stocks that act as reference for a large set of other stocks. To get more insight in the structure of this high k region we show a rank plot of the degree both for real data and for the considered models in the inset of Figure 5. For the real market it is evident the presence of a region of power law extending for more than one decade. On the other hand, for the random model many nodes have a similar value of the degree which is ranging for less than an order of magnitude. This is due to the fact that there is no hierarchy in the random model. The rank plot of the degree of the MST for the one factor model has not a scale free behavior. Indeed, there is a single highly connected node (the center) and a rapidly decaying degree as a function of the rank. This fact corresponds to the simple one-center hierarchy of the MST of the one-factor model.

A discrepancy between real data and models is also observed in the frequency distribution of the in-degree component. **Figure 6** shows the frequency distribution of the in-degree component for real and surrogate data. The inset of Figure 6 shows the rank plot of the same data. In all three cases the in-degree component distribution has a power law shape. This is particularly clear for the MST of the random uncorrelated time series where the power law last for more than two decades with an exponent of approximately -1.6. It is known that for critical random trees the probability distribution of tree size decays as a power-law with an exponent $3/2$ [36]. A critical random tree is a tree in which the mean number of sons of each node is one. In a MST the mean degree is exactly equal to $2n/(n-1) \simeq 2$. Hence when we orient the MST from the root to the leaves we have a tree with one son for each node. Our result shows that the in-degree component of the MST arising from random uncorrelated time series has properties similar to the one of a critical random tree. This is not the case for the one-factor model where the power law has greater absolute slope due to the star-like structure of the tree. Neither models is actually able to catch the oriented structure of real data whose in-degree component distribution is in between the two models. The same arguments are also valid for the region of high values of a as is evident from the rank plot in the inset.

3.4 Comparison between data and models

These results show that the topology of the MST for the real and for the considered artificial markets is different for node with both high and low degree. If we define the importance of a node as its degree (or its in-degree component), from our analysis emerges that the real market has a hierarchical distribution of importance of the nodes whereas the considered models are not able to catch such a hierarchical complexity. Specifically, in the random model the fluctuations select randomly few nodes and assign them small values of degree. Thus the MST of the random model is essentially non hierarchical. On the other hand the MST of the one factor model shows a simple one-center hierarchy. The MST of real market shows a more structured hierarchy of the importance of the stocks which is not captured by the considered models. The topology of stock return correlation based MST shows large scale correlation properties characteristic of complex networks in the native as well as in an oriented form. Such properties cannot be reproduced at all, even as a first approximation, by simple models as a random model or the widespread one-factor model.

4 The Stock investment network

We now consider a pretty different network that can be defined in a financial environment. This network is formed by the companies traded in a stock market and by the corresponding shareholders, and can therefore be considered as a social network. A directed link is drawn from the vertex representing a company to the vertex representing a shareholder of the company itself. While this investment relationship graph has in principle a bipartite nature (vertices can be assigned to two classes, companies and investors), it happens frequently that some shareholders of a certain company are themselves companies whose shares are traded in the market. Therefore the resulting network is in general a directed one where a significant fraction of listed companies are owners of other listed companies, with no well-defined bipartite structure. It is possible to consider the subnet restricted to the owners which are listed companies themselves (hereafter the "restricted" net). This yields the structure reported in Figs. 7,8,9 for different markets, providing a description of the interconnections among stocks. Whenever considering the whole investment relationships we will instead refer to the "extended net". We define the "portfolio diversification" as the in-degree k_i^{in} of the investor i, corresponding to the number of different assets in its portfolio. Vertices with zero in-degree are listed companies holding no shares of other stocks. The out-degree of a vertex is the number of shareholders of the corresponding asset. Since the data are obviously restricted to a limited number of investors per each asset, the out-degree of a company is a biased quantity and we cannot deal with its statistical description. We note that a weight can be assigned to each link, defined as the fraction s_{ij} of the shares outstanding of asset j held by i multiplied by the market capitalization c_j of the asset j. We define "portfolio volume" the quantity $v_i = \sum_j s_{ij} c_j$ representing the total wealth in the portfolio of i.

4.1 Data analysis

The data in our analysis report the shareholders of all stocks traded in the New York Stock Exchange (NYSE), in the National Association of Security Dealers Automated Quotations (NASDAQ) (both in the year 2000), and in the Italian stock market (MIB) in the year 2002. The corresponding number M of assets in the markets is 2053, 3063 and 240 respectively.

On both the extended and the restricted nets, we consider the statistical distribution $P'(k^{in})$ of the number of vertices with in-degree greater than or equal to k^{in} (see Fig. 10a). In all the extended nets the distribution always displays a power-law tail of the form $P'(k^{in}) \propto (k^{in})^{1-\gamma}$. The corresponding probability density is $P(k^{in}) \propto (k^{in})^{-\gamma}$, where the values of the exponent γ are given by $\gamma_{nys} = 2.37$, $\gamma_{nas} = 2.22$, $\gamma_{mib} = 2.97$. Market investments are therefore characterized by a scale-free topology which resembles that displayed by many other complex networks. By contrast, the restricted nets display no power-law behaviour (see Fig. 10b). This means that the relevant contribution to the scale-free nature of market investments comes from the investors outside the market.

It is also possible to compute the cumulative distribution $\theta'(v)$ describing the number of investors with portfolio volume greater than or equal to v. As shown in Fig. 11a, the tail of the distribution displays again a power-law behaviour $\theta'(v) \propto v^{1-\alpha}$. The corresponding probability density is $\theta(v) \propto v^{-\alpha}$, with $\alpha_{nys} = 1.95, \alpha_{nas} = 2.09, \alpha_{mib} = 2.24$. Interestingly, since v represents the invested wealth, the observed power-law tails generalize to a market investment context the well-known Pareto tails describing the right part of the wealth distribution of different economies[37–40]. The left part of the portfolio volume distribution also reflects the (functionally controversial) form of many observed wealth distributions. In the following we are only interested in the right tails of the distribution, so that the characterization of the remaining left part is irrelevant.

Although the scale-free character of the degree is already known to be a widespread topological feature, power-law distributions describing the sum of vertex weights have only been addressed theoretically in the field of complex networks [41]. Therefore our mapping of Pareto distributions (well established in the economic context) in a topological framework provides an empirical basis for the investigation of these specific properties of weighted networks.

4.2 Portfolio diversification and portfolio volume

It is interesting to note that the above empirical results are in contrast with the well-known Capital Asset Pricing Model[24] (CAPM). The latter predicts that the optimal (risk-minimizing) portfolio includes all the M assets traded in the market and is such that the amount of wealth invested in each asset i is proportional to the market capitalization c_i of i. In other words, if i decides to invest a total volume v_i, the best choice is to invest in each asset j a capital

$v_i c_j / \sum_{k=1}^{M} c_k$. In our framework this would clearly imply a constant in-degree $k_i^{in} = M$ for each agent i, even if the total invested volume v varies greatly among the investors. One could suspect that the (very different) observed form of $P(k^{in})$ is a biased result, since the investors of each asset j recorded in the data are only those who invested in j an amount of money larger than a certain threshold. However, since this threshold is usually of the form λc_j (for example, $\lambda = 0.0005$ in the Italian data), the recorded shareholdings are those such that:

$$v_i c_j / \sum_{k=1}^{M} c_k > \lambda c_j$$

which is satisfied if the volume v_i is such that $v_i > \lambda \sum_k c_k$ independently of c_j. In other words, the shareholdings of an investor behaving according to the CAPM model are either all observed (if v_i is sufficiently large) or all unobserved. This means that, also taking the above bias into account, the observed $P(k^{in})$ would again be peaked at the single value $k = M$. Note that the above argument can also model the situation when an investor uses the CAPM model to estimate the optimal weights, and however chooses to invest in a given asset only if, by doing so, he will hold a significant (subjective) fraction λ of shares of that asset. These arguments, although quite simplified, suggest that the non-trivial form of the in-degree distribution can genuinely witness the deviation of the investors behaviour from the ideal scenario explored by the CAPM. The above analysis provides therefore additional evidence of the inadequacy of the model, whose predictions are already commonly thought to be unrealistic (for instance by analysing price trends, see the preceding section).

The ideal scenario at the basis of the CAPM model is that all agents are equally informed about the market and process this information in the same way. Clearly, both hypotheses are unrealistic and result in a series of aforementioned predictions which differ from the empirical findings. In the following section we try to relax the above hypotheses into a stochastic model allowing the agents to have heterogeneous availabilities of information and to make different choices even when equally informed. One basic idea of the model is that, since the investors in our data are large long-term investors (and not short-term speculators), their choices depend on the detailed (often private) information that they need to gather concerning, for instance, the budgets and management strategies of companies. For this information to be acquired, investors have to face significant costs. Clearly, diversifying a portfolio by adding the asset j in it is convenient when the cost of acquiring information about j is smaller than the expected profit (in terms of risk reduction) associated to j. As a consequence, large-volume portfolios (corresponding to holders who can face large information acquisition costs) are therefore likely to display a large diversification as well. The second ingredient of the model is to allow for two equally wealthy agents to make different choices (due for instance to different preferred investment sectors), even if assets with better expected long-term performance are statistically more likely to be chosen.

4.3 Fitness model driven by Pareto's law

The above simple ideas (heterogeneous choices of the investors and information acquisition costs) can be directly implemented by generalizing a recent model[10] of network formation to the directed case. We assume that the probability f_{ij} that the shareholder i invests in the asset j is a function of two quantities, namely the total volume v_i that i decides to invest and a second quantity y_j (which can also be a vector) characterizing j (such as its price history, expected trend, etc.). By assuming the simplest separable form

$$f_{ij} = g(v_i)h(y_j)$$

we can directly express the expected in-degree k^{in} of an investor with volume v as

$$k^{in}(v) = g(v)h_{tot}$$

where h_{tot} is the total value of $h(y)$ computed over all assets. If the above expression can be inverted to yield $v(k^{in})$, it is possible to compute the corresponding in-degree distribution:

$$P(k^{in}) = \theta[v(k^{in})]\frac{d}{dk^{in}}v(k^{in})$$

which clearly depends on the volume distribution $\theta(v)$. The corresponding expressions for $k^{out}(v)$ and $P(k^{out})$ can be easily obtained, however (as we mentioned above) our information concerning k^{out} is incomplete and we cannot therefore use it to validate the model. Similarly, any hypothesis on the form of $h(y)$ would yield results that cannot be tested on the data. For this reason, we avoid any rigorous definition of y and deal only with the quantities related to v.

We therefore proceed by suggesting an explicit form for the quantities $\theta(v)$ and $g(v)$. The volume distribution $\theta(v)$ is chosen to display a power-law tail as the observed one. We therefore set $\theta(v) \propto v^{-\alpha}$. The form of the attachment probability $g(v)$ can instead be chosen by analogy with the traditional *preferential attachment*[9] mechanism. In the latter scheme, the attachment probability is chosen as an increasing function of the pre-existing vertex degree. While the common choice in the models is the linear one[**?**], its functional form can be measured on real networks[**?**] and is found to be proportional to k^β. The case $\beta = 1$ is the *linear* preferential attachment case, while $\beta > 1$ and $\beta < 1$ are the *superlinear* and *sublinear* cases respectively. In the system under consideration, the analogous choice would be $g(v) \propto v^\beta$. Substituting $\theta(v) \propto v^{-\alpha}$ and $g(v) \propto v^\beta$ in the expression for $P(k^{in})$ yields (for large k^{in})

$$P(k^{in}) \propto (k^{in})^{-\gamma} \qquad \gamma = (\alpha + \beta - 1)/\beta$$

We therefore recover, as in the original model[10], that the scale-free degree distribution can be obtained by letting the connection probability f_{ij} depend on a power-law distributed quantity v.

An independent test of the model can be performed by noting that the above hypotheses also result in the following expectation for $k^{in}(v)$:

$$k^{in}(v) \propto v^{\beta}$$

where β, once the values of γ and α are fixed to the empirical ones, is constrained to the value $\beta = (1-\alpha)/(1-\gamma)$. In Fig. 11b we superimpose the above prediction to the points obtained from the data for each network. Indeed, the model expectations are rigorously verified by the data, except for the low k^{in} region in MIB. A possible explanation for this anomaly is that, as we checked, these points correspond to those investors holding a very large fraction (about 50%) of the shares of an asset, whose portfolio has therefore a large volume even if its diversification is small. Clearly, these investors are the effective controllers of a company. While in both US markets the fraction of links in the network corresponding to such a large weight is of the order of 10^{-4} (so that their effect is irrelevant on the plot of Fig. 11b), in MIB it equals the extraordinarily larger value 0.13. This determines the peak at small k^{in} superimposed to the power-law trend in the Italian market, and singles out an important difference between MIB and the US markets. This suggests that the proposed mechanism fits well the investors' behaviour, apart from that of the effective holders of a company.

Interestingly, in all cases $\beta > 1$, corresponding to a *superlinear* attachment mechanism. It is however worth noting that, while the traditional preferential attachment rule yields scale-free topologies only in the linear case[**?**], here we observe power-law degree distributions in the nonlinear case as well. This is a remarkable result, since in order to obtain the empirical forms of $P(k^{in})$ the exponent β does not need to be fine tuned, and the results are therefore more robust under modification of the model hypotheses.

4.4 Further topological features of the shareholder networks

Besides proposing a model for the portfolio diversification, we report the study of other topological features of the shareholder networks allowing to characterize the different markets under study.

Differently from social networks, shareholding networks present low or absent clusterisation. Only 1 %. of nodes have non null Clustering Coefficient in NYSE, versus 12 %. in MIB. In Nasdaq there is no clustering at all. In fact, US markets consist of few stars with hundreds of leaves, the centers typically being investors non quoted on the market.

Shareholding relationships, represented as arcs departing from the owned stock and pointing to the shareholder, are characterized by the following quantities: the in-degree or portfolio diversification K_1 of the stock (if this also a shareholder of other stocks), the in-degree K_2 of the shareholder, the market capitalization C_1 of stock, the market capitalization C_2 of shareholder (if this is also a quoted company), the percentage of shares W owned by the shareholder.

Distribution of the amount of shares W are shown in **figure 12.** The scale is linear-log. NYSE and NASD have similar distributions, i.e. initial exponential decay, then power law. MIB shows a bump for values of shares just above 50 %, meaning that a fraction of stocks are controlled each by a single holder. This is another obvious difference between the US markets and the Italian market. Correlation between amount of share and portfolio size will make this difference more clear.

The average neighbor degree Knn as a function of K reveals (data not shown) that holders with larger portfolio size tend to own shares of holders with smaller portfolio size. This is common to the three market with the difference that portfolio size ranges up to 19 for MIB SSN and up to around 200 in NASDAQ and NYSE SSN's.

The distribution $P(K_2, W)$ of the number of shareholding relationships involving a shareholder with portfolio size K_2 and amount of shares W are shown in **Figure 13** for NYSE and MIB SN. In US markets largest shares are held by holders of any in-degree. In MIB largest shares are held by holders with low in-degree, meaning that holders that have total control of a company tend to own few stocks. In particular the picture shows an obvious cluster of relationships in which the holder owns more than 40% of the shares but has a low portfolio diversification. The cluster involves the 24% of the shareholders. These holders are typically (95%) non quoted companies. This means that there is a consistent fraction of shareholders which control a quoted company and are not themselves quoted on the market.

In all three markets there is a trend for companies to own companies with smaller market capitalization (data not shown). We don't find any correlation between market capitalization and amount of owned shares. Values of market capitalization were available only for quoted companies and not for all holders.

4.5 Effective Control Indexes

Until now we dealt with a weighted nature of the complex networks under study, but we didn't take into account the relative importance of a shareholder of a stock with respect to the other shareholders of that same stock. It is clear that the concentration of the ownerships plays a crucial role in financial strategy. We thus compute two indexes to capture the fact that a 10 % shareholder holds much more control if the other shareholders hold 1 per cent each, than if they hold 10 % each. This information is not contained in the amount of share alone nor in the W distribution over all nodes. We define the following quantities. For each stock SI:

$$SI = \frac{\sum_{i \in holders} w_i^2}{(\sum_{i \in holders} w_i)^2} \tag{6}$$

gives the effective number of holders. SI is close to 1 when there is a dominating holder. SI is equal to 1/N when there are N equally important holders. For each holder j and each stock i we compute:

$$h = \frac{w_{ij}^2}{(\sum_{k \in holders} w_{ik})^2} \tag{7}$$

Then for each holder we sum the above quantity for each of the stocks in his portfolio.

$$HI = \frac{\sum_{i \in stocks.owned.by.j} w_{ij}^2}{(\sum_{k \in holders.of.stock.i} w_{ik})^2} \tag{8}$$

HI gives the effective number of stocks controlled by a holder.

We report the distributions of SI and HI in **Figures 14, 15, 16**. While in the US markets SI distribution has an exponential decay, in MIB SI distribution is flat. As for HI, the distribution decreases very fast in US market: seemingly first exponentially then power-law. In MIB there is a nearly flat region in HI distribution. These results shows that in MIB the concentration of power among holders is distributed in a very different way from US markets. In MIB there as many companies controlled by a single holder than by several ones. In the US markets the large majority of companies is controlled by more than 5 holders.

Finally, in order to investigated whether holders controlling effectively several stocks tend to have special topological properties in the network we have searched for possible correlation between the indexes defined above and properties as in-degree, cluster coefficient and betweenness centrality. We find that SI tend to decrease with the value of betweenness centrality. This means that companies that are more central tend to be owned more evenly by several holders. To better take into account the fact that some links are weak and other are strong, we define a weighted version of the cluster coefficient (CCW). CCW is the sum of the weights (in [0 1]) of links, if any, between neighbors of a given node, normalized by the total number of possible links among them. In MIB we find a positive correlation between HI and ccw which suggest that holders that holds stocks linked by shareholding relationships tend to be able to accumulate effective control over those companies. As we look only at the in-degree in computing CCW, this means that if a holder A has positive ccw, then a company B in A's portfolio owns share of another company C in A's portfolio. But this means that on top of the shares of C, A has an indirect control of C through B. The indirect control could be the reason for the ability to concentrate power.

This result sounds fascinating because it shows a relationships between a local measure of power and a topological property. These two measures are independent. HI measures how much a holder actually counts in how many stocks. ccw is a measure of how strong the neighbors of a holder are related. Due to the small number of nodes with ccw>0, this measure should be repeated on other data sets.

4.6 Concluding remarks on shareholder networks

We proposed a topological characterization of the shareholdings corresponding to large investments in three real markets. We found that in all cases power-law distributions describe purely topological (portfolio diversification) as well as weighted (portfolio volume) quantities. A nontrivial scaling relation between these quantities holds. We finally recovered the observed properties within a simple stochastic model by assuming that the investment probability depends on the investor's wealth. In a network context, the above results support the hypothesis that Paretian distributions of non-topological quantities associated to the vertices may be at the basis of the emergence of scale-free topologies in a large number of real networks.

The analysis of the correlation between portfolio size and amount of shares has revealed that in the MIB network largest shares tend to be owned by holder with small portfolio. In the same network we have also found positive correlation between the number of companies effectively controlled by a holder and the cluster coefficient of the holder.

5 Acknowledgements

We would like to thank G. Bonanno, M. Castri, F. Lillo, R. Mantegna, L. Pietronero and E. Sciubba. Authors also acknowledge support from EU FET Open Project IST-2001-33555 COSIN.

References

1. Albert, R. and Barabsi, A.-L. Statistical mechanics of complex networks. Rev. Mod. Phys. 74, 47-97 (2002).
2. D. J. Watts and Strogatz, Nature **393**, 440 (1998).
3. P. Erdős, A. Rényi, Bull. Inst. Int. Stat. **38**, 343 (1961).
4. B.A. Huberman and L.A. Adamic *Nature* **399**, 130 (1999).
5. R. Albert, H. Jeong, and A. L. Barabási Nature **401**, 130 (1999).
6. G. Caldarelli, R. Marchetti and L.Pietronero, Europhysics Letters **52**, 386 (2000).
7. R. Pastor-Satorras, A. Vazquez and A. Vespignani, Phys. Rev. Lett. **87**, 258701 (2001).
8. M. E. J. Newman, D. J. Watts, and S. H. Strogatz, Proc. Natl. Acad. Sci. USA **99**, 2566 (2002).
9. A. L. Barabási and R. Albert, Science **286**, 509 (1999)

10. G. Caldarelli, A. Capocci, P. De Los Rios and M.A. Muñoz, Phys. Rev. Lett. 89, 278701 (2002)
11. M.E.J. Newman, M. Girvan *to appear in Proceedings of the XVIII Sitges Conference on Statistical Mechanic* **99**, 12583 (2003).
12. M.Boguna, R. Pastor-Satorras, A. Vespignani *ArXiv:cond-mat/0301149).*
13. M.E.J. Newman *Phys. Rev.* E 67, 026126 (2003).
14. G. Bianconi and A.-L. Barabási *Europhysics Letters* **54**, 436 (2001).
15. G. Caldarelli, A. Capocci, P. De Los Rios and M.A. Muñoz, *Physical Review Letters* 89, 258702 (2002).
16. D.S. Callaway, J.E. Hopcroft, J.M. Kleinberg, M.E.J. Newman and S.H. Strogatz *Phys.Rev.E* **64**, 041902 (2001).
17. Davis, G.F., Yoo, M., Baker, W.E., The small world of the American corporate elite, 1982-2001, *Strategic Organization* 1: 301-326 (2003).
18. M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Random graphs with arbitrary degree distributions and their applications, Phys. Rev. E 64, 026118 (2001).
19. M. E. J. Newman, Assortative mixing in networks, Phys. Rev. Lett. 89, 208701 (2002).
20. M. E. J. Newman and Juyong Park, Why social networks are different from other types of networks, Phys. Rev. E, in press.
21. M. Catanzaro, G. Caldarelli, L. Pietronero, Assortative model for social networks, cond-mat 0308073 v1
22. Battiston, S., Bonabeau, E., Weisbuch G., Decision making dynamics in corporate boards, *Physica A*, 322, 567 (2003).
23. Battiston, S., Weisbuch G., Bonabeau, E., Decision spread in the corporate board network, submitted.
24. Y. J. Campbell, A. W. Lo, A. C. Mackinlay *The Econometrics of Financial Markets*, (Princeton University Press, Princeton,1997) and references therein.
25. N Vandewalle, F Brisbois and X Tordoir Quantitative Finance **1**, 372 (2001).
26. K. V. Mardia, J. T. Kent and J. M. Bibby *Multivariate Analisys*, (CA: Academic, San Diego, 1979).
27. J. C. Gower, Biometrika **53**, 325 (1966).
28. R. N. Mantegna, Eur. Phys. J. B **11**, 193 (1999).
29. G. Bonanno, F. Lillo and R. N. Mantegna, Quantitative Finance **1**, 96 (2001).
30. The Standard Industrial Classification system can be found at http://www.osha.gov/oshstats/naics-manual.html
31. L. Laloux, P. Cizeau, J. P. Bouchaud and M. Potters, Phys. Rev. Letters **83**, 1467 (1999).
32. V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. Nunes Amaral and H. E. Stanley, Phys. Rev. Lett. **83**, 1471 (1999).
33. I. Rodriguez-Iturbe and A. Rinaldo, *Fractal River Basins*, (Cambridge University Press, Cambridge, 1997).
34. J.-P. Onnela, A. Chackraborti, K. Kaski, J. Kertsz ArXiv:cond-mat/0303579 and ArXiv:cond-mat/0302546
35. M. D. Penrose, The Annals of Probability **24**, 1903 (1996).
36. T.E. Harris, *The Theory of Branching Processes* (Dover, New York, 1989).
37. V. Pareto *Cours d'Économie Politique* (Macmillan, London, 1897). Reprinted in *Oeuvres Complétes* (Droz, Geneva, 1965).
38. W.W. Badger, *Mathematical models as a tool for the social science* (Gordon and Breach, New York, 1980).

39. C. Dagum, & M. Zenga, (eds.) *Income and Wealth Distribution, Inequality and Poverty* (Springer-Verlag, Berlin, 1990).
40. J. J. Persky, Pareto's law. Journal of Economic Perspectives **6**, 181-192 (1992).
41. Yook, S. H., Jeong, H., Barabsi, A.-L. and Tu, Y. Weighted evolving networks. Phys. Rev. Lett. 86, 5835-5838 (2001).
42. H. Markovitz, *Portfolio Selection: Efficient Diversification of Investments* (Wiley, New York, 1959).

Table 1. Results for assortativity coefficient r and clustering coefficient cc in the Italian director network in 1986 and 2002 (D86, D02)) and the Italian board network (B86, B02)

	$D86$	$D02$	$B86$	$B02$
r	0.131	0.121	0.250	0.322
cc	0.899	0.915	0.356	0.318

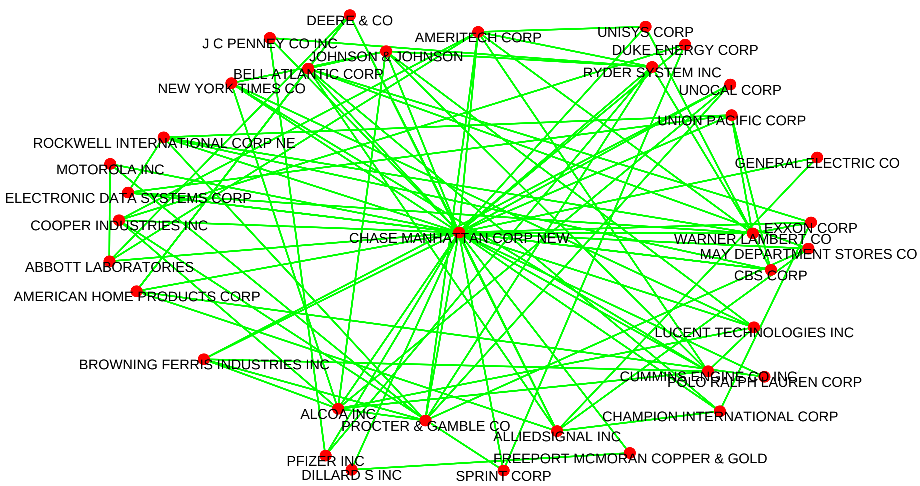

Fig. 1. The network of boards 1 degree of separation away from Chase Manhattan Bank's board.

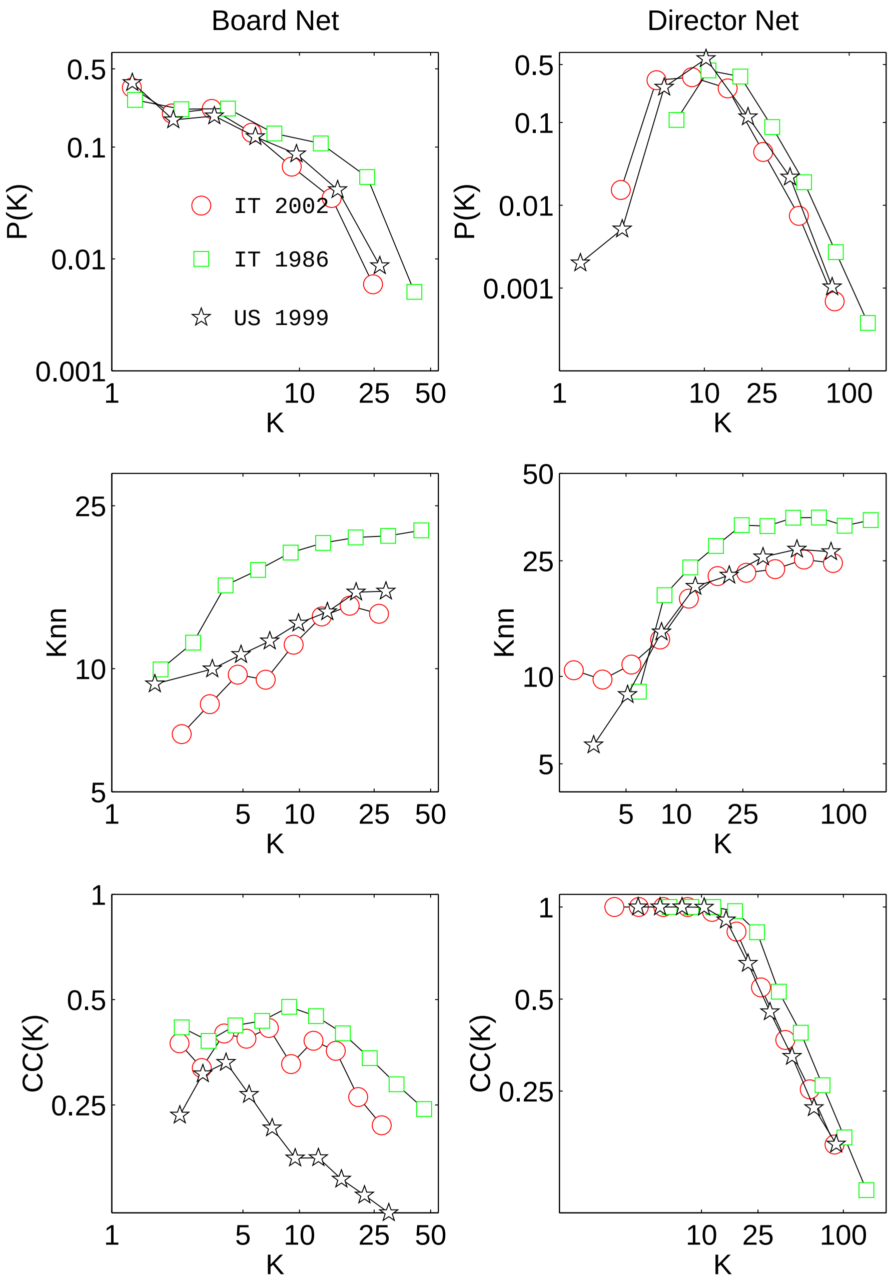

Fig. 2. Comparison of some properties of board (**left**) and director (**right**) networks in Fortune 1000 and MIB.
Top: degree distribution. **Middle**: average neighbors degree Knn as a function of the degree of the nodes. **Bottom**: clustering coefficient as function of the degree of the nodes

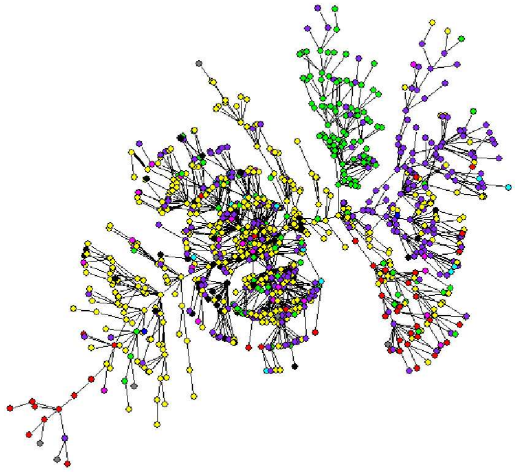

Fig. 3. Correlation based minimal spanning tree of real data from daily stock returns of 1071 stocks for the 12-year period 1987-1998 (3030 trading days). The node color is based on Standard Industrial Classification system. The correspondence is: red for mining - cyan for construction - yellow for manufacturing - green for transportation, communications, electric, gas and sanitary services - magenta for wholesale trade - black for retail trade - purple for finance, insurance and real estate - orange for service industries - light blue for public administration

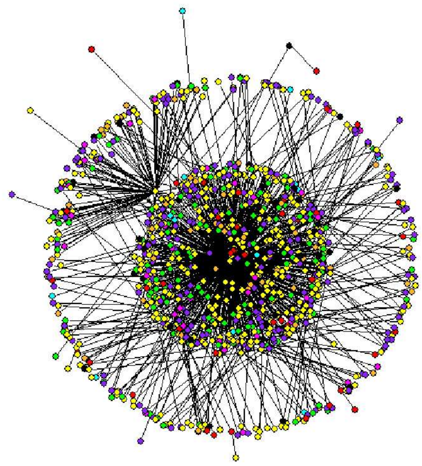

Fig. 4. Correlation based minimal spanning tree of a numerical simulation of the one factor model. The color codes are those used in Fig.1

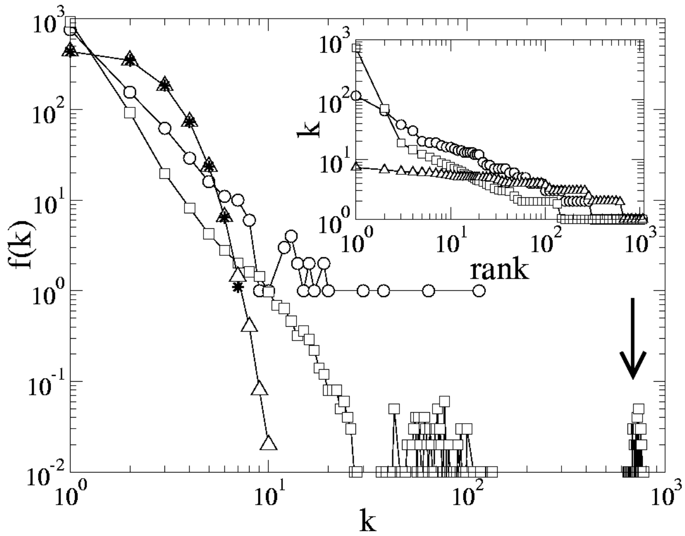

Fig. 5. Frequency distribution of the degree of the MST of real data (circle). We also show the mean degree distribution of random (triangle) and one-factor (square) model averaged over 100 numerical realizations of the MST. The stars are the theoretical values of the degree frequency for the random model in mean field limit. The inset shows the corresponding rank plot of the degree in the three cases.

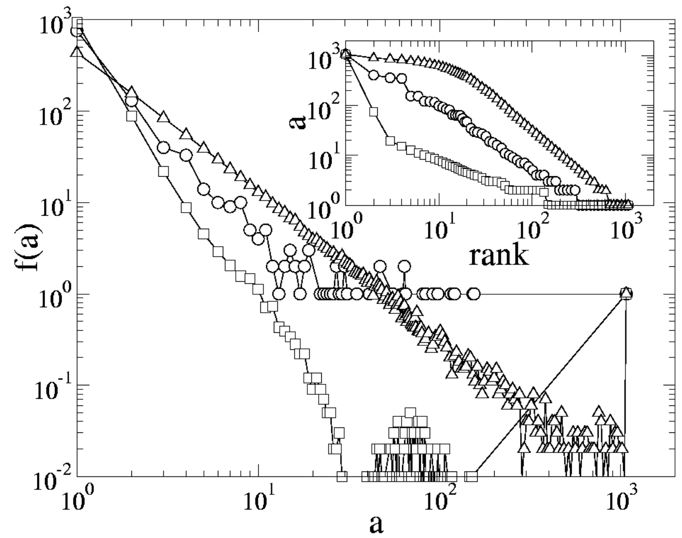

Fig. 6. Frequency distribution of the in-degree component of the MST of real data (circle). We also show the mean in-degree component distribution of random (triangle) and one-factor (square) model averaged over 100 numerical realizations of the MST. The inset shows the corresponding rank plot of the in-degree component for the three cases.

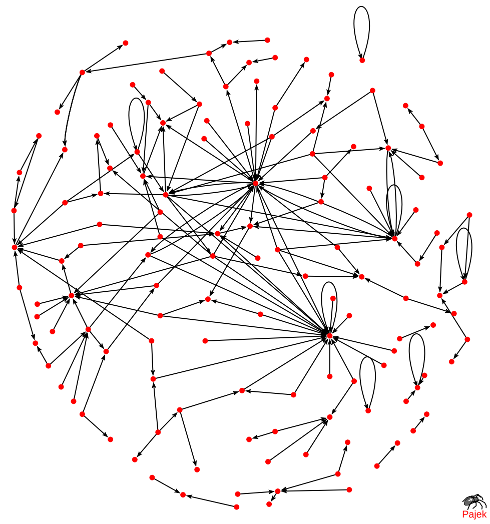

Fig. 7. Shareholder network for the Italian case.

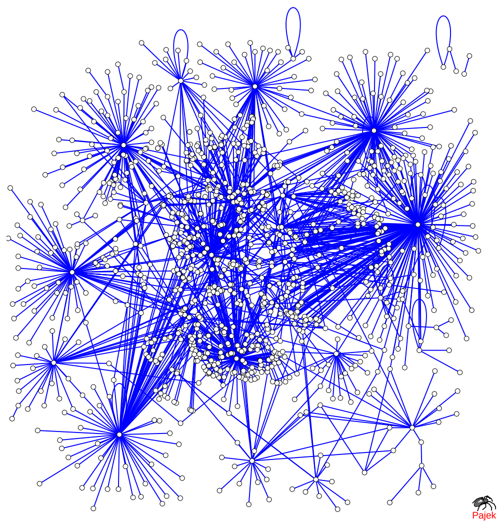

Fig. 8. Shareholder network for the NYSE case.

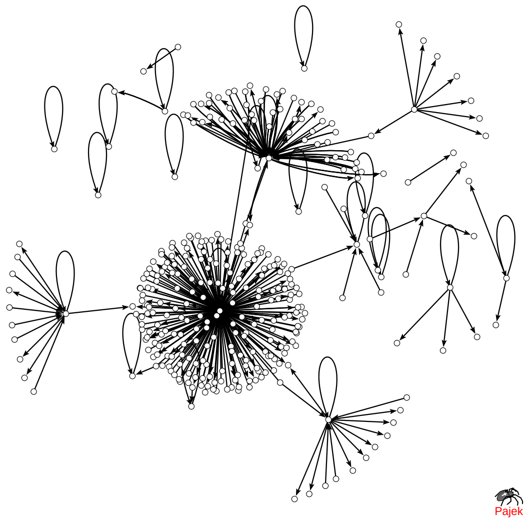

Fig. 9. Shareholder network for the Nasdaq case.

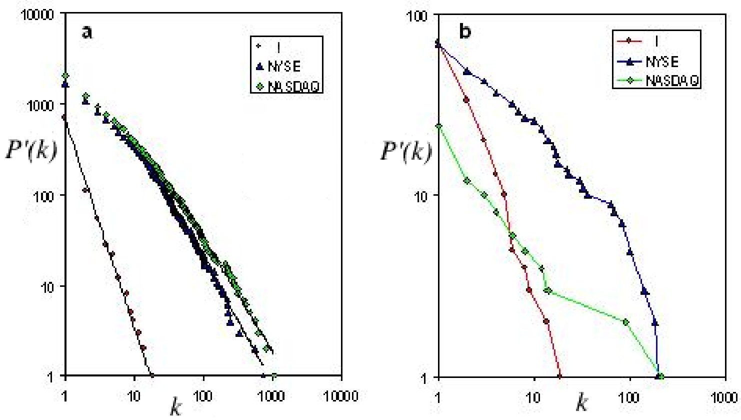

Fig. 10. Integrated degree distribution for the extended networks (**a** on the left) and the reduced one (**b** on the right).

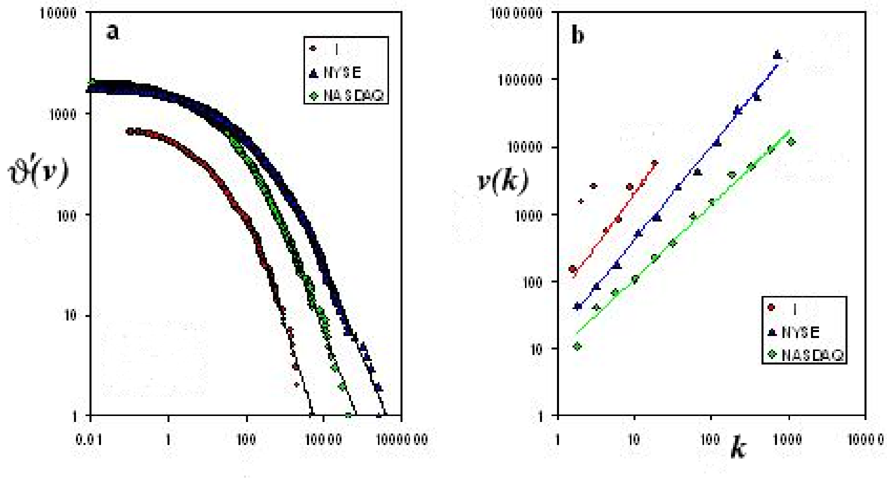

Fig. 11. Empirical properties of the portfolio volume. **a**) Cumulative volume distributions computed on the extended nets. **b**) Scaling of v against k^{in} in real networks (data points) and the power-law trend expected from the model (lines).

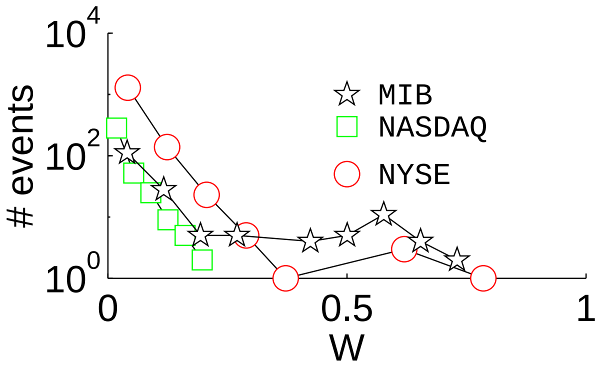

Fig. 12. Distribution of the amount of shares in the three shareholding networks

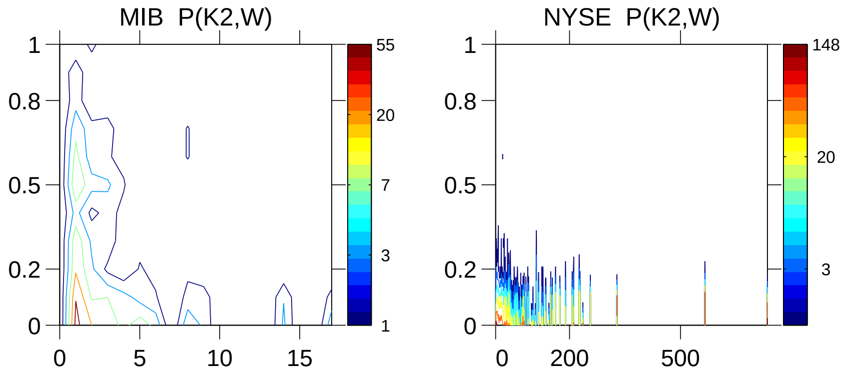

Fig. 13. horizontal axis: in-degree K_2 of the node at the end of an arc (the portfolio size of the holder), vertical axis: arc weight W (the amount of shares). P(K2,W) is the number of arcs entering a node with in-degree K_2 and weight W. **left**: distribution of P(K2,W) for NYSE. **right**: distribution of P(K2,W) for MIB. Logarithmic color scale.

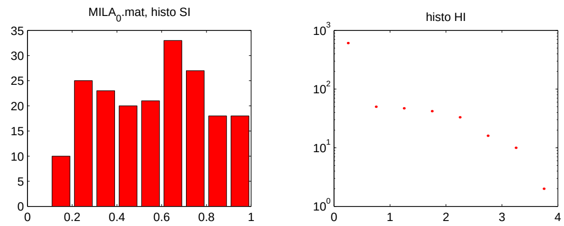

Fig. 14. MIB shareholding network.
SI measures the inverse of the effective number of holders for a stock. HI measures the number of stocks effectively controlled by a holder.

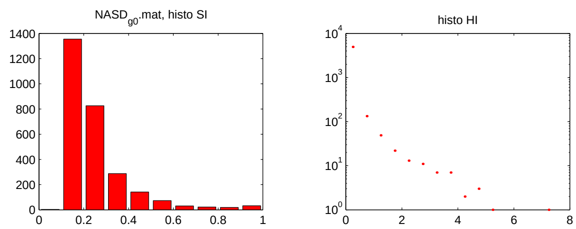

Fig. 15. NASDAQ shareholding network. SI and HI as in fig.14

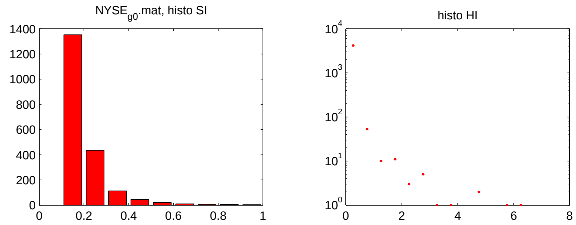

Fig. 16. NYSE shareholding network. SI and HI as in fig.14