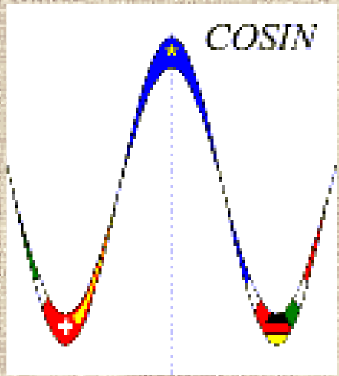
Minimal spanning tree networks in real and artificial financial markets

Overview

- Hierarchical structures can be detected in financial portfolios;
- One filtering procedure is a correlation-based clustering procedure;
- I discuss examples observed in the equity market and for financial markets located around the world;
- Artificial market models;
- Topology of MSTs in empirical data and market models;
- Conclusion.

The Collaboration

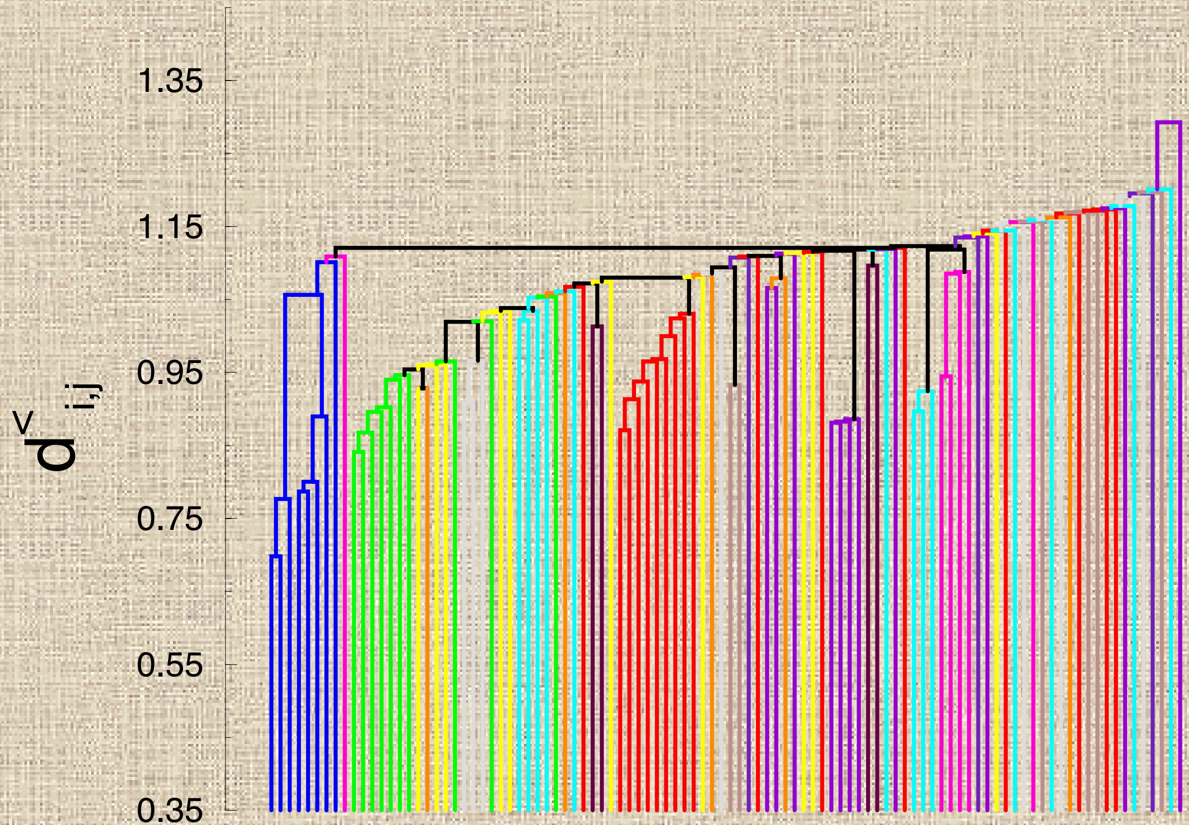


- Giovanni Bonanno
- Guido Caldarelli

- Fabrizio Lillo
- Salvatore Miccichè
- Rosario N. Mantegna

Hierarchical tree

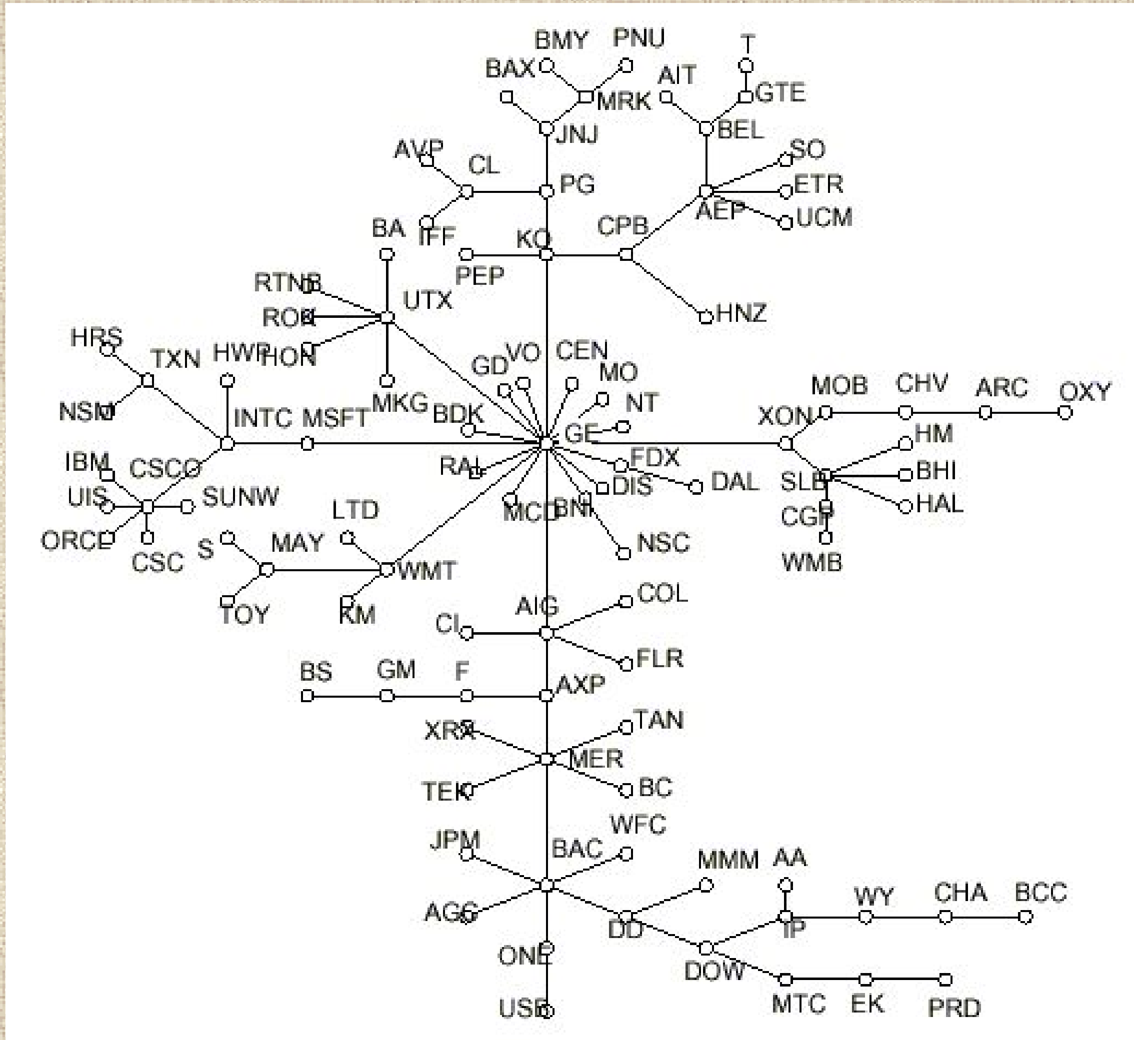
A meaningful taxonomy can be extracted from returns time series by using a correlation-based clustering procedure†:



A hierarchical tree is obtained

†R.N. Mantegna, Eur. Phys. J. B 11, 193-197 (1999)

Minimal spanning tree

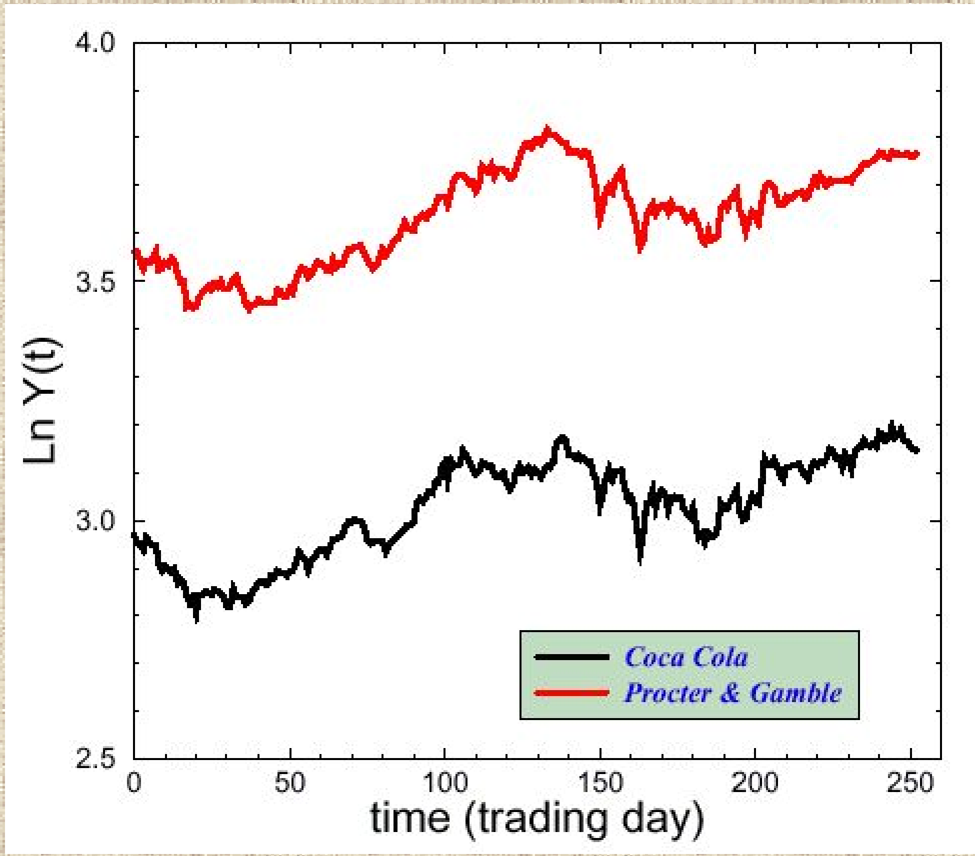


Together with a minimal spanning tree

This is the set of the 100 most capitalized US stocks in 1998.

Cross-correlation

Cross-correlation between stock returns are well-known

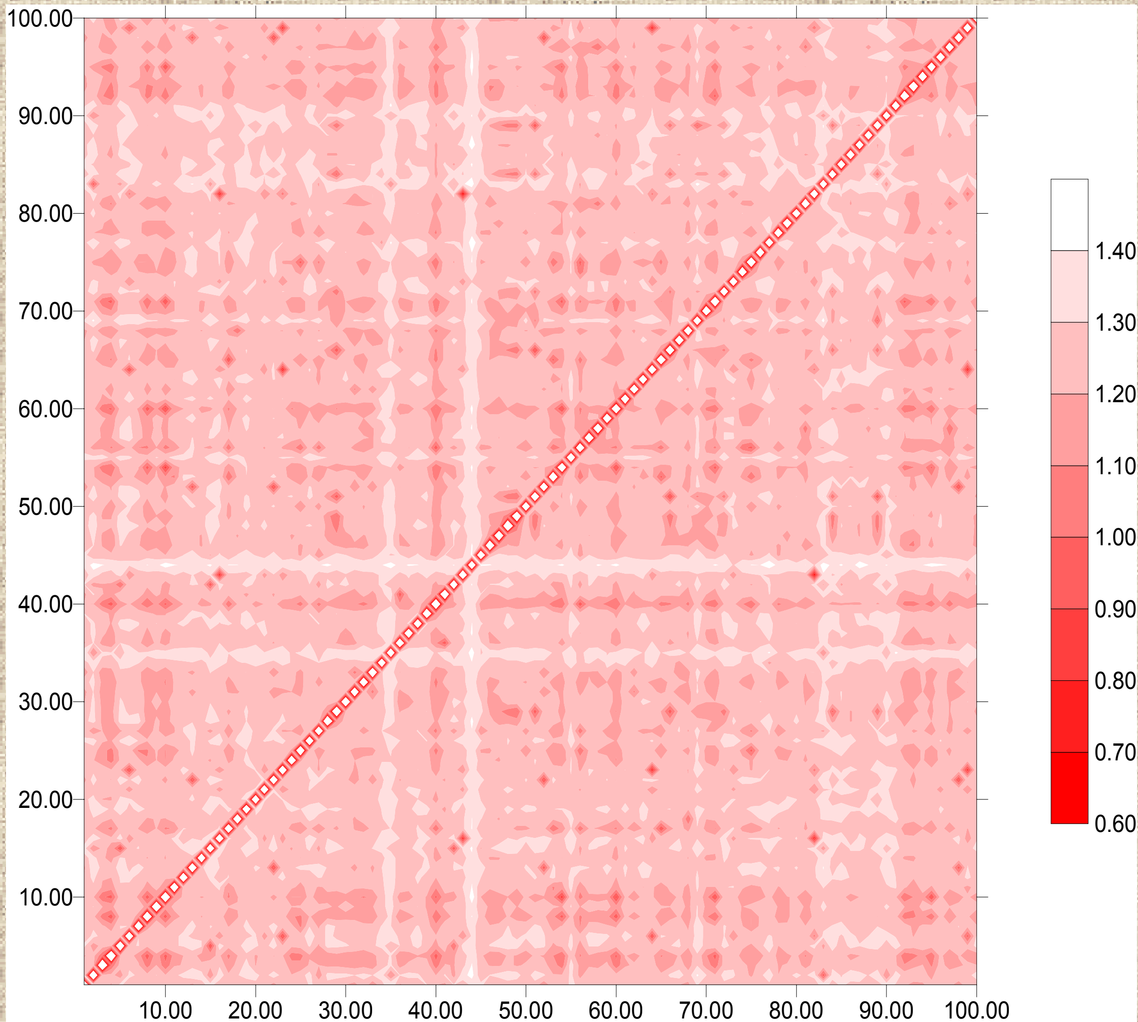


$$S_i \equiv \ln Y_i(t) - \ln Y_i(t-1) \cong R_i(t)$$

$$\rho_{ij} = \frac{\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle}{\sqrt{\left\langle S_i^2 - \langle S_i \rangle^2 \right\rangle \left\langle S_j^2 - \langle S_j \rangle^2 \right\rangle}}$$

They may be quantified by the correlation coefficient ρ_{ij}

A metric similarity measure



The function

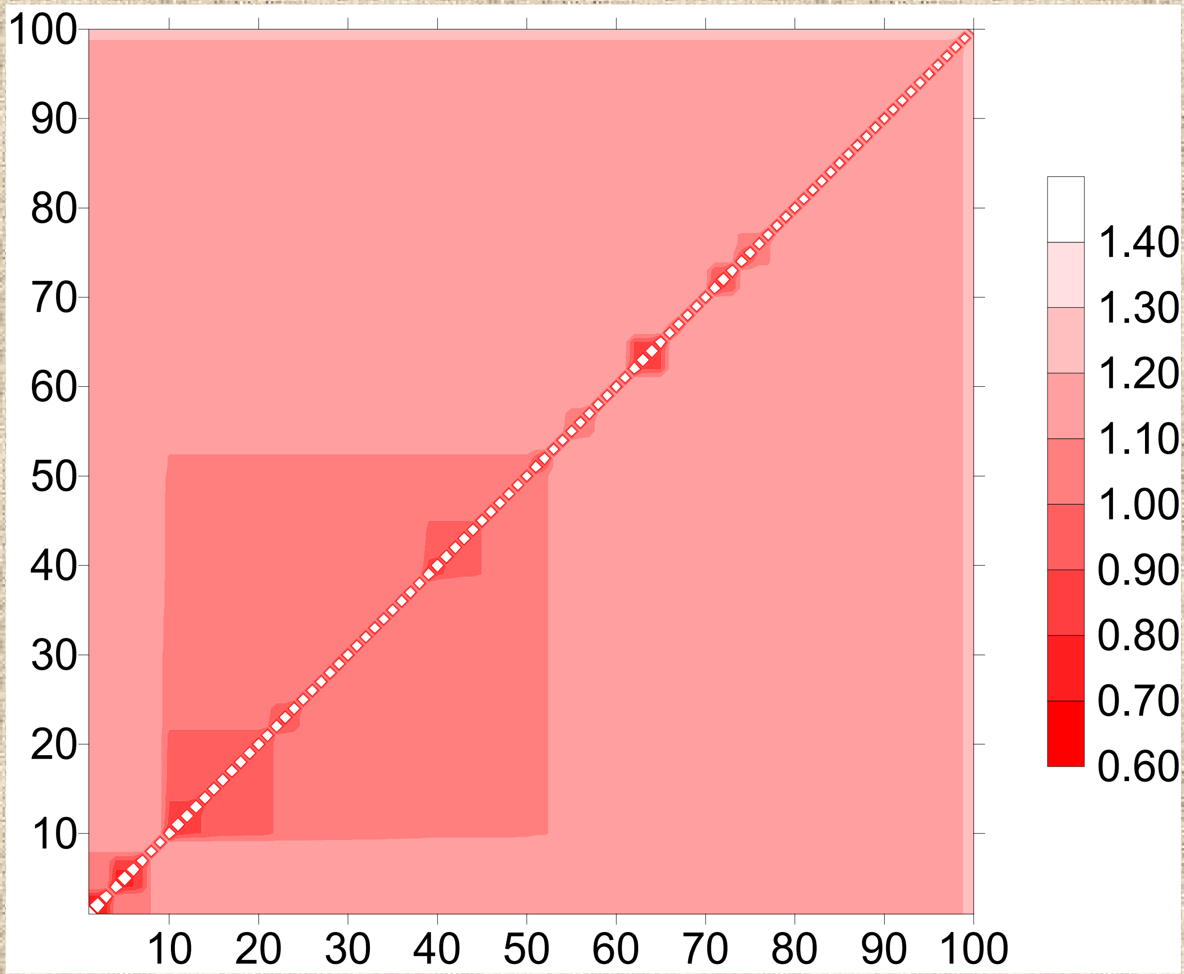
$$d_{ij} = \sqrt{2(1-\rho_{ij})}$$

verifies the axioms of a **metric distance.**

Filtering procedure

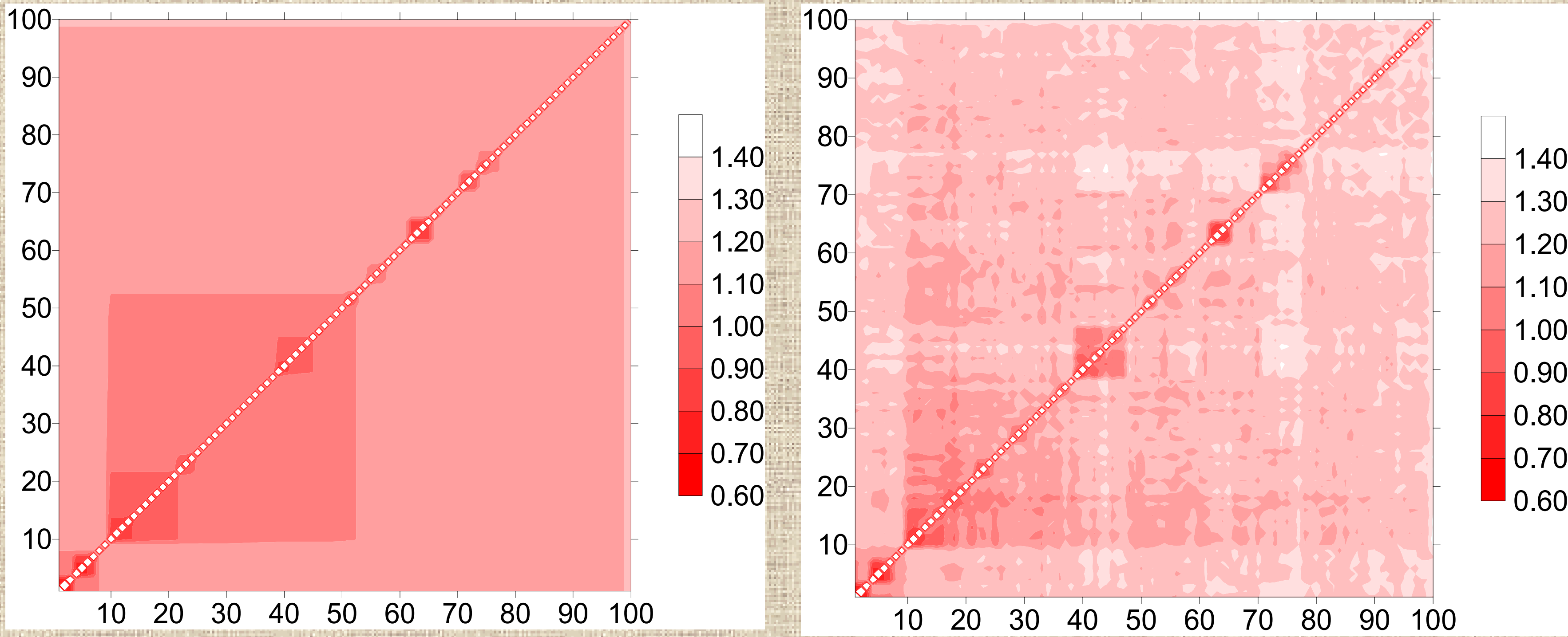
The use of the single linkage clustering method using the distance d_{ij} as a similarity measure is one possible filtering procedure selecting information about the hierarchical structure of the portfolio

The ultrametric distance matrix



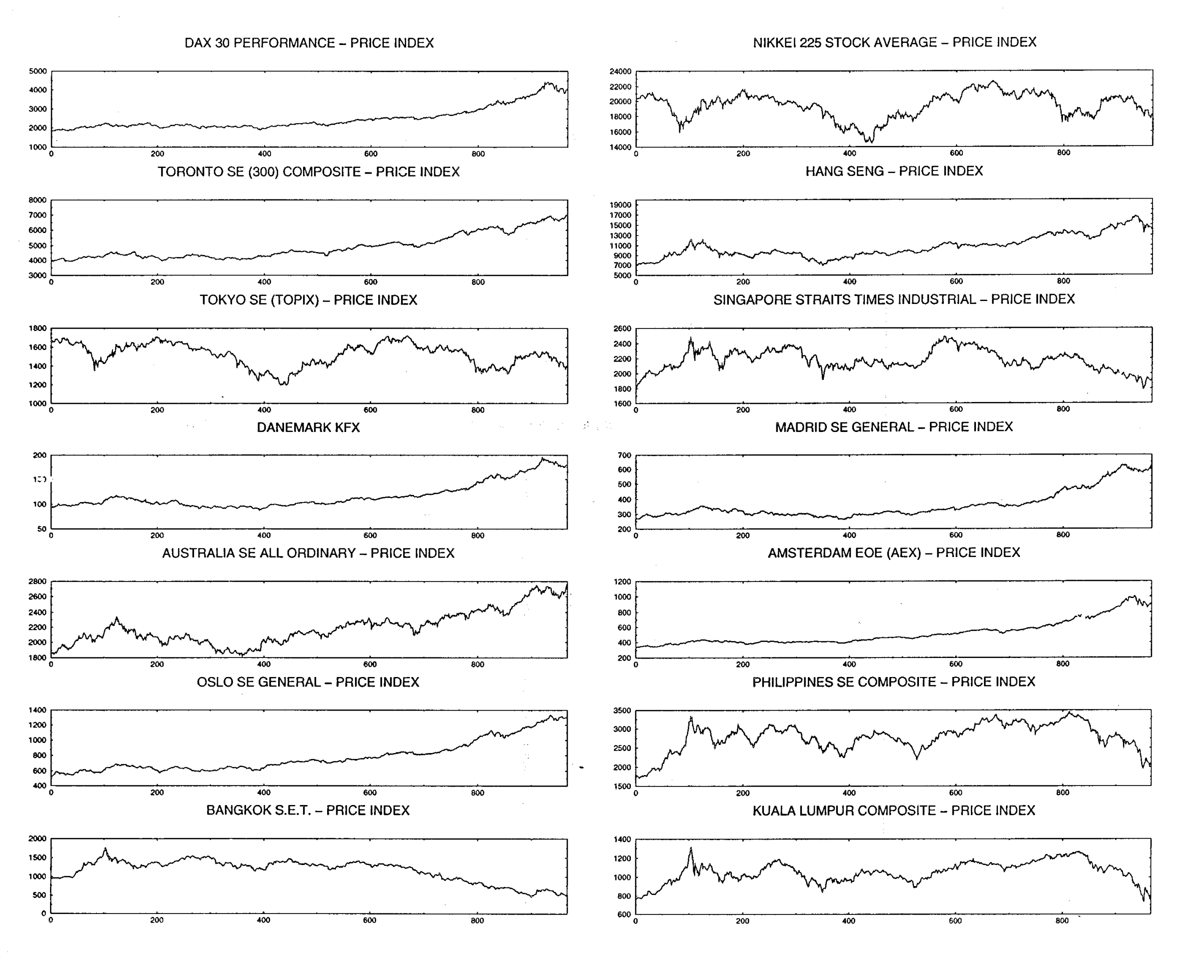
This filtering is rather severe by retaining only (n-1) correlation coefficients from the original n (n-1)/2

The (ordered) distance matrix



A direct comparison shows the amount of information not considered after the filtering procedure

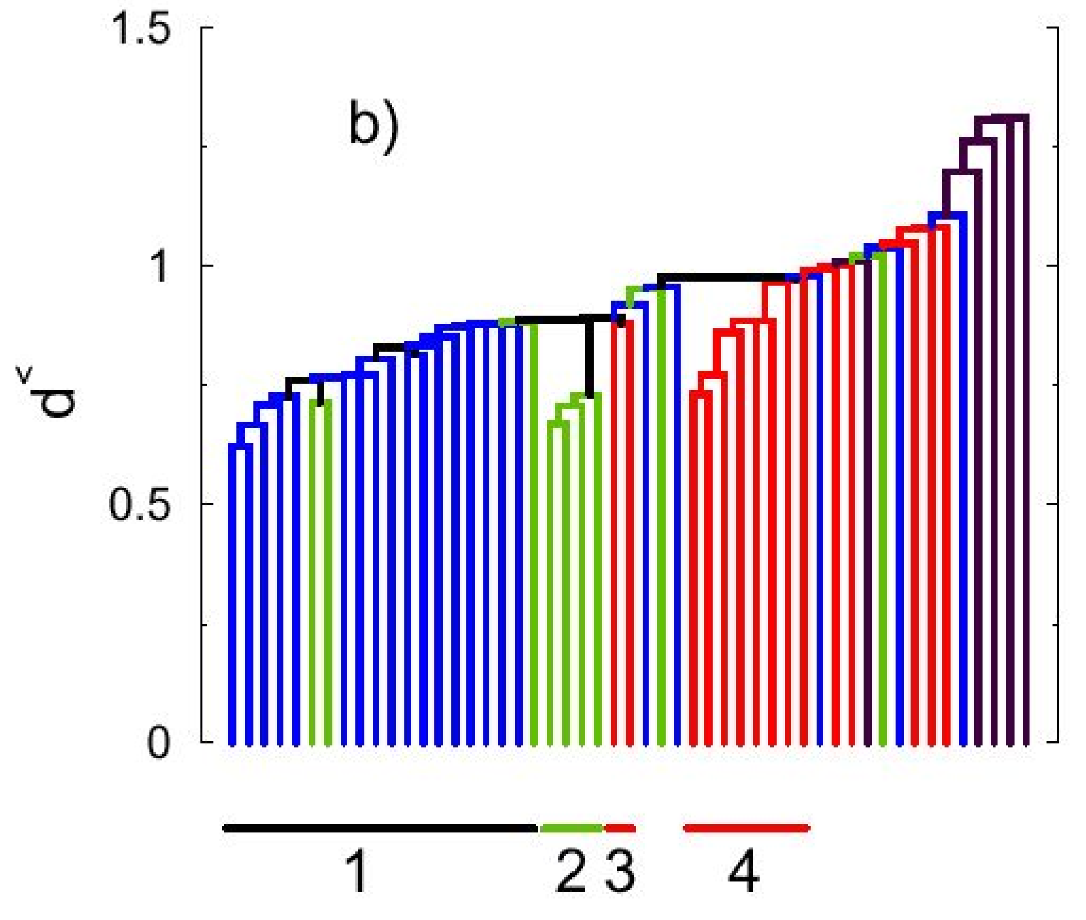
The global financial market



We investigate 51 MSCI stock indices both in local currency and in US dollars

[1] G. Bonanno, N. Vandewalle and R. N. Mantegna, Physical Review E**62**, R7615 (2000)

The global financial market

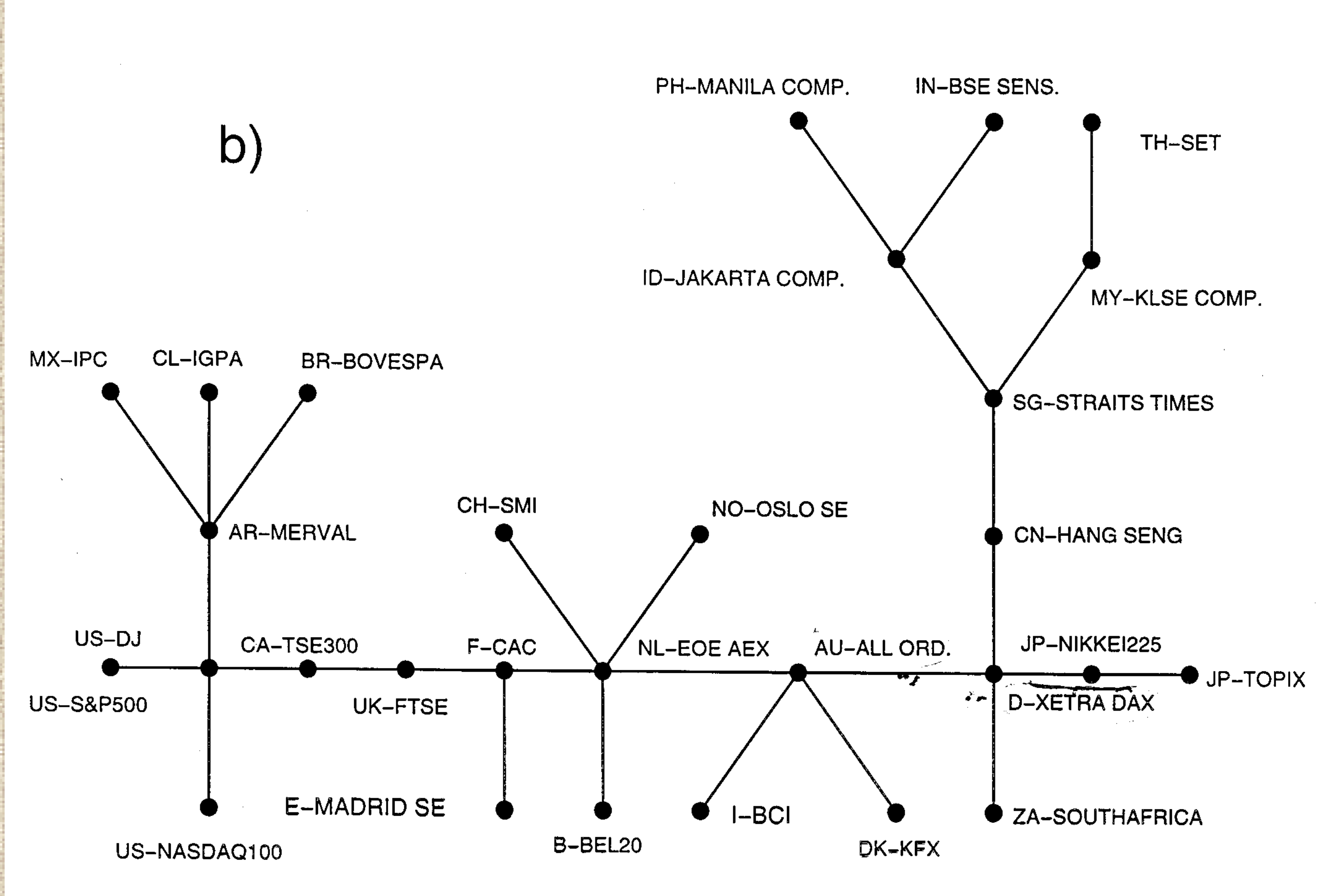


A hierarchical tree showing regional relations among stock indices of major stock exchanges has be detected[1]

1 **US**, **Canada** and **Europe**;
2 **South-America**;
3 **Japan**; and
4 **Asia**

[1] G. Bonanno, N. Vandewalle and R. N. Mantegna, Physical Review E**62**, R7615 (2000)

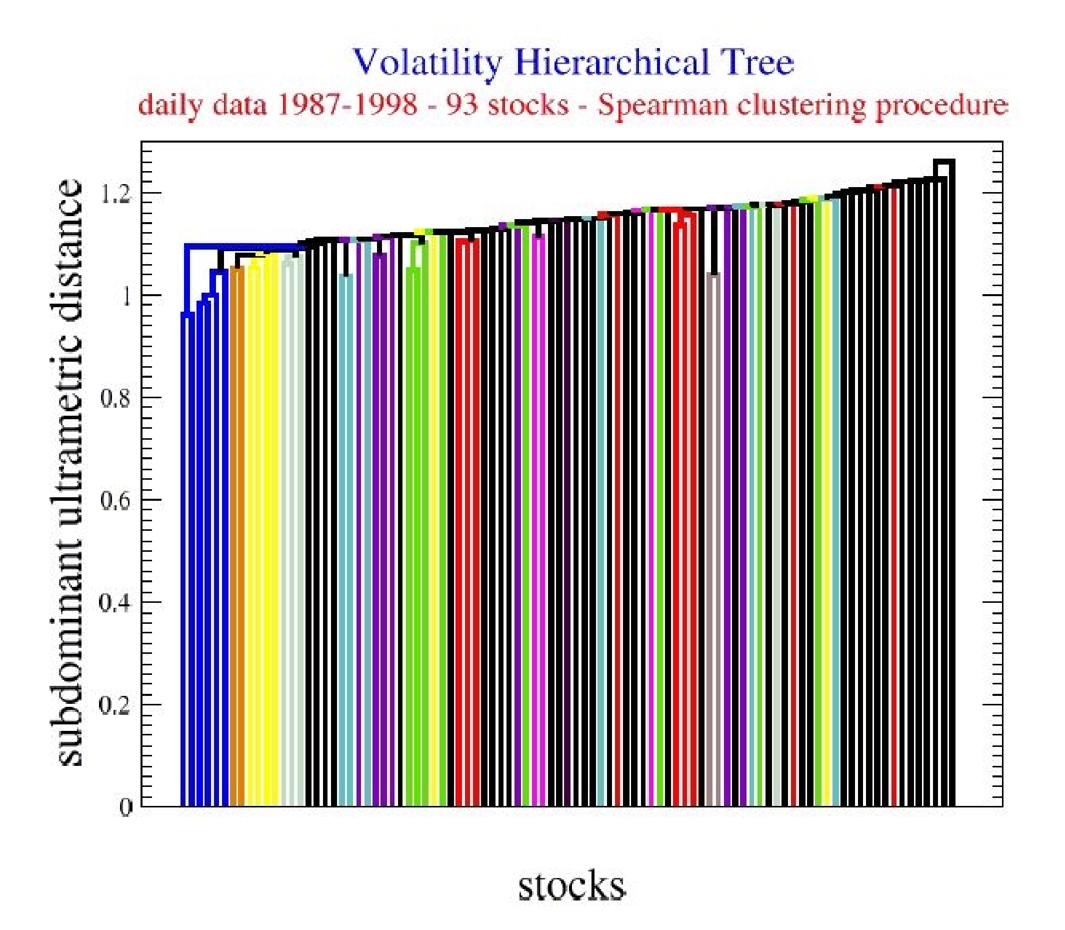
The global financial market



1996

MSTs show a clear regional clustering

Volatility

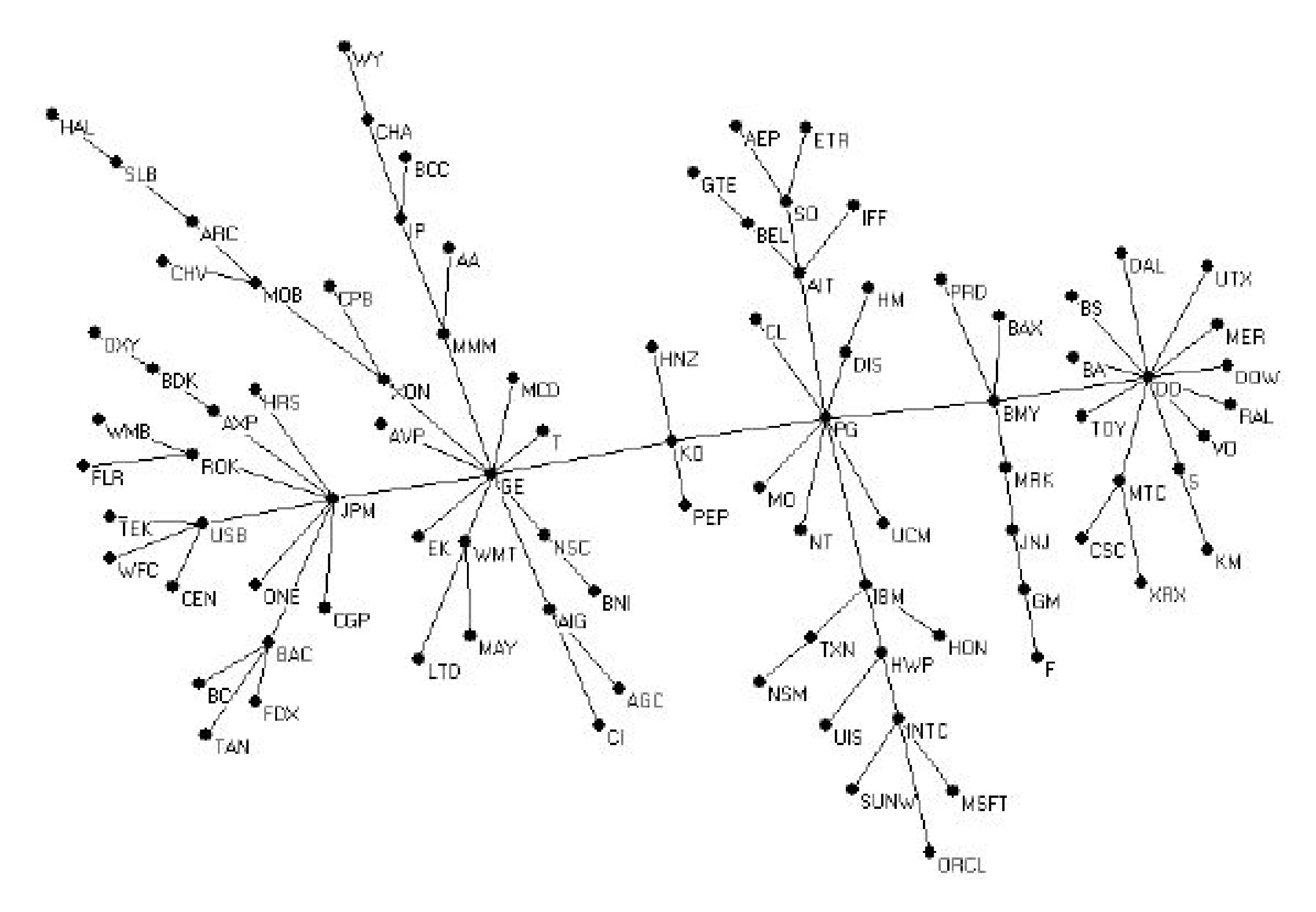


The same approach can of course be applied also at volatility time series[1]

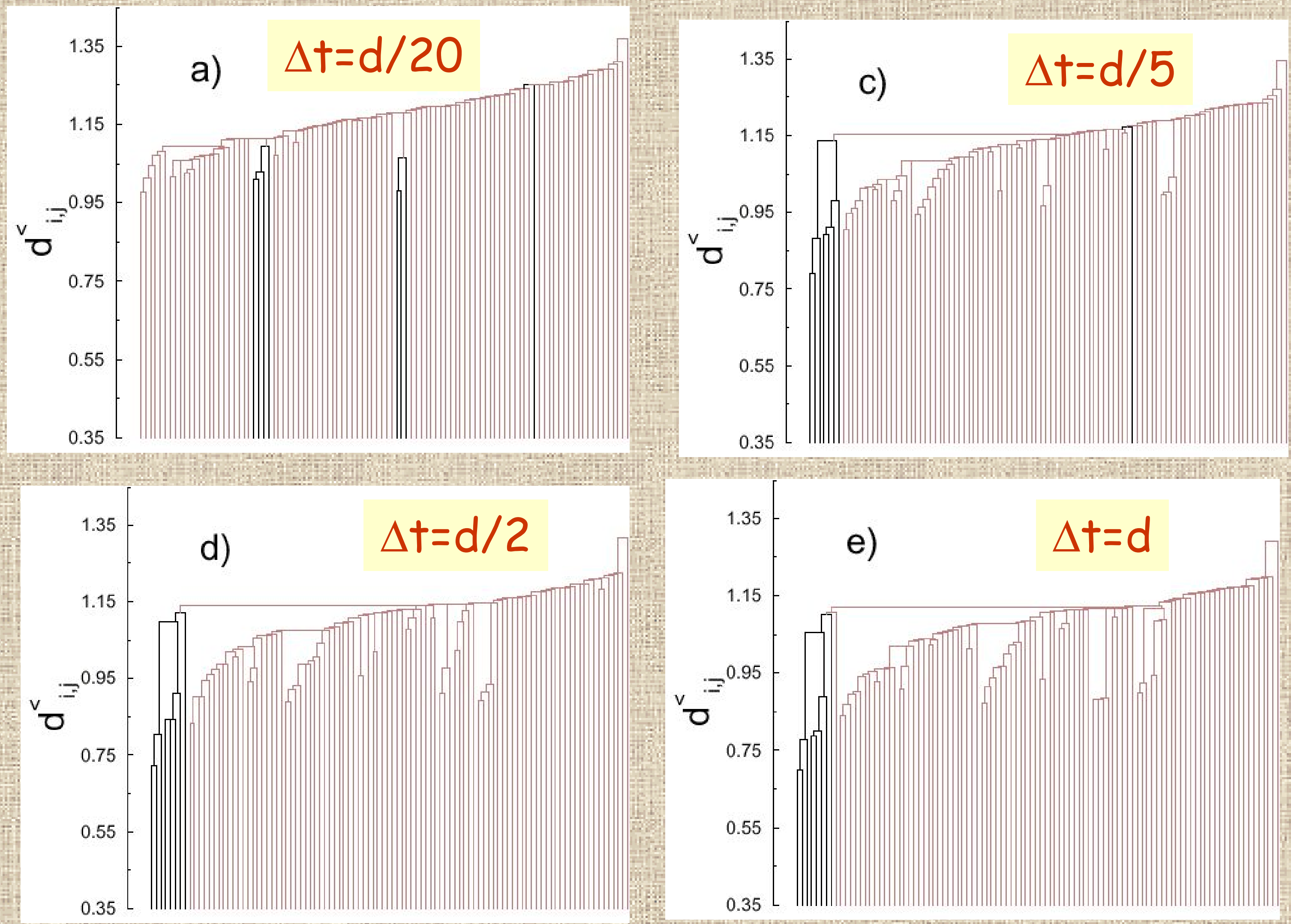
The set investigated is the set of 93 highly capitalized stocks traded at NYSE during the time period 87-98

1 Salvatore Miccichè Giovanni Bonanno, Fabrizio Lillo, Rosario N. Mantegna, Physica A **324**, 66-73 (2003)

Associated volatility MST

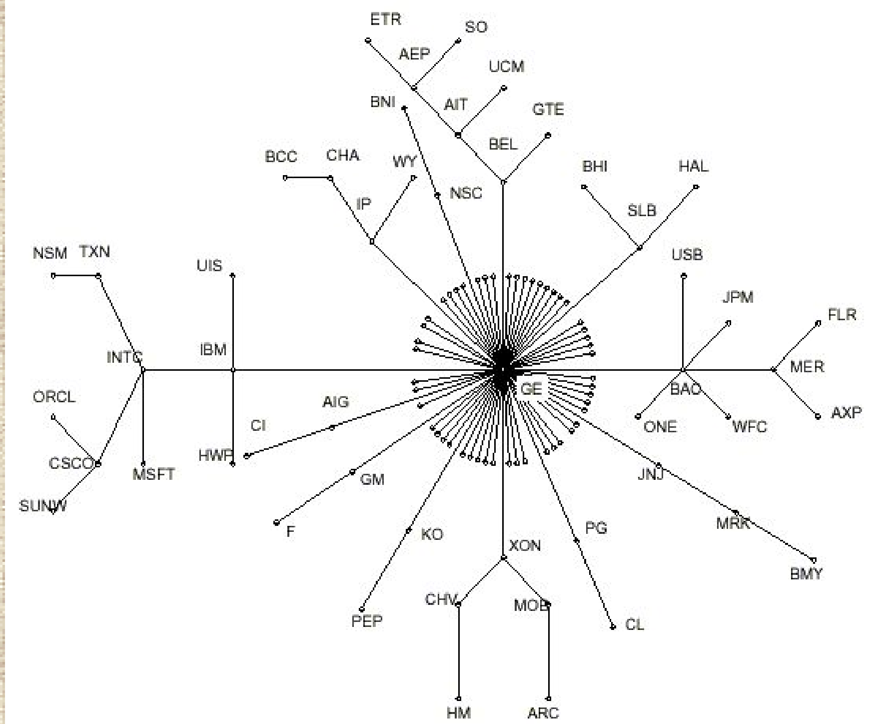


High frequency data† (stocks)

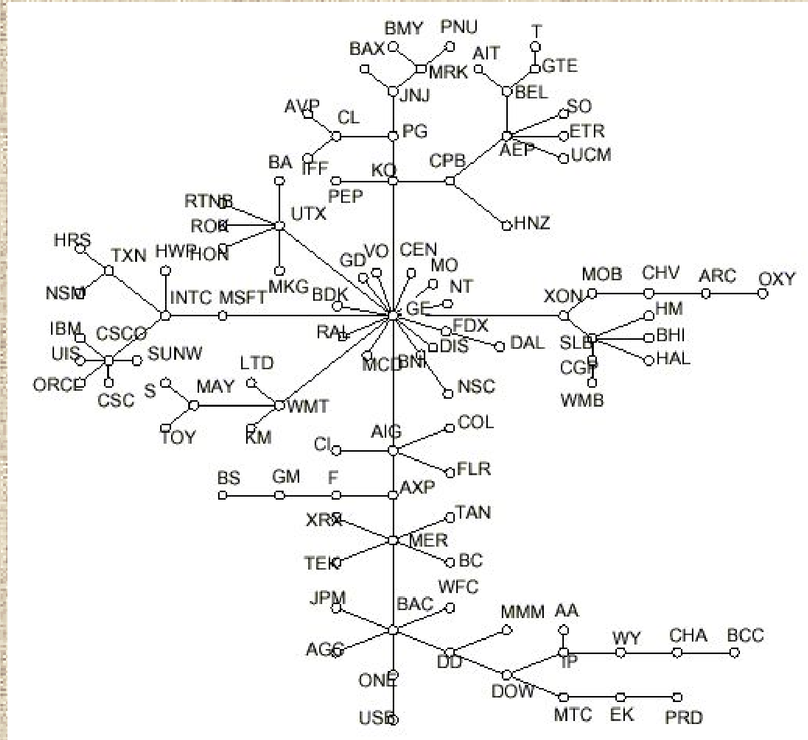


†Bonanno, Lillo, Mantegna, Quantitative Finance 1, 96-104 (2001)

MSTs at different time horizons†



Δt=19 min 30 sec



Δt=6 h 30 min

†Bonanno, Lillo, Mantegna, Quantitative Finance 1, 96-104 (2001)

Topology of MSTs

Topology† of MSTs in

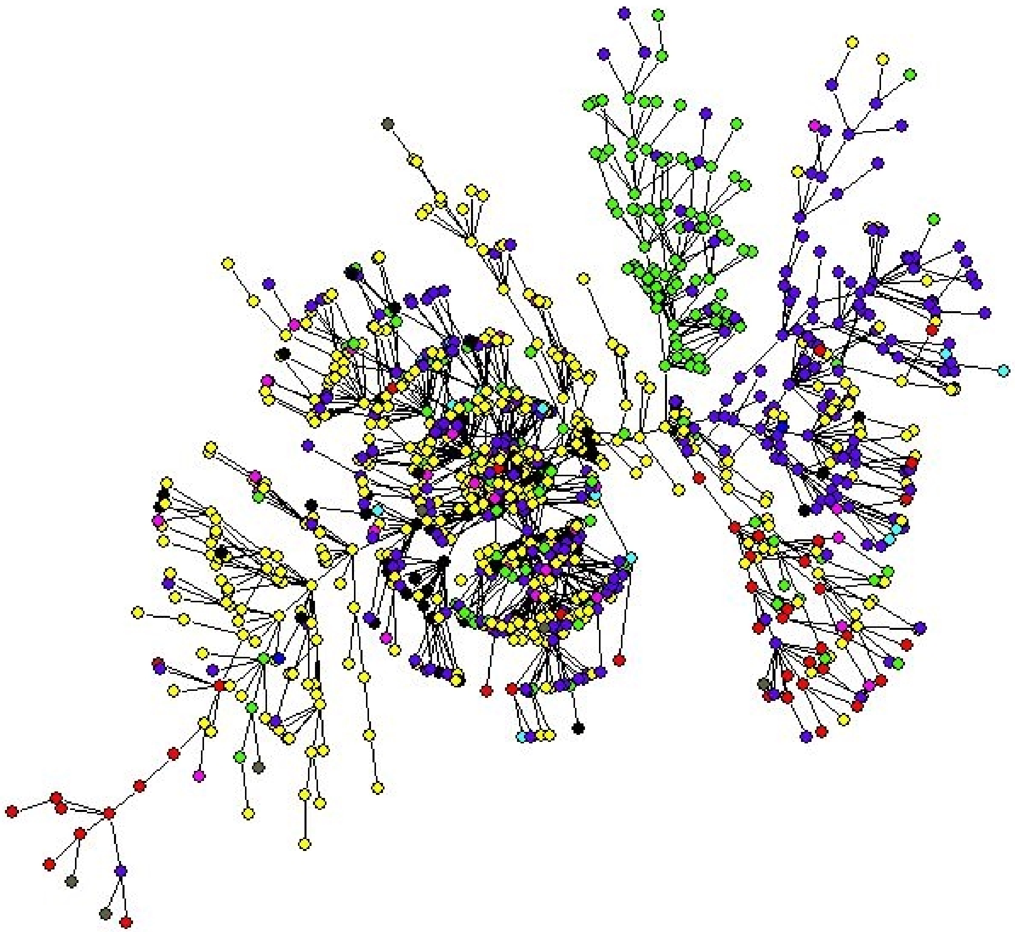
- empirical data;
- in an (unrealistic) random market

and

- in a (realistic) one-factor model

†Bonanno, Caldarelli, Lillo and Mantegna, cond-mat/0211546, PRE (in press 2003)

Empirical data

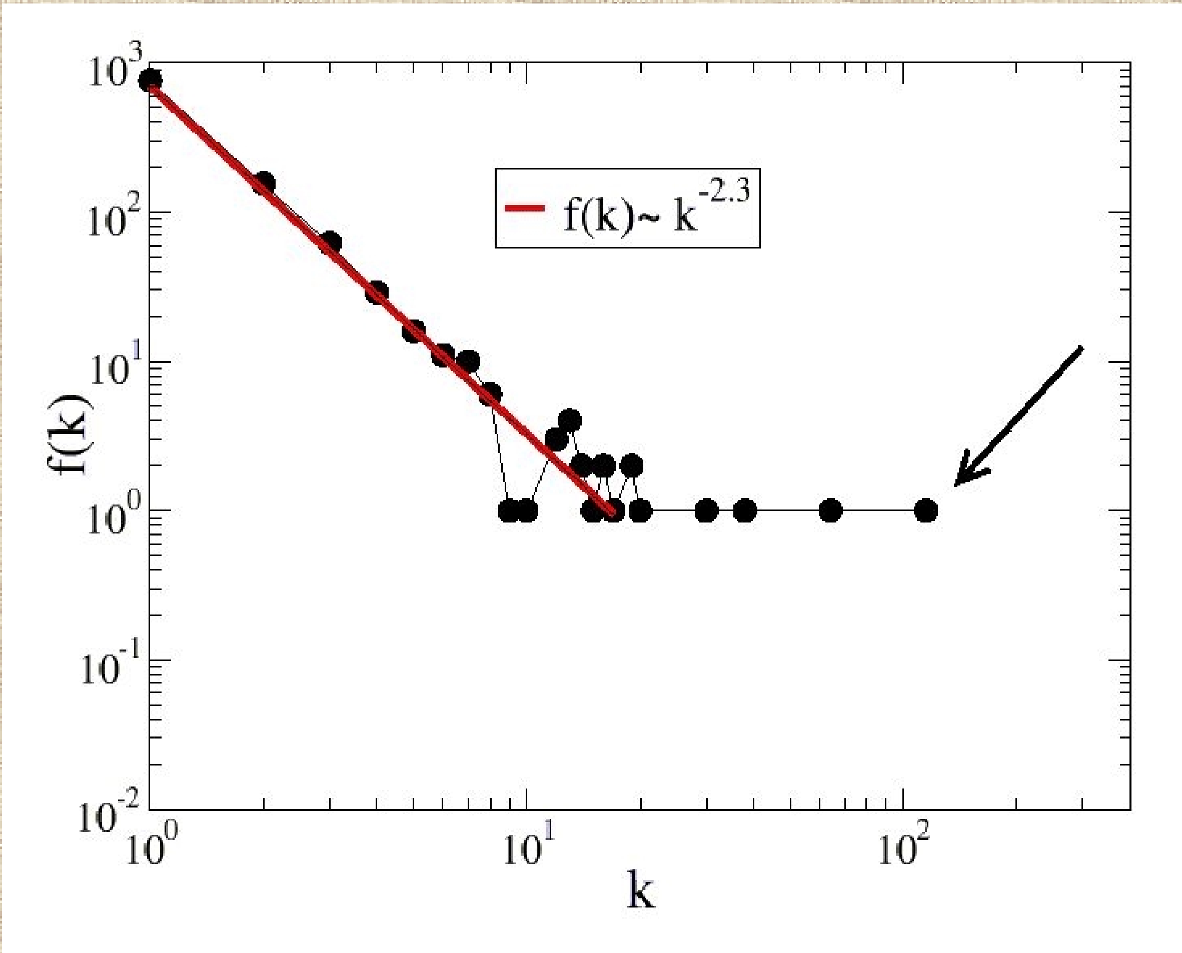


1071 stocks continuously traded at the NYSE during the period 1987-1998 (3030 trading days)

The color code refers to the SIC index:

- finance
- manufacturing
- construction
- utilities
- wholesale trade
- mining
- retail trade
- services
- public administration

Empirical data: degree distribution



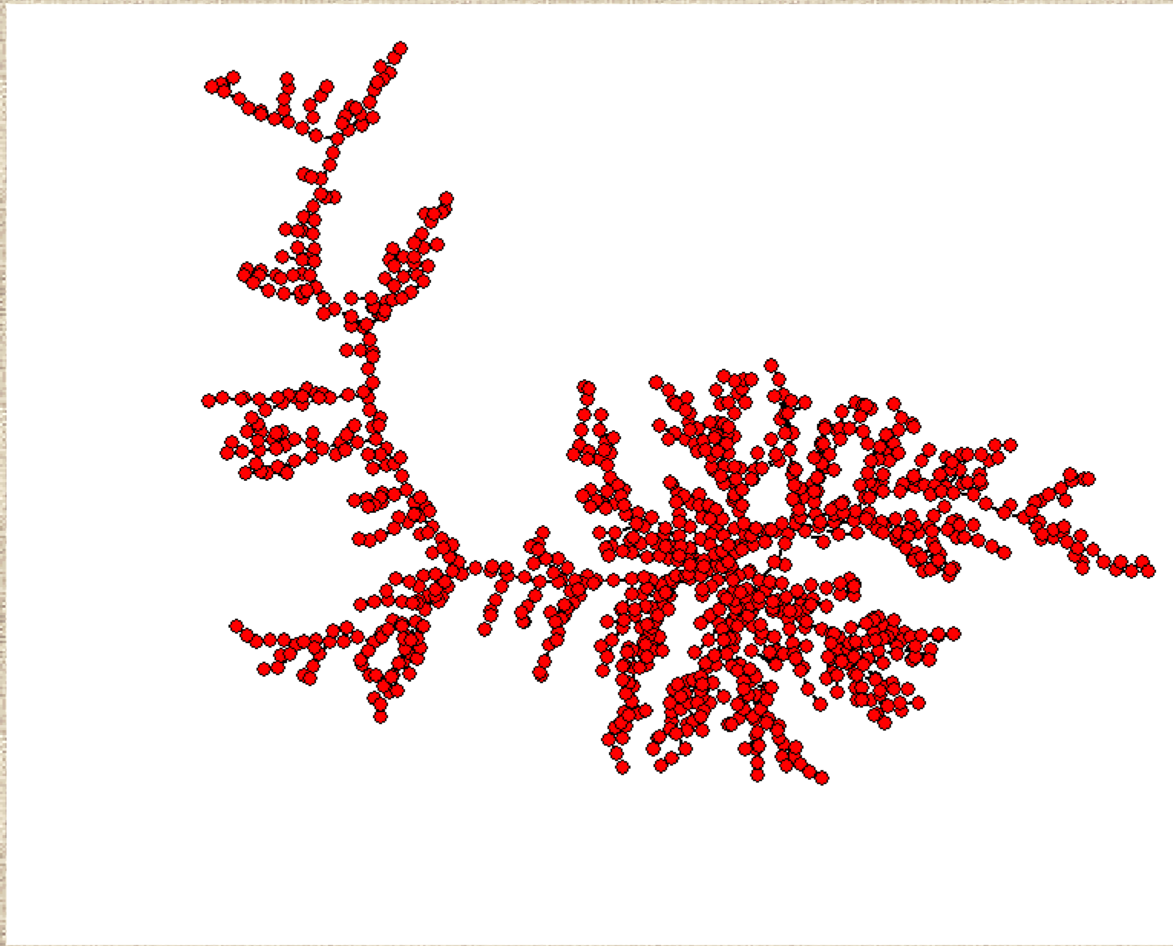
It is approximately power-law decaying for low k values and a few nodes with very high degree are present.

A random (unrealistic) model

$$R_i(t) = \varepsilon_i(t)$$

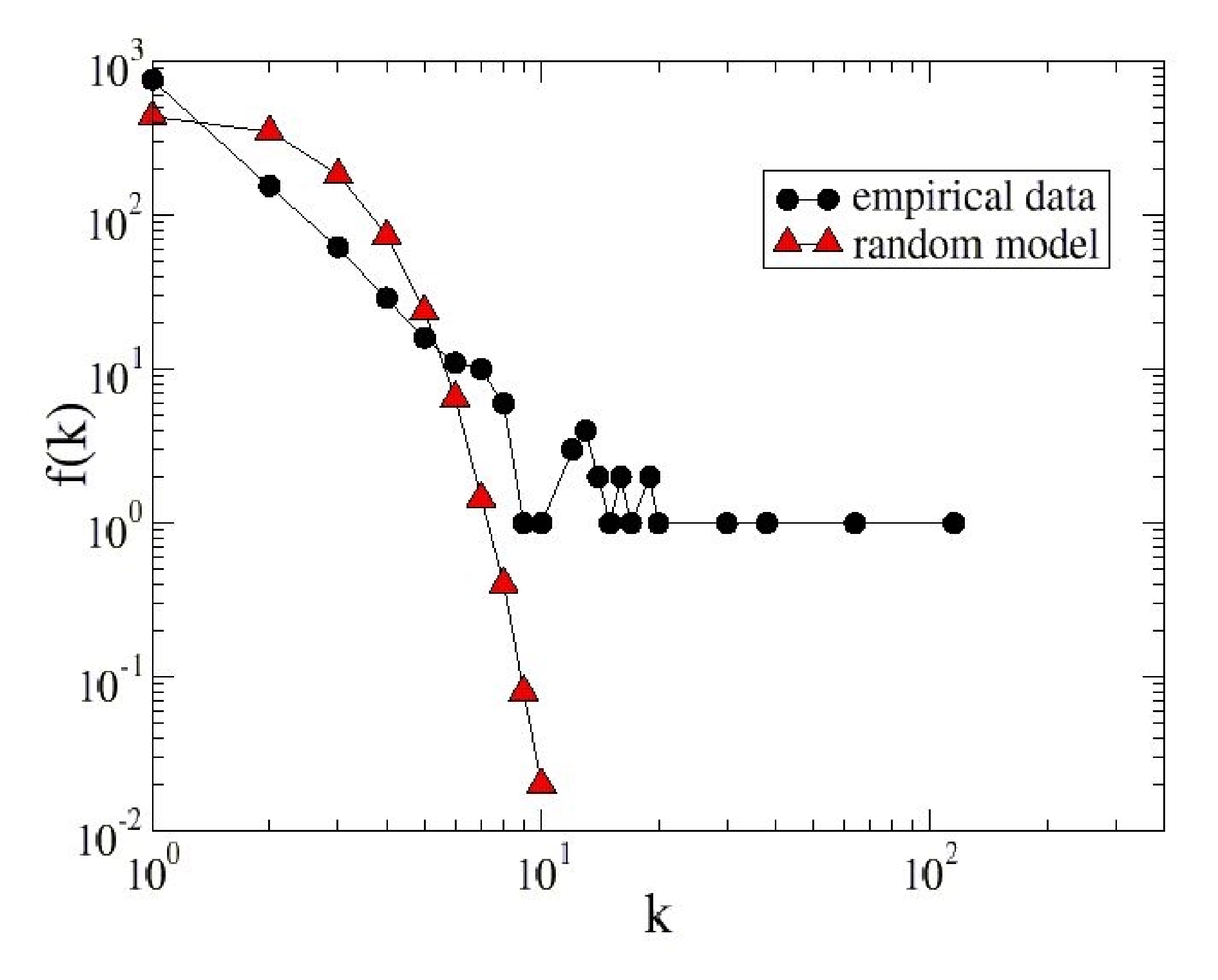
$\varepsilon_i(t)$ is a Gaussian zero-mean i.i.d. random variable

MST of a random artificial market

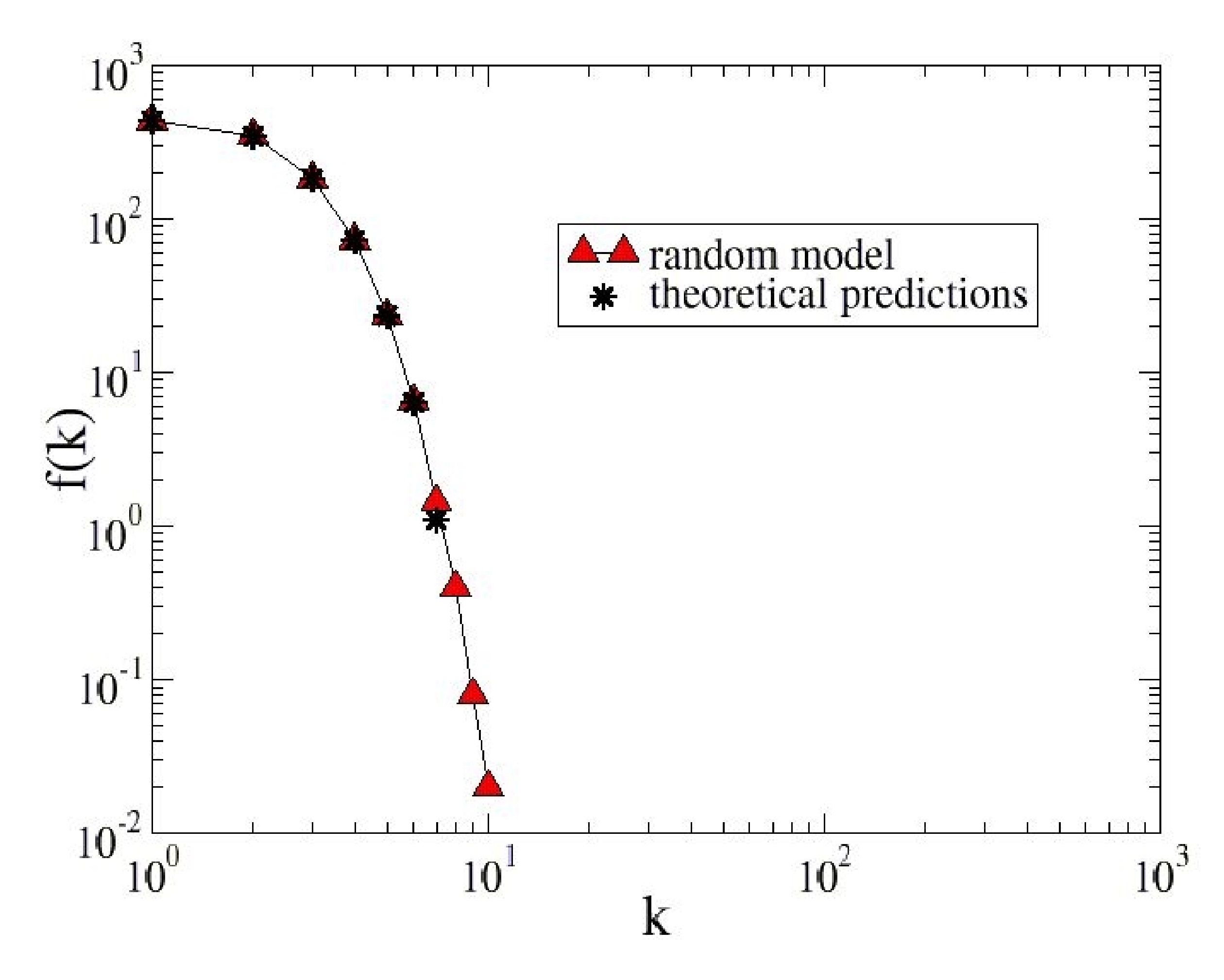


Topological properties of the random model

Degree distribution



Theoretical prediction of the random model



Theory developed in M.D.Penrose, The Annals of Probability **24**, 1903 (1996)

Eigenvalue spectrum

It may be worth mentioning that the random model describes a significant part of the spectrum of correlation matrix eigenvalues

From Laloux et al, PRL **83**, 1467 (1999)

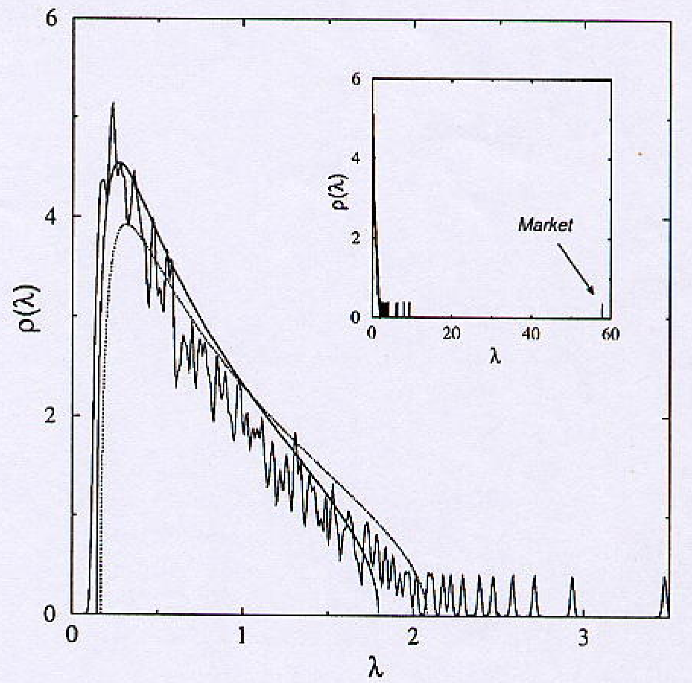


FIG. 1. Smoothed density of the eigenvalues of **C**, where the correlation matrix **C** is extracted from $N = 406$ assets of the S&P 500 during the years 1991–1996. For comparison we have plotted the density Eq. (3) for $Q = 3.22$ and $\sigma^2 = 0.85$: this is the theoretical value obtained assuming that the matrix is purely random except for its highest eigenvalue (dotted line). A better fit can be obtained with a smaller value of $\sigma^2 = 0.74$ (solid line), corresponding to 74% of the total variance. Inset: Same plot, but including the highest eigenvalue corresponding to the market, which is found to be 25 times greater than λ_{max}.

One-factor model

We compare our empirical results with numerical simulations based on the **one-factor model**

$$R_i(t) = \alpha_i + \beta_i R_M(t) + \varepsilon_i(t)$$

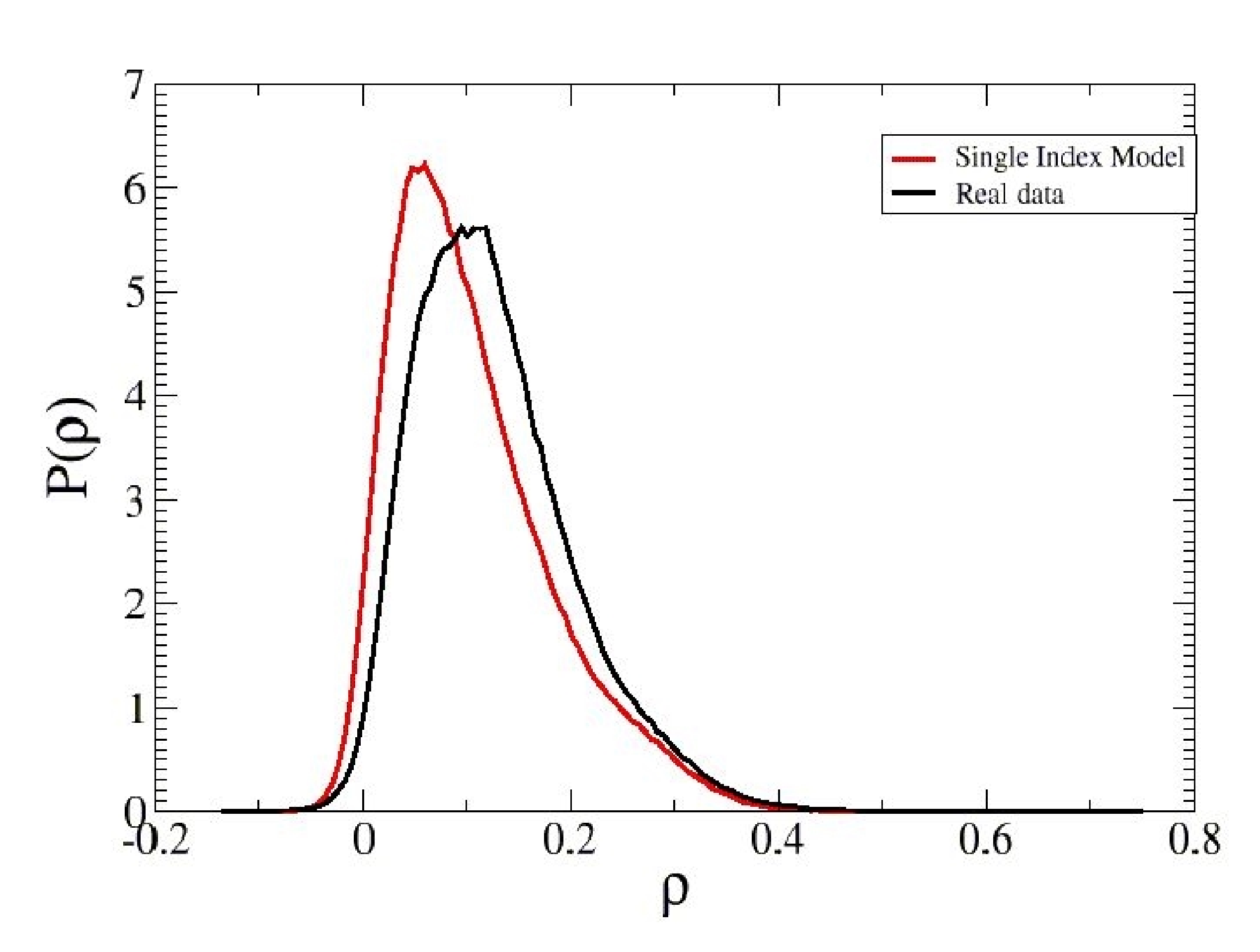
where α_i and β_i are two real parameters, ε_i is a zero-mean noise term characterized by a variance equal to σ_i^2

$R_M(t)$ is the market factor. We choose it as the SP 500 index.

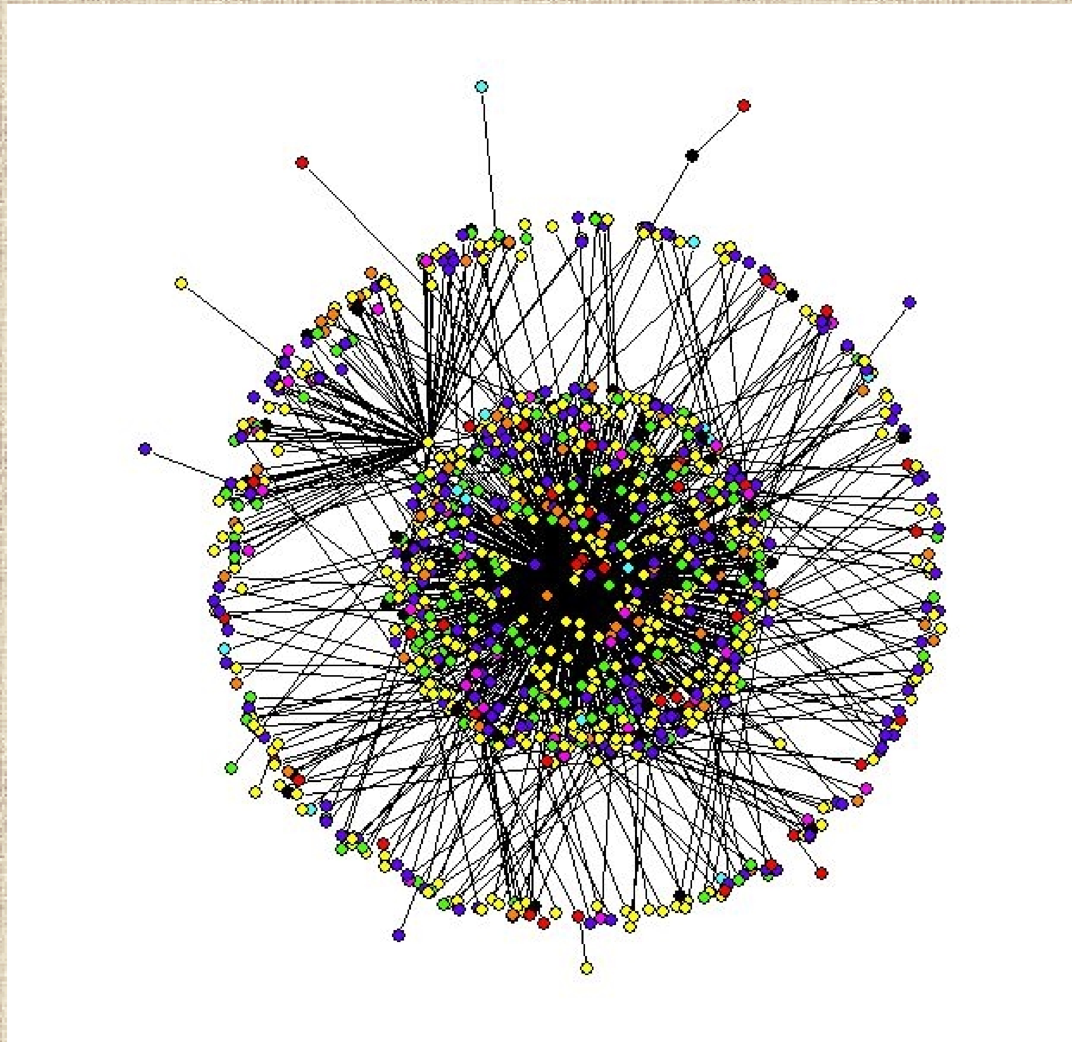
In our investigation, we estimate the model parameters and generate an artificial market.

Correlation in the one-factor model

The one-factor model explains more than 85% of the elements of correlation matrix

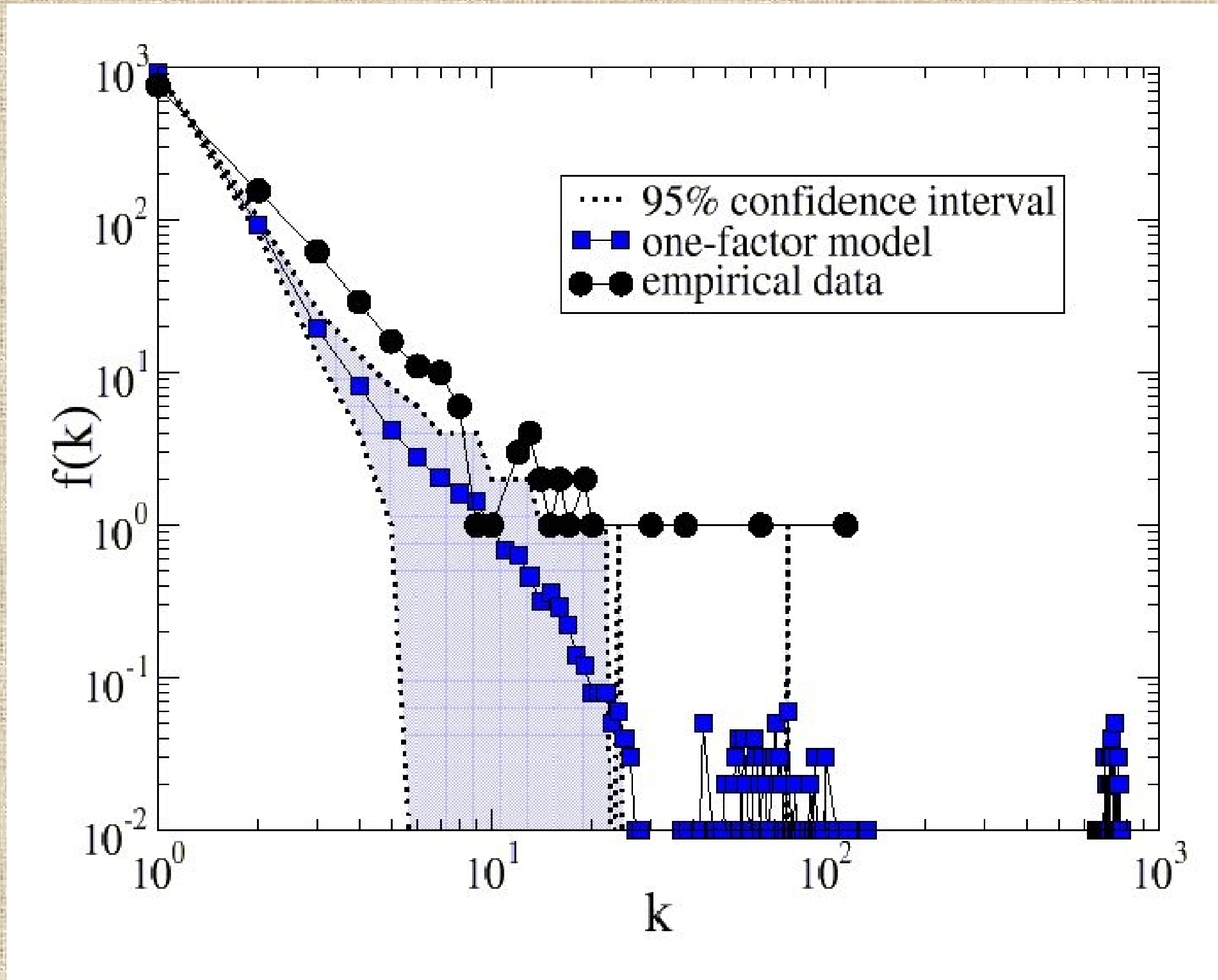


MST of a one-factor model



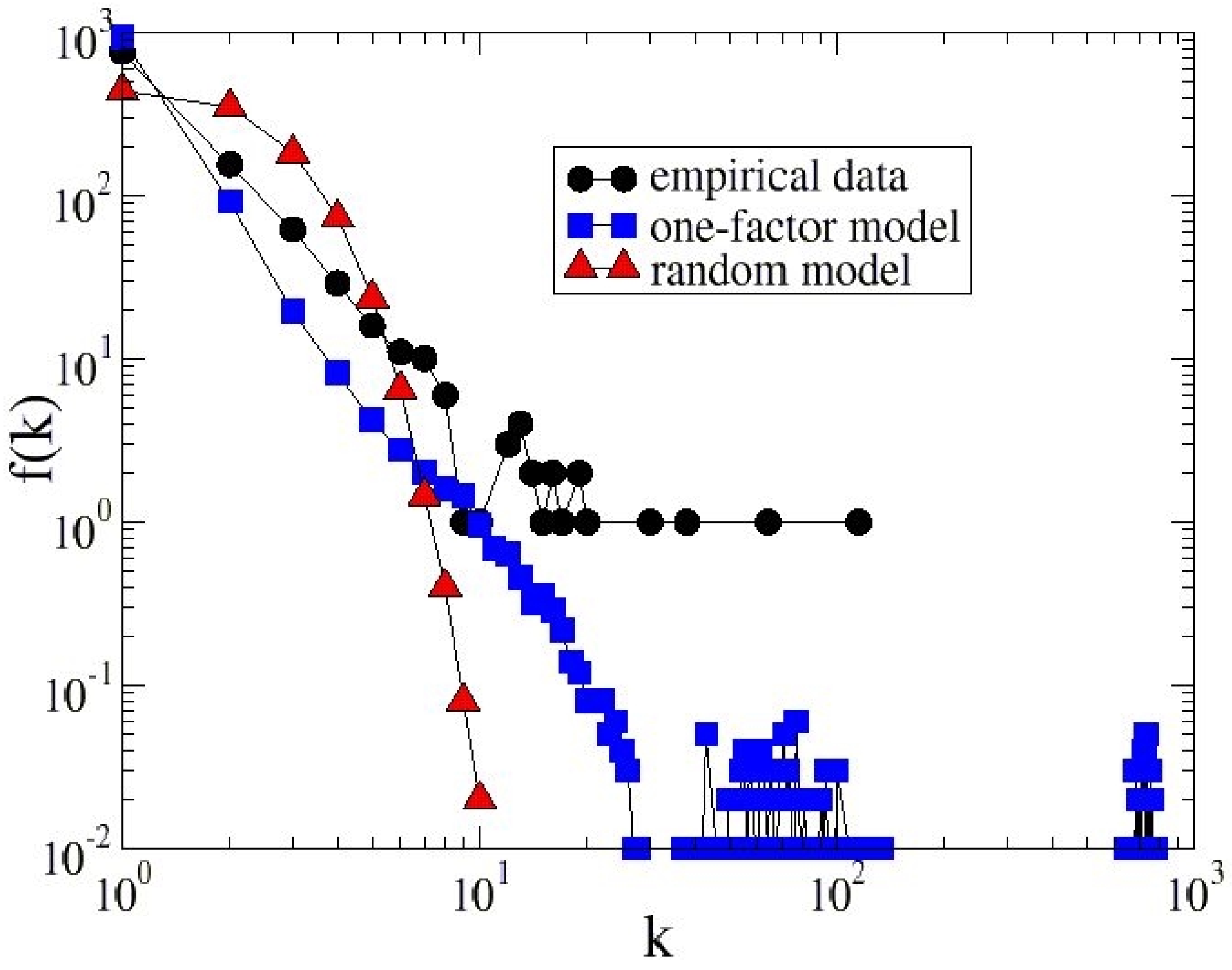
Topological properties of the one-factor model

Degree distribution



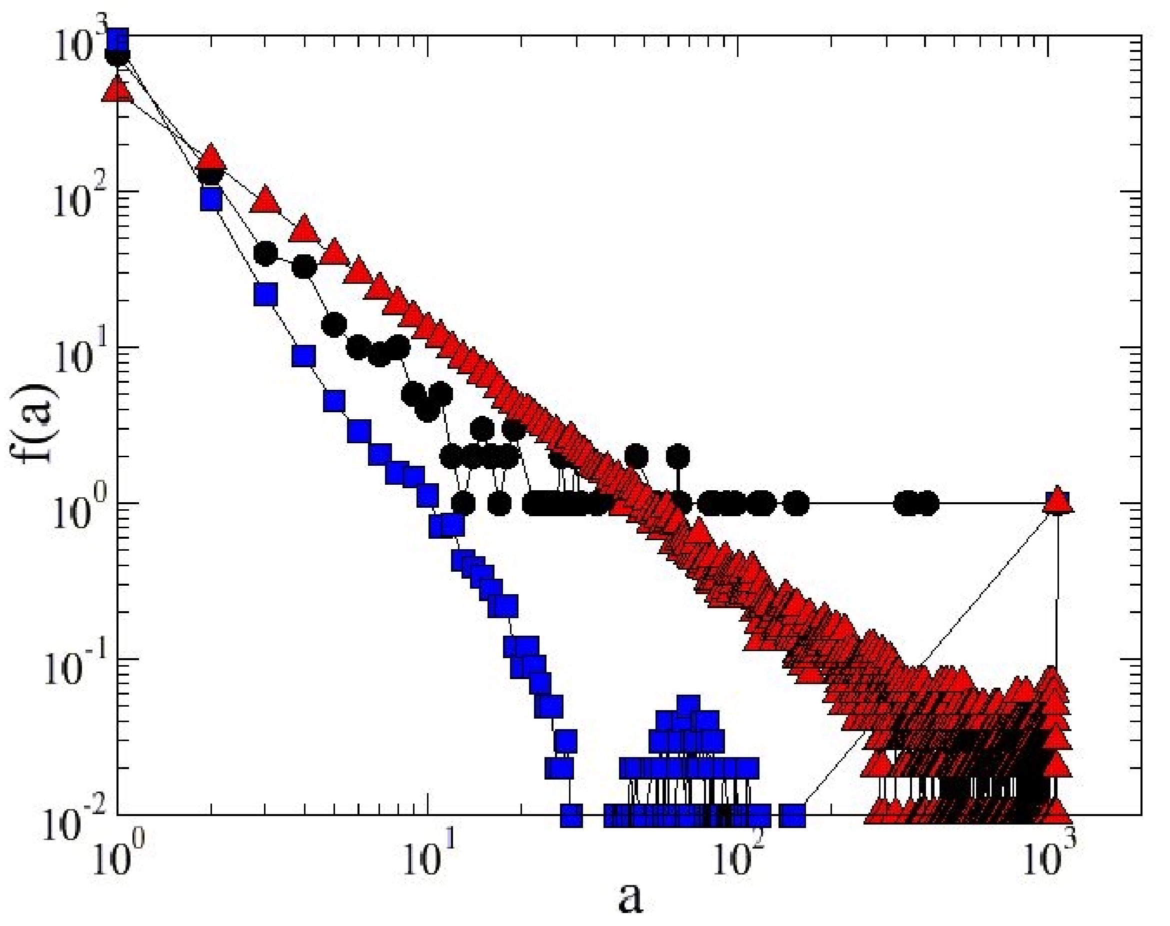
Minimal spanning tree networks in real and artificial markets

Topological properties of MSTs are different in the **degree pdf f(k)** and are enough to falsify a widespread financial model



Topological properties are able to falsify the model

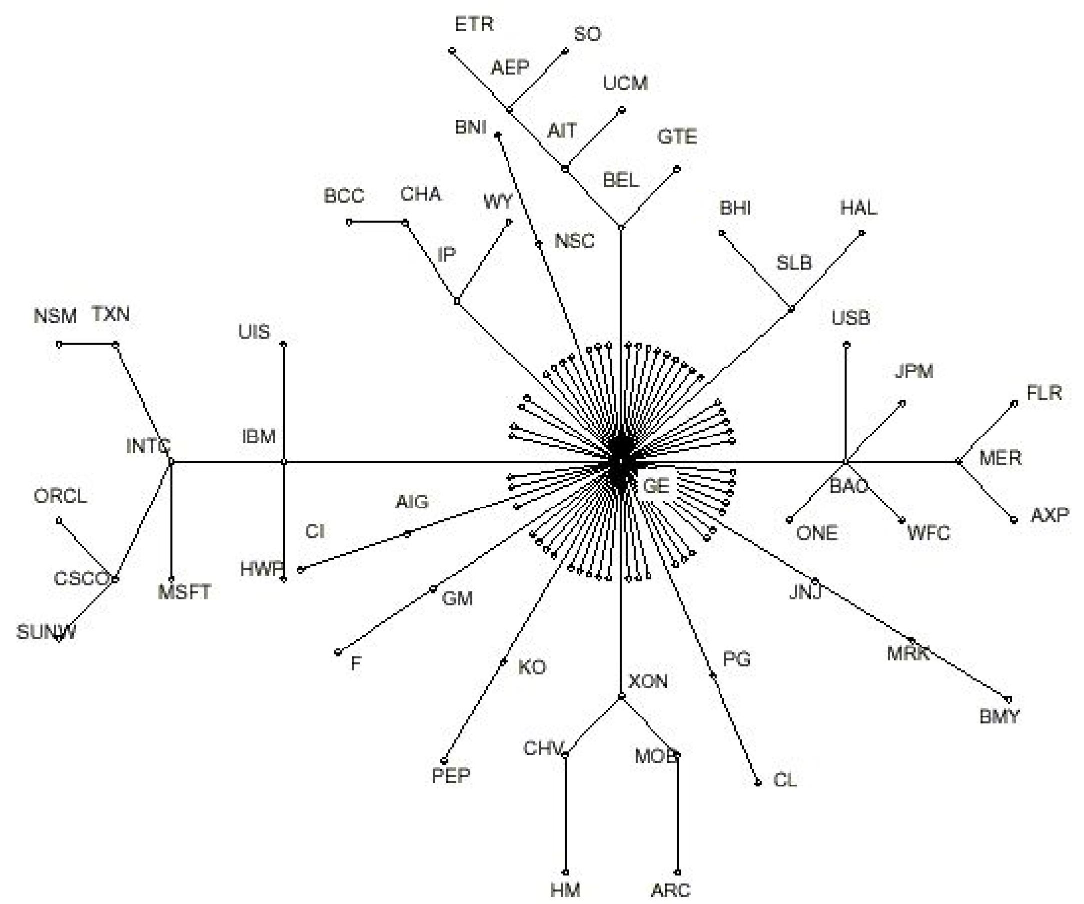
Other topological quantities, as for example, the **in-degree component** **f(a)** allow us to conclude the same way



Can we use our results for predictive purposes?

- Are empirical results suggesting that one-factor model better describes high-frequency returns rather than daily returns?

Δt
19 min
30 sec



A preliminary answer is **YES**!

Conclusion

- The topology of MSTs obtained by filtering the information present in the time dynamics of returns of an asset portfolio traded in a financial market allows to falsify simple models of markets dynamics.

- Topological properties of MSTs obtained at very short time horizons suggest that one-factor model is more appropriate for intraday time horizons rather than for daily time horizons of stock prices, weekly index returns of stock exchanges and volatility .

The OCS website

http://lagash.dft.unipa.it