Study of the correlations between stocks of different markets

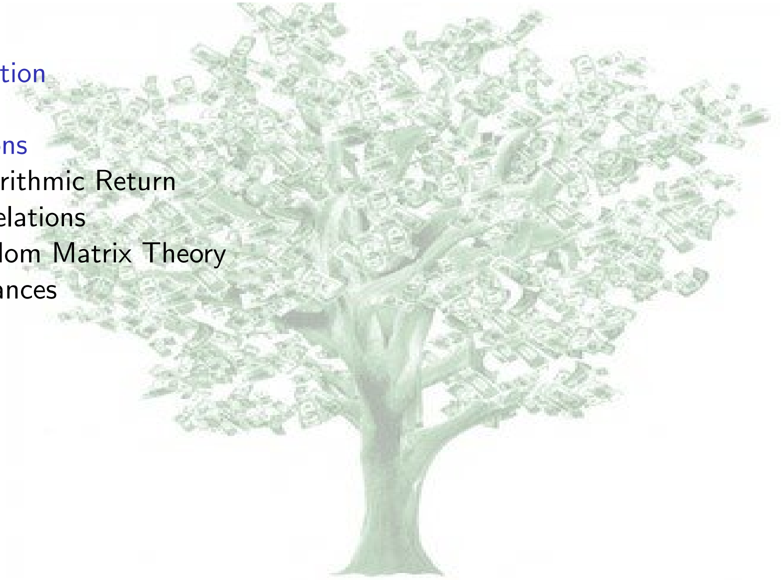
Ricardo Coelho

School of Physics
Trinity College Dublin

APFA 6 - Presentation, 5^{th} June 2007

Outline

Outline



Outline

Introduction

Definitions
- Logarithmic Return
- Correlations
- Random Matrix Theory
- Distances

Results
- Industrial clustering
- Geographical clustering
- Industrial vs Geographical clustering

Outline

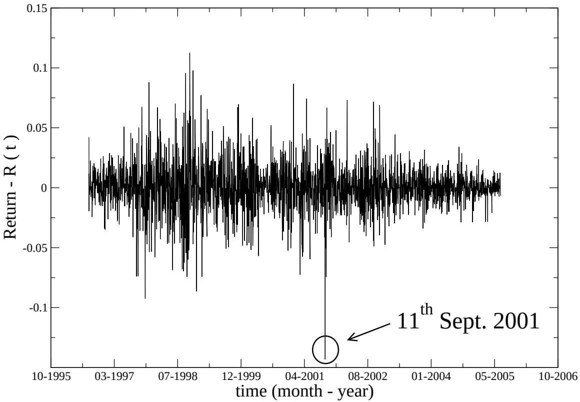
Introduction

We studied three different sets of data.

- A portfolio from the London Stock Exchange FTSE100 index:
  - 67 stocks
  - Daily closing price
  - From 2^{nd} August 1996 until 27^{th} June 2005
- Different World Indices
  - 53 countries' equity markets
  - Wednesday closing price (weekly returns)
  - From 8^{th} January 1997 until 1^{st} February 2006
- More than 6000 stocks from markets around the world
  - Daily closing price
  - From 30^{th} December 1994 until 1^{st} January 2007

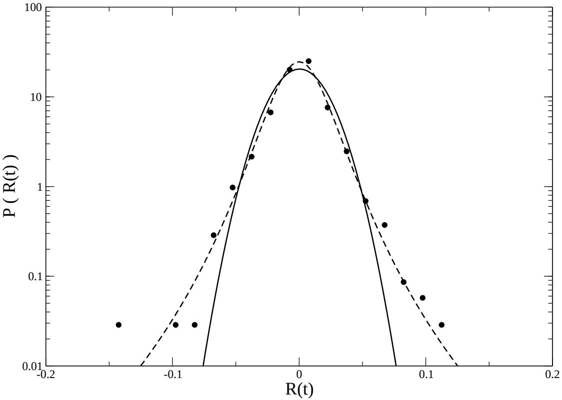
Log-return of stocks shows extreme events (e.g. HSBC)

$$R_i(t) = \ln P_i(t) - \ln P_i(t-1) \quad (1)$$



Distribution of the Log-return is non-Gaussian

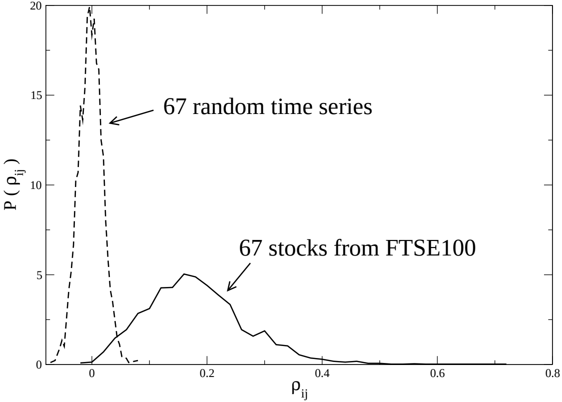
$k \sim 2.90$ and $q \sim 1.34$



Correlation between stocks i and j

$$\rho_{ij} = \frac{\langle \mathbf{R}_i \mathbf{R}_j \rangle - \langle \mathbf{R}_i \rangle \langle \mathbf{R}_j \rangle}{\sqrt{\left(\langle \mathbf{R}_i^2 \rangle - \langle \mathbf{R}_i \rangle^2 \right) \left(\langle \mathbf{R}_j^2 \rangle - \langle \mathbf{R}_j \rangle^2 \right)}} \qquad (2)$$

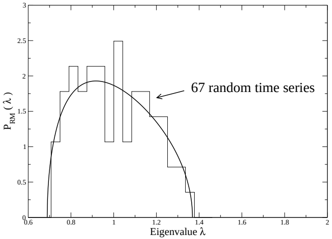
- ρ_{ij} form a symmetric $N \times N$ matrix; $-1 \leq \rho_{ij} \leq 1$; $\rho_{ii} = 1$.

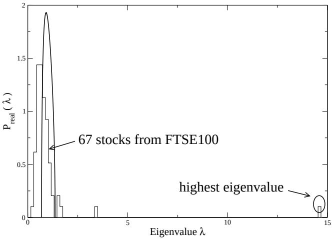


Matrix of correlations show non-randomness

Random Matrix Theory for $N \to \infty$ and $T \to \infty$, where $Q = T/N$ is fixed and bigger than 1.

$$P_{RM}(\lambda) = \frac{Q}{2\pi\sigma^2} \frac{\sqrt{(\lambda_{max} - \lambda)(\lambda - \lambda_{min})}}{\lambda} \tag{3}$$

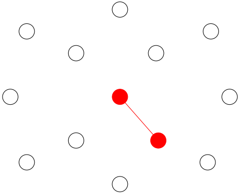




Distance between stocks i and j

$$d_{ij} = \sqrt{2(1 - \rho_{ij})} \tag{4}$$

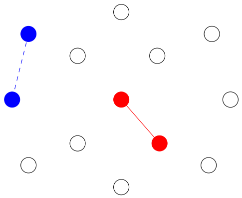
- $0 \leq d_{ij} \leq 2$, small values imply strong correlations.
- From distances we construct the Minimal Spanning Tree (MST).
- Choose the minimum distance between 2 stocks - one link between these 2.



Distance between stocks i and j

$$d_{ij} = \sqrt{2(1 - \rho_{ij})} \tag{4}$$

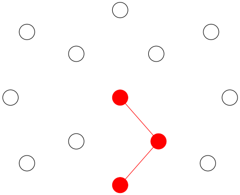
- $0 \leq d_{ij} \leq 2$, small values imply strong correlations.
- From distances we construct the Minimal Spanning Tree (MST).
- Choose the minimum distance between 2 stocks - one link between these 2.
- Find the next minimum distance. If the 2 stocks are not linked or both linked, choose the next minimum distance. If just 1 of the stocks has already links, link them.



Distance between stocks i and j

$$d_{ij} = \sqrt{2(1 - \rho_{ij})} \tag{4}$$

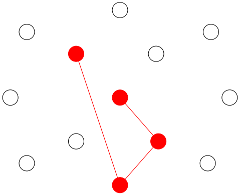
- $0 \leq d_{ij} \leq 2$, small values imply strong correlations.
- From distances we construct the Minimal Spanning Tree (MST).
- Choose the minimum distance between 2 stocks - one link between these 2.
- Find the next minimum distance. If the 2 stocks are not linked or both linked, choose the next minimum distance. If just 1 of the stocks has already links, link them.



Distance between stocks i and j

$$d_{ij} = \sqrt{2(1 - \rho_{ij})} \tag{4}$$

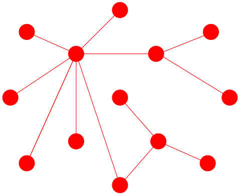
- $0 \leq d_{ij} \leq 2$, small values imply strong correlations.
- From distances we construct the Minimal Spanning Tree (MST).
- Choose the minimum distance between 2 stocks - one link between these 2.
- Find the next minimum distance. If the 2 stocks are not linked or both linked, choose the next minimum distance. If just 1 of the stocks has already links, link them.



Distance between stocks i and j

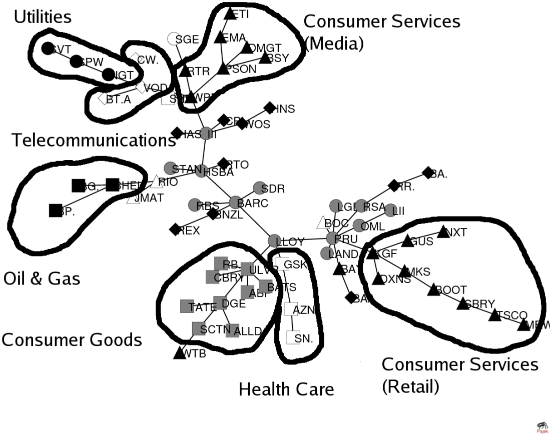
$$d_{ij} = \sqrt{2(1 - \rho_{ij})} \tag{4}$$

- $0 \leq d_{ij} \leq 2$, small values imply strong correlations.
- From distances we construct the Minimal Spanning Tree (MST).
- Choose the minimum distance between 2 stocks - one link between these 2.
- Find the next minimum distance. If the 2 stocks are not linked or both linked, choose the next minimum distance. If just 1 of the stocks has already links, link them.



Stocks cluster in industrial sectors (ICB classification)

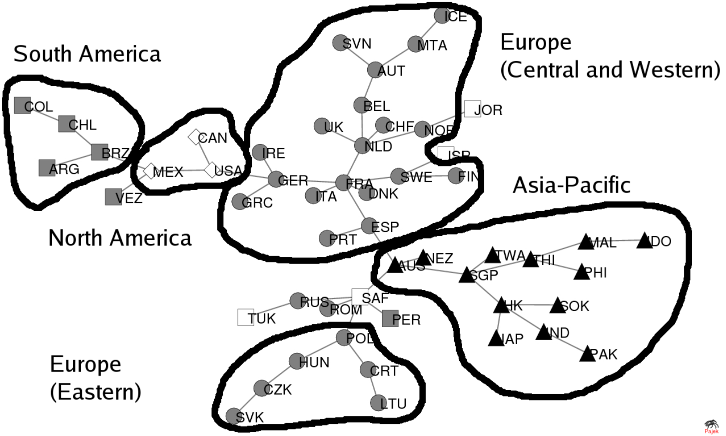
Full time series, $T = 2322$ days, for the FTSE100 stocks.



Physica A 373 (2007) 615-626

Markets organised by geographical location

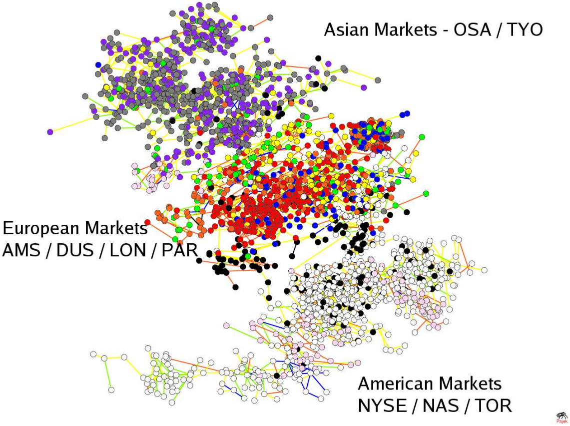
Full time series, $T = 475$ weeks, for the 53 World Indices.



Physica A 376 (2007) 455-466

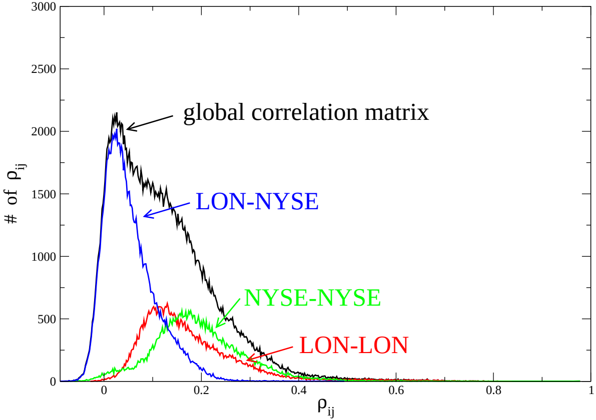
What about stocks from different countries???

Full time series, $T = 3127$ days, for 2500 stocks from different markets.



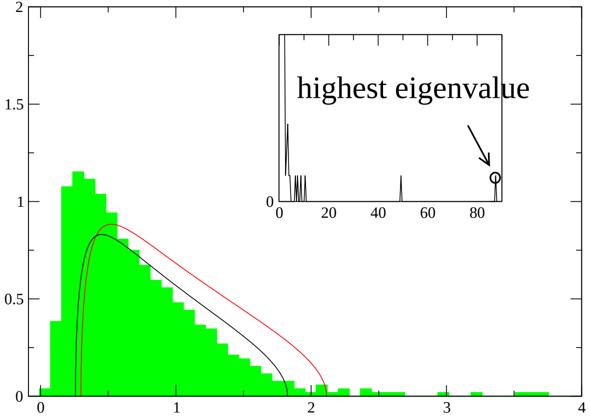
NYSE vs LON - Correlation between stocks of different markets

322 stocks from each market.

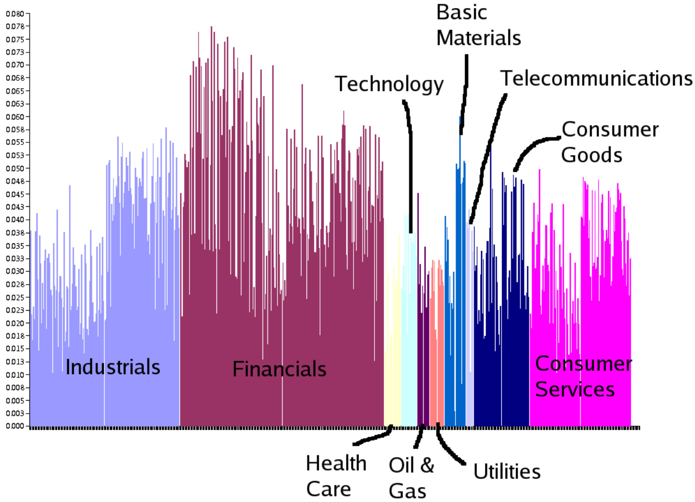


NYSE vs LON - Eigensystem information

The two largest eigenvalues could represent each market.

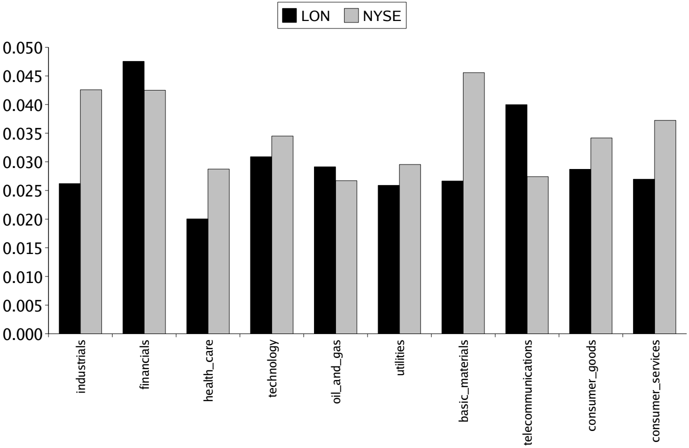


NYSE vs LON - Highest eigenvector information

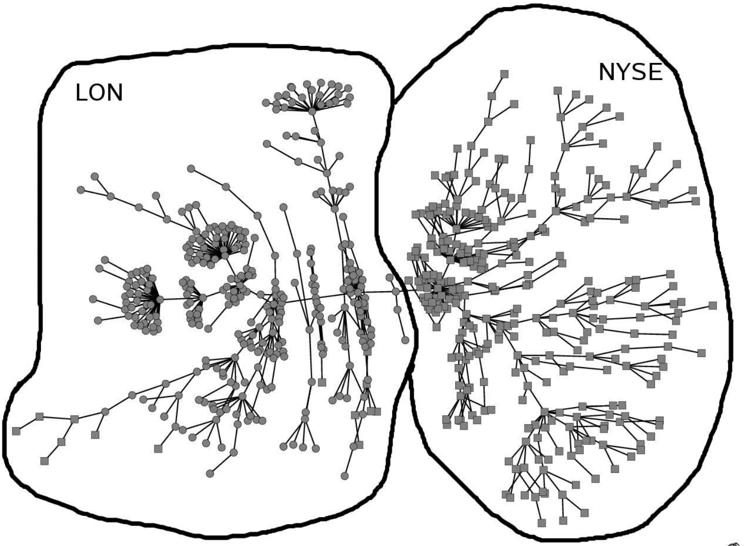


NYSE vs LON - Highest eigenvector information

Sum of elements from each sector/country - $\sum_i \lambda_i$



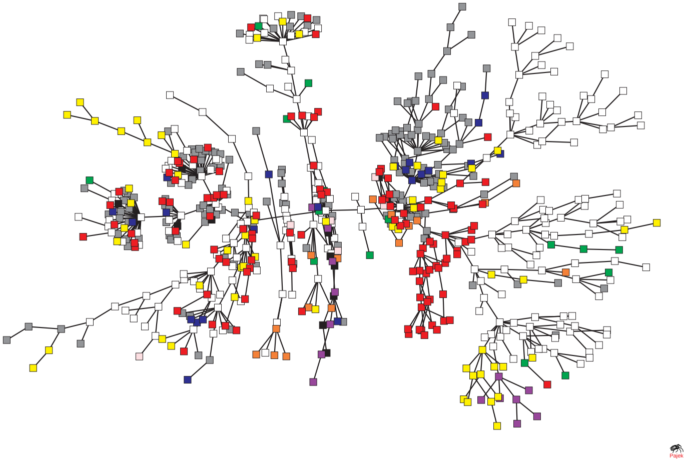
NYSE vs LON - Clustering of stocks in market



NYSE vs LON - Clustering in sector for NYSE

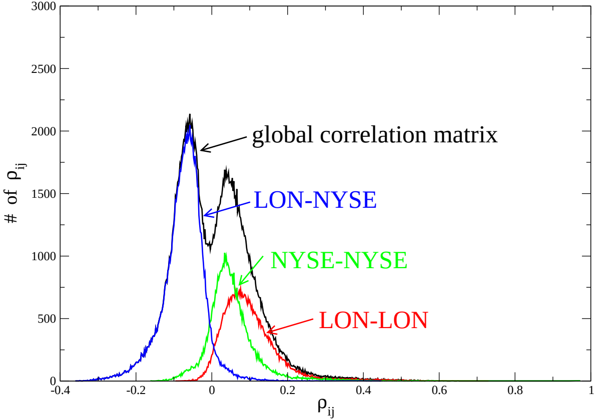
Different choice of stocks gives different clustering for LON.

Oil & Gas - black / Basic Materials - blue / Industrials - gray / Consumer Goods - yellow / Health Care - green
Consumer Services - red / Telecommunications - pink / Utilities - purple / Financials - white / Technology - orange



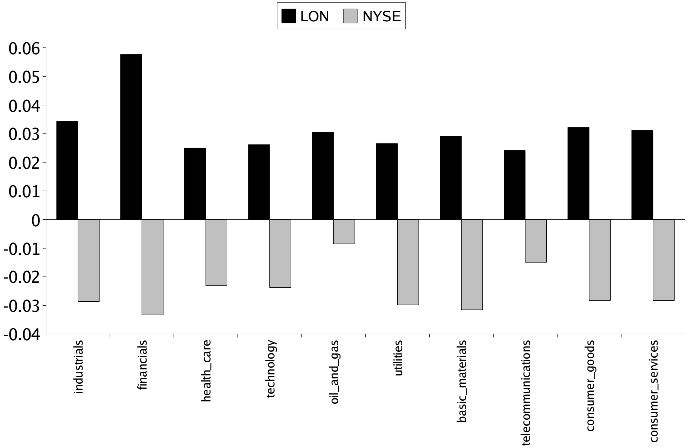
NYSE vs LON - Correlation (after filtering market mode)

Filtering affects NYSE more than LON.



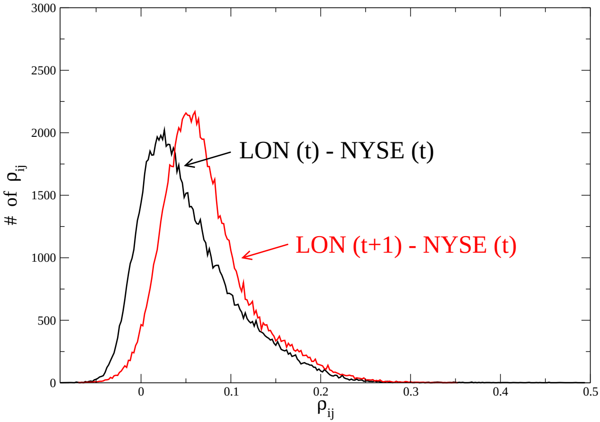
NYSE vs LON - Highest eigenvector (after filtering)

Sum of elements from each sector/country - $\sum_i \lambda_i$



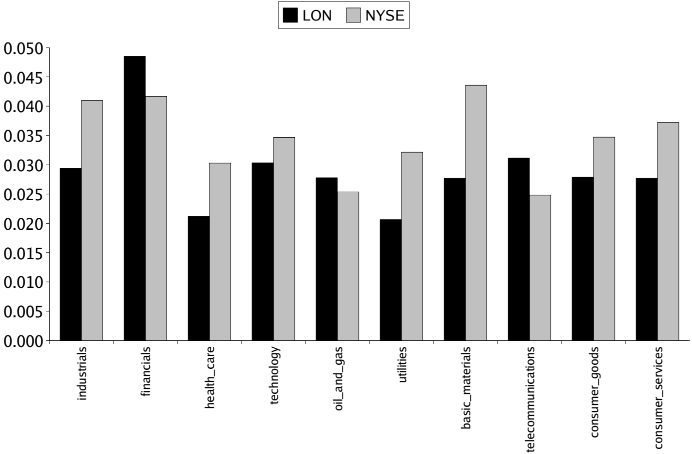
And if LON is one day ahead?

Correlation between NYSE stocks in time t and LON stocks in time $t + 1$.



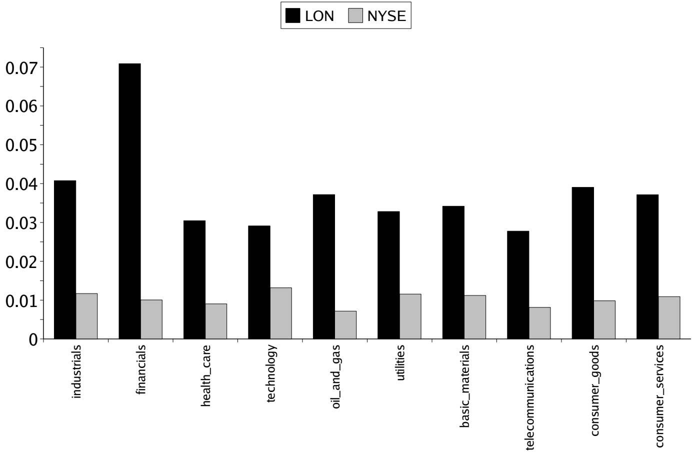
NYSE (t) vs LON (t+1) - Highest eigenvector

Sum of elements from each sector/country - $\sum_i \lambda_i$



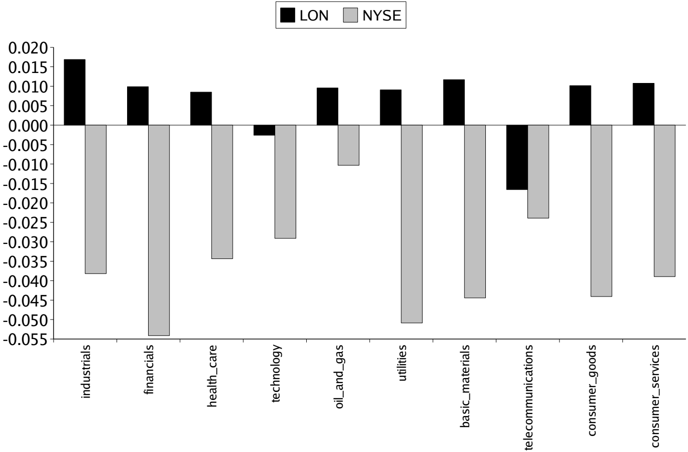
NYSE (t) vs LON (t+1) - Highest (after filtering)

Sum of elements from each sector/country - $\sum_i \lambda_i$



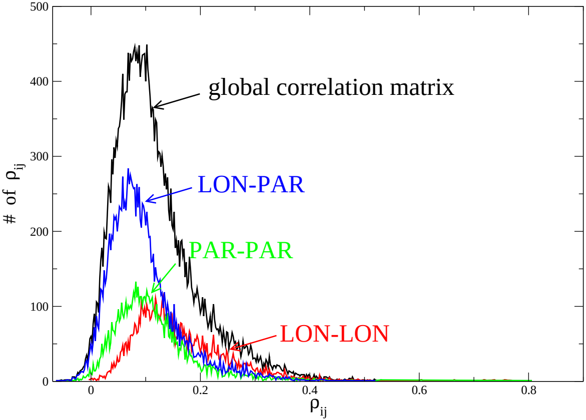
NYSE (t) vs LON (t+1) - 2^{th} highest (after filtering)

Sum of elements from each sector/country - $\sum_i \lambda_i$



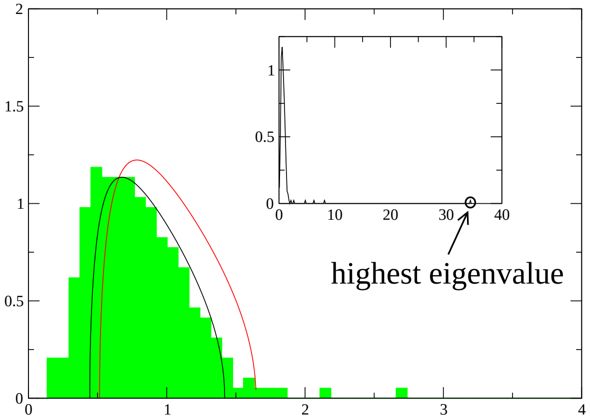
What about two markets from the same continent? PAR vs LON - Correlation

125 stocks from each market.



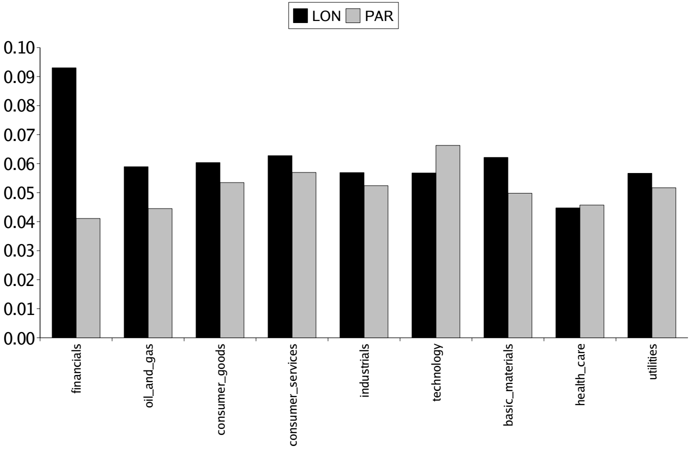
PAR vs LON - Eigensystem

One major eigenvalue in contrast with the NYSE-LON situation.



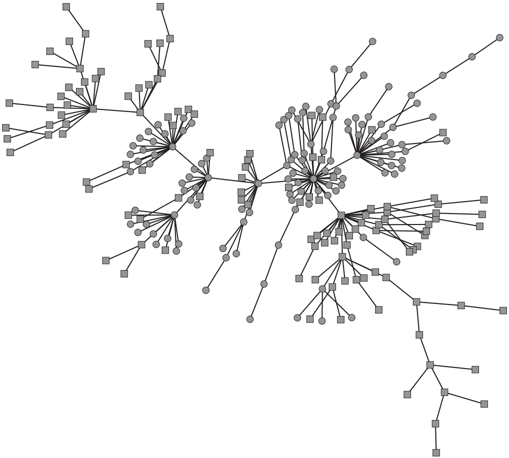
PAR vs LON - Highest eigenvector

Sum of elements from each sector/country - $\sum_i \lambda_i$



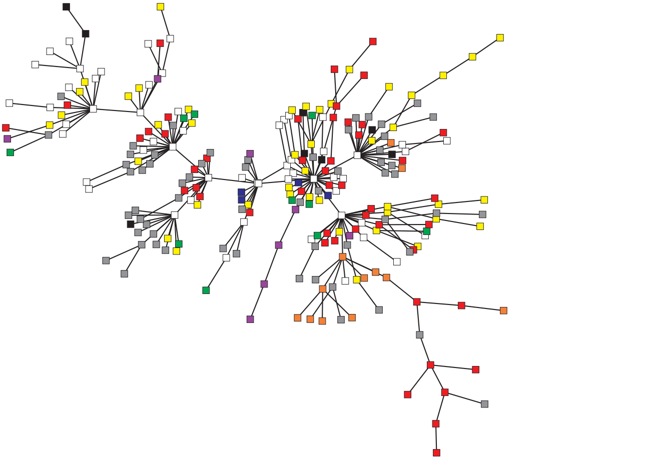
PAR vs LON - No clustering of stocks in market

Circles for LON and boxes for PAR.



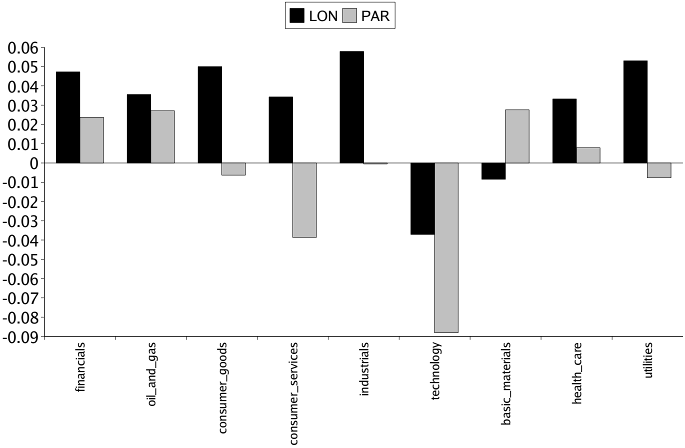
PAR vs LON - No clustering of stocks in sectors

Oil & Gas - black / Basic Materials - blue / Industrials - gray / Consumer Goods - yellow / Health Care - green
Consumer Services - red / Telecommunications - pink / Utilities - purple / Financials - white / Technology - orange

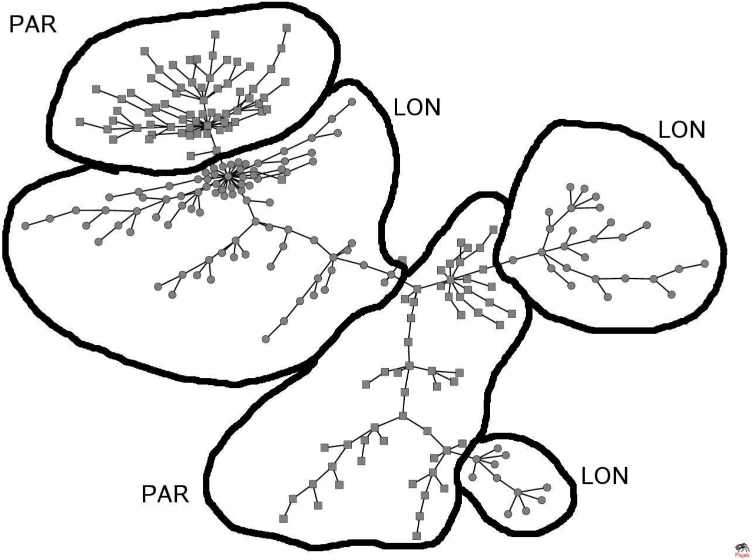


PAR vs LON - Highest eigenvector (after filtering)

Sum of elements from each sector/country - $\sum_i \lambda_i$

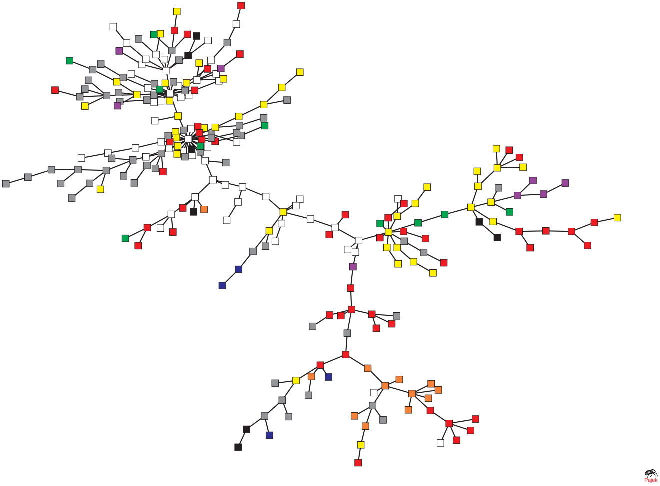


PAR vs LON - Clustering of stocks in market (a.f.)



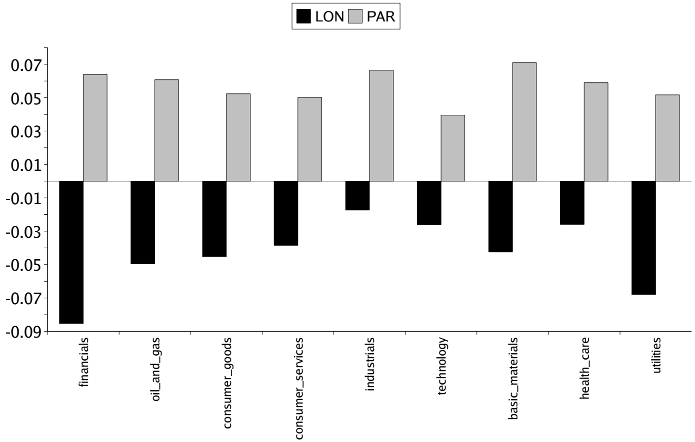
PAR vs LON - No clustering of stocks in sectors (a.f.)

Oil & Gas - black / Basic Materials - blue / Industrials - gray / Consumer Goods - yellow / Health Care - green
Consumer Services - red / Telecommunications - pink / Utilities - purple / Financials - white / Technology - orange



PAR vs LON - 2th highest eigenvector (after filtering)

Sum of elements from each sector/country - $\sum_i \lambda_i$



Conclusions

- Two different methodologies of analysing large amount of data give same results
- Similar behaviour for both the FTSE100 stocks and the World Indices:
  - MST show different clusters (Industrial sectors / Geography)
- The choice of the stocks can influence the clustering in sectors
- The stocks cluster in market before clustering in sectors
- The 3^{rd} eigenvalue for the PAR-LON and NYSE(t)-LON($t+1$) show the segregation of stocks in markets.

Future work

- Results for more than 2 markets from the same continent

Econophysics group of the Trinity College Dublin:

- Prof. Peter Richmond
- Dr. Stefan Hutzler
- Ricardo Coelho
- Joseph Barry

Collaborations:

- Prof. Brian Lucey (Business School of Trinity College Dublin)
- Prof. Claire Gilmore (McGowan School of Business of King's College, Pennsylvania)

Part of this work was published in Physica A:

`Physica A 373 (2007) 615-626`

`Physica A 376 (2007) 455-466`

Acknowledgments

- SFI - Science Foundation Ireland
- FCT - Fundação para a Ciência e a Tecnologia