Portfolio diversification through a network approach

F. Ricca, A. Scozzari

XIII Giornata della Ricerca MEMOTEF

27-28 giugno 2023

Outline

- Correlation network of the financial market
- Neighborhood separation constraints
- Assortative mixing in the correlation network
- Assortativity-based portfolio optimization models
- Results

Correlation network of the financial market

Correlation network of the financial market

Financial Network: $G = (N, E)$

- undirected
- **complete**
- edge-weighted

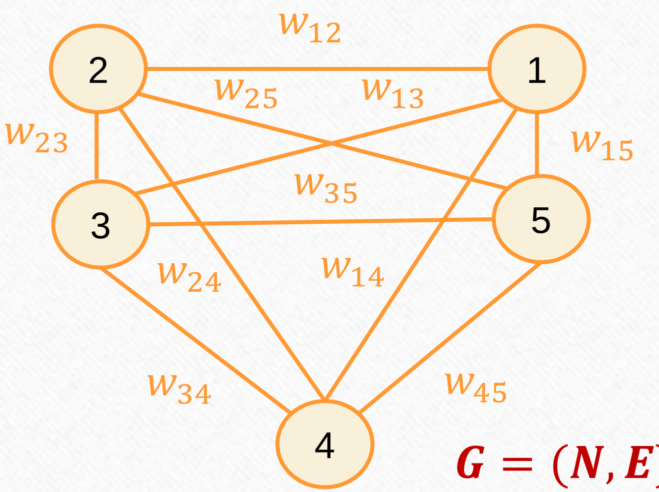
N set of nodes corresponding to assets, $|N| = n$

E set of edges between **any two** nodes, $|E| = m$

w_{ij} a weight for each edge (i,j)

$w_{ij} = \rho_{ij}$ Pearson correlation coeff. for returns of i and j

$w_{ij} = \tau_{ij}$ Kendall rank correlation for returns of i and j



$G = (N, E)$ with $n = 5$

We consider τ_{ij}, $(i, j) \in E$.

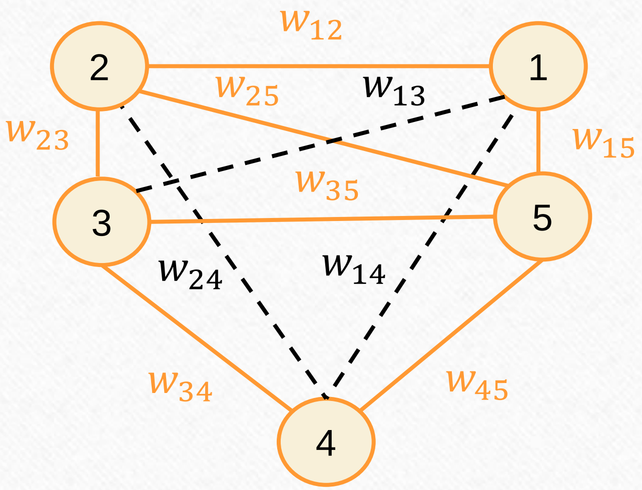
For R_i and R_j one has:

$$\tau_{ij} = \frac{2}{T(T-1)} \sum_{s<t} sgn(R_{is} - R_{it}) sgn(R_{js} - R_{jt}))$$

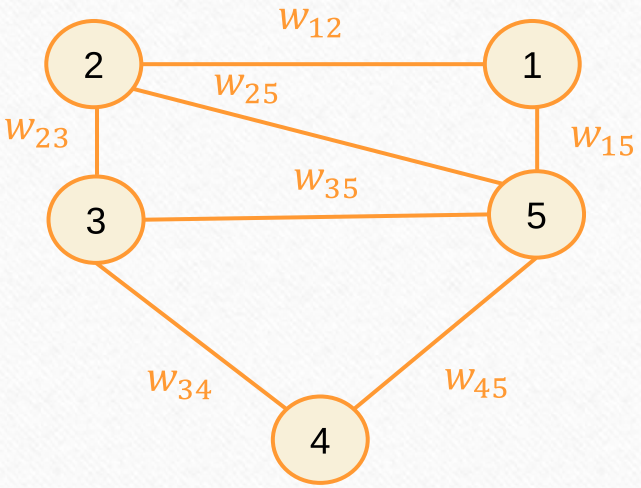
T is the number of time observations

Typically, in real markets we have $\tau_{ij} \geq 0$ for almost all pairs (i,j)

Correlation network of the financial market

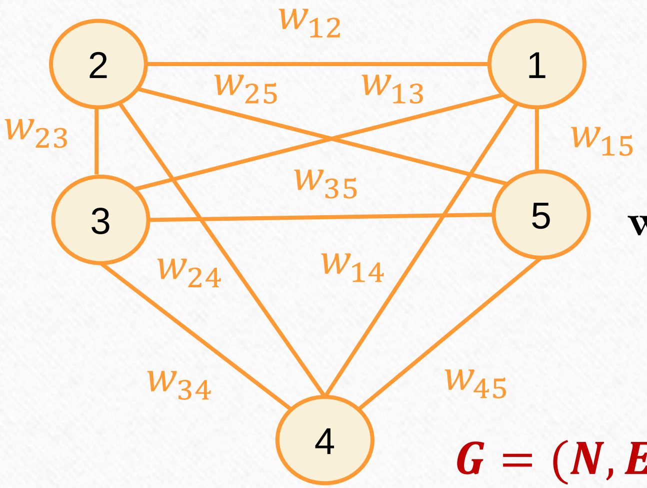


If $w_{13}, w_{14}, w_{24} < \theta$



$$\boldsymbol{G_\theta = (N, E_\theta)}$$

- Fix $\theta \in [-1,1]$
- $(i,j) \in E_\theta \quad \leftrightarrow \quad w_{ij} \geq \theta$



whole market

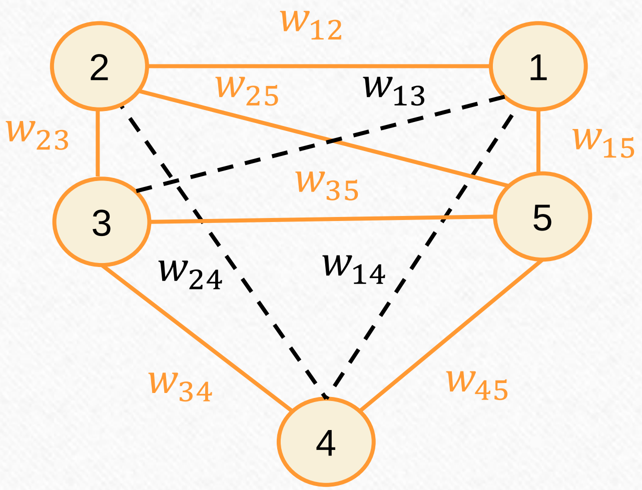
$\boldsymbol{G = (N, E)}$ **with** $\boldsymbol{n = 5}$

- The complete graph represents the **whole market correlation structure.**
- We are interested to **capture only** the structure of the **strongest links**.

➢ $G_\theta = (N, E_\theta)$ **filtered at a threshold** $\boldsymbol{\theta}$

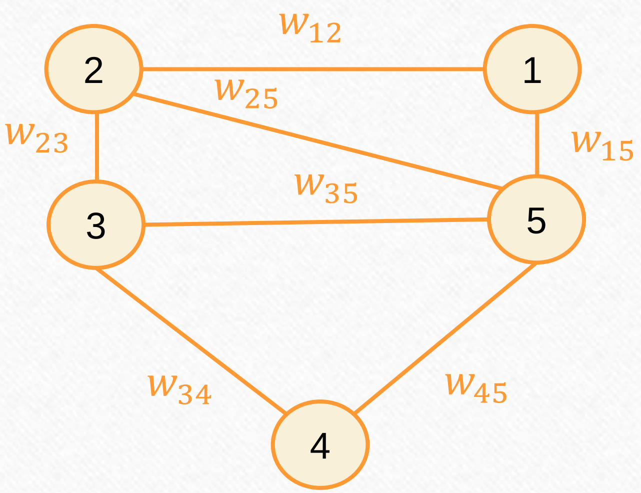
E_θ is a subset of E

Correlation network of the financial market



G: density = 1

If $w_{13}, w_{14}, w_{24} < \theta$



G_θ: density < 1

$$\boldsymbol{G_\theta = (N, E_\theta)}$$

- Fix $\theta \in [-1,1]$
- $(i,j) \in E_\theta \;\leftrightarrow\; w_{ij} \geq \theta$

Edge density of G_θ :

$$\frac{\boldsymbol{m_\theta}}{n(n-1)/2}$$

$m_\theta = |E_\theta|$

NOTE: As the **threshold** value **increases**, the edge **density** of the graph **G_θ decreases**.

Neighborhood separation constraints

Neighborhood separation constraints

$G_\theta = (N, E_\theta)$ Neighborhood of i $N(i) = \{j \in N: (\boldsymbol{i}, j) \in E_\theta\}$

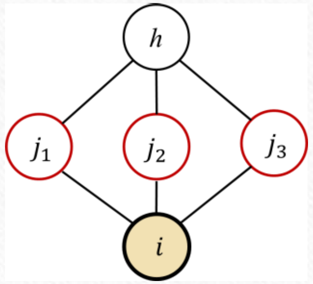
Neighborhood separation condition:

If i is **selected** → **no of its neighbors** can be chosen.

If i is **not selected** → **at most 1 neighbor of $\boldsymbol{i}$** can be chosen.

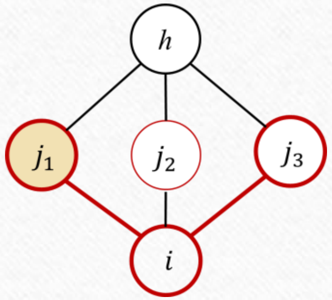
Example

If i is **selected**:



→ **no** among j_1, j_2, and j_3 can be selected

If i is **not selected**:



→ **at most 1** among j_1, j_2, and j_3 can be selected

One can include in the portfolio only asset-nodes **at distance at least 3** → Selected assets are **well-separated in the topology of $\boldsymbol{G_\theta}$**

Neighborhood separation constraints

Neighborhood of i $N(i) = \{j \in N: (\boldsymbol{i}, j) \in E_\theta\}$

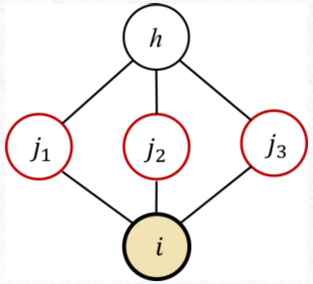
Neighborhood separation constraints:

$$y_i + \sum_{j \in N(i)} y_j \leq 1 \qquad i = 1, \ldots, n$$

n binary indicating variables:

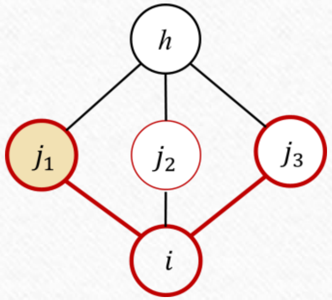
$$y_i = \begin{cases} 1 \; if \; asset \; i \; is \; in \; the \; portfolio \\ 0 \qquad otherwise \end{cases}$$

Example If i is **selected**:



→ no among j_1, j_2, and j_3 can be selected

If i is **not selected**:



→ at most one among j_1, j_2, and j_3 can be selected

One can include in the portfolio only asset-nodes **at distance at least 3**

Selected assets are **well-separated in the topology of G_θ**

Assortative mixing in the correlation network

Assortative mixing in the correlation network

$G_\theta = (N, E_\theta)$

Neighborhood of i $N(i) = \{j \in N: (\boldsymbol{i}, j) \in E_\theta\}$

Degree of i: $\boldsymbol{d_i = |N(i)|}$

Excess degree of i: $\boldsymbol{k_i = d_i - 1}$

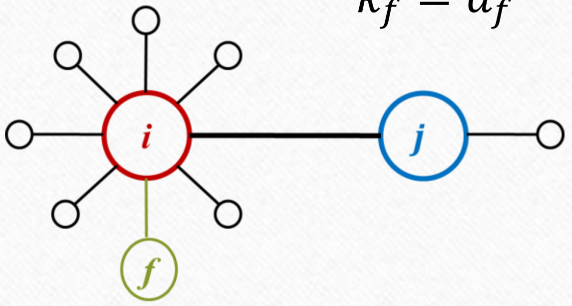
Assortative/disassortative edges

Example 1

$k_i = d_i - 1 = \mathbf{8} - 1 = 7$

$k_j = d_j - 1 = \mathbf{2} - 1 = 1$

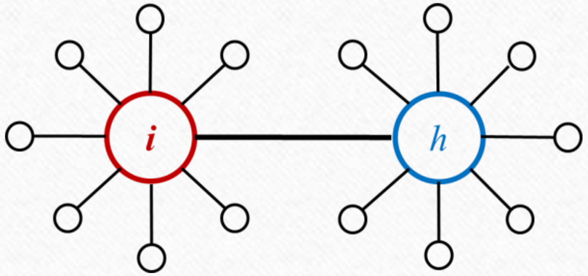
$k_f = d_f - 1 = \mathbf{1} - 1 = 0$



Edge (i, j) is **highly disassortative**: the **high**-degree node ***i*** is associated to the **low**-degree node ***j***

Example 2

$k_i = k_h$



Edge (i, h) is **highly assortative**: the **high**-degree node ***i*** is associated to the **high**-degree node ***h***

Assortative mixing in the correlation network

The **assortative mixing of a graph *G*** is computed as the **Pearson correlation coefficient of the excess degrees** of the nodes **at either ends of an edge** (Newmann, 2002) .

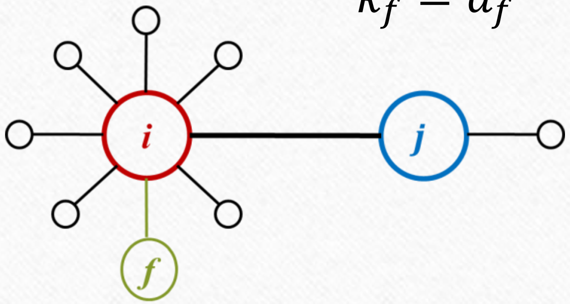
From **edge assortativity** → **Assortativity of G** → To **node contribution** to the assortativity of G (for **node-degree**: Piraveenan et al., 2008)

Example 1

$$k_i = d_i - 1 = \mathbf{8} - 1 = 7$$

$$k_j = d_j - 1 = \mathbf{2} - 1 = 1$$

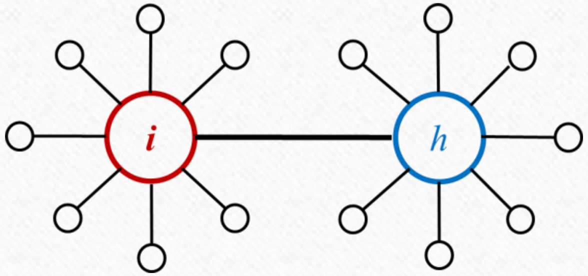
$$k_f = d_f - 1 = \mathbf{1} - 1 = 0$$



Edge (i, j) is **highly disassortative**: the **high**-degree node i is associated to the **low**-degree node $\boldsymbol{j}$

Example 2

$$k_i = k_h$$



Edge (i, h) is **highly assortative**: the **high**-degree node i is associated to the **high**-degree node $\boldsymbol{h}$

Assortativity-based portfolio optimization models

Assortativity-based portfolio selection models

Problem **domain**

VAR constraints

A target value R^{VAR} for the portfolio returns over time is fixed and this set of constraints allows that in **at most the δ% of the cases** this value is not reached.

γ is a fixed lower **bound** on the fractions invested in any asset

Neighborhood separation constraints

Portfolio constraint

$$\sum_{i=1}^{n} x_i = 1$$

$$\sum_{i=1}^{n} R_{it} x_i + M z_t \geq R^{VAR} \qquad t = 1,2,\ldots,T$$

$$\frac{1}{T}\sum_{t=1}^{T} z_t \leq \delta$$

$$\gamma y_i \leq x_i \leq y_i \qquad i = 1,2,\ldots,n$$

$$y_i + \sum_{j \in N(i)} y_j \leq 1 \qquad i = 1,2,\ldots,n$$

$$x_i \geq 0 \qquad i = 1,2,\ldots,n$$

$$y_i \in \{0,1\} \qquad i = 1,2,\ldots,n$$

$$z_t \in \{0,1\} \qquad t = 1,2,\ldots,T$$

Assortativity-based portfolio selection models

Problem **objective function**

$$\max_x \sum_{i=1,2,\dots,n} E(R_i)x_i - \sum_{i=1,2,\dots,n} \boldsymbol{r}_i y_i$$

$M_{D-AssMix}$ – Degree assortativity model

$$\max_x \sum_{i=1,2,\dots,n} E(R_i)x_i - \sum_{i=1,2,\dots,n} \boldsymbol{Sr}_i y_i$$

$M_{S-AssMix}$ – Strength assortativity model

$$\max_x \sum_{i=1}^{n} E(R_i)x_i$$

$\boldsymbol{M_{Bench}}$

$$\sum_{i=1}^{n} x_i = 1$$

$$\sum_{i=1}^{n} R_{it}x_i + Mz_t \geq R^{VAR} \qquad t = 1,2,\dots,T$$

$$\frac{1}{T}\sum_{t=1}^{T} z_t \leq \delta$$

$$\gamma y_i \leq x_i \leq y_i \qquad i = 1,2,\dots,n$$

$$y_i + \sum_{j\in N(i)} y_j \leq 1 \qquad i = 1,2,\dots,n$$

$$x_i \geq 0 \qquad i = 1,2,\dots,n$$

$$y_i \in \{0,1\} \qquad i = 1,2,\dots,n$$

$$z_t \in \{0,1\} \qquad t = 1,2,\dots,T$$

Results

Results: in-sample performance

90 in-sample days

30 out-of-sample days

S&P500: daily observations, rolling time windows, average $\boldsymbol{\mu_{in}}$ and $\boldsymbol{\sigma_{in}}$

θ	Density	AssMix)	Model	μ_{in}	σ_{in}	PTO	PSW	Av Num Assets	Max Num Assets
0.25	0.56	–	M_{Bench}	**0.0046**	0.0199	1.70	4.48	2.66	8
		$r = 0.044$	$M_{D-AssMix}$	0.0023	0.0148	1.75	6.64	3.5	19
		$Sr = 0.014$	$M_{S-AssMix}$	0.0023	**0.0147**	1.50	12.62	7.93	20
0.30	0.42	–	M_{Bench}	**0.0048**	0.0206	1.66	5.15	3.15	8
		$r = 0.060$	$M_{D-AssMix}$	0.0029	0.0164	1.66	9.01	4.98	20
		$Sr = 0.033$	$M_{S-AssMix}$	0.0019	**0.0132**	1.54	20.01	11.70	20
0.35	0.30	–	M_{Bench}	**0.0050**	0.0216	1.64	5.60	3.5	8
		$r = 0.085$	$M_{D-AssMix}$	0.0031	0.0170	1.66	11.68	6.28	20
		$Sr = 0.007$	$M_{S-AssMix}$	0.0017	**0.0124**	1.53	24.69	14.88	20
0.40	0.20	–	M_{Bench}	**0.0051**	0.0217	1.60	5.97	3.70	8
		$r = 0.14$	$M_{D-AssMix}$	0.0033	0.0169	1.60	14.55	8.13	20
		$Sr = 0.12$	$M_{S-AssMix}$	0.0014	**0.0121**	1.55	28.15	16.30	20

- M_{Bench} is obviously better for μ_{in}, but not for σ_{in}
- $M_{S-AssMix}$ and $M_{D-AssMix}$ tends to include more assets in the portfolio then M_{Bench}

Daily observations from 06/10/2006 to 31/12/2020 (3715 observations)

Results: out-of-sample performance

S&P500: daily observations, rolling time windows, average $\boldsymbol{\mu_{out}}$ and $\boldsymbol{\sigma_{out}}$

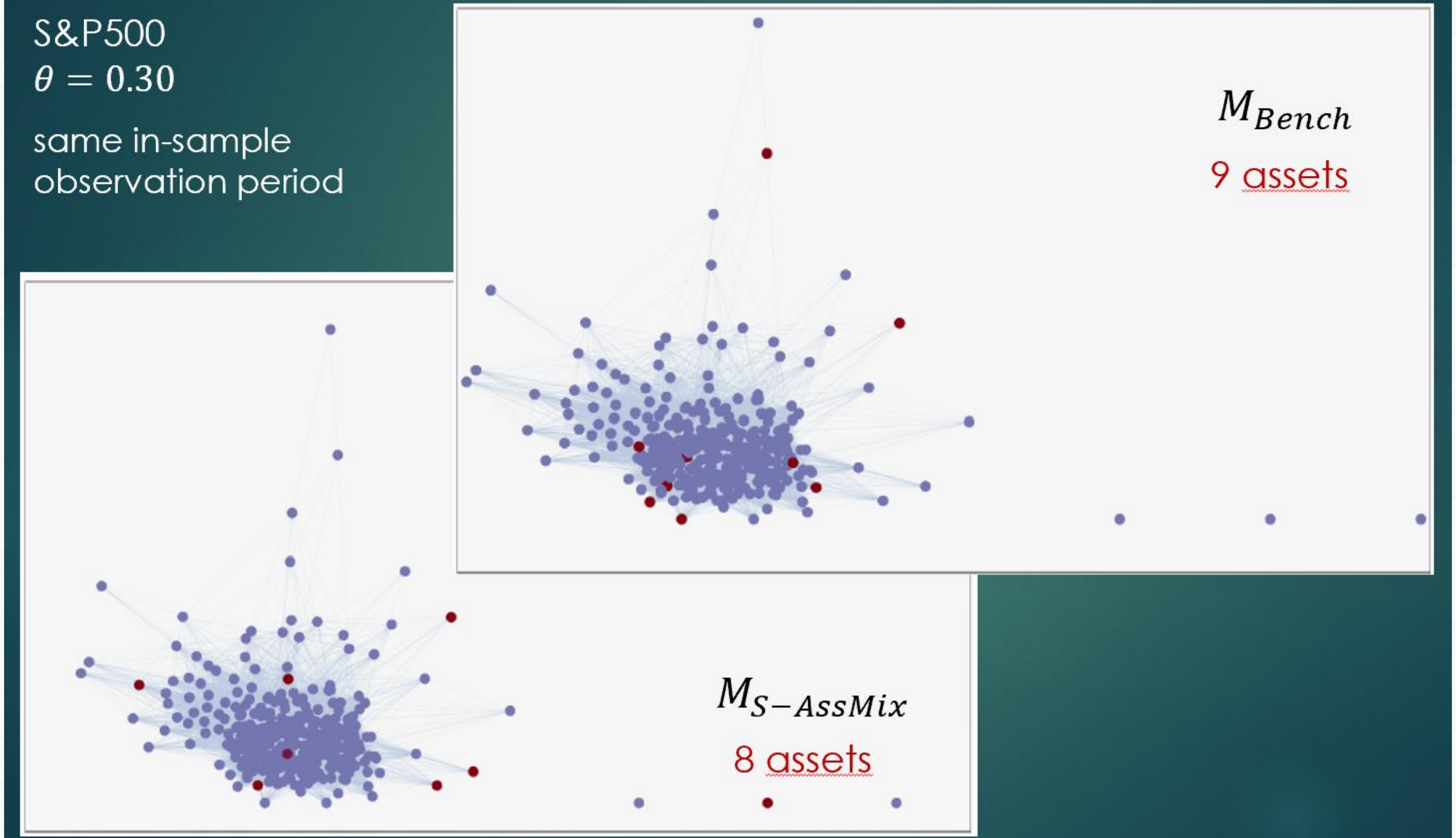
θ	Density	AssMix	Model	μ_{out}	σ_{out}	Sharpe	Time (Sec.)
0.25	0.56	–	M_{Bench}	0.0006	0.0175	0.0582	5.74
		$r = 0.044$	$M_{D-AssMix}$	0.0004	**0.0142**	0.0551	3.53
		$Sr = 0.014$	$M_{S-AssMix}$	**0.0009**	0.0146	**0.0829**	3.63
0.30	0.42	–	M_{Bench}	**0.0007**	0.0183	0.0559	14.80
		$r = 0.06$	$M_{D-AssMix}$	0.0006	0.0152	0.0638	3.76
		$Sr = 0.033$	$M_{S-AssMix}$	**0.0007**	**0.0127**	**0.1017**	3.76
0.35	0.30	–	M_{Bench}	0.0005	0.0187	0.0496	7.66
		$r = 0.085$	$M_{D-AssMix}$	**0.0007**	0.0159	0.0672	3.85
		$Sr = 0.007$	$M_{S-AssMix}$	0.0005	**0.0120**	**0.0911**	3.70
0.40	0.20	–	M_{Bench}	**0.0007**	0.0189	0.0559	5.78
		$r = 0.14$	$M_{D-AssMix}$	0.0005	0.0153	0.0596	4.51
		$Sr = 0.12$	$M_{S-AssMix}$	0.0005	**0.0115**	**0.0880**	3.10

The average indices of the optimal solutions found by our assortativity-based models are globally better than those of M_{Bench}

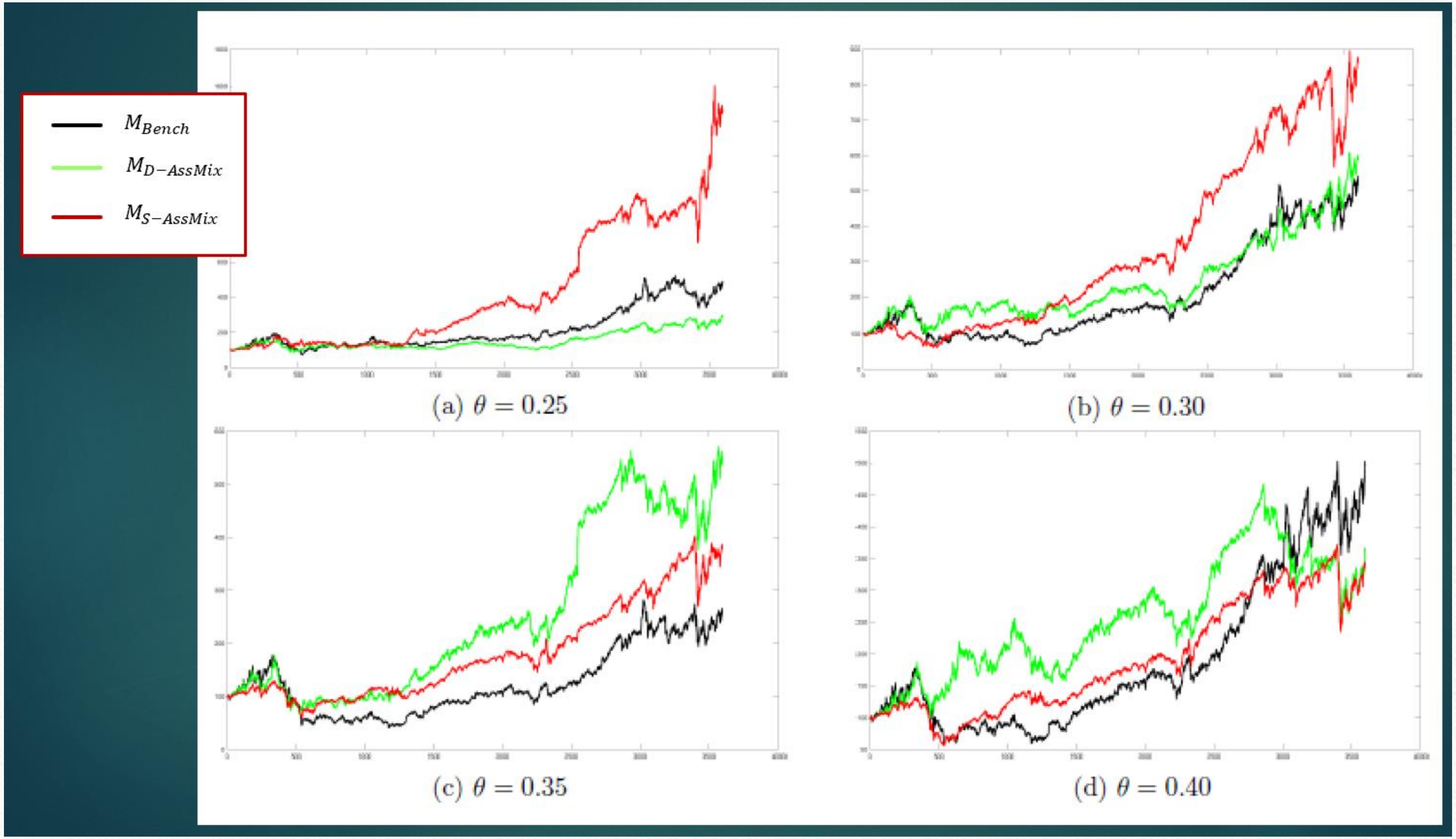
Results: optimal solutions

S&P500
$\theta = 0.30$

same in-sample observation period



Results: cumulated values



Conclusions

We introduce **new Mixed Integer Linear Programs** for portfolio optimization which maximize the portfolio expected return and prevent high positively correlated assets to be selected together:

- Network representation of the financial market $\boldsymbol{G_\theta}$
- Local assortativity coefficients for nodes of $\boldsymbol{G_\theta}$
- Neighborhood separation constraints

Our results show that the **combination** of these two new features is able to control portfolio returns variability and **favor diversification**, thus producing good out-of-sample performance of the optimal portfolios found.

Additional slides

Assortative mixing in the correlation network

Assortative mixing in the correlation network

The **assortative mixing of a graph *G*** is computed as the **Pearson correlation coefficient of the excess degrees** of the nodes **at either ends of an edge** (Newmann, 2002) .

From **edge assortativity** → **Assortativity of G** → To **node contribution** to the assortativity of G (for node-degree: Piraveenan et al., 2008)

(Newmann, 2002)

$$r = \frac{m^{-1}\sum_{(i,j)\in E} k_i k_j - \left[m^{-1}\sum_{(i,j)\in E}\frac{1}{2}(k_i+k_j)\right]^2}{m^{-1}\sum_{(i,j)\in E}(k_i^2+k_j^2) - \left[m^{-1}\sum_{(i,j)\in E}\frac{1}{2}(k_i+k_j)\right]^2}$$

Local assortativity coefficient for a node *i* of *G*
(Piraveenan, Prokopenko, Zomaya, 2008):

$$r_i = \frac{d_i k_i(\bar{k}_{N(i)} - \mu_{q(k)})}{2m\sigma^2_{q(k)}}$$

$\mu_{q(k)}$ and $\sigma^2_{q(k)}$:
mean and the variance of the node excess degree *k*.

Local assortativity coefficient for a node i of G
(Piraveenan, Prokopenko, Zomaya, 2008):

$$Sr_i = \frac{es_i\sum_{j\in N(i)} es_j - d_i\, es_i\, \mu_{q(s)}}{2m\sigma^2_{q(s)}} = \frac{d_i\, es_i\,(\bar{es}_{N(i)} - \mu_{q(s)})}{2m\sigma^2_{q(s)}}$$

$\mu_{q(s)}$ and $\sigma^2_{q(s)}$:
mean and the variance of the excess strength *s*.

Generalized assortativity coefficient

$$r^{\omega}_{(\alpha,\beta)} = \frac{\sum_{(i,j)\in E} w^{\beta}_{ij}\, es^{in}_{j}\, es^{out}_{i} - \Omega^{-1}\left(\sum_{(i,j)\in E} w^{\beta}_{ij}\, es^{in}_{j}\right)\left(\sum_{(i,j)\in E} w^{\beta}_{ij}\, es^{out}_{i}\right)}{\sqrt{\left[\sum_{(i,j)\in E} w^{\beta}_{ij}\,(es^{in}_{j})^2 - \Omega^{-1}\left(\sum_{(i,j)\in E} w^{\beta}_{ij}\, es^{in}_{j}\right)^2\right]\left[\sum_{(i,j)\in E} w^{\beta}_{ij}\,(es^{out}_{i})^2 - \Omega^{-1}\left(\sum_{(i,j)\in E} w^{\beta}_{ij}\, es^{out}_{i}\right)^2\right]}}$$

Generalized assortativity coefficient		
$r^{\omega}_{(\alpha,\beta)}$	$\beta = 0$	$\beta = 1$
$\alpha = 0$	Newmann's degree assortativity [36]	Weighted degree assortativity [26]
$\alpha = 1$	Our strenght assortativity	Weighted strenght assortativity [1]

Table 1: Generalized assortativity coefficient for all possible combinations of values for α and β.

- We **enhance and motivate** the use in portfolio selection of one specific strength assortativity index, called $\boldsymbol{r^{w}_{(1,0)}}$, which was introduced in the literature theoretically, but not actually applied in real-life problems (U. Pigorsch, M. Sabek, 2022).

Generalized assortativity coefficient

$$r^{\omega}_{(\alpha,\beta)} = \frac{\sum_{(i,j)\in E} w^{\beta}_{ij}\, es^{in}_{j}\, es^{out}_{i} - \Omega^{-1}(\sum_{(i,j)\in E} w^{\beta}_{ij}\, es^{in}_{j})(\sum_{(i,j)\in E} w^{\beta}_{ij}\, es^{out}_{i})}{\sqrt{[\sum_{(i,j)\in E} w^{\beta}_{ij}\,(es^{in}_{j})^2 - \Omega^{-1}(\sum_{(i,j)\in E} w^{\beta}_{ij}\, es^{in}_{j})^2][\sum_{(i,j)\in E} w^{\beta}_{ij}\,(es^{out}_{i})^2 - \Omega^{-1}(\sum_{(i,j)\in E} w^{\beta}_{ij}\, es^{out}_{i})^2]}}$$

Generalized assortativity coefficient		
$r^{\omega}_{(\alpha,\beta)}$	$\beta = 0$	$\beta = 1$
$\alpha = 0$	Newmann's degree assortativity [36]	Weighted degree assortativity [26]
$\alpha = 1$	Our strenght assortativity	Weighted strenght assortativity [1]

Table 1: Generalized assortativity coefficient for all possible combinations of values for α and β.

[1] A. Arcagni, R. Grassi, S. Stefani, A. Torriero (2021). Extending assortativity: an application to weighted social networks. Journal of Business Research, 129, 774–783.

[26] C.C. Leung; H.F. Chau (2007). Weighted assortative and disassortative networks model. Physica A: Statistical Mechanics and its Applications, 378, 591–602.

[36] M.E.J. Newmann (2002). Assortative mixing in networks. Physical review letters, 89, 208701.

Network analysis for portfolio selection: related literature

Network analysis for portfolio selection

In the literature of the last decade the use of **networks** in the analysis of the **financial market** has been extensively applied to study different aspects:

- **Caccioli et al. 2018** and **Neveu 2018**: the links of the network connecting financial institutions are viewed as channels for the propagation of risk and financial **systemic risk** and contagion in financial markets is studied through the analysis of the financial network.

- **Boginski et al. (2014)**: study **clustering dynamics** in the market correlation network by solving combinatorial optimization problems related to **cliques**, such as partitioning the market graph into a minimum number of distinct cliques, or finding the clique of maximum size.

Network analysis for portfolio selection

In particular, network structures and properties can be exploited for controlling portfolio volatility and enhancing **diversification**.

- **Peralta and Zareei (2016)**: establish a negative relationship between Markowitz's optimal portfolio fractions and the **centrality** of nodes representing assets in the financial market network.
- **Puerto et al. (2020)**: propose a Mixed Integer Linear Program for dealing with asset clustering and portfolio selection simultaneously. The financial network is endowed with a metric based on the correlation coefficients between assets' returns, and classical **location problems** on networks are implemented for clustering assets using such metric.
- **Clemente, Grassi, and Hitaj (2021)**: use the node **clustering coefficients** from network analysis to measure **global minimum interconnectedness between assets' returns**.

Network analysis for portfolio selection

Markowitz, 1952

GMV model

$$\min_x x^T \Sigma x$$
$$e^T x = 1$$
$$x \geq 0$$

x Portfolio fractions

Σ Covariance matrix

Clemente et al., 2021

GMH model

$$\min_x x^T \boldsymbol{H} x$$
$$e^T x = 1$$
$$x \geq 0$$

- Financial correlation network for the dependence of assets returns
- **interconnection matrix H** based on **node clustering coefficients**
- Portfolio expected return is not considered.

$$\boldsymbol{H} = \Delta^{\boldsymbol{T}} \boldsymbol{C} \Delta \qquad \Delta = diag(s_i)$$

$$c_{ij} = \begin{cases} C_i C_j & if\ i \neq j \\ 1 & otherwise \end{cases} \qquad s_i = \frac{\sigma_i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}}$$

- **Clemente, Grassi, and Hitaj (2021)**: use the node **clustering coefficients** from network analysis to measure **global minimum interconnectedness between assets' returns**.