An importance and frequency-based approach for sparse online portfolios

Gustavo Ribeiro de Oliveira Roque[†, a]
Carlos Cristiano Hasenclever Borges[‡, a]
João Daniel Madureira Yamim[*, a]

[a]Universidade Federal de Juiz de Fora

Abstract Portfolio optimization has advanced since Markowitz, yet dynamically adjusting weights remains a challenge. This study applies Online Gradient Descent (OGD) and Empirical Mode Decomposition (EMD) to construct and optimize portfolios in the Brazilian stock market. By decomposing financial time series into fast, medium, and slow modes, we build correlation-based networks and select stocks using centrality measures. Portfolios (5–30 stocks) are optimized via OGD and evaluated against a benchmark. Results show that small *RIS* portfolios consistently outperform large *RIS* portfolios in CAGR and VaR. The fast mode performs best, with all fast small *RIS* portfolios surpassing the benchmark. The mixed mode ($\gamma = 0.2, 0.35$) achieves superior results, with the best portfolios exceeding an 18% CAGR. Overall, 36% of mixed mode portfolios outperform the benchmark, with small *RIS* portfolios dominating the optimal region. These findings demonstrate the effectiveness of EMD-based stock selection and OGD for risk-adjusted returns.

Keywords: Portfolio optimization; Online Gradient Descent; Empirical Mode Decomposition; Risk measures; Brazilian stock market.

JEL codes: G11, C61.

1. Introduction

Portfolio optimization has encountered numerous challenges since Harry Markowitz established the field as a formal discipline in his seminal work (Markowitz, 1952). One of the key issues arising from Markowitz's model is how to dynamically adjust portfolio weights over time. His mean-variance approach determines allocations based on historical data, identifying the optimal past allocation but failing to address a crucial forward-looking question: given

How to cite: Roque, G. R. O., Borges, C. C. H. & Yamim, J. D. M. (2025). An importance and frequency-based approach for sparse online portfolios. *Revista Brasileira de Finanças*, *23*, e202505. https://doi.org/10.12660/rbfin.v23n1.2025.93034
Submitted on February 11, 2025. Published online on March 25, 2025. Editor in charge: Bruno Giovannetti.
[†]gustavo.roque@estudante.ufjf.br
[‡]carlos.cristiano@ufjf.br
[*]joao.yamim@engenharia.ufjf.br

©2025 Sociedade Brasileira de Finanças, under a Creative Commons Attribution 4.0 International license.

a set of available stocks, how should an investor allocate assets for the future? In this context, online optimization algorithms offer a compelling solution by enabling dynamic portfolio adjustments.

The application of machine learning techniques to portfolio optimization builds on foundational research (Bell and Cover, 1980, 1988; Kelly, 1956; Algoet and Cover, 1988) demonstrating that continuously rebalancing a portfolio can yield better results than selecting an allocation based on past performance and holding it unchanged. Furthermore, Cover (1991) introduced the concept of the Best Constant Rebalanced Portfolio (BCRP), a strategy that ensures the highest long-term wealth accumulation through constant portfolio reallocation. Online portfolio optimization algorithms aim to approximate the BCRP by dynamically adjusting allocations as new price data becomes available. Their performance is typically assessed using the regret function, as proposed by Zinkevich (2003), which quantifies the difference between the cumulative returns of an online algorithm and those of an idealized offline strategy that has access to future information.

This study tackles the problem by implementing the Online Gradient Descent (OGD) algorithm, introduced by Martin Zinkevich in 2003 (Zinkevich, 2003). Designed to handle sequential decision-making under uncertainty, OGD updates portfolio allocations iteratively using gradient-based adjustments. Its simplicity and efficiency make it a fundamental tool in finance and machine learning, where real-time decision-making is crucial.

Before diving into online optimization, another question arises: given a dataset containing numerous stocks, is it necessary to include all of them in the optimization process, or can a reduced selection yield comparable or even superior results? Research by Yang et al. (2019) suggests that portfolios can be constructed using Empirical Mode Decomposition (EMD), a technique first introduced by Norden Huang in 1998 (Huang et al., 1998). EMD decomposes financial time series into multiple intrinsic mode functions (IMFs), each representing a different frequency component. An IMF must satisfy two key conditions: (1) the number of zero crossings and extrema must be nearly equal, differing by at most one; (2) the mean value of the envelope formed by local maxima and minima must be zero at all points. While the first condition ensures balance between peaks and troughs, the second ensures symmetry between upper and lower envelopes.

After decomposing the data into IMFs, the approach in Yang et al. (2019) involves categorizing them into fast, medium, and slow frequency modes and computing Pearson correlation matrices for each category. These correlation matrices are then used to construct Planar Maximally Filtered Graphs (PM-

FGs), a technique for extracting essential information from complex financial networks (Tumminello et al., 2005).

Using PMFGs, we analyze the resulting networks and select stocks for portfolio construction based on in measure of the relevance of vertices connections. Portfolios of varying sizes (5, 10, 15, 20, 25, and 30 stocks) are formed by prioritizing stocks with the highest scores. These portfolios are then optimized using the Online Gradient Descent algorithm, and their performance is compared to a fully diversified portfolio to assess the benefits of stock selection. By experimenting with different portfolio sizes, we examine how risk and return evolve as the number of assets increases. Additionally, we investigate whether portfolios constructed from stocks with higher or lower scores achieve superior performance.

Prior research (Yang et al., 2019) has demonstrated that selecting stocks based on a centrality measure can influence risk, with expected shortfall—-a widely recognized coherent risk measure (Acerbi and Tasche, 2002)–exhibiting a dependency on portfolio size. As tested on portfolios ranging from 5 to 60 stocks, increasing diversification tends to lower risk, reinforcing the idea that diversification is a key strategy for mitigating losses. In this study, we explore how Compound Annual Growth Rate (CAGR) and Value at Risk (VaR) behave under similar conditions but in a different market and dataset, analyzing how these metrics impact portfolio performance.

2. Methodology

2.1 Online Gradient Descent

Online Gradient Descent (OGD) extends the conventional gradient descent method into an online framework and is utilized as a convex optimization technique. Given a convex function $f_t(x)$ at time step t, OGD updates its solution based on the following iteration:

$$x_{t+1} = \Pi_D\big(x_t - \eta_t \nabla_t f_t(x_t)\big), \tag{1}$$

where in portfolio optimization, x_t represents the weight vector, $\nabla f_t(x_t)$ is the gradient of the function at x_t, and η_t is the learning rate at time t. After adjusting the weights in the negative gradient direction, the OGD algorithm applies a projection function to ensure the updated weights remain within a convex feasible space. Initially, the weight vector b_0 is set as $1/n$, indicating equal distribution among all assets, and optimization proceeds from $t = 1$ to T, ultimately outputting x_{T+1} as the suggested allocation for the next round.

A key characteristic of OGD is that Zinkevich demonstrated that the Regret

function increases sublinearly over the number of iterations. This is crucial in online convex optimization, where multiple sequential optimizations occur, and deviation from the optimal solution is a concern. With an appropriate learning rate choice, we can ensure that Regret remains controlled, confirming that the algorithm effectively learns over time.

The Online Gradient Descent algorithm is outlined as follows:

Algorithm 1 Online Gradient Descent

Require: η

1: $b_1 \leftarrow 1/n$
2: **for** $t = 2$ to T **do**
3: $\quad x_t \leftarrow \Pi_D \left(x_t - \eta_t \frac{r_{t-1}}{x_{t-1}^T r_{t-1}} \right)$
4: **end for**
5: **Where** $\Pi_D(y) = \arg\min_{x \in D} ||y - x||$ and D is a convex set
6: **Return** x_{T+1}

2.2 EMD correlation-based network

As previously mentioned, asset selection in our portfolio involves applying Empirical Mode Decomposition (EMD) to each time series in our dataset, breaking it down into multiple Intrinsic Mode Functions (IMFs). These IMFs are categorized by frequency: the sum of the first three IMFs defines the fast mode, the fourth to sixth IMFs define the medium mode, and the remaining IMFs define the slow mode. This results in three distinct datasets—fast, medium, and slow—enabling the computation of Pearson cross-correlation matrices for each mode. These matrices are then processed using the Planar Maximally Filtered Graph (PMFG), which helps extract significant information while maintaining a manageable level of complexity. The resulting three complex systems are referred to as EMD correlation-based networks.

To illustrate the PMFG filtering process, consider a complex system with n elements, where S represents the similarity between pairs of elements—here, the Pearson correlation coefficient between assets. Constructing a descending ordered list S_1 of node pairs (i, j) based on the Pearson coefficient ρ_{ij} allows for systematic network extraction. Research in Mantegna (1999) highlights that using the Minimum Spanning Tree (MST) to retain only the strongest correlations offers valuable insights into financial systems. However, this approach excessively reduces the network structure, potentially discarding useful information. Thus, an alternative filtering method, the PMFG, is employed.

Like the MST, the PMFG constructs networks by prioritizing the highest correlation values but imposes an additional constraint: the graph must remain planar, meaning it can be embedded on a surface with genus $g = 0$. The algorithm follows the ordered list S_1 and adds edges between nodes if the resulting graph maintains this planarity condition. The final PMFG contains $3(n-2)$ edges, considerably more than the $(n-1)$ edges of the MST, preserving more meaningful relationships within the dataset.

Once the EMD correlation-based networks are built, they serve as a basis for portfolio selection using centrality metrics derived from graph theory. Two centrality measures introduced in Pozzi et al. (2013) are used:

$$X = \frac{C_D^w + C_D^u + C_{BC}^w + C_{BC}^u - 4}{4(n-1)} \tag{2}$$

$$Y = \frac{C_E^w + C_E^u C_C^w + C_C^u + C_{EC}^w + C_{EC}^u}{6(n-1)}, \tag{3}$$

where the superscripts w and u indicate weighted and unweighted measures, respectively. The subscripts denote specific centrality types: D for degree, BC for betweenness centrality, C for closeness, E for eccentricity, and EC for eigenvector centrality. These hybrid measures distinguish between different node connectivity scenarios:

- Small X, small Y: highly connected nodes linked to other highly connected nodes.
- Small X, large Y: highly connected nodes linked to sparsely connected nodes.
- Large X, small Y: sparsely connected nodes linked to highly connected nodes.
- Large X, large Y: sparsely connected nodes linked to other sparsely connected nodes.

To guide portfolio construction, we consider the difference of these measures, and call it Relative Influence Score: $RIS = X - Y$ (large for nodes with few important connections, small for those with many unimportant connections), while Yang et al. (2019) considers the sum. This is a critical difference between the works, as the sum represents a centrality measure, dividing the stocks into central and peripheral. Our approach is more related to the importance of stocks.

A detailed explanation of these centrality measures is available in Pozzi et al. (2013). Based on this methodology, we construct portfolios for small

and large values of beta with 5, 10, 15, 20, 25 and 30 assets, and analyze their evolution related to returns (CAGR) and risk (VaR).

3. Results

We utilized data from the Brazilian stock market, specifically focusing on a dataset comprising 84 stocks listed on the IBOVESPA (Índice da Bolsa de Valores de São Paulo). The IBOVESPA is the benchmark index of the B3 (Brasil Bolsa Balcão), the main stock exchange in Brazil, and serves as a key indicator of the performance of the Brazilian equity market. It is composed of the most liquid and representative stocks traded on the exchange, covering a diverse range of sectors such as finance, energy, commodities, and consumer goods. The index is weighted by market capitalization and trading volume, ensuring that it reflects the overall market trends and economic conditions in Brazil.

Firstly, we have selected the portfolios according to the technique described in section 2, using the Empirical Mode Decomposition to build complex networks where the vertices are the stocks in our dataset. After that, we have selected portfolios with 5, 10, 15, 20, 25 and 30 stocks based on the value of the measure *RIS*. With all the datasets ready, we performed online optimization via Online Gradient Descent on each dataset, and now we analyze the results.

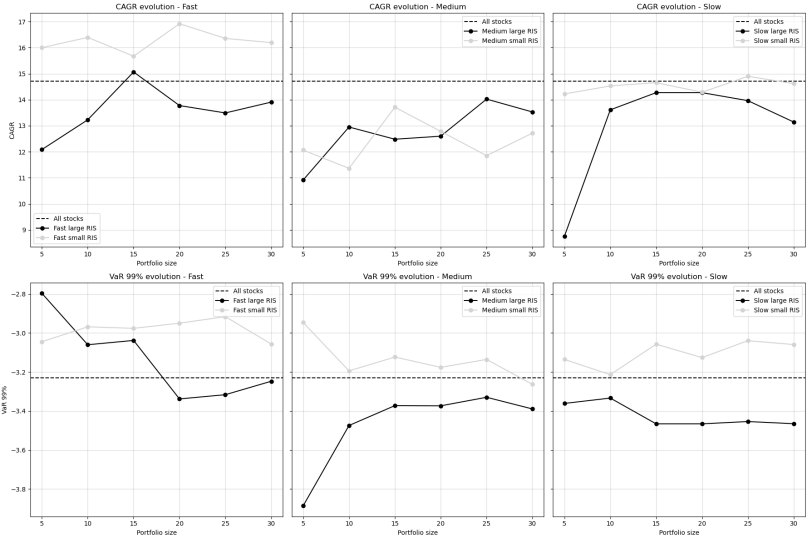
3.1 Fast, medium and slow modes

In Figure 1, we observe the evolution of the Compound Annual Growth Rate (CAGR) and Value at Risk (VaR), which represent the returns and risks of the portfolios, respectively. The results indicate that portfolios with smaller values of *RIS* consistently outperform those with larger values, both in terms of returns and risk management. Specifically, smaller *RIS* portfolios exhibit higher CAGR and lower VaR, demonstrating their superior performance and reduced risk exposure compared to portfolios with larger *RIS* values.

The fast mode small *RIS* portfolios demonstrated strong performance across all portfolio sizes, with the best-performing portfolio consisting of 20 stocks and the weakest performance observed in the 10-stock portfolio. Notably, all fast mode small *RIS* portfolios achieved a higher Compound Annual Growth Rate (CAGR) compared to the reference portfolio, which includes all stocks in the dataset. This suggests that the small *RIS* strategy effectively selects stocks that outperform the broader market.

In contrast, the fast mode large *RIS* portfolios exhibited relatively weaker performance. Among these, only the 10-stock portfolio surpassed the reference

Figure 1
Evolution of CAGR and VaR with the portfolio size.

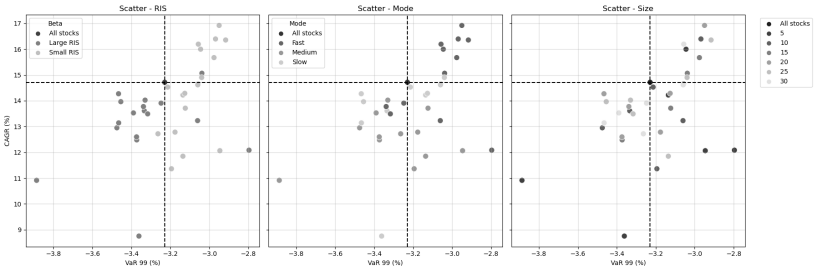


portfolio in terms of returns, while the 5-stock portfolio delivered the poorest results. An intriguing pattern emerged as the Value at Risk (VaR) became more negative, indicating higher risk levels, in tandem with improved returns for larger portfolio sizes. This implies that as the portfolio size increases, the fast mode large *RIS* strategy tends to generate higher returns but at the cost of greater risk exposure.

In the medium mode, all portfolios underperform in terms of returns compared to the benchmark. Regarding risk, small *RIS* portfolios tend to be less risky, but the risk increases as the portfolio size grows. In contrast, large *RIS* portfolios exhibit the opposite trend, with risk decreasing as the portfolio size increases. In the slow mode, small *RIS* portfolios demonstrate better performance in both returns and risk, closely aligning with the benchmark's performance while maintaining lower risk levels.

In Figure 2, it is evident that small *RIS* portfolios outperform large *RIS* portfolios, with only one large *RIS* portfolio located in the optimal region (top right), representing higher returns with a favorable risk profile. Overall, small *RIS* portfolios tend to be less risky and are concentrated in the right regions of the plot, while large *RIS* portfolios are riskier and deliver lower returns, clustering in the bottom left region.

Figure 2
Scatter plots for VaR x CAGR.



Regarding the modes, the fast mode demonstrates the best performance, with seven portfolios positioned in the optimal region. The slow mode follows as the second-best option, with some portfolios achieving comparable returns but with lower risk. As for portfolio size, no clear trend emerges. Well-performing portfolios, such as fast small *RIS* and slow small *RIS*, tend to maintain consistent levels of CAGR and VaR as the portfolio size increases.

3.2 Mixed Mode Results

Based on the results, we have introduced a new mode, referred to as the mixed mode, which combines the characteristics of both the fast and slow modes. The mixed mode is defined as the following convex combination:

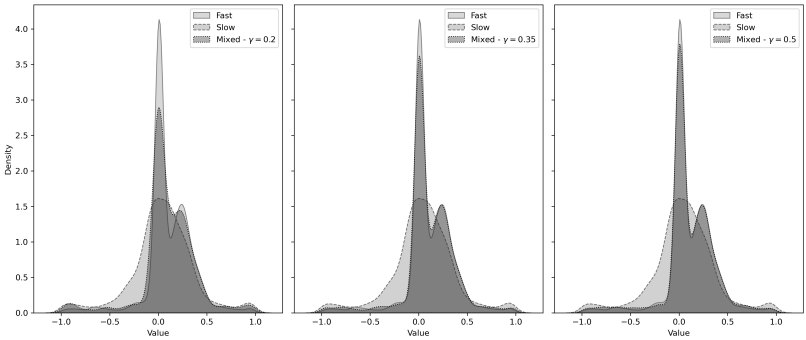
$$\text{Mixed mode} = \frac{(\gamma \cdot \text{Fast mode} + (1-\gamma) \cdot \text{Slow mode})}{2}$$

Here, γ is a weighting parameter that determines the relative contribution of each mode. In our analysis, we selected $\gamma = 0.2$, $\gamma = 0.35$ and $\gamma = 0.5$.

As seen in Figure 3, the fast mode exhibits a correlation matrix with a prominent peak centered around 0, a smaller peak to the right, and a slight tail. In contrast, the slow mode displays a more flattened distribution without significant peaks and features a broader tail. To effectively capture the combined behavior of both modes in the mixed-mode distribution, it was essential to assign a higher weight to the slow mode by choosing smaller values of γ. This ensures that the mixed mode retains the flattened profile and extended tail of the slow mode while incorporating the peaked structure of the fast mode.

This behavior is evident when choosing $\gamma = 0.2$. However, as γ increases to 0.35 and 0.5, the fast mode becomes more dominant. This shift occurs because the fast mode contains more concentrated values. Most values in the fast mode correlation matrix fall between 0 and 0.5, whereas the slow mode

Figure 3
Density of values in the correlation matrices for the fast, slow and mixed modes.



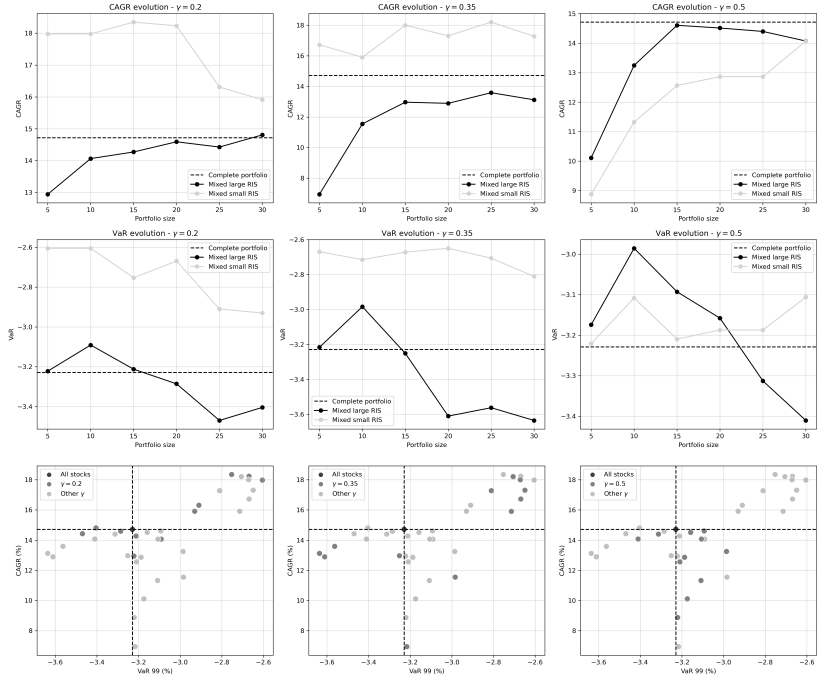
values are more widely spread within the interval $[-1,1]$. This distribution pattern pulls the mixed mode values closer to the fast mode, emphasizing its characteristics.

Figure 4 compares the results of the mixed mode with the fast and slow modes. The mixed mode demonstrates strong performance, particularly for small γ values, such as $\gamma = 0.2$ and $\gamma = 0.35$. These portfolios outperform the benchmark in both CAGR and VaR, even surpassing the results of the fast mode.

For $\gamma = 0.5$, the behavior reverses: portfolios with larger *RIS* values achieve better returns than those with smaller *RIS*. However, in this case, none of the selected portfolios outperform the benchmark. One possible explanation is that medium γ values may not be ideal. Smaller values produce outstanding portfolios, while values closer to 1 shift the distribution toward the fast mode, potentially improving performance. In terms of risk, up to 20 assets, all portfolios remain less risky than the benchmark. Beyond that point, small *RIS* portfolios continue to outperform, whereas large *RIS* portfolios underperform.

The scatter plot illustrates the mixed-mode portfolios for both large and small *RIS* values. While the plot does not explicitly differentiate between them, the results clearly show that the underperforming portfolios within the mixed mode for $\gamma = 0.2$ and $\gamma = 0.35$ correspond to large *RIS* values, which cluster in the lower region of the plot in terms of CAGR. Conversely, for these values of γ, the top-right region of the scatter plot contains six mixed-mode portfolios, highlighting their strong performance in both CAGR and VaR. This is a highly encouraging result, underscoring the effectiveness of the mixed-mode approach.

Figure 4
Results for the mixed mode.



On the other hand, for $\gamma = 0.5$, only two portfolios achieve similar levels of return with lower risk, but neither reaches the optimal region.

4. Conclusion

From the results, it is evident that the mixed mode with $\gamma = 0.2$ and $\gamma = 0.35$ produces the best portfolios in our analysis, with a CAGR value larger than 18%, while the best portfolio for the fast, medium and slow mode has achieved a CAGR of 17%.

Overall, 13 mixed mode portfolios were generated that outperformed the benchmark across all assets, representing a significant 36% of all portfolios created with this mode. Notably, all but two portfolios that surpassed the benchmark were small *RIS* portfolios. Among the 36 small *RIS* portfolios, 19 were located in the optimal region, outperforming the benchmark in both risk and return, a percentage of 52.78%. When focusing specifically on the

fast mode and mixed mode with small *RIS*, this percentage increases to 100% and 66.67%, respectively. These results suggest that concentrating on small *RIS* portfolios within these modes can lead to superior outcomes. Also, all the small *RIS* portfolios that did not surpassed the benchmark were mixed mode portfolios with $\gamma = 0.5$, indicating that smaller (and, perhaps, larger values) of γ produces better portfolios.

Acknowledgments The authors would like to thank CAPES for the financial support and the Postgraduate Program in Computational Modeling at UFJF for providing the infrastructure that made this work possible.

Conflict of interest The authors declare no conflict of interest.

Artificial Intelligence This research utilized AI tools to assist in data analysis, manuscript drafting, and figure generation. All AI-generated content was critically reviewed and validated by the authors to ensure accuracy and alignment with the scientific integrity of the study. The use of AI adhered to ethical guidelines, ensuring transparency and compliance with academic standards. Any biases or limitations inherent to the AI tools were carefully considered in the interpretation of results. The authors affirm that the AI tools did not compromise the originality or integrity of the work.

References

Acerbi, C. and Tasche, D. (2002). Expected shortfall: A natural coherent alternative to value at risk, *Economic Notes*, *31*(2), 379–388.

Algoet, P. H. and Cover, T. M. (1988). Asymptotic optimality and asymptotic equipartition properties of log-optimum investment, *The Annals of Probability*, pp. 876–898.

Bell, R. and Cover, T. M. (1988). Game-theoretic optimal portfolios, *Management Science*, *34*(6), 724–733.

Bell, R. M. and Cover, T. M. (1980). Competitive optimality of logarithmic investment, *Mathematics of Operations Research*, *5*(2), 161–166.

Cover, T. M. (1991). Universal portfolios, *Mathematical Finance*, *1*(1), 1–29.

Huang, N. E. et al. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proceedings*

of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, *454*(1971), 903–995.

Kelly, J. L. (1956). A new interpretation of information rate, *The Bell System Technical Journal*, *35*(4), 917–926.

Mantegna, R. N. (1999). Hierarchical structure in financial markets, *The European Physical Journal B-Condensed Matter and Complex Systems*, *11*, 193–197.

Markowitz, H. (1952). The utility of wealth, *Journal of Political Economy*, *60*(2), 151–158.

Pozzi, F., Di Matteo, T. and Aste, T. (2013). Spread of risk across financial markets: Better to invest in the peripheries, *Scientific Reports*, *3*(1), 1665.

Tumminello, M. et al. (2005). A tool for filtering information in complex systems, *Proceedings of the National Academy of Sciences*, *102*(30), 10421–10426.

Yang, L., Zhao, L. and Wang, C. (2019). Portfolio optimization based on empirical mode decomposition, *Physica A: Statistical Mechanics and its Applications*, *531*, 121813.

Zinkevich, M. (2003). Online convex programming and generalized infinitesimal gradient ascent, *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*.