ORIGINAL ARTICLE

Network Risk Parity: graph theory-based portfolio construction

Vito Ciciretti[1] · Alberto Pallotta[2] (ORCID)

Abstract

This study presents network risk parity, a graph theory-based portfolio construction methodology that arises from a thoughtful critique of the clustering-based approach used by hierarchical risk parity. Advantages of network risk parity include: the ability to capture one-to-many relationships between securities, overcoming the one-to-one limitation; the capacity to leverage the mathematics of graph theory, which enables us, among other things, to demonstrate that the resulting portfolios is less concentrated than those obtained with mean-variance; and the ability to simplify the model specification by eliminating the dependency on the selection of a distance and linkage function. Performance-wise, due to a better representation of systematic risk within the minimum spanning tree, network risk parity outperforms hierarchical risk parity and other competing methods, especially as the number of portfolio constituents increases.

Keywords Portfolio construction · Graph theory · Hierarchical clustering · Eigenvalues

JEL classification G11

Introduction

The seminal asset allocation model of Markowitz (1952) has been the stronghold of portfolio construction since 1952. However, extensive research documents three main limitations, namely producing unstable, concentrated, and underperforming portfolios. Michaud (1989) provides a detailed exploration of these issues. The finance literature has proposed numerous solutions, including the works of Black and Litterman (1992) and Ledoit and Wolf (2004), among others.

The literature on portfolio construction has recently focused on two fields of application: clustering and graph theory. On the clustering side, the main representative is the hierarchical risk parity (HRP) of de Prado (2016), which applies hierarchical clustering to securities based on correlations. This groups similar investments while distancing dissimilar ones, with optimal allocation achieved via inverse-variance allocation. Raffinot (2017) further provided a comprehensive framework for portfolio construction employing hierarchical clustering.

Peralta and Zareei (2016) marked the pioneering step with the inclusion of graph theory in the portfolio construction literature. They demonstrated the close relationship between graph centrality measures and optimal portfolio weights. Furthermore, they designed portfolios by equally distributing capital to the most central securities in low-volatility periods and rebalancing to the least central securities during high-volatility periods. Vỳrost et al. (2019) constructed optimal portfolios based on four graphical representations of securities: a complete graph, a minimum spanning tree, a planar maximally filtered graph, and a threshold significance graph.

This paper aims at introducing network risk parity (NRP), a novel graph theory-based portfolio construction method that produces fully invested, long-only portfolios. We derive network risk parity by drawing a parallel with hierarchical risk parity. Indeed, while hierarchical risk parity is based on hierarchical clustering, Network risk parity builds portfolios with graph theory. The connection point between both methodologies is that they calculate portfolio weights proportionally to the inverse of the eigenvalues of a modified covariance matrix. In hierarchical risk parity, the optimal

✉ Alberto Pallotta
a.pallotta@mdx.ac.uk

Vito Ciciretti
vciciretti8@gmail.com

1 Berlin, Germany

2 Middlesex University, The Burroughs, Hendon, London NW4 4BT, UK

weights are calculated as the inverse of each security's variance. This variance is obtained from the main diagonal of a quasi-diagonal covariance matrix, which is modified to incorporate the hierarchical structure derived using hierarchical clustering. We demonstrate that the weighting system employed by HRP is equivalent to taking the inverse of the eigenvalues of the same covariance matrix, after an essential quasi-diagonalization step. On the other hand, network risk parity operates on a different principle. It utilizes an adjacency matrix based on covariances, where one set of eigenvectors is determined by the eigenvector centrality, a measure of the influence of a node in a network. In NRP, the optimal portfolio weights are calculated as the inverse of the eigenvector centrality. A softmax normalization is applied to these weights to ensure a fully invested, long-only portfolio.

Three advantages of network risk parity stem from transposing clustering-based methodologies into a graph theory framework. Firstly, NRP, grounded on the principle of eigenvector centrality, encapsulates relationships among securities in a one-to-many fashion, resulting from the definition of eigenvector centrality that embeds the importance of neighboring nodes too. Hierarchical clustering, on the other hand, captures only one-to-one relationships due to its agglomerative clustering approach, which looks at pairwise distances. Secondly, hierarchical clustering depends on the determination of a distance and a linkage function. NRP, instead, is based on minimum spanning tree (MST), which solely depends on the function used to convert correlations into distances. Hierarchical risk parity in particular has been criticized for its adoption of a single linkage function, which differentiates securities based on the distance of the nearest points within clusters, thereby causing a chaining effect that expands the tree and impacts portfolio weights (Papenbrock 2011). Third, leveraging on graph theory, we can prove that the portfolio weights of NRP and HRP[1] have a lower bound larger than zero, thus assigning a positive weight to each portfolio constituent and improving portfolio diversification, a pitfall of the classic mean-variance approach.

Using a bootstrapping approach, we compare the Sharpe ratio of NRP with that of HRP, risk parity (RP), Markowitz's minimum variance optimization (MVO), and equally weighted (EW) portfolios. While NRP consistently outperforms the other methods, compared to HRP, its performance depends on the number of stocks in the portfolio, outperforming HRP as the portfolio size increases.

The rest of the paper is organized as follows: Section 2 introduces NRP and compares it to HRP. Section 3 presents the empirical results in terms of bootstrapped Sharpe ratio and weights. Section 4 concludes and discusses future work.

Methodology

Covariance

The starting point of network risk parity, similar to most portfolio construction methodologies, is to estimate the covariance matrix of asset returns. We use hourly security prices and define log-returns as:

$$r_{i,\tau} = \ln\left(\frac{p_{i,\tau}}{p_{i,\tau-1}}\right) \quad \forall \tau = 1, \ldots, T; \quad \forall i = 1, \ldots, n$$

where $r_{i,\tau}$ is the return of the i-th asset at time τ, $p_{i,\tau}$ is the price of the i-th asset at time τ, and T is the number of hours in each month. We estimate the conditional covariance by employing the methodology of Barndorff-Nielsen and Shephard (2004), who show that conditional covariances can be estimated through non-parametric realized covariances, RC_t, which converge in probability to the quadratic variation of the price process under very general assumptions. Thus, we calculate monthly realized covariances, RC_t, as the aggregation of cross-products of hourly returns, such that:

$$RC_t = \sum_{\tau=1}^{N_t} r_\tau r_\tau'. \tag{1}$$

This ensures positive definite realized covariance matrices and makes covariance fully observable and moldable with any time series model. The use of hourly returns in constructing our portfolios helps to provide a higher number of observations, which in turn increases the robustness of our covariance matrix estimation.[2] It should be noted, however, that network risk parity is a covariance-agnostic portfolio construction method. While accurate covariance estimation is crucial for any portfolio construction approach, the primary focus of this paper is not to provide an improved estimate of covariances.

Review of Hierarchical Risk Parity

Hierarchical Risk Parity produces portfolios from a three-step process. Firstly, hierarchical clustering is performed on correlation-based distance measures, starting from the correlation matrix of asset returns. Since clustering requires a distance measure to encapsulate the relationship between

---

[1] As the process of clustering a correlation matrix can be re-written as a tree on a complete digraph, the weight floor applies to HRP as well.

[2] In this study, we aggregate hourly returns into monthly covariances, and we assume the portfolios are rebalanced monthly.

securities, we transform correlations into Euclidean metrics.[3] For this transformation, Mantegna ([1999](#)) defines the distance $d_{i,j}$ as $d_{i,j} = 1 - \rho_{i,j}^2$.[4] The AGNES hierarchical clustering algorithm is then applied to these distances, separating securities into clusters organized in a linkage matrix. In the second step, the linkage matrix undergoes quasi-diagonalization so that the largest values align along the main diagonal. Finally, in the third step, optimal weights are calculated as the inverse of each security's variance, which is located on the main diagonal after the quasi-diagonalization step.

Graph theory background

For a comprehensive introduction to graph theory, refer to Bollobás ([1998](#), [2001](#)). Consider a directed weighted graph $G = (V, E, W)$ formed by a finite set of vertices V[5], a set of directed edges $E \subset V \times V$, where each $(e_x, e_y) \in E$ represents a link from $v_x \in V$ to $v_y \in V$, and a set of weights $W : E \to \mathbb{R}++$ defined on each edge. Two nodes $(v_x, v_y) \in V$ are said to be adjacent if there exists an edge $(e_x, e_y) \in E$. The adjacency matrix of the graph A_G is defined as the square matrix whose entries are $ai, j = w_{i,j}$ if $v_i, v_j \in E$ and $w_{i,j}$ are the weight of the edge, $a_{i,j} = 0$ otherwise.

In a financial setup, a graph can be used to represent a financial market, wherein a security is represented by a vertex, and the relationship between each pair of securities is represented by edges. The simplest way of measuring the relationships among securities is to use linear correlations. As in clustering, the adjacency matrix must be defined on a metric; hence, we apply the same $d_{i,j} = 1 - \rho_{i,j}^2$ transformation. Moreover, the diagonal of the adjacency matrix is set to zero to avoid self-loops[6]. As such, a non-zero entry in the adjacency matrix indicates the existence of a financial relationship between pairs of securities with strength $d_{i,j}$.

A graph built on a correlation matrix is a complete digraph, which is a graph where all pairs of vertices are connected by a pair of unique edges. This is the source of the link of HRP to graph theory. However, correlation matrices lack the notion of hierarchy. Simon ([1991](#)) argues that

complex systems, such as financial markets, can be arranged in a natural hierarchy comprising nested substructures. The goal of codependence analysis is choosing which cross-security relations really matter. From a tree representation standpoint, this means choosing which links in the tree are significant and removing the others. de Prado ([2016](#)) argues that the lack of hierarchical structure makes portfolio weights vary in unintended ways in an asset allocation problem. For this reason, complete digraphs do not add additional information compared to correlation matrices, while other subgraphs – such as spanning trees[7] – can better serve financial needs by incorporating a hierarchical representation and choosing the links between securities that really matter.

We employ the minimum spanning tree (MST) (see Appendix B for a graphical representation), a subset of a complete digraph that includes all vertices but selects the minimum possible number of edges by solving $\min_S \sum_{e \in s} W(e)$, where $S \leq E$ is the number of links in the MST and e represents each link in each realization, $s \in S$. To find minimum spanning tees, we employ the algorithm by Kruskal ([1956](#)).[8] It is worth noting that there are several parallels between Kruskal's algorithm for minimum spanning trees and AGNES algorithm for hierarchical clustering. First, both algorithms start with a set of fully disconnected nodes and iteratively build clusters. Second, both algorithms use a greedy approach to form clusters. In Kruskal's algorithm, the edges of the graph are sorted by weight based on the distance between securities, and at each step, the algorithm adds the next edge (with the lowest weight) that connects two previously unconnected clusters. Similarly, in AGNES, at each step, the algorithm merges the two clusters that are closest to each other, based on some distance metric. Finally, both algorithms produce a hierarchy of clusters, which can be visualized in the form of an MST or a dendrogram, respectively.

A minimum spanning tree allows quantifying how securities influence each other by means of a centrality measure. A centrality measure $C : V \to \mathbb{R}_+$ is a function that assigns a non-negative value to each node such that the higher the value, the more the node is connected to others. One such centrality measure is the eigenvector centrality, according to which a security displays a high centrality either by direct links to other securities or by being connected to other securities that are themselves highly connected. As such, the higher the eigenvector centrality, the more central a security

---

[3] The calculation of a distance measure requires that the quantities are Euclidean metrics, which correlations are not.

[4] According to this distance measure, two pairs of securities with the same linear correlation but with the opposite sign are considered equally distant. While this is a general limitation, we regard it as insignificant as long as the focus is on equity returns, which generally exhibit positive linear correlations.

[5] in this manuscript, we interchangeably use the terms "nodes" and "vertices" to refer to the elements of the set V.

[6] A self-loop is a link that connects a vertex to itself. We eliminate self-loops since self-relations are insignificant in portfolio construction.

[7] A spanning graph is a subgraph that contains all vertices of the complete graph. A tree is an acyclical and connected graph, where all nodes are connected by a single edge. It can be proved that every tree is a graph and every non-trivial graph contains at least one tree.

[8] Starting from a tree in each vertex, the Kruskal's algorithm removes any link with a minimum weight between the vertices, combining the trees for which the link has been removed.

is in the tree. Eigenvector centrality is based on the idea that a node's importance is determined by the importance of the nodes that it is connected to. In formula, the eigenvector centrality of a node $\zeta(v)$ is given:

$$\zeta(v) = \frac{1}{\lambda_{\max}} \sum_{v' \in N(v)} \zeta(v') \tag{2}$$

where $N(v)$ is the set of neighbors of node v and $\lambda_{\max}$ is the largest eigenvalue.

Network Risk Parity

We derive the network risk parity methodology from a parallel with hierarchical risk parity, assuming that its covariance matrix C is an n-dimensional diagonalized matrix with full-rank. Applying the spectral decomposition theorem to C yields:

$$C = U\Lambda U^T = \sum_{j=1}^{n} \lambda_j \mathbf{u}_j \mathbf{u}_j^T \tag{3}$$

where $\mathbf{u}_j$ is the j-th eigenvector associated with the eigenvalue λ_j of matrix C. This relation leads to the eigenvector equation $C\mathbf{u}_i = \lambda_i \mathbf{u}_i$, for all $i = 1, ..., n$. The eigenvalues λ_i and eigenvectors $\mathbf{u}_i$ are obtained by solving the characteristic equation $\det(C - \lambda I) = 0$. Solving for the eigenvalues yields (see Appendix A for the proof):

$$\lambda_i = \sigma_i, \quad \forall i = 1, ..., n \tag{4}$$

where σ_i are the elements of the diagonal matrix C. In other words, the optimal inverse-variance allocation of Hierarchical Risk Parity corresponds to the eigenvalues of the quasi-diagonal covariance matrix C, assuming that the diagonalization step effectively results in a diagonal matrix.

Rearranging the eigenvector centrality definition of the MST in matrix form, we have:

$$A\zeta = \lambda_{\max}\zeta \tag{5}$$

where ζ is the eigenvector of the adjacency matrix A associated with its largest eigenvalue $\lambda_{\max}$. In NRP we take a somewhat similar approach to HRP by using the eigenvector centrality and calculate the portfolio weights as:

$$w_i^* = \frac{1}{\zeta_i} \quad \forall i = 1, ..., n. \tag{6}$$

To obtain fully invested portfolios such that $\sum w = 1$, we apply the softmax normalization $\sigma(w) = \frac{e^w}{\sum(e^w)}$. As Laloux (1999) showed that the largest eigenvalue of a correlation matrix can be seen as a representative of systematic risk, and as the eigenvector centrality is associated with the largest eigenvalue of the adjacency matrix, the eigenvector

centrality can be understood as a gage of a security's contribution to systematic market risk. In the NRP approach, these securities are assigned lower weights in the portfolio, thereby aiming to minimize the portfolio's exposure to systematic risk.

Despite both HRP and NRP use the same distance metrics calculated as $d_{(i,j)} = 1 - \rho_{i,j}^2$, where $\rho_{i,j}$ is the Pearson correlation coefficient, they result in different portfolios due to the different ways the notion of hierarchy is imposed on the correlation matrix. Hierarchical risk parity uses hierarchical clustering with Euclidean distance and a single linkage criterion. Network risk parity, instead, uses Kruskal's algorithm to build a minimum spanning tree and select the meaningful interconnections between securities. Three benefits are associated with the latter approach.

First, in NRP relationships among securities are one-to-many rather than one-to-one as in HRP. In fact, the eigenvector centrality takes into account the importance of neighboring nodes too. On the other hand, hierarchical clustering uses a one-to-one approach to cluster securities.

Second, HRP depends on the specification of a distance function to transform linear correlation into metrics, and of a distance function and a linkage function for clustering purposes. NRP, on the other hand, depends only on the first. In particular, HRP has been criticized for using the single linkage function, which separates objects depending on the distance between the two closest points within clusters. This causes a chaining effect that widens the tree and results in an unequal distribution of the portfolio weights (Papenbrock 2011).

Third, the usage of graph theory allows us to analytically prove the level of concentration of the optimal portfolios, as will be shown.

Portfolio weights lower bound

Leveraging the concept of the degree $d(v)$ of a vertex v in a graph G, defined as the number of vertices in G that are adjacent to v, it can be established – using the Gershgorin circle theorem[9] - that the maximum eigenvalue $\lambda_{\max}$ of the

---

[9] The Gershgorin circle theorem provides a way to bound the eigenvalues of a square matrix with the sum of the absolute values of the matrix's entries along its rows or columns. Specifically, the theorem states that each eigenvalue of a matrix A lies within at least one of the n Gershgorin disks, which are defined as follows.

For $i = 1, 2, \ldots, n$, let R_i denote the sum of the absolute values of the off-diagonal entries in the ith row of A, and let a_{ii} denote the diagonal entry of A in the ith row. Then, the ith Gershgorin disk is the closed disk in the complex plane with center a_{ii} and radius R_i. That is,

$$D_i = z \in \mathbb{C} : |z - a_{ii}| \leq R_i.$$

The Gershgorin circle theorem then states that every eigenvalue of A lies in at least one of the Gershgorin disks, that is,

adjacency matrix A is bounded above by the maximum degree $d(v)$ of the graph. As $w = \zeta^{-1}$ and ζ corresponds to the largest eigenvalue λ, then $w \geq \max(d)^{-1}$, meaning that all portfolio assets have a weight that is greater than the inverse of the maximum degree of the MST. In particular, as $d(v) > 0$ for all non-empty graphs, the optimal weights of NRP are lower bound by the softmax-normalized degree of the MST. As HRP is based on the inverse-variance allocation of a quasi-diagonalized correlation matrix which we show to be proportional to λ^{-1},[10] given that the Gershgorin theorem holds for any square matrix, and given that the correlation matrix can be represented as a complete digraph, the portfolio weights of HRP have a lower bound greater than zero as well.

Conversely, the spectral radius of the graph $\rho(G)$[11] is bounded by $\rho(G) \geq \sqrt{\lambda_{max} - 1}$, where λ_{max} is the largest eigenvalue of adjacency matrix. This suggests that the spectral radius can be seen as the degree of concentration that can occur, as the weights are inversely proportional to the eigenvalues, which grow larger in highly volatile periods (hence, resulting in more spread out portfolio weights). This lower bound mechanism contributes to creating more diversified, less concentrated portfolios compared to those derived from mean-variance optimization methods.

Of note is that the largest eigenvalue of a covariance matrix, a measure that varies over time, serves as a proxy for systematic risk (Laloux 1999), and it tends to increase during high-volatility regimes. Viewing this from the perspective of the minimum spanning tree (MST), the spectral radius tends to have a higher upper bound during highly volatile periods as the maximum eigenvalue grows higher. This means that as volatility increases, the spectral radius also increases, thereby leading to a decrease in the optimal portfolio weights and resulting in more diversified portfolios. This ability to adapt to changing market conditions is a key strength of NRP and HRP. As MVO propagates the errors of the ill-estimated covariance matrix through its inversion, its performance becomes relatively worse during highly volatile periods, due to additional concentration. Whereas MVO fails to perform optimally during high-volatility periods (when most needed), NRP and HRP employ a self-defensive mechanism resulting in better-diversified

Table 1 Average monthly Sharpe ratios of bootstrapped portfolios for different numbers of constituents (best performing in boldface)

n constituents	NRP	HRP	RP	EW	MV
$n = 20$	0.853	**0.892**	0.653	0.817	0.430
$n = 50$	0.863	**0.882**	0.712	0.809	0.576
$n = 100$	**0.886**	0.845	0.725	0.799	0.434
$n = 200$	**0.891**	0.866	0.723	0.794	0.396

portfolios, thanks to the dynamic adjustment of the spectral radius of the graph.

Empirical results

In this section, we illustrate some empirical results regarding the application of network risk parity. In this section, we consider long-only,[12] fully invested portfolios rebalanced at monthly frequency.[13] We use the hourly prices of all the stocks composing the S &P 500 index from January 2010 to March 2023, aggregated to a monthly frequency in the process of calculating conditional covariances. To ensure no look-ahead bias in our study, we implemented a rolling window approach. For each monthly rebalancing, we utilized the past 2 years' data. To obtain a robust result, we use a bootstrapping approach that randomly selects $n = (20, 50, 100, 200)$ stocks – without replacement – out of the entire sample and iterates through 10000 simulations. We compare the performances of the competing methods in terms of the Sharpe ratio.[14]

Table 1 reports the average Sharpe ratios of the bootstrapped portfolios for the different numbers of constituents.

Footnote 9 (continued)

$$\lambda(A) \subseteq \bigcup_{i=1}^{n} D_i,$$

where $\lambda(A)$ denotes the set of eigenvalues of A.

[10] Or exactly equal in case of a fully diagonalized matrix.

[11] The spectral radius of a matrix is the maximum absolute value of its eigenvalues. More formally, if a matrix A is a square matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, the spectral radius $\rho(A)$ is defined as: $\rho(A) = \max_{1 \leq i \leq n} |\lambda_i|$.

[12] To obtain a long-short application, the adjacency matrix would need to be adjusted to account for short positions, the normalization of weights would have to be revisited to cater for the possibility of negative weights, and the risk contribution calculation would require an overhaul to incorporate the potential negative contribution of short positions.

[13] We aggregate hourly data into monthly realized covariance matrices to ensure a more robust estimation of the covariance matrix. However, one must notice that NRP is agnostic to the choice of the covariance estimator. The different frequency granularity between hourly and monthly data can result in deviations from normality. However, the use of non-normal returns influences the portfolio optimization process only for methodologies that depend on the assumption of normally distributed returns, such as Markowitz's minimum variance. Hierarchical risk parity and network risk parity methodologies, which are the focus of this study, do not rely on this assumption. Additionally, the use of bootstrapping in our simulations helps to address the potential issues stemming from non-normality, as it allows us to capture the empirical distribution of the data.

[14] Sharpe ratios are calculated following the monthly rebalancing scheme, hence using monthly portfolio returns and monthly standard deviations recorded at the time of rebalancing.

Network risk parity and hierarchical risk parity outperform the other methodologies in all sample sizes. In particular, hierarchical risk parity achieves the highest Sharpe ratio for smaller portfolio sizes, such as $n = (20, 50)$, while network risk parity is the best performing as the number of portfolio constituents increases.[15] This behavior is due to the fact that the higher the number of portfolio constituents, the more the minimum spanning tree resembles a financial market in terms of the ability to place an asset at the center of the tree rather than in the periphery. This is explained by the fact that the MST structure used in NRP is an improved representation of systematic risk and financial markets. MST preserves the most significant relationships between assets, effectively filtering out less meaningful correlations. This approach enables us to uncover the market's inherent structure by identifying clusters of assets that exhibit similar behavior, offering a more comprehensive understanding of systematic risk. Furthermore, financial markets are complex networks characterized by multidimensional inter-dependencies between assets. These connections extend beyond pairwise relationships, something the MST structure, with its graph representation, can capture more accurately. Therefore, the larger the number of portfolio constituents, the more effectively the real asset connections are represented, as the MST can better link assets based on similarities.

Network risk parity outperforms the other competing methods in all portfolio sizes, while MVO is the worst performer across all sample sizes. In fact, MVO often leans toward portfolios that capitalize on idiosyncratic risk, due to the inversion of an ill-estimated covariance matrix that propagates the errors. This often results in concentrated portfolios with large weights assigned to a few assets exhibiting lower volatility. On the other hand, NRP and HRP strive for more balanced portfolios as they are specifically designed to distribute risk equally across all assets in the portfolio. This more equal risk distribution inherently aligns the portfolio to the broader, systematic market movements rather than the idiosyncratic movements of individual assets. By reducing exposure to idiosyncratic risk, these strategies are less susceptible to the unexpected performance of a small subset of assets and therefore tend to be more stable. For NRP and HRP, this is further supported by the spectral radius acting as a self-defensive mechanism, as during periods of high volatility, the spectral radius also increases, thereby resulting in more evenly spread portfolio weights and better diversification. MVO, instead, during periods of high volatility, results in even more concentrated portfolios, as the covariance matrix is even more ill-estimated.

Figure 1 illustrates the histograms of the bootstrapped Sharpe ratios for each asset allocation strategy against the performance of the equally weighted portfolio[16] for the case $n = 100$.

To provide some context regarding the most severe downturns across the bootstrapped samples, for the worst bootstrap simulation of each strategy, we computed the maximum drawdown using $n = 20$. It was found that all portfolios reached their maximum drawdowns during the height of the COVID-19 crisis in March 2020. In particular, the NRP and HRP witnessed maximum drawdowns of 42% and 39%. RP and EW portfolios suffered a maximum drawdown of 48% and 46%, respectively, while MVO had the largest drawdown of 61%. For comparison, the S & P 500 experienced a maximum drawdown of 34% during the same COVID-19 period.

Finally, Figure 2 plots the boxplots of weights on the bootstrapping runs with $n = 20$. We highlighted in red the parts corresponding to a portfolio weight of 0, as this is the first red flag to identify ill-concentrated portfolios. As it is visible, only the mean-variance allocation incurs in portfolio weights equal to 0 as well as several times reaches portfolio weights well above 60%, hinting at poor diversification and major concentration. On the other hand, the weights of NRP and HRP appear to be better diversified and less concentrated, as well as exhibit a similar pattern, in light of the common lower bound mechanism. Risk parity, on the other hand, is better diversified than mean-variance, but it reaches peaks up to 60%.

Appendix 7 provides a commentary on the average bootstrapped performances of the competing methods against the market environments that characterized the periods analyzed.

Conclusions

In this manuscript, we have presented network risk parity (NRP), a novel portfolio construction methodology based on graph theory. NRP serves as the counterpart to hierarchical risk parity (HRP), providing a complementary approach to portfolio allocation. By utilizing the concept of eigenvector

---

[15] We run a hypothesis testing on the null hypothesis that the Sharpe ratios of NRP are equal to those of HRP. The null hypothesis is rejected for all numbers of constituents at 1% significance level, except for $n = 40$, which is only rejected at 5%.

[16] Our intent is not to use the equally weighted portfolio as a benchmark against which an active portfolio manager is measured. Instead, we regard it as a fundamental performance threshold that any asset allocation strategy should surpass, given that it represents a simple allocation approach in contrast to the intricacies of competing methods.
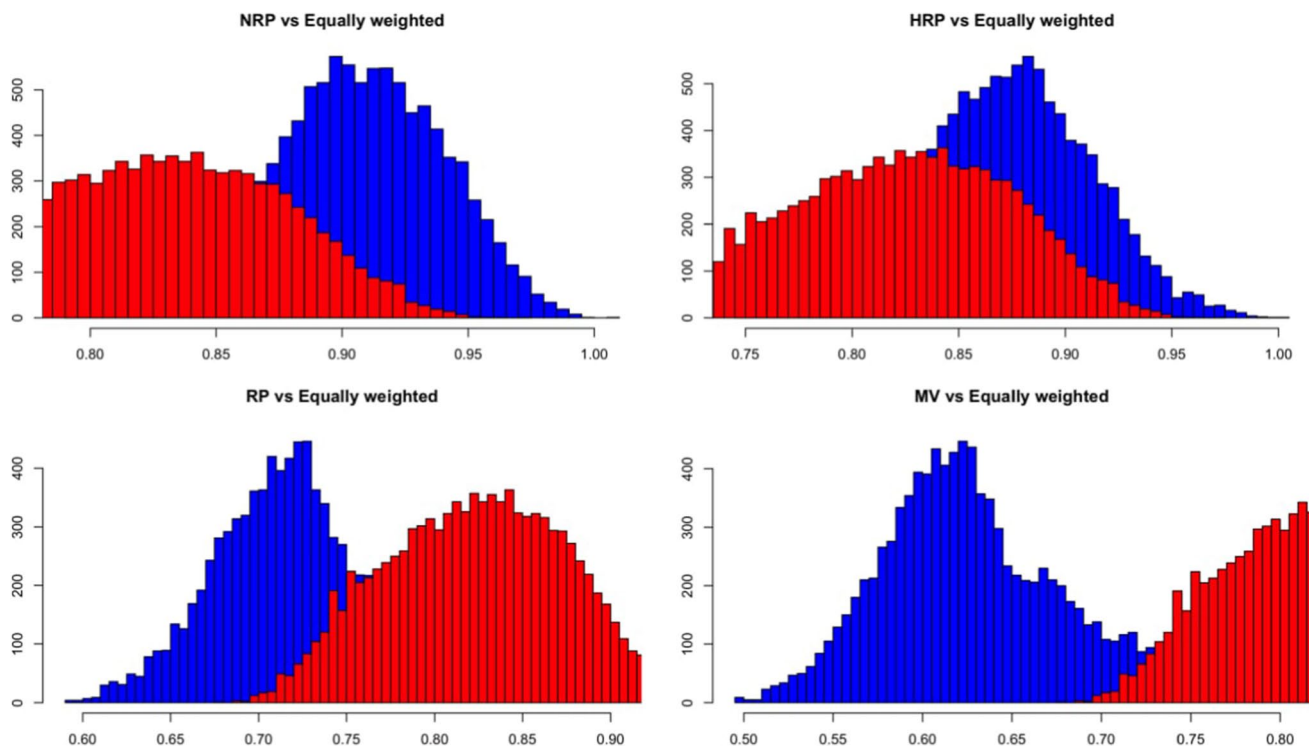
Fig. 1 Histograms of bootstrapped Sharpe ratios against equally weighted. Note: in red, we report the distribution of the Sharpe ratios achieved by the equally weighted portfolio, while in blue, the distributions of the Sharpe ratios achieved by each asset allocation strategy across the bootstrapped samples. As it is visible, network risk parity and hierarchical risk parity are more skewed toward higher Sharpe ratio values, while the opposite holds for risk parity and Markowitz's minimum variance
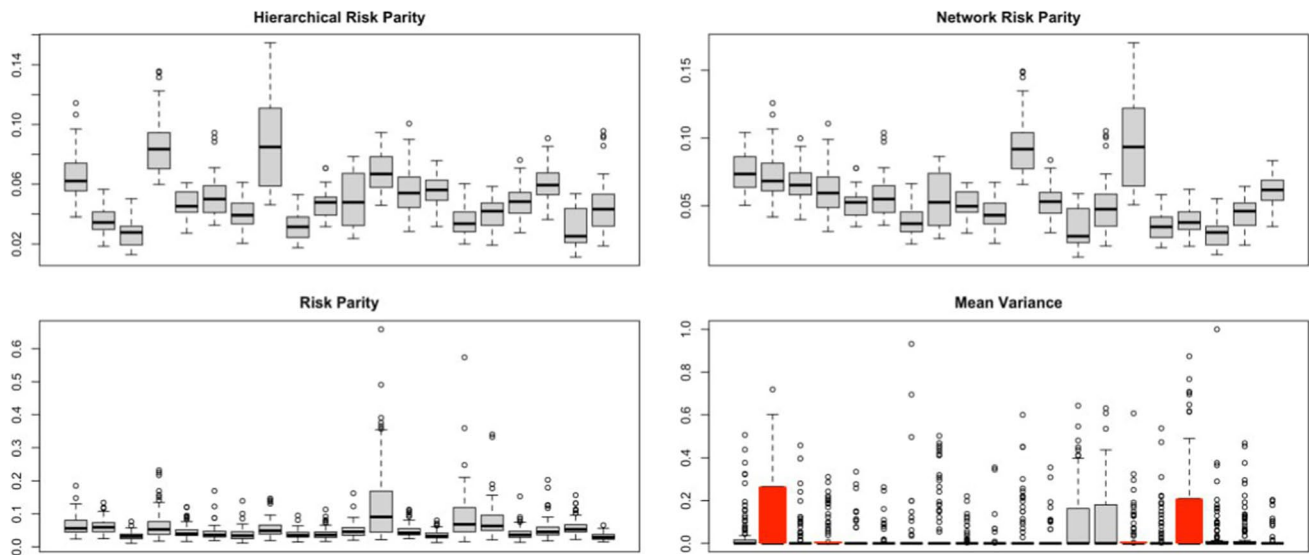
Fig. 2 Boxplots of bootstrapped weights to show the major diversification benefits associated with NRP and HRP. Note: The boxplots of the portfolio weights across the bootstrap run with $n = 20$ (chosen to preserve the readability of the plot). In red, we highlight the points where the boxplot touches a value of zero, a result that is not achievable in light of the lower bound rule applicable to graph-based allocations. The trivial equally weighted case is not included

centrality and the minimum spanning tree (MST), NRP offers several advantages over traditional methods.

One key advantage of NRP is its ability to capture one-to-many relationships between securities. Unlike one-to-one relationships captured by hierarchical clustering in HRP, NRP considers the importance of neighboring nodes as well. This broader perspective allows for a more comprehensive understanding of the interconnectedness and dependencies within a financial market.

Another advantage is the simplicity of the graph-based approach compared to the distance and linkage functions used in hierarchical clustering. NRP relies solely on the adjacency matrix and the MST, reducing the complexity of the methodology while still achieving effective risk allocation.

Furthermore, the graph theory framework of NRP allows for an analytically provable lower bound on the optimal weights of the portfolio. This lower bound ensures that the portfolio remains well-diversified, mitigating the risk of concentration on a few securities and avoiding the pitfalls associated with Markowitz's minimum-variance approach.

Empirical results based on bootstrapped samples of S & P 500 stocks demonstrate the superiority of NRP over risk parity, Markowitz's minimum-variance, and equally weighted portfolios in terms of the Sharpe ratio. Notably, the performance of NRP and HRP varies with the number of stocks in the portfolio, with HRP showing strength in smaller portfolios and NRP excelling as the number of constituents increases.

Future research should test different subgraph representations, with major focus posed on how the covariance matrix is estimated or forecast as well as on the inclusion of different asset classes other than equity. Furthermore, future research could focus on an exploration of covariance mis-estimations and the potential improvements brought by distance and adjacency matrices compared to sample covariances. Finally, as the current formulation of network risk parity assumes a long-only portfolio structure, the extension to long-short portfolios could be an interesting avenue for future research.

A proof of equation (4)

In this appendix, we briefly illustrate the algebrical steps to prove that the optimal weights generated by hierarchical risk parity are proportional to the inverse of the eigenvalues of the modified covariance matrix.

Let's assume that the modified covariance matrix C is squared, full-rank, and fully diagonal. For instance, in the two-dimensional case:

$$C = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$$

Applying the spectral decomposition theorem yields:

$$C = U\Lambda U^T = \sum_{0<j<k} \lambda_j \mathbf{u_j} \mathbf{u_j^T}$$

which results in $= C\mathbf{u_i} = \lambda_i \mathbf{u_i}, \forall i = j$. The eigenvalues and eigenvectors are obtained by solving $det(C - \lambda I)\underline{\mathbf{u}} = \underline{\mathbf{0}}$. Without loss of generality, let's assume a two-dimensional covariance matrix C, from where the eigenvalues are given by:

$$det(C - \lambda I) = det\left(\begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) = 0$$

which results in:

$$det\begin{bmatrix} \sigma_1 - \lambda & 0 \\ 0 & \sigma_2 - \lambda \end{bmatrix} = 0$$

and since in a $2x2$ matrix the solution of the determinant is given by the product of the diagonal terms minus the product of the anti-diagonal ones, one gets:

$$(\sigma_1 - \lambda)(\sigma_2 - \lambda) = \lambda^2 - \lambda(\sigma_1 + \sigma_2) + \sigma_1\sigma_2 = 0.$$

finally, solving the second degree equation:

$$\lambda_{1,2} = \frac{\sigma_1 + \sigma_2 \pm \sqrt{(\sigma_1 + \sigma_2)^2 - 4\sigma_1\sigma_2}}{2}$$

$$= \frac{\sigma_1 + \sigma_2 \pm \sqrt{\sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2 - 4\sigma_1\sigma_2}}{2}$$

$$= \frac{\sigma_1 + \sigma_2 \pm \sqrt{(\sigma_1 - \sigma_2)^2}}{2}.$$

Hence, the two solutions are:

$$\lambda_1 = \frac{\sigma_1 + \sigma_2 + \sigma_1 - \sigma_2}{2} = \sigma_1 \lor \lambda_2 = \frac{\sigma_1 + \sigma_2 - \sigma_1 + \sigma_2}{2} = \sigma_2.$$

Generalizing to the $i = 1, 2, ..., n$-dimensional matrix C yields:

$$\lambda_i = \sigma_i, \forall i \in (1, n).$$

$\square$
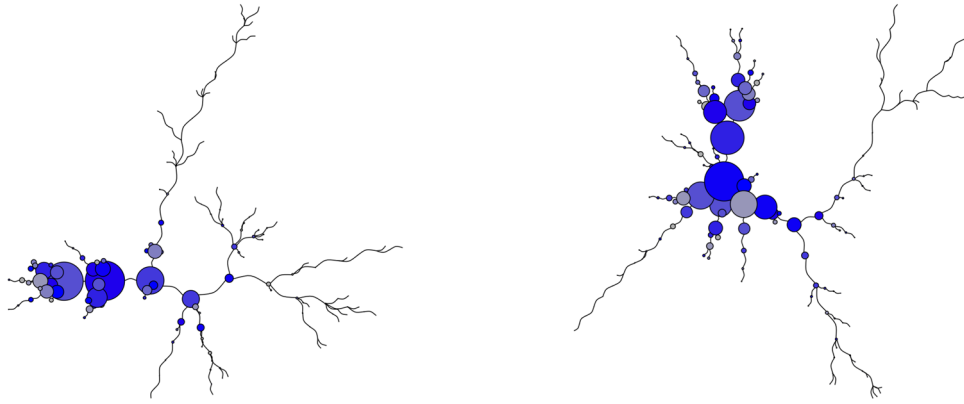
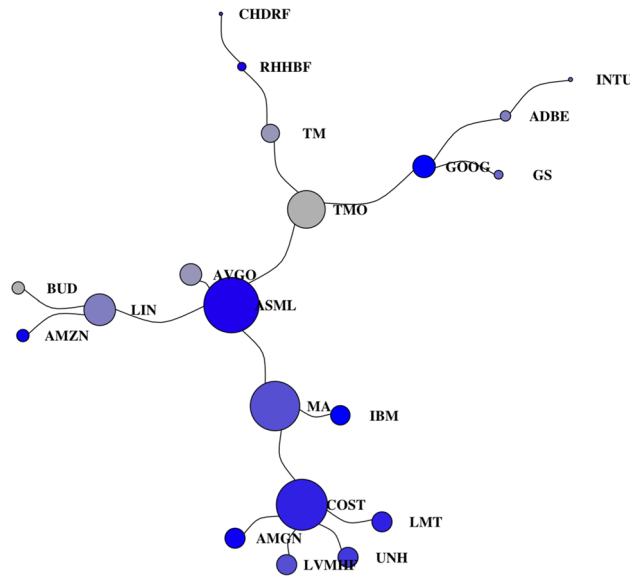# B Plots of minimum spanning tree

See Figs. .

Fig. 3 Minimum spanning trees. Note: On the left side, a minimum spanning tree obtained using the 500 components of S & P 500 in a low-volatility regime (November 22, 2019). On the right side (using the same scale), the same tree during the fallout of the Silicon Valley Bank in March 2023 (March 13, 2023), which triggered high volatility amid concern over systematic risk. Each node in these MSTs represents a stock, with the size of the node indicating its eigenvector centrality. This measure represents the influence of each node within the network, calculated considering both its direct connections and its indirect influence through its neighbors. It assigns relative scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of a node than equal connections to low-scoring nodes. Thus, a larger node size in the MST plot represents higher eigenvector centrality, i.e., a greater influence within the network. As such, in the center of the tree, highly connected stocks are typically found. The observable differences in the distribution of node sizes between the low-volatility and the high-volatility trees are due to the differences in eigenvector centralities. During periods of low volatility, markets tend to be more stable, and correlations between assets are generally lower and more evenly distributed. This results in a more uniform distribution of eigenvector centrality. Consequently, the network importance is more evenly spread out. In contrast, during periods of high volatility, correlations between assets often increase significantly


Fig. 4 Small minimum spanning tree. Note: An illustrative minimum spanning tree representation for a set of 20 securities as of March 10, 2023. The Kruskal algorithm for the construction of the MST first considers the full graph, where each node represents a single security and every possible pair of securities is connected by an edge. Each edge is assigned a weight based on the distance (correlation-based) between the corresponding pair of securities. Typically, a lower correlation results in a higher weight. The MST is then built by selecting edges such that the sum of the weights is minimized, and the graph remains connected without any cycles. In the resulting MST, each node (or stock) is connected to the closest node, thereby providing a visualization of the strongest relationships within the portfolio. This visualization helps to highlight the principle that securities that are more closely related, i.e., have higher correlations, are placed closer together in the MST, often forming clusters that may indicate groups of securities that might be affected similarly by market movements

C Commentary of the average performance of the competing methods

See Fig. 5.
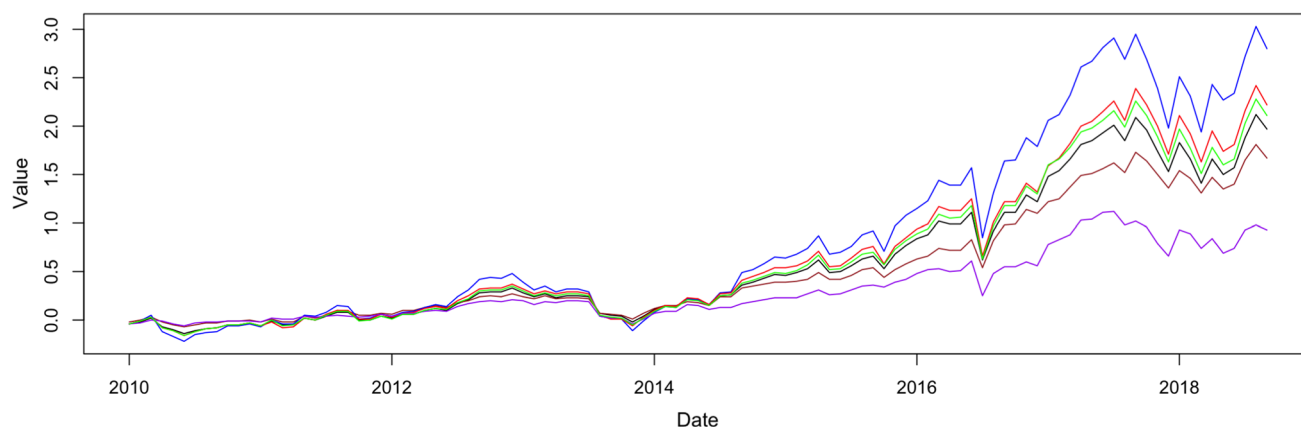
Fig. 5 Average bootstrapped equity plot for each strategy. Note: This figure shows the equity cumulative profit function of each competing method as well as the performance a buy and hold on the S & P 500. To ease readability of the figure, we have only plotted the average bootstrapped scenario for the case of $n = 100$ constituents. Specifically – NRP: blue, HRP: red, EW: green, RP: brown, MVO: purple, and S & P 500: black. The period from 2010 to 2023 was marked by several global financial events that significantly influenced portfolio performances, such as the Euro-Sovereign Crisis (2010-2012), the Chinese crisis in 2015, Brexit, the US-China trade war, COVID-19, the Russo-Ukraine 2022 war and inflationary environment, failures of US banks and UBS takeover of Credit Suisse. All portfolio strategies were impacted due to increased volatility in these periods, in line with the performance of the S & P 500. However, as it is visible from the plot of the equity lines below, strategies heavily reliant on covariance matrix estimations, such as risk parity and Markowitz, have faced heightened adversity due to increased asset correlations and volatility. On the contrary, hierarchical risk parity and network risk parity strategies, have relatively mitigated some risks during this uncertainty. Simultaneously, equally weighted portfolios have seen substantial declines due to the broad-based nature of the crisis

D Comparison between NRP and HRP

See Table 2.

Table 2 .

	NRP	HRP
Underlying theory	Graph theory (MST)	Hierarchical clustering
Starting quantity	Covariance of asset returns	Covariance of asset returns
Weight calculation	Inverse of eigenvector centrality	Inverse of variance
Specification dependencies	Correlations as metric	Correlation as metric, distance and linkage function
Minimum weight	Yes	Yes, as a matrix of distances is a square matrix that can be seen as a complete digraph
Relationships captured	One-to-many due to eigenvector centrality	One-to-one

Declarations

Conflict of interest The authors of this manuscript report no conflict of interest with an entity - with a financial or non financial interest - in the subject matter or materials discussed in this manuscript.

References

Barndorff-Nielsen, Ole E., and Neil Shephard. 2004. Econometric analysis of realized covariation: high frequency based covariance, regression, and correlation in financial economics. *Econometrica* 72: 885–925. https://doi.org/10.2139/ssrn.305583.

Black, Fischer, and Robert Litterman. 1992. Global portfolio optimization. *Financial Analysts Journal* 48 (5): 28–43.

Bollobás, Béla. 1998. *Random graphs*, In Modern graph theory. 21500252. Springer.

Bollobás, Béla. 2001. *Random graphs*, vol. 73. Cambridge University Press.

de Prado, Marcos, and Lopez. 2016. Building portfolios that outperform out-of-sample. *The Journal of Portfolio Management*. https://doi.org/10.3905/jpm.2016.42.4.059.

Kruskal, Joseph B. 1956. On the shortest spanning subtree of a graph and the traveling salesman problem. In Proceedings of the American Mathematical Society vol. 7, 48–50.

Laloux, Laurent, et al. 1999. Noise dressing of financial correlation matrices. *Physical Review Letters* 83 (7): 1467.

Ledoit, O., and M. Wolf. 2004. Honey, i shrunk the sample covariance matrix. *The Journal of Portfolio Management* 30 (4): 110–119.

Mantegna, Rosario N. 1999. Hierarchical structure in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems* 11 (1): 193–197.

Markowitz, Harry. 1952. Portfolio Selection. *The Journal of Finance* 7 (1): 77–91. https://doi.org/10.2307/2975974'www.jstor.org/stable/2975974'.

Michaud, Richard O. 1989. The markowitz optimization enigma: Is optimized optimal? *Financial Analysts Journal* 45 (1): 31–42. https://doi.org/10.2469/faj.v45.n1.31.

Papenbrock, Jochen. 2011. Asset clusters and asset networks in financial risk management and portfolio optimization. PhD thesis. Karlsruher Institut für Technologie (KIT).

Peralta, Gustavo, and Abalfazl Zareei. 2016. A network approach to portfolio selection. *Journal of Empirical Finance* 38: 157–180.

Raffinot, Thomas. 2017. Hierarchical clustering-based asset allocation. *The Journal of Portfolio Management* 44 (2): 89–99.

Simon, Herbert A. 1991. The architecture of complexity. In Facets of Systems Science. Springer, pp. 457–476.

Výrost, Tomas, Stefan Lyócsa, and Eduard Baumöhl. 2019. Network-based asset allocation strategies. *The North American Journal of Economics and Finance* 47: 516–536.

Vito Ciciretti earned a MSc in Quantitative Finance from Bocconi University in 2018. Currently, he works as a Quant in a major banking institution. He is member of the Italian Econometrics Association. His research interests include portfolio optimization, graph theory, unsupervised learning and regime detection.

Alberto Pallotta graduated with a degree in Engineering from the University of Perugia, Italy, in 2009. He furthered his education by obtaining an MSc in Artificial Intelligence with distinction from the University of Hertfordshire in 2021. In 2017, he earned a Certificate in Quantitative Finance (CQF) with distinction from Fitch Learning. Currently, Alberto serves as a lecturer at Middlesex University in the Department of economics and finance. He is pursuing a PhD at Middlesex University, focusing on quantitative finance. Throughout his career, Alberto has provided his expertise as an advisor in the field of quantitative finance to major financial institutions and private companies. His research interests include portfolio optimization, volatility forecasting, stock picking, and algorithmic trading.