Diversifying SnP500

Dhruv Gupta   Harshit Joshi

July 18, 2025

1 Introduction

The S&P 500 index exhibits significant concentration risk, with the top 10 stocks accounting for nearly 40% of the total market capitalization as of mid-2025 [1]. This concentration, primarily due to large-cap technology firms, reduces the overall diversification of the index and increases exposure to idiosyncratic risk. In this report, we apply a systematic and non-parametric method to diversify the S&P 500. The approach does not involve hyperparameter tuning and allocates risk based on the hierarchical structure among asset returns.

2 Hierarchical Risk Parity

The Hierarchical Risk Parity (HRP) [2] algorithm constructs a diversified portfolio in three stages, without relying on quadratic optimization. Consider a set of N assets with covariance matrix $\Sigma \in \mathbb{R}^{N\times N}$ and correlation matrix $\rho \in \mathbb{R}^{N\times N}$. We focus only on Stage 1 (Tree Clustering), since it alone provides the necessary hierarchical structure in our algorithm.

Stage 1: Tree Clustering

The hierarchical structure among assets is identified using the following steps:

1. **Correlation Distance:** Compute a distance matrix $D \in \mathbb{R}^{N\times N}$ using the correlation matrix ρ, where the distance between assets i and j is defined as:

$$D_{ij} = \sqrt{\frac{1}{2}(1-\rho_{ij})} \tag{1}$$

   This defines a proper metric that quantifies dissimilarity between asset returns.

2. **Linkage Metric:** Construct another matrix $\tilde{D} \in \mathbb{R}^{N\times N}$ by computing the Euclidean distance between the i-th and j-th column vectors of D:

$$\tilde{D}_{ij} = \sqrt{\sum_{k=1}^{N}(D_{ki}-D_{kj})^2} \tag{2}$$

   A lower $\tilde{D}_{ij}$ indicates that assets i and j have similar correlation profiles with the rest of the asset universe.

3. **Hierarchical Clustering:** Apply a hierarchical clustering algorithm (e.g., single linkage) to the matrix $\tilde{D}$. The output is a linkage matrix that defines a dendrogram over the assets.

3 Methodology

We construct a diversified portfolio from the S&P 500 asset universe, as the composition of the S&P 500 index is dynamic; to ensure a consistent universe for our analysis, we utilize daily price data for the 501 constituents as of July 03, 2025 [1] over a 10-year period. This dataset yields a percentage returns (not Log) matrix $\mathbf{X} \in \mathbb{R}^{T \times N}$, with $T = 2515$ trading days. From $\mathbf{X}$, we compute a pairwise asset distance matrix, which serves as the input for our analysis.

Traditional diversification methods face notable limitations. For instance, k-means clustering requires the *a priori* specification of the number of clusters, k, introducing a parameter-dependent bias. Hierarchical Risk Parity (HRP) avoids this by using a dendrogram, but it remains a heuristic algorithm. As such, HRP is not a formal optimization framework and cannot readily incorporate portfolio-level constraints, such as bounds on asset weights or sector exposures.

To overcome these limitations, we create a optimization framework that leverages the complete topological structure of the asset hierarchy. We first apply Ward's minimum variance method to the asset distance matrix $\mathbf{D} \in \mathbb{R}^{N \times N}$. This procedure generates a dendrogram that encodes a nested sequence of N distinct partitions of the assets. We denote the partition at level $\ell \in \{1, \ldots, N\}$ as $\mathcal{P}_\ell = \{C_{\ell,1}, \ldots, C_{\ell,k_\ell}\}$, where $k_\ell = N+1-\ell$ is the number of disjoint clusters at that level.

The core of our proposal is to formulate the portfolio construction as a min-max optimization problem. We want a single weight vector, $\mathbf{w}$, that is robustly diversified across the entire hierarchy. This is achieved by minimizing the maximum total intra-cluster variance, where the maximum is taken over all partitions encoded at each level of the dendrogram. For any cluster $C \subseteq \{1, \ldots, N\}$, let $\mathbf{w}_C$ be the sub-vector of weights and $\mathbf{\Sigma}_C$ be the covariance sub-matrix for the assets in C. The total intra-cluster variance for a given partition $\mathcal{P}_\ell$ is thus $\sum_{C \in \mathcal{P}_\ell} \mathbf{w}_C^\top \mathbf{\Sigma}_C \mathbf{w}_C$.

The final portfolio weights are obtained by solving the following Quadratically Constrained Quadratic Program (QCQP):

$$
\begin{aligned}
\underset{\mathbf{w} \in \mathbb{R}^N, z \in \mathbb{R}}{\text{minimize}} \quad & z \\
\text{subject to} \quad & \mathbf{1}^\top \mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0} \\
& z \geq \sum_{C \in \mathcal{P}_\ell} \mathbf{w}_C^\top \mathbf{\Sigma}_C \mathbf{w}_C, \quad \forall \ell \in \{1, \ldots, N\}
\end{aligned}
\tag{3}
$$

3.1 Analysis of the Formulation

Computational Complexity The optimization problem in (3) is a QCQP involving $N+1$ variables (N weights and the scalar z) and is subject to N quadratic constraints. The pres-

[1] Obtained from `https://github.com/fja05680/sp500`

ence of N quadratic constraints makes the problem computationally demanding, and direct solution for large-scale applications (e.g., $N = 501$) is typically intractable with standard solvers. This necessitates the use of specialized algorithms or principled model relaxations.

Numerical Stability A key advantage of this formulation is its numerical stability. The objective function and constraints are defined directly by the covariance matrix $\mathbf{\Sigma}$ and its sub-matrices, not by its inverse $\mathbf{\Sigma}^{-1}$. This circumvents the need for matrix inversion, a process which is often unstable for the large, ill-conditioned, and error-prone covariance matrices typical of financial data. Consequently, this model should be more robust than classical mean-variance optimization frameworks that rely on $\mathbf{\Sigma}^{-1}$.

3.2 Implicit Control of Inter-Cluster Risk

An important feature of this formulation is the implicit control of inter-cluster risk. While the constraints in (3) are defined by the sum of intra-cluster variances at each level ℓ, the nested structure of the partitions ensures that total portfolio risk is effectively managed. This is most apparent at the highest level of the hierarchy ($\ell = N$), where a single cluster contains all assets. The constraint for this level, $z \geq \mathbf{w}^\top \mathbf{\Sigma} \mathbf{w}$, directly bounds the total variance of the portfolio.

To analyze how the model manages risk between clusters, we examine the merger of two disjoint clusters, A and B, into a new cluster $C = A \cup B$. The covariance matrix $\mathbf{\Sigma}_C$ and the corresponding weight sub-vector $\mathbf{w}_C$ for this merged cluster can be expressed in block form:

$$\mathbf{\Sigma}_C = \begin{pmatrix} \mathbf{\Sigma}_A & \mathbf{\Sigma}_{AB} \\ \mathbf{\Sigma}_{AB}^\top & \mathbf{\Sigma}_B \end{pmatrix}, \qquad \mathbf{w}_C = \begin{pmatrix} \mathbf{w}_A \\ \mathbf{w}_B \end{pmatrix}$$

Here, $\mathbf{\Sigma}_A$ and $\mathbf{\Sigma}_B$ are the covariance matrices for assets within clusters A and B respectively, and $\mathbf{\Sigma}_{AB}$ is the inter-cluster covariance matrix, whose elements $[\mathbf{\Sigma}_{AB}]_{ij}$ represent the covariance between asset $i \in A$ and asset $j \in B$.

The variance of the sub-portfolio corresponding to cluster C is given by the quadratic form $\mathrm{Var}(\mathbf{w}_C) = \mathbf{w}_C^\top \mathbf{\Sigma}_C \mathbf{w}_C$. Expanding this expression yields:

$$\begin{aligned} \mathrm{Var}(\mathbf{w}_C) &= \begin{pmatrix} \mathbf{w}_A^\top & \mathbf{w}_B^\top \end{pmatrix} \begin{pmatrix} \mathbf{\Sigma}_A & \mathbf{\Sigma}_{AB} \\ \mathbf{\Sigma}_{AB}^\top & \mathbf{\Sigma}_B \end{pmatrix} \begin{pmatrix} \mathbf{w}_A \\ \mathbf{w}_B \end{pmatrix} \\ &= \mathbf{w}_A^\top \mathbf{\Sigma}_A \mathbf{w}_A + \mathbf{w}_A^\top \mathbf{\Sigma}_{AB} \mathbf{w}_B + \mathbf{w}_B^\top \mathbf{\Sigma}_{AB}^\top \mathbf{w}_A + \mathbf{w}_B^\top \mathbf{\Sigma}_B \mathbf{w}_B. \end{aligned}$$

Since the term $\mathbf{w}_B^\top \mathbf{\Sigma}_{AB}^\top \mathbf{w}_A$ is a scalar, it is equal to its transpose, $\mathbf{w}_A^\top \mathbf{\Sigma}_{AB} \mathbf{w}_B$. This allows for a precise decomposition of the variance of the combined cluster C:

$$\mathrm{Var}(\mathbf{w}_C) = \underbrace{\mathbf{w}_A^\top \mathbf{\Sigma}_A \mathbf{w}_A}_{\text{Intra-cluster Var(A)}} + \underbrace{\mathbf{w}_B^\top \mathbf{\Sigma}_B \mathbf{w}_B}_{\text{Intra-cluster Var(B)}} + \underbrace{2\mathbf{w}_A^\top \mathbf{\Sigma}_{AB} \mathbf{w}_B}_{\text{Inter-cluster Risk}}$$

The term $2\mathbf{w}_A^\top \mathbf{\Sigma}_{AB} \mathbf{w}_B$ is the **inter-cluster risk**. It quantifies the contribution of the pairwise covariances between assets in clusters A and B to the total variance of the combined portfolio. A negative value for this term, arising from negative correlations between the clusters, indicates a diversification benefit. The optimization must therefore balance the intra-cluster variances at one level with the total combined variance, which includes this inter-cluster term, at the next level up in the hierarchy.

3.3 Constraint Reduction for Computational Tractability

The primary challenge of the formulation in (3) is the computational burden of solving a QCQP with $O(N)$ constraints. We noticed, however, that a significant number of these constraints are redundant, allowing for a principled reduction of the problem size. This redundancy arises from the nested structure of the hierarchy.

Consider a merger at level $\ell + 1$ where two clusters, A and B from the partition $\mathcal{P}_\ell$, form a new cluster $C = A \cup B$. All other clusters remain unchanged. The constraints from (3) for levels ℓ and $\ell + 1$ are:

$$z \geq \sum_{C' \in \mathcal{P}_\ell \setminus \{A,B\}} \mathrm{Var}(\mathbf{w}_{C'}) + \mathrm{Var}(\mathbf{w}_A) + \mathrm{Var}(\mathbf{w}_B) \tag{4}$$

$$z \geq \sum_{C' \in \mathcal{P}_{\ell+1} \setminus \{C\}} \mathrm{Var}(\mathbf{w}_{C'}) + \mathrm{Var}(\mathbf{w}_C) \tag{5}$$

The summation terms in both inequalities are identical, as they cover the same set of clusters that are not involved in the merger. The key difference lies in the remaining terms. By decomposing the variance of the merged cluster, $\mathrm{Var}(\mathbf{w}_C) = \mathrm{Var}(\mathbf{w}_A) + \mathrm{Var}(\mathbf{w}_B) + 2\mathbf{w}_A^\top \mathbf{\Sigma}_{AB} \mathbf{w}_B$, we can establish a direct relationship between the right-hand sides (RHS) of the two constraints:

$$\mathrm{RHS}_{\ell+1} = \mathrm{RHS}_\ell + 2\mathbf{w}_A^\top \mathbf{\Sigma}_{AB} \mathbf{w}_B \tag{6}$$

A constraint is redundant if it is dominated by another across the entire feasible set. Given the long-only constraint ($\mathbf{w} \geq \mathbf{0}$), which implies $\mathbf{w}_A \geq \mathbf{0}$ and $\mathbf{w}_B \geq \mathbf{0}$, we can determine whether a constraint is redundant *a priori* by examining the elements of the inter-cluster covariance matrix $\mathbf{\Sigma}_{AB}$.

- **Case I: All elements of $\mathbf{\Sigma}_{AB}$ are non-negative.**
  If every element $\sigma_{ij} \geq 0$ for $i \in A, j \in B$, then the quadratic form $\mathbf{w}_A^\top \mathbf{\Sigma}_{AB} \mathbf{w}_B = \sum_{i \in A, j \in B} w_i \sigma_{ij} w_j$ is guaranteed to be non-negative. This implies $\mathrm{RHS}_{\ell+1} \geq \mathrm{RHS}_\ell$. The constraint at level $\ell + 1$ is therefore more restrictive. Any feasible solution satisfying the constraint at level $\ell + 1$ automatically satisfies the one at level ℓ. Thus, the constraint from level ℓ is redundant and can be removed from the optimization problem.

- **Case II: All elements of $\mathbf{\Sigma}_{AB}$ are non-positive.**
  If every element $\sigma_{ij} \leq 0$ for $i \in A, j \in B$, the quadratic form $\mathbf{w}_A^\top \mathbf{\Sigma}_{AB} \mathbf{w}_B$ is guaranteed to be non-positive. This implies $\mathrm{RHS}_{\ell+1} \leq \mathrm{RHS}_\ell$. The constraint at level ℓ is now more restrictive. Any feasible solution satisfying the constraint at level ℓ automatically satisfies the one at level $\ell + 1$. In this situation, the constraint from level $\ell + 1$ is redundant and can be removed.

- **Case III: $\mathbf{\Sigma}_{AB}$ contains both positive and negative elements.**
  If the inter-cluster covariances have mixed signs, the sign of the quadratic form $\mathbf{w}_A^\top \mathbf{\Sigma}_{AB} \mathbf{w}_B$ cannot be determined *a priori*, as it depends on the specific values of the optimization variables $\mathbf{w}$. We cannot conclude that one constraint consistently dominates the other across the entire feasible set. Therefore, no simplification based on this analysis is possible, and both constraints—at level ℓ and $\ell + 1$—must be retained.

This analysis provides a rigorous method for model reduction prior to optimization. In practice, financial asset returns are predominantly positively correlated, and hierarchical clustering algorithms like Ward's linkage tend to group such assets together. As a result, Case I is the most frequent scenario, allowing for the systematic pruning of a majority of the pre-merger constraints. The final problem is thus reduced to a tractable size, containing only the final all-encompassing constraint and the pairs of constraints corresponding to the infrequent mergers with mixed-sign inter-cluster covariances (Case III).

This now allows us to reformulate the problem using only the set of non-redundant, or "active," constraints. This significantly reduces the computational burden while preserving the objective of the original model.

3.4 The Reduced Minimax Total Variance Problem

Let $\mathcal{L}_{active}$ be the set of indices of all non-redundant constraints identified through the pruning process. The original optimization problem from Eq. (3) can now be expressed in its reduced, computationally tractable form:

$$\begin{aligned} \underset{\mathbf{w}\in\mathbb{R}^N, z\in\mathbb{R}}{\text{minimize}} \quad & z \\ \text{subject to} \quad & \mathbf{1}^\top \mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0} \\ & z \geq \sum_{C\in\mathcal{P}_\ell} \mathbf{w}_C^\top \mathbf{\Sigma}_C \mathbf{w}_C, \quad \forall \ell \in \mathcal{L}_{active} \end{aligned} \tag{7}$$

This formulation is mathematically equivalent to the original problem but contains far fewer quadratic constraints, making it feasible to solve for large-scale applications.

3.5 Relaxation to Minimax Individual Cluster Variance

For further computational ease, or to adopt an alternative risk objective, we can modify the structure of the risk constraints. Instead of constraining z by the *sum* of intra-cluster variances, we can constrain it by the variance of the *single worst-performing cluster*. This constitutes a relaxation of the problem in Eq. (7).

The optimization problem for this relaxed formulation is:

$$\begin{aligned} \underset{\mathbf{w}\in\mathbb{R}^N, z\in\mathbb{R}}{\text{minimize}} \quad & z \\ \text{subject to} \quad & \mathbf{1}^\top \mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0} \\ & z \geq \mathbf{w}_C^\top \mathbf{\Sigma}_C \mathbf{w}_C, \quad \forall \ell \in \mathcal{L}_{active}, \ \forall C \in \mathcal{P}_\ell \end{aligned} \tag{8}$$

This formulation is a relaxation because its feasible region is a superset of the feasible region of the reduced problem in Eq. (7). Consider any level $\ell \in \mathcal{L}_{active}$. The constraint from the reduced problem is on the sum of variances, while the relaxed problem constrains z by the maximum of those individual variances.

Since each variance term $\mathbf{w}_C^\top \mathbf{\Sigma}_C \mathbf{w}_C$ is non-negative, it is a fundamental property that:

$$\sum_{C\in\mathcal{P}_\ell} \mathbf{w}_C^\top \mathbf{\Sigma}_C \mathbf{w}_C \geq \max_{C\in\mathcal{P}_\ell} (\mathbf{w}_C^\top \mathbf{\Sigma}_C \mathbf{w}_C) \tag{9}$$

This inequality demonstrates that the constraint in Eq. (7) is tighter. Any feasible solution $(\mathbf{w}, z)$ for the reduced problem satisfies $z \geq \sum \mathrm{Var}(C)$, which automatically implies $z \geq \max \mathrm{Var}(C)$. Thus, any feasible solution for the reduced problem is also feasible for the relaxed one, but the converse is not true. This larger feasible set may lead to a different optimal portfolio $\mathbf{w}^*$ and a lower objective value z^*.

3.6 Incorporating a Minimum Return Target

To ensure the optimized portfolio meets a specified performance objective, either of the tractable formulations can be augmented with a linear constraint on its expected return. This requires the portfolio's expected return, $\mathbb{E}[R_p] = \boldsymbol{\mu}^\top \mathbf{w}$, to exceed that of a benchmark, such as the Equally Weighted Portfolio (EWP), by a given margin δ. The EWP return is the simple average of individual asset returns, $\bar{\mu} = \frac{1}{N}\mathbf{1}^\top \boldsymbol{\mu}$. The constraint is formally expressed as:

$$\boldsymbol{\mu}^\top \mathbf{w} \geq \bar{\mu} + \delta \tag{10}$$

where $\boldsymbol{\mu} \in \mathbb{R}^N$ is the vector of mean historical returns and $\delta > 0$ is a constant representing the minimum target excess return. The inclusion of this constraint transforms the problem from pure risk minimization to a constrained optimization that seeks a portfolio satisfying the minimum return requirement. (NOTE: δ is a DAILY return target)

4 Final Optimization Model

The final formulated model seeks to minimize the maximum variance found in any single cluster across a pre-determined set of non-redundant hierarchical levels, subject to achieving a minimum expected return, it is a Quadratically Constrained Quadratic Program (QCQP) defined as follows:

$$\begin{aligned}
\underset{\mathbf{w}\in\mathbb{R}^N, z\in\mathbb{R}}{\text{minimize}} \quad & z \\
\text{subject to} \quad & \mathbf{1}^\top \mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0} && \text{(Budget and Long-Only)} \\
& \boldsymbol{\mu}^\top \mathbf{w} \geq \bar{\mu} + \delta && \text{(Minimum Return Target)} \\
& z \geq \mathbf{w}_C^\top \boldsymbol{\Sigma}_C \mathbf{w}_C, \quad \forall \ell \in \mathcal{L}_{active},\ \forall C \in \mathcal{P}_\ell && \text{(Minimax Cluster Risk)}
\end{aligned} \tag{11}$$

The components of the model are:

- **Decision Variables:** The portfolio weight vector $\mathbf{w} \in \mathbb{R}^N$ and a scalar auxiliary variable $z \in \mathbb{R}$ that represents the maximum individual cluster variance.
- **Objective Function:** The objective is to minimize z, effectively minimizing the risk of the single worst-performing (i.e., highest variance) cluster.
- **Constraints:**
    1. The standard portfolio constraints, ensuring weights are non-negative and sum to one.

2. The minimum return constraint, which mandates that the portfolio's expected return $\boldsymbol{\mu}^\top \mathbf{w}$ must exceed the return of the Equally Weighted Portfolio ($\bar{\mu}$) by a specified hurdle rate δ.
3. The set of quadratic risk constraints. These are applied for each cluster C within each partition $\mathcal{P}_\ell$ for all levels ℓ in the active set $\mathcal{L}_{active}$. This ensures that z is greater than or equal to the variance of every individual cluster being monitored.

The inputs to this model are the asset covariance matrix $\boldsymbol{\Sigma}$, the vector of mean asset returns $\boldsymbol{\mu}$, the set of active hierarchical levels $\mathcal{L}_{active}$, and the return hurdle rate δ. The resulting QCQP is solved using a numerical optimization solver to find the optimal portfolio weights $\mathbf{w}^*$.

5 Results

5.1 Data and Model Inputs

We implement our model using daily adjusted closing prices for the 501 constituents of the S&P 500, with historical data sourced via the `yfinance` library up to 10 years prior to July 3, 2025. Daily percentage returns are computed, and this return series forms the basis for our analysis. The asset covariance matrix, Σ, is estimated using the Ledoit-Wolf shrinkage estimator to ensure it is well-conditioned and robust to estimation error. The correlation matrix, ρ, is derived directly from the return series, also chosen δ is 0

5.2 Hierarchical Structure and Constraint Reduction

We apply Ward's linkage hierarchical clustering to the asset distance matrix, yielding the dendrogram shown in Figure 1. This dendrogram defines the nested partitions used in our optimization framework.

A key element of our methodology was the *a priori* reduction of problem complexity. As detailed in Section 4.3, the number of active risk constraints depends on the sign of the inter-cluster covariances at each merger step. We find that the asset correlations are overwhelmingly positive, as illustrated by the distribution of off-diagonal correlation coefficients in Figure 2. Consequently, the vast majority of mergers fall under Case I, where the pre-merger constraint is redundant. This process prunes the number of active hierarchical levels from an initial 501 to just 9, dramatically improving the computational tractability of the optimization problem.

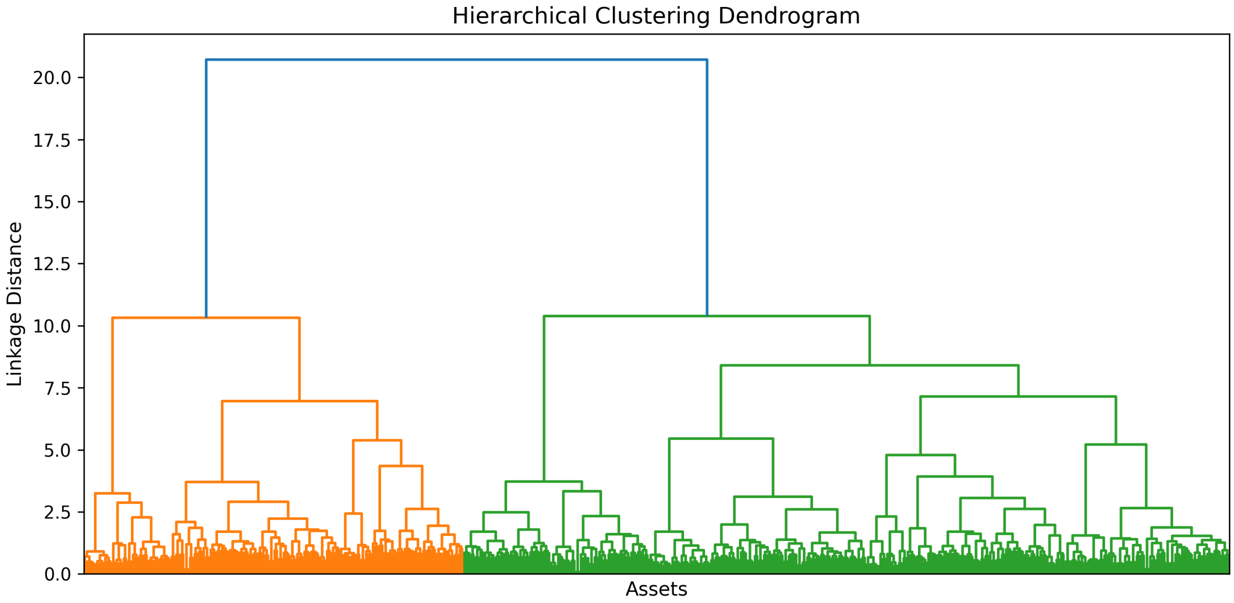

Figure 1: Asset dendrogram generated by Ward's linkage clustering on the correlation-based distance matrix.

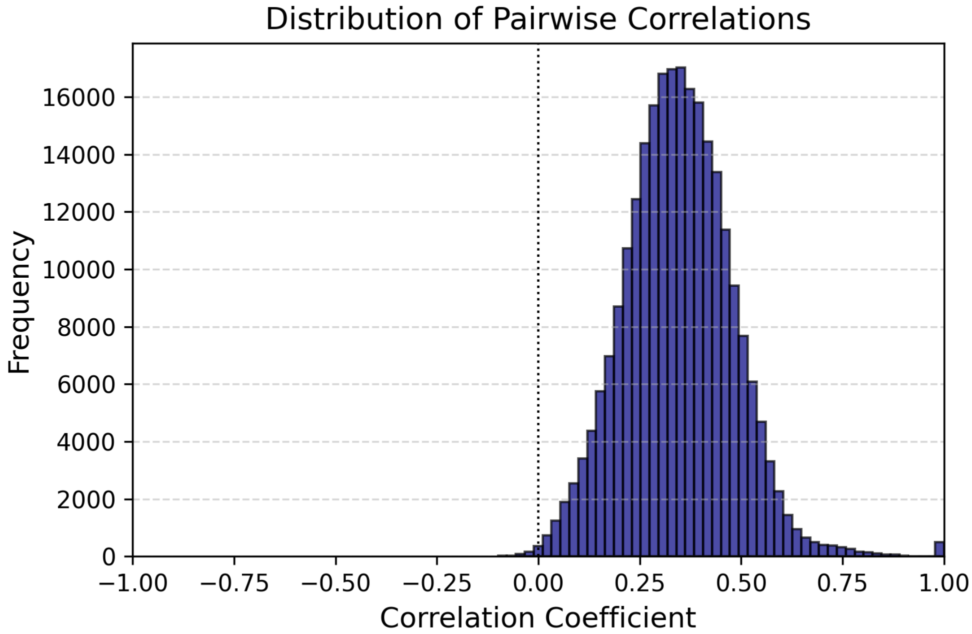

Figure 2: Distribution of off-diagonal elements in the asset correlation matrix. The high prevalence of positive correlations enables significant constraint reduction.

5.3 Portfolio Allocation and Performance Comparison

We compare the portfolio generated by our proposed model against the Hierarchical Risk Parity (HRP) algorithm implemented using the `PyPortfolioOpt` library. Table 1 summarizes the performance metrics for both portfolios.

Table 1: Performance comparison between our proposed model and the HRP benchmark.

Metric	**Our Model**	**HRP**
Expected Annual Return	24.03%	12.9%
Annual Volatility	20.97%	14.2%
Sharpe Ratio ($r_f = 0.00\%$)	1.15	0.91

The differing risk-return profiles can be better understood by the difference in top 10 and bottom 10 allocations in both the portfolios detailed in Table 2. Another visualization can be the scatter plot of distributions of both the weight vectors as shown in Figure 3

Table 2: Top and bottom 10 portfolio weights for our proposed model versus HRP.

Our Model		**HRP**	
Ticker	**Weight (%)**	**Ticker**	**Weight (%)**
Top 10 Holdings			
NVDA	3.276	VLTO	4.450
AMD	2.390	SW	3.361
ENPH	1.739	KVUE	3.220
SMCI	1.645	SOLV	3.010
TSLA	1.206	GEHC	1.805
AXON	0.981	GEV	0.740
ANET	0.973	JNJ	0.571
AVGO	0.949	KMB	0.537
MPWR	0.930	OTIS	0.523
KLAC	0.846	VZ	0.510
Bottom 10 Holdings			
VTRS	0.028	NCLH	0.033
WBA	0.031	APA	0.038
KHC	0.067	SMCI	0.039
PARA	0.070	ON	0.040
BIIB	0.070	CZR	0.042
KVUE	0.074	RCL	0.043
WBD	0.077	FANG	0.043
CAG	0.082	CCL	0.045
CPB	0.084	ENPH	0.045
SW	0.086	FCX	0.047

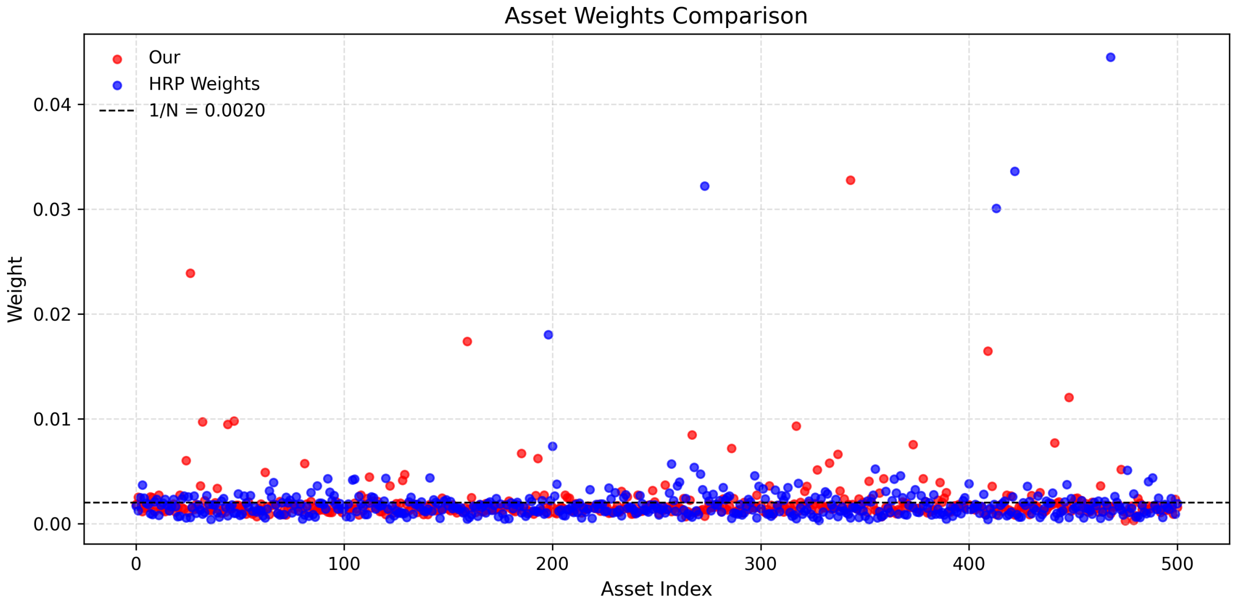

Figure 3: Scatter plot comparing the weight distributions of our model and HRP.

5.4 Sector-Level Diversification

We can see a broad Sector-Level Diversification using predefined GICS sectors, the pie-chart distributions for both the portfolios are as shown

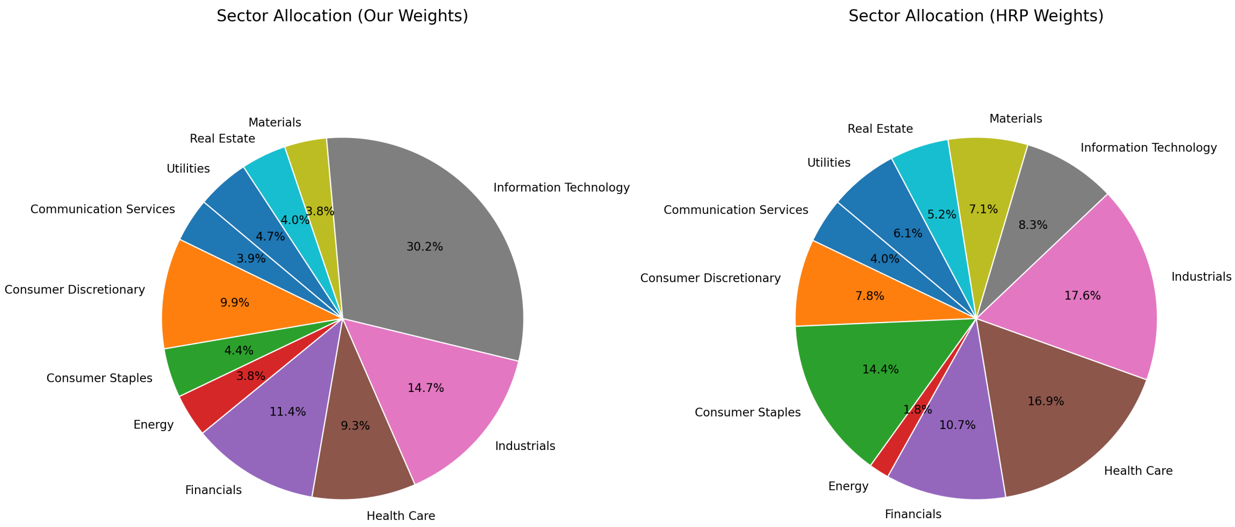

Figure 4: GICS sector allocation for the portfolio constructed by our model.

The within cluster (GICS) distribution for our portfolio weights are as follows

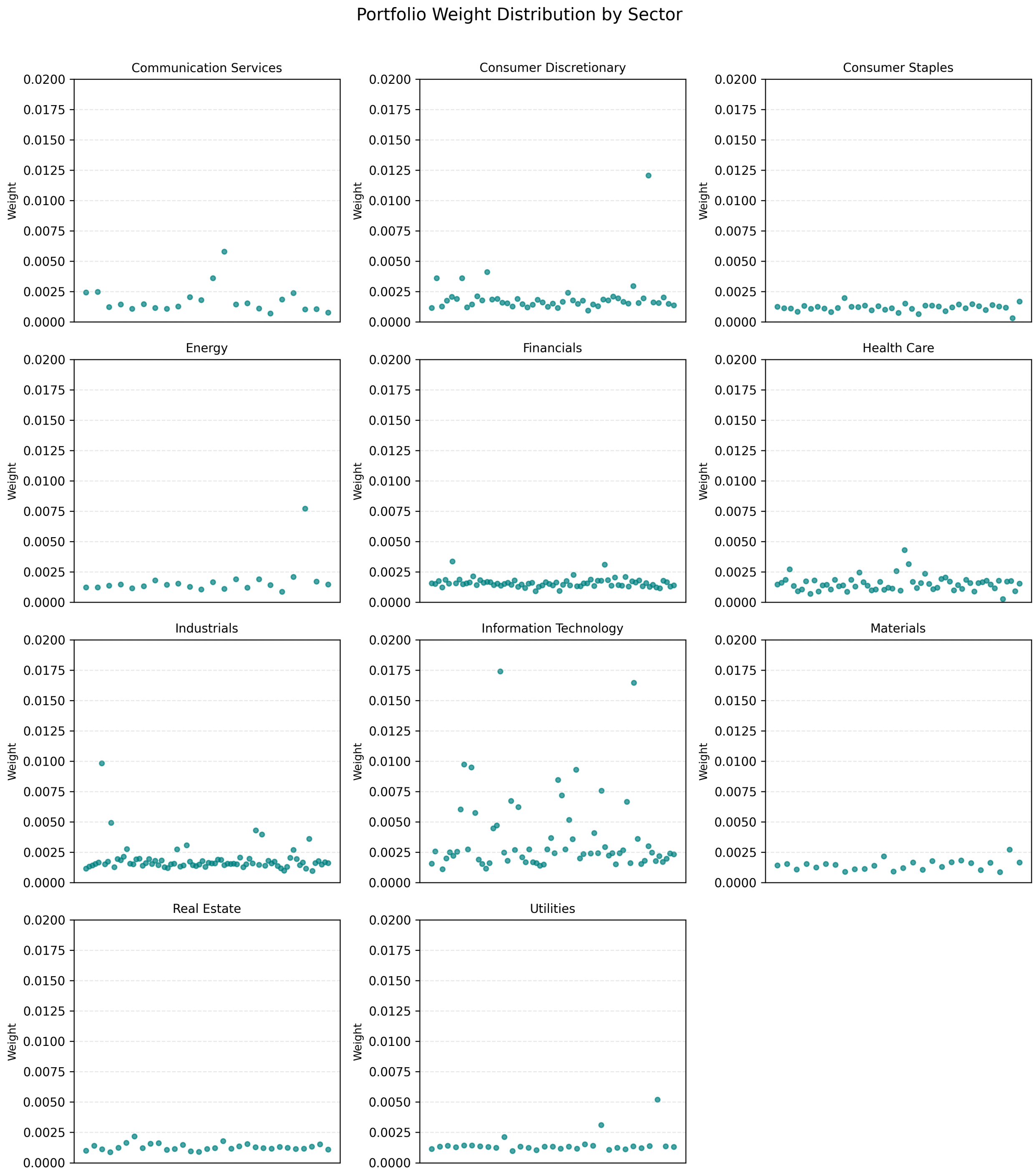

Figure 5: Our portfolio weight distribution organized by GICS sector.

References

[1] Apollo Academy. Extreme concentration in the s&p 500. `https://www.apolloacademy.com/extreme-concentration-in-the-sp-500-2/`, July 2025.

[2] Marcos López de Prado. Building diversified portfolios that outperform out-of-sample. *Journal of Portfolio Management*, 42(4):59–69, May 2016. Available at SSRN: `https://ssrn.com/abstract=2708678`.