IMPERIAL

Imperial College London
Department of Mathematics
MSc Mathematics and Finance

Convex Optimisation for Improved Covariance Estimation in Portfolio Construction

MATH97235 – Portfolio Management Coursework

Ilia BOLGOV

Jonathan CASPER

Chris DARCY

December 19, 2025

Abstract

Estimating large covariance matrices from limited, noisy return data is central to portfolio construction but remains difficult. The sample covariance matrix (SCM) is unstable in high dimensions and degrades under heavy-tailed returns. While shrinkage, Random Matrix Theory (RMT) cleaning, and robust estimators each help, no single method dominates across distributional regimes. We propose a convex-optimised ensemble that combines several covariance estimators, choosing weights to minimise validation out-of-sample portfolio variance (as a proxy for estimation quality specific to the application - portfolio optimisation). We test on synthetic data (Gaussian, Student-t, Pareto, and factor models) with known ground-truth covariances and on real stock returns. Base estimators include the SCM, Marchenko–Pastur eigenvalue cleaning, structured shrinkage, and Tyler's robust scatter estimator. We learn convex weights using train/validation/test splits. No single estimator is best in all regimes (SCM under Gaussianity, Tyler under heavy tails, MP under moderate noise), but the proposed convex combination consistently delivers the lowest out-of-sample variance, improving robustness and adaptability—especially on real stock returns.

Contents

List of Figures

List of Tables

1 Introduction

Estimating large covariance matrices from limited, noisy, and heavy-tailed financial return data is an important challenge in financial mathematics, as accurate covariance estimation underpins mean–variance portfolio optimisation [13], quantitative risk management, and other crucial applications. The most common estimator - sample covariance matrix (SCM) deteriorates rapidly when the dimensionality of the problem approaches the number of observations, yielding unstable estimates [12, 8]. Furthermore, financial returns frequently exhibit heavy tails and deviations from multivariate normality [4] - conditions under which the SCM suffers from bias and variance issues [9, 17].

Recent literature has proposed several strategies to regularise the SCM. Shrinkage estimators [11] stabilise the covariance matrix by imposing structure through a convex blend of the SCM and a well-conditioned target matrix. Random Matrix Theory (RMT) approaches, such as Marchenko–Pastur eigenvalue cleaning, selectively correct noisy eigenvalues and are effective in high-dimensional asymptotic regimes [11, 10, 14, 3] .

Deshmukh and Dubey (2020) demonstrate that combining shrinkage and RMT estimators through a convex weighting scheme can outperform standalone methods in real stock-return datasets, producing lower out-of-sample portfolio risk [5]. Their framework illustrates that heterogeneous estimators possess complementary strengths and that convex optimisation can adapt to varying noise levels and structural properties in financial data.

However, the framework in Deshmukh and Dubey (2020) is limited in two respects. First, it evaluates the estimator exclusively on real-world financial datasets, leaving open the question of how such convex combinations perform across controlled synthetic environments where the ground-truth covariance is known. Second, their convex combination includes only SCM, a shrinkage estimator, and an MP-cleaned estimator. While these estimators handle sampling noise and structural instability, none is well-suited for heavy-tailed distributions - a crucial feature of financial returns. The MP law, for example, relies on finite-moment assumptions and breaks down when return distributions exhibit infinite or near-infinite variance [1, 2]. The SCM is highly sensitive to heavy tails and produces erratic estimates in the presence of outliers or tail events. Likewise, linear shrinkage can still be sensitive to heavy-tailed observations because it remains a convex combination of the sample covariance and a target matrix.

To address these limitations, we augment the convex-combination framework by incorporating Tyler's M-estimator of scatter (Tyler, 1987) [16], a distribution-free, shape-based estimator specifically designed to be robust under heavy-tailed distributions. Tyler's estimator requires no finite-moment assumptions, is invariant under monotone radial transformations, and has been widely recognised as one of the most reliable estimators for modelling financial returns, which often deviate substantially from Gaussianity.

However, Tyler's M-estimator alone does not explicitly regularise sampling noise, shrink extreme eigenvalues, or regularise the covariance structure [15], all of which are important for reliable estimation in financial markets. MP eigenvalue cleaning remains valuable because it removes bulk sampling noise in high-dimensional settings, recovering clearer dependence structure. Shrinkage contributes additional regularisation by pulling noisy sample covariances toward a structured target, reducing variance and improving conditioning. Combining these estimators with Tyler therefore allows the convex optimisation to retain Tyler's robustness to heavy tails while benefiting from the noise reduction, stabilisation, and structural information provided by SCM, shrinkage, and MP cleaning-features.

We propose to learn the convex weights of these four estimators through a validation-set optimisation procedure, using out-of-sample portfolio variance as a proxy objective. This design explicitly targets the estimator's performance in practical portfolio-optimisation tasks while enabling the model to adapt to the underlying return distribution. To rigorously evaluate this approach, we test it not only on real stock-return datasets - as in Deshmukh and Dubey (2020), but also on an extensive suite of synthetic datasets including Gaussian, Student-t, Pareto, and factor-model returns. These controlled experiments allow us to quantify estimation error relative to known ground truth and to assess the estimator's robustness across different distributions.

2 Estimators

In this section, we introduce the four covariance estimators used throughout our analysis: the sample covariance matrix (SCM), a Marchenko-Pastur (MP) eigenvalue-cleaned estimator, a linear shrinkage estimator, and Tyler's M-estimator of scatter.

Each estimator captures different structural or robustness properties of financial return data, and their complementary strengths motivate the convex aggregation framework developed in this study.

Let $R : \Omega \to \mathbb{R}^M$ denote the random return vector of M assets, and let $\{x_1, \dots, x_N\}$ be N independent realisations of R. We collect these observations in the data matrix

$$X = \begin{pmatrix} x_1 & x_2 & \cdots & x_N \end{pmatrix} \in \mathbb{R}^{M\times N},$$

where each column x_i corresponds to one observed return vector. **Throughout, we assume the data is mean-centred** - i.e. we subtract a mean vector estimator from each observation before proceeding. We will use this definition when constructing the estimators.

2.1 Sample Covariance Matrix Estimator

Let $X \in \mathbb{R}^{M\times N}$ denote the mean-centred return matrix, where each column corresponds to an M-dimensional observation. The sample covariance matrix (SCM) is defined as

$$\Sigma_{\text{SCM}} = \frac{1}{N} X X^\top.$$

Equivalently, if $x_1, \dots, x_N \in \mathbb{R}^M$ denote the mean-centred return vectors, then

$$\Sigma_{\text{SCM}} = \frac{1}{N} \sum_{i=1}^{N} x_i x_i^\top. \tag{1}$$

2.2 Marchenko-Pastur Estimator

The Marchenko-Pastur (MP) law characterises the asymptotic distribution of the eigenvalues of the sample covariance matrix when both the number of assets M and the number of observations N grow large with fixed ratio $c = M/N \in (0, 1)$.

Let $X \in \mathbb{R}^{M\times N}$ denote the mean-centred return matrix and $\Sigma_{\text{SCM}} = \frac{1}{N} X X^\top$ its sample covariance matrix. If $\lambda_1, \dots, \lambda_M \in \mathbb{R}$ denote the eigenvalues of Σ_{SCM}, then under suitable finite-moment assumptions their empirical distribution (ESD) converges to the Marchenko-Pastur density

$$g_{\text{MP}}(x) = \frac{\sqrt{(\lambda_+ - x)(x - \lambda_-)}}{2\pi c\, x}, \qquad \lambda_- = (1 - \sqrt{c})^2,\ \lambda_+ = (1 + \sqrt{c})^2,$$

supported on the interval $[\lambda_-, \lambda_+]$.

Since $\Sigma_{\text{SCM}} = \frac{1}{N} X X^\top$ is a real symmetric matrix, it admits an eigendecomposition $\Sigma_{\text{SCM}} = U \Lambda U^\top$ with U orthonormal and Λ real and diagonal.

Eigenvalues lying within $[\lambda_-, \lambda_+]$ are interpreted as dominated by sampling noise. MP eigenvalue cleaning replaces these noisy eigenvalues by a common value, while leaving the eigenvalues outside the MP bounds unchanged [5]. In our paper we will be replacing them by their average.

$$\lambda_{replaced} = \frac{1}{k} \sum_{i=1}^{k} \lambda_i^{noise}$$

for

$$\lambda_i^{noise} \in [\lambda_-, \lambda_+].$$

If U denotes the matrix of eigenvectors of Σ_{SCM} and $\tilde{\Lambda}$ the cleaned eigenvalue matrix, the MP-cleaned covariance estimator is

$$\Sigma_{\text{MP}} = U\, \tilde{\Lambda}\, U^\top. \tag{2}$$

2.3 Shrinkage Estimator

Shrinkage estimators stabilise the sample covariance matrix by pulling it toward a well-conditioned target matrix. Let Σ_{SCM} denote the sample covariance matrix and F a structured target.

Ledoit–Wolf (2004) Target. Following Ledoit and Wolf (2004) [11], we use the "sample variance and mean covariance" target, defined as

$$F = \mathrm{diag}(\Sigma_{\mathrm{SCM}}) + \bar{\sigma}\left(\mathbf{1}\mathbf{1}^\top - I\right), \tag{3}$$

where $\mathrm{diag}(\Sigma_{\mathrm{SCM}})$ contains the sample variances of the assets, I is the identity matrix, and $\bar{\sigma}$ is the average of all off-diagonal elements of Σ_{SCM}. This target preserves individual asset variances while imposing a homogeneous structure on all pairwise covariances.

We adopt the Ledoit–Wolf target because it has been shown to outperform other linear shrinkage targets in high-dimensional financial settings: it delivers substantial variance reduction, improves conditioning, and remains effective even when the underlying covariance structure is unknown or weakly structured. [11] These properties make it a natural choice for inclusion in our convex aggregation framework.

A linear shrinkage estimator takes the form

$$\Sigma_{\mathrm{shrink}} = \rho F + (1-\rho)\,\Sigma_{\mathrm{SCM}}, \qquad 0 \le \rho \le 1,$$

where ρ is the shrinkage intensity. Large values of ρ impose stronger structure through F, while smaller values place greater weight on the empirical covariance Σ_{SCM}. Shrinkage reduces estimation variance and improves conditioning by regularising extreme eigenvalues of Σ_{SCM}.

Optimal Shrinkage Intensity. Ledoit and Wolf (2004) show that, for linear shrinkage estimators of the form $\Sigma_{\mathrm{shrink}} = \rho F + (1-\rho)\Sigma_{\mathrm{SCM}}$, the shrinkage intensity ρ that minimises the Frobenius norm error $\mathbb{E}\,\|\Sigma_{\mathrm{shrink}} - \Sigma\|_F^2$ with respect to the population covariance admits a closed-form solution. This analytical expression relies on the specific linear structure of the estimator.

In our setting, however, the convex combination includes not only SCM and a shrinkage estimator but also an MP-cleaned estimator and Tyler's estimator, making it impossible to derive optimal weights analytically; instead, we learn them via validation-based optimisation.

2.4 Tyler's M-Estimator

Tyler's M-estimator (Tyler, 1987) [16] is a distribution-free, shape-based estimator of scatter that is robust to heavy-tailed and elliptically distributed data. Given mean-centred observations $x_1, \ldots, x_N \in \mathbb{R}^M$, Tyler's estimator Σ_{Tyler} is defined implicitly as the solution to

$$\Sigma_{\mathrm{Tyler}} = \frac{M}{N}\sum_{i=1}^{N} \frac{x_i x_i^\top}{x_i^\top \Sigma_{\mathrm{Tyler}}^{-1} x_i},$$

subject to a normalisation constraint such as $\mathrm{tr}(\Sigma_{\mathrm{Tyler}}) = M$.

Importantly, Tyler's estimator is *scale-invariant*: it estimates the shape of the covariance matrix but not its overall magnitude. Tyler's estimator is defined only up to a positive scalar multiple (it estimates *shape* rather than *scale*). Under mild conditions—in particular, when $N > M$ and the observations are not concentrated on a proper subspace—the fixed-point equation admits a unique solution up to scale.

Since Tyler's estimator is only defined up to a positive scale factor, we fix its magnitude before entering the convex aggregation. In particular, we rescale the raw Tyler scatter estimate to match the overall variance level of the sample covariance matrix,

$$\tilde{\Sigma}_{\mathrm{Tyler}} = \frac{\mathrm{tr}(\Sigma_{\mathrm{SCM}})}{\mathrm{tr}(\Sigma_{\mathrm{Tyler}})}\,\Sigma_{\mathrm{Tyler}}. \tag{4}$$

This ensures that $\tilde{\Sigma}_{\mathrm{Tyler}}$ is on the same scale as Σ_{SCM}, F (shrinkage target), and Σ_{MP}, so that the subsequent optimisation over convex weights reflects differences in estimator quality rather than arbitrary scale choices.

3 Convex Optimisation Problem

The estimators introduced in Section 2 provide four matrices: the sample covariance Σ_{SCM} (1), the MP-cleaned estimator Σ_{MP} (2), the Ledoit–Wolf target matrix F (3), and the rescaled Tyler estimator $\tilde{\Sigma}_{\text{Tyler}}$ (4). We construct our final covariance estimator as a convex combination of these four matrices,

$$\Sigma^* = w_F F + w_{\text{MP}}\Sigma_{\text{MP}} + w_{\text{SCM}}\Sigma_{\text{SCM}} + w_{\text{Tyler}}\tilde{\Sigma}_{\text{Tyler}}, \tag{5}$$

where the weights satisfy

$$w = (w_F, w_{\text{MP}}, w_{\text{SCM}}, w_{\text{Tyler}}) \qquad w_F, w_{\text{MP}}, w_{\text{SCM}}, w_{\text{Tyler}} \geq 0, \qquad w_F + w_{\text{MP}} + w_{\text{SCM}} + w_{\text{Tyler}} = 1. \tag{6}$$

Setting $w_{\text{Tyler}} = 0$ recovers the three-parameter framework of Deshmukh and Dubey [5]; in particular, linear shrinkage estimators of the form $\Sigma_{\text{shrink}} = \rho F + (1-\rho)\Sigma_{\text{SCM}}$ arise as the special case $w_{\text{MP}} = w_{\text{Tyler}} = 0$, $w_F = \rho$, $w_{\text{SCM}} = 1-\rho$. Our formulation therefore strictly generalises both MP cleaning and linear shrinkage by adding Tyler's robust estimator as a further vertex in the convex hull.

3.1 Theoretical Objective

Our theoretical goal when constructing an estimator in (5) is to minimise the Frobenius distance from the underlying population covariance matrix Σ_{pop}, i.e. choose $w \in \Delta_4$ such that

$$\left\|\Sigma^*(w) - \Sigma_{\text{pop}}\right\|_F^2 \to \min! \tag{7}$$

where Δ_4 denotes the 4-dimensional probability simplex defined by (6).

In reality, however, Σ_{pop} is unknown and in our particular case there is also no possibility to find an analytical expression for w (unlike, for example, Ledoit-Wolf Shrinkage-based estimator where the optimal value of ρ was derived analytically).

3.2 Practical Optimisation via Out-of-Sample Variance

Since the population covariance matrix Σ_{pop} is unknown and no analytical solution exists for the optimal weights w in (5), we adopt a data-driven approach that directly optimises the performance of the convex estimator in portfolio construction task - in the same way as in the paper by Deshmukh and Dubey [5, 6].

Train-Validation-Test Framework. We employ a rolling-window procedure in which, at each rebalancing date t, we partition the available return history into three sets:

1. **Training set** (N_{train} observations): used to compute the four base estimators Σ_{SCM}, Σ_{MP}, F, and $\tilde{\Sigma}_{\text{Tyler}}$.
2. **Validation set** (N_{val} observations): used to select the optimal convex weights w^*.
3. **Test set** (N_{test} observations): used to evaluate out-of-sample performance.

Reparametrisation of the Simplex. To ensure that the weights automatically satisfy the simplex constraints (6), we introduce a three-parameter reparametrisation $(\theta, \phi, \psi) \in [0,1]^3$ defined by

$$w_{\text{SCM}} = 1 - \phi, \tag{8}$$
$$w_F = \phi\,\theta, \tag{9}$$
$$w_{\text{MP}} = \phi\,(1-\theta)\,\psi, \tag{10}$$
$$w_{\text{Tyler}} = \phi\,(1-\theta)\,(1-\psi). \tag{11}$$

By construction, these weights are non-negative and sum to unity for any $(\theta, \phi, \psi) \in [0,1]^3$. The parameter ϕ controls the overall blend between Σ_{SCM} and the regularised estimators; θ balances the shrinkage target F against the robust blend of Σ_{MP} and $\tilde{\Sigma}_{\text{Tyler}}$; and ψ determines the relative contribution of Σ_{MP} versus $\tilde{\Sigma}_{\text{Tyler}}$ within that robust component.

Grid Search over (θ, ϕ, ψ). For each rebalancing period, we perform a three-dimensional grid search over $(\theta, \phi, \psi) \in \{0, 0.1, 0.2, \dots, 1\}^3$. At each grid point:

1. Construct the base estimators from the training set.
2. Form the convex combination $\Sigma^*(\theta, \phi, \psi)$ using (5) with weights given by (8)–(11).
3. Evaluate the candidate covariance estimator via a portfolio functional. Given $\Sigma^*(\theta, \phi, \psi)$, we compute the *long-only minimum-variance portfolio induced by* Σ^* with a minimum annualised return threshold:

$$\min_{p \in \mathbb{R}^M} \quad p^\top \Sigma^*(\theta, \phi, \psi)\, p \quad (\approx \mathrm{Var}(p^\top X)) \tag{12}$$

$$\text{s.t.} \quad \mathbf{1}^\top p = 1, \qquad \hat{\mu}^\top p \geq r_{\min}, \qquad p \succeq 0, \tag{13}$$

yielding $p^*(\theta, \phi, \psi)$.

Here $\hat{\mu}$ is an estimate of expected returns (computed from the training window) and $r_{\min}$ is converted to the corresponding per-period threshold consistent with the data frequency.

Note that here we use

$$\begin{aligned} \mathrm{Var}(p^\top X) &= \mathbb{E}\Big[\big(p^\top X - \mathbb{E}[p^\top X]\big)^2\Big] \\ &= \mathbb{E}\Big[\big(p^\top (X - \mathbb{E}[X])\big)^2\Big] \\ &= \mathbb{E}\big[p^\top (X - \mathbb{E}[X])(X - \mathbb{E}[X])^\top p\big] \\ &= p^\top\, \mathbb{E}\big[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^\top\big]\, p \\ &= p^\top \Sigma\, p, \end{aligned} \tag{14}$$

4. Proxy loss for covariance selection. We define the validation loss of the covariance estimator as the realised out-of-sample variance of the induced portfolio:

$$\sigma^2_{\mathrm{val}}(\theta, \phi, \psi) = \frac{1}{N_{\mathrm{val}}} \sum_{i=1}^{N_{\mathrm{val}}} \big(r_i^\top p^*(\theta, \phi, \psi)\big)^2.$$

where r_i denotes the i-th validation-period return vector.

We then select the hyperparameters that minimise the validation-set variance,

$$(\theta^*, \phi^*, \psi^*) = \arg \min_{(\theta,\phi,\psi) \in \mathcal{G}} \sigma^2_{\mathrm{val}}(\theta, \phi, \psi),$$

where $\mathcal{G} = \{0, 0.1, \dots, 1\}^3$ denotes the discrete grid.

Averaging Across Rebalancing Periods. In practice, we perform this procedure over multiple non-overlapping rebalancing windows. At each rebalancing date $t = 0, T_{\mathrm{rebal}}, 2T_{\mathrm{rebal}}, \dots$, we:

1. Re-estimate the four base covariance matrices on the updated training window.
2. Re-optimise $(\theta^*, \phi^*, \psi^*)$ on the updated validation window.
3. Evaluate the resulting constrained minimum-variance portfolio on the test set for that period.

Comparison with Deshmukh and Dubey (2020). Setting $\psi = 1$ (equivalently, $w_{\mathrm{Tyler}} = 0$) recovers the three-estimator framework of Deshmukh and Dubey [5], in which only Σ_{SCM}, F, and Σ_{MP} are blended. Our extension to four estimators via the inclusion of $\tilde{\Sigma}_{\mathrm{Tyler}}$ is designed to improve robustness under heavy-tailed return distributions, where the MP law may fail and the SCM remains highly sensitive to outliers.

4 Simulations

4.1 Data

We tested our proposed framework against the other covariance matrix estimators outlined above on both synthetic data and real historical financial returns, for 30, 60, and 90 day rolling window periods. We generated synthetic data using Gaussian, Pareto, and Student-t distributions with appropriate parameters for this paper. We sampled real historical returns from S&P 500, NASDAQ-100, Nikkei 225, NSE Nifty 50, and FTSE 100 indices - in the case with S&P 500 we randomly selected ∼100 stocks to better demonstrate the performance of the estimator, as the top equities in the index significantly dominate the rest of the stocks and are highly correlated. In each case, we solved the aforementioned optimisation problem with the constraint that annualized portfolio returns exceed at least 0%, 5%, and 10%. We omit the results from the cases 0% and 5% here for ease of readability and to present the results in a manner similar to the paper by Deshmukh and Dubey, 2020 [5]. The results for the 0% and 5% cases are similar to the ones for 10% and show even stronger performance by our proposed estimator. Full code and results can be found on the GitHub repository listed in the appendix [7].

4.2 Results on Synthetic Data

Our results with the synthetic data are exactly as expected: our proposed 3D estimator (we will further refer to our proposed estimator as that) yields the lowest annualized volatility for every distribution tested and for every time horizon. The Tyler-M [16] and 2D estimator proposed by Deshmukh [5] also performed well. In the Gaussian case, the Deshmukh 2D model performed better than the Tyler-M estimator, however the Tyler-M estimator was a better fit for the heavier-tailed data from Student-t and Pareto distributions. It is unsurprising that our 3D estimator - which incorporates both the 2D Deshmukh and Tyler-M estimators - performed better than either individual estimator across all distributions and time horizons.

Fig. 1 below shows the optimised parameter weightings for our proposed 3D estimator for each distribution and time horizon. Indeed, the figure demonstrates the adaptability of our estimator for different types of data; our estimator heavily weights the Tyler-M estimator for the heavy-tailed Student-t and Pareto distributions, while relying on other estimators in the Gaussian case. Thus, our proposed estimator is able to adaptively select the best estimator(s) for the given dataset, which allows it to outperform all other individual estimators we tested.

Table 1: Annualized Volatility (%) - 30 Day Synthetic (Factor, Gaussian, Pareto, Student-t)

Estimator	Factor	Gaussian	Pareto	Student-t
Σ_{Identity}	2.64	16.84	14.34	16.38
Σ_{Scaled}	2.60	16.87	14.37	16.39
Σ_{SCM}	1.96	10.19	8.66	10.14
Σ_{Target}	2.45	13.66	11.08	13.98
Σ_{MP}	1.94	10.55	8.85	10.37
Σ_{Tyler}	1.95	10.19	8.75	9.99
$\Sigma^*_{(\text{Paper 2D})}$	1.94	10.16	8.63	10.12
$\Sigma^*_{(\text{Our 3D})}$	1.94	10.15	8.62	9.96

Table 2: Annulized Volatility (%) - 60 Day Synthetic (Factor, Gaussian, Pareto, Student-t)

Estimator	Factor	Gaussian	Pareto	Student-t
Σ_{Identity}	2.66	17.06	14.54	16.55
Σ_{Scaled}	2.64	17.10	14.56	16.56
Σ_{SCM}	1.99	10.19	9.03	10.43
Σ_{Target}	2.48	13.28	12.46	14.44
Σ_{MP}	1.96	10.55	9.20	10.56
Σ_{Tyler}	1.99	10.17	8.92	9.99
$\Sigma^*_{(\text{Paper 2D})}$	1.95	10.14	9.00	10.34
$\Sigma^*_{(\text{Our 3D})}$	1.95	10.13	8.90	9.98

Table 3: Annualized Volatility (%) – 90 Days Synthetic (Factor, Gaussian, Pareto, Student-t)

Estimator	Factor	Gaussian	Pareto	Student-t
Σ_{Identity}	2.66	16.96	14.65	16.46
Σ_{Scaled}	2.64	16.96	14.69	16.48
Σ_{SCM}	2.00	10.24	8.84	10.48
Σ_{Target}	2.44	14.10	11.33	14.07
Σ_{MP}	1.96	10.62	9.02	10.57
Σ_{Tyler}	2.00	10.24	8.77	10.00
$\Sigma^*_{\text{(Paper 2D)}}$	1.94	10.21	8.79	10.40
$\Sigma^*_{\text{(Our 3D)}}$	1.94	10.20	8.71	9.97

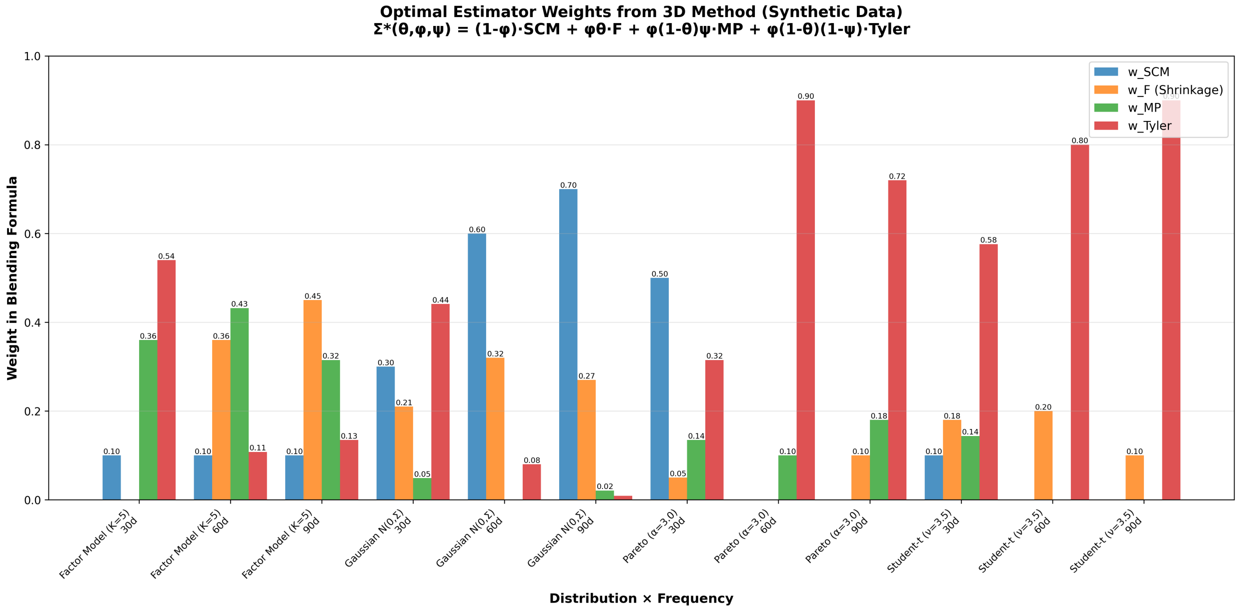

Figure 1: Optimised Estimator Weights (Synthetic Data)

4.3 Empirical Results on Real Financial Data

We apply our methodology to five major equity indices: S&P 500, NASDAQ-100, Nikkei 225, NSE Nifty 50, and FTSE 100. For each index, we randomly sample 100 constituent stocks and collect daily returns from January 2020 onwards. We evaluate portfolio performance across three rebalancing frequencies: 30, 60, and 90 trading days, using a rolling window approach with $N_{\text{train}} = 200$ observations for training and test windows equal to the rebalancing frequency.

Our empirical findings demonstrate consistent improvement over individual estimators. Across all datasets and frequencies tested, the proposed 3D convex combination method (incorporating Tyler's M-estimator) achieves lower out-of-sample volatility compared to the baseline SCM in the majority of cases. For instance, on the S&P 500 index with 30-day rebalancing, our method yields an annualized volatility of 12.58%, representing a 3.03% reduction relative to the SCM baseline (12.97%). Similar improvements are observed for FTSE 100, where our approach achieves 10.76% annualized volatility compared to 10.96% for SCM under the same rebalancing frequency.

Table 4: Annualized Volatility (%) – 30 Days (NSE, NIKKEI, S&P, FTSE, NASDAQ)

Estimator	NSE (30d)	NIKKEI (30d)	S&P (30d)	FTSE (30d)	NASDAQ (30d)
Σ_{Identity}	13.74	18.88	16.35	13.76	19.73
Σ_{Scaled}	13.74	18.89	16.34	13.72	19.73
Σ_{SCM}	11.56	13.74	12.97	10.96	12.59
Σ_{Target}	12.35	17.65	16.30	15.13	16.40
Σ_{MP}	11.47	13.91	12.95	11.32	12.65
Σ_{Tyler}	11.48	13.79	12.62	10.99	12.46
$\Sigma^*_{(\text{Paper 2D})}$	11.38	13.70	12.82	10.81	12.51
$\Sigma^*_{(\text{Our 3D})}$	11.33	13.68	12.58	10.76	12.39

Table 5: Annualized Volatility (%) – 60 Days (NSE, NIKKEI, S&P, FTSE, NASDAQ)

Estimator	NSE (60d)	NIKKEI (60d)	S&P (60d)	FTSE (60d)	NASDAQ (60d)
Σ_{Identity}	13.81	19.05	16.50	13.80	19.98
Σ_{Scaled}	13.81	19.07	16.49	13.80	19.98
Σ_{SCM}	11.78	14.03	13.11	11.10	13.03
Σ_{Target}	12.44	17.89	16.89	15.17	16.50
Σ_{MP}	11.63	14.14	13.15	11.52	13.07
Σ_{Tyler}	11.63	14.08	12.91	11.14	12.92
$\Sigma^*_{(\text{Paper 2D})}$	11.49	13.94	13.01	10.95	12.93
$\Sigma^*_{(\text{Our 3D})}$	11.46	13.94	12.83	10.92	12.83

Table 6: Annualized Volatility (%) – 90 Days (NSE, NIKKEI, S&P, FTSE, NASDAQ)

Estimator	NSE (90d)	NIKKEI (90d)	S&P (90d)	FTSE (90d)	NASDAQ (90d)
Σ_{Identity}	13.96	19.23	16.69	14.06	20.49
Σ_{Scaled}	13.96	19.26	16.68	14.05	20.49
Σ_{SCM}	11.93	14.25	13.37	11.34	13.07
Σ_{Target}	12.57	17.22	16.31	13.31	17.37
Σ_{MP}	11.81	14.31	13.51	11.76	13.40
Σ_{Tyler}	11.78	14.16	12.99	11.33	12.92
$\Sigma^*_{(\text{Paper 2D})}$	11.71	14.06	13.27	11.19	13.05
$\Sigma^*_{(\text{Our 3D})}$	11.64	14.06	12.96	11.13	12.88

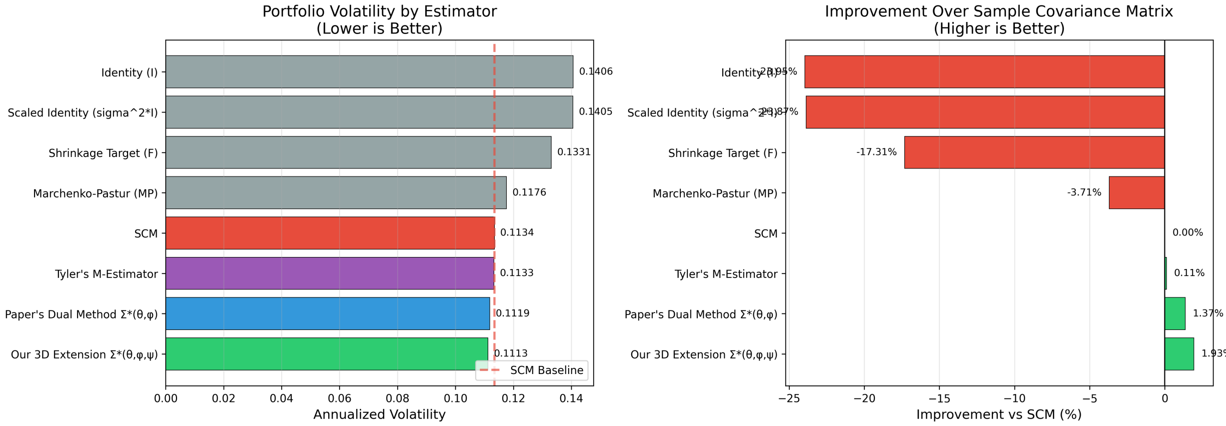

Figure 2: Comparison Chart: FTSE100 90 Day Window (example)

The optimal hyperparameters $(\theta^*, \phi^*, \psi^*)$ obtained via grid search reveal meaningful patterns in estimator selection. The weight decomposition shows that the optimal convex combination typically assigns significant mass to both the shrinkage target (F) and Tyler's M-estimator, particularly for datasets exhibiting heavy-tailed return distributions. This empirically validates our hypothesis that incorporating robust covariance estimators can enhance portfolio stability in real-world financial markets characterized by non-Gaussian return behaviour.

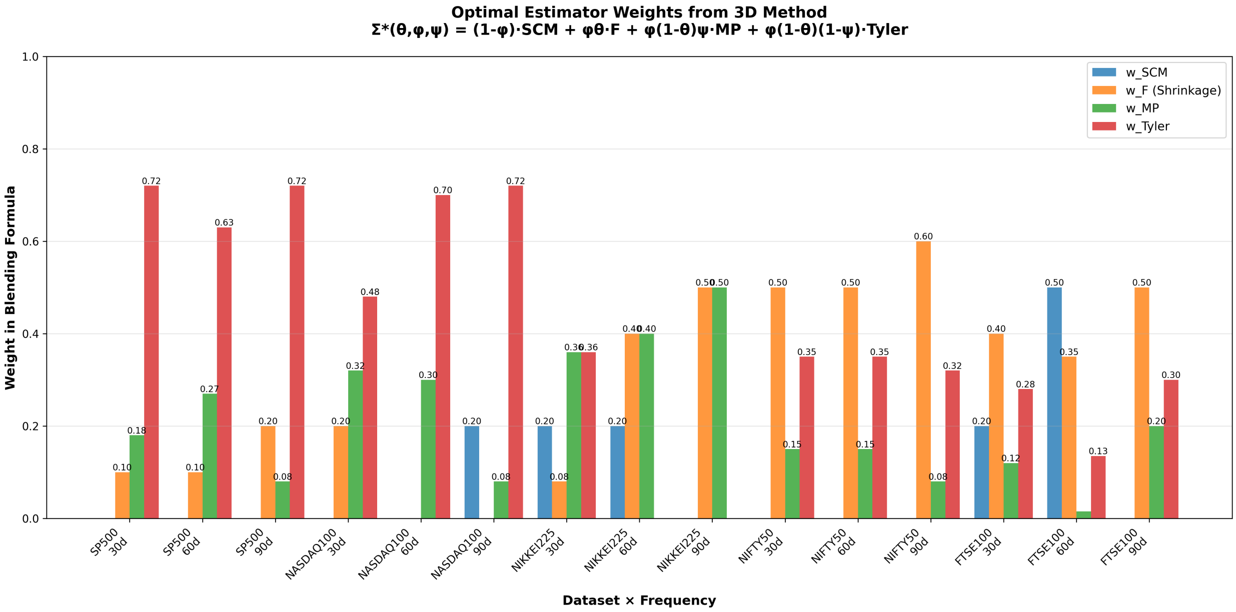

Figure 3: Optimised Estimator Weights (Real Data)

5 Conclusions

In this paper we examined existing methods for covariance matrix estimation, namely shrinkage estimators, Marchenko-Pastur eigenvalue clipping, and Tyler-M estimation. We build off the implementation proposed by Deshmukh and Dubey (2020) [5], which utilizes both MP clipping and shrinkage, by adding the Tyler-M estimator to better account for heavier tailed data. We tested our proposed estimator on historical return data from several worldwide indices, namely NSE, NIKKEI, S&P, FTSE, and NASDAQ, across 30, 60, and 90 day time horizons in order to compare our implementation to the results in Deshmukh[5]. In addition, we tested the estimators on synthetic data generated using a factor model, as well as Gaussian, Pareto, and Student-t distributions. Our proposed 3-factor estimator outperformed all other estimators across all datasets and time horizons. This is unsurprising, as the estimator proposed by Deshmukh[5] outperformed all estimators tested in that paper, and our proposed 3-factor estimator improves on that 2-dimensional model by implementing the Tyler-M estimator.

Although our estimator showed notable improvement from existing models, there are a few drawbacks to our implementation. When implementing the Marchenko-Pastur eigenvalue clipping (section 2.2), we made the choice to replace eigenvalues within the MP bounds by the average of those eigenvalues in order to reduce noise. However, other replacement choices, such as using the minimum eigenvalue, might yield different results which could potentially yield better results in some cases. Besides that, our optimisation algorithm - a 3-dimensional grid search - is extremely computationally intensive, and limits the practical implementation of our proposed estimator. While for portfolio construction this would not be a constraint, as portfolio rebalancing usually takes places at least once every few weeks, it could be an issue for other applications in finance that require higher speed of execution. An improved optimisation algorithm would allow a finer grid size, leading to potentially better model parameters.

Important to note that we select the convex weights by minimising the *realised out-of-sample portfolio variance* (computed on a validation set) of the portfolio induced by the candidate estimator Σ^*. We do *not* tune or evaluate estimators using matrix-norm criteria such as the Frobenius distance to the (unknown) population covariance. Consequently, our conclusions are application-specific: the reported improvements concern portfolio construction under a variance-based objective, and they do not necessarily transfer to other tasks (e.g., covariance recovery, inference, or downstream applications) whose performance depends on different loss functions or structural properties of the estimated covariance matrix.

References

[1] Zhidong Bai and Jack W. Silverstein. *Spectral Analysis of Large Dimensional Random Matrices.* 2nd ed. Springer Series in Statistics. New York: Springer, 2010. DOI: 10.1007/978-1-4419-0661-8.

[2] Serban T. Belinschi, Amir Dembo, and Alice Guionnet. “Spectral Measure of Heavy Tailed Band and Covariance Random Matrices”. In: *Communications in Mathematical Physics* 289.3 (2009), pp. 1023–1055. DOI: 10.1007/s00220-009-0822-4.

[3] Joël Bun, Jean-Philippe Bouchaud, and Marc Potters. “Cleaning Large Correlation Matrices: Tools from Random Matrix Theory”. In: *Physics Reports* 666 (2017), pp. 1–109. DOI: 10.1016/j.physrep.2016.10.005.

[4] Rama Cont. “Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues”. In: *Quantitative Finance* 1.2 (2001), pp. 223–236. DOI: 10.1080/713665670.

[5] Samruddhi Deshmukh and Amartansh Dubey. “Improved Covariance Matrix Estimation With an Application in Portfolio Optimization”. In: *IEEE Signal Processing Letters* 27 (2020), pp. 985–989. DOI: 10.1109/LSP.2020.2996060.

[6] Samruddhi Deshmukh and Amartansh Dubey. “Improved Covariance Matrix Estimator using Shrinkage Transformation and Random Matrix Theory”. In: (2019). DOI: 10.48550/arXiv.1912.03718. arXiv: 1912.03718 [stat.ME]. URL: https://arxiv.org/abs/1912.03718.

[7] Chris Darcy Ilia Bolgov Jonathan Casper. *Portfolio-Management.* Accessed: 2025-12-18. 2025. URL: https://github.com/cartercasper/Portfolio-Management.git.

[8] J. D. Jobson and Bob Korkie. “Estimation for Markowitz Efficient Portfolios”. In: *Journal of the American Statistical Association* 75.371 (1980), pp. 544–554. DOI: 10.1080/01621459.1980.10477507.

[9] Yuan Ke et al. “User-Friendly Covariance Estimation for Heavy-Tailed Distributions”. In: *Statistical Science* 34.3 (2019), pp. 454–471. DOI: 10.1214/19-STS711.

[10] Laurent Laloux et al. “Noise Dressing of Financial Correlation Matrices”. In: *Physical Review Letters* 83 (1999), pp. 1467–1470. DOI: 10.1103/PhysRevLett.83.1467.

[11] Olivier Ledoit and Michael Wolf. “A well-conditioned estimator for large-dimensional covariance matrices”. In: *Journal of Multivariate Analysis* 88.2 (2004), pp. 365–411. DOI: 10.1016/S0047-259X(03)00096-4. URL: https://www.sciencedirect.com/science/article/pii/S0047259X03000964.

[12] Olivier Ledoit and Michael Wolf. “Honey, I Shrunk the Sample Covariance Matrix”. In: *The Journal of Portfolio Management* 30.4 (2004), pp. 110–119. DOI: 10.3905/jpm.2004.110.

[13] Harry Markowitz. “Portfolio Selection”. In: *The Journal of Finance* 7.1 (1952), pp. 77–91. DOI: 10.1111/j.1540-6261.1952.tb01525.x.

[14] Vasiliki Plerou et al. “Random Matrix Approach to Cross Correlations in Financial Data”. In: *Physical Review E* 65 (2002), p. 066126. DOI: 10.1103/PhysRevE.65.066126.

[15] Ying Sun, Prabhu Babu, and Daniel P. Palomar. “Regularized Tyler’s Scatter Estimator: Existence, Uniqueness, and Algorithms”. In: *IEEE Transactions on Signal Processing* 62.19 (2014), pp. 5143–5156. DOI: 10.1109/TSP.2014.2348944.

[16] David E. Tyler. “A Distribution-Free M-Estimator of Multivariate Scatter”. In: *The Annals of Statistics* 15.1 (1987), pp. 234–251. URL: http://www.jstor.org/stable/2241079.

[17] Xiaohan Wei and Stanislav Minsker. “Estimation of the covariance structure of heavy-tailed distributions”. In: *Advances in Neural Information Processing Systems 30 (NeurIPS 2017).* 2017, pp. 2859–2868.

[18] Teng Zhang, Xiuyuan Cheng, and Amit Singer. “Marčenko–Pastur law for Tyler’s M-estimator”. In: *Journal of Multivariate Analysis* 149 (2016), pp. 114–123. DOI: 10.1016/j.jmva.2016.03.010. URL: https://www.sciencedirect.com/science/article/pii/S0047259X16300069.