ÖREBRO UNIVERSITY
SCHOOL OF BUSINESS

WORKING PAPER

1/2018

Tangency portfolio weights for singular covariance matrix in small and large dimensions: estimation and test theory

Taras Bodnar, Stepan Mazur, Krzysztof Podgórski,

Joanna Tyrcha
Statistics

ISSN 1403-0586

http://www.oru.se/Institutioner/Handelshogskolan-vid-Orebro-universitet/Forskning/Publikationer/Working-papers/
Örebro University School of Business
701 82 Örebro
SWEDEN

Tangency portfolio weights for singular covariance matrix in small and large dimensions: estimation and test theory

Taras Bodnar[a], Stepan Mazur[b,1], Krzysztof Podgórski[c], Joanna Tyrcha[a]

[a] *Department of Mathematics, Stockholm University, SE-10691 Stockholm, Sweden*

[b] *Department of Statistics, Örebro University School of Business, SE-70182 Örebro, Sweden*

[c] *Department of Statistics, Lund University, SE-22007 Lund, Sweden*

Abstract

In this paper we derive the finite-sample distribution of the estimated weights of the tangency portfolio when both the population and the sample covariance matrices are singular. These results are used in the derivation of a statistical test on the weights of the tangency portfolio where the distribution of the test statistic is obtained under both the null and the alternative hypotheses. Moreover, we establish the high-dimensional asymptotic distribution of the estimated weights of the tangency portfolio when both the portfolio dimension and the sample size increase to infinity. The theoretical findings are implemented in an empirical application dealing with the returns on the stocks included into the S&P 500 index.

JEL Classification: C10, C44

ASM Classification: 62H10, 62H12, 91G10

Keywords: tangency portfolio; singular Wishart distribution; singular covariance matrix; high-dimensional asymptotics; hypothesis testing

[1]* Corresponding author. E-mail address: stepan.mazur@oru.se. Stepan Mazur and Krzysztof Podgórski acknowledge financial support from the Swedish Research Council Grant 2008-5382 and Riksbankens Jubileumsfond Dnr: P13-1024:1.

1 Introduction

The fundamental goal of portfolio theory is to optimally allocate investments between different assets. The mean-variance optimization is a quantitative tool which allows the making of this allocation through considering the trade-off between the risk of portfolio and its return.

The foundations of modern portfolio theory have been developed in Markowitz [11], where a mean-variance portfolio optimization procedure was introduced. In this, investors incorporate their preferences towards the risk and the expectation of return to seek the 'best' allocation of wealth. This means the portfolios are selected to maximize anticipated profit subject to achieving a specified level of risk or, equivalently, to minimize variance subject to achieving a specified level of expected gain. The risk aversion strategy in the absence of risk free assets (bonds) leads to the minimal variance portfolio. This has to be changed in the presence of bonds and the tangency portfolio is a component for a portfolio that hedges stocks with bond investment or, for higher returns levels, borrows to invest in the stocks. Nowadays, the mean-variance analysis of Markowitz remains an important though basic tool for both practitioners and researchers in the financial sector. Therefore, having a complete understanding of the tangency portfolio properties under all realistic conditions is of great importance for any financial strategist.

To implement these optimal portfolios in practice, the inverse of covariance matrix of asset returns need to be estimated. Traditionally, the sample covariance matrix has been used for this purpose under the assumption of non-singular true (population) covariance matrix. However, the problem of potential multicollinearity and strong correlations of asset returns results in clear limitations in taking such an approach due to latent singularity or near singularity of the population covariance. In addition, the number of assets in a portfolio can be larger than the number of observations and create another cause for singularity. There is a wide literature concerning cases with very high-dimensional datasets and a small sample size relative to dimension, see Pruzek [15] and Stein et al. [16] among many others. However, consequences of this particular problem for the portfolio theory have not been addressed until recently. Singular population covariance matrix and small sample size relative to the portfolio size was first discussed in Bodnar et al. [1] and later results were extended in Bodnar et al. [2]. In the last paper, the authors analyzed the *global minimum variance portfolio* for small sample and singular covariance matrix. Here we tackle properties of the *tangency portfolio* (TP) under these two singularity conditions. We derive a stochastic representation of the tangency portfolio weights estimator as well as a linear test for the portfolio weights. We also establish the asymptotic distribution of the estimated portfolio weights under a high-dimensional asymptotic regime. Theoretical results are applied to an illustrative example based on 440 actual weekly stock returns.

We conclude this introduction with pointing out importance and some difficulties in practical handling of the singularity and small samples size problem. Firstly, the singularity of the data typically cannot be observed due to observational or computational noise. As far as the authors are aware, a detection problem or formal statistical tests for singularity have not been thoroughly studied in the literature and deserve a separate treatment. Secondly, the issue of non-normality and time dependence in the singular data will affect the accuracy of the results based on the normality assumption. Unfortunately, we are not aware of any statistical test for normality under the singularity assumption and developing a methodology for such a test deserves a separated treatment. Here, we heuristically assume that averaging over time blocks that are long enough brings us to independence and normality of the data. We estimate the rank of the covariance matrix using a test developed by Nadakuditi and Eldeman [13], where the singularity problem was discussed from the signal processes perspective. We use the method to verify that the rank of the covariance matrix is less than the portfolio size.

The rest of the paper is structured as follows. In the next section, we discuss the problem of the estimation of the TP weights and derive a very useful stochastic representation for the estimated weights of this portfolio which fully characterizes their distribution. This stochastic representation is used to establish the distributional properties of the estimated TP weights, including their mean vector and covariance matrix in Section 2.2 and in the derivation of a statistical test on the TP weights in Section 2.3. In Section 3, the high-dimensional asymptotic distribution of the estimated weights is present and a high-dimensional asymptotic test on the TP weights is derived. The theoretical findings are applied to real data in Section 4. Several supplementary theoretical results are formulated and proved in the Appendix.

2 Distributional properties of the estimated TP weights

Let $\mathbf{x}_i = (x_{1i}, ..., x_{ki})^T$ denote the k-dimensional vector of log-returns of some k assets at time $i = 1, ..., n$ and $\mathbf{w} = (w_1, ..., w_k)^T$ be the vector of weights, i.e., the parts of the investor wealth invested into those assets. In this section, we assume that $\mathbf{x}_1, ..., \mathbf{x}_n$ are independently and identically distributed with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ which is a non-negative definite matrix with

$rank(\mathbf{\Sigma}) = r < n$.

Following the mean-variance analysis suggested in the seminal paper of Markowitz [11], the optimal portfolio in the case of absence a risk-free asset is obtained by minimizing the portfolio variance for a given level of the expected return under the constraint $\mathbf{w}^T \mathbf{1}_k = 1$ where $\mathbf{1}_k$ denotes the vector of ones. The set of all optimal portfolio following Markowitz's approach constitutes the efficient frontier. If a risk-free asset is available to an investor, then part of the investor's wealth can be invested into the risk-free asset and may reduce the variance, whereas the rest of the wealth is invested into the tangency portfolio located on the efficient frontier. This portfolio is obtained by drawing a tangency line from the location of the risk-free asset (bonds) in the mean-variance space to the efficient frontier. Then any portfolio from this tangent line is efficient for the given return level and is obtained by a combination of the tangent portfolio and the investment (borrowing) in bonds is efficient. In general, the weights of the portfolio are given by

$$\mathbf{w}_{TP} = \alpha^{-1} \mathbf{\Sigma}^{-1} (\boldsymbol{\mu} - r_f \mathbf{1}_k) \tag{1}$$

if the covariance matrix matrix is positive definite where r_f denotes the return of the risk-free asset and the coefficient α describes the investor's attitude towards risk. Changing $\alpha \in (0, \infty)$, we obtain all portfolios from the tangent line. It is also noted that the vector $\mathbf{w}_{TP}$ in (1) determines only the structure of the portfolio which corresponds to risky asset only, whereas $1 - \mathbf{w}_{TP}^T \mathbf{1}_k$ is the part of the wealth which is invested into the risk-free asset.

If $\mathbf{\Sigma}$ is singular, then the weights of the tangency portfolio are calculated by replacing the inverse with the Moore-Penrose inverse. This leads to

$$\mathbf{w}_{TP} = \alpha^{-1} \mathbf{\Sigma}^{+} (\boldsymbol{\mu} - r_f \mathbf{1}_k) \,. \tag{2}$$

It is noted that the Moore-Penrose inverse has already been applied in the portfolio theory. For instance, Pappas et al. [14] and Bodnar et al. [1] used the Moore-Penrose inverse in the estimation of the mean-variance optimal portfolio, whereas Bodnar et al. [2] considered the application of the Moore-Penrose inverse in the estimation of the weights of the global minimum variance portfolio. Moreover, Pappas et al. [14] showed that the estimator of the mean-variance portfolio weights which is based on the Moore-Penrose inverse possesses the minimal Euclidean norm and is an unique solution of the corresponding optimization problem. Similar results also hold in the case of $\mathbf{w}_{TP}$ following Theorem 2.2 of Pappas et al. [14].

The optimal weights $\mathbf{w}_{TP}$ depend on the unknown parameters $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$. In order to estimate them we use the corresponding sample estimators expressed as

$$\overline{\mathbf{x}} = \frac{1}{n}\sum_{j=1}^{n}\mathbf{x}_j \quad \text{and} \quad \mathbf{S} = \frac{1}{n-1}\sum_{j=1}^{n}(\mathbf{x}_j - \overline{\mathbf{x}})(\mathbf{x}_j - \overline{\mathbf{x}})^T.$$

Replacing $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ with $\overline{\mathbf{x}}$ and $\mathbf{S}$ in (2), we get the sample estimator of optimal tangency portfolio weights which is given by

$$\hat{\mathbf{w}}_{TP} = \alpha^{-1}\mathbf{S}^{+}(\overline{\mathbf{x}} - r_f\mathbf{1}_k) \tag{3}$$

The distribution of $\hat{\mathbf{w}}_{TP}$ appears to be singular since the covariance matrix of $\hat{\mathbf{w}}_{TP}$ is singular (see Theorem 3 below). For that reason, we consider linear combinations of $\hat{\mathbf{w}}_{TP}$ and derive their finite-sample distribution. Since the characteristic function of a random vector $\mathbf{Y}$ given by $\phi_{\mathbf{Y}}(\mathbf{t}) = E(\exp(i\mathbf{t}^T\mathbf{Y}))$ determines uniquely the distribution of $\mathbf{Y}$, a stochastic representation obtained for an arbitrary linear combination of $\hat{\mathbf{w}}_{TP}$ can be used to characterize the distribution of $\hat{\mathbf{w}}_{TP}$.

We consider a more general case by deriving the joint distribution of several linear transformations of $\mathbf{w}_{TP}$. Let

$$\boldsymbol{\theta} = \mathbf{L}\mathbf{w}_{TP} = \alpha^{-1}\mathbf{L}\boldsymbol{\Sigma}^{+}(\boldsymbol{\mu} - r_f\mathbf{1}_k),$$

where $\mathbf{L}$ is a non-random $p \times k$ matrix of $rank\ p < r$ such that $rank(\mathbf{L}\boldsymbol{\Sigma}) = p$. The latter condition ensures that the distribution of the sample estimator of $\boldsymbol{\theta}$ expressed as

$$\hat{\boldsymbol{\theta}} = \mathbf{L}\hat{\mathbf{w}}_{TP} = \alpha^{-1}\mathbf{L}\mathbf{S}^{+}(\overline{\mathbf{x}} - r_f\mathbf{1}_k) \tag{4}$$

is non-singular.

2.1 Stochastic representation

It is convenient to represent the distribution of $\hat{\boldsymbol{\theta}} = \mathbf{L}\hat{\mathbf{w}}_{TP}$ in an explicit stochastic form. From Theorem 4 of Bodnar et al. [1] we have that $\overline{\mathbf{x}} \sim \mathcal{N}_k(\boldsymbol{\mu}, 1/n\boldsymbol{\Sigma})$, $\mathbf{V} = (n-1)\mathbf{S} \sim \mathcal{W}_k(n-1, \boldsymbol{\Sigma})$, and $\overline{\mathbf{x}}$, $(n-1)\mathbf{S}$ are independent. Then

$$\hat{\boldsymbol{\theta}} = \mathbf{L}\hat{\mathbf{w}}_{TP} = \alpha^{-1}\mathbf{L}\mathbf{S}^{+}(\overline{\mathbf{x}} - r_f\mathbf{1}_k) = \frac{n-1}{\alpha}\mathbf{L}\mathbf{V}^{+}(\overline{\mathbf{x}} - r_f\mathbf{1}_k).$$

From that the stochastic representation presented in the following result which is a direct consequence of Lemma 1 and Corollary 1 both given in the Appendix.

Theorem 1. *Let* $\mathbf{x}_1, ..., \mathbf{x}_n$ *be i.i.d. random vectors with* $\mathbf{x}_1 \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $k > n-1$, *and let* $rank(\boldsymbol{\Sigma}) = r \le n-1$. *Consider* $\mathbf{L}$ *a* $p \times k$ *non-random matrix of rank* $p < r$ *such that* $rank(\mathbf{L}\boldsymbol{\Sigma}) = p$. *Additionally, let* $\mathbf{S}_1 = (\mathbf{L}\boldsymbol{\Sigma}^{+}\mathbf{L}^T)^{-1/2}\mathbf{L}\boldsymbol{\Sigma}^{+1/2}$ *and* $\mathbf{Q}_1 = \mathbf{S}_1^T\mathbf{S}_1$. *Then the stochastic representation of* $\widehat{\boldsymbol{\theta}}$ *is given by*

$$\hat{\boldsymbol{\theta}} \stackrel{d}{=} \frac{n-1}{\alpha}\xi^{-1}\left(\mathbf{L}\boldsymbol{\Sigma}^{+1/2}\bar{\mathbf{z}} + \sqrt{\frac{1}{n-r+1}}(\mathbf{L}\boldsymbol{\Sigma}^{+}\mathbf{L}^T)^{1/2}\right.$$
$$\left.\times\left[\sqrt{\bar{\mathbf{z}}^T\bar{\mathbf{z}}}\mathbf{I}_p - \frac{\sqrt{\bar{\mathbf{z}}^T\bar{\mathbf{z}}} - \sqrt{\bar{\mathbf{z}}^T(\mathbf{I}_p - \mathbf{Q}_1)\bar{\mathbf{z}}}}{\bar{\mathbf{z}}^T\mathbf{Q}_1\bar{\mathbf{z}}}\mathbf{S}_1\bar{\mathbf{z}}\bar{\mathbf{z}}^T\mathbf{S}_1^T\right]\mathbf{t}_0\right),$$

where $\xi \sim \chi^2_{n-r}$, $\bar{\mathbf{z}} \sim \mathcal{N}_k\left(\boldsymbol{\Sigma}^{+1/2}(\boldsymbol{\mu} - r_f\mathbf{1}_k), \frac{1}{n}\mathbf{P}\mathbf{P}^T\right)$, *and* $\mathbf{t}_0 \sim t_p(n-r+1; \mathbf{0}, \mathbf{I}_p)$. *Moreover,* ξ, $\bar{\mathbf{z}}$, *and* $\mathbf{t}_0$ *are mutually independent. The* $k \times r$ *matrix* $\mathbf{P}$ *is the semi-orthogonal matrix of the eigenvectors of* $\boldsymbol{\Sigma}$ *such that* $\mathbf{P}^T\mathbf{P} = \mathbf{I}_r$.

In particular, the special case of $p = 1$ *and* $\mathbf{L} = \mathbf{l}^T$ *is given by*

$$\hat{\theta} \stackrel{d}{=} \frac{n-1}{\alpha}\xi^{-1}\left(\mathbf{l}^T\boldsymbol{\Sigma}^{+}(\boldsymbol{\mu} - r_f\mathbf{1}_k) + \sqrt{\left(\frac{1}{n} + \frac{r-1}{n(n-r+1)}u\right)\mathbf{l}^T\boldsymbol{\Sigma}^{+}\mathbf{l}}\,z_0\right),$$

where $\xi \sim \chi^2_{n-r}$, $z_0 \sim \mathcal{N}(0,1)$, *and* $u \sim \mathcal{F}(r-1, n-r+1, n(\boldsymbol{\mu} - r_f\mathbf{1}_k)^T\mathbf{R}_\mathbf{l}(\boldsymbol{\mu} - r_f\mathbf{1}_k))$ *with* $\mathbf{R}_\mathbf{l} = \boldsymbol{\Sigma}^{+} - \boldsymbol{\Sigma}^{+}\mathbf{l}\mathbf{l}^T\boldsymbol{\Sigma}^{+}/\mathbf{l}^T\boldsymbol{\Sigma}^{+}\mathbf{l}$; ξ, u, *and* z_0 *are mutually independently distributed.*

2.2 Mean vector and covariance matrix

In the next theorem, we derive the first two moments of the estimator of the TP weights by applying the stochastic representation of Theorem 1.

Theorem 2. *Let* $\mathbf{x}_1, ..., \mathbf{x}_n$ *be i.i.d. random vectors with* $\mathbf{x}_1 \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $k > n-1$ *and let* $rank(\boldsymbol{\Sigma}) = r \le n-1$. *Then it holds that*

$$E(\hat{\mathbf{w}}_{TP}) = \frac{n-1}{n-r-2}\mathbf{w}_{TP},$$

and

$$Var(\hat{\mathbf{w}}_{TP}) = c_1\mathbf{w}_{TP}\mathbf{w}_{TP}^T + c_2\boldsymbol{\Sigma}^{+}$$

with

$$c_1 = \frac{(n-r)(n-1)^2}{(n-r-1)(n-r-2)^2(n-r-4)}$$
$$c_2 = \frac{(n-1)^2(n-2+n(\boldsymbol{\mu}-r_f\mathbf{1}_k)^T\boldsymbol{\Sigma}^+(\boldsymbol{\mu}-r_f\mathbf{1}_k))}{n(n-r-1)(n-r-2)(n-r-4)\alpha^2}.$$

Proof. First, we derive the expected value of $\hat{\theta}$. From Theorem 1, we get for an arbitrary $\mathbf{l}$ that

$$E(\hat{\theta}) = \frac{n-1}{\alpha}E(\xi^{-1}) \times \times \left[\mathbf{l}^T\boldsymbol{\Sigma}^+(\boldsymbol{\mu}-r_f\mathbf{1}_k) + E\left(\sqrt{\left(\frac{1}{n}+\frac{r-1}{n(n-r-1)}u\right)\mathbf{l}^T\boldsymbol{\Sigma}^+\mathbf{l}}\right)\cdot E(z_0)\right].$$

Since $E(z_0)=0$ and $E(\xi^{-1}) = \frac{1}{n-r-2}$, we obtain that

$$E(\hat{\theta}) = \frac{n-1}{\alpha}\frac{1}{n-r-2}\mathbf{l}^T\boldsymbol{\Sigma}^+(\boldsymbol{\mu}-r_f\mathbf{1}_k) = \frac{n-1}{n-r-2}\theta.$$

Next, we derive the variance of $\hat{\theta}$. It holds that

$$\begin{aligned} Var(\hat{\theta}) &= E(\hat{\theta}^2) - [E(\hat{\theta})]^2 \\ &= \left(\frac{n-1}{\alpha}\right)^2 E(\xi^{-2})\left[(\mathbf{l}^T\boldsymbol{\Sigma}^+(\boldsymbol{\mu}-r_f\mathbf{1}_k))^2 + E\left(\frac{1}{n}+\frac{r-1}{n(n-r+1)}u\right)\right. \\ &\quad \left.\times \mathbf{l}^T\boldsymbol{\Sigma}^+\mathbf{l}\cdot E(z_0^2)\right] - \left(\frac{n-1}{n-r-2}\right)^2\theta^2 \\ &= c_1\theta^2 + c_2\mathbf{l}^T\boldsymbol{\Sigma}^+\mathbf{l} \end{aligned}$$

with c_1 and c_2 as in the formulation of the theorem. The last result is obtained using the facts that $E(z_0^2)=1$, $E(\xi^{-2}) = \frac{1}{(n-r-2)(n-r-4)}$, and $E(u) = \frac{(n-r+1)(r-1+\tilde{\lambda})}{(r-1)(n-r-1)}$ with $\tilde{\lambda} = n(\boldsymbol{\mu}-r_f\mathbf{1}_k)^T\mathbf{R}_\mathbf{l}(\boldsymbol{\mu}-r_f\mathbf{1}_k))$.

Finally, using the fact that $\mathbf{l}$ was an arbitrary vector we get that

$$E(\hat{\mathbf{w}}_{TP}) = \frac{n-1}{n-r-2}\mathbf{w}_{TP}$$

and

$$Var(\hat{\mathbf{w}}_{TP}) = c_1\mathbf{w}_{TP}\mathbf{w}_{TP}^T + c_2\boldsymbol{\Sigma}^+.$$

The theorem is proved. □

2.3 Inference procedures

We propose statistical testing procedures for the optimal weights of the tangency portfolio. The hypotheses of the test are given by

$$H_0 : \mathbf{l}^T \mathbf{w}_{TP} = 0 \quad \text{against} \quad H_1 : \mathbf{l}^T \mathbf{w}_{TP} = \rho_1 \neq 0 \tag{5}$$

with the following test statistics which extends the one introduced by Bodnar and Okhrin [4] to the case of singular covariance matrix

$$T = \sqrt{\frac{n-r}{n-1}} \frac{\alpha \mathbf{l}^T \hat{\mathbf{w}}_{TP}}{\sqrt{\mathbf{l}^T \mathbf{S}^+ \mathbf{l}} \sqrt{\frac{1}{n} + \frac{1}{n-1} \overline{\mathbf{y}}^T \hat{\mathbf{R}}_{\mathbf{l}} \overline{\mathbf{y}}}} \tag{6}$$

where $\hat{\mathbf{R}}_{\mathbf{l}} = \mathbf{S}^+ - \mathbf{S}^+ \mathbf{l}\mathbf{l}^T \mathbf{S}^+ / \mathbf{l}^T \mathbf{S}^+ \mathbf{l}$ and $\overline{\mathbf{y}} = \overline{\mathbf{x}} - r_f \mathbf{1}_k \sim \mathcal{N}_k \left(\boldsymbol{\mu} - r_f \mathbf{1}_k, \frac{1}{n} \boldsymbol{\Sigma} \right)$.

The next theorem provides the distribution of T both under the null and under the alternative hypotheses.

Theorem 3. *Let $\mathbf{x}_1, ..., \mathbf{x}_n$ be i.i.d. random vectors with $\mathbf{x}_1 \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $k > n-1$ and with $rank(\boldsymbol{\Sigma}) = r \leq n-1$. Also, let $\mathbf{l}$ be a k-dimensional vector of constants. Then*

(a) the density function of T is given by

$$f_T(x) = \frac{n(n-r+1)}{(r-1)(n-1)} \int_0^\infty f_{t_{n-r,\vartheta(y)}}(x) f_{\mathcal{F}_{r-1,n-r+1,ns}} \left(\frac{n(n-r+1)}{(r-1)(n-1)} y \right) \mathbf{d}y,$$

where $\vartheta(y) = \frac{\alpha \rho_1}{\sqrt{\mathbf{l}^T \boldsymbol{\Sigma}^+ \mathbf{l} \left(\frac{1}{n} + \frac{1}{n-1} y \right)}}$ and $s = (\boldsymbol{\mu} - r_f \mathbf{1}_k)^T \mathbf{R}_{\mathbf{l}} (\boldsymbol{\mu} - r_f \mathbf{1}_k)$ with $\mathbf{R}_{\mathbf{l}} = \boldsymbol{\Sigma}^+ - \boldsymbol{\Sigma}^+ \mathbf{l}\mathbf{l}^T \boldsymbol{\Sigma}^+ / \mathbf{l}^T \boldsymbol{\Sigma}^+ \mathbf{l}$; the symbol $t_{n-r,\vartheta(y)}$ denotes a non-central t-distribution with $n-r$ degrees of freedom and non-centrality parameter $\vartheta(y)$.

(b) under the null hypothesis it holds that $T \sim t_{n-r}$.

Proof. Because $\overline{\mathbf{y}}$ and $\mathbf{S}$ are independently distributed (c.f., Theorem 4 in Bodnar et al. [1]), it follows that the conditional distribution of $\mathbf{l}^T \mathbf{S}^+ \overline{\mathbf{y}}$ given $\overline{\mathbf{y}} = \overline{\mathbf{y}}^*$ is equal to the distribution of $\mathbf{l}^T \mathbf{S}^+ \overline{\mathbf{y}}^*$. Let $\widetilde{\mathbf{L}} = (\mathbf{l}, \overline{\mathbf{y}}^*)^T$ and define

$$\widetilde{\mathbf{S}} = \widetilde{\mathbf{L}}((n-1)\mathbf{S})^+ \widetilde{\mathbf{L}}^T = (n-1)^{-1} \begin{pmatrix} \mathbf{l}^T \mathbf{S}^+ \overline{\mathbf{y}}^* & \mathbf{l}^T \mathbf{S}^+ \overline{\mathbf{y}}^* \\ \overline{\mathbf{y}}^{*T} \mathbf{S}^+ \mathbf{l} & \overline{\mathbf{y}}^{*T} \mathbf{S}^+ \overline{\mathbf{y}}^* \end{pmatrix}$$

and

$$\widetilde{\boldsymbol{\Sigma}} = \widetilde{\mathbf{L}} \mathbf{S}^+ \widetilde{\mathbf{L}}^T = \begin{pmatrix} \mathbf{l}^T \boldsymbol{\Sigma}^+ \overline{\mathbf{y}}^* & \mathbf{l}^T \boldsymbol{\Sigma}^+ \overline{\mathbf{y}}^* \\ \overline{\mathbf{y}}^{*T} \boldsymbol{\Sigma}^+ \mathbf{l} & \overline{\mathbf{y}}^{*T} \boldsymbol{\Sigma}^+ \overline{\mathbf{y}}^* \end{pmatrix}.$$

From Theorem 1 of Bodnar et al. [1] and Theorem 3.4.1 of Gupta and Nagar [10] we get that $\widetilde{\mathbf{S}} \sim \mathcal{IW}(n-r+4, \widetilde{\mathbf{\Sigma}})$. Then the application of Theorem 3 in Bodnar and Okhrin [3] leads to

$$\frac{\mathbf{l}^T\mathbf{S}^+\overline{\mathbf{y}}^*}{\mathbf{l}^T\mathbf{S}^+\mathbf{l}}\Big|\overline{\mathbf{y}}^{*T}\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}}^* \sim \mathcal{N}\left(\frac{\mathbf{l}^T\mathbf{\Sigma}^+\overline{\mathbf{y}}^*}{\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}}, (n-1)^{-1}\overline{\mathbf{y}}^{*T}\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}}^*(\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l})^{-1}\right), \tag{7}$$

$$(n-1)^{-1}\overline{\mathbf{y}}^{*T}\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}}^* \sim \mathcal{IW}(n-r+3, \overline{\mathbf{y}}^{*T}\mathbf{R}_\mathbf{l}\overline{\mathbf{y}}^*), \tag{8}$$

and $\mathbf{l}^T\mathbf{S}^+\mathbf{l}$ is independent of $\dfrac{\mathbf{l}^T\mathbf{S}^+\overline{\mathbf{y}}^*}{\mathbf{l}^T\mathbf{S}^+\mathbf{l}}$ and $\overline{\mathbf{y}}^{*T}\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}}^*$. Since $\mathbf{l}^T\mathbf{S}^+\mathbf{l}$ does not depend on $\overline{\mathbf{y}}$, we also obtain that $\mathbf{l}^T\mathbf{S}^+\mathbf{l}$ is independent of $\dfrac{\mathbf{l}^T\mathbf{S}^+\overline{\mathbf{y}}}{\mathbf{l}^T\mathbf{S}^+\mathbf{l}}$ and $\overline{\mathbf{y}}^T\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}}$.

Using (7) we get the following stochastic representation

$$\frac{\mathbf{l}^T\mathbf{S}^+\overline{\mathbf{y}}}{\mathbf{l}^T\mathbf{S}^+\mathbf{l}} \stackrel{d}{=} \frac{\mathbf{l}^T\mathbf{\Sigma}^+\overline{\mathbf{y}}}{\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}} + \sqrt{\frac{(n-1)^{-1}\overline{\mathbf{y}}^T\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}}}{\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}}}z_0,$$

where $z_0 \sim \mathcal{N}(0,1)$ which is independent of $\mathbf{l}^T\mathbf{\Sigma}^+\overline{\mathbf{y}}$ and $\overline{\mathbf{y}}^T\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}}$.

From the proof of Corollary 1 given in the Appendix we obtain $\mathbf{l}^T\mathbf{\Sigma}^+\overline{\mathbf{y}} \sim \mathcal{N}(\mathbf{l}^T\mathbf{\Sigma}^+\overline{\mathbf{y}}, \mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}/n)$,

$$\frac{n(n-r+1)}{(n-1)(r-1)}\overline{\mathbf{y}}^T\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}} \sim \mathcal{F}_{r-1,n-r+1,ns}$$

with $s = (\boldsymbol{\mu} - r_f\mathbf{1}_k)^T\mathbf{R}_\mathbf{l}(\boldsymbol{\mu} - r_f\mathbf{1}_k)$, and $\mathbf{l}^T\mathbf{\Sigma}^+\overline{\mathbf{y}}$, $\overline{\mathbf{y}}^T\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}}$ are independent. Consequently,

$$\frac{\mathbf{l}^T\mathbf{S}^+\overline{\mathbf{y}}}{\mathbf{l}^T\mathbf{S}^+\mathbf{l}}|\overline{\mathbf{y}}^T\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}} \sim \mathcal{N}\left(\frac{\mathbf{l}^T\mathbf{\Sigma}^+\boldsymbol{\mu}}{\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}}, \left(\frac{1}{n} + \frac{\overline{\mathbf{y}}^T\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}}}{n-1}\right)(\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l})^{-1}\right).$$

Hence,

$$\left(\frac{1}{n} + \frac{\overline{\mathbf{y}}^T\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}}}{n-1}\right)^{-1/2}\frac{\mathbf{l}^T\mathbf{S}^+\overline{\mathbf{y}}}{\mathbf{l}^T\mathbf{S}^+\mathbf{l}}|\overline{\mathbf{y}}^T\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}} \sim \mathcal{N}\left(\left(\frac{1}{n} + \frac{\overline{\mathbf{y}}^T\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}}}{n-1}\right)^{-1/2}\frac{\mathbf{l}^T\mathbf{\Sigma}^+\boldsymbol{\mu}}{\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}}, (\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l})^{-1}\right).$$

Finally, the application of $\dfrac{\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}}{(n-1)^{-1}\mathbf{l}^T\mathbf{S}^+\mathbf{l}} \sim \chi^2_{n-r}$ and the independence of $\mathbf{l}^T\mathbf{S}^+\mathbf{l}$ and $(\mathbf{l}^T\mathbf{S}^+\overline{\mathbf{y}}, \overline{\mathbf{y}}^T\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}})$ leads to

$$T|\overline{\mathbf{y}}^T\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}} = y \sim t_{n-r,\vartheta(y)}$$

with $\vartheta(y) = \frac{\alpha\rho_1}{\sqrt{\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}\left(\frac{1}{n}+\frac{1}{n-1}y\right)}}$. Now, the result of Theorem 3.(a) follows from the fact that $\frac{n(n-r+1)}{(n-1)(r-1)}\overline{\mathbf{y}}^T\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}} \sim \mathcal{F}_{r-1,n-r+1,ns}$, whereas the statement of Theorem 3.(b) is obtained by noting that $\vartheta(y) = 0$ under H_0. □

In the statistical test (5), we check if a linear combination of the TP weights is equal to zero under the null hypothesis. This procedure, in particular, can be used to test if the TP weight of a selected asset is equal to zero or deviates considerably from zero. On the other hand, it cannot be applied if the investor wants to verify if the TP weight of a selected asset is equal to a given value, for example, when the aim is to decide whether to keep the current value of the weight or to reallocate the wealth into other assets. The situation is more complicated in the latter case and it is discussed next. We consider the following test problem

$$H_0 : \mathbf{l}^T\mathbf{w}_{TP} = \rho_0 \quad \text{against} \quad H_1 : \mathbf{l}^T\mathbf{w}_{TP} = \rho_1 \neq \rho_0,$$

where $\rho_0, \rho_1 \in \mathbb{R}$ are preselected constants.

Using the proof of Theorem 3 we obtain

$$\frac{\alpha\mathbf{l}^T\hat{\mathbf{w}}_{TP}}{\mathbf{l}^T\mathbf{S}^+\mathbf{l}}|\overline{\mathbf{y}}^T\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}} = y \sim \mathcal{N}\left(\frac{\alpha\mathbf{l}^T\mathbf{w}_{TP}}{\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}}, \left(\frac{1}{n} + \frac{1}{n-1}y\right)\frac{1}{\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}}\right)$$

and, hence, under H_0

$$\sqrt{\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}}\frac{\frac{\alpha\mathbf{l}^T\hat{\mathbf{w}}_{TP}}{\mathbf{l}^T\mathbf{S}^+\mathbf{l}} - \frac{\alpha\rho_0}{\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}}}{\sqrt{\frac{1}{n} + \frac{1}{n-1}\overline{\mathbf{y}}^T\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}}}} \sim \mathcal{N}(0,1).$$

The last statistic depends on the unknown quantity $\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}$, which is a nuisance parameter. In order to derive an unconditional test without dependence on $\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}$, we consider an additional hypothesis given by

$$H_0 : \mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l} = v_0 \quad \text{against} \quad H_1 : \mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l} = v_1 \neq v_0$$

with the test statistic expressed as

$$T_2 = (n-1)\frac{v_0}{\mathbf{l}^T\mathbf{S}^+\mathbf{l}}.$$

Under the null hypothesis $H_0 : \mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l} = v_0$ it holds that T_2 has a χ^2 distribution with $n-r$ degrees of freedom. Putting these two marginal tests together, we get a two dimensional test with statistics $\tilde{T} = (T_1^*, T_2)^T$ where

$$T_1^* = \sqrt{v_0}\frac{\frac{\alpha\mathbf{l}^T\hat{\mathbf{w}}_{TP}}{\mathbf{l}^T\mathbf{S}^+\mathbf{l}} - \frac{\alpha\rho_0}{v_0}}{\sqrt{\frac{1}{n} + \frac{1}{n-1}\overline{\mathbf{y}}^T\hat{\mathbf{R}}_\mathbf{l}\overline{\mathbf{y}}}}$$

and the global hypothesis is expressed as

$$H_0 : \mathbf{l}^T\mathbf{w}_{TP} = \rho_0,\ \mathbf{l}^T\boldsymbol{\Sigma}^{+}\mathbf{l} = v_0 \quad \text{against} \quad H_1 : \mathbf{l}^T\mathbf{w}_{TP} = \rho_1 \neq \rho_0,\ \mathbf{l}^T\boldsymbol{\Sigma}^{+}\mathbf{l} = v_1 \neq v_0.$$

The distribution of the test statistics $\tilde{T}$ is derived in the next theorem.

Theorem 4. *Let $\mathbf{x}_1, ..., \mathbf{x}_n$ be i.i.d. random vectors with $\mathbf{x}_1 \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $k > n-1$ and with rank$(\boldsymbol{\Sigma}) = r \leq n-1$. Then*

(a) the density function of $\tilde{T}$ is given by

$$\begin{aligned} f_{\tilde{T}}(x,z) &= \frac{n(n-r+1)}{(r-1)(n-1)}\beta^{-1} f_{\chi^2_{n-r}}\left(\frac{z}{\beta}\right)\int_0^\infty f_{\mathcal{N}(\delta(y),\beta)}(x) \\ &\quad \times f_{\mathcal{F}_{r-1,n-r+1,ns}}\left(\frac{n(n-r+1)}{(r-1)(n-1)}y\right)dy, \end{aligned}$$

where $\delta(y) = \alpha\lambda_1/\sqrt{\frac{1}{n}+\frac{1}{n-1}\omega}$, $\beta = v_0/v_1$, and $\lambda_1 = \sqrt{v_0}(\rho_1/v_1 - \rho_0/v_0)$.

(b) under the null hypothesis it holds that $T_1^* \sim N(0,1)$, $T_2 \sim \chi^2_{n-r}$, and T_1^* is independent of T_2.

Proof. Following the proof of Theorem 3 we get that

- $\dfrac{\mathbf{l}^T\boldsymbol{\Sigma}^{+}\mathbf{l}}{(n-1)^{-1}\mathbf{l}^T\mathbf{S}^{+}\mathbf{l}} \sim \chi^2_{n-r}$;
- $\frac{n(n-r+1)}{(n-1)(r-1)}\overline{\mathbf{y}}^T\hat{\mathbf{R}}_{\mathbf{l}}\overline{\mathbf{y}} \sim \mathcal{F}_{r-1,n-r+1,ns}$;
- $T_1^*|\overline{\mathbf{y}}^T\hat{\mathbf{R}}_{\mathbf{l}}\overline{\mathbf{y}} \sim \mathcal{N}(\delta(y),\beta)$;
- $\frac{\mathbf{l}^T\hat{\mathbf{w}}_{TP}}{\mathbf{l}^T\mathbf{S}^{+}\mathbf{l}}$ and $\overline{\mathbf{y}}^T\hat{\mathbf{R}}_{\mathbf{l}}\overline{\mathbf{y}}$ are independent of $\mathbf{l}^T\mathbf{S}^{+}\mathbf{l}$,

from which part (a) of the theorem follows. The statement of part (b) is obtained by noting that $\delta(y) = 0$ under the null hypothesis. □

From Theorem 4, a joint $(1-\gamma)$ confidence region for (ρ_0, v_0) is given by

$$\begin{aligned} \mathcal{A} = \Bigg\{ (\rho, v) : \rho \in v\frac{\mathbf{l}^T\hat{\mathbf{w}}_{TP}}{\mathbf{l}^T\mathbf{S}^{+}\mathbf{l}} + \frac{z_{1-\tilde{\gamma}/2}\sqrt{v}}{\alpha}\sqrt{\frac{1}{n}+\frac{1}{n-1}\overline{\mathbf{y}}^T\hat{\mathbf{R}}_{\mathbf{l}}\overline{\mathbf{y}}}\,(-1,1)\,, \\ v \in \frac{\mathbf{l}^T\mathbf{S}^{+}\mathbf{l}}{n-1}\left(\chi^2_{n-r;\tilde{\gamma}/2}, \chi^2_{n-r;1-\tilde{\gamma}/2}\right) \Bigg\}, \end{aligned}$$

where $\tilde{\gamma} = 1-(1-\gamma)^2$ is the Sidak correction of a multiple test; $z_{1-\tilde{\gamma}/2}$ and $\chi^2_{n-r;1-\tilde{\gamma}/2}$ are $1-\tilde{\gamma}/2$ quantiles of the standard normal and the χ^2-distribution with $n-r$ degrees of freedom respectively.

3 High-dimensional asymptotics

In this section we investigate the asymptotic behavior of the estimated TP weights and the corresponding inference procedure under the high-dimensional asymptotic regime. We treat the rank $r = r_n$ of the covariance matrix $\mathbf{\Sigma}$ as the "actual" dimension of the data-generating process and assume that $r_n/n \to c \in [0, 1)$ as $n \to \infty$. It is noted that we impose no relationship between k and n with the exception that $k > n$ and, in reality, the portfolio dimension k can growth to infinity much faster than the sample size n does, e.g., an exponential growth can be considered which is very popular in economic literature.

In Theorem 5 we derive the high-dimensional asymptotic distribution of a linear combination of the estimated TP weights that follows from Theorem 1.

Theorem 5. *Let $\mathbf{x}_1, ..., \mathbf{x}_n$ be i.i.d. random vectors with $\mathbf{x}_1 \sim \mathcal{N}_k(\boldsymbol{\mu}, \mathbf{\Sigma})$, $k > n-1$ and with rank$(\mathbf{\Sigma}) = r_n \leq n-1$. Assume that there exists $\gamma > 0$ such that $r_n^{-\gamma}(\boldsymbol{\mu} - r_f\mathbf{1}_k)^T\mathbf{\Sigma}^+(\boldsymbol{\mu} - r_f\mathbf{1}_k) < \infty$ uniformly on r_n. Also, let $\mathbf{l}$ be a k-dimensional vector of constants such that $r_n^{-\gamma}\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l} < \infty$ uniformly on r_n. Define $c_n = r_n/n$. Then, under Assumption (A1), we get the following high-dimensional asymptotic distribution*

$$\sqrt{n-r_n}\sigma_\gamma^{-1}\left(\mathbf{l}^T\hat{\mathbf{w}}_{TP} - \frac{n-1}{n-r_n}\mathbf{l}^T\mathbf{w}_{TP}\right) \xrightarrow{d} \mathcal{N}(0, 1), \tag{9}$$

where

$$\sigma_\gamma^2 = \frac{\alpha^{-2}}{(1-c_n)^2}\left(\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l} + \left(\alpha\mathbf{l}^T\mathbf{w}_{TP}\right)^2 + \mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}(\boldsymbol{\mu} - r_f\mathbf{1}_k)^T\mathbf{\Sigma}^+(\boldsymbol{\mu} - r_f\mathbf{1}_k)\right). \tag{10}$$

Proof. The application of the stochastic representation from Corollary 1 leads to

$$\begin{aligned}\left(\mathbf{l}^T\hat{\mathbf{w}}_{TP} - \frac{n-1}{n-r_n}\mathbf{l}^T\mathbf{w}_{TP}\right) &= \alpha^{-1}\mathbf{l}^T\mathbf{\Sigma}^+(\boldsymbol{\mu} - r_f\mathbf{1}_k)\frac{n-1}{\xi}\left(1 - \frac{\xi}{n-r_n}\right) \\ &+ \alpha^{-1}\frac{n-1}{\xi}\sqrt{\left(1 + \frac{r_n-1}{n-r_n+1}u\right)\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}\frac{z_0}{n}}\end{aligned}$$

where $\xi \sim \chi^2_{n-r_n}$, $z_0 \sim \mathcal{N}(0, 1)$, and $u \sim \mathcal{F}(r_n - 1, n - r_n + 1, ns)$, $s = (\boldsymbol{\mu} - r_f\mathbf{1}_k)^T\mathbf{R}_\mathbf{l}(\boldsymbol{\mu} - r_f\mathbf{1}_k)$ with $\mathbf{R}_\mathbf{l} = \mathbf{\Sigma}^+ - \mathbf{\Sigma}^+\mathbf{l}\mathbf{l}^T\mathbf{\Sigma}^+/\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}$; ξ, u, and z_0 are mutually independently distributed.

In using Lemma 3 from Bodnar and Reiß [5], we get

$$\frac{\xi}{n-r_n} - 1 \stackrel{a.s.}{\to} 0 \quad \text{and} \quad \sqrt{n-r_n}\left(\frac{\xi}{n-r_n} - 1\right) \stackrel{d}{\to} \mathcal{N}(0,2).$$

The application of the stochastic representation of a non-central F-distributed random variables, i.e., $u = \dfrac{\eta_1/(r_n-1) + (z_1 + \sqrt{n}\sqrt{s})^2/(r_n-1)}{\eta_2/(n-r_n+1)}$ with independent variables $\eta_1 \sim \chi^2_{r_n-2}$, $\eta_2 \sim \chi^2_{n-r_n+1}$, and $z_1 \sim \mathcal{N}(0,1)$, leads to

$$\begin{aligned} u - 1 - \frac{ns}{r_n-1} &= \left(\frac{\eta_1/(r_n-1) + (z_1 + \sqrt{n}\sqrt{s})^2/(r_n-1)}{\eta_2/(n-r_n+1)} - 1 - \frac{ns}{r_n-1}\right) \\ = \ & \frac{\eta_1/(r_n-1) + (z_1 + \sqrt{n}\sqrt{s})^2/(r_n-1) - 1 - \frac{ns}{r_n-1}}{\eta_2/(n-r_n+1)} \\ - \ & \left(1 + \frac{ns}{r_n-1}\right)\frac{n-r_n+1}{\eta_2}\left(1 - \frac{\eta_2}{n-r_n+1}\right) \stackrel{a.s.}{\to} 0\,. \end{aligned}$$

Putting these results together and using that ξ and z_0 are independent, we obtain the statement of the theorem. □

From Theorem 5 we obtain the distribution of any finite combination of the estimated TP weights. Namely, for the q linear combinations of $\hat{\mathbf{w}}_{TP}$ presented by matrix $\mathbf{L}$, it holds that

$$\sqrt{n-r_n}\mathbf{\Omega}_\gamma^{-1/2}\left(\mathbf{L}\hat{\mathbf{w}}_{TP} - \frac{n-1}{n-r_n}\mathbf{L}\mathbf{w}_{TP}\right) \stackrel{d}{\to} \mathcal{N}_q(\mathbf{0}, \mathbf{I}),$$

where

$$\begin{aligned} \mathbf{\Omega}_\gamma &= \frac{\alpha^{-2}}{(1-c_n)^2}\Bigg(\mathbf{L}\mathbf{\Sigma}^+\mathbf{L}^T + \mathbf{L}\mathbf{\Sigma}^+(\boldsymbol{\mu} - r_f\mathbf{1}_k)(\boldsymbol{\mu} - r_f\mathbf{1}_k)^T\mathbf{\Sigma}^+\mathbf{L}^T \\ &+ (\boldsymbol{\mu} - r_f\mathbf{1}_k)^T\mathbf{\Sigma}^+(\boldsymbol{\mu} - r_f\mathbf{1}_k)\mathbf{L}\mathbf{\Sigma}^+\mathbf{L}^T\Bigg). \end{aligned}$$

It is remarkable that, we impose no assumption on the eigenvalues of the covariance matrix $\mathbf{\Sigma}$ in Theorem 5. The derived asymptotic results are also valid when $\mathbf{\Sigma}$ possesses unbounded spectrum, i.e. when its largest eigenvalue tends to infinity as $r_n \to \infty$. Consequently, our findings can also be applied in the case of a factor model which is a popular model for the asset returns with an unbounded spectrum (see, e.g., Chamberlain [6]; Fan et al. [7, 9, 8]; Bodnar and Reiß [5]). Finally, the constant γ is a technical one and has only a minor influence on the results presented in Theorems 5 and 6. It only controls the growth rate of the quadratic form and we mainly require that they are of the same order for $r_n \to \infty$ as $n \to \infty$.

Next, we derive the joint distribution of the test statistics T presented in (6) both under the null and under the alternative hypotheses.

Theorem 6. *Let $\mathbf{x}_1, ..., \mathbf{x}_n$ be i.i.d. random vectors with $\mathbf{x}_1 \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $k > n-1$ and with rank$(\boldsymbol{\Sigma}) = r_n \le n-1$. Assume that there exists $\gamma > 0$ such that $r_n^{-\gamma}(\boldsymbol{\mu} - r_f \mathbf{1}_k)^T \boldsymbol{\Sigma}^+ (\boldsymbol{\mu} - r_f \mathbf{1}_k) < \infty$ uniformly on r_n. Also, let $\mathbf{l}$ be a k-dimensional vector of constants such that $r_n^{-\gamma} \mathbf{l}^T \boldsymbol{\Sigma}^+ \mathbf{l} < \infty$ uniformly on r_n. Define $c_n = r_n/n$. Then, under Assumption (A1), we get*

(a)

$$\sigma_T^{-1} \left(T - \frac{\sqrt{n}\alpha\rho_1}{\sqrt{\mathbf{l}^T \boldsymbol{\Sigma}^+ \mathbf{l} \left(1 + \frac{r_n - 1}{n - r_n + 1}\left(1 + \frac{n}{r_n - 1} s\right)\right)}} \right) \xrightarrow{d} \mathcal{N}(0, 1)\,, \qquad (11)$$

where

$$\sigma_T^2 \;=\; 1 + \frac{(\alpha\rho_1)^2}{\mathbf{l}^T \boldsymbol{\Sigma}^+ \mathbf{l}(1+s)} \left(2 + \frac{s^2 + c + 2s}{2(1+s)^2} \right),$$

with $s = (\boldsymbol{\mu} - r_f \mathbf{1}_k)\mathbf{R}_\mathbf{l}(\boldsymbol{\mu} - r_f \mathbf{1}_k)^T$ with $\mathbf{R}_\mathbf{l} = \boldsymbol{\Sigma}^+ - \boldsymbol{\Sigma}^+ \mathbf{l}\mathbf{l}^T \boldsymbol{\Sigma}^+ / \mathbf{l}^T \boldsymbol{\Sigma}^+ \mathbf{l}$;

(b) under the null hypothesis it holds that $T \sim \mathcal{N}(0, 1)$.

Proof. From the proof of Theorem 3, we get that

$$T | \bar{\mathbf{y}}^T \hat{\mathbf{R}}_\mathbf{l} \bar{\mathbf{y}} = y \sim t_{n - r_n, \vartheta(y)}$$

where $\vartheta(y) = \frac{\alpha\rho_1}{\sqrt{\mathbf{l}^T \boldsymbol{\Sigma}^+ \mathbf{l}\left(\frac{1}{n} + \frac{1}{n-1} y\right)}}$. Consequently, the stochastic representation of T is given by

$$T \stackrel{d}{=} \sqrt{\frac{n - r_n}{\xi}} \left(\frac{\sqrt{n}\alpha\rho_1}{\sqrt{\mathbf{l}^T \boldsymbol{\Sigma}^+ \mathbf{l}\left(1 + \frac{r_n - 1}{n - r_n + 1} u\right)}} + z_0 \right)$$

where $z_0 \sim \mathcal{N}(0, 1)$, $\xi \sim \chi^2_{n - r_n}$ and $u = \frac{n(n - r_n + 1)}{(n-1)(r_n - 1)} \bar{\mathbf{y}}^T \hat{\mathbf{R}}_\mathbf{l} \bar{\mathbf{y}} \sim \mathcal{F}_{r_n - 1, n - r_n + 1, ns}$; z_0, ξ, and u are independent.

It holds that

$$
\begin{aligned}
& T - \frac{\sqrt{n}\alpha\rho_1}{\sqrt{\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}\left(1+\frac{r_n-1}{n-r_n+1}\left(1+\frac{n}{r_n-1}s\right)\right)}} \\
= & \sqrt{\frac{n-r_n}{\xi}}z_0 + \frac{\alpha\rho_1}{\sqrt{\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}\left(1+\frac{r_n-1}{n-r_n+1}u\right)}} \\
\times & \left(\sqrt{n}\left(\sqrt{\frac{n-r_n}{\xi}}-1\right)+\sqrt{n}\left(1-\frac{\sqrt{1+\frac{r_n-1}{n-r_n+1}u}}{\sqrt{1+\frac{r_n-1}{n-r_n+1}\left(1+\frac{n}{r_n-1}s\right)}}\right)\right),
\end{aligned}
$$

where

$$
\begin{aligned}
& \sqrt{1+\frac{r_n-1}{n-r_n+1}\left(1+\frac{n}{r_n-1}s\right)}-\sqrt{1+\frac{r_n-1}{n-r_n+1}u} \\
= & \frac{\frac{r_n-1}{n-r_n+1}\left(1+\frac{n}{r_n-1}s-u\right)}{\sqrt{1+\frac{r_n-1}{n-r_n+1}u}+\sqrt{1+\frac{r_n-1}{n-r_n+1}\left(1+\frac{n}{r_n-1}s\right)}}.
\end{aligned}
$$

Next, we use the stochastic properties of the non-central F-distribution, i.e., there are exist independent random variables $\eta_1 \sim \chi^2_{r_n-2}$, $\eta_2 \sim \chi^2_{n-r_n+1}$, and z_1 such that

$$
u \stackrel{d}{=} \frac{\eta_1/(r_n-1)+(z_1+\sqrt{n}\sqrt{s})^2/(r_n-1)}{\eta_2/(n-r_n+1)}.
$$

It leads to

$$
\begin{aligned}
& \sqrt{n}\left(1+\frac{n}{r_n-1}s-u\right) \\
= & \sqrt{n}\left(1+\frac{n}{r_n-1}s-\frac{\eta_1/(r_n-1)+(z_1+\sqrt{n}\sqrt{s})^2/(r_n-1)}{\eta_2/(n-r_n+1)}\right) \\
= & \left(1+\frac{n}{r_n-1}s\right)\sqrt{n}\left(1-\frac{n-r_n+1}{\eta_2}\right)+\frac{n-r_n+1}{\eta_2}\sqrt{n}\left(1-\frac{\eta_1}{r_n-1}\right) \\
+ & \frac{n-r_n+1}{\eta_2}\sqrt{n}\frac{z_1^2}{r_n-1}+2z_1\frac{n}{r_n-1}\sqrt{s}\frac{n-r_n+1}{\eta_2}.
\end{aligned}
$$

Since $z_0, z_1, \xi, \eta_1, \eta_2$ are independent, we get

$$\sigma_T^{-1}\left(T - \frac{\sqrt{n}\alpha\rho_1}{\sqrt{\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}\left(1+\frac{r_n-1}{n-r_n+1}\left(1+\frac{n}{r_n-1}s\right)\right)}}\right) \xrightarrow{d} \mathcal{N}(0,1)$$

with

$$\begin{aligned}\sigma_T^2 &= 1 + \frac{(\alpha\rho_1)^2}{\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}(1+s)}(1-c) \\ &\times \left(2\frac{1}{1-c} + \frac{(1-c)^2}{4(1+s)^2}\frac{c^2}{(1-c)^2}\left(2\left(1+\frac{s}{c}\right)^2\frac{1}{1-c} + \frac{2}{c} + 4\frac{s}{c^2}\right)\right) \\ &= 1 + \frac{(\alpha\rho_1)^2}{\mathbf{l}^T\mathbf{\Sigma}^+\mathbf{l}(1+s)}\left(2 + \frac{s^2+c+2s}{2(1+s)^2}\right).\end{aligned}$$

The statement of Theorem 6.(b) is obtained by noting that $\rho_1 = 0$ under H_0. □

4 Empirical study

In this section the results of an empirical study are presented. It is shown how we can apply the theory from the previous sections to real data. We consider the log return weekly data from S&P 500 of 440 stocks for the period from the 30th of April, 2007 to the 21st of April, 2017 resulting in 513 observations. The weekly returns on the three-month US treasury bill are used as the risk-free rate. The risk aversion coefficient α is set to 100.

One of the main issues considered in the empirical study is that the singularity of the covariance matrix and, consequently, in the data generating process can never be observed in the strict sense, since the observational noise will prevent us from obtaining the determinant of the estimated covariance matrix exactly equal to zero. We assume that our observations are coming from the model

$$\mathbf{Y} = \mathbf{X} + \mathbf{E},$$

where $\mathbf{X}$ is following the singular model that is discussed in our work, while $\mathbf{E}$ is an observational noise. In order to verify that the data generating model of $\mathbf{X}$ is indeed singular, we conduct the test developed in Nadakuditi and Eldeman [13] to find the rank of $\mathbf{\Sigma}$. Figure 1 presents the behaviour of the estimator for the rank of the covariance matrix by using the rolling estimator with the estimation window of 300 weeks. We observe a relatively stable behaviour in the resulting estimators for the rank lying between 115 and 140.

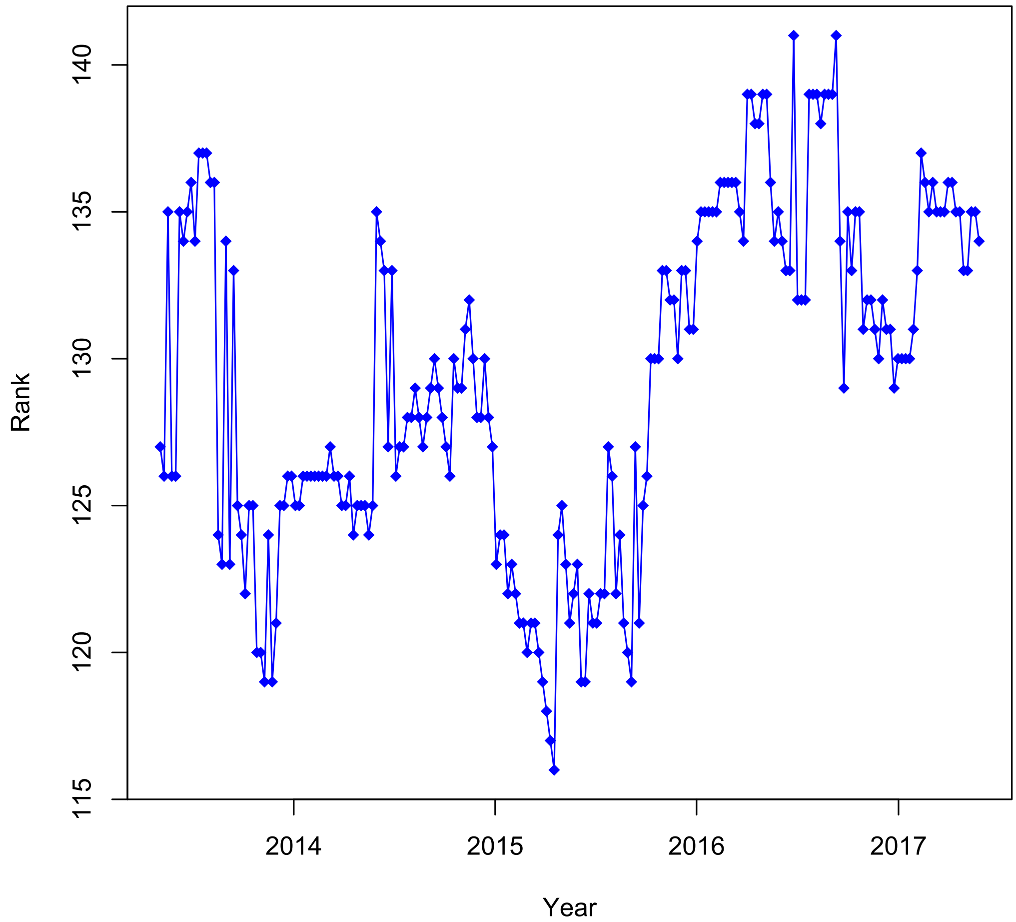

Figure 1: The rolling estimator for the rank of the covariance matrix with the estimation window of 300 weeks.

The smallest value of the estimated ranks is present at the beginning of 2015 following a slight increase in 2016 and 2017. To this end, all obtained ranks are considerably smaller than the sample size $n = 300$. The latter finding leads to the conclusion that a large amount of noise is present in the considered financial data which influences both the estimation of the covariance matrix and the determination of the structure of optimal portfolios. Moreover, the obtained empirical result motivates the application of the derived theoretical findings to the considered data.

Next, we assess the quality of the asymptotic results presented in Theorem 5 by comparing the asymptotic distribution of the estimated TP weights derived under the high-dimensional asymptotic regime to the corresponding exact ones obtained by applying the stochastic representation of Theorem 1. In order to make this comparison for practically relevant values of the mean vector and

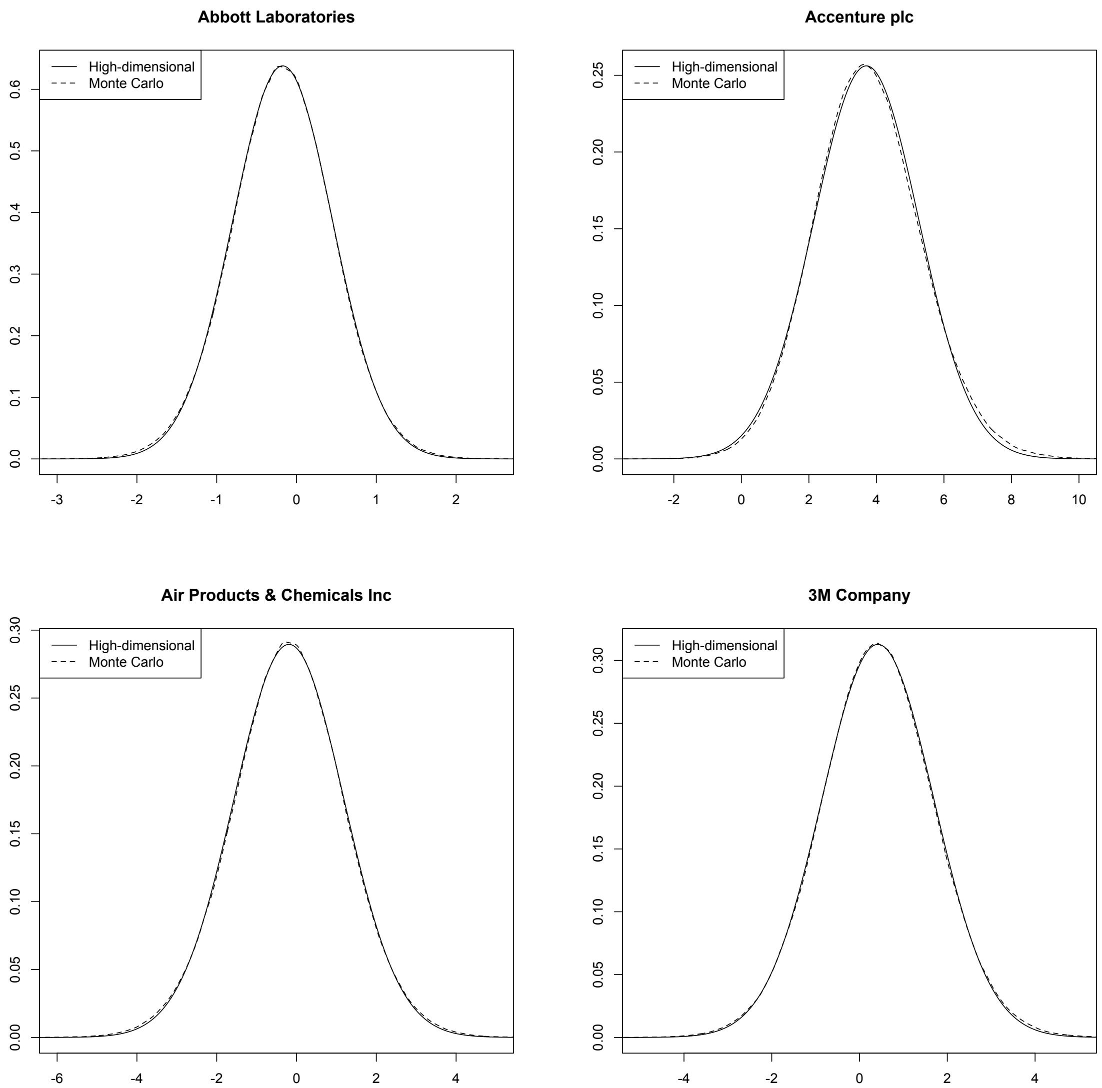

Figure 2: The densities of the sampling and the high-dimensional asymptotic distributions of the TP portfolio weights for four stocks: Abbott Laboratories, Accenture plc, Air Products & Chemicals Inc, and 3M Company.

the covariance matrix, we replace these quantities in the expressions provided by Theorems 1 and 5 with the corresponding sample counterparts obtained by using the most recent 300 observation vectors of the returns.

In Figure 2, both the high-dimensional asymptotic and the finite-sample densities of the TP weights corresponding to Abbott Laboratories, Accenture plc, Air Products & Chemicals Inc, and 3M Company are present. The asymptotic densities are calculated following Theorem 5, while for the finite-sample densities we apply the stochastic representation of Theorem 1 which is used to generate samples of $N = 10^5$ independent realizations of the estimated TP weights. Based on these simulated data, the sampling densities are obtained by

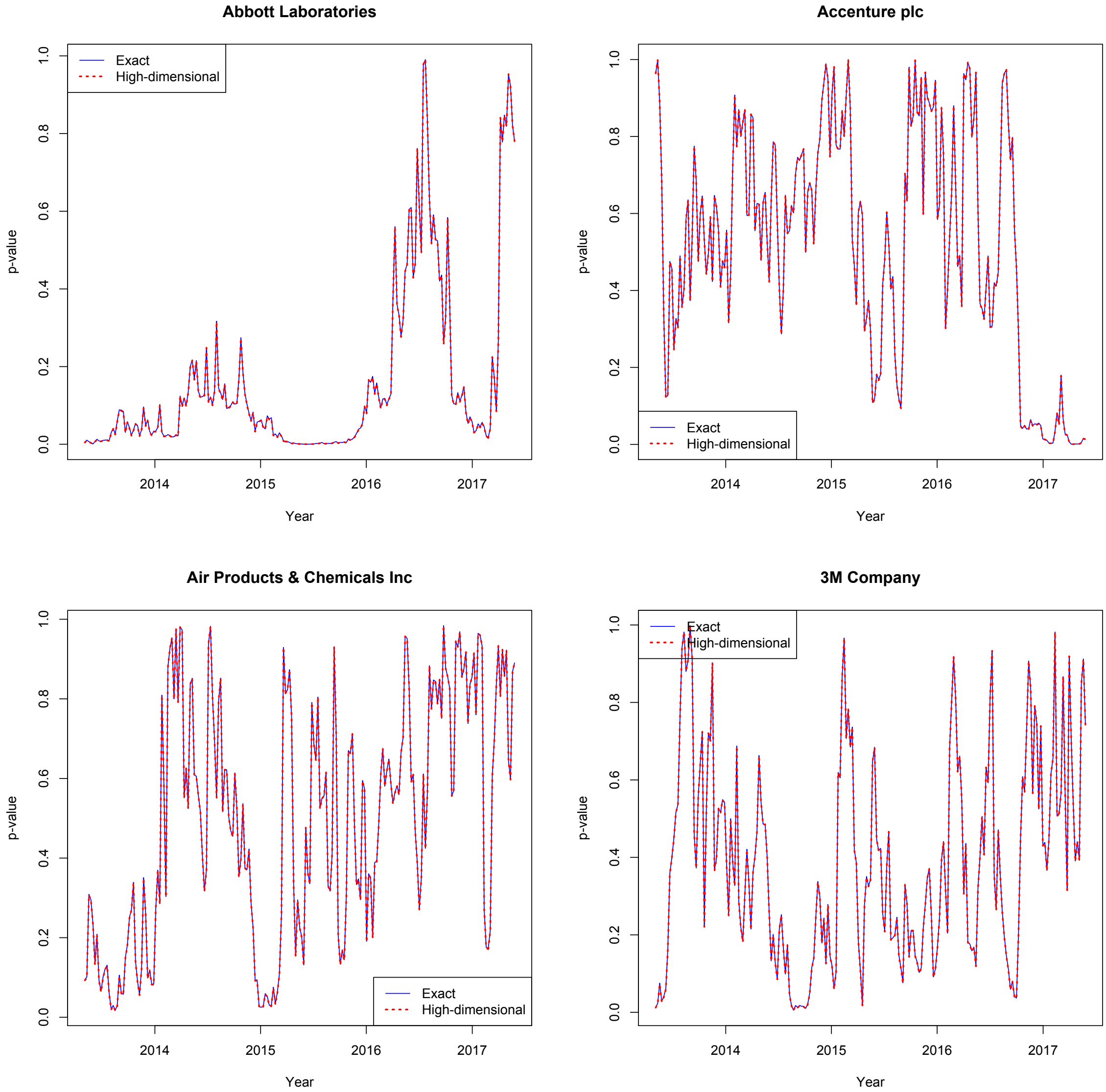

Figure 3: p-values of the exact and the asymptotic tests on the tangency portfolio weights applied to four stocks: Abbott Laboratories, Accenture plc, Air Products & Chemicals Inc, and 3M Company.

the kernel density estimation. We observe that the high-dimensional asymptotic results provide a very good approximation even for the moderate sample size, like $n = 300$. There are only minor differences in the two densities presented in each out of four plots in Figure 2. Moreover, we conclude that only the estimated TP weight in the case of Accenture plc deviates considerably from zero, while the estimated TP weights of all other considered stocks have a large amount of probability mass around zero.

Since some estimated TP weights possess a positive probability mass around zero, we run statistical tests to verify if their population counterparts are equal to zero. In practice, this procedure might reduce the size of the portfolio and, consequently, the transaction costs. The results obtained in Theorems 3 and 6 are used for this purpose. In Figure 3, we present the time series behaviour of the p-values obtained from the exact test and the asymptotic test on the hypothesis (5). The same rolling window procedure as in Figure 1 is used here. The four plots in Figure 3 show the results for the four stocks: Abbott Laboratories, Accenture plc, Air Products & Chemicals Inc, and 3M Company. We observe that the p-values are relatively large in almost all of the considered cases leading to the conclusion that the null hypothesis (5) cannot be rejected. The deviations from zero are present mainly for the TP weight of Abbott Laboratories in 2015 and of Accenture plc at the end of 2016 and at beginning of 2017.

5 Summary

Since recognizing the impact of the estimation error on the behaviour of optimal portfolios, estimating and testing the structure of optimal portfolios have become important topics in financial literature. Since the expressions of optimal portfolio weights involve both the mean vector and the covariance matrix, the common statistical tests on the mean vector and on the covariance matrix cannot be applied in this situation and new approaches have to be developed.

We deal with this problem under the additional challenge of singularity which is present in both the population covariance matrix and the sample covariance matrix. The finite-sample distributions of the estimated TP weights are characterized by deriving their stochastic representation employing which the mean vector and the covariance matrix of the estimated TP weights are obtained. Furthermore, we develop exact tests on the structure of the TP and derive the distribution functions of the test statistics under both the null hypothesis and the alternative hypothesis.

Another point treated in the paper deals with the problem of high-dimensional optimal portfolios which have risen in popularity in financial literature recently. We contribute by deriving the asymptotic distributions of the estimated TP weights under the high-dimensional asymptotic regime, i.e. when both the portfolio dimension and the sample size increase. Additionally, the high-dimensional asymptotic distribution of the considered test statistic is obtained under both hypotheses.

The theoretical findings are applied to real data. In using the method suggested by Nadakuditi and Eldeman [13] to determine the rank of the covariance matrix, we find that the covariance matrix of the returns on the stocks included into the S&P 500 index is singular. This empirical result motivates the application of the obtained theoretical findings to the considered data. Finally, we find that some weights of the TP do not deviate significantly from zero, which allows the reduction of the portfolio size and, consequently, can lead to smaller transaction costs.

Appendix

In this section, we consider the distributional properties of the product of the singular inverse Wishart random matrix and a normal vector, which are used in the derivation of our main results presented in Section 2.

Let $\mathbf{A} \sim \mathcal{W}_k(n, \mathbf{\Sigma})$, i.e., the random matrix $\mathbf{A}$ has a k-dimensional singular Wishart distribution with n degrees of freedom and covariance matrix $\mathbf{\Sigma}$ which is non-negative definite with $rank(\mathbf{\Sigma}) = r \leq n < k$. Also, let $\mathbf{I}_k$ be the $k \times k$ identity matrix and the symbol $\overset{d}{=}$ stands for the equality in distribution.

Since $\mathbf{A}$ and $\mathbf{\Sigma}$ are singular, then the Moore-Penrose inverse $\mathbf{W}^+$ of a matrix $\mathbf{W}$ will be employed as an important tool of analysis. Let us recall that a matrix $\mathbf{W}^+$ is the Moore-Penrose inverse of $\mathbf{W}$ if the following conditions hold

(I) $\mathbf{W}\mathbf{W}^+\mathbf{W} = \mathbf{W}$,

(II) $\mathbf{W}^+\mathbf{W}\mathbf{W}^+ = \mathbf{W}^+$,

(III) $(\mathbf{W}\mathbf{W}^+)^T = \mathbf{W}\mathbf{W}^+$,

(IV) $(\mathbf{W}^+\mathbf{W})^T = \mathbf{W}^+\mathbf{W}$.

Next, we define the square root of the Moore-Penrose inverse $\mathbf{W}^+$ of a symmetric singular matrix $\mathbf{W}$ with $rank(\mathbf{W}) = r$. Let $\mathbf{W} = \mathbf{H}\mathbf{D}\mathbf{H}^T$ be the singular value decomposition of $\mathbf{W}$ where $\mathbf{D} = diag(d_1, \ldots, d_r)$ is the $r \times r$ matrix of non-zero eigenvalues and $\mathbf{H} : k \times r$ is the semi-orthogonal matrix of the corresponding eigenvectors such that $\mathbf{H}^T\mathbf{H} = \mathbf{I}_r$. Then $(\mathbf{W}^+)^{1/2} = \mathbf{H}\mathbf{D}^{-1/2}\mathbf{H}^T$ with $\mathbf{D}^{-1/2} = diag(d_1^{-1/2}, ..., d_r^{-1/2})$.

In Lemma 1, we derive the stochastic representation of linear symmetric transformations of the product of a singular Wishart matrix and a normal vector.

Lemma 1. *Let* $\mathbf{A} \sim \mathcal{W}_k(n, \mathbf{\Sigma})$, $\mathbf{z} \sim \mathcal{N}_k(\boldsymbol{\mu}, \lambda\mathbf{\Sigma})$ *with* $\lambda > 0$, *and* $rank(\mathbf{\Sigma}) = r \leq n < k$. *Let* $\mathbf{A}$ *and* $\mathbf{z}$ *be independent. Consider* $\mathbf{L}$ *a* $p \times k$ *non-random matrix of rank* $p < r$ *such that* $rank(\mathbf{L}\mathbf{\Sigma}) = p$. *Additionally, let* $\mathbf{S}_1 = (\mathbf{L}\mathbf{\Sigma}^+\mathbf{L}^T)^{-1/2}\mathbf{L}\mathbf{\Sigma}^{+1/2}$ *and* $\mathbf{Q}_1 = \mathbf{S}_1^T\mathbf{S}_1$. *Then the stochastic representation of*

$\mathbf{LA}^+\mathbf{z}$ *is given by*

$$\mathbf{LA}^+\mathbf{z} \overset{d}{=} \xi^{-1}\left(\mathbf{L\Sigma}^{+1/2}\mathbf{y} + \sqrt{\frac{1}{n-r+2}}(\mathbf{L\Sigma}^+\mathbf{L}^T)^{1/2}\right.$$
$$\left.\times\left[\sqrt{\mathbf{y}^T\mathbf{y}}\mathbf{I}_p - \frac{\sqrt{\mathbf{y}^T\mathbf{y}} - \sqrt{\mathbf{y}^T(\mathbf{I}_k - \mathbf{Q}_1)\mathbf{y}}}{\mathbf{y}^T\mathbf{Q}_1\mathbf{y}}\mathbf{S}_1\mathbf{y}\mathbf{y}^T\mathbf{S}_1^T\right]\mathbf{t}_0\right),$$

where $\xi \sim \chi^2_{n-r+1}$*,* $\mathbf{t}_0 \sim t_p(n-r+2;\mathbf{0},\mathbf{I}_p)$*, and* $\mathbf{y} \sim \mathcal{N}_k\left(\mathbf{\Sigma}^{+1/2}\boldsymbol{\mu}, \lambda\mathbf{PP}^T\right)$ *where* $\mathbf{P} : k \times r$ *is the semi-orthogonal matrix of the eigenvectors of* $\mathbf{\Sigma}$ *such that* $\mathbf{P}^T\mathbf{P} = \mathbf{I}_r$*. Moreover,* ξ*,* $\mathbf{y}$*, and* $\mathbf{t}_0$ *are mutually independent. The symbol* $t_p(d;\mathbf{a},\mathbf{B})$ *stands for the p-dimensional multivariate t-distribution with d degrees of freedom, location parameter* $\mathbf{a}$*, and dispersion matrix* $\mathbf{B}$*.*

Proof. Since $\mathbf{z}$ and $\mathbf{A}$ are independently distributed, it follows that the conditional distribution of $\mathbf{LA}^+\mathbf{z}$ given $\mathbf{z} = \mathbf{z}^*$ is equal to the distribution of $\mathbf{LA}^+\mathbf{z}^*$ which can be expressed in the following form

$$\mathbf{LA}^+\mathbf{z}^* = \mathbf{z}^{*T}\mathbf{\Sigma}^+\mathbf{z}^*\frac{\mathbf{LA}^+\mathbf{z}^*}{\mathbf{z}^{*T}\mathbf{A}^+\mathbf{z}^*}\frac{\mathbf{z}^{*T}\mathbf{A}^+\mathbf{z}^*}{\mathbf{z}^{*T}\mathbf{\Sigma}^+\mathbf{z}^*}. \quad (12)$$

Next, we show that $\mathbf{z}^{*T}\mathbf{A}^+\mathbf{z}^*/\mathbf{z}^{*T}\mathbf{\Sigma}^+\mathbf{z}^*$ is independent of $\mathbf{z}^{*T}\mathbf{\Sigma}^+\mathbf{z}^*\mathbf{LA}^+\mathbf{z}^*/\mathbf{z}^{*T}\mathbf{A}^+\mathbf{z}^*$, and derive their distributions.

Let $\widetilde{\mathbf{L}} = (\mathbf{L}^T, \mathbf{z}^*)^T$ such that $rank(\widetilde{\mathbf{L}}) = p+1 \le r$ and define $\widetilde{\mathbf{A}} = \widetilde{\mathbf{L}}\mathbf{A}^+\widetilde{\mathbf{L}}^T = \left\{\widetilde{\mathbf{A}}_{ij}\right\}_{i,j=1,2}$ with $\widetilde{\mathbf{A}}_{11} = \mathbf{LA}^+\mathbf{L}^T$, $\widetilde{\mathbf{A}}_{12} = \mathbf{LA}^+\mathbf{z}^*$, $\widetilde{\mathbf{A}}_{21} = \mathbf{z}^{*T}\mathbf{A}^+\mathbf{L}^T$, and $\widetilde{A}_{22} = \mathbf{z}^{*T}\mathbf{A}^+\mathbf{z}^*$. Similarly, we define $\widetilde{\mathbf{\Sigma}} = \widetilde{\mathbf{L}}\mathbf{\Sigma}^+\widetilde{\mathbf{L}}^T = \left\{\widetilde{\mathbf{\Sigma}}_{ij}\right\}_{i,j=1,2}$ with $\widetilde{\mathbf{\Sigma}}_{11} = \mathbf{L\Sigma}^+\mathbf{L}^T$, $\widetilde{\mathbf{\Sigma}}_{12} = \mathbf{L\Sigma}^+\mathbf{z}^*$, $\widetilde{\mathbf{\Sigma}}_{21} = \mathbf{z}^{*T}\mathbf{\Sigma}^+\mathbf{L}^T$, and $\widetilde{\Sigma}_{22} = \mathbf{z}^{*T}\mathbf{\Sigma}^+\mathbf{z}^*$. Since $\mathbf{A} \sim \mathcal{W}_k(n,\mathbf{\Sigma})$ and $rank(\widetilde{\mathbf{L}}) = p+1 \le r$ we get from Theorem 1 of Bodnar et al. [1] and Theorem 3.4.1 of Gupta and Nagar [10] that the random matrix $\widetilde{\mathbf{A}} = \left\{\widetilde{\mathbf{A}}_{ij}\right\}_{i,j=1,2}$ has the $(p+1)$-variate inverse Wishart distribution with $(n-r+2p+3)$ degrees of freedom and the non-singular covariance matrix $\widetilde{\mathbf{\Sigma}}$, i.e. $\widetilde{\mathbf{A}} \sim \mathcal{IW}(n-r+2p+3,\widetilde{\mathbf{\Sigma}})$. Then from Theorem 3 of Bodnar and Okhrin [3] we get that $\mathbf{z}^{*T}\mathbf{A}^+\mathbf{z}^*$ is independent of $\mathbf{LA}^+\mathbf{z}^*/\mathbf{z}^{*T}\mathbf{A}^+\mathbf{z}^*$ for given $\mathbf{z}^*$. Hence, $\mathbf{z}^{*T}\mathbf{A}^+\mathbf{z}^*/\mathbf{z}^{*T}\mathbf{\Sigma}^+\mathbf{z}^*$ is independent of $\mathbf{z}^{*T}\mathbf{\Sigma}^+\mathbf{z}^*\mathbf{LA}^+\mathbf{z}^*/\mathbf{z}^{*T}\mathbf{A}^+\mathbf{z}^*$.

The application of Corollary 1 in Bodnar et al. [1] leads to

$$\frac{\mathbf{z}^{*T}\mathbf{\Sigma}^+\mathbf{z}^*}{\mathbf{z}^{*T}\mathbf{A}^+\mathbf{z}^*} \sim \chi^2_{n-r+1}, \quad (13)$$

and is independent of $\mathbf{z}^*$. Consequently, $\xi = \dfrac{\mathbf{z}^T\mathbf{\Sigma}^+\mathbf{z}}{\mathbf{z}^T\mathbf{A}^+\mathbf{z}} \sim \chi^2_{n-r+1}$ and it is independent of $\mathbf{z}$. Moreover, $\mathbf{z}^T\mathbf{A}^+\mathbf{z}/\mathbf{z}^T\mathbf{\Sigma}^+\mathbf{z}$ and $\mathbf{z}^T\mathbf{\Sigma}^+\mathbf{z}\mathbf{LA}^+\mathbf{z}/\mathbf{z}^T\mathbf{A}^+\mathbf{z}$ are independent as well.

Finally, using the proof of Theorem 5 of Bodnar et al. [1] we obtain

$$\mathbf{z}^{*T}\mathbf{\Sigma}^{+}\mathbf{z}^{*}\frac{\mathbf{L}\mathbf{A}^{+}\mathbf{z}^{*}}{\mathbf{z}^{*T}\mathbf{A}^{+}\mathbf{z}^{*}} \sim t_p\left(n-r+2; \mathbf{L}\mathbf{\Sigma}^{+}\mathbf{z}^{*}, \frac{\mathbf{z}^{*T}\mathbf{\Sigma}^{+}\mathbf{z}^{*}}{n-r+2}\mathbf{L}\mathbf{R}_{\mathbf{z}^{*}}\mathbf{L}^{T}\right), \tag{14}$$

where $\mathbf{R}_{\mathbf{a}} = \mathbf{\Sigma}^{+} - \mathbf{\Sigma}^{+}\mathbf{a}\mathbf{a}^{T}\mathbf{\Sigma}^{+}/\mathbf{a}^{T}\mathbf{\Sigma}^{+}\mathbf{a}$ for a k-dimensional vector $\mathbf{a}$.

Substituting (13) and (14) in (12) we get the following stochastic representation of $\mathbf{L}\mathbf{A}^{+}\mathbf{z}$

$$\mathbf{L}\mathbf{A}^{+}\mathbf{z} \stackrel{d}{=} \xi^{-1}\left(\mathbf{L}\mathbf{\Sigma}^{+}\mathbf{z} + \sqrt{\frac{\mathbf{z}^{T}\mathbf{\Sigma}^{+}\mathbf{z}}{n-r+2}}(\mathbf{L}\mathbf{R}_{\mathbf{z}}\mathbf{L}^{T})^{1/2}\mathbf{t}_0\right),$$

where $\xi \sim \chi^2_{n-r+1}$, $\mathbf{z} \sim \mathcal{N}_k(\boldsymbol{\mu}, \lambda\mathbf{\Sigma})$, and $\mathbf{t}_0 \sim t_p(n-r+2; \mathbf{0}, \mathbf{I})$. Moreover, ξ, $\mathbf{z}$, and $\mathbf{t}_0$ are mutually independent.

Next, we calculate the square root of $\mathbf{L}\mathbf{R}_{\mathbf{z}}\mathbf{L}^{T}$ using the equality

$$(\mathbf{D} - \mathbf{b}\mathbf{b}^{T})^{1/2} = \mathbf{D}^{1/2}(\mathbf{I}_p - c\mathbf{D}^{-1/2}\mathbf{b}\mathbf{b}^{T}\mathbf{D}^{-1/2})$$

with $c = \frac{1-\sqrt{1-\mathbf{b}^{T}\mathbf{D}^{-1}\mathbf{b}}}{\mathbf{b}^{T}\mathbf{D}^{-1}\mathbf{b}}$, $\mathbf{D} = \mathbf{z}^{T}\mathbf{\Sigma}^{+}\mathbf{z}\mathbf{L}\mathbf{\Sigma}^{+}\mathbf{L}^{T}$, and $\mathbf{b} = \mathbf{L}\mathbf{\Sigma}^{+}\mathbf{z}$.

This leads to

$$\mathbf{L}\mathbf{A}^{+}\mathbf{z} \stackrel{d}{=} \xi^{-1}\Bigg(\mathbf{L}\mathbf{\Sigma}^{+}\mathbf{z} + \sqrt{\frac{1}{n-r+2}}(\mathbf{L}\mathbf{\Sigma}^{+}\mathbf{L}^{T})^{1/2}$$
$$\times\left[\sqrt{\mathbf{z}^{T}\mathbf{\Sigma}^{+}\mathbf{z}}\mathbf{I}_p - \frac{\sqrt{\mathbf{z}^{T}\mathbf{\Sigma}^{+}\mathbf{z}} - \sqrt{\mathbf{z}^{T}(\mathbf{\Sigma}^{+} - \mathbf{\Sigma}^{+1/2}\mathbf{Q}_1\mathbf{\Sigma}^{+1/2})\mathbf{z}}}{\mathbf{z}^{T}\mathbf{\Sigma}^{+1/2}\mathbf{Q}_1\mathbf{\Sigma}^{+1/2}\mathbf{z}}\mathbf{S}_1\mathbf{\Sigma}^{+1/2}\mathbf{z}\mathbf{z}^{T}\mathbf{\Sigma}^{+1/2}\mathbf{S}_1^{T}\right]\mathbf{t}_0\Bigg)$$

with $\mathbf{S}_1 = (\mathbf{L}\mathbf{\Sigma}^{+}\mathbf{L}^{T})^{-1/2}\mathbf{L}\mathbf{\Sigma}^{+1/2}$ and $\mathbf{Q}_1 = \mathbf{S}_1^{T}\mathbf{S}_1$.

Let $\mathbf{\Sigma} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^{T}$ be the singular value decomposition of $\mathbf{\Sigma}$ where $\mathbf{\Lambda} = diag(\delta_1, \dots, \delta_r)$ is the $r \times r$ matrix of non-zero eigenvalues and $\mathbf{P} : k \times r$ is the semi-orthogonal matrix of the corresponding eigenvectors such that $\mathbf{P}^{T}\mathbf{P} = \mathbf{I}_r$. Then after the transformation $\mathbf{y} = \mathbf{\Sigma}^{+1/2}\mathbf{z} \sim \mathcal{N}_k\left(\mathbf{\Sigma}^{+1/2}\boldsymbol{\mu}, \lambda\mathbf{P}\mathbf{P}^{T}\right)$ we get the statement of the theorem. □

In the next corollary, we consider an important special case of Lemma 1 when $p = 1$ and $\mathbf{L} = \mathbf{l}^{T}$.

Corollary 1. *Let* $\mathbf{A} \sim \mathcal{W}_k(n, \mathbf{\Sigma})$, *and* $\mathbf{z} \sim \mathcal{N}_k(\boldsymbol{\mu}, \lambda\mathbf{\Sigma})$ *with* $\lambda > 0$ *and* $rank(\mathbf{\Sigma}) = r \le n < k$ *with* $r \ge 2$. *Let* $\mathbf{A}$ *and* $\mathbf{z}$ *be independent. Also, let* $\mathbf{l}$ *be a* k*-dimensional vector of constants. Then the stochastic representation of* $\mathbf{l}^{T}\mathbf{A}^{+}\mathbf{z}$ *is given by*

$$\mathbf{l}^{T}\mathbf{A}^{+}\mathbf{z} \stackrel{d}{=} \xi^{-1}\left(\mathbf{l}^{T}\mathbf{\Sigma}^{+}\boldsymbol{\mu} + \sqrt{\lambda\left(1 + \frac{(r-1)}{n-r+2}u\right)\mathbf{l}^{T}\mathbf{\Sigma}^{+}\mathbf{l}}\,z_0\right), \tag{15}$$

where $\xi \sim \chi^2_{n-r+1}$*,* $z_0 \sim \mathcal{N}(0,1)$*, and* $u \sim \mathcal{F}(r-1, n-r+2; \boldsymbol{\mu}^T\mathbf{R}_\mathbf{l}\boldsymbol{\mu}/\lambda)$ *(non-central* F*-distribution with* $r-1$ *and* $n-r+2$ *degrees of freedom and non-centrality parameter* $\boldsymbol{\mu}^T\mathbf{R}_\mathbf{l}\boldsymbol{\mu}/\lambda$*) with* $\mathbf{R}_\mathbf{l} = \boldsymbol{\Sigma}^+ - \boldsymbol{\Sigma}^+\mathbf{l}\mathbf{l}^T\boldsymbol{\Sigma}^+/\mathbf{l}^T\boldsymbol{\Sigma}^+\mathbf{l}$*;* ξ*,* u*, and* z_0 *are mutually independently distributed.*

Proof. From the proof of Lemma 1 we get the following stochastic representation of $\mathbf{l}^T\mathbf{A}^+\mathbf{z}$ expressed as

$$\mathbf{l}^T\mathbf{A}^+\mathbf{z} \stackrel{d}{=} \xi^{-1}\left(\mathbf{l}^T\boldsymbol{\Sigma}^{+1/2}\mathbf{y} + \sqrt{\frac{\mathbf{l}^T\boldsymbol{\Sigma}^+\mathbf{l}}{n-r+2}\eta} t_0\right),$$

where $\eta = \mathbf{y}^T(\mathbf{I}_k - \mathbf{Q})\mathbf{y}$ with $\mathbf{Q} = \boldsymbol{\Sigma}^{+1/2}\mathbf{l}\mathbf{l}^T\boldsymbol{\Sigma}^{+1/2}/\mathbf{l}^T\boldsymbol{\Sigma}^+\mathbf{l}$, $\xi \sim \chi^2_{n-r+1}$, $t_0 \sim t_{n-r+2}$ (t-distribution with $n-r+2$ degrees of freedom), and $\mathbf{y} \sim \mathcal{N}_k\left(\boldsymbol{\Sigma}^{+1/2}\boldsymbol{\mu}, \lambda\mathbf{P}\mathbf{P}^T\right)$.

Next we prove that $\lambda^{-1}\eta \sim \chi^2_{r-1;\delta^2}$, where $\delta^2 = \boldsymbol{\mu}^T\mathbf{R}_\mathbf{l}\boldsymbol{\mu}/\lambda$. First, we note that $\boldsymbol{\Sigma}^{+1/2} = \mathbf{P}\boldsymbol{\Delta}^{-1/2}\mathbf{P}^T$, where $\boldsymbol{\Delta}$ and $\mathbf{P}$ are the same as in the proof of Theorem 1, and $\boldsymbol{\Delta}^{-1/2} = diag(\delta_1^{-1/2}, \ldots, \delta_r^{-1/2})$. Second, we have that the matrix $(\mathbf{I}_k - \mathbf{Q})$ is an idempotent matrix, since $\mathbf{Q}$ is an idempotent. Furthermore,

(i) $\operatorname{tr}\left[\left(\lambda^{-1}(\mathbf{I}_k - \mathbf{Q})\right)\left(\lambda\mathbf{P}\mathbf{P}^T\right)\right] = \operatorname{tr}[\mathbf{P}^T\mathbf{P}] - \operatorname{tr}[\mathbf{Q}] = r-1$;

(ii) $\left(\lambda\mathbf{P}\mathbf{P}^T\right)\left(\lambda^{-1}(\mathbf{I}_k - \mathbf{Q})\right)\left(\lambda\mathbf{P}\mathbf{P}^T\right)\left(\lambda^{-1}(\mathbf{I}_k - \mathbf{Q})\right)\left(\lambda\mathbf{P}\mathbf{P}^T\right) = \left(\lambda\mathbf{P}\mathbf{P}^T\right)\left(\lambda^{-1}(\mathbf{I}_k - \mathbf{Q})\right)\left(\lambda\mathbf{P}\mathbf{P}^T\right)$;

(iii) $\boldsymbol{\mu}^T\boldsymbol{\Sigma}^{+1/2}\left(\lambda^{-1}(\mathbf{I}_k - \mathbf{Q})\right)\left(\lambda\mathbf{P}\mathbf{P}^T\right)\left(\lambda^{-1}(\mathbf{I}_k - \mathbf{Q})\right)\boldsymbol{\Sigma}^{+1/2}\boldsymbol{\mu} = \boldsymbol{\mu}^T\boldsymbol{\Sigma}^{+1/2}\left(\lambda^{-1}(\mathbf{I}_k - \mathbf{Q})\right)\boldsymbol{\Sigma}^{+1/2}\boldsymbol{\mu}$;

(iv) $\boldsymbol{\mu}^T\boldsymbol{\Sigma}^{+1/2}\left[\left(\lambda^{-1}(\mathbf{I}_k - \mathbf{Q})\right)\left(\lambda\mathbf{P}\mathbf{P}^T\right)\right]^2 = \boldsymbol{\mu}^T\boldsymbol{\Sigma}^{+1/2}\left(\lambda^{-1}(\mathbf{I}_k - \mathbf{Q})\right)\left(\lambda\mathbf{P}\mathbf{P}^T\right)$.

Now, the application of Theorem 5.1.3 of Mathai and Provost [12] shows that $\lambda^{-1}\eta$ is a χ^2_{r-1,δ^2} variable with $\delta^2 = \boldsymbol{\mu}^T\mathbf{R}_\mathbf{l}\boldsymbol{\mu}/\lambda$. Since $(\mathbf{I}_k - \mathbf{Q})\left(\lambda\mathbf{P}\mathbf{P}^T\right)\boldsymbol{\Sigma}^{+1/2}\mathbf{l} = \mathbf{0}$, the application of Theorem 5.4.3 in Mathai and Provost [12] proves that $\mathbf{l}^T\boldsymbol{\Sigma}^{+1/2}\mathbf{y}$ and η are independently distributed.

Finally, we note that the random variable $t_0 \sim t(n-r+2, 0, 1)$ has the following stochastic representation

$$t_0 \stackrel{d}{=} z_0\sqrt{\frac{n-r+2}{\zeta}},$$

where $z_0 \sim \mathcal{N}(0,1)$ and $\zeta \sim \chi^2_{n-r+2}$; z_0 and ζ are independent. Hence,

$$\begin{aligned}
\mathbf{l}^T\boldsymbol{\Sigma}^{+1/2}\mathbf{y} + \sqrt{\frac{\mathbf{l}^T\boldsymbol{\Sigma}^+\mathbf{l}}{n-r+2}\eta} t_0 | \eta, \zeta &\sim \mathcal{N}\left(\mathbf{l}^T\boldsymbol{\Sigma}^+\boldsymbol{\mu}, \lambda\mathbf{l}^T\boldsymbol{\Sigma}^+\mathbf{l}\left(1 + \frac{\lambda^{-1}\eta}{\zeta}\right)\right) \\
&= \mathcal{N}\left(\mathbf{l}^T\boldsymbol{\Sigma}^+\boldsymbol{\mu}, \lambda\mathbf{l}^T\boldsymbol{\Sigma}^+\mathbf{l}\left(1 + \frac{r-1}{n-r+2}u\right)\right),
\end{aligned}$$

where

$$u = \frac{\lambda^{-1}\eta/(r-1)}{\zeta/(n-r+2)} \sim \mathcal{F}(r-1, n-r+2, \boldsymbol{\mu}^T\mathbf{R}_\mathbf{l}\boldsymbol{\mu}/\lambda).$$

Putting all together we get the statement of the corollary. □

References

[1] Bodnar, T., Mazur, S., and Podgórski, K. (2016). Singular inverse Wishart distribution and its application to portfolio theory. Journal of Multivariate Analysis, 143:314–326.

[2] Bodnar, T., Mazur, S., and Podgórski, K. (2017). A test for the global minimum variance portfolio for small sample and singular covariance. AStA Advances in Statistical Analysis, 101:253–265.

[3] Bodnar, T. and Okhrin, Y. (2008). Properties of the singular, inverse and generalized inverse partitioned Wishart distributions. Journal of Multivariate Analysis, 99:2389– 2405.

[4] Bodnar, T. and Okhrin, Y. (2011). On the product of inverse Wishart and normal distributions with applications to discriminant analysis and portfolio theory. Scandinavian Journal of Statistics, 38:311–331.

[5] Bodnar, T. and Reiß, M. (2016). Exact and asymptotic tests on a factor model in low and large dimensions with applications. Journal of Multivariate Analysis, 150:125–151.

[6] Chamberlain, G. (1983). Funds, factors, and diversification in arbitrage pricing models. Econometrica, 51:1305–1323.

[7] Fan, J., Fan, Y., and Lv, J. (2008). High dimensional covariance matrix estimation using a factor model. Journal of Econometrics, 147:186–197.

[8] Fan, J., Liao, Y., and Mincheva, M. (2013). Large covariance estimation by thresholding principal orthogonal complements. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 75:603–680.

[9] Fan, J., Zhang, J., and Yu, K. (2012). Vast portfolio selection with gross-exposure constraints. Journal of the American Statistical Association, 107:592–606.

[10] Gupta, A. and Nagar, D. (2000). Matrix variate distributions. Chapman and Hall/CRC, Boca Raton.

[11] Markowitz, H. (1952). Mean-variance analysis in portfolio choice and capital markets. Journal of Finance, 7:77–91.

[12] Mathai, A. and Provost, S. B. (1992). Quadratic Forms in Random Variables. Marcel Dekker.

[13] Nadakuditi, R. R. and Eldeman, A. (2008). Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples. IEEE Trans. Signal Processing, 56:2625–2638.

[14] Pappas, D., Kiriakopoulos, K., and Kaimakamis, G. (2010). Optimal portfolio selection with singular covariance matrix. International Mathematical Forum, 5:2305–2318.

[15] Pruzek, R. (1994). High dimensional covariance estimation: Avoiding the 'curse of dimensionality'. In et al. (eds), B. H., editor, Proceedings of the First US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach., Dordrecht. Springer.

[16] Stein, C., Efron, B., and Morris, C. (1972). Improving the usual estimator of a normal covariance matrix,. Technical Report 37, Dept. of Statistics, Stanford University.