Mathematical Statistics
Stockholm University

Optimal shrinkage-based portfolio selection in high dimensions

Taras Bodnar
Yarema Okhrin
Nestor Parolya

Research Report 2017:1

ISSN 1650-0377

Postal address:

Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

Internet:

http://www.math.su.se

Mathematical Statistics
Stockholm University
Research Report **2017:1**,
http://www.math.su.se

Optimal shrinkage-based portfolio selection in high dimensions

February 2017

Taras Bodnar[a], Yarema Okhrin[b] and Nestor Parolya[c, 1]

[a] *Department of Mathematics, Stockholm University, Roslagsvägen 101, SE-10691 Stockholm, Sweden*

[b] *Department of Statistics, University of Augsburg, Universitätsstr. 16, D-86159 Augsburg, Germany*

[c] *Institute of Statistics, Leibniz University Hannover, Königsworther Platz 1, D-30167 Hannover, Germany*

Abstract

In this paper we estimate the mean-variance (MV) portfolio in the high-dimensional case using the recent results from the theory of random matrices. We construct a linear shrinkage estimator which is distribution-free and is optimal in the sense of maximizing with probability 1 the asymptotic out-of-sample expected utility, i.e., mean-variance objective function. Its asymptotic properties are investigated when the number of assets p together with the sample size n tend to infinity such that $p/n \to c \in (0, +\infty)$. The results are obtained under weak assumptions imposed on the distribution of the asset returns, namely the existence of the fourth moments is only required. Thereafter we perform numerical and empirical studies where the small- and large-sample behavior of the derived estimator is investigated. The suggested estimator shows significant improvements over the naive diversification and it is robust to the deviations from normality.

JEL Classification: G11, C13, C14, C58, C65
Keywords: expected utility portfolio, large-dimensional asymptotics, covariance matrix estimation, random matrix theory.

1 Introduction

In the seminal paper of Markowitz (1952) the author suggests to determine the optimal composition of a portfolio of financial assets by minimizing the portfolio variance assuming that the

[1]Corresponding author. E-mail address: parolya@statistik.uni-hannover.de

expected portfolio return attains some prespecified fixed value. By varying this value we obtain the whole efficient frontier in the mean-standard deviation space. Despite of its simplicity, this approach justifies the advantages of diversification and is a standard technique and benchmark in asset management. Equivalently (see, Tobin (1958), Bodnar et al. (2013)) we can obtain the same portfolios by maximizing the expected quadratic utility (EU) with the optimization problem given by

$$\mathbf{w}'\boldsymbol{\mu}_n - \frac{\gamma}{2}\mathbf{w}'\boldsymbol{\Sigma}_n\mathbf{w} \to max \quad \text{subject to} \quad \mathbf{w}'\mathbf{1}_p = 1\,, \tag{1.1}$$

where $\mathbf{w} = (\omega_1, \ldots, \omega_p)'$ is the vector of portfolio weights, $\mathbf{1}_p$ is the p-dimensional vector of ones, $\boldsymbol{\mu}_n$ and $\boldsymbol{\Sigma}_n$ are the p-dimensional mean vector and the $p \times p$ covariance matrix of asset returns, respectively. The quantity $\gamma > 0$ determines the investor's behavior towards risk. It must be noted that the maximization of the mean-variance objective function (1.1) is equivalent to the maximization of the exponential utility (CARA) function under the assumption of normality of the asset returns. In this case γ equals the investor's absolute risk aversion coefficient (see, e.g., Pratt (1964)).

The solution of the optimization problem (1.1) is well known and it is given by

$$\mathbf{w}_{EU} = \mathbf{w}_{GMV} + \gamma^{-1}\mathbf{Q}_n\boldsymbol{\mu}_n\,, \tag{1.2}$$

where

$$\mathbf{Q}_n = \boldsymbol{\Sigma}_n^{-1} - \frac{\boldsymbol{\Sigma}_n^{-1}\mathbf{1}_p\mathbf{1}_p'\boldsymbol{\Sigma}_n^{-1}}{\mathbf{1}_p'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}_p} \tag{1.3}$$

and

$$\mathbf{w}_{GMV} = \frac{\boldsymbol{\Sigma}_n^{-1}\mathbf{1}_p}{\mathbf{1}_p'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}_p} \tag{1.4}$$

is the vector of the weights of the global minimum variance portfolio.

In practice, however, the above mentioned approach of constructing an optimal portfolio frequently shows poor out-of-sample performance in terms of various performance measures. Even naive portfolio strategies, e.g., equally weighted portfolio (see, DeMiguel et al. (2009)), often outperform the mean-variance strategy. One of the reasons is the estimation risk. The unknown parameters $\boldsymbol{\mu}_n$ and $\boldsymbol{\Sigma}_n$ have to be estimated using historical data on asset returns. This results in the "plug-in" estimator of the EU portfolio (1.2) which is a traditional and simple way to evaluate the portfolio in practice. This estimator is constructed by replacing in (1.2) the mean vector $\boldsymbol{\mu}_n$ and the covariance matrix $\boldsymbol{\Sigma}_n$ with their sample counterparts. Okhrin and Schmid (2006) derive the expected return and the variance of the sample portfolio weights under the assumption that the asset returns follow a multivariate normal distribution, while Bodnar and Schmid (2011) obtain the exact finite-sample distribution. Recently, Bodnar et al. (2016) extended these results to the case $n < p$.

The estimation of the parameters has a negativ impact on the performance of the asset allocation strategy. This was noted in a series of papers with Merton (1980), Best and Grauer (1991), Chopra and Ziemba (1993) among others. Several approaches have arisen to reduce the consequences of estimation risk. One strand of research opted for the Bayesian framework and using appropriate priors take the estimation risk into account already while building the portfolio. The second strand relied on the shrinkage techniques and is related to the method exploited in this paper. A straightforward way to improve the properties of the estimators for $\boldsymbol{\mu}_n$ and $\boldsymbol{\Sigma}_n$ is to use the shrinkage approach (see, Jorion (1986), Ledoit and Wolf (2004)). Alternatively, one may apply the shrinkage estimation to the portfolio weights directly. Golosnoy and Okhrin (2007) consider the multivariate shrinkage estimator by shrinking the portfolios with and without the riskless asset to an arbitrary static portfolio. A similar technique is used by Frahm and Memmel (2010), who constructed a feasible shrinkage estimator for the GMV portfolio which dominates the traditional one. At last, Bodnar et al. (2017) suggest a shrinkage estimator for the GMV portfolio which is feasible even for the singular sample covariance matrix.

An important issue nowadays is, however, the asset allocation for large portfolios. The sample estimators work well only in the case when the number of assets p is fixed and substantially smaller than the sample size n. This case is known as the standard asymptotics in statistics (see, Le Cam and Lo Yang (2000)). Under this asymptotics the traditional sample estimator is a consistent estimator for the EU portfolio. But what happens when the dimension p and the sample size n are comparable of size, say $p = 900$ and $n = 1000$? Technically, here we are in the situation when both the number of assets p and the sample size n tend to infinity. In the case when p/n tends to some concentration ratio $c > 0$ this asymptotics is known as high-dimensional asymptotics or "Kolmogorov" asymptotics (see, e.g., Bai and Silverstein (2010)). If c is close to one the sample covariance matrix tends to be close to a singular one and when $c > 1$ it becomes singular. The sample estimator of the EU portfolio behaves badly in this case both from the theoretical and practical points of view (see, e.g., El Karoui (2010)). The sample covariance matrix is very unstable and tends to under- or overestimate the true parameters for c smaller but close to 1 (see, Bai and Shi (2011)). For $c > 1$ the inverse sample covariance does not exist and the portfolio cannot be constructed in the traditional way. Taking the above mentioned information into account the aim of the paper is to construct a feasible and simple shrinkage estimator of the EU portfolio which is optimal in an asymptotic sense and is additionally distribution-free. The estimator is developed using the fast growing branch of probability theory, namely random matrix theory. The main result of this theory is proved by Marčenko and Pastur (1967) and further extended under very general conditions by Silverstein (1995). Now it is called Marčenko-Pastur equation. Its importance arises in many areas of science because it shows how the true covariance matrix and its sample estimator are connected asymptotically. Knowing this we can build suitable estimators for high-dimensional quantities which depend on $\boldsymbol{\Sigma}_n$. In our case this refers to the shrinkage intensities. Note however, that the optimal shrinkage intensity depends again on the unknown characteristics of the asset returns. To overcome this problem we derive consistent estimators for specific functions (quadratic and bilinear forms) of the inverse sample covariance matrix and succeed to provide consistent estimators for the optimal shrinkage intensities.

The rest of paper is organized as follows. In the next section, we construct a shrinkage estimator for the optimal portfolio weights obtained by shrinking the mean-variance weights to an arbitrary target portfolio. The oracle shrinkage intensity and the corresponding feasible bona-fide estimators for $c < 1$ and $c > 1$ are established as well. The derived results are evaluated in Section 3 in extensive simulation and empirical studies. All proofs are moved to in the Appendix.

2 Optimal shrinkage estimator of mean-variance portfolio

Let $\mathbf{Y}_n = (\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n)$ be the $p \times n$ data matrix which consist of n vectors of the returns on $p \equiv p(n)$ assets. Let $E(\mathbf{y}_i) = \boldsymbol{\mu}_n$ and $Cov(\mathbf{y}_i) = \mathbf{\Sigma}_n$ for $i \in 1, ..., n$. We assume that $p/n \to c \in (0, +\infty)$ as $n \to \infty$. This type of limiting behavior is known as "the large dimensional asymptotics" or "the Kolmogorov asymptotics". In this case the traditional sample estimators perform poor or even very poor and tend to over/underestimate the unknown parameters of the asset returns, e.g., the mean vector and the covariance matrix.

Throughout the paper it is assumed that there exists a $p \times n$ random matrix $\mathbf{X}_n$ which consists of independent and identically distributed (i.i.d.) real random variables with zero mean and unit variance such that

$$\mathbf{Y}_n = \boldsymbol{\mu}_n \mathbf{1}_n' + \mathbf{\Sigma}_n^{\frac{1}{2}} \mathbf{X}_n \,. \tag{2.1}$$

It must be noted that the observation matrix $\mathbf{Y}_n$ has the dependent rows but independent columns. Broadly speaking, this means that we allow arbitrary cross-sectional correlations of the asset returns but assume their independence over time. Although this assumption looks quite restrictive for financial applications, there exist stronger results from random matrix theory which show that the model can be extended to (weakly) depending variables by demanding more complicated conditions on the elements of $\mathbf{Y}_n$ (see, Bai and Zhou (2008)) or by controlling the number of dependent entries as dimension increases (see, Hui and Pan (2010), Friesen et al. (2013), Wei et al. (2016)). Nevertheless, this will only make the proofs more technical, but leave the results unchanged. For that reason we assume independent asset returns over time only to simplify the proofs of the main theorems.

The three assumptions which are used throughout the paper are

(A1) The covariance matrix of the asset returns $\mathbf{\Sigma}_n$ is a nonrandom p-dimensional positive definite matrix.

(A2) The elements of the matrix $\mathbf{X}_n$ have uniformly bounded $4 + \varepsilon$ moments for some $\varepsilon > 0$.

(A3) There exist $M_l, M_u \in (0, +\infty)$ such that $M_l \leq \mathbf{1}_p' \mathbf{\Sigma}_n^{-1} \mathbf{1}_p, \quad \boldsymbol{\mu}_n' \mathbf{\Sigma}_n^{-1} \boldsymbol{\mu}_n \leq M_u$.

All of these regularity assumptions are general enough to fit many real world situations. The assumption (A1) together with (2.1) are usual for financial and statistical problems and impose no strong restrictions. The assumption (A2) is a technical one and can be relaxed for practical purposes (see, section with simulations). The assumption (A3) requires that the quantities which are used in the calculations are finite. This assumption is quite general and imposes no additional constrains neither on the mean vector $\boldsymbol{\mu}_n$, like its Euclidean norm is bounded, nor on the covariance matrix $\boldsymbol{\Sigma}_n$, like its eigenvalues lie in the compact interval. The last point allows us to assume a factor model for the data matrix $\mathbf{Y}_n$ which implies that the largest eigenvalue of $\boldsymbol{\Sigma}_n$ is of order p (c.f. Fan et al. (2008), Fan et al. (2012), Fan et al. (2013)). Finally, assumption (A3) ensures that $\mathbf{1}_p'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\mu}_n$ is finite as well which follows directly from the Cauchy-Schwarz inequality. This implies that the variance and the expected return of the global minimum variance portfolio are uniformly bounded in p.

The sample covariance matrix is given by

$$\mathbf{S}_n = \frac{1}{n}\mathbf{Y}_n(\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n')\mathbf{Y}_n' = \frac{1}{n}\boldsymbol{\Sigma}_n^{\frac{1}{2}}\mathbf{X}_n(\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n')\mathbf{X}_n'\boldsymbol{\Sigma}_n^{\frac{1}{2}}\,, \tag{2.2}$$

where the symbol $\mathbf{I}_n$ stands for the n-dimensional identity matrix. The sample mean vector becomes

$$\bar{\mathbf{y}}_n = \frac{1}{n}\mathbf{Y}_n\mathbf{1}_n = \boldsymbol{\mu}_n + \boldsymbol{\Sigma}_n^{\frac{1}{2}}\bar{\mathbf{x}}_n \;\; \text{with } \bar{\mathbf{x}}_n = \frac{1}{n}\mathbf{X}_n\mathbf{1}_n\,. \tag{2.3}$$

2.1 Oracle estimator. Case $c < 1$

In this section we consider the optimal shrinkage estimator for the EU portfolio weights presented in the introduction by optimizing the shrinkage parameter α and fixing some target portfolio $\mathbf{b}$.

The resulting estimator for $c < 1$ is given by

$$\hat{\mathbf{w}}_{GSE} = \alpha_n\hat{\mathbf{w}}_S + (1-\alpha_n)\mathbf{b} \;\; \text{with } \mathbf{b}'\mathbf{1}_p = 1 \text{ and } \mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{b} \le \infty\,, \tag{2.4}$$

where the vector $\hat{\mathbf{w}}_S$ is the sample estimator of the EU portfolio given in (1.2), namely

$$\hat{\mathbf{w}}_S = \frac{\mathbf{S}_n^{-1}\mathbf{1}_p}{\mathbf{1}_p'\mathbf{S}_n^{-1}\mathbf{1}_p} + \gamma^{-1}\hat{\mathbf{Q}}_n\bar{\mathbf{y}}_n \tag{2.5}$$

with

$$\hat{\mathbf{Q}}_n = \mathbf{S}_n^{-1} - \frac{\mathbf{S}_n^{-1}\mathbf{1}_p\mathbf{1}_p'\mathbf{S}_n^{-1}}{\mathbf{1}_p'\mathbf{S}_n^{-1}\mathbf{1}_p}\,. \tag{2.6}$$

The target portfolio $\mathbf{b} \in \mathbb{R}^p$ is a given nonrandom (and further random independent of $\mathbf{Y}_n$) vector with $\mathbf{b}'\mathbf{1}_p = 1$ and uniformly bounded variance. No assumption is imposed on the shrinkage intensity α_n which is the object of our interest.

The aim is now to find the optimal shrinkage intensity for a given nonrandom target portfolio $\mathbf{b}$ which maximizes the mean-variance objective function (1.1). The maximization problem is given by

$$U = \hat{\mathbf{w}}'_{GSE}(\alpha_n)\boldsymbol{\mu}_n - \frac{\gamma}{2}\hat{\mathbf{w}}'_{GSE}(\alpha_n)\boldsymbol{\Sigma}_n\hat{\mathbf{w}}_{GSE}(\alpha_n) \longrightarrow max \ \text{ with respect to } \ \alpha_n\,. \tag{2.7}$$

The expected utility function (2.7) can be rewritten as

$$U = \alpha_n\hat{\mathbf{w}}'_S\boldsymbol{\mu}_n + (1-\alpha_n)\mathbf{b}'\boldsymbol{\mu}_n - \frac{\gamma}{2}\left(\alpha_n^2\hat{\mathbf{w}}'_S\boldsymbol{\Sigma}_n\hat{\mathbf{w}}_S + 2\alpha_n(1-\alpha_n)\mathbf{b}'\boldsymbol{\Sigma}_n\hat{\mathbf{w}}_S + (1-\alpha_n)^2\mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{b}\right) \to max$$
$$\text{with respect to } \ \alpha_n\,. \tag{2.8}$$

Next, taking the derivative of U with respect to α_n and setting it equal to zero we get

$$\frac{\partial U}{\partial \alpha_n} = (\hat{\mathbf{w}}_S - \mathbf{b})'\boldsymbol{\mu}_n - \gamma\left(\alpha_n\hat{\mathbf{w}}'_S\boldsymbol{\Sigma}_n\hat{\mathbf{w}}_S + (1-2\alpha_n)\mathbf{b}'\boldsymbol{\Sigma}_n\hat{\mathbf{w}}_S - (1-\alpha_n)\mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{b}\right)\,.$$

From the last equation it is easy to find the optimal shrinkage intensity α_n^* given by

$$\alpha_n^* = \gamma^{-1}\frac{(\hat{\mathbf{w}}_S - \mathbf{b})'(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b})}{(\hat{\mathbf{w}}_S - \mathbf{b})'\boldsymbol{\Sigma}_n(\hat{\mathbf{w}}_S - \mathbf{b})}\,. \tag{2.9}$$

To ensure that α_n^* is the unique maximizer of (2.7) the second derivative of U must be negative, which is always fulfilled. Indeed, it follows from the positive definitiveness of the matrix $\boldsymbol{\Sigma}_n$, namely

$$\frac{\partial^2 U}{\partial \alpha_n^2} = -\gamma(\hat{\mathbf{w}}_S - \mathbf{b})'\boldsymbol{\Sigma}_n(\hat{\mathbf{w}}_S - \mathbf{b}) < 0\,. \tag{2.10}$$

In the next theorem we show the asymptotic properties of the optimal shrinkage intensity α_n^* under large dimensional asymptotics.

Theorem 2.1. *Assume (A1)-(A3). Then it holds that*

$$|\alpha_n^* - \alpha^*| \xrightarrow{a.s.} 0 \ \ \textit{for} \ \ \frac{p}{n} \to c \in (0,1) \ \ \textit{as} \ \ n \to \infty$$

with

$$\alpha^* = \gamma^{-1}\frac{(R_{GMV} - R_b)\left(1 + \dfrac{1}{1-c}\right) + \gamma(V_b - V_{GMV}) + \dfrac{\gamma^{-1}}{1-c}s}{\dfrac{1}{1-c}V_{GMV} - 2\left(V_{GMV} + \frac{\gamma^{-1}}{1-c}(R_b - R_{GMV})\right) + \gamma^{-2}\left(\dfrac{s}{(1-c)^3} + \dfrac{c}{(1-c)^3}\right) + V_b}\,, \tag{2.11}$$

where $R_{GMV} = \dfrac{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\mu}_n}{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}}$ *and* $V_{GMV} = \dfrac{1}{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}}$ *are the expected return and the variance of the true global minimum variance portfolio, while* $s = \boldsymbol{\mu}_n'\mathbf{Q}_n\boldsymbol{\mu}_n$ *is the slope parameter of the efficient frontier. The quantities* $R_b = \mathbf{b}'\boldsymbol{\mu}_n$ *and* $V_b = \mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{b}$ *denote the expected return and the variance of the target portfolio* $\mathbf{b}$.

Next we assess the performance of the classical estimator of the portfolio weights $\hat{\mathbf{w}}_S$ and the optimal shrinkage weights $\hat{\mathbf{w}}_{GSE}$. As a measure of performance we consider the relative increase in the utility of the portfolio return compared to the portfolio based on true parameters of asset returns. The results are summarized in the following corollary.

Corollary 2.1. *(a) Let U_{EU} and U_S be the expected quadratic utilities for the true EU portfolio and its traditional estimator. Then under the assumptions of Theorem 2.1, the relative loss of the traditional estimator of the EU portfolio is given by*

$$R_S = \frac{U_{EU} - U_S}{U_{EU}} \xrightarrow{a.s.} \frac{\frac{\gamma}{2}\left(\frac{1}{1-c} - 1\right) \cdot V_{GMV} + \gamma^{-1}\left(-\frac{1}{2} - \frac{1}{(1-c)} + \frac{1}{2(1-c)^3}\right) \cdot s + \frac{\gamma^{-1}}{2} \cdot \frac{c}{(1-c)}}{R_{GMV} - \frac{\gamma^{-1}}{2} \cdot s - \frac{\gamma}{2} V_{GMV}} \tag{2.12}$$

for $\frac{p}{n} \to c \in (0,1)$ as $n \to \infty$.

(b) Let U_{GSE} be the expected quadratic utility for optimal shrinkage estimator of the EU portfolio. Under the assumptions of Theorem 2.1, the relative loss of the optimal shrinkage estimator is given by

$$R_{GSE} = \frac{U_{EU} - U_{GSE}}{U_{EU}} \xrightarrow{a.s.} (\alpha^*)^2 R_S + (1-\alpha^*)^2 R_{\mathbf{b}} \quad \textit{for} \quad \frac{p}{n} \to c \in (0,1) \ \textit{as} \ n \to \infty. \tag{2.13}$$

2.2 Oracle estimator. Case $c > 1$.

Here, similarly as in Bodnar et al. (2017), we will use the generalized inverse of the sample covariance matrix $\mathbf{S}_n$. Particularly, we use the following generalized inverse of the sample covariance matrix $\mathbf{S}_n$

$$\mathbf{S}_n^* = \mathbf{\Sigma}_n^{-1/2} \left(\frac{1}{n}\mathbf{X}_n\mathbf{X}_n' - \bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\right)^+ \mathbf{\Sigma}_n^{-1/2}, \tag{2.14}$$

where $'+'$ denotes the Moore-Penrose inverse. It can be shown that $\mathbf{S}_n^*$ is a generalized inverse of $\mathbf{S}_n$ satisfying $\mathbf{S}_n^*\mathbf{S}_n\mathbf{S}_n^* = \mathbf{S}_n^*$ and $\mathbf{S}_n\mathbf{S}_n^*\mathbf{S}_n = \mathbf{S}_n$. However, $\mathbf{S}_n^*$ is not exactly equal to the Moore-Penrose inverse because it does not satisfy the conditions $(\mathbf{S}_n^*\mathbf{S}_n)' = \mathbf{S}_n^*\mathbf{S}_n$ and $(\mathbf{S}_n\mathbf{S}_n^*)' = \mathbf{S}_n\mathbf{S}_n^*$. In case $c < 1$ the generalized inverse $\mathbf{S}_n^*$ coincides with the usual inverse $\mathbf{S}_n^{-1}$. Moreover, if $\mathbf{\Sigma}_n$ is a multiple of identity matrix then $\mathbf{S}_n^*$ is equal to the Moore-Penrose inverse $\mathbf{S}_n^+$. In this section, $\mathbf{S}_n^*$ is used only to determine an oracle estimator for the weights of the MV portolio. The bona fide estimator is constructed in the next section.

Thus, the oracle estimator for $c > 1$ is given by

$$\hat{\mathbf{w}}_{GSE}^* = \alpha_n^+ \hat{\mathbf{w}}_{S^*} + (1-\alpha_n^+)\mathbf{b} \ \text{ with } \mathbf{b}'\mathbf{1}_p = 1 \text{ and } \mathbf{b}'\mathbf{\Sigma}_n\mathbf{b} < \infty, \tag{2.15}$$

where the vector $\hat{\mathbf{w}}_{S^*}$ is the sample estimator of the EU portfolio given in (1.2), namely

$$\hat{\mathbf{w}}_{S^*} = \frac{\mathbf{S}_n^*\mathbf{1}_p}{\mathbf{1}_p'\mathbf{S}_n^*\mathbf{1}_p} + \gamma^{-1}\hat{\mathbf{Q}}_n^*\bar{\mathbf{y}}_n \tag{2.16}$$

with

$$\hat{\mathbf{Q}}_n^* = \mathbf{S}_n^* - \frac{\mathbf{S}_n^*\mathbf{1}_p\mathbf{1}_p'\mathbf{S}_n^*}{\mathbf{1}_p'\mathbf{S}_n^*\mathbf{1}_p}. \tag{2.17}$$

Again, the shrinkage intensity α_n^+ is the object of our interest. In order to save place we skip the optimization procedure for α_n^+ as it is only slightly different from the case $c < 1$. Thus, the optimal shrinkage intensity α_n^+ in case $c > 1$ is given by

$$\alpha_n^+ = \gamma^{-1}\frac{(\hat{\mathbf{w}}_{S^*} - \mathbf{b})'(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b})}{(\hat{\mathbf{w}}_{S^*} - \mathbf{b})'\boldsymbol{\Sigma}_n(\hat{\mathbf{w}}_{S^*} - \mathbf{b})}\,. \tag{2.18}$$

In the next theorem we find the asymptotic equivalent quantity for α_n^+ in the case $p/n \to c \in (1, +\infty)$ as $n \to \infty$.

Theorem 2.2. *Assume (A1)-(A3). Then it holds that*

$$\left|\alpha_n^+ - \alpha^+\right| \xrightarrow{a.s.} 0 \quad for \quad \frac{p}{n} \to c \in (1, +\infty) \quad as \quad n \to \infty$$

with

$$\alpha^+ = \gamma^{-1}\frac{(R_{GMV} - R_b)\left(1 + \dfrac{1}{c(c-1)}\right) + \gamma(V_b - V_{GMV}) + \dfrac{\gamma^{-1}}{c(c-1)}s}{\dfrac{c^2}{(c-1)}V_{GMV} - 2\left(V_{GMV} + \frac{\gamma^{-1}}{c(c-1)}(R_b - R_{GMV})\right) + \dfrac{\gamma^{-2}}{(c-1)^3}(s+c) + V_b}\,, \tag{2.19}$$

where $R_{GMV} = \dfrac{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\mu}_n}{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}}$ *and* $V_{GMV} = \dfrac{1}{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}}$ *are the expected return and the variance of the true global minimum variance portfolio, while* $s = \boldsymbol{\mu}_n'\boldsymbol{\Sigma}_n\boldsymbol{\mu}_n$ *is the slope parameter of the efficient frontier. The quantities* $R_b = \mathbf{b}'\boldsymbol{\mu}_n$ *and* $V_b = \mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{b}$ *denote the expected return and the variance of the target portfolio* $\mathbf{b}$.

Similarly as for the case $c < 1$ we provide here the expression for the relative losses.

Corollary 2.2. *(a) Let* U_{EU} *and* U_S *be the expected quadratic utilities for the true EU portfolio and its traditional estimator. Then under the assumptions of Theorem 2.2, the relative loss of the traditional estimator of the EU portfolio is given by*

$$R_S = \frac{U_{EU} - U_S}{U_{EU}} \xrightarrow{a.s.} \frac{\frac{\gamma}{2}\left(\frac{c^2}{c-1} - 1\right)\cdot V_{GMV} + \gamma^{-1}\left(-\frac{1}{2} - \frac{1}{c(c-1)} + \frac{1}{2(c-1)^3}\right)\cdot s + \frac{\gamma^{-1}}{2}\cdot\frac{c}{(1-c)^3}}{R_{GMV} - \frac{\gamma^{-1}}{2}\cdot s - \frac{\gamma}{2}V_{GMV}} \tag{2.20}$$

for $\frac{p}{n} \to c \in (1, +\infty)$ *as* $n \to \infty$.

(b) Let U_{GSE} *be the expected quadratic utility for the optimal shrinkage estimator of the EU portfolio. Under the assumptions of Theorem 2.2, the relative loss of the optimal shrinkage estimator is given by*

$$R_{GSE} = \frac{U_{EU} - U_{GSE}}{U_{EU}} \xrightarrow{a.s.} (\alpha^+)^2 R_S + (1-\alpha^+)^2 R_{\mathbf{b}_n} \quad for \quad \frac{p}{n} \to c \in (1, +\infty) \ as \ n \to \infty. \tag{2.21}$$

2.3 Estimation of unknown parameters. Bona fide estimator

The limiting shrinkage intensity α^* is not feasible in practice, since the quantities it depends on, i.e. R_{GMV}, V_{GMV}, s, R_b and V_b, are unknown. In this subsection we derive the corresponding consistent estimators. The results are summarized in two propositions dealing with the cases $c \in (0,1)$ and $c \in (1,\infty)$ respectively. The statements follow directly from the proofs of Theorems 2.1 and 2.2.

Proposition 2.1. *The consistent estimators of R_{GMV}, V_{GMV}, s, R_b and V_b under large dimensional asymptotics $p/n \to c < 1$ as $n \to \infty$ are given by*

$$\begin{aligned}
\hat{R}_c &= \hat{R}_{GMV} \xrightarrow{a.s.} R_{GMV} && (2.22)\\
\hat{V}_c &= \frac{1}{1-p/n}\hat{V}_{GMV} \xrightarrow{a.s.} V_{GMV} && (2.23)\\
\hat{s}_c &= (1-p/n)\hat{s} - p/n \xrightarrow{a.s.} s && (2.24)\\
\hat{R}_b &= \mathbf{b}'\bar{\mathbf{y}}_n \xrightarrow{a.s.} R_b && (2.25)\\
\hat{V}_b &= \mathbf{b}'\mathbf{S}_n\mathbf{b} \xrightarrow{a.s.} V_b\,, && (2.26)
\end{aligned}$$

where $\hat{R}_{GMV}$, $\hat{V}_{GMV}$ and $\hat{s}$ are traditional plug-in estimators.

Using Proposition 2.1 we can immediately construct a bona-fide estimator for expected utility portfolio weights in case $c < 1$. It holds that

$$\hat{\mathbf{w}}_{BFGSE} = \widehat{\alpha}^* \left(\frac{\mathbf{S}_n^{-1}\mathbf{1}_p}{\mathbf{1}_p'\mathbf{S}_n^{-1}\mathbf{1}_p} + \gamma^{-1}\hat{\mathbf{Q}}_n\bar{\mathbf{y}}_n \right) + (1-\widehat{\alpha}^*)\mathbf{b} \tag{2.27}$$

with

$$\widehat{\alpha}^* = \gamma^{-1}\frac{(\hat{R}_c - \hat{R}_b)\left(1 + \dfrac{1}{1-p/n}\right) + \gamma(\hat{V}_b - \hat{V}_c) + \dfrac{\gamma^{-1}}{1-p/n}\hat{s}_c}{\dfrac{1}{1-p/n}\hat{V}_c - 2\left(\hat{V}_c + \frac{\gamma^{-1}}{1-p/n}(\hat{R}_b - \hat{R}_c)\right) + \gamma^{-2}\left(\dfrac{\hat{s}_c}{(1-p/n)^3} + \dfrac{p/n}{(1-p/n)^3}\right) + \hat{V}_b} \tag{2.28}$$

where $\hat{R}_c$, $\hat{V}_c$, $\hat{s}_c$, $\hat{R}_b$ and $\hat{V}_b$ are given above in (2.1). The expression (2.27) is the optimal shrinkage estimator for a given target portfolio $\mathbf{b}$ in sense that the shrinkage intensity $\widehat{\alpha}^*$ tends almost surely to its optimal value α^* for $p/n \to c \in (0,1)$ as $n \to \infty$.

The situation is more complex in case $c > 1$. Here we can present only oracle estimators for the unknown quantities R_{GMV}, V_{GMV} and s.

Proposition 2.2. *The consistent oracle estimators of R_{GMV}, V_{GMV}, s under large dimensional asymptotics $p/n \to c > 1$ as $n \to \infty$ are given by*

$$\begin{aligned}
\hat{R}_c^o &= \hat{R}_{GMV} \xrightarrow{a.s.} R_{GMV}\\
\hat{V}_c^o &= \frac{1}{p/n(p/n-1)}\hat{V}_{GMV} \xrightarrow{a.s.} V_{GMV}\\
\hat{s}_c^o &= p/n[(p/n-1)\hat{s} - 1] \xrightarrow{a.s.} s\,,
\end{aligned}$$

where $\hat{R}_{GMV}$, $\hat{V}_{GMV}$ and $\hat{s}$ are the traditional plug-in estimators.

Here, the quantities from Proposition 2.2 are not the bona fide estimators, since the matrix $\mathbf{S}_n^*$ depends on the unknown quantities. Thus, we propose a reasonable approximation using the application of the Moore-Penrose inverse $\mathbf{S}_n^+$. It is easy to verify that in case of $\mathbf{\Sigma}_n = \sigma^2\mathbf{I}_p$ for any $\sigma > 0$ the considered approximation becomes the exact one.

Taking into account the above discussion and the result of Theorem 2.2, the bona fide estimators of the quantities R_{GMV}, V_{GMV} and x in case $c > 1$ is approximated by

$$\hat{R}_c^+ \approx \frac{\bar{\mathbf{y}}_n'\mathbf{S}_n^+\mathbf{1}_p}{\mathbf{1}_p'\mathbf{S}_n^+\mathbf{1}_p} \quad \text{for} \quad c \in (1, +\infty) \tag{2.29}$$

$$\hat{V}_c^+ \approx \frac{1}{p/n(p/n-1)}\frac{1}{\mathbf{1}_p'\mathbf{S}_n^+\mathbf{1}_p} \quad \text{for} \quad c \in (1, +\infty) \tag{2.30}$$

$$\hat{s}_c^+ \approx p/n[(p/n-1)\bar{\mathbf{y}}_n'\mathbf{Q}_n^+\bar{\mathbf{y}}_n - 1] \quad \text{for} \quad c \in (1, +\infty)\,. \tag{2.31}$$

The application of (2.29) leads to the bona fide optimal shrinkage estimator of the GMV portfolio in case $c > 1$ expressed as

$$\hat{\mathbf{w}}_{BFGSE}^+ = \hat{\alpha}^+ \left(\frac{\mathbf{S}_n^+\mathbf{1}_p}{\mathbf{1}_p'\mathbf{S}_n^+\mathbf{1}_p} + \gamma^{-1}\hat{\mathbf{Q}}_n^+\bar{\mathbf{y}}_n \right) + (1 - \hat{\alpha}^+)\mathbf{b}_n\,, \tag{2.32}$$

with

$$\hat{\alpha}^+ = \gamma^{-1} \frac{(\hat{R}_c^+ - \hat{R}_b)\left(1 + \frac{1}{p/n(p/n-1)}\right) + \gamma(\hat{V}_b - \hat{V}_c^+) + \frac{\gamma^{-1}}{p/n(p/n-1)}\hat{s}_c^+}{\frac{(p/n)^2}{p/n-1}\hat{V}_c^+ - 2\left(\hat{V}_c^+ + \frac{\gamma^{-1}}{p/n(p/n-1)}(\hat{R}_b - \hat{R}_c^+)\right) + \frac{\gamma^{-2}}{(p/n-1)^3}\left(\hat{s}_c^+ + p/n\right) + \hat{V}_b}\,, \tag{2.33}$$

where $\hat{R}_b$ and $\hat{V}_b$ are given in (2.25) and (2.26), respectively; $\mathbf{Q}_n^+ = \mathbf{S}_n^+ - \dfrac{\mathbf{S}_n^+\mathbf{1}\mathbf{1}'\mathbf{S}_n^+}{\mathbf{1}'\mathbf{S}_n^+\mathbf{1}}$ and $\mathbf{S}_n^+$ is the Moore-Penrose pseudo-inverse of the sample covariance matrix $\mathbf{S}_n$.

3 Simulation and empirical studies

In this section we illustrate the performance and the advantages of the derived results using simulated and real data. Particularly we address the estimation precision of the shrinkage coefficient and compare the traditional estimator with the asymptotic intensity and its consistent estimator.

3.1 Simulation study

For simulation purposes we select the structure of the spectrum of the covariance matrix and of the mean vector to make it consistent with the characteristics of the empirical data. Particularly, for each dimension p we select the expected returns equally spread on the interval -0.3 to 0.3, capturing a typical spectrum of daily returns measured in percent. The covariance matrix has a strong impact on the properties of the shrinkage intensity and for this reason we consider several structures of its spectra. Replicating the properties of empirical data we generate covariance matrices with eigenvalues satisfying the equation $\lambda_i = 0.1e^{\delta c \cdot (i-1)/p}$ for $i = 1, ..., p$. Thus the smallest eigenvalue is 0.1 and by selecting appropriate values for c we control the largest eigenvalue and thus the condition index of the covariance matrix. Large condition indices imply ill-conditioned covariance matrices, with the eigenvalues very sensitive to changes of the elements. We choose δ to attain the condition indices of 150, 1000 and 8000. The target portfolio weights are set equal to the weights of the equally weighted portfolio, i.e. $b_i = 1/p$ for $i = 1, ..., p$.

First we assess the general behavior of the oracle and bona-fide shrinkage intensities as functions of c. The oracle shrinkage intensities are computed using expressions in (2.11) and (2.19) for the cases $c < 1$ and $c > 1$, respectively. The parameters these expressions depend on are computed using the true mean vector and the true covariance matrix. For the bona-fide shrinkage intensities we estimate these parameters consistently and thus use expressions in (2.28) and (2.33). The results are illustrated for different condition indices in Figure 1. We observe that in all cases the shrinkage intensity falls to zero if $c \to 1_-$ and increases with c if $c > 1$. Thus if c is small the shrinkage estimator puts higher weight on the classical estimator of the portfolio weights, due to lower estimation risk. If c tends to 1 the system becomes unstable because of eigenvalues which are close to zero. In this case the portfolio weights collapse to the target portfolio weights. With c further increasing the shrinkage intensity increases too, implying that the pseudo-inverse covariance matrix can be evaluated in a proper way. The fraction of the EU portfolio increases with c in this case. Furthermore, the discrepancy between the oracle and bona-fide estimators naturally increases with higher condition indices. However, the shrinkage intensities are smaller for lower condition indices. This can be explained by the specific impact of the condition indices on the portfolio characteristics V_{GMV} and s. Simulations showed that V_{GMV} increases in the condition index and s decreases. Thus their complex interaction in the expression for the limit of α_n^+ results in the decreasing behavior of the shrinkage intensity as a function of the condition index.

In a similar fashion we analyze the relative losses of portfolios based on the traditional, the oracle and the bona-fide estimators. As a benchmark, we take the equally weighted portfolio which is also the target portfolio of the shrinkage estimator. The relative losses as functions of c for fixed $p = 150$ are plotted in Figure 2. For $c < 1$ the losses of the traditional estimator show explosive behavior and are comparable to the shrinkage-based estimators only for very small values of c. Thus the traditional estimator is reliable only if the sample size is at least three times larger than the dimension. The performance of the two shrinkage-based estimators is similar and stable over the whole range of c excluding a small neighborhood of one. It is important that the shrinkage-based portfolio clearly dominates the equally weighted benchmark. However, the losses are not monotonous and attain the maximum for c between 0.5 and 0.8. For c between 0.9 and 1.2 the shrinkage-based estimators lead to discontinuous and unstable losses, disadvising the investors from using samples sizes comparable with the dimension. For $c > 1$ the situation is similar with an extremely small difference between oracle- and bona-fide-based losses also in the limit. The traditional estimators attains the minimal loss for c between two and three and increase steadily for larger intensities.

The behavior of losses as functions of the dimension p is illustrated in Figure 3. The fraction c is set to 0.2, 0.5, 0.8 and 2 from left to right, while the condition index equals 150, 1000 and 8000 from top to bottom. From financial perspective it is important to note that the traditional estimator outperforms the equally weighted portfolio only for small values of c (in the particular setup for $c < 0.3$), thus when the classical estimators are stable and robust. For larger c the losses of the traditional estimator increase dramatically and are comparable with the remaining estimators only for $c > 1$. As before the oracle and the bona-fide estimators show similar results and clearly beat the benchmarks. Furthermore, the performance is stable for a wide range of dimensions. Note that all losses are decreasing functions of the dimension. Several exceptions can be found for the shrinkage-based strategies, but these can be due to poor validity of the asymptotic results for relatively small dimensions.

3.2 Empirical study

The data sets used in this study covers daily data on 445 S&P500 constituents available for the whole period from 01.01.2004 till 10.10.2014. The investor allocates his wealth to the constituents for prespecified values of p and c with daily reallocation. For simplicity we neglect the transaction costs in the below discussion. For illustration purposes we consider the first $p = 50$ and 100 assets in alphabetic ordering and $c = 0.2$, 0.5, 0.8 and 2. The risk aversion coefficient equals 10. In contrary to the simulation study the true parameters are not known for real data. Thus the oracle shrinkage coefficient is computed using the plug-in and not the consistent estimators of the portfolio characteristics.

The time series of estimated shrinkage coefficients (oracle and bona fide) is given in Figure 4. We observe that for small values of c and thus a low estimation risk the shrinkage intensities are close to one. The behavior is very stable, but mimics the periods of high and low volatility of financial markets, particularly with spikes in 2009 and 2011. Thus high volatility on financial markets causes higher shrinkage coefficients and thus larger fraction of the mean-variance portfolio. This can be justified by stronger effects of diversification during turmoil periods. With larger c the certainty in the classical portfolio diminishes. This results in lower and more volatile the shrinkage intensities. The investors prefers equally weighted portfolio due large estimation risk. The bona fide estimator is smaller than the oracle counterpart for moderate values of c, but becomes larger if c tends to one. For $c > 2$ the shrinkage coefficients are unstable, with a better behavior for larger portfolio consisting of 100 assets.

To illustrate the economic advantages of the derived theoretical results we consider the certainty equivalent, the Sharpe ratio, the Value-at-risk and the expected shortfall (both at 0.01 and 0.05 levels) as performance measures of real-data trading strategies. To robustify the procedure we randomly draw 1000 portfolios of size $p = 50$, 100, 300 from the population of size 455. For each of the portfolios we implement dynamic trading strategies for the last 200 days. We compute the optimal portfolio weights using the most recent n observations obtained for $c = 0.2$, 0.5, 0.8 and 2. The realized returns for each portfolio are used to compute the above mentioned performance measures. The averages over all 1000 portfolios of given size are summarized in Tables 1-3. For small c values, e.g. $c = 0.2$ the traditional estimator performs well and dominates all the alternatives. This coincides with the conclusions in the simulation study. For $c = 0.5$ the trading strategies based on the oracle and bona-fide estimators become superior. For $c = 0.8$ the evidence is mixed if we take the benchmarks into account. For some parameter constellations and measures the equally weighted portfolio is dominating, while for other the shrinkage-based strategies stay superior. For $c = 2$ the equally weighted portfolio dominate in the majority of the cases. The traditional estimator obviously fails and provides frequently useless results. Additionally to the evidence from the simulation study in the previous section, we conclude that the shrinkage-based strategies are economically advantageous too if $c < 1$ and not close to one. This conclusion is robust with respect to the portfolio composition and the chosen performance measure. For large c the shrinkage-based portfolios still improve the traditional estimator, but cannot outperform the benchmark.

4 Summary

In this paper we consider the portfolio selection in high-dimensional framework. Particularly, we assume that the number of assets p and the sample size n tend to infinity, but their ratio p/n tends to constant c. Note that the c maybe larger than one, implying that we have more assets than observations. Because of the large estimation risk we suggest a shrinkage-based estimator of the portfolio weights, which shrinks the mean-variance portfolio to the equally weighted portfolio. For the established shrinkage intensity we derive the limiting value. It depends on c and on the characteristics of the efficient frontier. Unfortunately the result is only an oracle value and is not feasible in practice, since it depends on unknown quantities. Thus we suggest a bona-fide estimator which overcomes this problem. From the technical point of view we rely on the theory of random matrices and work with the asymptotic behavior of linear and quadratic forms in mean vector and (pseudo)-inverse covariance matrix. In extensive simulation and empirical studies we evaluate the performance of established results with artificial and real data. If the sample size is smaller than or comparable to the dimension, then the performance is poor and some simple alternatives, e.g. equally weighted portfolio, should be favored.

5 Appendix A: Proofs

Here the proofs of the theorems are given. Recall that the sample mean vector and the sample covariance matrix are given by

$$\bar{\mathbf{y}}_n = \frac{1}{n}\mathbf{Y}_n\mathbf{1}_n = \boldsymbol{\mu}_n + \boldsymbol{\Sigma}_n^{\frac{1}{2}}\bar{\mathbf{x}}_n \quad \text{with} \quad \bar{\mathbf{x}}_n = \frac{1}{n}\mathbf{X}_n\mathbf{1}_n \tag{5.1}$$

and

$$\mathbf{S}_n = \frac{1}{n}\mathbf{Y}_n(\mathbf{I}-\frac{1}{n}\mathbf{1}\mathbf{1}')\mathbf{Y}_n' = \mathbf{\Sigma}_n^{\frac{1}{2}}\mathbf{V}_n\mathbf{\Sigma}_n^{\frac{1}{2}} \quad \text{with} \quad \mathbf{V}_n = \frac{1}{n}\mathbf{X}_n(\mathbf{I}-\frac{1}{n}\mathbf{1}\mathbf{1}')\mathbf{X}_n', \tag{5.2}$$

respectively. Later on, we also make use of $\tilde{\mathbf{V}}_n$ defined by

$$\tilde{\mathbf{V}}_n = \frac{1}{n}\mathbf{X}_n\mathbf{X}_n' \tag{5.3}$$

and the formula for the 1-rank update of usual inverse given by (c.f., Horn and Johnsohn (1985))

$$\mathbf{V}_n^{-1} = (\tilde{\mathbf{V}}_n - \bar{\mathbf{x}}_n\bar{\mathbf{x}}_n')^{-1} = \tilde{\mathbf{V}}_n^{-1} + \frac{\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}}{1-\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n} \tag{5.4}$$

as well as the formula for the 1-rank update of Moore-Penrose inverse (see,Meyer (1973)) expressed as

$$\begin{aligned}\mathbf{V}_n^{+} &= \left(\tilde{\mathbf{V}}' - \bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\right)^{+}\\ &= \tilde{\mathbf{V}}_n^{+} - \frac{\tilde{\mathbf{V}}_n^{+}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^2 + (\tilde{\mathbf{V}}_n^{+})^2\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})}{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^2\bar{\mathbf{x}}_n} + \frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^3\bar{\mathbf{x}}_n}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^2\bar{\mathbf{x}}_n)^2}\tilde{\mathbf{V}}_n^{+}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{+}. \end{aligned}\tag{5.5}$$

First, we present an important lemma which is a special case of Theorem 1 in Rubio and Mestre (2011).

Lemma 5.1. *Assume (A2). Let a nonrandom $p\times p$-dimensional matrix $\mathbf{\Theta}_p$ and a nonrandom $n\times n$-dimensional matrix $\mathbf{\Theta}_n$ possess a uniformly bounded trace norms (sum of singular values). Then it holds that*

$$\left|tr\left(\mathbf{\Theta}_p(\tilde{\mathbf{V}}_n - z\mathbf{I}_p)^{-1}\right) - m(z)tr\left(\mathbf{\Theta}_p\right)\right| \xrightarrow{a.s.} 0 \tag{5.6}$$

$$\left|tr\left(\mathbf{\Theta}_n(1/n\mathbf{X}_n'\mathbf{X}_n - z\mathbf{I}_n)^{-1}\right) - \underline{m}(z)tr\left(\mathbf{\Theta}_n\right)\right| \xrightarrow{a.s.} 0 \tag{5.7}$$

for $p/n \longrightarrow c \in (0,+\infty)$ as $n\to\infty$, where

$$m(z) = (x(z)-z)^{-1} \quad \text{and} \quad \underline{m}(z) = -\frac{1-c}{z} + cm(z) \tag{5.8}$$

with

$$x(z) = \frac{1}{2}\left(1-c+z+\sqrt{(1-c+z)^2-4z}\right). \tag{5.9}$$

Proof of Lemma 5.1: The application of Theorem 1 in Rubio and Mestre (2011) leads to (5.6) where $x(z)$ is a unique solution in $\mathbb{C}^+$ of the following equation

$$\frac{1-x(z)}{x(z)} = \frac{c}{x(z)-z}. \tag{5.10}$$

The two solutions of (5.10) are given by

$$x_{1,2}(z) = \frac{1}{2}\left(1-c+z\pm\sqrt{(1-c+z)^2-4z}\right). \tag{5.11}$$

In order to decide which of two solutions is feasible, we note that $x_{1,2}(z)$ is the Stieltjes transform with a positive imaginary part. Thus, without loss of generality, we can take $z = 1 + c + i2\sqrt{c}$ and get

$$\mathbf{Im}\{x_{1,2}(z)\} = \mathbf{Im}\left\{\frac{1}{2}\left(2 + i2\sqrt{c} \pm i2\sqrt{2c}\right)\right\} = \mathbf{Im}\left\{1 + i\sqrt{c}(1 \pm \sqrt{2})\right\} = \sqrt{c}\left(1 \pm \sqrt{2}\right), \tag{5.12}$$

which is positive only if the sign " + " is chosen. Hence, the solution is given by

$$x(z) = \frac{1}{2}\left(1 - c + z + \sqrt{(1-c+z)^2 - 4z}\right). \tag{5.13}$$

The second assertion of the lemma follows directly from Bai and Silverstein (2010). $\square$

Second, we will need the following technical lemmas.

Lemma 5.2. *Assume (A2). Let $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$ be universal nonrandom vectors with bounded Euclidean norms. Then it holds that*

$$\left|\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^{-1}\boldsymbol{\theta} - (1-c)^{-1}\boldsymbol{\xi}'\boldsymbol{\theta}\right| \xrightarrow{a.s.} 0\,, \tag{5.14}$$

$$\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n \xrightarrow{a.s.} c\,, \tag{5.15}$$

$$\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\boldsymbol{\theta} \xrightarrow{a.s.} 0\,, \tag{5.16}$$

$$\left|\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^{-2}\boldsymbol{\theta} - (1-c)^{-3}\boldsymbol{\xi}'\boldsymbol{\theta}\right| \xrightarrow{a.s.} 0\,, \tag{5.17}$$

$$\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-2}\bar{\mathbf{x}}_n \xrightarrow{a.s.} \frac{c}{(1-c)}\,, \tag{5.18}$$

$$\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-2}\boldsymbol{\theta} \xrightarrow{a.s.} 0 \tag{5.19}$$

for $p/n \longrightarrow c \in (0,1)$ as $n \to \infty$.

Proof of Lemma 5.2: Since the trace norm of $\boldsymbol{\theta}\boldsymbol{\xi}'$ is uniformly bounded, i.e.

$$||\boldsymbol{\theta}\boldsymbol{\xi}'||_{tr} \le \sqrt{\boldsymbol{\theta}'\boldsymbol{\theta}}\sqrt{\boldsymbol{\xi}'\boldsymbol{\xi}} < \infty,$$

we get from Lemma 5.1 that

$$|tr((\tilde{\mathbf{V}}_n - z\mathbf{I}_p)^{-1}\boldsymbol{\theta}\boldsymbol{\xi}') - m(z)tr(\boldsymbol{\theta}\boldsymbol{\xi}')| \xrightarrow{a.s.} 0 \ \text{ for } p/n \to c < 1 \text{ as } n \to \infty$$

Furthermore, the application of $m(z) \to (1-c)^{-1}$ as $z \to 0$ leads to

$$|\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^{-1}\boldsymbol{\theta} - (1-c)^{-1}\boldsymbol{\xi}'\boldsymbol{\theta}| \xrightarrow{a.s.} 0 \ \text{ for } p/n \to c < 1 \text{ as } n \to \infty\,,$$

which proves (5.14).

For deriving (5.15) we consider

$$\begin{aligned}\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n &= \lim_{z\to 0^+}\operatorname{tr}\left[\frac{1}{\sqrt{n}}\mathbf{X}_n'\left(\frac{1}{n}\mathbf{X}_n\mathbf{X}_n'-z\mathbf{I}_p\right)^{-1}\frac{1}{\sqrt{n}}\mathbf{X}_n\left(\frac{\mathbf{1}_n\mathbf{1}_n'}{n}\right)\right]\\ &= \lim_{z\to 0^+}\operatorname{tr}\left[\left(\frac{\mathbf{1}_n\mathbf{1}_n'}{n}\right)\right]+z\operatorname{tr}\left[\left(\frac{1}{n}\mathbf{X}_n'\mathbf{X}_n-z\mathbf{I}_n\right)^{-1}\left(\frac{\mathbf{1}_n\mathbf{1}_n'}{n}\right)\right],\end{aligned}$$

where the last equality follows from the Woodbury formula (e.g., Horn and Johnsohn (1985)). The application of Lemma 5.1 leads to

$$\operatorname{tr}\left[\left(\frac{\mathbf{1}_n\mathbf{1}_n'}{n}\right)\right]+z\operatorname{tr}\left[\left(\frac{1}{n}\mathbf{X}_n'\mathbf{X}_n-z\mathbf{I}_n\right)^{-1}\left(\frac{\mathbf{1}_n\mathbf{1}_n'}{n}\right)\right]\xrightarrow{a.s.}[1+(c-1)+czm(z)]\operatorname{tr}\left[\left(\frac{\mathbf{1}_n\mathbf{1}_n'}{n}\right)\right]$$

for $p/n \longrightarrow c<1$ as $n\to\infty$ where $m(z)$ is given by (5.8). Setting $z\to 0^+$ and taking into account $\lim_{z\to 0^+} m(z)=\frac{1}{1-c}$ we get

$$\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n \xrightarrow{a.s.} 1+c-1=c \text{ for } \frac{p}{n}\longrightarrow c\in(0,1) \text{ as } n\to\infty\,.$$

The result (5.16) was derived in Pan (2014) (see, p. 673 of this reference).

Next, we prove (5.17). It holds that

$$\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^{-2}\boldsymbol{\theta} \;=\; \frac{\partial}{\partial z}\operatorname{tr}\left[\left(\tilde{\mathbf{V}}_n-z\mathbf{I}_p\right)^{-1}\boldsymbol{\theta}\boldsymbol{\xi}'\right]\Big|_{z=0} = \frac{\partial}{\partial z}\zeta_n(z)\Big|_{z=0}$$

where $\zeta_n(z)=\operatorname{tr}\left[\left(\tilde{\mathbf{V}}_n-z\mathbf{I}\right)^{-1}\boldsymbol{\theta}\boldsymbol{\xi}'\right]$. From Lemma 5.1 $\zeta_n(z)$ tends a.s. to $m(z)\boldsymbol{\xi}'\boldsymbol{\theta}$ as $n\to\infty$. Furthermore,

$$\frac{\partial}{\partial z}m(z)\Big|_{z=0}=\frac{\partial}{\partial z}\frac{1}{x(z)-z}\Big|_{z=0}=-\frac{x'(z)-1}{(x(z)-z)^2}\Big|_{z=0}=-\left.\frac{\frac{1}{2}\left(1-\frac{1+c-z}{\sqrt{(1-c+z)^2-4z}}\right)-1}{(x(z)-z)^2}\right|_{z=0}=\frac{1}{(1-c)^3}\,. \tag{5.20}$$

Consequently,

$$|\boldsymbol{\xi}'\mathbf{S}_n^{-2}\boldsymbol{\theta}-(1-c)^{-3}\boldsymbol{\xi}'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\theta}|\xrightarrow{a.s.}0 \;\text{ for } p/n\to c<1 \text{ as } n\to\infty\,.$$

Let $\eta_n(z)=\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n-z\mathbf{I})^{-1}\bar{\mathbf{x}}_n$ and $\boldsymbol{\Theta}_n=\left(\frac{\mathbf{1}_n\mathbf{1}_n'}{n}\right)$. Then

$$\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-2}\bar{\mathbf{x}}_n=\frac{\partial}{\partial z}\eta_n(z)\Big|_{z=0}\,,$$

where

$$\begin{aligned}\eta_n(z) &= \operatorname{tr}\left[\frac{1}{\sqrt{n}}\mathbf{X}_n'\left(\frac{1}{n}\mathbf{X}_n\mathbf{X}_n' - z\mathbf{I}_p\right)^{-1}\frac{1}{\sqrt{n}}\mathbf{X}_n\boldsymbol{\Theta}_n\right]\\ &= \operatorname{tr}(\boldsymbol{\Theta}_n) + z\operatorname{tr}\left[(1/n\mathbf{X}_n'\mathbf{X}_n - z\mathbf{I}_n)^{-1}\boldsymbol{\Theta}_n\right] \xrightarrow{a.s.} 1 + z\underline{m}(z) = c + czm(z)\end{aligned}$$

for $\frac{p}{n} \to c \in (0,1)$ as $n \to \infty$. Hence,

$$\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-2}\bar{\mathbf{x}}_n \xrightarrow{a.s.} cm(0) + cz\frac{\partial}{\partial z}m(z)\Big|_{z=0} = \frac{c}{1-c}$$

for $\frac{p}{n} \to c \in (0,1)$ as $n \to \infty$.

Finally, we get

$$\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-2}\boldsymbol{\theta} = \frac{\partial}{\partial z}\operatorname{tr}\left[\bar{\mathbf{x}}_n'\left(\tilde{\mathbf{V}}_n - z\mathbf{I}_p\right)^{-1}\boldsymbol{\theta}\right]\Big|_{z=0} \xrightarrow{a.s.} 0$$

for $\frac{p}{n} \to c \in (0,1)$ as $n \to \infty$. □

Lemma 5.3. *Assume (A2). Let $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$ be universal nonrandom vectors with bounded Euclidean norms. Then it holds that*

$$\begin{aligned}\left|\boldsymbol{\xi}'\mathbf{V}_n^{-1}\boldsymbol{\theta} - (1-c)^{-1}\boldsymbol{\xi}'\boldsymbol{\theta}\right| &\xrightarrow{a.s.} 0\,, &(5.21)\\ \bar{\mathbf{x}}_n'\mathbf{V}_n^{-1}\bar{\mathbf{x}}_n &\xrightarrow{a.s.} \frac{c}{1-c}\,, &(5.22)\\ \bar{\mathbf{x}}_n'\mathbf{V}_n^{-1}\boldsymbol{\theta} &\xrightarrow{a.s.} 0\,, &(5.23)\\ \left|\boldsymbol{\xi}'\mathbf{V}_n^{-2}\boldsymbol{\theta} - (1-c)^{-3}\boldsymbol{\xi}'\boldsymbol{\theta}\right| &\xrightarrow{a.s.} 0\,, &(5.24)\\ \bar{\mathbf{x}}_n'\mathbf{V}_n^{-2}\bar{\mathbf{x}}_n &\xrightarrow{a.s.} \frac{c}{(1-c)^3}\,, &(5.25)\\ \bar{\mathbf{x}}_n'\mathbf{V}_n^{-2}\boldsymbol{\theta} &\xrightarrow{a.s.} 0 &(5.26)\end{aligned}$$

for $p/n \longrightarrow c \in (0,1)$ as $n \to \infty$.

Proof of Lemma 5.3: From (5.4) we obtain

$$\boldsymbol{\xi}'\mathbf{V}_n^{-1}\boldsymbol{\theta} = \boldsymbol{\xi}'\tilde{\mathbf{V}}_n^{-1}\boldsymbol{\theta} + \frac{\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\boldsymbol{\theta}}{1 - \bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n} \xrightarrow{a.s.} (1-c)^{-1}\boldsymbol{\xi}'\boldsymbol{\theta}$$

for $p/n \longrightarrow c \in (0,1)$ as $n \to \infty$ following (5.14)-(5.16). Similarly, we get (5.22) and (5.23).

In case of (5.23), we get

$$\begin{aligned}\boldsymbol{\xi}'\mathbf{V}_n^{-2}\boldsymbol{\theta} &= \boldsymbol{\xi}'\tilde{\mathbf{V}}_n^{-2}\boldsymbol{\theta} + \frac{\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^{-2}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\boldsymbol{\theta}}{1-\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n} + \frac{\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-2}\boldsymbol{\theta}}{1-\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n}\\ &+ \bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-2}\bar{\mathbf{x}}_n\frac{\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\boldsymbol{\theta}}{(1-\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n)^2} \xrightarrow{a.s.} (1-c)^{-1}\boldsymbol{\xi}'\boldsymbol{\theta}\end{aligned}$$

for $p/n \longrightarrow c \in (0,1)$ as $n \to \infty$. Similarly,

$$\begin{aligned}\bar{\mathbf{x}}_n'\mathbf{V}_n^{-2}\bar{\mathbf{x}}_n &= \bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-2}\bar{\mathbf{x}}_n + \frac{\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-2}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n}{1-\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n} + \frac{\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-2}\bar{\mathbf{x}}_n}{1-\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n}\\ &+ \bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-2}\bar{\mathbf{x}}_n\frac{\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n}{(1-\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n)^2} = \frac{\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-2}\bar{\mathbf{x}}_n}{(1-\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n)^2} \xrightarrow{a.s.} \frac{c}{(1-c)^3}\end{aligned}$$

and

$$\begin{aligned}\bar{\mathbf{x}}_n'\mathbf{V}_n^{-2}\boldsymbol{\theta} &= \bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-2}\boldsymbol{\theta} + \frac{\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-2}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\boldsymbol{\theta}}{1-\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n} + \frac{\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-2}\boldsymbol{\theta}}{1-\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n}\\ &+ \bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-2}\bar{\mathbf{x}}_n\frac{\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\boldsymbol{\theta}}{(1-\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n)^2} \xrightarrow{a.s.} 0\end{aligned}$$

for $p/n \longrightarrow c \in (0,1)$ as $n \to \infty$. □

Lemma 5.4. *Assume (A2). Let $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$ be universal nonrandom vectors with bounded Euclidean norms and let $\mathbf{P}_n = \mathbf{V}_n^{-1} - \frac{\mathbf{V}_n^{-1}\boldsymbol{\eta}\boldsymbol{\eta}'\mathbf{V}_n^{-1}}{\boldsymbol{\eta}'\mathbf{V}_n^{-1}\boldsymbol{\eta}}$ where $\boldsymbol{\eta}$ is a universal nonrandom vectors with bounded Euclidean norm. Then it holds that*

$$\boldsymbol{\xi}'\mathbf{P}_n\boldsymbol{\theta} \xrightarrow{a.s.} (1-c)^{-1}\left(\boldsymbol{\xi}'\boldsymbol{\theta} - \frac{\boldsymbol{\xi}'\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{\theta}}{\boldsymbol{\eta}'\boldsymbol{\eta}}\right), \tag{5.27}$$

$$\bar{\mathbf{x}}_n'\mathbf{P}_n\bar{\mathbf{x}}_n \xrightarrow{a.s.} \frac{c}{1-c}, \tag{5.28}$$

$$\bar{\mathbf{x}}_n'\mathbf{P}_n\boldsymbol{\theta} \xrightarrow{a.s.} 0, \tag{5.29}$$

$$\boldsymbol{\xi}'\mathbf{P}_n^2\boldsymbol{\theta} \xrightarrow{a.s.} (1-c)^{-3}\left(\boldsymbol{\xi}'\boldsymbol{\theta} - \frac{\boldsymbol{\xi}'\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{\theta}}{\boldsymbol{\eta}'\boldsymbol{\eta}}\right), \tag{5.30}$$

$$\bar{\mathbf{x}}_n'\mathbf{P}_n^2\bar{\mathbf{x}}_n \xrightarrow{a.s.} \frac{c}{(1-c)^3}, \tag{5.31}$$

$$\bar{\mathbf{x}}_n'\mathbf{P}_n^2\boldsymbol{\theta} \xrightarrow{a.s.} 0 \tag{5.32}$$

for $p/n \longrightarrow c \in (0,1)$ as $n \to \infty$.

Proof of Lemma 5.4: It holds that

$$\boldsymbol{\xi}'\mathbf{P}_n\boldsymbol{\theta} = \boldsymbol{\xi}'\mathbf{V}_n^{-1}\boldsymbol{\theta} - \frac{\boldsymbol{\xi}'\mathbf{V}_n^{-1}\boldsymbol{\eta}\boldsymbol{\eta}'\mathbf{V}_n^{-1}\boldsymbol{\theta}}{\boldsymbol{\eta}'\mathbf{V}_n^{-1}\boldsymbol{\eta}} \xrightarrow{a.s.} (1-c)^{-1}\left(\boldsymbol{\xi}'\boldsymbol{\theta} - \frac{\boldsymbol{\xi}'\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{\theta}}{\boldsymbol{\eta}'\boldsymbol{\eta}}\right)$$

for $p/n \longrightarrow c \in (0,1)$ as $n \to \infty$ following (5.21). Similarly, we get

$$\bar{\mathbf{x}}_n'\mathbf{P}_n\bar{\mathbf{x}}_n = \bar{\mathbf{x}}_n'\mathbf{V}_n^{-1}\bar{\mathbf{x}}_n - \frac{\bar{\mathbf{x}}_n'\mathbf{V}_n^{-1}\boldsymbol{\eta}\boldsymbol{\eta}'\mathbf{V}_n^{-1}\bar{\mathbf{x}}_n}{\boldsymbol{\eta}'\mathbf{V}_n^{-1}\boldsymbol{\eta}} \xrightarrow{a.s.} \frac{c}{1-c}$$

and

$$\bar{\mathbf{x}}_n'\mathbf{P}_n\boldsymbol{\theta} = \bar{\mathbf{x}}_n'\mathbf{V}_n^{-1}\boldsymbol{\theta} - \frac{\bar{\mathbf{x}}_n'\mathbf{V}_n^{-1}\boldsymbol{\eta}\boldsymbol{\eta}'\mathbf{V}_n^{-1}\boldsymbol{\theta}}{\boldsymbol{\eta}'\mathbf{V}_n^{-1}\boldsymbol{\eta}} \xrightarrow{a.s.} 0$$

for $p/n \longrightarrow c \in (0,1)$ as $n \to \infty$.

The rest of the proof follows from the equality

$$\mathbf{P}_n^2 = \mathbf{V}_n^{-2} - \frac{\mathbf{V}_n^{-2}\boldsymbol{\eta}\boldsymbol{\eta}'\mathbf{V}_n^{-1}}{\boldsymbol{\eta}'\mathbf{V}_n^{-1}\boldsymbol{\eta}} - \frac{\mathbf{V}_n^{-1}\boldsymbol{\eta}\boldsymbol{\eta}'\mathbf{V}_n^{-2}}{\boldsymbol{\eta}'\mathbf{V}_n^{-1}\boldsymbol{\eta}} + \boldsymbol{\eta}'\mathbf{V}_n^{-2}\boldsymbol{\eta}\frac{\mathbf{V}_n^{-1}\boldsymbol{\eta}\boldsymbol{\eta}'\mathbf{V}_n^{-1}}{(\boldsymbol{\eta}'\mathbf{V}_n^{-1}\boldsymbol{\eta})^2}$$

and Lemma 5.3. $\square$

Proof of Theorem 2.1: The optimal shrinkage intensity can be rewritten in the following way

$$\begin{aligned} \alpha_n^* &= \gamma^{-1}\frac{\hat{\mathbf{w}}_S'(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b}) - \mathbf{b}'(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b})}{\hat{\mathbf{w}}_S'\boldsymbol{\Sigma}_n\hat{\mathbf{w}}_S - 2\mathbf{b}'\boldsymbol{\Sigma}_n\hat{\mathbf{w}}_S + \mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{b}} \qquad (5.33)\\ &= \gamma^{-1}\frac{\frac{\mathbf{1}'\mathbf{S}_n^{-1}(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b})}{\mathbf{1}'\mathbf{S}_n^{-1}\mathbf{1}} + \gamma^{-1}\bar{\mathbf{y}}_n'\hat{\mathbf{Q}}_n(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b}) - \mathbf{b}'(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b})}{\frac{\mathbf{1}'\mathbf{S}_n^{-1}\boldsymbol{\Sigma}_n\mathbf{S}_n^{-1}\mathbf{1}}{(\mathbf{1}'\mathbf{S}_n^{-1}\mathbf{1})^2} + 2\gamma^{-1}\frac{\bar{\mathbf{y}}_n'\hat{\mathbf{Q}}_n\boldsymbol{\Sigma}_n\mathbf{S}_n^{-1}\mathbf{1}}{\mathbf{1}'\mathbf{S}_n^{-1}\mathbf{1}} + \gamma^{-2}\bar{\mathbf{y}}_n'\hat{\mathbf{Q}}_n\boldsymbol{\Sigma}_n\hat{\mathbf{Q}}_n\bar{\mathbf{y}}_n - 2\frac{\mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{S}_n^{-1}\mathbf{1}}{\mathbf{1}'\mathbf{S}_n^{-1}\mathbf{1}} - 2\gamma^{-1}\mathbf{b}'\boldsymbol{\Sigma}_n\hat{\mathbf{Q}}_n\bar{\mathbf{y}}_n + \mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{b}} \end{aligned}$$

and $\bar{\mathbf{y}}_n = \boldsymbol{\mu}_n + \boldsymbol{\Sigma}_n^{1/2}\bar{\mathbf{x}}_n$ and $\mathbf{S}_n = \boldsymbol{\Sigma}_n^{1/2}\mathbf{V}_n\boldsymbol{\Sigma}_n^{1/2}$.

From Assumption (A3), we get that the following vectors $\boldsymbol{\Sigma}_n^{-1/2}\mathbf{1}$, $\boldsymbol{\Sigma}_n^{-1/2}(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b})$, $\boldsymbol{\Sigma}_n^{-1/2}\boldsymbol{\mu}_n$, and $\boldsymbol{\Sigma}_n^{1/2}\mathbf{b}$ possess bounded Euclidean norms. As a result, the application of Lemma 5.2 leads to

$$\begin{aligned} \mathbf{1}'\mathbf{S}_n^{-1}\mathbf{1} &= \mathbf{1}'\boldsymbol{\Sigma}_n^{-1/2}\mathbf{V}_n^{-1}\boldsymbol{\Sigma}_n^{-1/2}\mathbf{1} \xrightarrow{a.s.} (1-c)^{-1}\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1},\\ \mathbf{1}'\mathbf{S}_n^{-1}(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b}) &= \mathbf{1}'\boldsymbol{\Sigma}_n^{-1/2}\mathbf{V}_n^{-1}\boldsymbol{\Sigma}_n^{-1/2}(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b}) \xrightarrow{a.s.} (1-c)^{-1}\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b}),\\ \mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{S}_n^{-1}\mathbf{1} &= \mathbf{b}'\boldsymbol{\Sigma}_n\boldsymbol{\Sigma}_n^{-1/2}\mathbf{V}_n^{-1}\boldsymbol{\Sigma}_n^{-1/2}\mathbf{1} \xrightarrow{a.s.} (1-c)^{-1}\mathbf{b}'\boldsymbol{\Sigma}_n\boldsymbol{\Sigma}_n^{-1}\mathbf{1} = (1-c)^{-1},\\ \mathbf{1}'\mathbf{S}_n^{-1}\boldsymbol{\Sigma}_n\mathbf{S}_n^{-1}\mathbf{1} &= \mathbf{1}'\boldsymbol{\Sigma}_n^{-1/2}\mathbf{V}_n^{-2}\boldsymbol{\Sigma}_n^{-1/2}\mathbf{1} \xrightarrow{a.s.} (1-c)^{-3}\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1} \end{aligned}$$

for $p/n \longrightarrow c \in (0,1)$ as $n \to \infty$.

Finally, from Lemma 5.3 and 5.4 as well as by using the equalities

$$\hat{\mathbf{Q}}_n = \boldsymbol{\Sigma}_n^{-1/2}\mathbf{P}_n\boldsymbol{\Sigma}_n^{-1/2} \quad \text{and} \quad \mathbf{P}_n\mathbf{V}_n^{-1} = \mathbf{V}_n^{-2} - \frac{\mathbf{V}_n^{-1}\boldsymbol{\eta}\boldsymbol{\eta}'\mathbf{V}_n^{-2}}{\boldsymbol{\eta}'\mathbf{V}_n^{-1}\boldsymbol{\eta}}$$

with $\boldsymbol{\eta} = \boldsymbol{\Sigma}_n^{-1/2}\mathbf{1}$ we obtain

$$
\begin{aligned}
\bar{\mathbf{y}}_n'\hat{\mathbf{Q}}_n(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b}) &= \boldsymbol{\mu}_n'\boldsymbol{\Sigma}_n^{-1/2}\mathbf{P}_n\boldsymbol{\Sigma}_n^{-1/2}(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b}) + \bar{\mathbf{x}}_n'\mathbf{P}_n\boldsymbol{\Sigma}_n^{-1/2}(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b}) \\
&\xrightarrow{a.s.} (1-c)^{-1}\boldsymbol{\mu}_n'\mathbf{Q}_n(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b}), \\
\bar{\mathbf{y}}_n'\hat{\mathbf{Q}}_n\boldsymbol{\Sigma}_n\mathbf{S}_n^{-1}\mathbf{1} &= \boldsymbol{\mu}_n'\boldsymbol{\Sigma}_n^{-1/2}\mathbf{P}_n\mathbf{V}_n^{-1}\boldsymbol{\Sigma}_n^{-1/2}\mathbf{1} + \bar{\mathbf{x}}_n'\mathbf{P}_n\mathbf{V}_n^{-1}\boldsymbol{\Sigma}_n^{-1/2}\mathbf{1} \\
&\xrightarrow{a.s.} (1-c)^{-3}\boldsymbol{\mu}_n'\mathbf{Q}_n\mathbf{1} = 0, \\
\bar{\mathbf{y}}_n'\hat{\mathbf{Q}}_n\boldsymbol{\Sigma}_n\hat{\mathbf{Q}}_n\bar{\mathbf{y}}_n &= \bar{\mathbf{x}}_n'\mathbf{P}_n^2\bar{\mathbf{x}}_n + 2\boldsymbol{\mu}_n'\boldsymbol{\Sigma}_n^{-1/2}\mathbf{P}_n^2\bar{\mathbf{x}}_n + \boldsymbol{\mu}_n'\boldsymbol{\Sigma}_n^{-1/2}\mathbf{P}_n^2\boldsymbol{\Sigma}_n^{-1/2}\boldsymbol{\mu}_n \\
&\xrightarrow{a.s.} \frac{c}{(1-c)^3} + (1-c)^{-3}\boldsymbol{\mu}_n'\mathbf{Q}_n\boldsymbol{\mu}_n, \\
\mathbf{b}'\boldsymbol{\Sigma}_n\hat{\mathbf{Q}}_n\bar{\mathbf{y}}_n &= \mathbf{b}'\boldsymbol{\Sigma}_n^{1/2}\mathbf{P}_n\bar{\mathbf{x}}_n + \mathbf{b}'\boldsymbol{\Sigma}_n^{1/2}\mathbf{P}_n\boldsymbol{\Sigma}_n^{-1/2}\boldsymbol{\mu}_n \\
&\xrightarrow{a.s.} (1-c)^{-1}\left(\mathbf{b}'\boldsymbol{\mu}_n - \frac{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\mu}_n}{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}}\right)
\end{aligned}
$$

for $p/n \longrightarrow c \in (0,1)$ as $n \to \infty$.

Substituting the above results into the expression of the shrinkage intensity, we get $\alpha_n^* \xrightarrow{a.s.} \alpha^*$, where

$$
\alpha^* = \gamma^{-1}\frac{\dfrac{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b})}{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}} + \dfrac{\gamma^{-1}}{1-c}\boldsymbol{\mu}_n'\mathbf{Q}_n(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b}) - \mathbf{b}'(\boldsymbol{\mu}_n - \gamma\boldsymbol{\Sigma}_n\mathbf{b})}{\dfrac{1}{1-c}\dfrac{1}{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}} + \gamma^{-2}\left(\dfrac{1}{(1-c)^3}\boldsymbol{\mu}_n'\mathbf{Q}_n\boldsymbol{\mu}_n + \dfrac{c}{(1-c)^3}\right) - 2\dfrac{1}{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}} - 2\frac{\gamma^{-1}}{1-c}\left(\mathbf{b}'\boldsymbol{\mu}_n - \dfrac{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\mu}_n}{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}}\right) + \mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{b}} \tag{5.34}
$$

Denoting now $V_{GMV} = \dfrac{1}{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}}$, $R_{GMV} = \dfrac{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\mu}_n}{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}}$, $s = \boldsymbol{\mu}_n'\mathbf{Q}_n\boldsymbol{\mu}_n$, $R_b = \mathbf{b}'\boldsymbol{\mu}_n$ and making some technical manipulations we get the statement of Theorem 2.1.

□

For the proof of Theorem 2.2 we need several results about the properties of Moore-Penrose inverse which are summarized in the following three lemmas.

Lemma 5.5. *Assume (A2). Let $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$ be universal nonrandom vectors with bounded Eu-*

clidean norms. Then it holds that

$$\begin{aligned}
\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^+\boldsymbol{\theta} &\xrightarrow{a.s.} c^{-1}(c-1)^{-1}\boldsymbol{\xi}'\boldsymbol{\theta}\,, &(5.35)\\
\boldsymbol{\xi}'(\tilde{\mathbf{V}}_n^+)^2\boldsymbol{\theta} &\xrightarrow{a.s.} (c-1)^{-3}\boldsymbol{\xi}'\boldsymbol{\theta}\,, &(5.36)\\
\bar{\mathbf{x}}'\tilde{\mathbf{V}}_n^+\bar{\mathbf{x}}_n &= 1\,, &(5.37)\\
\bar{\mathbf{x}}'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n &\xrightarrow{a.s.} \frac{1}{c-1}\,, &(5.38)\\
\bar{\mathbf{x}}'(\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n &\xrightarrow{a.s.} \frac{c}{(c-1)^3}\,, &(5.39)\\
\bar{\mathbf{x}}'(\tilde{\mathbf{V}}_n^+)^4\bar{\mathbf{x}}_n &\xrightarrow{a.s.} \frac{c(c+1)}{(c-1)^5}\,, &(5.40)\\
\bar{\mathbf{x}}'\tilde{\mathbf{V}}_n^+\boldsymbol{\theta} &\xrightarrow{a.s.} 0\,, &(5.41)\\
\bar{\mathbf{x}}'(\tilde{\mathbf{V}}_n^+)^2\boldsymbol{\theta} &\xrightarrow{a.s.} 0\,, &(5.42)\\
\bar{\mathbf{x}}'(\tilde{\mathbf{V}}_n^+)^3\boldsymbol{\theta} &\xrightarrow{a.s.} 0 &(5.43)
\end{aligned}$$

for $p/n \longrightarrow c \in (1, +\infty)$ as $n \to \infty$.

Proof of Lemma 5.5: It holds that

$$\tilde{\mathbf{V}}^+ = \left(\frac{1}{n}\mathbf{X}_n\mathbf{X}_n'\right)^+ = \frac{1}{\sqrt{n}}\mathbf{X}_n\left(\frac{1}{n}\mathbf{X}_n'\mathbf{X}_n\right)^{-2}\frac{1}{\sqrt{n}}\mathbf{X}_n'$$

and, similarly,

$$(\tilde{\mathbf{V}}^+)^i = \frac{1}{\sqrt{n}}\mathbf{X}_n\left(\frac{1}{n}\mathbf{X}_n'\mathbf{X}_n\right)^{-(i+1)}\frac{1}{\sqrt{n}}\mathbf{X}_n' \quad \text{for} \quad i = 2, 3, 4.$$

Let $\boldsymbol{\Theta} = \boldsymbol{\theta}\boldsymbol{\xi}'$. It holds that

$$\begin{aligned}
\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^+\boldsymbol{\theta} &= \mathrm{tr}\left[\frac{1}{\sqrt{n}}\mathbf{X}_n\left(\frac{1}{n}\mathbf{X}_n'\mathbf{X}_n\right)^{-2}\frac{1}{\sqrt{n}}\mathbf{X}_n'\boldsymbol{\Theta}\right] = \frac{\partial}{\partial z}\mathrm{tr}\left[\frac{1}{\sqrt{n}}\mathbf{X}_n\left(\frac{1}{n}\mathbf{X}_n'\mathbf{X}_n - z\mathbf{I}_n\right)^{-1}\frac{1}{\sqrt{n}}\mathbf{X}_n'\boldsymbol{\Theta}\right]\Bigg|_{z=0}\,,\\
\boldsymbol{\xi}'(\tilde{\mathbf{V}}_n^+)^2\boldsymbol{\theta} &= \mathrm{tr}\left[\frac{1}{\sqrt{n}}\mathbf{X}_n\left(\frac{1}{n}\mathbf{X}_n'\mathbf{X}_n\right)^{-3}\frac{1}{\sqrt{n}}\mathbf{X}_n'\boldsymbol{\Theta}\right] = \frac{1}{2}\frac{\partial^2}{\partial z^2}\mathrm{tr}\left[\frac{1}{\sqrt{n}}\mathbf{X}_n\left(\frac{1}{n}\mathbf{X}_n'\mathbf{X}_n - z\mathbf{I}_n\right)^{-1}\frac{1}{\sqrt{n}}\mathbf{X}_n'\boldsymbol{\Theta}\right]\Bigg|_{z=0}\,.
\end{aligned}$$

The application of Woodbury formula (matrix inversion lemma, see, e.g., Horn and Johnsohn (1985)),

$$\frac{1}{\sqrt{n}}\mathbf{X}_n\left(\frac{1}{n}\mathbf{X}_n'\mathbf{X}_n - z\mathbf{I}_n\right)^{-1}\frac{1}{\sqrt{n}}\mathbf{X}_n' = \mathbf{I}_p + z\left(\frac{1}{n}\mathbf{X}_n\mathbf{X}_n' - z\mathbf{I}_p\right)^{-1} \tag{5.44}$$

leads to

$$\begin{aligned}
\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^+\boldsymbol{\theta} &= \frac{\partial}{\partial z}z\mathrm{tr}\left[\left(\frac{1}{n}\mathbf{X}_n\mathbf{X}_n' - z\mathbf{I}_p\right)^{-1}\boldsymbol{\Theta}\right]\Bigg|_{z=0}\,,\\
\boldsymbol{\xi}'(\tilde{\mathbf{V}}_n^+)^2\boldsymbol{\theta} &= \frac{1}{2}\frac{\partial^2}{\partial z^2}z\mathrm{tr}\left[\left(\frac{1}{n}\mathbf{X}_n\mathbf{X}_n' - z\mathbf{I}_p\right)^{-1}\boldsymbol{\Theta}\right]\Bigg|_{z=0}\,.
\end{aligned}$$

From the proof of Lemma 5.2 we know that the matrix $\boldsymbol{\Theta}$ possesses the bounded trace norm.

Then the application of Lemma 5.1 leads to

$$\begin{aligned}\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^{+}\boldsymbol{\theta} &\xrightarrow{a.s.} \frac{\partial}{\partial z}\frac{z}{x(z)-z}\bigg|_{z=0}\boldsymbol{\xi}'\boldsymbol{\theta}\,,\\ \boldsymbol{\xi}'(\tilde{\mathbf{V}}_n^{+})^2\boldsymbol{\theta} &\xrightarrow{a.s.} \frac{1}{2}\frac{\partial^2}{\partial z^2}\frac{z}{x(z)-z}\bigg|_{z=0}\boldsymbol{\xi}'\boldsymbol{\theta}\end{aligned}$$

for $p/n \to c > 1$ as $n \to \infty$, where $x(z)$ is given in (5.9).

Let us make the following notations

$$\theta(z) = \frac{z}{x(z)-z} \quad \text{and} \quad \phi(z) = \frac{x(z)-zx'(z)}{z^2}\,.$$

Then the first and the second derivatives of $\theta(z)$ are given by

$$\theta'(z) = \theta^2(z)\phi(z) \quad \text{and} \quad \theta^{''}(z) = 2\theta(z)\theta'(z)\phi(z) + \theta^2(z)\phi'(z)\,. \tag{5.45}$$

Using L'Hopital's rule, we get

$$\theta(0) = \lim_{z\to 0^+}\theta(z) = \lim_{z\to 0^+}\frac{z}{x(z)-z} = \lim_{z\to 0^+}\frac{1}{(x'(z)-1)} = \frac{1}{\frac{1}{2}\left(1-\frac{1+c}{|1-c|}\right)-1} = -\frac{c-1}{c}\,, \tag{5.46}$$

$$\phi(0) = \lim_{z\to 0^+}\phi(z) = \lim_{z\to 0^+}\frac{x(z)-zx'(z)}{z^2} = -\frac{1}{2}\lim_{z\to 0^+}x^{''}(z) = -\frac{1}{2}\lim_{z\to 0^+}\frac{-2c}{((1-c+z)^2-4z)^{3/2}} = \frac{c}{(c-1)^3}\,, \tag{5.47}$$

and

$$\begin{aligned}\lim_{z\to 0^+}\phi'(z) &= -\lim_{z\to 0^+}\frac{2(x(z)-zx'(z))+z^2x^{''}(z)}{z^2}\\ &= -\lim_{z\to 0^+}\frac{2\phi(z)+x^{''}(z)}{z} = -\lim_{z\to 0^+}\left(2\phi'(z)+x^{'''}(z)\right),\end{aligned} \tag{5.48}$$

which implies

$$\phi'(0) = \lim_{z\to 0^+}\phi'(z) = -\frac{1}{3}\lim_{z\to 0^+}x^{'''}(z) = -\frac{1}{3}\lim_{z\to 0^+}\frac{6c(z-c-1)}{((1-c+z)^2-4z)^{5/2}} = \frac{2c(c+1)}{(c-1)^5}\,. \tag{5.49}$$

Combining (5.45), (5.46), (5.47), and (5.49), we get

$$\theta^{'}(0) = \lim_{z\to 0^+}\theta^{'}(z) = \theta^2(0)\phi(0) = \frac{1}{c(c-1)}$$

and

$$\theta^{''}(0) = \lim_{z\to 0^+}\theta^{''}(z) = 2\theta^3(0)\phi^2(0) + \theta^2(0)\phi'(0) = \frac{2}{(c-1)^3}\,.$$

Hence,

$$\begin{aligned}
\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^{+}\boldsymbol{\theta} &\xrightarrow{a.s.} \frac{1}{c(c-1)}\boldsymbol{\xi}'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\theta} \ \text{ for } p/n \to c > 1 \ \text{ as } n \to \infty, \\
\boldsymbol{\xi}'(\tilde{\mathbf{V}}_n^{+})^2\boldsymbol{\theta} &\xrightarrow{a.s.} \frac{1}{(c-1)^3}\boldsymbol{\xi}'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\theta} \ \text{ for } p/n \to c > 1 \ \text{ as } n \to \infty\,.
\end{aligned}$$

Taking into account that

$$\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{+}\bar{\mathbf{x}}_n = \frac{1}{n}\mathbf{1}_n'\mathbf{X}_n'\mathbf{X}_n(\mathbf{X}_n'\mathbf{X}_n)^{-2}\mathbf{X}_n'\mathbf{X}_n\mathbf{1}_n = \frac{1}{n}\mathbf{1}_n'\mathbf{1}_n = 1\,.$$

we get (5.37). Similarly, using

$$\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^i\bar{\mathbf{x}}_n \ = \ 1/n\mathbf{1}_n'(1/n\mathbf{X}_n'\mathbf{X}_n)^{-(i-1)}\mathbf{1}_n \ \text{ for } \ i = 2, 3, 4$$

we get

$$\begin{aligned}
1/n\mathbf{1}_n'(1/n\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{1}_n &= \lim_{z\to 0^+}\mathrm{tr}[(1/n\mathbf{X}_n'\mathbf{X}_n - z\mathbf{I})^{-1}\boldsymbol{\Theta}] \xrightarrow{a.s.} \underline{m}(0)\,, \\
1/n\mathbf{1}_n'(1/n\mathbf{X}_n'\mathbf{X}_n)^{-2}\mathbf{1}_n &= \lim_{z\to 0^+}\frac{\partial}{\partial z}\mathrm{tr}[(1/n\mathbf{X}_n'\mathbf{X}_n - z\mathbf{I})^{-1}\boldsymbol{\Theta}] \xrightarrow{a.s.} \underline{m}'(0)\,, \\
1/n\mathbf{1}_n'(1/n\mathbf{X}_n'\mathbf{X}_n)^{-2}\mathbf{1}_n &= \frac{1}{2}\lim_{z\to 0^+}\frac{\partial^2}{\partial z^2}\mathrm{tr}[(1/n\mathbf{X}_n'\mathbf{X}_n - z\mathbf{I})^{-1}\boldsymbol{\Theta}] \xrightarrow{a.s.} \frac{1}{2}\underline{m}''(0)
\end{aligned}$$

for $p/n \to c > 1$ as $n \to \infty$, where $\boldsymbol{\Theta}_n = 1/n\mathbf{1}_n\mathbf{1}_n'$.

The application of the equality (c.f., Bai and Silverstein (2010))

$$\underline{m}(z) = -\left(z - \frac{c}{1+\underline{m}(z)}\right)^{-1}.$$

leads to[2]

$$\underline{m}(0) \equiv \lim_{z\to 0^+}\underline{m}(z) = \frac{1}{c}(1+\underline{m}(0))\,,$$

which is linear in $\underline{m}(0)$ and results

$$\underline{m}(0) = \frac{1}{c-1}\,. \tag{5.50}$$

For the next one we investigate the first derivative of $\underline{m}(z)$ with respect to z, namely

$$\underline{m}'(z) = \left[z - \frac{c}{1+\underline{m}(z)}\right]^{-2}\left(1 + \frac{c\underline{m}'(z)}{(1+\underline{m}(z))^2}\right).$$

[2]Note that $\underline{m}(z)$ is bounded as $z \to 0^+$ because for any contour $C = \{x + iy | y \in [0, y_0]\}$ the Stieltjes transform $\underline{m}(z)$ is analytic on C and, thus, there exists $\delta > 0$ s.t. $\sup_{z\in C}|\underline{m}^{(k)}(z)| < \delta^{k+1}k!$ (see, e.g., Bai and Silverstein (2004), p. 585).

Now for $z \to 0^+$ we obtain

$$\underline{m}'(0) = \left[\frac{c}{1+\underline{m}(0)}\right]^{-2}\left(1+\frac{c\underline{m}'(0)}{(1+\underline{m}(0))^2}\right) = \left(\frac{1+\underline{m}(0)}{c}\right)^2 + \frac{\underline{m}'(0)}{c},$$

which is again a linear equation in $\underline{m}'(0)$. Thus, using (5.50) we have

$$(1-c^{-1})\underline{m}'(0) = \left(\frac{1+\dfrac{1}{c-1}}{c}\right)^2$$

As a result, we get

$$\underline{m}'(0) = \frac{c}{(c-1)^3}.$$

Finally, we calculate

$$\begin{aligned}
\underline{m}''(z) &= -2\left[z-\frac{c}{1+\underline{m}(z)}\right]^{-3}\left(1+\frac{c\underline{m}'(z)}{(1+\underline{m}(z))^2}\right)^2 \\
&+ \left[z-\frac{c}{1+\underline{m}(z)}\right]^{-2}\left(\frac{c\underline{m}''(z)((1+\underline{m}(z))^2)-2c(\underline{m}'(z))^2(1+\underline{m}(z))}{(1+\underline{m}(z))^4}\right)
\end{aligned}$$

which leads to

$$\underline{m}''(0) = 2\frac{(\underline{m}'(0))^2}{\underline{m}(0)} + (\underline{m}(0))^2\left(\frac{c\underline{m}''(0)}{(1+\underline{m}(0))^2} - \frac{2c(\underline{m}'(0))^2}{(1+\underline{m}(0))^3}\right)$$

and, consequently,

$$\underline{m}''(0) = 2\frac{c(c+1)}{(c-1)^5}.$$

As a result, we obtain (5.38)-(5.40).

For (5.41) we consider

$$\bar{\mathbf{x}}'\tilde{\mathbf{V}}_n^+\boldsymbol{\theta} = \mathrm{tr}\left[\mathbf{I}_p + z\left(\frac{1}{n}\mathbf{X}_n\mathbf{X}_n' - z\mathbf{I}_p\right)^{-1}\boldsymbol{\theta}\bar{\mathbf{x}}'\right] = \bar{\mathbf{x}}'\boldsymbol{\theta} + z\bar{\mathbf{x}}'\left(\tilde{\mathbf{V}}_n - z\mathbf{I}_p\right)^{-1}\boldsymbol{\theta}.$$

Because of (5.16), it holds that $\bar{\mathbf{x}}'\left(\tilde{\mathbf{V}}_n - z\mathbf{I}_p\right)^{-1}\boldsymbol{\theta}$ is uniformly bounded as $z \longrightarrow 0$. Moreover, $\bar{\mathbf{x}}'\boldsymbol{\theta} \xrightarrow{a.s.} 0$ as $p \longrightarrow \infty$ following Kolmogorov's strong law of large numbers (c.f., Sen and Singer (1993, Theorem 2.3.10), since $\boldsymbol{\theta}$ has a bounded Euclidean norm. Hence, $\bar{\mathbf{x}}'\tilde{\mathbf{V}}_n^+\boldsymbol{\theta} \xrightarrow{a.s.} 0$ for $p/n \longrightarrow c \in (1,+\infty)$ as $n \to \infty$.

Finally, in the case of (5.42) and (5.43), we get

$$\begin{aligned}
\bar{\mathbf{x}}'(\tilde{\mathbf{V}}_n^+)^2\boldsymbol{\theta} &= \frac{\partial}{\partial z}z\mathrm{tr}\left[\left(\frac{1}{n}\mathbf{X}_n\mathbf{X}_n' - z\mathbf{I}_p\right)^{-1}\boldsymbol{\theta}\bar{\mathbf{x}}'\right]\Bigg|_{z=0} \xrightarrow{a.s.} 0, \\
\bar{\mathbf{x}}'(\tilde{\mathbf{V}}_n^+)^3\boldsymbol{\theta} &= \frac{1}{2}\frac{\partial^2}{\partial z^2}z\mathrm{tr}\left[\left(\frac{1}{n}\mathbf{X}_n\mathbf{X}_n' - z\mathbf{I}_p\right)^{-1}\boldsymbol{\theta}\bar{\mathbf{x}}'\right]\Bigg|_{z=0} \xrightarrow{a.s.} 0.
\end{aligned}$$

for $p/n \longrightarrow c \in (1, +\infty)$ as $n \to \infty$. $\square$

Lemma 5.6. *Assume (A2). Let* $\boldsymbol{\theta}$ *and* $\boldsymbol{\xi}$ *be universal nonrandom vectors with bounded Euclidean norms. Then it holds that*

$$\begin{aligned}
\boldsymbol{\xi}'\mathbf{V}_n^{+}\boldsymbol{\theta} &\xrightarrow{a.s.} c^{-1}(c-1)^{-1}\boldsymbol{\xi}'\boldsymbol{\theta}\,, &(5.51)\\
\bar{\mathbf{x}}_n'\mathbf{V}_n^{+}\bar{\mathbf{x}}_n &\xrightarrow{a.s.} \frac{1}{c-1}\,, &(5.52)\\
\bar{\mathbf{x}}_n'\mathbf{V}_n^{+}\boldsymbol{\theta} &\xrightarrow{a.s.} 0\,, &(5.53)\\
\boldsymbol{\xi}'(\mathbf{V}_n^{+})^2\boldsymbol{\theta} &\xrightarrow{a.s.} (c-1)^{-3}\boldsymbol{\xi}'\boldsymbol{\theta}\,, &(5.54)\\
\bar{\mathbf{x}}_n'(\mathbf{V}_n^{+})^2\bar{\mathbf{x}}_n &\xrightarrow{a.s.} \frac{c}{(c-1)^3}\,, &(5.55)\\
\bar{\mathbf{x}}_n'(\mathbf{V}_n^{+})^2\boldsymbol{\theta} &\xrightarrow{a.s.} 0 &(5.56)
\end{aligned}$$

for $p/n \longrightarrow c \in (1, +\infty)$ *as* $n \to \infty$.

Proof of Lemma 5.6: From (5.5) we get

$$\begin{aligned}
\boldsymbol{\xi}'\mathbf{V}_n^{+}\boldsymbol{\theta} &= \boldsymbol{\xi}'\tilde{\mathbf{V}}_n^{+}\boldsymbol{\theta} - \frac{\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^{+}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^2\boldsymbol{\theta} + \boldsymbol{\xi}'(\tilde{\mathbf{V}}_n^{+})^2\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{+}\boldsymbol{\theta}}{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^2\bar{\mathbf{x}}_n}\\
&+ \frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^3\bar{\mathbf{x}}_n}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^2\bar{\mathbf{x}}_n)^2}\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^{+}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{+}\boldsymbol{\theta} \xrightarrow{a.s.} c^{-1}(c-1)^{-1}\boldsymbol{\xi}'\boldsymbol{\theta}
\end{aligned}$$

for $p/n \longrightarrow c \in (1, +\infty)$ as $n \to \infty$ following (5.35)-(5.37). Similarly, we get

$$\begin{aligned}
\bar{\mathbf{x}}_n'\mathbf{V}_n^{+}\bar{\mathbf{x}}_n &= \bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{+}\bar{\mathbf{x}}_n - \frac{\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{+}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^2\bar{\mathbf{x}}_n + \bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^2\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{+}\bar{\mathbf{x}}_n}{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^2\bar{\mathbf{x}}_n}\\
&+ \frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^3\bar{\mathbf{x}}_n}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^2\bar{\mathbf{x}}_n)^2}\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{+}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{+}\bar{\mathbf{x}}_n \xrightarrow{a.s.} \frac{1}{c-1}
\end{aligned}$$

and

$$\begin{aligned}
\bar{\mathbf{x}}_n'\mathbf{V}_n^{+}\boldsymbol{\theta} &= \bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{+}\boldsymbol{\theta} - \frac{\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{+}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^2\boldsymbol{\theta} + \bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^2\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{+}\boldsymbol{\theta}}{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^2\bar{\mathbf{x}}_n}\\
&+ \frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^3\bar{\mathbf{x}}_n}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^{+})^2\bar{\mathbf{x}}_n)^2}\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{+}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^{+}\boldsymbol{\theta} \xrightarrow{a.s.} 0
\end{aligned}$$

for $p/n \longrightarrow c \in (1, +\infty)$ as $n \to \infty$.

Now, we consider the equality

$$\begin{aligned}
\left[(\tilde{\mathbf{V}}_n - \bar{\mathbf{x}}_n\bar{\mathbf{x}}_n')^+\right]^2 &= \left(\tilde{\mathbf{V}}_n^+ - \frac{\tilde{\mathbf{V}}_n^+\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2 + (\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)}{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n} + \frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n)^2}\tilde{\mathbf{V}}_n^+\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+\right)^2 \\
&= (\tilde{\mathbf{V}}_n^+)^2 + \left[\frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n)^2}\tilde{\mathbf{V}}_n^+\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+ - \frac{(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+ + \tilde{\mathbf{V}}_n^+\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2}{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n}\right]^2 \\
&- \frac{2(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2 + \tilde{\mathbf{V}}_n^+\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3 + (\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+}{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n} \\
&+ \frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n)^2}(\tilde{\mathbf{V}}_n^+\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2 + (\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+) \\
&= (\tilde{\mathbf{V}}_n^+)^2 + \frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^4\bar{\mathbf{x}}_n}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n)^2}\tilde{\mathbf{V}}_n^+\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+ - \frac{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n)^2}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n)^3}\tilde{\mathbf{V}}_n^+\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+ \\
&+ \frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n)^2}\left[(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+ + \tilde{\mathbf{V}}_n^+\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\right] \\
&- \frac{(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2 + \tilde{\mathbf{V}}_n^+\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3 + (\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+}{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n}
\end{aligned}$$

Hence,

$$\begin{aligned}
\boldsymbol{\xi}'(\mathbf{V}_n^+)^2\boldsymbol{\theta} &= \boldsymbol{\xi}'(\tilde{\mathbf{V}}_n^+)^2\boldsymbol{\theta} + \frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^4\bar{\mathbf{x}}_n}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n)^2}\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^+\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+\boldsymbol{\theta} - \frac{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n)^2}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n)^3}\boldsymbol{\xi}'\tilde{\mathbf{V}}_n^+\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+\boldsymbol{\theta} \\
&+ \frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n)^2}\left[\boldsymbol{\xi}'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+\boldsymbol{\theta} + \boldsymbol{\xi}'\tilde{\mathbf{V}}_n^+\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\boldsymbol{\theta}\right] \\
&- \frac{\boldsymbol{\xi}'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\boldsymbol{\theta} + \boldsymbol{\xi}'\tilde{\mathbf{V}}_n^+\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3\boldsymbol{\theta} + \boldsymbol{\xi}'(\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+\boldsymbol{\theta}}{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n} \xrightarrow{a.s.} (c-1)^{-3}\boldsymbol{\xi}'\boldsymbol{\theta}\,,
\end{aligned}$$

$$\begin{aligned}
\bar{\mathbf{x}}_n'(\mathbf{V}_n^+)^2\bar{\mathbf{x}}_n &= \bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n + \frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^4\bar{\mathbf{x}}_n}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n)^2} - \frac{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n)^2}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n)^3} + 2\frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n}{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n} \\
&- \bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n - 2\frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n}{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n} \\
&= \frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^4\bar{\mathbf{x}}_n}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n)^2} - \frac{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n)^2}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n)^3} \xrightarrow{a.s.} \frac{c}{(c-1)^3}\,,
\end{aligned}$$

and

$$\begin{aligned}
\bar{\mathbf{x}}_n'(\mathbf{V}_n^+)^2\boldsymbol{\theta} &= \bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\boldsymbol{\theta} + \frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^4\bar{\mathbf{x}}_n}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n)^2}\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+\boldsymbol{\theta} - \frac{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n)^2}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n)^3}\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+\boldsymbol{\theta} \\
&+ \frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n}{(\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n)^2}\left[\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+\boldsymbol{\theta} + \bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\boldsymbol{\theta}\right] \\
&- \frac{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\boldsymbol{\theta} + \bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3\boldsymbol{\theta} + \bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^3\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n'\tilde{\mathbf{V}}_n^+\boldsymbol{\theta}}{\bar{\mathbf{x}}_n'(\tilde{\mathbf{V}}_n^+)^2\bar{\mathbf{x}}_n} \xrightarrow{a.s.} 0
\end{aligned}$$

for $p/n \longrightarrow c \in (1, +\infty)$ as $n \to \infty$. □

Lemma 5.7. *Assume (A2). Let $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$ be universal nonrandom vectors with bounded Euclidean norms and let $\mathbf{P}_n^+ = \mathbf{V}_n^+ - \frac{\mathbf{V}_n^+\boldsymbol{\eta}\boldsymbol{\eta}'\mathbf{V}_n^+}{\boldsymbol{\eta}'\mathbf{V}_n^+\boldsymbol{\eta}}$ where $\boldsymbol{\eta}$ is a universal nonrandom vectors with bounded Euclidean norm. Then it holds that*

$$\boldsymbol{\xi}'\mathbf{P}_n^+\boldsymbol{\theta} \xrightarrow{a.s.} c^{-1}(c-1)^{-1}\left(\boldsymbol{\xi}'\boldsymbol{\theta} - \frac{\boldsymbol{\xi}'\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{\theta}}{\boldsymbol{\eta}'\boldsymbol{\eta}}\right), \tag{5.57}$$

$$\bar{\mathbf{x}}_n'\mathbf{P}_n^+\bar{\mathbf{x}}_n \xrightarrow{a.s.} \frac{1}{c-1}, \tag{5.58}$$

$$\bar{\mathbf{x}}_n'\mathbf{P}_n^+\boldsymbol{\theta} \xrightarrow{a.s.} 0, \tag{5.59}$$

$$\boldsymbol{\xi}'(\mathbf{P}_n^+)^2\boldsymbol{\theta} \xrightarrow{a.s.} (c-1)^{-3}\left(\boldsymbol{\xi}'\boldsymbol{\theta} - \frac{\boldsymbol{\xi}'\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{\theta}}{\boldsymbol{\eta}'\boldsymbol{\eta}}\right), \tag{5.60}$$

$$\bar{\mathbf{x}}_n'(\mathbf{P}_n^+)^2\bar{\mathbf{x}}_n \xrightarrow{a.s.} \frac{c}{(c-1)^3}, \tag{5.61}$$

$$\bar{\mathbf{x}}_n'(\mathbf{P}_n^+)^2\boldsymbol{\theta} \xrightarrow{a.s.} 0 \tag{5.62}$$

for $p/n \longrightarrow c \in (1, +\infty)$ as $n \to \infty$.

Proof of Lemma 5.7: It holds that

$$\boldsymbol{\xi}'\mathbf{P}_n^+\boldsymbol{\theta} = \boldsymbol{\xi}'\mathbf{V}_n^+\boldsymbol{\theta} - \frac{\boldsymbol{\xi}'\mathbf{V}_n^+\boldsymbol{\eta}\boldsymbol{\eta}'\mathbf{V}_n^+\boldsymbol{\theta}}{\boldsymbol{\eta}'\mathbf{V}_n^+\boldsymbol{\eta}} \xrightarrow{a.s.} c^{-1}(c-1)^{-1}\left(\boldsymbol{\xi}'\boldsymbol{\theta} - \frac{\boldsymbol{\xi}'\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{\theta}}{\boldsymbol{\eta}'\boldsymbol{\eta}}\right)$$

for $p/n \longrightarrow c \in (1, +\infty)$ as $n \to \infty$ following (5.35). Similarly, we get

$$\bar{\mathbf{x}}_n'\mathbf{P}_n^+\bar{\mathbf{x}}_n = \bar{\mathbf{x}}_n'\mathbf{V}_n^+\bar{\mathbf{x}}_n - \frac{\bar{\mathbf{x}}_n'\mathbf{V}_n^+\boldsymbol{\eta}\boldsymbol{\eta}'\mathbf{V}_n^+\bar{\mathbf{x}}_n}{\boldsymbol{\eta}'\mathbf{V}_n^+\boldsymbol{\eta}} \xrightarrow{a.s.} \frac{1}{c-1}$$

and

$$\bar{\mathbf{x}}_n'\mathbf{P}_n^+\boldsymbol{\theta} = \bar{\mathbf{x}}_n'\mathbf{V}_n^+\boldsymbol{\theta} - \frac{\bar{\mathbf{x}}_n'\mathbf{V}_n^+\boldsymbol{\eta}\boldsymbol{\eta}'\mathbf{V}_n^+\boldsymbol{\theta}}{\boldsymbol{\eta}'\mathbf{V}_n^+\boldsymbol{\eta}} \xrightarrow{a.s.} 0$$

for $p/n \longrightarrow c \in (1, +\infty)$ as $n \to \infty$.

The rest of the proof follows from the equality

$$(\mathbf{P}_n^+)^2 = (\mathbf{V}_n^+)^2 - \frac{(\mathbf{V}_n^+)^2\boldsymbol{\eta}\boldsymbol{\eta}'\mathbf{V}_n^+}{\boldsymbol{\eta}'\mathbf{V}_n^+\boldsymbol{\eta}} - \frac{\mathbf{V}_n^+\boldsymbol{\eta}\boldsymbol{\eta}'(\mathbf{V}_n^+)^2}{\boldsymbol{\eta}'\mathbf{V}_n^+\boldsymbol{\eta}} + \boldsymbol{\eta}'(\mathbf{V}_n^+)^2\boldsymbol{\eta}\frac{\mathbf{V}_n^+\boldsymbol{\eta}\boldsymbol{\eta}'\mathbf{V}_n^+}{(\boldsymbol{\eta}'\mathbf{V}_n^+\boldsymbol{\eta})^2}$$

and Lemma 5.6. $\square$

Proof of Theorem 2.2: In case of $c > 1$, the optimal shrinkage intensity is given by

$$\begin{aligned}
\alpha_n^+ &= \gamma^{-1}\frac{\hat{\mathbf{w}}_{S^*}'(\boldsymbol{\mu}_n-\gamma\boldsymbol{\Sigma}_n\mathbf{b})-\mathbf{b}'(\boldsymbol{\mu}_n-\gamma\boldsymbol{\Sigma}_n\mathbf{b})}{\hat{\mathbf{w}}_{S^*}'\boldsymbol{\Sigma}_n\hat{\mathbf{w}}_{S^*}-2\mathbf{b}'\boldsymbol{\Sigma}_n\hat{\mathbf{w}}_{S^*}+\mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{b}}\\
&= \gamma^{-1}\frac{\dfrac{\mathbf{1}'\mathbf{S}_n^*(\boldsymbol{\mu}_n-\gamma\boldsymbol{\Sigma}_n\mathbf{b})}{\mathbf{1}'\mathbf{S}_n^*\mathbf{1}}+\gamma^{-1}\bar{\mathbf{y}}_n'\hat{\mathbf{Q}}_n^*(\boldsymbol{\mu}_n-\gamma\boldsymbol{\Sigma}_n\mathbf{b})-\mathbf{b}'(\boldsymbol{\mu}_n-\gamma\boldsymbol{\Sigma}_n\mathbf{b})}{\dfrac{\mathbf{1}'\mathbf{S}_n^*\boldsymbol{\Sigma}_n\mathbf{S}_n^*\mathbf{1}}{(\mathbf{1}'\mathbf{S}_n^*\mathbf{1})^2}+2\gamma^{-1}\dfrac{\bar{\mathbf{y}}_n'\hat{\mathbf{Q}}_n^*\boldsymbol{\Sigma}_n\mathbf{S}_n^*\mathbf{1}}{\mathbf{1}'\mathbf{S}_n^*\mathbf{1}}+\gamma^{-2}\bar{\mathbf{y}}_n'\hat{\mathbf{Q}}_n^*\boldsymbol{\Sigma}_n\hat{\mathbf{Q}}_n^*\bar{\mathbf{y}}_n-2\dfrac{\mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{S}_n^*\mathbf{1}}{\mathbf{1}'\mathbf{S}_n^*\mathbf{1}}-2\gamma^{-1}\mathbf{b}'\boldsymbol{\Sigma}_n\hat{\mathbf{Q}}_n^*\bar{\mathbf{y}}_n+\mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{b}},
\end{aligned}$$

where $\bar{\mathbf{y}}_n=\boldsymbol{\mu}_n+\boldsymbol{\Sigma}_n^{1/2}\bar{\mathbf{x}}_n$ and $\mathbf{S}_n^*=\boldsymbol{\Sigma}_n^{-1/2}\mathbf{V}_n^+\boldsymbol{\Sigma}_n^{-1/2}$.

From Assumption (A3), we get that the following vectors $\boldsymbol{\Sigma}_n^{-1/2}\mathbf{1}$, $\boldsymbol{\Sigma}_n^{-1/2}(\boldsymbol{\mu}_n-\gamma\boldsymbol{\Sigma}_n\mathbf{b})$, $\boldsymbol{\Sigma}_n^{-1/2}\boldsymbol{\mu}_n$, and $\boldsymbol{\Sigma}_n^{1/2}\mathbf{b}$ possess bounded Euclidean norms. As a result, the application of Lemma 5.2 leads to

$$\begin{aligned}
\mathbf{1}'\mathbf{S}_n^*\mathbf{1} &= \mathbf{1}'\boldsymbol{\Sigma}_n^{-1/2}\mathbf{V}_n^+\boldsymbol{\Sigma}_n^{-1/2}\mathbf{1}\xrightarrow{a.s.}c^{-1}(c-1)^{-1}\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1},\\
\mathbf{1}'\mathbf{S}_n^*(\boldsymbol{\mu}_n-\gamma\boldsymbol{\Sigma}_n\mathbf{b}) &= \mathbf{1}'\boldsymbol{\Sigma}_n^{-1/2}\mathbf{V}_n^+\boldsymbol{\Sigma}_n^{-1/2}(\boldsymbol{\mu}_n-\gamma\boldsymbol{\Sigma}_n\mathbf{b})\xrightarrow{a.s.}c^{-1}(c-1)^{-1}\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}(\boldsymbol{\mu}_n-\gamma\boldsymbol{\Sigma}_n\mathbf{b}),\\
\mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{S}_n^*\mathbf{1} &= \mathbf{b}'\boldsymbol{\Sigma}_n\boldsymbol{\Sigma}_n^{-1/2}\mathbf{V}_n^+\boldsymbol{\Sigma}_n^{-1/2}\mathbf{1}\xrightarrow{a.s.}c^{-1}(c-1)^{-1}\mathbf{b}'\boldsymbol{\Sigma}_n\boldsymbol{\Sigma}_n^{-1}\mathbf{1}=c^{-1}(c-1)^{-1},\\
\mathbf{1}'\mathbf{S}_n^*\boldsymbol{\Sigma}_n\mathbf{S}_n^*\mathbf{1} &= \mathbf{1}'\boldsymbol{\Sigma}_n^{-1/2}(\mathbf{V}_n^+)^2\boldsymbol{\Sigma}_n^{-1/2}\mathbf{1}\xrightarrow{a.s.}(c-1)^{-3}\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}
\end{aligned}$$

for $p/n\longrightarrow c\in(1,+\infty)$ as $n\to\infty$.

Finally, from Lemma 5.3 and 5.4 as well as by using the equalities

$$\hat{\mathbf{Q}}_n^*=\boldsymbol{\Sigma}_n^{-1/2}\mathbf{P}_n^+\boldsymbol{\Sigma}_n^{-1/2}\quad\text{and}\quad \mathbf{P}_n^+\mathbf{V}_n^+=(\mathbf{V}_n^+)^2-\frac{\mathbf{V}_n^+\boldsymbol{\eta}\boldsymbol{\eta}'(\mathbf{V}_n^+)^2}{\boldsymbol{\eta}'\mathbf{V}_n^+\boldsymbol{\eta}}$$

with $\boldsymbol{\eta}=\boldsymbol{\Sigma}_n^{-1/2}\mathbf{1}$ we obtain

$$\begin{aligned}
\bar{\mathbf{y}}_n'\hat{\mathbf{Q}}_n^*(\boldsymbol{\mu}_n-\gamma\boldsymbol{\Sigma}_n\mathbf{b}) &= \boldsymbol{\mu}_n'\boldsymbol{\Sigma}_n^{-1/2}\mathbf{P}_n^+\boldsymbol{\Sigma}_n^{-1/2}(\boldsymbol{\mu}_n-\gamma\boldsymbol{\Sigma}_n\mathbf{b})+\bar{\mathbf{x}}_n'\mathbf{P}_n\boldsymbol{\Sigma}_n^{-1/2}(\boldsymbol{\mu}_n-\gamma\boldsymbol{\Sigma}_n\mathbf{b})\\
&\xrightarrow{a.s.}c^{-1}(c-1)^{-1}\boldsymbol{\mu}_n'\mathbf{Q}_n(\boldsymbol{\mu}_n-\gamma\boldsymbol{\Sigma}_n\mathbf{b}),\\
\bar{\mathbf{y}}_n'\hat{\mathbf{Q}}_n^*\boldsymbol{\Sigma}_n\mathbf{S}_n^*\mathbf{1} &= \boldsymbol{\mu}_n'\boldsymbol{\Sigma}_n^{-1/2}\mathbf{P}_n^+\mathbf{V}_n^+\boldsymbol{\Sigma}_n^{-1/2}\mathbf{1}+\bar{\mathbf{x}}_n'\mathbf{P}_n^+\mathbf{V}_n^+\boldsymbol{\Sigma}_n^{-1/2}\mathbf{1}\\
&\xrightarrow{a.s.}(c-1)^{-3}\boldsymbol{\mu}_n'\mathbf{Q}_n\mathbf{1}=0,\\
\bar{\mathbf{y}}_n'\hat{\mathbf{Q}}_n^*\boldsymbol{\Sigma}_n\hat{\mathbf{Q}}_n^*\bar{\mathbf{y}}_n &= \bar{\mathbf{x}}_n'(\mathbf{P}_n^+)^2\bar{\mathbf{x}}_n+2\boldsymbol{\mu}_n'\boldsymbol{\Sigma}_n^{-1/2}(\mathbf{P}_n^+)^2\bar{\mathbf{x}}_n+\boldsymbol{\mu}_n'\boldsymbol{\Sigma}_n^{-1/2}(\mathbf{P}_n^+)^2\boldsymbol{\Sigma}_n^{-1/2}\boldsymbol{\mu}_n\\
&\xrightarrow{a.s.}\frac{c}{(c-1)^3}+(c-1)^{-3}\boldsymbol{\mu}_n'\mathbf{Q}_n\boldsymbol{\mu}_n,\\
\mathbf{b}'\boldsymbol{\Sigma}_n\hat{\mathbf{Q}}_n^*\bar{\mathbf{y}}_n &= \mathbf{b}'\boldsymbol{\Sigma}_n^{1/2}\mathbf{P}_n^+\bar{\mathbf{x}}_n+\mathbf{b}'\boldsymbol{\Sigma}_n^{1/2}\mathbf{P}_n^+\boldsymbol{\Sigma}_n^{-1/2}\boldsymbol{\mu}_n\\
&\xrightarrow{a.s.}c^{-1}(c-1)^{-1}\left(\mathbf{b}'\boldsymbol{\mu}_n-\frac{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\mu}_n}{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}}\right)
\end{aligned}$$

for $p/n\longrightarrow c\in(1,+\infty)$ as $n\to\infty$.

Hence, $\alpha_n^+\xrightarrow{a.s.}\alpha^+$, where

$$\alpha^+=\gamma^{-1}\frac{\dfrac{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}(\boldsymbol{\mu}_n-\gamma\boldsymbol{\Sigma}_n\mathbf{b})}{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}}+\dfrac{\gamma^{-1}}{c(c-1)}\boldsymbol{\mu}_n'\mathbf{Q}_n(\boldsymbol{\mu}_n-\gamma\boldsymbol{\Sigma}_n\mathbf{b})-\mathbf{b}'(\boldsymbol{\mu}_n-\gamma\boldsymbol{\Sigma}_n\mathbf{b})}{\dfrac{c^2}{(c-1)}\dfrac{1}{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}}+\dfrac{\gamma^{-2}}{(c-1)^3}(\boldsymbol{\mu}_n'\mathbf{Q}_n\boldsymbol{\mu}_n+c)-2\dfrac{1}{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}}-2\frac{\gamma^{-1}}{c(c-1)}\left(\mathbf{b}'\boldsymbol{\mu}_n-\dfrac{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\mu}_n}{\mathbf{1}'\boldsymbol{\Sigma}_n^{-1}\mathbf{1}}\right)+\mathbf{b}'\boldsymbol{\Sigma}_n\mathbf{b}}$$

This completes the proof of Theorem 2.2. □

References

Bai, J. and Shi, S. (2011). Estimating high dimensional covariance matrices and its applications. Annals of Economics and Finance, 12:199–215.

Bai, Z. and Zhou, W. (2008). Large sample covariance matrices without independence structures in columns. Statistica Sinica, 18(2):425–442.

Bai, Z. D. and Silverstein, J. W. (2004). CLT for linear spectral statistics of large dimensional sample covariance matrices. Annals of Probability, 32:553–605.

Bai, Z. D. and Silverstein, J. W. (2010). Spectral Analysis of Large Dimensional Random Matrices. Springer, New York.

Best, M. J. and Grauer, R. R. (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. The Review of Financial Studies, 4:315–342.

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

Bodnar, T., Parolya, N., and Schmid, W. (2013). On the equivalence of quadratic optimization problems commonly used in portfolio theory. European Journal of Operational Research, 229:637–644.

Bodnar, T., Parolya, N., and Schmid, W. (2017). Estimation of the global minimum variance portfolio in high dimensions. European Journal of Operational Research, under revision.

Bodnar, T. and Schmid, W. (2011). On the exact distribution of the estimated expected utility portfolio weights: Theory and applications. Statistics & Risk Modeling, 28:319–342.

Chopra, V. K. and Ziemba, W. T. (1993). The effect of errors in means, variances, and covariances on optimal portfolio choice. The Review of Financial Studies, 19:6–11.

DeMiguel, V., L., G., and U., R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? The Review of Financial Studies, 22:1915–1953.

El Karoui, N. (2010). High-dimensionality effects in the Markowitz problem and other quadratic programs with linear constraints: Risk underestimation. Annals of Statistics, 38:3487–3566.

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

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

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

Frahm, G. and Memmel, C. (2010). Dominating estimators for minimum-variance portfolios. Journal of Econometrics, 159:289–302.

Friesen, O., Löwe, M., and Stolz, M. (2013). Gaussian fluctuations for sample covariance matrices with dependent data. Journal of Multivariate Analysis, 114:270–287.

Golosnoy, V. and Okhrin, Y. (2007). Multivariate shrinkage for optimal portfolio weights. European Journal of Finance, 13:441–458.

Horn, R. A. and Johnsohn, C. R. (1985). Matrix Analysis. Cambridge University Press, Cambridge.

Hui, J. and Pan, G. (2010). Limiting spectral distribution for large sample covariance matrices with m-dependent elements. Communications in Statistics - Theory and Methods, 39(6):935–941.

Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. The Journal of Financial and Quantitative Analysis, 21:279–292.

Le Cam, L. and Lo Yang, G. (2000). Asymptotics in statistics: some basic concepts. Springer, New York.

Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. Journal of Multivariate Analysis, 88:365–411.

Marčenko, V. A. and Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. Sbornik: Mathematics, 1:457–483.

Markowitz, H. (1952). Portfolio selection. The Journal of Finance, 7:77–91.

Merton, C. R. (1980). On estimating the expected return on the market: An exploratory investigation. Journal of Financial Economics, 8:323–361.

Meyer, C. R. (1973). Generalized inversion of modified matrices. SIAM Journal of Applied Mathematics, 24:315–323.

Okhrin, Y. and Schmid, W. (2006). Distributional properties of portfolio weights. Journal of Econometrics, 134:235–256.

Pan, G. M. (2014). Comparison between two types of large sample covariance matrices. Ann. Inst. H. Poincaré Probab. Statist., 50:655–677.

Pratt, J. W. (1964). Risk aversion in the small and in the large. Econometrica, 32:122–136.

Rubio, F. and Mestre, X. (2011). Spectral convergence for a general class of random matrices. Statistics & Probability Letters, 81:592–602.

Silverstein, J. W. (1995). Strong convergence of the empirical distribution of eigenvalues of large-dimensional random matrices. Journal of Multivariate Analysis, 55:331–339.

Tobin, J. (1958). Liquidity preference as behavior towards risk. Review of Economic Studies, 25:65–86.

Wei, M., Yang, G., and Yang, L. (2016). The limiting spectral distribution for large sample covariance matrices with unbounded m-dependent entries. Communications in Statistics - Theory and Methods, 45(22):6651–6662.

6 Appendix B

Tables and Figures

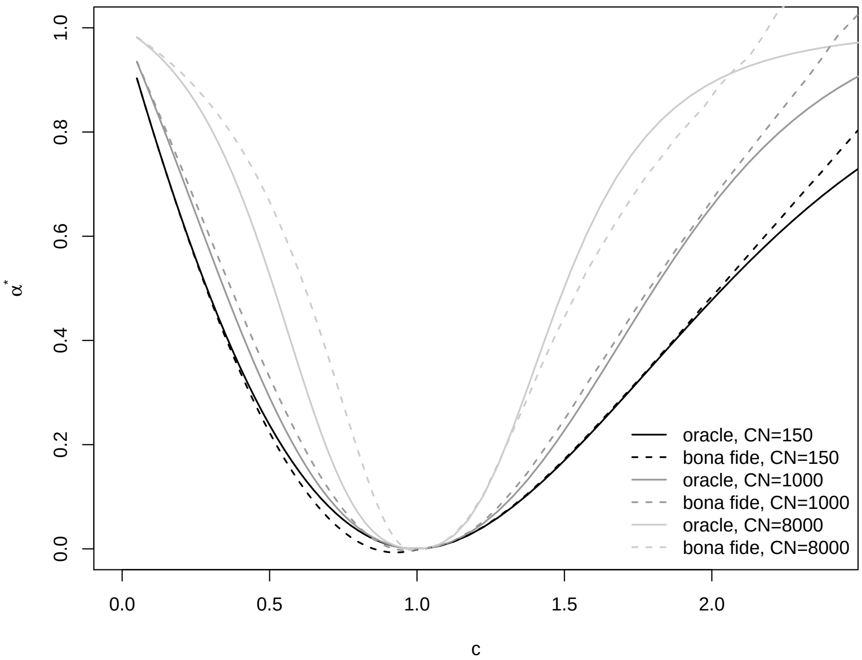

Figure 1: The asymptotic optimal shrinkage intensity as a function of c (left) and of the sample size n (right).

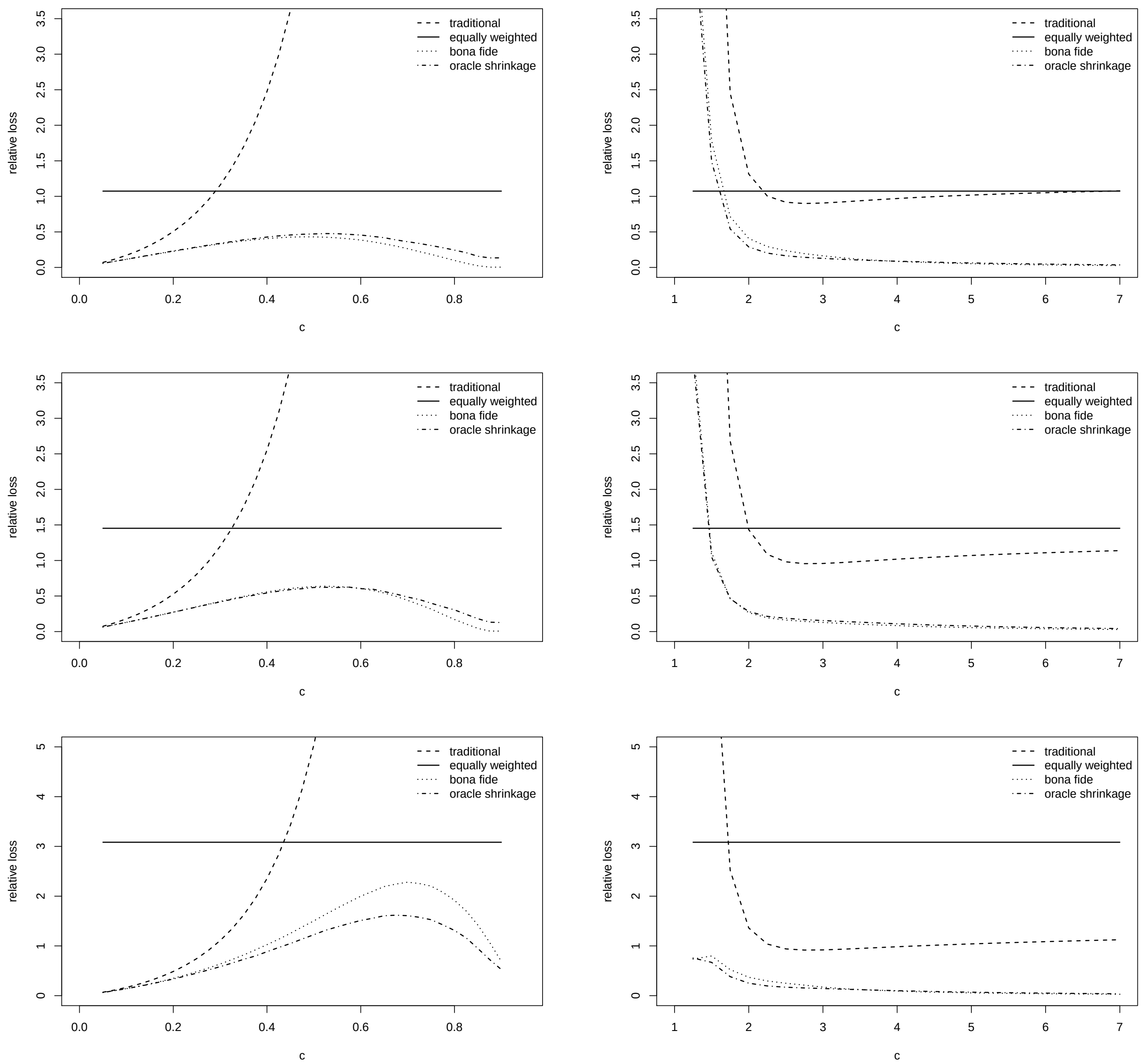

Figure 2: The relative losses for the portfolios based on optimal shrinkage estimator, the traditional estimator, the bona-fide estimator and the equally weighted portfolio as a function of c for different values of c, fixed $p = 150$ and different values of the condition index (150, 1000, 8000 from top to bottom).

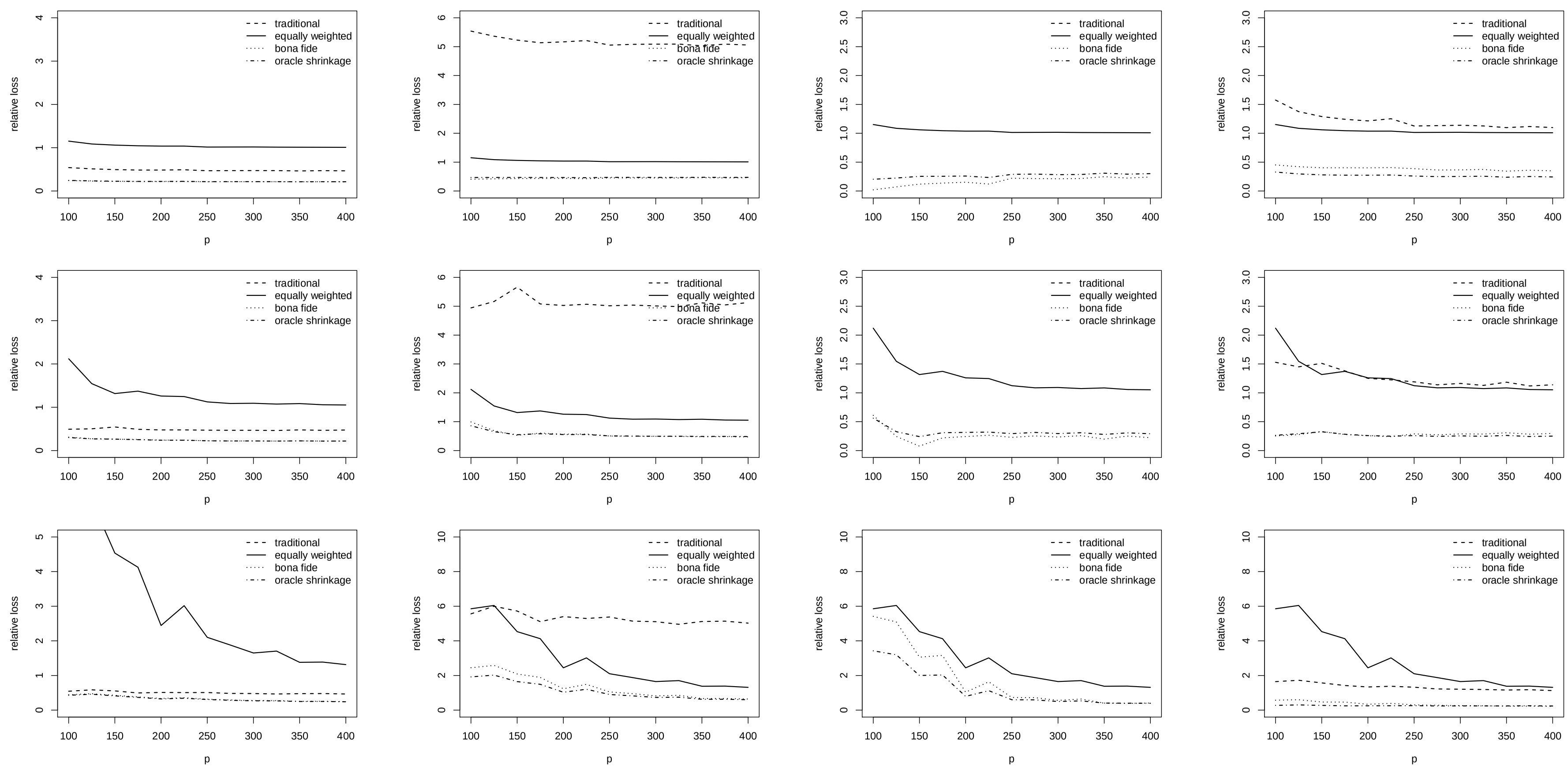

Figure 3: The relative losses for the portfolios based in the optimal shrinkage estimator, the traditional estimator, the bona-fide estimator and the equally weighted portfolio as a function of the dimension p, for different values of the condition index (150, 1000, 8000 from top to bottom) and different values of c $(0.2,\ 0.5,\ 0.8, 2$ from left to right).

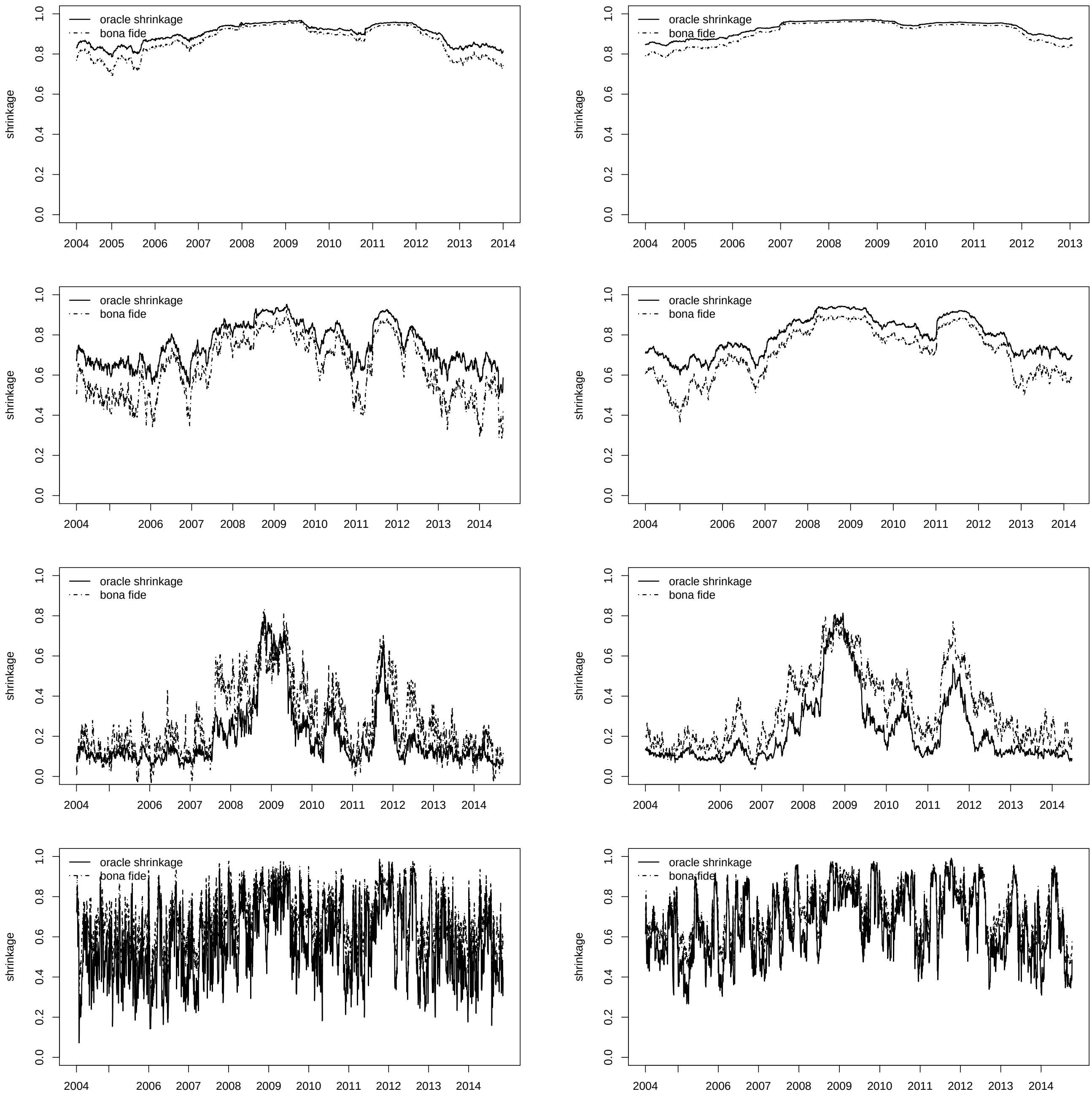

Figure 4: The oracle and the bona fide shrinkage coefficients for the first 50 assets (left) and 100 assets (right) and $c = 0.2$, 0.5, 0.8 and 2 (from top to bottom).

$p = 50$ and $c = 0.2$								
	CE		SR		VaR		ES	
	average	median	average	median	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
trad	-1.76555	-1.71384	0.04102	0.04408	0.02672	0.34622	-1.24698	-2.22581
oracle	-1.66281	-1.6279	0.03836	0.04062	-0.95527	-1.44478	-0.93466	-1.40688
bona fide	-1.64487	-1.61976	0.03703	0.03898	-0.95752	-1.46251	-1.298	-1.75999
equal	-2.46546	-2.45708	0.01747	0.01667	-0.96238	-1.48418	-1.30493	-1.75763
gmv	-1.7044	-1.66003	0.04474	0.04742	-1.24698	-2.22581	-1.31805	-1.77074

$p = 50$ and $c = 0.5$								
	CE		SR		VaR		ES	
	average	median	average	median	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
trad	-3.10895	-2.86948	0.04032	0.04331	0.03955	0.55934	-1.25216	-2.2334
oracle	-2.19538	-2.1336	0.03413	0.03582	-1.20093	-1.86062	-1.1323	-1.74002
bona fide	-2.05219	-2.01325	0.03172	0.03291	-1.0931	-1.69468	-1.67068	-2.3468
equal	-2.47703	-2.46523	0.0176	0.01776	-1.08567	-1.76208	-1.50844	-2.05655
gmv	-2.75714	-2.59012	0.0523	0.05292	-1.25216	-2.2334	-1.53213	-2.08616

$p = 50$ and $c = 0.8$								
	CE		SR		VaR		ES	
	average	median	average	median	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
trad	-18.80143	-17.72463	0.02055	0.02023	0.04277	1.74382	-1.2472	-2.22427
oracle	-2.39894	-2.3866	0.02031	0.01865	-2.879	-4.87552	-1.97098	-3.21846
bona fide	-2.6238	-2.60074	0.01976	0.01916	-1.22955	-2.06654	-4.3227	-6.63222
equal	-2.47045	-2.45915	0.01769	0.01775	-1.26809	-2.0478	-1.76774	-2.369
gmv	-8.67635	-8.15525	0.03241	0.03245	-1.2472	-2.22427	-1.78173	-2.40348

$p = 50$ and $c = 2$								
	CE		SR		VaR		ES	
	average	median	average	median	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
trad	-24.44784	-23.7383	-0.01239	-0.01112	0.00525	1.25693	-1.24935	-2.22743
oracle	-9.38714	-9.17254	-0.01261	-0.01219	-3.51657	-5.97855	-1.8153	-2.84675
bona fide	-12.18357	-11.86633	-0.01316	-0.01154	-2.23072	-3.59898	-5.30814	-8.08356
equal	-2.47412	-2.46527	0.01819	0.01805	-2.52278	-4.14614	-3.22776	-4.80051
gmv	-6.27938	-6.13782	0.00498	0.00615	-1.24935	-2.22743	-3.6993	-5.53501

Table 1: Certainty equivalent, Sharpe ratio, Value-at-Risk and expected shortfall averaged over 1000 randomly chosen portfolios of size 50 and different values of c. The trading period is fixed at 200 days from 01.01.2014 to 10.10.2014.

$p = 100$ and $c = 0.2$

	CE		SR		VaR		ES	
	average	median	average	median	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
trad	-1.50488	-1.42101	0.07882	0.07822	0.04445	0.29945	-1.24151	-2.23059
oracle	-1.43113	-1.37694	0.07404	0.07361	-0.83142	-1.22887	-0.81285	-1.19665
bona fide	-1.41588	-1.3674	0.07227	0.07182	-0.84239	-1.25081	-1.11574	-1.51356
equal	-2.40043	-2.40332	0.01838	0.01793	-0.84853	-1.26252	-1.13456	-1.53018
gmv	-1.45281	-1.37389	0.08036	0.07994	-1.24151	-2.23059	-1.14451	-1.53959

$p = 100$ and $c = 0.5$

	CE		SR		VaR		ES	
	average	median	average	median	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
trad	-2.49811	-2.46517	0.0491	0.04971	0.03676	0.45061	-1.23692	-2.22751
oracle	-1.91599	-1.90005	0.04445	0.045	-1.11946	-1.71233	-1.03912	-1.58793
bona fide	-1.81656	-1.80127	0.04215	0.04214	-1.0366	-1.57957	-1.5364	-2.13475
equal	-2.3951	-2.39566	0.01829	0.01801	-1.02879	-1.61281	-1.40842	-1.91279
gmv	-2.21631	-2.18907	0.05454	0.05278	-1.23692	-2.22751	-1.42153	-1.91837

$p = 100$ and $c = 0.8$

	CE		SR		VaR		ES	
	average	median	average	median	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
trad	-13.80483	-12.91635	0.01926	0.02058	0.05212	1.30538	-1.23559	-2.21778
oracle	-2.3052	-2.30347	0.01972	0.01965	-2.51969	-3.98799	-1.72483	-2.70894
bona fide	-2.47104	-2.45862	0.01965	0.01888	-1.2176	-2.09414	-3.56849	-5.13616
equal	-2.39085	-2.38622	0.01833	0.01824	-1.24107	-2.05829	-1.76804	-2.34152
gmv	-6.47477	-6.21022	0.0456	0.04673	-1.23559	-2.21778	-1.76762	-2.3518

$p = 100$ and $c = 2$

	CE		SR		VaR		ES	
	average	median	average	median	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
trad	-63.01361	-61.25124	-0.00829	-0.00975	0.00518	1.43255	-1.24095	-2.22977
oracle	-26.16811	-25.54021	-0.01134	-0.01164	-5.38551	-9.20075	-1.88699	-2.96903
bona fide	-30.1065	-29.23551	-0.0116	-0.01258	-3.57836	-5.9384	-8.07706	-12.15231
equal	-2.39875	-2.39528	0.01783	0.01755	-3.83868	-6.30492	-5.22286	-7.63462
gmv	-7.15755	-7.02123	0.00445	0.00353	-1.24095	-2.22977	-5.56355	-8.10265

Table 2: Certainty equivalent, Sharpe ratio, Value-at-Risk and expected shortfall averaged over 1000 randomly chosen portfolios of size 100 and different values of c. The trading period is fixed at 200 days from 01.01.2014 to 10.10.2014.

	$p = 300$ and $c = 0.2$							
	CE		SR		VaR		ES	
	average	median	average	median	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
trad	-1.46987	-1.46758	0.0976	0.09865	0.05329	0.30154	-1.24017	-2.25125
oracle	-1.45133	-1.44916	0.09577	0.09693	-0.78201	-1.33434	-0.77598	-1.33958
bona fide	-1.44735	-1.44579	0.09533	0.09653	-0.77761	-1.32319	-1.19482	-1.99672
equal	-2.35709	-2.35673	0.01836	0.0184	-0.77692	-1.32071	-1.19054	-2.0004
gmv	-1.45443	-1.45252	0.09732	0.09887	-1.24017	-2.25125	-1.18975	-2.0013

	$p = 300$ and $c = 0.5$							
	CE		SR		VaR		ES	
	average	median	average	median	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
trad	-1.79542	-1.76418	0.07044	0.07	0.04827	0.31888	-1.23951	-2.24925
oracle	-1.51127	-1.50021	0.06434	0.0643	-0.8869	-1.29331	-0.82115	-1.2044
bona fide	-1.46261	-1.45231	0.06235	0.06242	-0.85692	-1.2443	-1.16634	-1.52194
equal	-2.35432	-2.35385	0.01839	0.01837	-0.85904	-1.2505	-1.12257	-1.45355
gmv	-1.54615	-1.51707	0.08519	0.08328	-1.23951	-2.24925	-1.12727	-1.46092

	$p = 300$ and $c = 0.8$							
	CE		SR		VaR		ES	
	average	median	average	median	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
trad	-9.06343	-8.96082	0.01513	0.01544	0.02813	0.84991	-1.23959	-2.24892
oracle	-2.20963	-2.20671	0.01901	0.01914	-2.16156	-3.34203	-1.46594	-2.24366
bona fide	-2.4478	-2.43266	0.01839	0.01848	-1.25743	-2.04462	-2.99941	-4.21906
equal	-2.35577	-2.35489	0.01824	0.01837	-1.29321	-1.9974	-1.77034	-2.29253
gmv	-4.22143	-4.22323	0.0306	0.03216	-1.23959	-2.24892	-1.77521	-2.31476

	$p = 300$ and $c = 2$							
	CE		SR		VaR		ES	
	average	median	average	median	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
trad	-85.42486	-84.6437	-0.03032	-0.02978	0.05349	1.30579	-1.23807	-2.2519
oracle	-38.05023	-37.83635	-0.02152	-0.02114	-6.7731	-10.05307	-1.80992	-2.84784
bona fide	-43.39013	-43.0466	-0.02558	-0.02517	-4.52543	-6.66035	-9.06877	-12.26569
equal	-2.35459	-2.35355	0.01836	0.01846	-4.84062	-7.13362	-6.0214	-8.08897
gmv	-6.47548	-6.42183	0.04717	0.04649	-1.23807	-2.2519	-6.44415	-8.65646

Table 3: Certainty equivalent, Sharpe ratio, Value-at-Risk and expected shortfall averaged over 1000 randomly chosen portfolios of size 300 and different values of c. The trading period is fixed at 200 days from 01.01.2014 to 10.10.2014.