Dynamic shrinkage estimation of the high-dimensional minimum-variance portfolio

Taras Bodnar[1], Nestor Parolya[2], and Erik Thorsén[1]

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

[2]Department of Applied Mathematics, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands

***Abstract*—In this paper, new results in random matrix theory are derived which allow us to construct a shrinkage estimator of the global minimum variance (GMV) portfolio when the shrinkage target is a random object. More specifically, the shrinkage target is determined as the holding portfolio estimated from previous data. The theoretical findings are applied to develop theory for dynamic estimation of the GMV portfolio, where the new estimator of its weights is shrunk to the holding portfolio at each time of reconstruction. Both cases with and without overlapping samples are considered in the paper. The non-overlapping samples corresponds to the case when different data of the asset returns are used to construct the traditional estimator of the GMV portfolio weights and to determine the target portfolio, while the overlapping case allows intersections between the samples. The theoretical results are derived under weak assumptions imposed on the data-generating process. No specific distribution is assumed for the asset returns except from the assumption of finite $4+\varepsilon$, $\varepsilon > 0$, moments. Also, the population covariance matrix with unbounded spectrum can be considered. The performance of new trading strategies is investigated via an extensive simulation. Finally, the theoretical findings are implemented in an empirical illustration based on the returns on stocks included in the S&P 500 index.**

***Keywords*:** Shrinkage estimator; high-dimensional covariance matrix; random matrix theory; minimum variance portfolio; parameter uncertainty; dynamic decision making

I. Introduction

Global minimum-variance (GMV) portfolio is the one of the mostly used investment strategies by both practitioners and researchers in finance. This portfolio possesses the smallest variance among all optimal portfolios obtained as solutions of Markowitz's mean-variance optimization problem (cf., Markowitz (1952)). It solves the following problem

$$\mathbf{w}^\top \mathbf{\Sigma}\mathbf{w} \to min \quad \text{with} \quad \mathbf{w}^\top \mathbf{1}_p = 1, \tag{I.1}$$

where $\mathbf{w}$ denotes the vector of the portfolio weights which determines the structure of the investor portfolio, the symbol $\mathbf{1}_p$ stands for the p-dimensional vector of ones, and $\mathbf{\Sigma}$ is the covariance matrix of the p-dimensional vector of asset returns $\mathbf{y} = (y_1, ..., y_p)^\top$.

The solution of the optimization problem (I.1) is given by

$$\mathbf{w}_{GMV} = \frac{\mathbf{\Sigma}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}. \tag{I.2}$$

The weights of the GMV portfolio have several nice properties, which simplify its applicability in practice and, thus, make it a popular investment strategy. The weights of the GMV portfolio do not depend on the mean vector of the asset returns, which we will denote by $\boldsymbol{\mu}$ in the following. This is the only mean-variance optimal portfolio whose weights are independent of $\boldsymbol{\mu}$. Moreover, the GMV portfolio has a special location on the set of the mean-variance optimal portfolios, which is a parabola in the mean-variance space and is known as the efficient frontier (cf., Merton (1972)). Its mean and variance determines the location of the vertex of this parabola (see, e.g, Kan and Smith (2008), Bodnar and Schmid (2009)).

The application of (I.2) requires the knowledge of $\mathbf{\Sigma}$ in practice, which is usually not provided. The covariance matrix $\mathbf{\Sigma}$ has to be estimated by using historical data of the asset returns, before the GMV portfolio can be constructed. The quality of the estimator of $\mathbf{\Sigma}$ has a large impact on the stochastic properties of the holding GMV portfolio and it leads to further uncertainty in the investor decision problem, known as the estimation uncertainty. The estimation uncertainty can have a great impact on the constructed portfolio which could be larger than the one induced by the model uncertainty included in the optimization problem (I.1). The effect becomes even stronger, when the portfolio dimension is comparable to the sample size used to estimate $\mathbf{\Sigma}$.

Traditionally, the covariance matrix is estimated by its sample counterpart given by

$$\begin{aligned}\mathbf{S}_n &= \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{y}_i - \bar{\mathbf{y}}_n)(\mathbf{y}_i - \bar{\mathbf{y}}_n)^\top \\ &= \frac{1}{n-1}\mathbf{Y}_n\left(\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top\right)\mathbf{Y}_n^\top \end{aligned} \tag{I.3}$$

with $\bar{\mathbf{y}}_n = \frac{1}{n}\sum_{i=1}^{n}\mathbf{y}_i$, where $\mathbf{y}_1, .., \mathbf{y}_n$ denotes the sample of asset returns and $\mathbf{Y}_n = (\mathbf{y}_1, .., \mathbf{y}_n)$ denotes the data matrix. The symbol $\mathbf{I}_n$ stands for the n-dimensional identity matrix. Then, the sample (also known) as the traditional estimator of $\mathbf{w}_{GMV}$ is obtained as

$$\hat{\mathbf{w}}_S = \frac{\mathbf{S}_n^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_n^{-1}\mathbf{1}_p}. \tag{I.4}$$

The distributional properties of $\hat{\mathbf{w}}_S$ have extensively been studied in statistical and econometric literature. Jobson and Korkie (1980) derive the asymptotic distribution of $\hat{\mathbf{w}}_S$ assuming that the asset returns are independent and normally distributed and the portfolio dimension is considerably smaller than the sample size. Okhrin and Schmid (2006) obtain the exact distribution of the sample estimator of the GMV portfolio weights assuming normality, while Bodnar and Schmid (2008) extend these results to elliptically contoured distribution and develop a statistical test theory on the GMV portfolio weights.

However, when the portfolio dimension is comparable to the sample size, the results derived under the classical asymptotic regime, that is when p is considerably smaller than n, cannot longer be used. Moreover, the effect of dimensionality considerably influences the estimation of the covariance matrix needed to determine the weights of the GMV portfolio. Using the recent results of the random matrix theory several improved estimator for the weights of the GMV portfolio has been suggested when the portfolio dimension is comparable to the sample size, i.e., under the large-dimensional asymptotic regime (see, e.g., Bai and Silverstein (2010)). The properties of high-dimensional optimal portfolio weights are also studied by Fan et al. (2012), Hautsch et al. (2015), Ao et al. (2019), Kan et al. (2019), Bodnar et al. (2022a), Cai et al. (2020), Ding et al. (2021), among others.

Shrinkage approach is one of the mostly used methods to construct an improved estimator for the weights of the GMV portfolio. Shrinkage-type estimators were first proposed by Stein (1956) with the aim to reduce the estimation error present in the sample mean vector computed for a sample from a multivariate normal distribution. Recently, this procedure has also been applied in the construction of the improved estimators of the high-dimensional mean vector (cf, Chételat and Wells (2012), Wang et al. (2014), Bodnar et al. (2019b)), covariance matrix (see, e.g., Ledoit and Wolf (2004), Ledoit and Wolf (2012), Bodnar et al. (2014)), inverse of the covariance matrix (see, e.g., Wang et al. (2015), Bodnar et al. (2016)), as well as of the optimal portfolio weights (see, Golosnoy and Okhrin (2007), Frahm and Memmel (2010), Ledoit and Wolf (2017), Bodnar et al. (2018), Bodnar et al. (2022b)). Interval shrinkage estimators of optimal portfolio weights have recently been derived by Bodnar et al. (2019a), Bodnar et al. (2021a).

The shrinkage estimator for the weights of the GMV portfolio are obtained as a linear combination of the sample estimator $\hat{\mathbf{w}}_S$ and the target portfolio $\mathbf{b}$ with $\mathbf{b}^\top \mathbf{1} = 1$. The estimator is expressed as (see, Bodnar et al. (2018))

$$\hat{\mathbf{w}}_{SH} = \hat{\psi}_n \hat{\mathbf{w}}_S + (1 - \hat{\psi}_n)\mathbf{b} \tag{I.5}$$

where

$$\hat{\psi}_n = \frac{(1 - c_n)\hat{R}_{\mathbf{b}}}{c_n + (1 - c_n)\hat{R}_{\mathbf{b}}}, \tag{I.6}$$

with $\hat{R}_{\mathbf{b}} = (1 - c_n)\, \mathbf{b}^\top \mathbf{S}_n \mathbf{b} \mathbf{1}_p^\top \mathbf{S}_n^{-1} \mathbf{1}_p - 1$ and $c_n = p/n$. Bodnar et al. (2018) show that the shrinkage estimator outperforms the sample estimator of the GMV portfolio weights in terms of minimizing the out-of-sample portfolio variance and the difference becomes drastic when p approaches n. Moreover, the shrinkage estimator of the GMV portfolio weights (I.5) provides a simple and a promising procedure how the one-period portfolio choice problem based on minimizing the portfolio variance can be solved in practice.

Once an optimal portfolio is determined, an investor faces with the problem of optimal portfolio reallocation in the next period of time. One of the important decision to be made by the investor is to decide whether the holding portfolio is optimal or has to be adjusted (see, e.g., Bodnar (2009)), while Golosnoy et al. (2019) consider the exponential smoothing method to predict the weights of the GMV portfolio over some periods of time. In the current paper we contribute to the literature by developing a dynamic GMV portfolio based on the shrinkage approach. At each time point of the portfolio reconstruction the traditional estimator of the GMV portfolio weights is shrunk towards the weights of the holding portfolio, which by construction are the shrinkage estimator of the GMV portfolio from the previous period. The practical advantage of the new dynamic trading strategy is two-fold: (i) First, it diminishes the transaction costs required for the reconstruction of the holding portfolio; (ii) Second, it reduces the out-of-sample variance of the constructed GMV portfolio by applying the shrinkage approach in the estimation of the portfolio weights.

From the perspectives of statistical theory, we develop new results that allow us to use the shrinkage estimators with a random target. These estimators are obtained under weak conditions imposed on the data-generating process. In particular, only the existence of the fourth moments is needed without explicit specification of the probability distribution assumed for the asset returns. Moreover, no assumption about the spectrum of the population covariance matrix is imposed in the paper. The eigenvalues of the covariance matrix can be as large as in the factor models (see, e.g., Fan et al. (2008), Fan et al. (2013), Ding et al. (2021)). We only require that the ratio of the variances of the target portfolio and the GMV portfolio is bounded. Moreover, the derived theoretical results allow the application of the overlapping samples in the determination of the target portfolio and in the construction of the traditional portfolio used in the specification of the shrinkage GMV portfolio. Finally, the obtained findings are implemented in the R-package *DOSPortfolio* (see, Bodnar et al. (2021b)).

It is remarkable that the statistical methods developed in the paper can be linked to the approaches applied in statistical signal processing by noting that the GMV portfolio is related to the Capon or minimum variance spatial filter in signal processing literature (cf, Verdú (1998) and Van Trees (2002)). Rubio et al. (2012) and Yang et al. (2018) investigate the estimation risk of the high-dimensional minimum variance beamformer, while Li et al. (2004) study its constrained versions. The applications of random matrix theory to signal processing and portfolio optimization are presented in Feng and Palomar (2016), Bodnar et al. (2019a), Bodnar et al. (2021a).

The rest of the paper is organized as follows. In Section II, the main theoretical findings of the paper are provided. The dynamic shrinkage estimator for the weights of the GMV portfolio is derived in the case of non-overlapping samples in Section II-A, while Section II-B presents the results in the overlapping case. The performance of the new trading strategies is investigated in Section III via an extensive simulation study, where the approaches are also compared to the existing ones. In Section IV, the new approaches to estimate the GMV portfolio are implemented to the real data consisting of the returns on stocks included in S&P 500 index. Concluding remarks are given in Section V, while the technical proofs are moved to the supplementary material.

II. Dynamic estimation of GMV portfolio

Throughout the paper we assume that the GMV portfolio is constructed at time point t_1 by using the sample of size n_1, and then the investor updates the constructed GMV portfolio as new information arrives on the capital market. The information set is presented as a sequence of asset returns taken between time point t_{i-1} and t_i for $i = 2, ..., T$. Between each pairs (t_{i-1}, t_i) it is assumed that n_i vectors of asset returns are available which are collected into the data matrix $\mathbf{Y}_{n_i}$ that is assumed to possess the following stochastic representation:

$$\mathbf{Y}_{n_i} = \boldsymbol{\mu}\mathbf{1}_{n_i}^\top + \boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{X}_{n_i}, \tag{II.1}$$

where $\mathbf{X}_{n_i}$ is a $p \times n_i$ matrix which consists of independent and identically distributed (i.i.d.) real random variables with zero mean and unit variance. Also, we assume that the entries of $\mathbf{X}_{n_i}$, $i = 1, ..., T$, possess the $4 + \varepsilon$, $\epsilon > 0$, moments, while no specific distributional assumption is imposed on the element of $\mathbf{X}_{n_i}$. To this end, it is assumed that $\mathbf{Y}_{n_i}$, $i = 1, ..., T$, are independent random matrices.

We consider an investor who opts on the shrinkage estimation of the GMV portfolio weights in each period of time t_i. Namely, after constructing the shrinkage estimator of the GMV portfolio as defined in (I.5) at time point t_1, the investor updates the GMV portfolio weights by shrinking their sample estimator computed at each time point t_i to the holding GMV portfolio determined at time point t_{i-1}. Two estimation strategies are developed in this section, which are based on non-overlapping and overlapping samples, respectively. The first procedure can be related to the rolling window estimation but with probably different sample sizes. The main advantage here is that smaller sample sizes are used in the construction of the sample weights of the GMV portfolio and, thus, the extreme observation observed in the asset returns will sooner be detected. Such a strategy might be recommendable during the turbulent period on the capital market, since it allows a faster adjustment of the holding portfolio. In contrary, when the stable period on the capital market is present, then the investor would prefer to use all available information, which leas to the extended window estimation strategy. In this case the part of data used in the construction of the GMV portfolio has already been used to determine the currently holding portfolio to which the new estimator is shrunk and, consequently, we have the case with overlapping samples. Both situations require completely different techniques from random matrix theory to be developed in order to derive the stochastic properties of the estimation procedures, which are developed in the consequent two subsections.

A. Dynamic GMV portfolio with non-overlapping samples

Under the non-overlapping scenario, the investor uses the sample of asset returns collected in $\mathbf{Y}_{n_i}$ to construct the sample estimator of the GMV portfolio at each time point t_i expressed as

$$\hat{\mathbf{w}}_{S;n_i} = \frac{\mathbf{S}_{n_i}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_{n_i}^{-1}\mathbf{1}_p} \tag{II.2}$$

where

$$\mathbf{S}_{n_i} = \frac{1}{n_i - 1}\mathbf{Y}_{n_i}\left(\mathbf{I}_{n_i} - \frac{1}{n_i}\mathbf{1}_{n_i}\mathbf{1}_{n_i}^\top\right)\mathbf{Y}_{n_i}^\top. \tag{II.3}$$

The shrinkage estimator of the GMV portfolio is then obtained at time point t_i by shrinking (II.2) to the weights of the holding portfolio, i.e., to the shrinkage estimator of the GMV portfolio $\hat{\mathbf{w}}_{SH;n_{i-1}}$ constructed in the previous period, by minimizing the loss function determined as the out-of-sample variance with respect to the shrinkage intensity ψ_{n_i} in the following way:

$$\min_{\psi_i} L_i(\psi_i) = \min_{\psi_i} \hat{\mathbf{w}}_{SH;i}^\top \boldsymbol{\Sigma}\hat{\mathbf{w}}_{SH;i} \tag{II.4}$$

with

$$\hat{\mathbf{w}}_{SH;n_i} = \psi_i\hat{\mathbf{w}}_{S;n_i} + (1-\psi_i)\hat{\mathbf{w}}_{SH;n_{i-1}}, \tag{II.5}$$

where $\hat{\mathbf{w}}_{SH;n_0} = \mathbf{b}$ is the shrinkage target used for the construction of the shrinkage estimator for the GMV portfolio weights at time point t_1.

Rewriting (II.4) we get

$$\begin{aligned} L_i(\psi_i) =& \psi_i^2\hat{\mathbf{w}}_{S;n_i}^\top\boldsymbol{\Sigma}\hat{\mathbf{w}}_{S;n_i} \\ &+ 2\psi_i(1-\psi_i)\hat{\mathbf{w}}_{S;n_i}^\top\boldsymbol{\Sigma}\hat{\mathbf{w}}_{SH;n_{i-1}} \\ &+ (1-\psi_i)^2\hat{\mathbf{w}}_{SH;n_{i-1}}^\top\boldsymbol{\Sigma}\hat{\mathbf{w}}_{SH;n_{i-1}}, \end{aligned} \tag{II.6}$$

which is minimized at

$$\psi_{n_i}^* = \frac{\hat{\mathbf{w}}_{SH;n_{i-1}}^\top\boldsymbol{\Sigma}\left(\hat{\mathbf{w}}_{SH;n_{i-1}} - \hat{\mathbf{w}}_{S;n_i}\right)}{\left(\hat{\mathbf{w}}_{SH;n_{i-1}} - \hat{\mathbf{w}}_{S;n_i}\right)^\top\boldsymbol{\Sigma}\left(\hat{\mathbf{w}}_{SH;n_{i-1}} - \hat{\mathbf{w}}_{S;n_i}\right)}. \tag{II.7}$$

In Theorem II.1 we derive the asymptotic equivalent to $\psi_{n_i}^*$ which can be used in the construction of the shrinkage estimator at time point t_i.

Theorem II.1. *Let $\mathbf{Y}_{n_i}$ possess the stochastic representation as in* (II.1) *and let $\mathbf{b}$ be the deterministic shrinkage target for $i = 1$. Assume that the relative loss of portfolio $\mathbf{b}$ given by*

$$r_0 = \frac{V_{\mathbf{b}}}{V_{GMV}} - 1 = \mathbf{1}_p^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}_p\mathbf{b}^\top\boldsymbol{\Sigma}\mathbf{b} - 1 \tag{II.8}$$

is uniformly bounded in p, where

$$V_{\mathbf{b}} = \mathbf{b}^\top\boldsymbol{\Sigma}\mathbf{b} \text{ and } V_{GMV} = \mathbf{w}_{GMV}^\top\boldsymbol{\Sigma}\mathbf{w}_{GMV} = \frac{1}{\mathbf{1}_p^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}_p} \tag{II.9}$$

are the variances of the target portfolio $\mathbf{b}$ and of the population GMV portfolio, respectively. Then it holds that

$$\left|\psi_{n_i}^* - \psi_i^*\right| \xrightarrow{a.s.} 0 \quad \text{with} \quad \psi_i^* = \frac{(1-c_i)r_{i-1}}{(1-c_i)r_{i-1} + c_i} \tag{II.10}$$

for $p/n_i \to c_i \in (0,1)$ as $n \to \infty$ where r_i is the asymptotic equivalent of the relative loss $r_{\hat{\mathbf{w}}_{SH;n_i}} = \mathbf{1}_p^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}_p\hat{\mathbf{w}}_{SH;n_i}^\top\boldsymbol{\Sigma}\hat{\mathbf{w}}_{SH;n_i} - 1$ of the portfolio with weights $\hat{\mathbf{w}}_{SH;n_i}$ given by

$$r_i = (\psi_i^*)^2\frac{c_i}{1-c_i} + (1-\psi_i^*)^2 r_{i-1} \tag{II.11}$$

for $i = 1, ..., T$.

The proof of Theorem II.1 is given in the supplementary material. Its results provide a simple recursive algorithm how the shrinkage intensities have to be computed in practice. Independently of the number of portfolio reallocations, T, the only unknown quantity in the algorithm is the relative loss of of the target portfolio $\mathbf{b}$ used in the construction of the shrinkage estimator at time $i = 1$. Using the sample $\mathbf{Y}_{n_1}$ its consistent estimator is given by

$$\hat{r}_0 = \left(1 - \frac{p}{n_1}\right)\mathbf{1}_p^\top\mathbf{S}_{n_1}^{-1}\mathbf{1}_p\mathbf{b}^\top\mathbf{S}_{n_1}\mathbf{b} - 1. \tag{II.12}$$

Then, the resulting (bona fide) shrinkage estimator of the GMV portfolio at time i is given by

$$\hat{\mathbf{w}}_{BF;n_i} = \hat{\psi}_i^*\hat{\mathbf{w}}_{S;n_i} + (1-\hat{\psi}_i^*)\hat{\mathbf{w}}_{BF;n_{i-1}} \tag{II.13}$$

where $\hat{\psi}_i^* = \frac{(n_i-p)r_{i-1}}{(n_i-p)r_{i-1}+p}$ and $\hat{r}_i$ is computed recursively by

$$\hat{r}_i = (\hat{\psi}_i^*)^2\frac{p}{n_i - p} + (1-\hat{\psi}_i^*)^2\hat{r}_{i-1} \tag{II.14}$$

with $\hat{r}_0$ as in (II.12) and $\hat{\mathbf{w}}_{BF;n_0} = \mathbf{b}$.

We conclude this section with several important remarks:

Remark II.2. The deterministic target portfolio $\mathbf{b}$ can also be replaced by the sample GMV portfolio computed by using data available before the sample $\mathbf{Y}_{n_1}$ is taken. If we denote these data by $\mathbf{Y}_{n_0}$, then the target weights $\mathbf{b}$ are replaced by

$$\hat{\mathbf{w}}_{S;n_0} = \frac{\mathbf{S}_{n_0}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top\mathbf{S}_{n_0}^{-1}\mathbf{1}_p} \tag{II.15}$$

In this case the relative loss r_0 does not longer depend on the population covariance matrix $\boldsymbol{\Sigma}$ and following the proof of Theorem II.1 it is given by

$$\tilde{r}_0 = \frac{c_0}{1-c_0} \approx \frac{p}{n_0 - p}.$$

As a result, the (bona fide) shrinkage estimator of the GMV portfolio weights is obtained as in (II.13) and (II.14) with $\hat{r}_0$ replaced by $\tilde{r}_0$ and $\hat{\mathbf{w}}_{BF;n_0} = \hat{\mathbf{w}}_{S;n_0}$. In a similar way other random targets can be employed into our model, e.g., nonlinear shrinkage Ledoit and Wolf (2012), but then the asymptotics and estimation of r_0 becomes highly nontrivial and one needs to handle every of those targets separately. Because of the large number of possible target portfolios $\mathbf{b}$ we leave this interesting topic for the future research. For the sake of brevity concentrate ourselves on the naive equally weighted target $\mathbf{b} = \mathbf{1}_p/p$ in our simulation and empirical studies.

Remark II.3. The results of Theorem II.1 are derived under very week conditions which require the existent of $4+\varepsilon$, $\varepsilon > 0$, moments only. No structural assumption on $\mathbf{\Sigma}$ neither on $\mathbf{b}$ are imposed.

Remark II.4. Other consistent estimators for r_0 can be constructed. For instance, we can update our estimator at each time point t_i as soon as new data of the asset returns become available. Let $N_i = \sum_{j=1}^{i} n_j$ be the total number of asset return vectors available at time point t_i and let $\mathbf{Y}_{N_i}$ be the $p \times N_i$ matrix of the asset returns up to time N_i that is $\mathbf{Y}_{N_i} = (\mathbf{Y}_{n_1}\, \mathbf{Y}_{n_2}\, ...\, \mathbf{Y}_{n_i})$. Then, at time point i, a consistent estimator for r_0 is obtained by

$$\hat{r}_{0;i} = \left(1 - \frac{p}{N_i}\right) \mathbf{1}_p^\top \mathbf{S}_{N_i}^{-1} \mathbf{1}_p \mathbf{b}^\top \mathbf{S}_{N_i} \mathbf{b} - 1. \tag{II.16}$$

where $\mathbf{S}_{N_i}$ is the sample covariance matrix based on the data matrix $\mathbf{Y}_{N_i}$ as given in (I.3) with $n = N_i$. Then, the (bona fide) shrinkage estimator of the GMV portfolio weights is computed following (II.13) and (II.14) with $\hat{r}_0$ replaced by $\hat{r}_{0;i}$. Since larger dataset is used to estimate r_0, we expect that this approach will perform better as the one suggested in (II.12) - (II.14). On the other side, the new method is more time demanding, since the recursion in (II.14) has to be started from the beginning at each time t_i.

B. Dynamic GMV portfolio with overlapping samples

In this section, we present the shrinkage estimator for the GMV portfolio which is constructed based on the overlapping samples. In Remark II.4 it is suggested to use all available data $\mathbf{Y}_{n_1}, \mathbf{Y}_{n_2}, ..., \mathbf{Y}_{n_i}$ up to time point t_i to determine a consistent estimator for the relative loss r_0 of portfolio $\mathbf{b}$. Here, we use the similar idea in the construction of the sample estimator of the GMV portfolio weights at time t_i. Such an approach possesses an advantage that we only require $n_1 > p$, while the other sample sizes $n_2, ..., n_T$ can also be smaller than the portfolio dimension p.

Using the notations N_i, $\mathbf{Y}_{N_i}$, and $\mathbf{S}_{N_i}$ introduced in Remark II.4, we define

$$\hat{\mathbf{w}}_{S;N_i} = \frac{\mathbf{S}_{N_i}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_{N_i}^{-1}\mathbf{1}_p}. \tag{II.17}$$

as the sample estimator of the GMV portfolio weights based on data of the asset returns included in $\mathbf{Y}_{N_i}$. Substituting $\hat{\mathbf{w}}_{S;N_i}$ instead of $\hat{\mathbf{w}}_{S;n_i}$ in (II.5), the loss function $L_i(\psi_i)$ in (II.4) is maximized at

$$\Psi_{N_i}^* = \frac{\hat{\mathbf{w}}_{SH;N_{i-1}}^\top \mathbf{\Sigma} \left(\hat{\mathbf{w}}_{SH;N_{i-1}} - \hat{\mathbf{w}}_{S;N_i}\right)}{\left(\hat{\mathbf{w}}_{SH;N_{i-1}} - \hat{\mathbf{w}}_{S;N_i}\right)^\top \mathbf{\Sigma} \left(\hat{\mathbf{w}}_{SH;N_{i-1}} - \hat{\mathbf{w}}_{S;N_i}\right)}. \tag{II.18}$$

In Theorem II.5 we derive an iterative procedure for computing the deterministic equivalents to $\Psi_{N_i}^*$ for $i = 1, ..., T$. The proof of Theorem II.5 is given in the supplementary material.

Theorem II.5. *Let $\mathbf{Y}_{n_i}$ possess the stochastic representation as in* (II.1) *and let $\mathbf{b}$ be the deterministic shrinkage target for $i = 1$. Assume that the relative loss of portfolio $\mathbf{b}$ given by $R_0 = \mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p \mathbf{b}^\top \mathbf{\Sigma}\mathbf{b} - 1$ is uniformly bounded in p. Then it holds that*

$$\left|\Psi_{N_i}^* - \Psi_i^*\right| \overset{a.s.}{\to} 0$$

for $p/N_j \to C_j \in (0,1)$ as $N_j \to \infty$, $j = 1, ..., i$ and $i = 1, ..., T$ where

$$\Psi_i^* = \frac{(R_{i-1}+1) - K_i}{(R_{i-1}+1) + (1-C_i)^{-1} - 2K_i}, \tag{II.19}$$

$$K_i = \beta_{i-1;0}^* + \sum_{j=1}^{i-1} \beta_{i-1;j}^* D_{j,i}, \tag{II.20}$$

$$R_i = (\Psi_i^*)^2 \frac{C_i}{1-C_i} + (1-\Psi_i^*)^2 R_{i-1} + 2\Psi_i^*(1-\Psi_i^*)(K_i - 1), \tag{II.21}$$

(II.22)

with $\beta_{0;0}^* = 1$, $\beta_{i-1;i-1}^* = \Psi_{i-1}^*$ and

$$\beta_{i-1;k}^* = (1-\Psi_{i-1}^*)\beta_{i-2;k}^*, \tag{II.23}$$

for $k = 0, ...i-2$. Finally, $D_{j,i}$ is given by

$$D_{j,i} = 1 - \frac{2(1-C_j)}{(1-C_j) + (1-C_i)\frac{C_j}{C_i} + \sqrt{\left(1-\frac{C_j}{C_i}\right)^2 + 4(1-C_i)\frac{C_j}{C_i}}}. \tag{II.24}$$

Similarly, to the case with non-overlapping samples, the recursive procedure derived in Theorem II.5 depends only on a univariate unobservable quantity R_0, which is the relative loss of the target portfolio $\mathbf{b}$ at time point t_1. Both approaches suggested in Section II-A can be used to construct a consistent estimator for R_0, and hence to obtain a (bona fide) estimator of the GMV portfolio weights. These procedures are the following:

- We estimate R_0 by

$$\hat{R}_0 = \hat{r}_0 = \left(1 - \frac{p}{N_1}\right) \mathbf{1}_p^\top \mathbf{S}_{N_1}^{-1} \mathbf{1}_p \mathbf{b}^\top \mathbf{S}_{N_1} \mathbf{b} - 1 \tag{II.25}$$

  as in (II.12). In this case the estimator for R_0 is constructed by using the first sample $\mathbf{Y}_{N_1}$ only and the recursive procedure of Theorem II.5 is then used leading to the (bona fide) optimal shrinkage estimators for the weights at each time point t_i, $i \in 1, ..., T$ expressed as

$$\hat{\mathbf{w}}_{BF;N_i} = \hat{\Psi}_i^* \hat{\mathbf{w}}_{S;N_i} + (1-\hat{\Psi}_i^*)\hat{\mathbf{w}}_{BF;N_{i-1}} \tag{II.26}$$

  where $\hat{\Psi}_i^*$ is computed recursively as in Theorem II.5 with R_0 replaced by $\hat{R}_0$ and using the empirical counterpart for C_i given by $C_{N_i} = p/N_i$.
- At each time point i, we use all available information to estimate R_0, i.e.,

$$\hat{R}_{0;i} = \hat{r}_{0;i} = \left(1 - \frac{p}{N_i}\right) \mathbf{1}_p^\top \mathbf{S}_{N_i}^{-1} \mathbf{1}_p \mathbf{b}^\top \mathbf{S}_{N_i} \mathbf{b} - 1 \tag{II.27}$$

  and recompute the recursion of Theorem II.5 at each time point t_i. Since a larger dataset is used to estimate R_0, better results are expected although the computations becomes more time demanding in the second case.

To this end, we note that the deterministic target portfolio $\mathbf{b}$ can be replaced by the sample GMV portfolio computed

by using data $\mathbf{Y}_{n_0}$ available before the first sample $\mathbf{Y}_{N_1}$ as in (II.15) of Remark II.2. In this case we get

$$\widetilde{R}_0 = \frac{p}{n_0 - p},$$

which is used in the iterative computation of Theorem II.5 instead of R_0. Since no unknown quantities are present in the definition of $\widetilde{R}_0$, the iterative procedure of Theorem II.5 becomes deterministic.

III. Finite-sample performance

A. Benchmark strategies and the setup of the simulation study

The suggested dynamic estimation strategies are compared to several benchmark strategies via en extensive simulation study in this section, while the results of the empirical illustration are provided in Section IV. The performance of the following eight dynamic trading strategies will be established:

Strategy 1:
Bona fide shrinkage estimator of the GMV portfolio (II.13) with (II.14) following Theorem II.1 where r_0 is estimated from the first sample as in (II.12);

Strategy 2:
Bona fide shrinkage estimator of the GMV portfolio (II.26) where $\hat{\Psi}_i^*$ is computed recursively as in Theorem II.5 and R_0 is estimated from the first sample as in (II.25);

Strategy 3:
Bona fide shrinkage estimator of the GMV portfolio (II.13) with (II.14) following Theorem II.1 where r_0 is recomputed as in (II.16) when a new sample becomes available;

Strategy 4:
Bona fide shrinkage estimator of the GMV portfolio (II.26) where $\hat{\Psi}_i^*$ is computed recursively as in Theorem II.5 and R_0 is recomputed as in (II.27) when a new sample becomes available;

Strategy 5:
Sample estimator of the GMV portfolio computed at each time t_i, $i = 1, 2, ..., T$, i.e., $\psi_i = 1$ in (II.5) for $i = 1, 2, ..., T$;

Strategy 6:
Target portfolio $\mathbf{b}$ used at each time t_i, $i = 1, 2, ..., T$, i.e., $\psi_i = 0$ in (II.5) for $i = 1, 2, ..., T$;

Strategy 7:
One-period shrinkage estimator of the GMV portfolio (I.5) with (I.6) reconstructed at each time t_i, $i = 1, 2, ..., T$ with equally-weighted portfolio as the target portfolio.

Strategy 8:
Ledoit-Wolf nonlinear shrinkage estimator of the GMV portfolio (see, Ledoit and Wolf (2017)) computed at each time t_i, $i = 1, 2, ..., T$;

The first four strategies are based on the theoretical results derived in Sections II-A and II-B where two different methods for constructing bona fide shrinkage estimators of the GMV portfolio weights are explored following the discussion after Theorems II.1 and II.5, respectively. **Strategies 5 to 7** are the benchmark strategies which are based on the traditional estimator of the GMV portfolio, on the target portfolio, and on the one-period shrinkage estimator. The last **Strategy 8** is the recent state-of-the-art method of Ledoit and Wolf (2017), which efficiently applies the nonlinear shrinkage estimator of the covariance matrix on the GMV portfolio weights.

Since the GMV portfolio is the solution of the portfolio optimization problem with the aim to minimize the portfolio variance, the relative loss in the out-of-sample variance is used as a performance measure in this comparison study which for the portfolio with the estimated weights $\hat{\mathbf{w}}$ is expressed as:

$$\begin{aligned}\text{Relative loss}\,(\mathbf{w}) =& \frac{\hat{\mathbf{w}}^\top \boldsymbol{\Sigma} \hat{\mathbf{w}} - V_{GMV}}{V_{GMV}} \\ =& \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1} \hat{\mathbf{w}}^\top \boldsymbol{\Sigma} \hat{\mathbf{w}} - 1, \end{aligned} \tag{III.1}$$

where we use the formula for the global minimum variance V_{GMV} given in (II.9).

In the simulation study, we will look at two investment horizons $T = 10$ and $T = 20$. For each segment we let $n_i = 250$, which would correspond to an investor who rebalances the holding portfolio on a yearly basis. The parameters of the listed below models are simulated according to $\boldsymbol{\mu} = (\mu_1, ..., \mu_p)^\top$ with $\mu_i \sim U(-0.2, 0.2)$ and the covariance matrix $\boldsymbol{\Sigma}$ is configured such that 20% of the eigenvalues are equal to 0.2, 40% equal to one and 40% equal to 4, whereas the eigenvectors are generated from the Haar distribution. Following this simulation setup $\boldsymbol{\Sigma}$ will have the same spectral distribution for all considered values of the concentration ratio c_n.

Four different stochastic models for the data-generating process will be considered, which are listed below:

Scenario 1: t-distribution
The elements of $\mathbf{x}_t$ are drawn independently from t-distribution with 5 degrees of freedom, i.e., $x_{tj} \sim t(5)$ for $j = 1, ..., p$, while $\mathbf{y}_t$ is constructed according to (II.1).

Scenario 2: CAPM
The vector of asset returns $\mathbf{y}_t$ is generated according to the CAPM (Capital Asset Pricing Model), i.e.,

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\beta} z_t + \boldsymbol{\Sigma}^{1/2} \mathbf{x}_t,$$

with independently distributed $z_t \sim N(0, 1)$ and $\mathbf{x}_t \sim N_p(\mathbf{0}, \mathbf{I})$. The elements of vector $\boldsymbol{\beta}$ are drawn from the uniform distribution, that is $\beta_i \sim U(-1, 1)$ for $i = 1, ..., p$.

Scenario 3: CCC-GARCH model of Bollerslev (1990)
The asset returns are simulated according to

$$\mathbf{y}_t | \boldsymbol{\Sigma}_t \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}_t)$$

where the conditional covariance matrix is specified by

$$\boldsymbol{\Sigma}_t = \mathbf{D}_t^{1/2} \mathbf{C} \mathbf{D}_t^{1/2} \;\text{ with }\; \mathbf{D}_t = \text{diag}(h_{1,t}, h_{2,t}, ..., h_{p,t}),$$

where

$$h_{j,t} = \alpha_{j,0} + \alpha_{j,1} (\mathbf{y}_{j,t-1} - \boldsymbol{\mu}_j)^2 + \beta_{j,1} h_{j,t-1},$$

for $j = 1, 2, ..., p$, and $t = 1, 2, ..., n_i,\ i = 1, ..., T$. The coefficients of the CCC model are sampled according to $\alpha_{j,1} \sim U(0, 0.1)$ and $\beta_{j,1} \sim U(0.6, 0.7)$ which implies that the stationarity conditions, $\alpha_{j,1} + \beta_{j,1} < 1$, are always fulfilled. The intercept $\alpha_{j,0}$ is thereafter chosen such that the unconditional covariance matrix is equal to $\boldsymbol{\Sigma}$.

Scenario 4: VARMA model
The vector of asset returns $\mathbf{y}_t$ is simulated according to a

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\Gamma}(\mathbf{y}_{t-1} - \boldsymbol{\mu}) + \boldsymbol{\Sigma}^{1/2} \mathbf{x}_t, \;\text{ with }\; \mathbf{x}_t \sim N_p(\mathbf{0}, \mathbf{I})$$

for $t = 1, ..., n_i,\ i = 1, ..., T$, where $\boldsymbol{\Gamma} = \text{diag}(\gamma_1, \gamma_2, ..., \gamma_p)$ where $\gamma_i \sim U(-0.9, 0.9)$ for $i = 1, ..., p$.

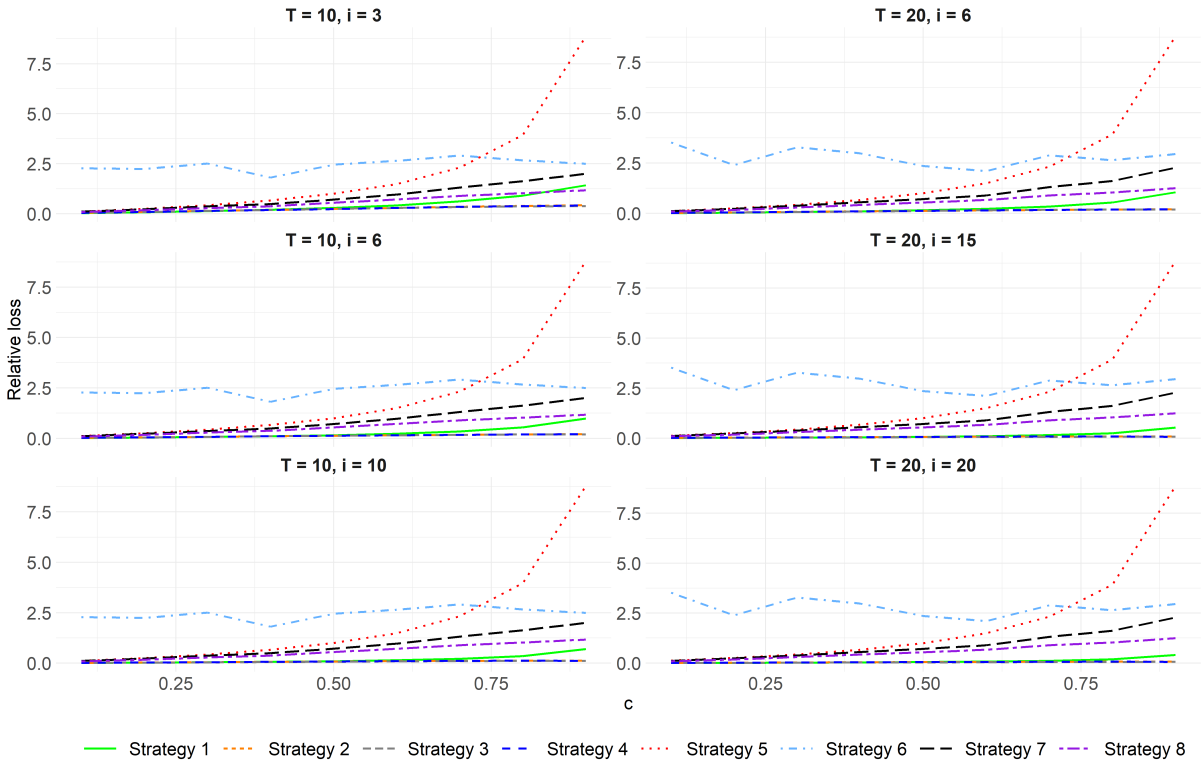

Fig. 1: Relative losses for the different time steps i and investment horizons T. Data were simulated following **Scenario 1** for different values of c.

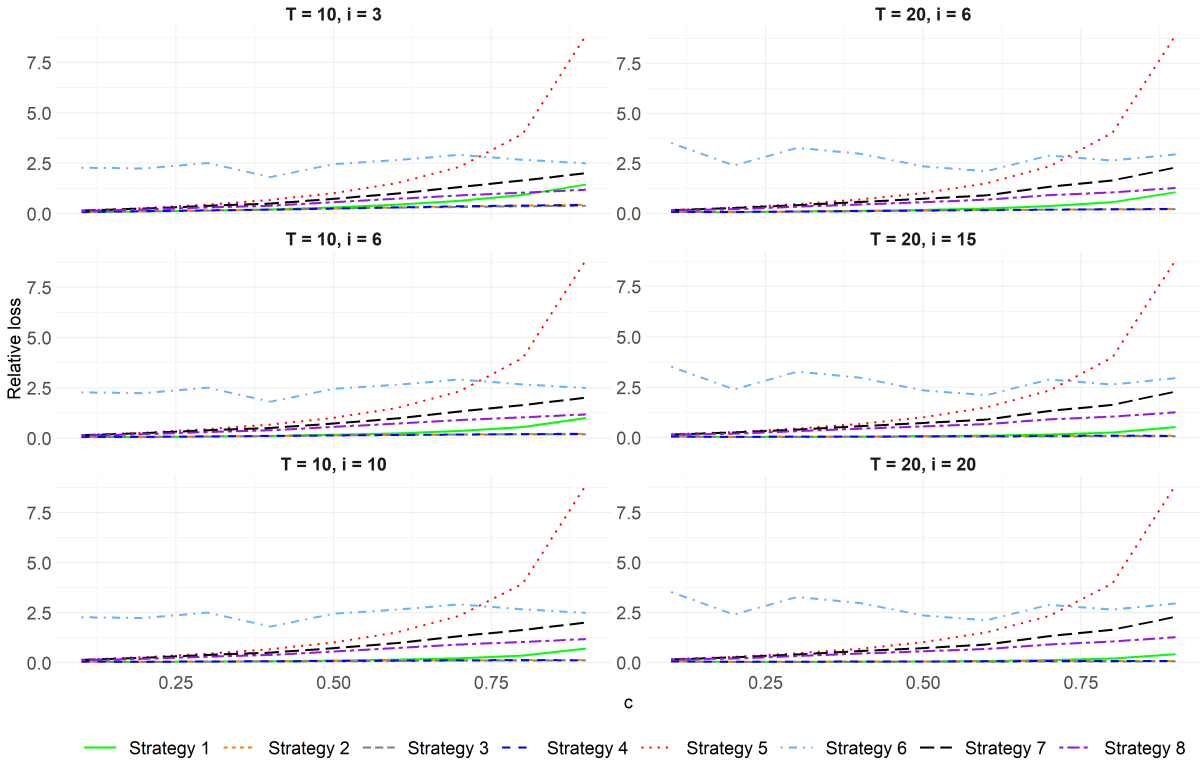

Fig. 2: Relative losses for the different time steps i and investment horizons T. Data were simulated following **Scenario 2** for different values of c.

Scenario 1 and **Scenario 2** fulfill the conditions imposed on the data-generating model in Section II. The application

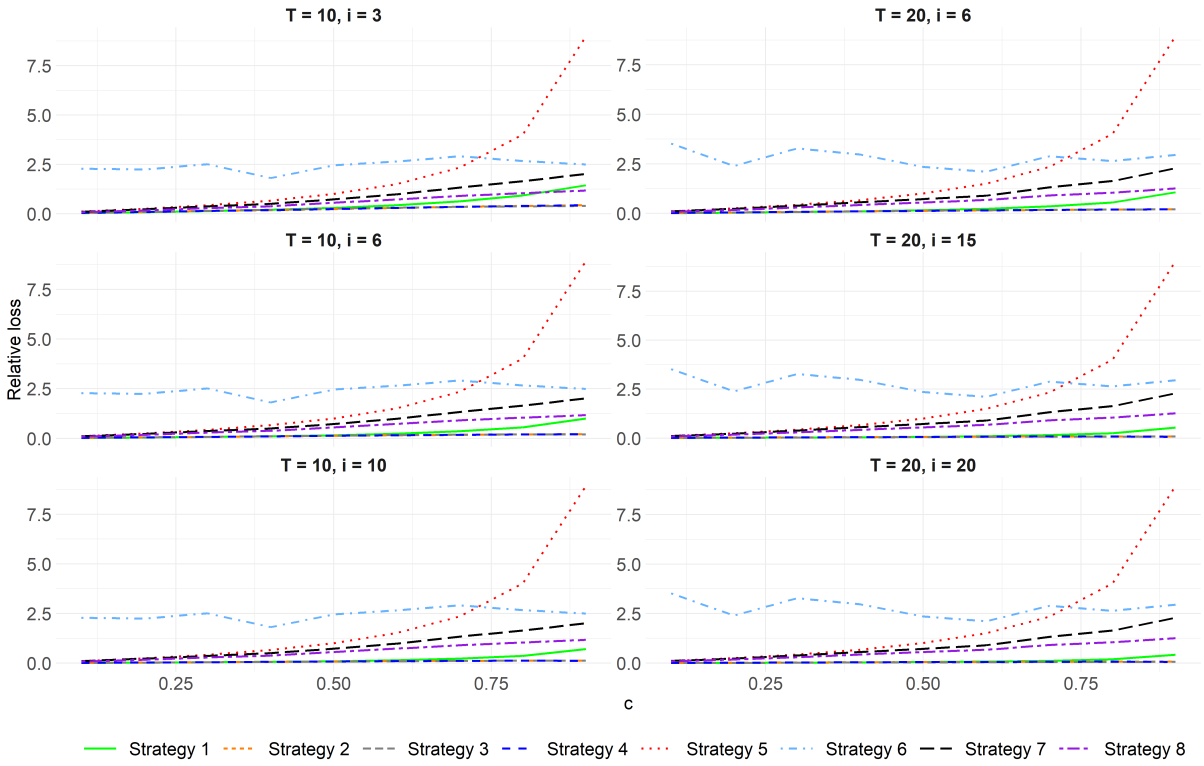

Fig. 3: Relative losses for the different time steps i and investment horizons T. Data were simulated following **Scenario 3** for different values of c.

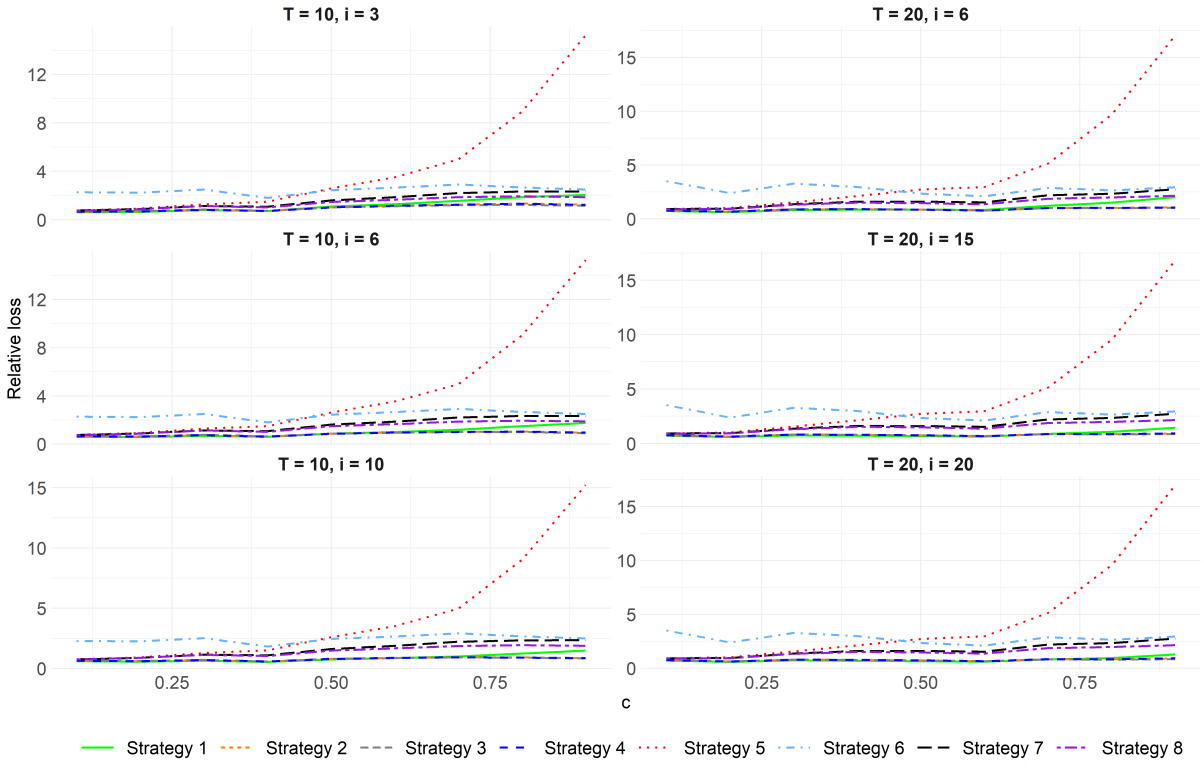

Fig. 4: Relative losses for the different time steps i and investment horizons T. Data were simulated following **Scenario 4** for different values of c.

of both scenarios result in samples that consist of independent random vectors with finite $4 + \varepsilon$, $\varepsilon > 0$, moments.

Furthermore, the covariance matrix possesses finite eigenvalues in **Scenario 1**, while it has an unbounded spectrum in **Scenario 2** (cf., Fan et al. (2013)). On the other side, the samples obtained following **Scenario 3** and **Scenario 4** consist of dependent observations. In **Scenario 3** the random vector are uncorrelated although a non-linear dependence is present in the time series structure of the model, while the elements of the samples obtained from **Scenario 4** are strongly linearly dependent.

For each segment of the time partition we generate a new sample of $n = 250$ observations, which is applied in the computation of $\hat{\boldsymbol{\mu}}$, $\hat{\mathbf{S}}_{n_i}$, $\hat{\mathbf{S}}_{N_i}$, $\hat{\mathbf{w}}_{S;n_i}$, and $\hat{\mathbf{w}}_{SH;n_i}$. As a target portfolio we use the equally weighted portfolio with the weights $\mathbf{b} = \mathbf{1}_p/p$. The results of the simulation study are based on 5000 independent runs from which the average relative loss is computed for each scenario, strategy and several values of the concentration ratio c.

B. Performance of the trading strategies

Figures 1 to 4 present the results of the simulation study for $i = 3, i = 6$, and $i = 10$ when $T = 10$ and for $i = 6, i = 13$ and $i = 20$ when $T = 20$. Interestingly, the computed average loss show a similar behaviour independently of the data-generating model used to draw the samples. This observation also holds in the case of **Scenario 3** and **Scenario 4**, which by construction do not fulfill the assumptions imposed on the data-generating model in the derivation of the theoretical results. As such, one can conclude that the presence of non-linear dependence structure between the observation vectors or even strong linear dependence has only minor impact on the validity of the results derived in Theorem II.1 and Theorem II.5.

The best performance is obtained for **Strategy 2**, **Strategy 3**, and **Strategy 4**, which are followed by **Strategy 1**. The differences between the computed values for **Strategies 2 to 4** are very small and are present at the fourth decimal. All dynamic estimation strategies considerably outperform the four considered benchmark strategies, independently of the scenario used to generate samples. On the third place we rank the nonlinear shrinkage estimator, i.e., **Strategy 8**, while the single-period shrinkage estimator is ranked on the place four. Finally, we note that for **Scenarios 1-4** the traditional estimator performs better than the portfolio strategy based on the target portfolio, when the concentration ratio c is smaller than 0.75, while it produces extremely large values of relative losses, when c approaches one. This observation becomes even more prominent in case of the **Scenario 4**, where a strong autocorrelation was employed. Here already for $c = 0.5$ the target portfolio starts outperforming the traditional estimator. To this end, we conclude that the dynamic re-estimation of the relative loss of the target portfolio $\mathbf{b}$ shows a significant improvement when non-overlapping samples are used and the concentration ratio c is relative large. In contrast, the application of the dynamic re-estimation of the relative loss in the case of overlapping samples leads to the considerably large computation time without large improvements. Finally, the increase of the trading horizon T has only a minor impact on the plots presented in Figures 1 to 4. The larger value of T slightly reduces the computed average relative losses in the case of **Strategy 1**, while they become a slightly large for the single-period shrinkage approach.

IV. Application to stocks from S&P 500

In this section we will apply the suggested new approaches and the benchmark strategies presented in Section III on daily market data. The computation is performed by using the R-package *DOSPortfolio* (see, Bodnar et al. (2021b)).

A. Data description

We will use daily returns on 348 stocks included in the S&P500 index from March 2011 up until March 2020. The stocks were chosen by the availability of their price data during the trading period. Two portfolios of size $p = 100$ (high-dimensional case) and $p = 50$ (low-dimensional case) are considered, where 50 stocks are chosen randomly from 348 stocks included in the empirical study for the first portfolio, while the second one contains the first 50 stocks in the alphabetic order of the former portfolio. We set $c = 0.25$ or $c = 0.5$ and will therefore use $n_i = 200$ trading days for each year i.

Figure 5 presents the descriptive statistics computed for the univariate series of the randomly chosen 100 stocks. Namely, the first four (centralized) sample moments of the univariate time series of asset returns are depicted in the figure, which are computed for the whole period from March 2011 up until March 2021. The returns are on average positive over this period, while the sample skewness is on average negative with the sample distribution to be skewed to the left and with a few assets having very large (negative) values. The largest negative skewness corresponds to the Mondelez International, Inc. (ticker MDLZ) which is also among those assets with highest kurtosis. Finally, we note that the computed kurtosis are relatively large for the considered stocks, showing that the assumption of normality might not be fulfilled for the considered data.

B. Results of the empirical illustration

A consequence of the exponential weighting schemes to which the shrinkage estimators belong to, is that the portfolio structure changes by smaller increments. As a result, we expect that the portfolio turnover of the dynamic (estimated) GMV portfolios based on the introduced shrinkage approaches to be smaller in comparison to the unconstrained strategy ($\psi_i = 1$), but to be larger in comparison to the static portfolio choice ($\psi_i = 0$). For each strategy k introduced in Section III-A, let $\mathbf{w}^{(k)}$ denote the vector of the weights induced by the kth strategy and let $w_{i,j}^{(k)}$ stand for the weight for the j-th asset after the i-th portfolio rebalancing.

For each strategy k the turnover is defined by (see e.g. Golosnoy et al. (2019))

$$\text{Turnover}^{(k)} = \frac{1}{T}\sum_{i=1}^{T} ||\mathbf{w}_i^{(k)} - \mathbf{w}_{i-1}^{(k)}||_1. \tag{IV.1}$$

The turnover can be seen as the cost for transitioning from one portfolio to another, given that the transaction costs are constant for all assets and time periods. The amount of turnover will affect the development of wealth of the portfolio. Moreover, following Golosnoy et al. (2019), we will compute the average absolute values of holding portfolio weights, the average minimum and maximum portfolio weights, the average sum of negative weights in the portfolio, and the average fraction of negative weights in the portfolio

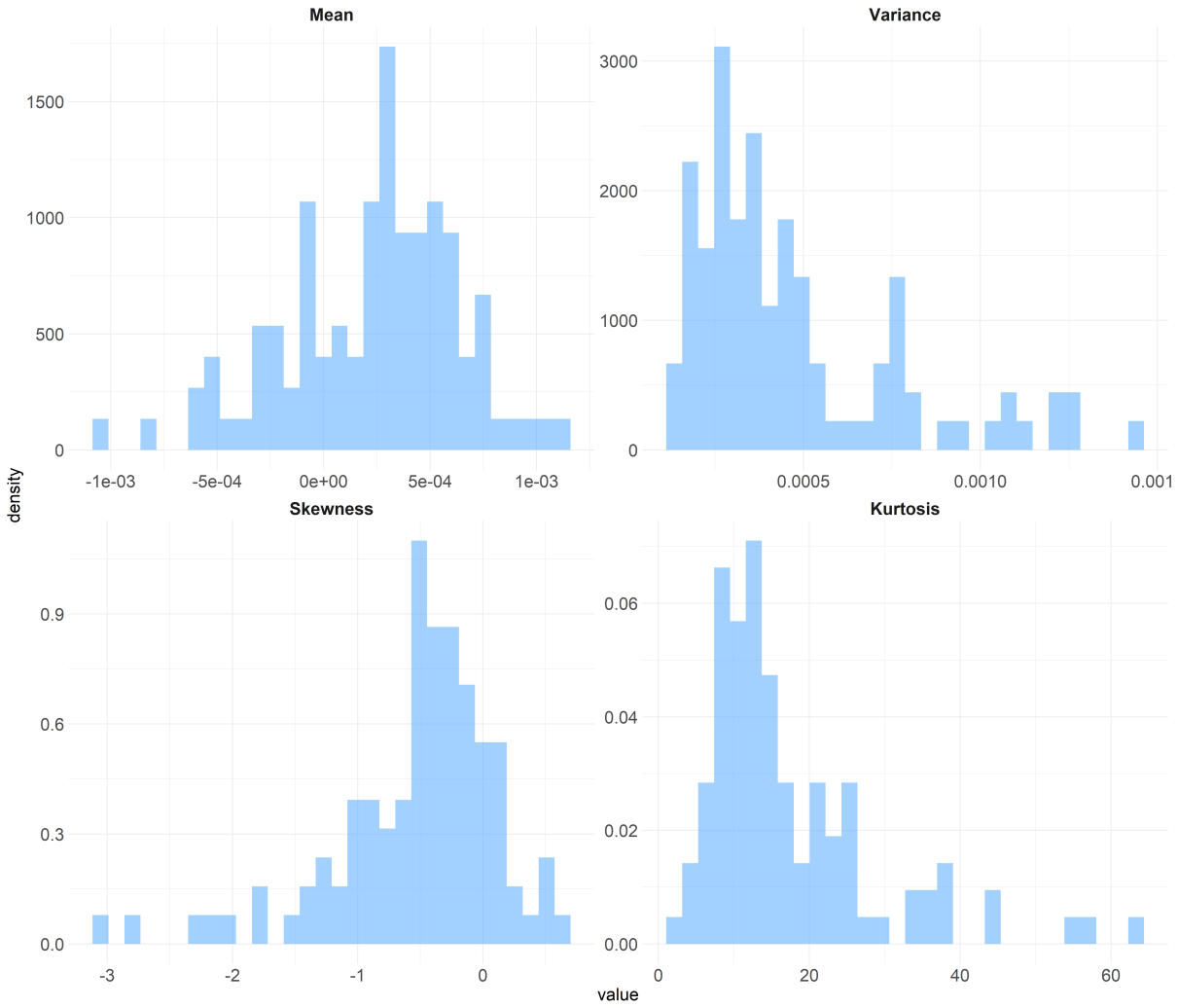

Fig. 5: First four moments for the univariate time series of the asset returns from 100 randomly selected assets of the S&P500 index. The data consist of daily returns from March of 2011 to March of 2021.

as further performance measures. They are given by

$$|\mathbf{w}^{(k)}| = \frac{1}{Tp}\sum_{i=1}^{T}\sum_{j=1}^{p}|w_{i,j}^{(k)}|, \tag{IV.2}$$

$$\max \mathbf{w}^{(k)} = \frac{1}{T}\sum_{i=1}^{T}\left(\max_{j} w_{i,j}^{(k)}\right), \tag{IV.3}$$

$$\min \mathbf{w}^{(k)} = \frac{1}{T}\sum_{i=1}^{T}\left(\min_{j} w_{i,j}^{(k)}\right), \tag{IV.4}$$

$$\mathbf{w}_i^{(k)}\mathbb{1}(\mathbf{w}_i^{(k)} < 0) = \frac{1}{T}\sum_{i=1}^{T}\sum_{j=1}^{p} w_{i,j}^{(k)}\mathbb{1}(w_{i,j}^{(k)} < 0), \tag{IV.5}$$

$$\mathbb{1}(\mathbf{w}_i^{(k)} < 0) = \frac{1}{Tp}\sum_{i=1}^{T}\sum_{j=1}^{p}\mathbb{1}(w_{i,j}^{(k)} < 0). \tag{IV.6}$$

Moreover, we also consider important classical portfolio performance measures: total excess portfolio return, out-of-sample variance and average Sharpe ratio. The computed values of the introduced performance measures are summarized in Table I for **Strategies 1 to 8** over the entire period. The corresponding values for **Strategy 6** are also included in the table but many of its entries are obviously equal to zero since the weights are all equal. In fact, **Strategy 6** has the smallest values of them all for the first 5 performance measures. Similarly, no negative weights are present in the target portfolio since each asset will be assigned a weight of $1/p$. Due to this fact we have chosen to highlight the second to best strategy, which is **Strategy 8**. If the investor is interested in the portfolio weights, then **Strategy 8** is the best as it takes the smallest positions, short the smallest amount of stocks and so forth. The average maximum weights, mean of all shorted and the proportion shorted seem to increase slightly with p. It seems as though when p increases, **Strategy 8** decrease in the other metrics regarding the portfolio weights. The same is not true for the other strategies, especially **Strategy 5**. This can most likely be attributed to the nonlinear shrinkage of the eigenvalues, which seem to have a direct impact on the magnitude and direction of the weights. None of the other strategies has the same flexibility and can not compete with it. The other strategies seem to have natural ordering in terms of the first five metrics. **Strategy 1** through **4** are better than **Strategy 7**, which in turn is better than **Strategy 5**. One interesting feature of the re-estimation strategies, **Strategy 3** and **4**, does not imply a large change in the metrics. One of the largest changes can be seen in the mean out of the shorted weights for **Strategy 4** when $p = 200$. When we re-estimate

the initial relative loss, we change the shrinkage coefficients and all dynamic GMV portfolios thereafter. For the portfolio weights, that seem to imply an increase in certain metrics.

The last performance measures, seen in the four last rows of Table I are based on the portfolio return. **Strategy 3** and **Strategy 4** have the largest portfolio return. Since we only present up to 4 decimals both are highlighted. In reality, they differ on the sixth decimal, which can of course make a big difference depending on the capital invested. In contrast, **Strategy 1** and **Strategy 2** beat all of other strategies in terms of portfolio variance when $p = 150$ and **Strategy 2** has the least amount of variance when $p = 200$. We can also see that as p increases, the variances of **Strategies 5 to 8** increase. The dynamic shrinkage becomes an increasingly important feature in this context. It does not suffice to optimize for the out-of-sample variance, which **Strategies 1 to 4** and **Strategy 7** does, but also the dynamic nature of the investment process. These decrease with p. In turn, for $p = 150$ **Strategy 4** gives the highest Sharpe Ratio and for $p = 200$ **Strategy 2** gives the highest Sharpe Ratio. From the out-of-sample variance it is not surprising to see that **Strategy 5** provides the worst Sharpe ratio.

The last performance measure in the last row is turnover. As one can expect, **Strategy 5** is worst, generating the largest turnover. This classic strategy has the most flexibility. However, this flexibility often leads to unnecessary reconstruction of the holding portfolio. This becomes apparent in more extreme case when p/n is close to one. The smallest turnover is given by **Strategy 1**. The dynamic shrinkage force small movements between reallocations. However, in **Strategy 3** and **Strategy 4** we change the perception of the initial relative loss. This implies that algorithm change its opinion of what the optimal weights should have been. It causes the weights to change more relative to **Strategy 1** and **Strategy 2**. The re-estimation caused a larger return for $p = 150$, a higher variance but also, in turn, a higher turnover. The dynamic shrinkage approach does not optimize towards decreasing turnover but it is a consequence of enforcing small movements. This also implies that other strategies can be close to the dynamic shrinkage approach. **Strategy 8** shows promise of this feature. When p increases the nonlinear shrinkage estimator becomes more stable (between reallocation periods) and increases its turnover slightly.

The development of the investors wealth for the eight trading strategies introduced in Section III-A over 10 years is depicted in Figures 6 and 7. The wealth is computed according to a buy-and-hold strategy until next reallocation period. That is, given that n_i days has passed, we use these to estimate the portfolio weights and then rebalance the holding portfolio to the new portfolio according to the different strategies. The wealth is accumulated on a daily basis which corresponds to the frequency of data used to construct the portfolio.

The portfolios are the same as presented in I, e.g., one set of assets with $p = 150$, $p/n_i = 0.6$ and another with $p = 200$, $p/n_i = 0.8$. In both scenarios the trading strategies based on the dynamic shrinkage approach using the re-estimation technique produce the largest wealth at the end of the investing period.

In Figure 6 where $p = 150$ the worst is **Strategy 5** in terms of the wealth. However, there is a number of strategies that have close to the same performance. These are **Strategy 7** and **Strategy 8**. Throughout the whole period they are very close to each other where **Strategy 8** deviates the most from the other two. The next grouping is **Strategy 1**, **Strategy 2** and **Strategy 6**, the equally weighted portfolio. **Strategy 6** provides slightly less increase in wealth. The other two are very close in terms of their final wealth. They also seem to move in tandem. Their wealth is similar but from I we can see that **Strategy 1**, **Strategy 2** perform slightly different. The classical benchmark, **Strategy 6**, starts of accumulating a lot of wealth, but decrease just prior to 2016. In this figure, it also becomes clear why **Strategy 1** and **Strategy 2** had the smallest variance in Table I. They vary very little in contrast to other strategies.

In Figure 7 where $p = 200$, we have a slightly different scenario. **Strategy 3** and **4** still provide the largest final wealth and **Strategy 1** and **2** second to that. Thereafter there seem to be a clearer separation between the rest in terms of final wealth. **Strategy 6** is still the best out of **Strategies 5 through 8**. However, **Strategy 8** is very close to deliver the same performance. The traditional GMV portfolio takes on both large turnover and does not provide any decent result. It almost cuts the wealth by half in the end of this period. The last strategy, **Strategy 7** is able to catch up and ends on a positive note.

All portfolios are hit quite heavily by COVID in the early 2020, which is indicated by a dashed line on the 1st of March, 2020[1]. Some portfolios are quick to adapt to the event and other are not. However, all portfolios seem to experience a very sharp increase in wealth post COVID. This can most likely be attributed to the very bullish scenario which was caused by banks all over the world pushing capital into the market.

In Table II we show the largest negative difference of the strategies wealth throughout the whole period. We can see that **Strategy 2** has the smallest decrease in wealth between days during this period where **Strategy 1** comes in second. These strategies seem to have been robust against the COVID crisis. Thereafter, **Strategy 3** and **Strategy 4** have the smallest decrease. However, these are almost twice the size. The rest of the strategies are on a similar level. They all exhibit close to a drop of 0.2 units in wealth.

These results are in line with the previous empirical findings of Bodnar et al. (2021a) who document that the equally weighted portfolio performs well in the stable period on the capital market, but its performance is very bad during the turbulent periods. To conclude, all four of the proposed dynamic shrinkage strategies show impressively good performance over the state-of-the-art static portfolios especially in case when p becomes close to n_i.

V. Summary

In many practical situation an investor after constructing an optimal portfolio faces the problem of the portfolio reallocation based on the new data which arrive on the capital market after the portfolio was built. We deal with challenging task in the current paper by developing several dynamic optimal shrinkage estimators for the weights of the GMV portfolio. In the derivation of the theoretical findings, new results in random matrix theory are deduced which allow us to obtained optimal shrinkage estimator in both important cases with and without overlapping samples. In the case of non-overlapping samples, the investor uses the data of asset returns after the last reconstruction of the portfolio, while the whole data might be used in the case of overlapping samples. It is remarkable that the two settings require different theoretical results in random matrix theory to be derived results and they result in quite different optimal shrinkage intensities. Moreover, only minor distribution assumption are imposed on the data-generating process, like the existence of $4+\varepsilon, \varepsilon > 0$, moments are required only. Also, the covariance matrix might have an unbounded spectrum.

The results of the simulation study show that the dynamic shrinkage procedures derived in the paper are robust against violations of the model assumptions. In particular,

[1]This is of course somewhat arbitrary since it is hard to specify a certain day that COVID hits the market.

TABLE I: Performance measures over $T = 8$ periods of portfolio rebalancing. The strategy which is most performant (in terms of its measurement) is highlighted in bold on each row.Each Strategy is abbreviated "S.".

Type	Portfolio size	S. 1	S. 2	S. 3	S. 4	S. 5	S. 6	S. 7	S. 8
$\lvert\mathbf{w}^{(k)}\rvert$	150	0.0245	0.0248	0.0255	0.0265	0.0477	0.0067	0.0298	**0.0209**
	200	0.0218	0.0228	0.0224	0.0248	0.0627	0.0050	0.0300	**0.0173**
$\max \mathbf{w}^{(k)}$	150	0.1107	0.1120	0.1568	0.1633	0.2434	0.0067	0.1537	**0.0836**
	200	0.1736	0.1812	0.1642	0.1823	0.5628	0.0050	0.2633	**0.0865**
$\min \mathbf{w}^{(k)}$	150	-0.1155	-0.1171	-0.1018	-0.1065	-0.1882	0.0067	-0.1134	**-0.0652**
	200	-0.0938	-0.0983	-0.0940	-0.1052	-0.3338	0.0050	-0.1554	**-0.0523**
$\mathbf{w}_i^{(k)}\mathbb{1}(\mathbf{w}_i^{(k)} < 0)$	150	-1.3387	-1.3617	-1.4102	-1.4903	-3.0742	0.0000	-1.7380	**-1.0648**
	200	-1.6802	-1.7767	-1.7367	-1.9807	-5.7679	0.0000	-2.4983	**-1.2292**
$\mathbb{1}(\mathbf{w}_i^{(k)} < 0)$	150	0.4592	0.4592	0.4392	0.4442	0.4750	0.0000	0.4392	**0.4050**
	200	0.4513	0.4519	0.4469	0.4656	0.4863	0.0000	0.4469	**0.4250**
$\bar{\mathbf{y}}_{\mathbf{w}^{(k)}}$	150	0.0003	0.0003	**0.0004**	**0.0004**	0.0001	0.0003	0.0002	0.0001
	200	0.0003	0.0003	**0.0004**	**0.0004**	-0.0001	0.0003	0.0001	0.0002
$\sigma_{\mathbf{w}^{(k)}}$	150	**0.0062**	**0.0062**	0.0083	0.0083	0.0123	0.0120	0.0112	0.0099
	200	0.0055	**0.0054**	0.0080	0.0078	0.0156	0.0122	0.0124	0.0102
$\mathrm{SR}^{(k)}$	150	0.0485	0.0486	0.0513	**0.0525**	0.0107	0.0250	0.0143	0.0138
	200	0.0510	**0.0514**	0.0464	0.0496	-0.0067	0.0208	0.0069	0.0180
$\mathrm{Turnover}^{(k)}$	150	**0.0462**	0.0507	1.4516	1.5016	8.6568	0.0000	5.2351	3.1340
	200	**0.1193**	0.1372	1.9099	2.0772	15.5228	0.0000	7.1262	3.4561

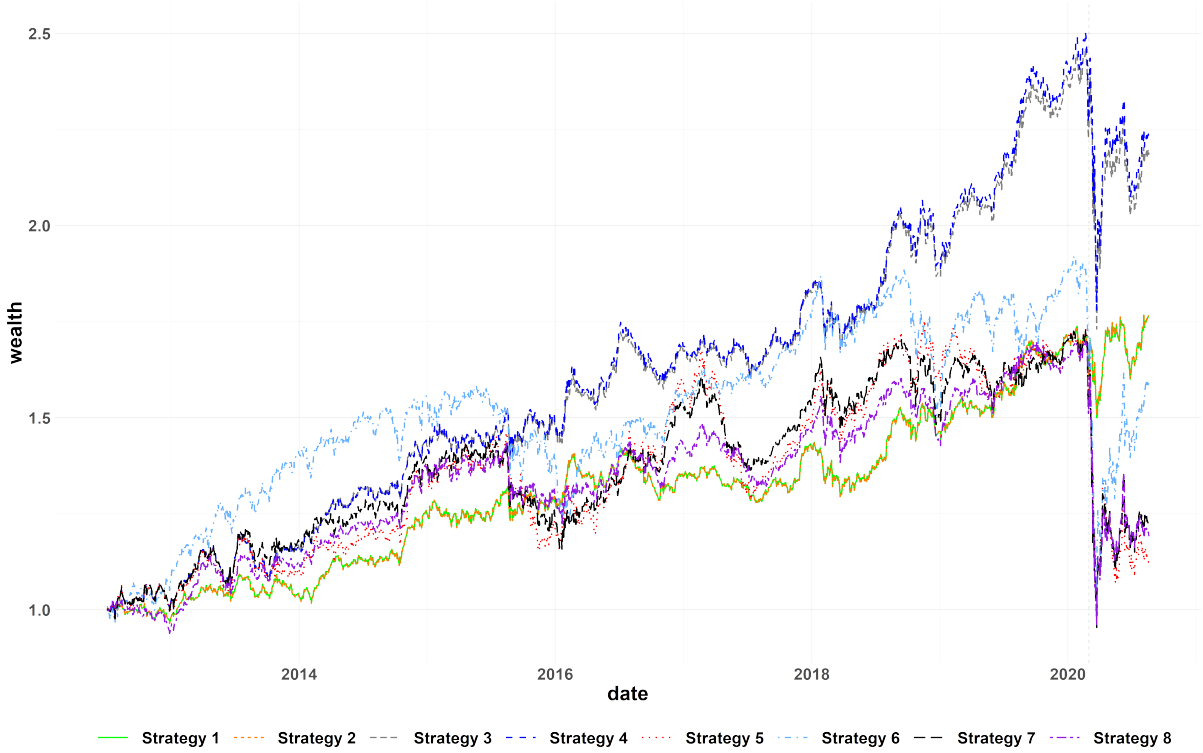

Fig. 6: Development of the investor wealth based on the dynamic trading strategies described in Section III-A. In this figure the portfolio size is equal to 150.

we conclude based on the results of the simulation study, that the performance of the suggested dynamic approach will not be strongly influenced when the asset returns are generated from a multivariate GARCH model and from a VARMA model. Although, both multivariate times series model assume that the asset returns are time dependent, it has only a minor influence the suggested trading strategies. Finally, we apply the new approaches to real data of returns on stocks included in the S&P 500 index and compare them with several benchmark approaches, consisting of investing into the target portfolio, the sample GMV portfolio, and the single-period GMV portfolio. Several performance measures are considered and it is shown that the dynamic shrinkage portfolio constructed by using overlapping samples possesses the best performance in terms of the turnover and the development of the portfolio weights.

The dynamic strategies based non-overlapping sample are simple to implement and they provide drastically less turnover in comparison to the benchmark approaches. Although the approaches based on the overlapping estimators are harder to implement, they decrease the turnover by 50% in comparison to the corresponding non-overlapping strategies with no significant loss in wealth. Furthermore, they require that the sample size is larger than the portfolio dimension only when the portfolio is constructed for the first time, while the non-overlapping approaches need the sample size to be larger than the portfolio dimension by each reconstruction of the portfolio.

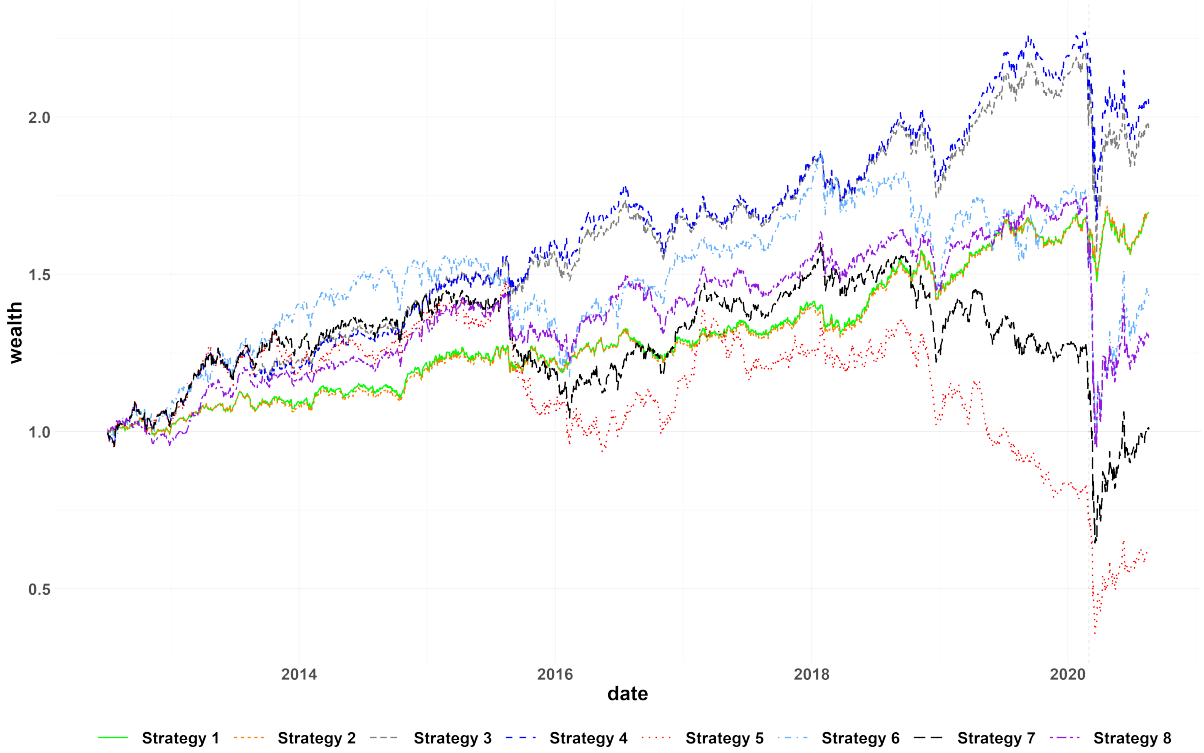

Fig. 7: Development of the investor wealth based on the dynamic trading strategies described in Section III-A. In this figure the portfolio size is equal to 200.

TABLE II: The largest loss in wealth between days throughout the whole period. The strategy which is most performant, e.g. smallest loss, is highlighted in bold on each row. Each Strategy is abbreviated "S.".

size	S. 1	S. 2	S. 3	S. 4	S. 5	S. 6	S. 7	S. 8
150	-0.0586	**-0.0580**	-0.1180	-0.1174	-0.1975	-0.2200	-0.2048	-0.1744
200	-0.0469	**-0.0442**	-0.1226	-0.1190	-0.2701	-0.2266	-0.2482	-0.2009

No portfolio is ever static. Making optimal transitions are therefore of great interest to any investor. These results provide a fully data-driven dynamic approaches how the GMV portfolio can be rebalanced. In many practical applications the investors might want to have more assets in their portfolios than the available sample size. This demands a special attention since the sample covariance matrix is singular in this case and its inverse does not exist any longer. This challenging problem has not been treated in the paper and is left for future research.

ACKNOWLEDGEMENT

This research was partly supported by the Swedish Research Council (VR) via the project "Bayesian Analysis of Optimal Portfolios and Their Risk Measures"

APPENDIX

In this section the proofs of the theoretical results are given. We first state several lemmas which will be used in the proofs of the theorems.

For any integer $n > 2$, we define

$$\mathbf{V}_n = \frac{1}{n-1}\mathbf{X}_n\left(\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top\right)\mathbf{X}_n^\top, \widetilde{\mathbf{V}}_n = \frac{1}{n-1}\mathbf{X}_n\mathbf{X}_n^\top, \tag{V.1}$$

where $\mathbf{X}_n$ is given in (II.1) for the special case $n = n_i$. Hence,

$$\mathbf{S}_n = \mathbf{\Sigma}^{1/2}\mathbf{V}_n\mathbf{\Sigma}^{1/2} = \mathbf{\Sigma}^{1/2}\widetilde{\mathbf{V}}_n\mathbf{\Sigma}^{1/2} - \frac{n}{n-1}\mathbf{\Sigma}^{1/2}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n^\top\mathbf{\Sigma}^{1/2}$$

with $\bar{\mathbf{x}}_n = \frac{1}{n}\mathbf{X}_n\mathbf{1}_n = \mathbf{\Sigma}^{-1/2}\bar{\mathbf{y}}_n$.

The statement of Lemma V.1 is derived as Lemma 5.3 in Bodnar et al. (2022b)

Lemma V.1. *Let $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$ be two nonrandom vectors with bounded Euclidean norms. Then it holds that*

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

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

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

Lemma V.2. *Let $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$ be two nonrandom vectors with bounded Euclidean norms and let $m, n > 1$. Then it holds that*

$$\left|\boldsymbol{\xi}^\top\widetilde{\mathbf{V}}_n^{-1}\widetilde{\mathbf{V}}_{n+m}^{-1}\boldsymbol{\theta} - d_{n,n+m}\boldsymbol{\xi}^\top\boldsymbol{\theta}\right| \xrightarrow{a.s.} 0, \tag{V.4}$$

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

$$d_{n,n+m} = \frac{b_{n,n+m}}{a_{n,n+m}}\Bigg(\left(1-\frac{p}{n}\right)^{-1} - 2\Bigg(1-\frac{p}{n}+a_{n,n+m} + \sqrt{\left(1-\frac{p}{n}-a_{n,n+m}\right)^2+4a_{n,n+m}}\Bigg)^{-1}\Bigg), \quad \text{(V.5)}$$

where

$$a_{n,n+m} = \frac{n+m-p}{n+m}\frac{n+m-1}{n-1}, b_{n,n+m} = \frac{n+m-1}{n-1}. \quad \text{(V.6)}$$

Proof of Lemma V.2. It holds that

$$\begin{aligned} & \left|\boldsymbol{\xi}^\top \tilde{\mathbf{V}}_n^{-1}\tilde{\mathbf{V}}_{n+m}^{-1}\boldsymbol{\theta} - d_{n,n+m}\boldsymbol{\xi}^\top\boldsymbol{\theta}\right| \quad \text{(V.7)} \\ \leq & \left|\frac{\boldsymbol{\xi}^\top \tilde{\mathbf{V}}_n^{-1}\tilde{\mathbf{V}}_{n+m}^{-1}\boldsymbol{\theta}}{\sqrt{\boldsymbol{\xi}^\top \tilde{\mathbf{V}}_n^{-2}\boldsymbol{\xi}}\sqrt{\boldsymbol{\theta}^\top\boldsymbol{\theta}}} - \frac{\boldsymbol{\xi}^\top \tilde{\mathbf{V}}_n^{-1}\left(\frac{n-1}{n+m-1}\tilde{\mathbf{V}}_n + \frac{n+m-p}{n+m}\mathbf{I}\right)^{-1}\boldsymbol{\theta}}{\sqrt{\boldsymbol{\xi}^\top \tilde{\mathbf{V}}_n^{-2}\boldsymbol{\xi}}\sqrt{\boldsymbol{\theta}^\top\boldsymbol{\theta}}}\right| \\ \times & \sqrt{\boldsymbol{\xi}^\top \tilde{\mathbf{V}}_n^{-2}\boldsymbol{\xi}}\sqrt{\boldsymbol{\theta}^\top\boldsymbol{\theta}} \\ + & \left|\boldsymbol{\xi}^\top \tilde{\mathbf{V}}_n^{-1}\left(\frac{n-1}{n+m-1}\tilde{\mathbf{V}}_n + \frac{n+m-p}{n+m}\mathbf{I}\right)^{-1}\boldsymbol{\theta} - d_{n,n+m}\boldsymbol{\xi}^\top\boldsymbol{\theta}\right|, \end{aligned}$$

where $\boldsymbol{\theta}^\top\boldsymbol{\theta} < \infty$ by the assumption and by Lemma 5.2 of Bodnar et al. (2022b) we get $\boldsymbol{\xi}^\top\tilde{\mathbf{V}}_n^{-2}\boldsymbol{\xi} \overset{a.s.}{\to} (1-c)^{-3}\boldsymbol{\xi}^\top\boldsymbol{\xi}$.

Let $\widetilde{\mathbf{V}}_{n+1:n+m} = \frac{1}{m-1}\mathbf{X}_{n+1:n+m}\mathbf{X}_{n+1:n+m}^\top$ where $\mathbf{X}_{n+1:n+m}$ stands for the submatrix of $\mathbf{X}_{n+m}$ consisting of its last m columns. Since $\mathbf{X}_n$ and $\mathbf{X}_{n+1:n+m}$ are independent, $\widetilde{\mathbf{V}}_n$ and $\widetilde{\mathbf{V}}_{n+1:n+m}$ are also independent. Moreover, we get

$$\widetilde{\mathbf{V}}_{n+m} = \frac{n-1}{n+m-1}\widetilde{\mathbf{V}}_n + \frac{m-1}{n+m-1}\widetilde{\mathbf{V}}_{n+1:n+m}.$$

Following Theorem 1 in Rubio and Mestre (2011) conditionally on $\widetilde{\mathbf{V}}_n$, we get that

$$\left|\frac{\boldsymbol{\xi}^\top \tilde{\mathbf{V}}_n^{-1}\tilde{\mathbf{V}}_{n+m}^{-1}\boldsymbol{\theta}}{\sqrt{\boldsymbol{\xi}^\top \tilde{\mathbf{V}}_n^{-2}\boldsymbol{\xi}}\sqrt{\boldsymbol{\theta}^\top\boldsymbol{\theta}}} - \frac{\boldsymbol{\xi}^\top \tilde{\mathbf{V}}_n^{-1}\left(\frac{n-1}{n+m-1}\tilde{\mathbf{V}}_n + k_{n,n+m}\mathbf{I}\right)^{-1}\boldsymbol{\theta}}{\sqrt{\boldsymbol{\xi}^\top \tilde{\mathbf{V}}_n^{-2}\boldsymbol{\xi}}\sqrt{\boldsymbol{\theta}^\top\boldsymbol{\theta}}}\right| \overset{a.s.}{\to} 0$$

for $p/m \to \tilde{c} \in (0,\infty)$ as $m \to \infty$ where $k_{n,n+m}$ is found as the solution of the equation

$$\begin{aligned} & \frac{m}{p}(k_{n,n+m}^{-1} - 1) \quad \text{(V.8)} \\ = & \frac{1}{p}\mathrm{tr}\left(\frac{m}{n+m-1}\left(\frac{n-1}{n+m-1}\tilde{\mathbf{V}}_n + k_{n,n+m}\frac{m}{n+m-1}\mathbf{I}\right)^{-1}\right) \\ \overset{a.s.}{\to} & 2\frac{m}{n}\Bigg(1-\frac{p}{n}+k_{n,n+m}\frac{m}{n} \\ + & \sqrt{\left(1-\frac{p}{n}-k_{n,n+m}\frac{m}{n}\right)^2+4k_{n,n+m}\frac{m}{n}}\Bigg)^{-1} \end{aligned}$$

Using that for $a, b \in \mathbb{R}$ we have

$$\left((a+b)+\sqrt{(a-b)^2+4b}\right)\left((a+b)-\sqrt{(a-b)^2+4b}\right) = 4b(a-1),$$

the identity (V.8) is equivalent to

$$\begin{aligned} 2\frac{m}{n}(k_{n,n+m}-1) = & 1-\frac{p}{n}+k_{n,n+m}\frac{m}{n} \\ - & \sqrt{\left(1-\frac{p}{n}-k_{n,n+m}\frac{m}{n}\right)^2+4k_{n,n+m}\frac{m}{n}} \end{aligned}$$

or

$$\begin{aligned} & \sqrt{\left(1-\frac{p}{n}-k_{n,n+m}\frac{m}{n}\right)^2+4k_{n,n+m}\frac{m}{n}} \\ = & 1-\frac{p}{n}+2\frac{m}{n}-\frac{m}{n}k_{n,n+m}, \end{aligned}$$

whose solution is given by

$$k_{n,n+m} = \frac{n+m-p}{n+m}.$$

For the second summand in (V.8) we get that

$$\begin{aligned} & \boldsymbol{\xi}^\top\tilde{\mathbf{V}}_n^{-1}\left(\frac{n-1}{n+m-1}\widetilde{\mathbf{V}}_n + \frac{n+m-p}{n+m}\mathbf{I}\right)^{-1}\boldsymbol{\theta} \\ = & b_{n,n+m}\boldsymbol{\xi}^\top\tilde{\mathbf{V}}_n^{-1}\left(\widetilde{\mathbf{V}}_n + a_{n,n+m}\mathbf{I}\right)^{-1}\boldsymbol{\theta} \\ = & \frac{b_{n,n+m}}{a_{n,n+m}}\left(\boldsymbol{\xi}^\top\tilde{\mathbf{V}}_n^{-1}\boldsymbol{\theta} - \boldsymbol{\xi}^\top\left(\widetilde{\mathbf{V}}_n + a_{n,n+m}\mathbf{I}\right)^{-1}\boldsymbol{\theta}\right) \\ \overset{a.s.}{\to} & \frac{b_{n,n+m}}{a_{n,n+m}}\Bigg(\left(1-\frac{p}{n}\right)^{-1} - 2\Bigg(1-\frac{p}{n}+a_{n,n+m} \\ + & \sqrt{\left(1-\frac{p}{n}-a_{n,n+m}\right)^2+4a_{n,n+m}}\Bigg)^{-1}\Bigg)\boldsymbol{\xi}^\top\boldsymbol{\theta} \end{aligned}$$

using Lemma 5.1 in Bodnar et al. (2022b) two times. □

Lemma V.3. *Let $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$ be two nonrandom vectors with bounded Euclidean norms and let $m, n > 1$. Then it holds that*

$$\left|\boldsymbol{\xi}^\top\mathbf{V}_n^{-1}\mathbf{V}_{n+m}^{-1}\boldsymbol{\theta} - d_{n,n+m}\boldsymbol{\xi}^\top\boldsymbol{\theta}\right| \overset{a.s.}{\to} 0, \quad \text{(V.9)}$$

for $p/n \to c \in (0,1)$ as $n \to \infty$ where $d_{n,n+m}$ is defined in (V.5)*.*

Proof of Lemma V.3. The application of the Sherman–Morrison formula leads to

$$\begin{aligned} & \boldsymbol{\xi}^\top\mathbf{V}_n^{-1}\mathbf{V}_{n+m}^{-1}\boldsymbol{\theta} = \boldsymbol{\xi}^\top\left(\widetilde{\mathbf{V}}_n - \frac{n}{n-1}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n^\top\right)^{-1} \\ \times & \left(\widetilde{\mathbf{V}}_{n+m} - \frac{n+m}{n+m-1}\bar{\mathbf{x}}_{n+m}\bar{\mathbf{x}}_{n+m}^\top\right)^{-1}\boldsymbol{\theta} \\ = & \boldsymbol{\xi}^\top\tilde{\mathbf{V}}_n^{-1}\tilde{\mathbf{V}}_{n+m}^{-1}\boldsymbol{\theta} + \frac{n+m}{n+m-1}\frac{\boldsymbol{\xi}^\top\tilde{\mathbf{V}}_n^{-1}\tilde{\mathbf{V}}_{n+m}^{-1}\bar{\mathbf{x}}_{n+m}\bar{\mathbf{x}}_{n+m}^\top\tilde{\mathbf{V}}_{n+m}^{-1}\boldsymbol{\theta}}{1-\frac{n+m}{n+m-1}\bar{\mathbf{x}}_{n+m}^\top\tilde{\mathbf{V}}_{n+m}^{-1}\bar{\mathbf{x}}_{n+m}} \\ + & \frac{n}{n-1}\frac{\boldsymbol{\xi}^\top\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n^\top\tilde{\mathbf{V}}_n^{-1}\tilde{\mathbf{V}}_{n+m}^{-1}\boldsymbol{\theta}}{1-\frac{n}{n-1}\bar{\mathbf{x}}_n^\top\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n} + \frac{n}{n-1}\frac{n+m}{n+m-1} \\ \times & \frac{\boldsymbol{\xi}^\top\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n^\top\tilde{\mathbf{V}}_n^{-1}\tilde{\mathbf{V}}_{n+m}^{-1}\bar{\mathbf{x}}_{n+m}\bar{\mathbf{x}}_{n+m}^\top\tilde{\mathbf{V}}_{n+m}^{-1}\boldsymbol{\theta}}{(1-\frac{n}{n-1}\bar{\mathbf{x}}_n^\top\tilde{\mathbf{V}}_n^{-1}\bar{\mathbf{x}}_n)(1-\frac{n+m}{n+m-1}\bar{\mathbf{x}}_{n+m}^\top\tilde{\mathbf{V}}_{n+m}^{-1}\bar{\mathbf{x}}_{n+m})}. \end{aligned}$$

The applications of Lemma 5.2 in Bodnar et al. (2022b) and Lemma V.2 completes the proof. □

Proof of Theorem II.1. Rewriting (II.7) we get

$$\psi_{n_i}^* = \frac{\hat{\mathbf{w}}_{SH;n_{i-1}}^\top\boldsymbol{\Sigma}\hat{\mathbf{w}}_{SH;n_{i-1}} - \frac{\hat{\mathbf{w}}_{SH;n_{i-1}}^\top\boldsymbol{\Sigma}\mathbf{S}_{n_i}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top\mathbf{S}_{n_i}^{-1}\mathbf{1}_p}}{\hat{\mathbf{w}}_{SH;n_{i-1}}^\top\boldsymbol{\Sigma}\hat{\mathbf{w}}_{SH;n_{i-1}} - 2\frac{\hat{\mathbf{w}}_{SH;n_{i-1}}^\top\boldsymbol{\Sigma}\mathbf{S}_{n_i}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top\mathbf{S}_{n_i}^{-1}\mathbf{1}_p} + \frac{\mathbf{1}_p^\top\mathbf{S}_{n_i}^{-1}\boldsymbol{\Sigma}\mathbf{S}_{n_i}^{-1}\mathbf{1}_p}{(\mathbf{1}_p^\top\mathbf{S}_{n_i}^{-1}\mathbf{1}_p)^2}} = \frac{h_{n_i}}{g_{n_i}}$$

with

$$\begin{aligned} h_{n_i} = & \mathbf{1}_p^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}_p\hat{\mathbf{w}}_{SH;n_{i-1}}^\top\boldsymbol{\Sigma}\hat{\mathbf{w}}_{SH;n_{i-1}} \\ - & \frac{\mathbf{1}_p^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top\mathbf{S}_{n_i}^{-1}\mathbf{1}_p}\hat{\mathbf{w}}_{SH;n_{i-1}}^\top\boldsymbol{\Sigma}\mathbf{S}_{n_i}^{-1}\mathbf{1}_p \end{aligned}$$

and

$$\begin{aligned} g_{n_i} = & \mathbf{1}_p^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}_p\hat{\mathbf{w}}_{SH;n_{i-1}}^\top\boldsymbol{\Sigma}\hat{\mathbf{w}}_{SH;n_{i-1}} \\ + & \frac{\mathbf{1}_p^\top\mathbf{S}_{n_i}^{-1}\boldsymbol{\Sigma}\mathbf{S}_{n_i}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}_p}\left(\frac{\mathbf{1}_p^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top\mathbf{S}_{n_i}^{-1}\mathbf{1}_p}\right)^2 \\ - & 2\frac{\mathbf{1}_p^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top\mathbf{S}_{n_i}^{-1}\mathbf{1}_p}\hat{\mathbf{w}}_{SH;n_{i-1}}^\top\boldsymbol{\Sigma}\mathbf{S}_{n_i}^{-1}\mathbf{1}_p \end{aligned}$$

The application of Lemma V.1 with $\boldsymbol{\xi} = \boldsymbol{\theta} = \dfrac{\boldsymbol{\Sigma}^{-1/2}\mathbf{1}_p}{\mathbf{1}_p^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}_p}$ leads to

$$\frac{\mathbf{1}_p^\top\mathbf{S}_{n_i}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}_p} \overset{a.s.}{\to} (1-c_i)^{-1}, \frac{\mathbf{1}_p^\top\mathbf{S}_{n_i}^{-1}\boldsymbol{\Sigma}\mathbf{S}_{n_i}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}_p} \overset{a.s.}{\to} (1-c_i)^{-3}$$

(V.10)

for $p/n_i \to c_i \in (0,1)$ as $n_i \to \infty$.

Since the variance of the GMV portfolio is equal to $1/(\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p)$, we have that

$$\begin{aligned}\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p \hat{\mathbf{w}}_{SH;n_{i-1}}^\top \mathbf{\Sigma}\hat{\mathbf{w}}_{SH;n_{i-1}} &= \frac{\hat{\mathbf{w}}_{SH;n_{i-1}}^\top \mathbf{\Sigma}\hat{\mathbf{w}}_{SH;n_{i-1}}}{1/\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p} \\ &= 1 + r_{\hat{\mathbf{w}}_{SH;n_{i-1}}} \geq 1,\end{aligned}$$

where $r_{\hat{\mathbf{w}}_{SH;n_{i-1}}}$ is the relative loss of the portfolio with weights $\hat{\mathbf{w}}_{SH;n_{i-1}}$ computed with respect to the variance of the GMV portfolio. Next, we recursively derive the asymptotic behaviour of $r_{\hat{\mathbf{w}}_{SH;n_{i-1}}}$.

For $i = 1$ the sample estimator of the GMV portfolio weights $\hat{\mathbf{w}}_{S;n_i}$ is shrunk to the deterministic target $\mathbf{b}$, i.e., $\hat{\mathbf{w}}_{SH;n_0} = \mathbf{b}$. In this case the relative loss is given by

$$r_{\hat{\mathbf{w}}_{SH;n_0}} = r_0 = \mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p \mathbf{b}^\top \mathbf{\Sigma}\mathbf{b} - 1,$$

which is bounded uniformly on p following the assumption of the theorem and

$$\begin{aligned}\frac{\mathbf{b}^\top \mathbf{\Sigma}\mathbf{S}_{n_1}^{-1}\mathbf{1}_p}{\sqrt{\mathbf{b}^\top \mathbf{\Sigma}\mathbf{b}}\sqrt{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}} &\overset{a.s.}{\to} (1-c_1)^{-1}\frac{1}{\sqrt{\mathbf{b}^\top \mathbf{\Sigma}\mathbf{b}}\sqrt{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}} \\ &= (1-c_1)^{-1}\frac{1}{\sqrt{r_0+1}},\end{aligned}$$

using Lemma V.1 with $\boldsymbol{\xi} = \dfrac{\mathbf{\Sigma}^{1/2}\mathbf{b}}{\sqrt{\mathbf{b}^\top \mathbf{\Sigma}\mathbf{b}}}$ and $\boldsymbol{\theta} = \dfrac{\mathbf{\Sigma}^{-1/2}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}$. Combining these results with (V.10) leads to

$$\psi_{n_1}^* \overset{a.s.}{\to} \psi_1^*$$

for $p/n_1 \to c_1 \in (0,1)$ as $n_1 \to \infty$ with

$$\psi_1^* = \frac{(r_0+1)-1}{(r_0+1)+(1-c_1)^{-1}-2} = \frac{(1-c_1)r_0}{(1-c_1)r_0 + c_1}.$$

Moreover, the relative loss of the portfolio $\hat{\mathbf{w}}_{SH;n_{i-1}}$ yields

$$\begin{aligned}r_{\hat{\mathbf{w}}_{SH;n_1}} &= \mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p \Bigg((\psi_{n_1}^*)^2 \hat{\mathbf{w}}_{S;n_1}^\top \mathbf{\Sigma}\hat{\mathbf{w}}_{S;n_1} \\ &+ 2\psi_{n_1}^*(1-\psi_{n_1}^*)\hat{\mathbf{w}}_{S;n_1}^\top \mathbf{\Sigma}\mathbf{b} + (1-\psi_{n_1}^*)^2 \mathbf{b}^\top \mathbf{\Sigma}\mathbf{b}\Bigg) - 1 \\ &= (\psi_{n_1}^*)^2 \frac{\mathbf{1}_p^\top \mathbf{S}_{n_1}^{-1}\mathbf{\Sigma}\mathbf{S}_{n_1}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}\left(\frac{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_{n_1}^{-1}\mathbf{1}_p}\right)^2 \\ &+ 2\psi_{n_1}^*(1-\psi_{n_1}^*)\mathbf{1}_p^\top \mathbf{S}_{n_1}^{-1}\mathbf{\Sigma}\mathbf{b}\frac{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_{n_1}^{-1}\mathbf{1}_p} \\ &+ (1-\psi_{n_1}^*)^2(r_0+1) - 1 \\ &\overset{a.s.}{\to} (\psi_1^*)^2(1-c_1)^{-1} + 2\psi_1^*(1-\psi_1^*) \\ &+ (1-\psi_1^*)^2(r_0+1) - 1 \\ &= (\psi_1^*)^2\frac{c_1}{1-c_1} + (1-\psi_1^*)^2 r_0 = r_1,\end{aligned}$$

for $p/n_1 \to c_1 \in (0,1)$ as $n_1 \to \infty$.

Using the last result we get for $i = 2$ that

$$\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p \hat{\mathbf{w}}_{SH;n_1}^\top \mathbf{\Sigma}\hat{\mathbf{w}}_{SH;n_1} \overset{a.s.}{\to} r_1 + 1.$$

Since non-overlaping samples $\mathbf{Y}_{n_1}$ $\mathbf{Y}_{n_2}$ are used in the computation of the $\hat{\mathbf{w}}_{SH;n_1}$ and $\mathbf{S}_{n_2}$, these two random objects are independent. The application of Lemma V.1 conditionally on $\mathbf{Y}_{n_1}$ leads to

$$\begin{aligned}&\left| \frac{\hat{\mathbf{w}}_{SH;n_1}^\top \mathbf{\Sigma}\mathbf{S}_{n_2}^{-1}\mathbf{1}_p}{\sqrt{\hat{\mathbf{w}}_{SH;n_1}^\top \mathbf{\Sigma}\hat{\mathbf{w}}_{SH;n_1}}\sqrt{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}} \right. \\ &\left. - \frac{(1-c_2)^{-1}}{\sqrt{\hat{\mathbf{w}}_{SH;n_1}^\top \mathbf{\Sigma}\hat{\mathbf{w}}_{SH;n_1}}\sqrt{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}} \right| \overset{a.s.}{\to} 0,\end{aligned}$$

and, hence,

$$\begin{aligned}&\left| \hat{\mathbf{w}}_{SH;n_1}^\top \mathbf{\Sigma}\mathbf{S}_{n_2}^{-1}\mathbf{1}_p - (1-c_2)^{-1} \right| \\ &= \left| \frac{\hat{\mathbf{w}}_{SH;n_1}^\top \mathbf{\Sigma}\mathbf{S}_{n_2}^{-1}\mathbf{1}_p}{\sqrt{\hat{\mathbf{w}}_{SH;n_1}^\top \mathbf{\Sigma}\hat{\mathbf{w}}_{SH;n_1}}\sqrt{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}} \right. \\ &\left. \times\sqrt{\hat{\mathbf{w}}_{SH;n_1}^\top \mathbf{\Sigma}\hat{\mathbf{w}}_{SH;n_1}}\sqrt{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p} - (1-c_2)^{-1} \right| \overset{a.s.}{\to} 0,\end{aligned}$$

since $\hat{\mathbf{w}}_{SH;n_1}^\top \mathbf{\Sigma}\hat{\mathbf{w}}_{SH;n_1}\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p = O_P(1)$ The last results together with (V.10) yields

$$\psi_{n_2}^* \overset{a.s.}{\to} \psi_2^* = \frac{(r_1+1)-1}{(r_1+1)+(1-c_2)^{-1}-2} = \frac{(1-c_2)r_1}{(1-c_2)r_1 + c_2}$$

for $p/n_2 \to c_2 \in (0,1)$ as $n_2 \to \infty$.

Finally, the relative loss of the portfolio $\hat{\mathbf{w}}_{SH;n_2}$ tends to

$$\begin{aligned}r_{\hat{\mathbf{w}}_{SH;n_2}} &= \mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p \Bigg((\psi_{n_2}^*)^2 \hat{\mathbf{w}}_{S;n_2}^\top \mathbf{\Sigma}\hat{\mathbf{w}}_{S;n_2} \\ &+ 2\psi_{n_2}^*(1-\psi_{n_2}^*)\hat{\mathbf{w}}_{S;n_2}^\top \mathbf{\Sigma}\hat{\mathbf{w}}_{SH;n_1} \\ &+ (1-\psi_{n_2}^*)^2 \hat{\mathbf{w}}_{SH;n_1}^\top \mathbf{\Sigma}\hat{\mathbf{w}}_{SH;n_1}\Bigg) - 1 \\ &= (\psi_{n_2}^*)^2 \frac{\mathbf{1}_p^\top \mathbf{S}_{n_2}^{-1}\mathbf{\Sigma}\mathbf{S}_{n_2}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}\left(\frac{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_{n_2}^{-1}\mathbf{1}_p}\right)^2 \\ &+ 2\psi_{n_2}^*(1-\psi_{n_2}^*)\mathbf{1}_p^\top \mathbf{S}_{n_2}^{-1}\mathbf{\Sigma}\hat{\mathbf{w}}_{SH;n_1}\frac{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_{n_2}^{-1}\mathbf{1}_p} \\ &+ (1-\psi_{n_2}^*)^2(r_{\hat{\mathbf{w}}_{SH;n_1}}+1) - 1 \\ &\overset{a.s.}{\to} (\psi_2^*)^2(1-c_2)^{-1} + 2\psi_2^*(1-\psi_2^*) + (1-\psi_2^*)^2(r_1+1) - 1 \\ &= (\psi_2^*)^2\frac{c_2}{1-c_2} + (1-\psi_2^*)^2 r_1 = r_2,\end{aligned}$$

for $p/n_2 \to c_2 \in (0,1)$ as $n_2 \to \infty$.

Repeating the above steps for $i = 3, ..., T$ leads to the statement of the theorem. □

Proof of Theorem II.5. We get

$$\Psi_{N_i}^* = \frac{\hat{\mathbf{w}}_{SH;N_{i-1}}^\top \mathbf{\Sigma}\hat{\mathbf{w}}_{SH;N_{i-1}} - \frac{\hat{\mathbf{w}}_{SH;N_{i-1}}^\top \mathbf{\Sigma}\mathbf{S}_{N_i}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_{N_i}^{-1}\mathbf{1}_p}}{\hat{\mathbf{w}}_{SH;N_{i-1}}^\top \mathbf{\Sigma}\hat{\mathbf{w}}_{SH;N_{i-1}} - 2\frac{\hat{\mathbf{w}}_{SH;N_{i-1}}^\top \mathbf{\Sigma}\mathbf{S}_{N_i}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_{N_i}^{-1}\mathbf{1}_p} + \frac{\mathbf{1}_p^\top \mathbf{S}_{N_i}^{-1}\mathbf{\Sigma}\mathbf{S}_{N_i}^{-1}\mathbf{1}_p}{(\mathbf{1}_p^\top \mathbf{S}_{N_i}^{-1}\mathbf{1}_p)^2}} = \frac{H_{N_i}}{G_{N_i}}$$

with

$$\begin{aligned}H_{N_i} &= \mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p \hat{\mathbf{w}}_{SH;N_{i-1}}^\top \mathbf{\Sigma}\hat{\mathbf{w}}_{SH;N_{i-1}} \\ &- \frac{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_{N_i}^{-1}\mathbf{1}_p}\hat{\mathbf{w}}_{SH;N_{i-1}}^\top \mathbf{\Sigma}\mathbf{S}_{N_i}^{-1}\mathbf{1}_p\end{aligned}$$

and

$$\begin{aligned}G_{N_i} &= \mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p \hat{\mathbf{w}}_{SH;N_{i-1}}^\top \mathbf{\Sigma}\hat{\mathbf{w}}_{SH;N_{i-1}} \\ &+ \frac{\mathbf{1}_p^\top \mathbf{S}_{N_i}^{-1}\mathbf{\Sigma}\mathbf{S}_{N_i}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}\left(\frac{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_{N_i}^{-1}\mathbf{1}_p}\right)^2 \\ &- 2\frac{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}_{N_i}^{-1}\mathbf{1}_p}\hat{\mathbf{w}}_{SH;N_{i-1}}^\top \mathbf{\Sigma}\mathbf{S}_{N_i}^{-1}\mathbf{1}_p\end{aligned}$$

From Lemma V.1 we have

$$\frac{\mathbf{1}_p^\top \mathbf{S}_{N_i}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p} \overset{a.s.}{\to} (1-C_i)^{-1}, \frac{\mathbf{1}_p^\top \mathbf{S}_{N_i}^{-1}\mathbf{\Sigma}\mathbf{S}_{N_i}^{-1}\mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{\Sigma}^{-1}\mathbf{1}_p} \overset{a.s.}{\to} (1-C_i)^{-3} \tag{V.11}$$

for $p/N_i \to C_i \in (0,1)$ as $N_i \to \infty$.

The recursive structure of $\hat{\mathbf{w}}_{SH;N_{i-1}}$ implies that

$$\hat{\mathbf{w}}_{SH;N_{i-1}} = \sum_{j=0}^{i-1} \beta^*_{N_{i-1};N_j} \hat{\mathbf{w}}_{N_j} \quad \text{with} \quad \hat{\mathbf{w}}_{N_0} = \mathbf{b},$$

where $\beta^*_{N_{i-1};N_j}$ are computed recursively by $\beta^*_{N_0;N_0} = 1$, $\beta^*_{N_{i-1};N_{i-1}} = \Psi^*_{N_{i-1}}$ and $\beta^*_{N_{i-1};N_j} = (1 - \Psi^*_{N_{i-1}})\beta^*_{N_{i-2};N_j}$, $j = 0, ...i-2$. Hence,

$$\begin{aligned}\hat{\mathbf{w}}^\top_{SH;N_{i-1}} \boldsymbol{\Sigma} \mathbf{S}^{-1}_{N_i} \mathbf{1}_p &= \beta^*_{N_{i-1};N_0} \mathbf{b}^\top \boldsymbol{\Sigma} \mathbf{S}^{-1}_{N_i} \mathbf{1}_p \\ &+ \sum_{j=1}^{i-1} \beta^*_{N_{i-1};N_j} \frac{\mathbf{1}_p^\top \mathbf{S}^{-1}_{N_j} \boldsymbol{\Sigma} \mathbf{S}^{-1}_{N_i} \mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}^{-1}_{N_j} \mathbf{1}_p}\end{aligned}$$

Let Ψ^*_{i-1} denote the deterministic asymptotic limit of $\Psi^*_{N_{i-1}}$, whose recursive computation is discussed below. We also define $\beta^*_{0;0} = 1, \beta^*_{i-1;i-1} = \Psi^*_{i-1}, \beta^*_{i-1;j} = (1 - \Psi^*_{i-1})\beta^*_{i-2;j}, j = 0, ...i-2$. Then, the application of Lemma V.3 yields

$$\hat{\mathbf{w}}^\top_{SH;N_{i-1}} \boldsymbol{\Sigma} \mathbf{S}^{-1}_{N_i} \mathbf{1}_p \xrightarrow{a.s.} (1 - C_i)^{-1} K_i$$

with $K_i = \beta^*_{i-1;0} + \sum_{j=1}^{i-1} \beta^*_{i-1;j} D_{j,i}$ for $p/N_j \to C_j \in (0,1)$ as $N_j \to \infty$, $j = 1, ..., i-1$, where

$$D_{j,i} = 1 - \frac{2(1 - C_j)}{(1 - C_j) + (1 - C_i)\frac{C_j}{C_i} + \sqrt{\left(1 - \frac{C_j}{C_i}\right)^2 + 4(1 - C_i)\frac{C_j}{C_i}}}.$$

Moreover, the relative loss of the portfolio with weights $\hat{\mathbf{w}}_{SH;N_i}$ is asymptotically given by

$$\begin{aligned}R_{\hat{\mathbf{w}}_{SH;N_i}} &= \mathbf{1}_p^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}_p \Bigg((\Psi^*_{N_i})^2 \hat{\mathbf{w}}^\top_{S;n_i} \boldsymbol{\Sigma} \hat{\mathbf{w}}_{S;N_i} \\ &+ 2\Psi^*_{N_i}(1 - \psi^*_{N_i}) \hat{\mathbf{w}}^\top_{S;N_i} \boldsymbol{\Sigma} \hat{\mathbf{w}}_{SH;N_{i-1}} \\ &+ (1 - \Psi^*_{N_i})^2 \hat{\mathbf{w}}^\top_{SH;N_{i-1}} \boldsymbol{\Sigma} \hat{\mathbf{w}}_{SH;N_{i-1}} \Bigg) - 1 \\ &= (\Psi^*_{N_i})^2 \frac{\mathbf{1}_p^\top \mathbf{S}^{-1}_{N_i} \boldsymbol{\Sigma} \mathbf{S}^{-1}_{N_i} \mathbf{1}_p}{\mathbf{1}_p^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}_p} \left(\frac{\mathbf{1}_p^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}^{-1}_{N_i} \mathbf{1}_p} \right)^2 \\ &+ 2\Psi^*_{N_i}(1 - \Psi^*_{N_i}) \mathbf{1}_p^\top \mathbf{S}^{-1}_{N_i} \boldsymbol{\Sigma} \hat{\mathbf{w}}_{SH;N_{i-1}} \frac{\mathbf{1}_p^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}_p}{\mathbf{1}_p^\top \mathbf{S}^{-1}_{N_i} \mathbf{1}_p} \\ &+ (1 - \Psi^*_{N_i})^2 (r_{\hat{\mathbf{w}}_{SH;N_{i-1}}} + 1) - 1 \\ &\xrightarrow{a.s.} R_i = (\Psi^*_i)^2 (1 - C_i)^{-1} + 2\Psi^*_i (1 - \Psi^*_i) K_i \\ &+ (1 - \Psi^*_i)^2 (R_{i-1} + 1) - 1 \\ &= (\Psi^*_i)^2 \frac{C_i}{1 - C_i} + (1 - \Psi^*_i)^2 R_{i-1} \\ &+ 2\Psi^*_i (1 - \Psi^*_i)(K_i - 1).\end{aligned}$$

Finally, we get

$$\Psi^*_{N_i} \xrightarrow{a.s.} \Psi^*_i, \Psi^*_i = \frac{(R_{i-1} + 1) - K_i}{(R_{i-1} + 1) + (1 - C_i)^{-1} - 2K_i},$$

for $p/N_j \to C_j \in (0,1)$ as $N_j \to \infty$, $j = 1, ..., i$. As a result, the computation of Ψ^*_i is performed in the following recursive way: (i) first, we compute $R_0 = \mathbf{1}_p^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}_p \mathbf{b}^\top \boldsymbol{\Sigma} \mathbf{b} - 1$ and $K_1 = (1 - C_1)^{-1}$ which are used to obtain Ψ^*_1; (ii) second, using R_0 and Ψ^*_1 we find R_1 and $K_1 = (1 - C_1)^2$ used in the computation of Ψ^*_2 and proceed the recursive procedure for $i = 3, ..., T$. The boundedness of R_0 ensures that all computed values are finite as well. □

REFERENCES

Ao, M., Yingying, L., and Zheng, X. (2019). Approaching mean-variance efficiency for large portfolios. *The Review of Financial Studies*, 32(7):2890–2919.

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

Bodnar, O. (2009). Sequential surveillance of the tangency portfolio weights. *International Journal of Theoretical and Applied Finance*, 12(06):797–810.

Bodnar, T., Dette, H., Parolya, N., and Thorsén, E. (2022a). Sampling distributions of optimal portfolio weights and characteristics in low and large dimensions. *Random Matrices: Theory and Applications*, 11:2250008.

Bodnar, T., Dmytriv, S., Okhrin, Y., Parolya, N., and Schmid, W. (2021a). Statistical inference for the expected utility portfolio in high dimensions. *IEEE Transactions on Signal Processing*, 69:1–14.

Bodnar, T., Dmytriv, S., Parolya, N., and Schmid, W. (2019a). Tests for the weights of the global minimum variance portfolio in a high-dimensional setting. *IEEE Transactions on Signal Processing*, 67(17):4479–4493.

Bodnar, T., Gupta, A. K., and Parolya, N. (2014). On the strong convergence of the optimal linear shrinkage estimator for large dimensional covariance matrix. *Journal of Multivariate Analysis*, 132:215–228.

Bodnar, T., Gupta, A. K., and Parolya, N. (2016). Direct shrinkage estimation of large dimensional precision matrix. *Journal of Multivariate Analysis*, 146:223–236.

Bodnar, T., Okhrin, O., and Parolya, N. (2019b). Optimal shrinkage estimator for high-dimensional mean vector. *Journal of Multivariate Analysis*, 170:63–79.

Bodnar, T., Okhrin, Y., and Parolya, N. (2022b). Optimal shrinkage-based portfolio selection in high dimensions. *Journal of Business & Economic Statistics*, (to appear).

Bodnar, T., Parolya, N., and Schmid, W. (2018). Estimation of the global minimum variance portfolio in high dimensions. *European Journal of Operational Research*, 266(1):371–390.

Bodnar, T., Parolya, N., and Thorsén, E. (2021b). *DOSPortfolio: Dynamic Optimal Shrinkage Portfolio*. R package version 0.1.0.

Bodnar, T. and Schmid, W. (2008). A test for the weights of the global minimum variance portfolio in an elliptical model. *Metrika*, 67(2):127–143.

Bodnar, T. and Schmid, W. (2009). Econometrical analysis of the sample efficient frontier. *The European Journal of Finance*, 15(3):317–335.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *The Review of Economics and Statistics*, 72(3):498–505.

Cai, T. T., Hu, J., Li, Y., and Zheng, X. (2020). High-dimensional minimum variance portfolio estimation based on high-frequency data. *Journal of Econometrics*, 214(2):482–494.

Chételat, D. and Wells, M. T. (2012). Improved multivariate normal mean estimation with unknown covariance when p is greater than n. *The Annals of Statistics*, 40:3137–3160.

Ding, Y., Li, Y., and Zheng, X. (2021). High dimensional minimum variance portfolio estimation under statistical factor models. *Journal of Econometrics*, 222(1):502–515.

Fan, J., Fan, Y., and Lv, J. (2008). High dimensional covariance matrix estimation using a factor model. *Journal of Econometrics*, 147(1):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, Statistical methodology*, 75(4):603–680.

Fan, J., Zhang, J., and Yu, K. (2012). Vast portfolio selection

with gross-exposure constraints. *Journal of the American Statistical Association*, 107(498):592–606.

Feng, Y. and Palomar, D. P. (2016). *A Signal Processing Perspective on Financial Engineering*, volume 9. now Publishers Inc.

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

Golosnoy, V., Gribisch, B., and Seifert, M. I. (2019). Exponential smoothing of realized portfolio weights. *Journal of Empirical Finance*, 53:222–237.

Golosnoy, V. and Okhrin, Y. (2007). Multivariate shrinkage for optimal portfolio weights. *The European Journal of Finance*, 13(5):441–458.

Hautsch, N., Kyj, L. M., and Malec, P. (2015). Do high-frequency data improve high-dimensional portfolio allocations? *Journal of Applied Econometrics*, 30:263–290.

Jobson, J. D. and Korkie, B. (1980). Estimation for markowitz efficient portfolios. *Journal of the American Statistical Association*, 75(371):544–554.

Kan, R. and Smith, D. R. (2008). The distribution of the sample minimum-variance frontier. *Management Science*, 54(7):1364–1380.

Kan, R., Wang, X., and Zheng, X. (2019). In-sample and out-of-sample sharpe ratios of multi-factor asset pricing models. *Available at SSRN 3454628*.

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

Ledoit, O. and Wolf, M. (2012). Nonlinear shrinkage estimation of large-dimensional covariance matrices. *Annals of Statistics*, 40:1024–1060.

Ledoit, O. and Wolf, M. (2017). Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *The Review of Financial Studies*, 30(12):4349–4388.

Li, J., Stoica, P., and Wang, Z. (2004). Doubly constrained robust capon beamformer. *IEEE Transactions on Signal Processing*, 52(9):2407–2423.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

Merton, R. C. (1972). An analytic derivation of the efficient portfolio frontier. *The Journal of Financial and Quantitive Analysis*, 7(4):1851–1872.

Okhrin, Y. and Schmid, W. (2006). Distributional properties of portfolio weights. *Journal of Econometrics*, 134(1):235–256.

Rubio, F. and Mestre, X. (2011). Spectral convergence for a general class of random matrices. *Statistics & Probability Letters*, 81(5):592–602.

Rubio, F., Mestre, X., and Palomar, D. P. (2012). Performance analysis and optimal selection of large minimum variance portfolios under estimation risk. *IEEE Journal of Selected Topics in Signal Processing*, 6(4):337–350.

Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, pages 197–206, Berkeley, Calif. University of California Press.

Van Trees, H. L. (2002). *Optimum Array Processing*. Wiley.

Verdú, S. (1998). *Multiuser Detection*. Cambridge Univ. Press.

Wang, C., Pan, G., Tong, T., and Zhu, L. (2015). Shrinkage estimation of large dimensional precision matrix using random matrix theory. *Statistica Sinica*, 25:993–1008.

Wang, C., Tong, T., Cao, L., and Miao, B. (2014). Nonparametric shrinkage mean estimation for quadratic loss functions with unknown covariance matrices. *Journal of Multivariate Analysis*, 125:222 – 232.

Yang, L., McKay, M. R., and Couillet, R. (2018). High-dimensional MVDR beamforming: Optimized solutions based on spiked random matrix models. *IEEE Transactions on Signal Processing*, 66(7):1933–1947.