



# Minimum variance portfolio optimization in the spiked covariance model

Liusha Yang, Romain Couillet, Matthew R. McKay

## ► To cite this version:

Liusha Yang, Romain Couillet, Matthew R. McKay. Minimum variance portfolio optimization in the spiked covariance model. IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2015) , Dec 2015, Cancun, Mexico. 10.1109/CAMSAP.2015.7383724 . hal-01262633

**HAL Id: hal-01262633**

**<https://hal.science/hal-01262633>**

Submitted on 26 Jan 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Minimum Variance Portfolio Optimization in the Spiked Covariance Model

Liusha Yang\*, Romain Couillet†, Matthew R. McKay\*

\*Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong

†CentraleSupélec-LSS-Université ParisSud, Gif sur Yvette, France

**Abstract**—We study the design of minimum variance portfolio when asset returns follow a low rank factor model. Using results from random matrix theory, an optimal shrinkage approach for the isolated eigenvalues of the covariance matrix is developed. The proposed portfolio optimization strategy is shown to have good performance on synthetic data but not always on real data sets. This leads us to refine the data model by considering time correlation between samples. By updating the shrinkage of the isolated eigenvalues accounting for the unknown time correlation, our portfolio optimization method is shown to have improved performance and achieves lower risk values than competing methods on real financial data sets.

## I. INTRODUCTION

The theory of portfolio optimization starts with the pathbreaking work of Markowitz [1], who developed the mean-variance optimization framework which requires perfect knowledge of the mean and covariance of the (stationary) time series of asset returns. But in practice, the estimation error of the mean and covariance matrix of the asset returns degrades the performance of the Markowitz model. It is argued in [2] that estimates of the covariance matrix are more accurate than those of the expected returns. Thus, many studies concentrate on improving the performance of the global minimum variance portfolio (GMVP), which provides the lowest possible portfolio risk and involves only the covariance matrix estimate.

The most often used sample covariance matrix (SCM) constructed on historical data shows good performance only when the number of (supposedly independent) samples  $n$  is much larger than the number of assets  $N$ . However, covariance estimates for portfolio optimization commonly involve few historical observations of sometimes up to a thousand assets, which strongly degrades the risk performance of SCM.

Many works based on random matrix theory have considered covariance matrix estimation in the  $n \simeq N$  regime, to get around the problem of the scarcity of samples. One approach inspired by the Ledoit-Wolf technique [3] is the shrinkage covariance estimation, a weighted average of two matrices, such as in [4], where a shrinkage Tyler’s estimator is proposed with the shrinkage parameter calibrated to minimize the portfolio risk. Another approach is to impose some factor structure on the covariance matrix estimator, which reduces the number of parameters to be estimated, such as in [5] where,

under various loss functions, optimal shrinkages of eigenvalues are developed in a so-called spiked covariance model.

By observing the empirical spectrum in the financial data, many studies [6] find indeed that while a majority of the eigenvalues fall within a “bulk”, several eigenvalues leak out, which are referred to as spikes. The real information is argued to be contained in these spikes, while the “bulk” essentially contains noise. To exploit the spiked structure of the financial data, we assume the population covariance matrix  $\mathbf{C}_N$  of stock returns follows the aforementioned spiked covariance model, in which all eigenvalues are equal to one but for a few outstanding ones.

Unlike recent works on the spiked covariance model [5], which developed the covariance estimators under loss functions such as the minimum expected Frobenius norm loss  $\|\hat{\mathbf{C}}_N - \mathbf{C}_N\|_F^2$  or expected Frobenius norm loss of precision matrices  $\|\hat{\mathbf{C}}_N^{-1} - \mathbf{C}_N^{-1}\|_F^2$ , this paper designs a novel covariance estimation technique in the spiked model for GMVP optimization. We first analyze the convergence of the realized risk in the double limit regime, where  $N, n \rightarrow \infty$ , with  $c_N = N/n \rightarrow c \in (0, 1)$ . Then we derive the optimal shrinkage of the spikes to minimize the portfolio risk in the limit. Finally, we construct our portfolio selection by the estimators of the optimized spikes.

The performance of the proposed strategy is tested both on synthetic and real data. While the strategy achieves superior performance over all the competing methods in the synthetic data simulation, it does not perform the best when the time correlation is not considered in the real data test. By accounting for time correlation and updating the shrinkage of the spikes, we update our method. The resulting performance gets improved and our method achieves lower risk values than the other spiked covariance estimation approaches.

## II. GMVP IN THE SPIKED COVARIANCE MODEL

### A. Data model and problem formulation

We consider a time series  $\mathbf{x}_t$ ,  $t = 1, \dots, n$  comprising logarithmic returns of  $N$  financial assets. It is modeled as  $\mathbf{x}_t \stackrel{i.i.d.}{\sim} N(\boldsymbol{\mu}, \mathbf{C}_N)$  where  $\boldsymbol{\mu}$  is the mean vector and the matrix  $\mathbf{C}_N$  is assumed to follow a spiked covariance model, that is,

$$\mathbf{C}_N = \mathbf{I}_N + t_1 \mathbf{v}_1 \mathbf{v}_1^T + t_2 \mathbf{v}_2 \mathbf{v}_2^T + \dots + t_r \mathbf{v}_r \mathbf{v}_r^T.$$

In this model, the population eigenvalues are given by  $(t_1 + 1, \dots, t_r + 1, 1, \dots, 1)$  where the number of “spikes”  $r$  and their

Yang and McKay’s work is supported by the Hong Kong Research Grants Council under grant number 16206914.

Couillet’s work is supported by the ERC MORE EC-120133.

amplitudes  $t_1 \geq \dots \geq t_r \geq 0$  are fixed independently of  $N$  and  $n$ . The vectors  $\mathbf{v}_1, \dots, \mathbf{v}_r$  are the top  $r$  eigenvectors.

The sample covariance matrix is defined as

$$\mathbf{S}_N = \frac{1}{n} \sum_{t=1}^n \tilde{\mathbf{x}}_t \tilde{\mathbf{x}}_t^T = \sum_{i=1}^N \lambda_i \mathbf{u}_i \mathbf{u}_i^T$$

with  $\tilde{\mathbf{x}}_t = \mathbf{x}_t - \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i$ , the eigenvalues  $\lambda_1 \geq \dots \geq \lambda_N$  and  $\mathbf{u}_i$  the eigenvector associated to  $\lambda_i$ .

Let  $\mathbf{h} \in \mathbb{R}^N$  denote the portfolio selection, i.e., the vector of asset holdings in units of currency normalized by the total outstanding wealth, such that  $\mathbf{h}^T \mathbf{1}_N = 1$ . Then the portfolio variance (or risk) over the investment period of interest is given by  $\sigma^2(\mathbf{h}) = E[|\mathbf{h}^T \mathbf{x}_t|^2] = \mathbf{h}^T \mathbf{C}_N \mathbf{h}$ . Accordingly, the GMVP selection problem can be formulated as the following quadratic optimization problem with a linear constraint:

$$\min_{\mathbf{h}} \quad \sigma^2(\mathbf{h}) \quad \text{s.t.} \quad \mathbf{h}^T \mathbf{1}_N = 1. \quad (1)$$

The solution to (1) is  $\mathbf{h}_{\text{GMVP}} = \frac{\mathbf{C}_N^{-1} \mathbf{1}_N}{\mathbf{1}_N^T \mathbf{C}_N^{-1} \mathbf{1}_N}$ , and the corresponding portfolio risk is

$$\sigma^2(\mathbf{h}_{\text{GMVP}}) = \frac{1}{\mathbf{1}_N^T \mathbf{C}_N^{-1} \mathbf{1}_N}. \quad (2)$$

Here, (2) represents the theoretical minimum portfolio risk bound, attained upon knowing the true precision matrix  $\mathbf{C}_N^{-1}$  exactly. In practice,  $\mathbf{C}_N^{-1}$  is unknown, and instead we form an estimate, denoted by  $\hat{\mathbf{C}}_N^{-1}$ . Thus, the GMVP selection based on  $\hat{\mathbf{C}}_N^{-1}$  is given by

$$\hat{\mathbf{h}}_{\text{GMVP}} = \frac{\hat{\mathbf{C}}_N^{-1} \mathbf{1}_N}{\mathbf{1}_N^T \hat{\mathbf{C}}_N^{-1} \mathbf{1}_N}.$$

Similar to  $\mathbf{h}_{\text{GMVP}}$ , the quality of  $\hat{\mathbf{h}}_{\text{GMVP}}$ , implemented based on the in-sample precision matrix prediction  $\hat{\mathbf{C}}_N^{-1}$ , can be measured by its achieved out-of-sample or realized portfolio risk as

$$\sigma^2(\hat{\mathbf{h}}_{\text{GMVP}}) = \frac{\mathbf{1}_N^T \hat{\mathbf{C}}_N^{-1} \mathbf{C}_N \hat{\mathbf{C}}_N^{-1} \mathbf{1}_N}{(\mathbf{1}_N^T \hat{\mathbf{C}}_N^{-1} \mathbf{1}_N)^2}. \quad (3)$$

The goal is to construct a good estimator  $\hat{\mathbf{C}}_N^{-1}$ , and consequently  $\hat{\mathbf{h}}_{\text{GMVP}}$ , which minimizes this quantity.

### B. Optimal spike shrinkage under the GMVP framework

We consider the precision matrix estimator  $\hat{\mathbf{C}}_N^{-1}$  to be of the following structure:

$$\hat{\mathbf{C}}_N^{-1} = \mathbf{I}_N + w_1 \mathbf{u}_1 \mathbf{u}_1^T + w_2 \mathbf{u}_2 \mathbf{u}_2^T + \dots + w_r \mathbf{u}_r \mathbf{u}_r^T \quad (4)$$

where  $\mathbf{u}_1, \dots, \mathbf{u}_r$  are the top  $r$  sample eigenvectors of  $\mathbf{S}_N$  and  $w_1, \dots, w_r \in \mathbb{R}$  are the shrinkage parameters to be optimized over.

By plugging the spiked structures of  $\mathbf{C}_N$  and  $\hat{\mathbf{C}}_N^{-1}$  in (3), the realized risk is  $\sigma^2(\mathbf{w})$ , a function of  $\mathbf{w}$ , where  $\mathbf{w} = [w_1, \dots, w_r]^T$ . Our goal is to find the optimal  $\mathbf{w}$  that minimizes  $\sigma^2(\mathbf{w})$ . However, this cannot be done directly since it involves  $\mathbf{C}_N$ , which is unobservable. Also note that the naïve approach of simply replacing  $\mathbf{C}_N^{-1}$  with  $\hat{\mathbf{C}}_N^{-1}$  in (3)

would yield the so-called ‘‘in-sample risk’’, which is known to underestimate the realized portfolio risk, leading to overly-optimistic investment decisions.

To tackle this problem, we first derive a deterministic asymptotic approximation of (3) as  $N, n \rightarrow \infty$  and obtain the optimal non-random  $\mathbf{w}^*$  depending on  $\mathbf{C}_N$ . Then we provide a consistent estimator  $\hat{\mathbf{w}}^*$  based on sample eigenvalues and the corresponding eigenvectors, which constructs the optimal precison matrix  $\hat{\mathbf{C}}_{\text{risk}}^{-1}$  and the portfolio selection  $\hat{\mathbf{h}}^*$ .

1) *Asymptotic deterministic equivalence  $\bar{\sigma}^2(\mathbf{w})$  of the realized portfolio risk  $\sigma^2(\mathbf{w})$* : For our asymptotic analysis, we assume the following:

#### Assumption 1.

- a. As  $N, n \rightarrow \infty$ ,  $N/n = c_N \rightarrow c \in (0, 1)$ .
- b.  $t_1 \geq \dots \geq t_r > \sqrt{c}$ .

Observe that  $\sigma^2(\mathbf{w})$  consists of quadratic terms, such as  $\frac{1}{N} \mathbf{1}_N^T \mathbf{u}_i \mathbf{u}_i^T \mathbf{1}_N$ ,  $\frac{1}{N} \mathbf{1}_N^T \mathbf{u}_i \mathbf{u}_i^T \mathbf{v}_i \mathbf{v}_i^T \mathbf{1}_N$  and  $\frac{1}{N} \mathbf{1}_N^T \mathbf{u}_i \mathbf{u}_i^T \mathbf{v}_i \mathbf{v}_i^T \mathbf{u}_i \mathbf{u}_i^T \mathbf{1}_N$ ,  $i = 1, \dots, r$ . Their asymptotic equivalents when  $N, n \rightarrow \infty$  can be derived based on the following lemma.

**Lemma 1.** [7, Proof of Theorem 9.5] Let Assumption 1 hold. Then, for  $\mathbf{a}_N, \mathbf{b}_N \in \mathbb{R}^N$  two vectors of bounded spectral norm,<sup>1</sup>

$$\mathbf{a}_N^T \mathbf{u}_i \mathbf{u}_i^T \mathbf{b}_N - s_i \mathbf{a}_N^T \mathbf{v}_i \mathbf{v}_i^T \mathbf{b}_N \xrightarrow{\text{a.s.}} 0$$

where  $s_i = \frac{1 - c/(t_i)^2}{1 + c/t_i}$ ,  $i = 1, \dots, r$ .

In addition, under Assumption 1,  $|\mathbf{u}_i^T \mathbf{v}_j| \xrightarrow{\text{a.s.}} 0$  when  $i \neq j$ . Applying Lemma 1, we obtain the deterministic equivalent of the risk  $\sigma^2(\mathbf{w})$  in Theorem 1,

**Theorem 1.** Let Assumption 1 hold. Denoting  $k_i = \frac{1}{N} \mathbf{1}_N^T \mathbf{v}_i \mathbf{v}_i^T \mathbf{1}_N$ ,  $i = 1, \dots, r$ , we have,

$$|N\sigma^2(\mathbf{w}) - N\bar{\sigma}^2(\mathbf{w})| \xrightarrow{\text{a.s.}} 0$$

where  $\bar{\sigma}^2(\mathbf{w}) = p/q$ , with

$$p = 1 + \sum_{i=1}^r (t_i s_i^2 k_i + s_i k_i) w_i^2 + 2 \sum_{i=1}^r (s_i k_i + t_i s_i k_i) w_i,$$

$$q = \left( \sum_{i=1}^r s_i k_i w_i + 1 \right)^2.$$

2) *Optimal  $\mathbf{w}^*$  that minimizes  $\bar{\sigma}^2(\mathbf{w})$* : The optimal  $\mathbf{w}^*$  is provided in the following proposition. It is a non-random function of  $t_i$ ,  $s_i$ , and  $k_i$ ,  $i = 1, \dots, r$ .

**Proposition 1.** Under the settings of Theorem 1,

$$\mathbf{w}^* \triangleq \arg \min_{\mathbf{w}} \bar{\sigma}^2(\mathbf{w})$$

is given by  $w_j^* = \frac{b_j f}{a_j d} - \frac{t_j + 1}{t_j s_j + 1}$ ,  $j = 1, \dots, r$ , with

$$a_j = t_j s_j^2 k_j + s_j k_j, \quad b_j = s_j k_j, \quad d = 1 - \sum_{i=1}^r \frac{t_i s_i k_i + s_i k_i}{t_i s_i + 1},$$

<sup>1</sup>  $\xrightarrow{\text{a.s.}}$  denotes almost sure convergence.

$$f = 1 + \sum_{i=1}^r t_i k_i - \sum_{i=1}^r \frac{(s_i k_i + t_i s_i k_i)^2}{t_i s_i^2 k_i + s_i k_i}.$$

3) *Consistent estimator  $\hat{\mathbf{w}}^*$  of  $\mathbf{w}^*$* : Since  $\mathbf{w}^*$  depends on the unobservable  $\mathbf{C}_N$ , for practical purposes, we need to estimate all necessary values to construct an estimator of it. Random matrix results on the spiked model provide us with the estimators of  $t_i$ ,  $s_i$ , and  $k_i$ , which in turn help to construct a consistent estimator  $\hat{\mathbf{w}}^*$  of the optimal deterministic  $\mathbf{w}^*$ .

**Proposition 2.** *Under the settings of Theorem 1, for  $j = 1, \dots, r$ ,*

$$|\hat{w}_j^* - w_j^*| \xrightarrow{\text{a.s.}} 0$$

where  $\hat{w}_j^*$  is obtained by plugging  $(\hat{t}_i, \hat{s}_i, \hat{k}_i)$ , the respective estimators of  $(t_i, s_i, k_i)$ ,  $i = 1, \dots, r$ , in the expression of  $w_j^*$ , with

$$\hat{t}_i = \frac{\lambda_i + 1 - c + \sqrt{(\lambda_i + 1 - c)^2 - 4\lambda_i}}{2} - 1,$$

$$\hat{s}_i = \frac{1 - c/(\hat{t}_i)^2}{1 + c/\hat{t}_i}, \quad \hat{k}_i = \frac{1}{\hat{s}_i} \frac{1}{N} \mathbf{1}_N^T \mathbf{u}_i \mathbf{u}_i^T \mathbf{1}_N.$$

4) *Optimized  $\hat{\mathbf{C}}_{\text{risk}}^{-1}$  and  $\hat{\mathbf{h}}^*$* : The optimized  $\hat{\mathbf{C}}_{\text{risk}}^{-1}$  under the GMVP framework is now constructed by plugging  $\hat{\mathbf{w}}^*$  in (4). We finally obtain  $\hat{\mathbf{h}}^*$  as  $\frac{\hat{\mathbf{C}}_{\text{risk}}^{-1} \mathbf{1}_N}{\mathbf{1}_N^T \hat{\mathbf{C}}_{\text{risk}}^{-1} \mathbf{1}_N}$ .

### C. Simulations

We use both synthetic data and real market data to show the performance of  $\hat{\mathbf{C}}_{\text{risk}}^{-1}$  compared to other competing methods. The benchmarks are  $\hat{\mathbf{C}}_{\text{clip}}$ , an eigenvalue clipping method [6], which keeps the spiked eigenvalues (and corresponding eigenvectors) while replacing the rest with a scaled identity matrix, and two estimators  $\hat{\mathbf{C}}_{\text{Fro}}$  and  $\hat{\mathbf{C}}_{\text{Froinv}}$  from [5] designed to minimize the expected Frobenius norm error  $\|\hat{\mathbf{C}}_N - \mathbf{C}_N\|_F^2$  and the precision matrix Frobenius norm error  $\|\hat{\mathbf{C}}_N^{-1} - \mathbf{C}_N^{-1}\|_F^2$  respectively. These estimators are all constructed assuming a spiked model.

In the synthetic data simulation, the data are generated i.i.d. by a normal distribution  $N(\boldsymbol{\mu}, \mathbf{C}_N)$ , where  $\mathbf{C}_N = \mathbf{I}_N + 14\mathbf{v}_1 \mathbf{v}_1^T + 9\mathbf{v}_2 \mathbf{v}_2^T + 4\mathbf{v}_3 \mathbf{v}_3^T$ ,  $\mathbf{v}_1 = \sqrt{3/N}[\mathbf{1}_{N/3}; \mathbf{0}_{2N/3}]$ ,  $\mathbf{v}_2 = \sqrt{3/N}[\mathbf{1}_{N/3}; \mathbf{0}_{N/3}; \mathbf{1}_{N/3}]$  and  $\mathbf{v}_3 = \sqrt{3/N}[\mathbf{0}_{2N/3}; \mathbf{1}_{N/3}]$ . The mean vector  $\boldsymbol{\mu}$  can be set arbitrarily since it is subtracted with the empirical mean, having no impact on the covariance estimates.

Fig. 1 shows the performance of the different estimation approaches in terms of the realized risk, averaged over 200 Monte Carlo simulations. As expected from our theoretical development,  $\hat{\mathbf{C}}_{\text{risk}}^{-1}$  outperforms other methods for all large  $n$  (with  $c_N = 1/2$ ).

We now investigate the out-of-sample portfolio performance of the different estimators with the real market data. We select 95 stocks comprising the S&P100. In particular, we use the dividend-adjusted daily closing prices downloaded from the Yahoo Finance database to obtain the continuously compounded (logarithmic) returns for the 95 constituents of the S&P100 over  $L = 1005$  working days, from Jan. 3, 2011

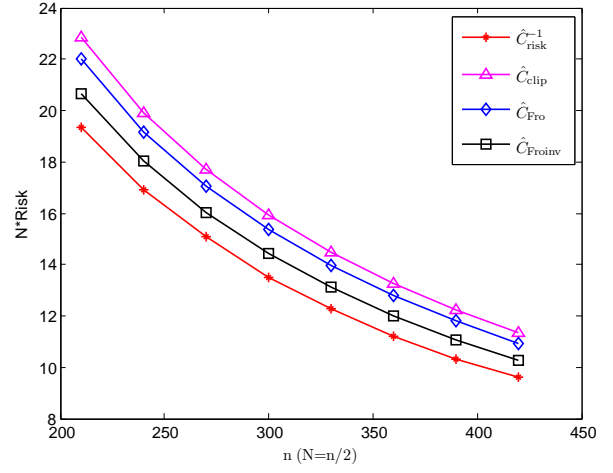


Fig. 1. The average realized portfolio risks of different methods in the GMVP framework using synthetic data.

to Dec. 31, 2014 (excluding weekends and public holidays). The out-of-sample evaluation is defined in terms of a rolling window method [4] and the performances of the methods are tested under different assumptions of the number of spikes  $r$ .

Fig. 2 shows that, unlike the synthetic data test,  $\hat{\mathbf{C}}_{\text{risk}}^{-1}$  does not perform the best. There are several factors that may degrade the performance of our proposed approach  $\hat{\mathbf{C}}_{\text{risk}}^{-1}$ . Firstly, the time series of asset returns are known to have time correlation which is not considered in our data model. Secondly, the distribution of the data may not conform with normal distribution (the data may be heavy-tailed distributed and outlier-contaminated). Thirdly, there may exist large estimation errors in the computation of the shrinkage parameters, which needs to be more finely taken care of. To improve the performance of our method, we here take into consideration the time correlation in our data model. The details are given in the next section.

## III. EXTENSION TO TIME CORRELATED DATA

### A. Data model

We consider the  $N \times n$  asset returns matrix to be  $\mathbf{Y}_N = \mathbf{X}_N \mathbf{T}^{1/2}$ , where  $\mathbf{X}_N = [\mathbf{x}_1, \dots, \mathbf{x}_n]$  and  $\mathbf{x}_i$ ,  $i = 1, \dots, n$  are based on the data model described in Section II-A. The matrix  $\mathbf{T}$  is Hermitian nonnegative and is required to follow the mild Assumption 3 in [8].<sup>2</sup> The new data model is still a spiked random matrix model, which has been less investigated in the literature, but in e.g. [8]. In this case, the parameter estimation of the spikes is still possible and allows us to obtain the following result.

**Proposition 3.** *Denote  $\lambda_1 \geq \dots \geq \lambda_N$  the sample eigenvalues of the SCM of  $\mathbf{Y}_N$  and  $\mathbf{u}_i$ ,  $i = 1, \dots, r$  the top  $r$  sample eigenvectors. Define  $\hat{m}(x) = \frac{1}{N-r} \sum_{j=r+1}^N \frac{1}{\lambda_j - x}$ ,*

<sup>2</sup>This assumption ensures that, when  $\mathbf{C}_N = \mathbf{I}_N$ , i.e.,  $r = 0$ , the spectrum of  $\mathbf{S}_N$  does not contain spikes induced by  $\mathbf{T}$  itself.

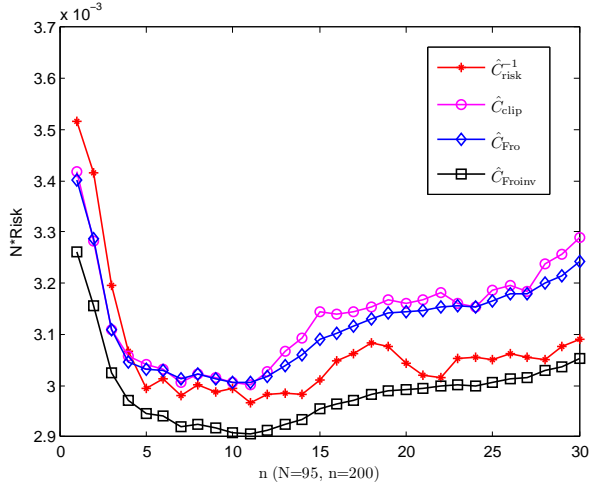


Fig. 2. Realized portfolio risks achieved out-of-sample over 1005 days of S&P100 real market data (from 2011 to 2014) under different number of assumed spikes.

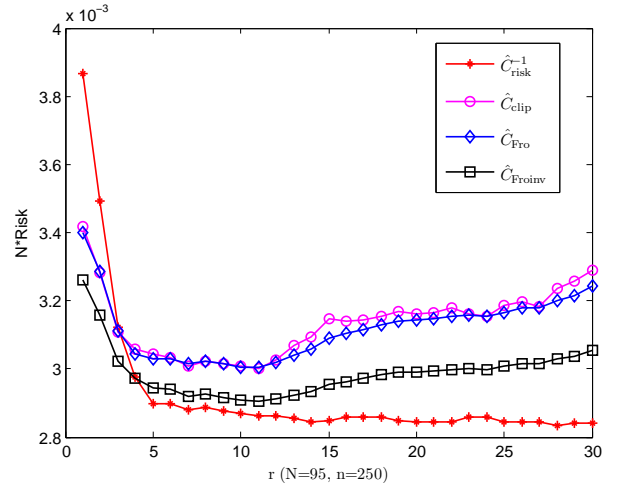


Fig. 3. Realized portfolio risks achieved out-of-sample over 1005 days of S&P100 real market data (from 2011 to 2014) under different number of assumed spikes (time correlation considered).

$\hat{g}(x) = \hat{m}(x)(x\hat{c}\hat{m}(x) + c - 1)$ . Under the (updated) settings of Theorem 1, for  $j = 1, \dots, r$ ,

$$|\hat{w}_j^* - w_j^*| \xrightarrow{\text{a.s.}} 0$$

where  $\hat{w}_j^*$  is obtained by plugging  $(\hat{t}_i, \hat{s}_i, \hat{k}_i)$ , the respective estimators of  $(t_i, s_i, k_i)$ ,  $i = 1, \dots, r$ , in the expression of  $w_j^*$  with

$$\begin{aligned} \hat{t}_i &= (\hat{g}(\lambda_i) \frac{1}{n} \text{tr}[\frac{1}{N} \mathbf{Y}_N^T \mathbf{Y}_N])^{-1}, \quad \hat{s}_i = \frac{\hat{m}(\lambda_i) \hat{g}(\lambda_i)}{\hat{g}'(\lambda_i)}, \\ \hat{k}_i &= \frac{1}{\hat{s}_i} \frac{1}{N} \mathbf{1}_N^T \mathbf{u}_i \mathbf{u}_i^T \mathbf{1}_N. \end{aligned}$$

Note fundamentally that the only required information about the unknown  $\mathbf{T}$  is in the linear eigenvalue statistics  $\hat{m}(x)$  (thus  $\mathbf{T}$  itself needs not be estimated). By Proposition 3, we then construct  $\hat{\mathbf{C}}_{\text{risk}}^{-1}$  with the updated  $\hat{\mathbf{w}}^*$ , and obtain the updated optimized minimum risk portfolio selection  $\hat{\mathbf{h}}^*$ .

#### B. Simulations

Under the same setting of Section II-C, we test the performance of the different spiked covariance estimators or precision matrix estimator with the same S&P100 data set. Fig. 3 shows that our proposed method  $\hat{\mathbf{C}}_{\text{risk}}^{-1}$  has a large improvement, achieving the lowest risk when the supposed number of spike is  $r \geq 5$ .

#### IV. CONCLUSIONS

Exploiting a factor covariance structure for the financial asset returns, we have proposed a novel minimum-variance portfolio optimization strategy based on a random matrix spiked covariance estimation framework. In the strategy, the “spikes” are calibrated to minimize the realized portfolio risk. In addition, taking into consideration the time correlation in the data improves the performance of our proposed portfolio

optimization strategy. It has been demonstrated that our approach outperforms more standard techniques in terms of the realized portfolio risk for synthetic data. For real historical stock returns from S&P100, by considering time correlation in the data model, our strategy with the updated precision matrix estimator  $\hat{\mathbf{C}}_{\text{risk}}^{-1}$  also shows best performance compared with other existing spiked covariance estimators.

It should be noted that our derivation technique is flexible and can be adopted other metrics, such as Sharpe ratio maximization.

The way to determine the number of spikes (assumed here known) in the real data is however an open question still under investigation. It would be also of interest to combine robust estimation methods with the present spiked model, making the portfolio both adapted to the factor model and robust to outliers and impulsiveness of the data. These considerations are left to future work.

#### REFERENCES

- [1] H. Markowitz, “Portfolio selection,” *J. Finance*, vol. 7, no. 1, pp. 77–91, Mar. 1952.
- [2] R. C. Merton, “On estimating the expected return on the market: An exploratory investigation,” *J. Financ. Econ.*, vol. 8, no. 4, pp. 323–361, Dec. 1980.
- [3] O. Ledoit and M. Wolf, “A well-conditioned estimator for large-dimensional covariance matrices,” *J. Multivar. Anal.*, vol. 88, no. 2, pp. 365–411, Feb. 2004.
- [4] L. Yang, R. Couillet, and M. R. McKay, “A robust approach to minimum variance portfolio optimization,” *arXiv preprint arXiv:1503.08013*, 2015.
- [5] D. L. Donoho, M. Gavish, and I. M. Johnstone, “Optimal shrinkage of eigenvalues in the spiked covariance model,” *arXiv preprint arXiv:1311.0851*, 2013.
- [6] L. Laloux, P. Cizeau, M. Potters, and J.-P. Bouchaud, “Random matrix theory and financial correlations,” *Int. J. Theor. Appl. Finance*, vol. 3, no. 03, pp. 391–397, 2000.
- [7] R. Couillet, M. Debbah *et al.*, *Random matrix methods for wireless communications*. Cambridge University Press Cambridge, MA, 2011.
- [8] J. Vinogradova, R. Couillet, and W. Hachem, “Statistical inference in large antenna arrays under unknown noise pattern,” *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5633–5645, 2013.