



# An Iterative Approach to Ill-Conditioned Optimal Portfolio Selection

Mårten Gulliksson<sup>1</sup> · Stepan Mazur<sup>2</sup> 

Accepted: 29 October 2019 / Published online: 18 November 2019  
© The Author(s) 2019

## Abstract

Covariance matrix of the asset returns plays an important role in the portfolio selection. A number of papers is focused on the case when the covariance matrix is positive definite. In this paper, we consider portfolio selection with a singular covariance matrix. We describe an iterative method based on a second order damped dynamical systems that solves the linear rank-deficient problem approximately. Since the solution is not unique, we suggest one numerical solution that can be chosen from the iterates that balances the size of portfolio and the risk. The numerical study confirms that the method has good convergence properties and gives a solution as good as or better than the solutions that are based on constrained least norm Moore–Penrose, Lasso, and naive equal-weighted approaches. Finally, we complement our result with an empirical study where we analyze a portfolio with actual returns listed in S&P 500 index.

**Keywords** Mean–variance portfolio · Singular covariance matrix · Linear ill-posed problems · Second order damped dynamical systems

## 1 Introduction

Modern portfolio theory has drawn much attention in the academic literature starting from 1952 when Harry Max Markowitz published his seminal paper about portfolio selection (see Markowitz 1952). He proposed efficient way of portfolio allocation that guarantees the lowest risk for a given level of the expected return.

A number of papers are devoted to questions like, e.g., how can an optimal portfolio be constructed, monitored, and/or estimated by using historical data (see, e.g., Alexander and Baptista 2004; Golosnoy and Okhrin 2009; Bodnar 2009; Bodnar et al.

---

✉ Stepan Mazur  
Stepan.Mazur@oru.se

<sup>1</sup> School of Science and Technology, Örebro University, 70182 Örebro, Sweden

<sup>2</sup> School of Business, Örebro University, 70182 Örebro, Sweden

2017a; Bauder et al. 2018), what is the influence of parameter uncertainty on the portfolio performance (cf., Okhrin and Schmid 2006; Bodnar and Schmid 2008; Javed et al. 2017), how do the asset returns influence the portfolio choice (see, e.g., Jondeau and Rockinger 2006; Mencia and Sentana 2009; Adcock 2010; Harvey et al. 2010; Amenguala and Sentana 2010), how is it possible to estimate the characteristics of the distribution of the asset returns (see, e.g., Jorion 1986; Wang 2005; Frahm and Memmel 2010), how can the structure of optimal portfolio be statistically justified (Gibbons et al. 1989; Britten-Jones 1999; Bodnar and Schmid 2009). Björk et al. (2014) studied the mean–variance portfolio optimization in continuous time, whereas Liesiö and Salo (2012) developed a portfolio selection framework which uses the set inclusion to capture incomplete information about scenario probabilities and utility functions. Chiarawongse et al. (2012) formulated a mean–variance portfolio selection problem that accommodates qualitative input about expected returns and provided an algorithm that solves the problem, while Levy and Levy (2014) analyzed the parameter estimation error in portfolio optimization.

There is another strand of research which focuses on factor models with applications to finance. In particular, factor models are commonly used in asset pricing theory and classical examples are the Capital Asset Pricing Model (CAPM) and the Arbitrage Pricing Theory (APT) (see Ross 1976; Sun et al. 2019 and references therein). Practically speaking, factor models can be used in the portfolio construction, the evaluation of performance of portfolio managers, the predicting asset returns, etc. (Meucci 2005; Chincarini and Kim 2006). Different estimation techniques for the covariance and precision matrices which are based on factor models in small and large dimensions with applications in portfolio theory are proposed by Ledoit and Wolf (2003), Fan et al. (2008), Fan et al. (2012), Fan et al. (2013), Bodnar and Reiss (2016), De Nard et al. (2019) and among others. There is also the well known Barra Risk Factor Analysis which is pioneered by Bar Rosenberg, founder of Barra Inc. (see Grinold and Kahn 2000; Christopherson et al. 2009; Connor and Korajczyk 2010). In its core the model involves a number of factors that can be utilized to predict and control risk.<sup>1</sup>

All above discussed papers are focused on the case when the covariance matrix of the asset returns is positive definite. In practice, covariance matrix is unknown, therefore, it needs to be estimated using historical data. Most common estimators are sample and maximum likelihood estimators. If the number of observations is greater than the number of assets in the portfolio then both estimators of the covariance matrix are positive definite. However, if the number of observations is less than the number of assets in the portfolio then both estimators of the covariance matrix are singular.<sup>2</sup> In this paper, we focus on the second case that leads us to the singular estimators of the covariance matrix. In practice, one could face this situation because of different

<sup>1</sup> Let us note that Barra Inc. delivered different practical methods that can also control risk via additional constraints and without using factor models (Barra 2007).

<sup>2</sup> Under the assumption of normally distributed data, the sample estimator of the covariance matrix has a (singular) Wishart distribution and its distributional properties are well studied by Muirhead (1982), Srivastava (2003), Bodnar and Okhrin (2008) and among others. Moreover, a product of a (singular) Wishart, or a (singular) inverse Wishart, matrix with a (singular) Gaussian vector characterizes the sample estimator of the tangency portfolio weights. Distributional properties of these products in different settings are analyzed by Bodnar and Okhrin (2011), Bodnar et al. (2013, 2014), Kotsiuba and Mazur (2015), Bodnar et al. (2018a), Bodnar et al. (2019b).

reasons which are discussed by Bodnar et al. (2016, 2017b) and Bodnar et al. (2019c). For example, it is very common to consider the sample size that is shorter over time period to avoid varying dependence between assets. Additionally, one can have a large number of assets in the portfolio: the S&P 500 index consists of 500 companies that are traded on the NYSE and NASDAQ. It is also possible to face the multicollinearity because of high correlation of assets within a specific industry branch.

Since optimal portfolio weights depend on the inverse of the covariance matrix, the original optimization problem formulated by Markowitz (1952) will have an infinite number of solutions when the covariance matrix is singular. In Pappas et al. (2010), the solution to the optimization problem with singular covariance matrix is obtained by replacing the inverse with the Moore–Penrose inverse. It leads us to the unique solution with the minimal Euclidean norm. Statistical properties of the optimal portfolio weights for small sample and singular covariance matrix are well studied by Bodnar et al. (2016, 2017b) and Bodnar et al. (2019c). There are also various regularization methods that can be used when the covariance matrix is singular. The most common techniques are the ridge-type approach, the Landweber–Fridman algorithm, the spectral cut-off method, and the Lasso-type approach. The ridge-type approach is constructed by adding a diagonal matrix to the covariance matrix (Tikhonov and Arsenin 1977), while the Landweber–Fridman algorithm delivers a convergent iterative scheme (Kress 1999). The spectral cut-off method is based on the removing the eigenvectors associated with the smallest eigenvalues (Chernousova and Golubev 2014), and the Lasso-type approach penalizes the  $l_1$  norm of the optimal portfolio weights (Brodie et al. 2009).

The main aim of the present paper is to deliver an alternative approach. In particular, we employ an iterative method that solves the linear ill-posed problem approximately. Iterative methods for linear ill-posed problems are certainly not new and include classical methods like Landweber iteration with Nesterov acceleration and Conjugate gradient methods, see Neubauer (2000, 2017). Very recently a new approach has been developed based on second order damped dynamical systems, see Gulliksson et al. (2019) for an introduction and overview and specifically Zhang and Hofmann (2018) for the ill-posed linear case. In Zhang and Hofmann (2018) it is shown that the method is a regularization method, i.e., loosely speaking for an ill-posed linear problem there exists a unique solution when the number of iterations tend to infinity and the error tends to zero. We have applied and extended this method to the rank-deficient and linearly constrained portfolio selection problem considered here. Specifically, we show that the method is convergent and how to choose optimal parameters (time step and damping). As seen in Sects. 3.1 and 3.2 the iterative method generally performs better in the sense of giving a smaller variance of the portfolio.

The rest of the paper is structured as follows. In Sect. 2, an iterative approach to ill-conditioned optimal portfolio selection is discussed. The results of numerical and empirical studies are discussed in details in Sect. 3, while Sect. 4 summarizes the paper.

## 2 Main Results

We consider a portfolio of  $k$  assets and let  $\mathbf{x}_i = (x_{1i}, \dots, x_{ki})^T$  be the  $k$ -dimensional vector of log-returns of these assets at time  $i = 1, \dots, N$ . We assume that the second moment of  $\mathbf{x}_i$  is finite. Let the mean vector of the asset returns be denoted by  $\boldsymbol{\mu}$  and the covariance matrix by  $\boldsymbol{\Sigma}$  such that  $\text{rank}(\boldsymbol{\Sigma}) = r \leq k$ , i.e.  $\boldsymbol{\Sigma}$  can also be singular matrix. Let  $\mathbf{w} = (w_1, \dots, w_k)^T$  be the vector of portfolio weights, where  $w_j$  denotes the weight of the  $j$ th asset, and let  $\mathbf{1}$  be the vector of ones while  $\mathbf{I}$  be the identity matrix.

The classical problem of portfolio selection is defined as

$$\min_{\mathbf{w}} \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w} \quad \text{s.t.} \quad \mathbf{w}^T \mathbf{1} = 1, \quad \mathbf{w}^T \boldsymbol{\mu} = q \tag{1}$$

where  $q$  is the expected rate of return that is required on the portfolio. In general, we allow for short sales and, consequently, for negative weights.<sup>3</sup>

If  $\boldsymbol{\Sigma}$  is a positive definite matrix then the optimization problem (1) has unique solution which is given by

$$\mathbf{w} = \frac{C - qB}{AC - B^2} \boldsymbol{\Sigma}^{-1} \mathbf{1} + \frac{qA - B}{AC - B^2} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \tag{2}$$

where  $A = \mathbf{1}^T \boldsymbol{\Sigma}^{-1} \mathbf{1}$ ,  $B = \mathbf{1}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$ ,  $C = \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$ . However, if  $\boldsymbol{\Sigma}$  is singular then the optimization problem (1) has an infinite number of solutions. Pappas et al. (2010) suggested the solution that appears to be unique with the minimal Euclidean norm and is obtained by replacing the inverse with the Moore–Penrose inverse

$$\mathbf{w} = \frac{C - qB}{AC - B^2} \boldsymbol{\Sigma}^+ \mathbf{1} + \frac{qA - B}{AC - B^2} \boldsymbol{\Sigma}^+ \boldsymbol{\mu} \tag{3}$$

with  $A = \mathbf{1}^T \boldsymbol{\Sigma}^+ \mathbf{1}$ ,  $B = \mathbf{1}^T \boldsymbol{\Sigma}^+ \boldsymbol{\mu}$ ,  $C = \boldsymbol{\mu}^T \boldsymbol{\Sigma}^+ \boldsymbol{\mu}$ .

In practice, both  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$  are unknown parameters and the investor cannot determine  $\mathbf{w}$ . Consequently, she/he should estimate  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$  using previous observations. The most common estimators of  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$  are given by

$$\hat{\boldsymbol{\mu}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i = \frac{1}{N} \mathbf{X} \mathbf{1}, \quad \text{and} \quad \hat{\boldsymbol{\Sigma}} = \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_i - \hat{\boldsymbol{\mu}})(\mathbf{x}_i - \hat{\boldsymbol{\mu}})^T = \frac{1}{N-1} \mathbf{X} \mathbf{V} \mathbf{X}^T, \tag{4}$$

where  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N)$ , and  $\mathbf{V} = \mathbf{I} - \frac{1}{N} \mathbf{1} \mathbf{1}^T$  is a symmetric and idempotent matrix, i.e.,  $\mathbf{V} = \mathbf{V}^T$  and  $\mathbf{V}^2 = \mathbf{V}$ . If  $N \geq k$  then sample covariance matrix  $\hat{\boldsymbol{\Sigma}}$  is positive definite, but  $\hat{\boldsymbol{\Sigma}}$  is singular when  $N < k$ . Hence, we can get the case when portfolio size is larger the sample size and, therefore, sample covariance matrix will be singular.

<sup>3</sup> The assumption of short selling might be not fulfilled on some capital markets. On the other side, the assumption of short selling is a common assumption in the literature (see, for example, Britten-Jones 1999; Kan and Smith 2008; Bodnar et al. 2018b, 2019a) and is a practically reliable assumption.

Then one can use the solution with smallest Euclidean norm that is defined in (3). Alternatively, one can use an iterative approach that is discussed in the next section.

### 2.1 The Discrete Functional Particle Method

In this section we give a brief description of the Discrete Functional Particle Method (DFPM) and describe how DFPM can be used to solve (1). For a more comprehensive discussion of DFPM see Gulliksson et al. (2019).

Let us first consider the unconstrained minimization problem

$$\min_{\mathbf{u} \in \mathcal{R}^n} V(\mathbf{u}), \tag{5}$$

where  $V : \mathcal{R}^n \rightarrow \mathcal{R}$  is at least a twice continuously differentiable convex function giving a unique solution  $\mathbf{u}^*$ . We use the conventional notation for inner-product in  $\mathcal{R}^n$  and norm as  $(\cdot, \cdot)$  and  $\|\cdot\|$ , respectively.

The main idea for solving (5) is to utilize the fact that the solution  $\mathbf{u}^*$  to (5) is also a stationary solution to the second order damped dynamical system

$$\ddot{\mathbf{u}}(t) + \eta \dot{\mathbf{u}}(t) = -\nabla V(\mathbf{u}(t)), \quad \eta > 0, \tag{6}$$

and this solution is unique and globally exponentially stable, see, e.g., references in Bégout et al. (2015). The dynamical system (6) is most efficiently used using a symplectic method such as, e.g., symplectic Runge–Kutta or Störmer–Verlet Hairer et al. (2006) applied on a reformulation of (6) as the first order system

$$\begin{aligned} \dot{\mathbf{u}} &= \mathbf{v} \\ \dot{\mathbf{v}} &= -\eta \mathbf{v} - \nabla V(\mathbf{u}). \end{aligned} \tag{7}$$

The approach of finding the solution to (5) by solving (6) with a symplectic method is DFPM. It is the combination of the damped dynamical system together with an efficient (fast, stable, accurate) symplectic solver that makes DFPM a very competitive method.

Let us now return to our main optimization problem (1) in order to apply DFPM. The constraints in (1) can be treated in mainly three different ways. Firstly, a projected symplectic method can be used enabling the iterates to stay on the constraint manifold. Secondly, additional damped equations can be formulated such that the constraints are asymptotically satisfied. Thirdly, since the constraints are linear they can be eliminated and an unconstrained problem can be solved by DFPM. Here, we choose the third alternative and leave the other two for future research.

Define

$$\mathbf{B} = \begin{bmatrix} \mathbf{1}^T \\ \boldsymbol{\mu}^T \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 1 \\ q \end{bmatrix},$$

then the constraints can be written as  $\mathbf{B}\mathbf{w} = \mathbf{c}$  and the solution is  $\mathbf{w} = \mathbf{Z}\mathbf{u} + \mathbf{g}$  where  $\mathbf{Z}$  spans the null space of  $\mathbf{B}$  and  $\mathbf{g}$  is any solution to  $\mathbf{B}\mathbf{w} = \mathbf{c}$ . After some algebra the minimization problem (1) can be written as

$$\min_{\mathbf{u}} \mathbf{u}^T \mathbf{Z}^T \Sigma \mathbf{Z} \mathbf{u} + 2\mathbf{g}^T \Sigma \mathbf{Z} \mathbf{u} + \mathbf{g}^T \Sigma \mathbf{g} := \Phi(\mathbf{u}). \tag{8}$$

The solution will be uniquely defined if  $\mathbf{M} = \mathbf{Z}^T \Sigma \mathbf{Z}$  is invertible. In the following we always choose  $\mathbf{g} = \mathbf{B}^+ \mathbf{c}$  where  $\mathbf{B}^+$  is the Moore–Penrose inverse of  $\mathbf{B}$ , i.e.,  $\mathbf{g}$  is the least norm solution to the constraint equations.

Consider again the unconstrained problem (8). If we define  $\mathbf{M} = \mathbf{Z}^T \Sigma \mathbf{Z}$ ,  $\mathbf{d} = \mathbf{Z}^T \Sigma \mathbf{g}$  we get

$$\min_{\mathbf{u} \in \mathcal{R}^{n-2}} \frac{1}{2} \mathbf{u}^T \mathbf{M} \mathbf{u} - \mathbf{u}^T \mathbf{d}, \quad \mathbf{M} \in \mathcal{R}^{(n-2) \times (n-2)}, \quad \mathbf{d} \in \mathcal{R}^{n-2}, \tag{9}$$

where we will assume that  $\mathbf{M}$  is positive semidefinite. In the general setting of the minimization problem (5) we can define

$$V(\mathbf{u}) = \frac{1}{2} \mathbf{u}^T \mathbf{M} \mathbf{u} - \mathbf{u}^T \mathbf{d} \tag{10}$$

and it is straightforward to formulate DFPM for  $V(\mathbf{u})$  in (10) as

$$\ddot{\mathbf{u}} + \eta \dot{\mathbf{u}} = \mathbf{d} - \mathbf{M} \mathbf{u}. \tag{11}$$

or the first order system

$$\begin{aligned} \dot{\mathbf{u}} &= \mathbf{v} \\ \dot{\mathbf{v}} &= -\eta \mathbf{v} + (\mathbf{d} - \mathbf{M} \mathbf{u}). \end{aligned} \tag{12}$$

Additionally we need initial conditions, say,  $\mathbf{u}(0) = \mathbf{u}_0$ ,  $\dot{\mathbf{u}}(0) = \mathbf{v}_0$ . Using symplectic Euler on (12) we get

$$\begin{aligned} \mathbf{v}_{k+1} &= (\mathbf{I} - \Delta t \eta) \mathbf{v}_k - \Delta t (\mathbf{d} - \mathbf{M} \mathbf{u}_k) \\ \mathbf{u}_{k+1} &= \mathbf{u}_k + \Delta t \mathbf{v}_{k+1} \end{aligned} \tag{13}$$

or, equivalently,

$$\mathbf{w}_{k+1} = \mathbf{G} \mathbf{w}_k + \mathbf{b}, \quad \mathbf{G} = \begin{bmatrix} \mathbf{I} - \Delta t^2 \mathbf{M} & \Delta t (1 - \Delta t \eta) \mathbf{I} \\ \Delta t \mathbf{I} & (1 - \Delta t \eta) \mathbf{I} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -\Delta t^2 \mathbf{d} \\ -\Delta t \mathbf{d} \end{bmatrix}, \tag{14}$$

where  $\Delta t$  is the time step,  $\eta$  the damping, and we have defined  $\mathbf{w}_k = [\mathbf{u}_k, \mathbf{v}_k]^T$ . Equation (14) defines DFPM for our problem. In the next section we derive values of  $\Delta t, \eta$  that will ensure fast convergence.

### 2.2 Convergence Analysis and Choice of Parameters

When  $\mathbf{M}$  in (9) has full rank  $\mathbf{u}(t)$  will converge to the unique solution  $\mathbf{u} = \mathbf{M}^{-1} \mathbf{d}$ . Turning to the rank-deficient case we consider the SVD of  $\mathbf{M} = \mathbf{U} \Sigma_M \mathbf{U}^T$  and transform (11) to

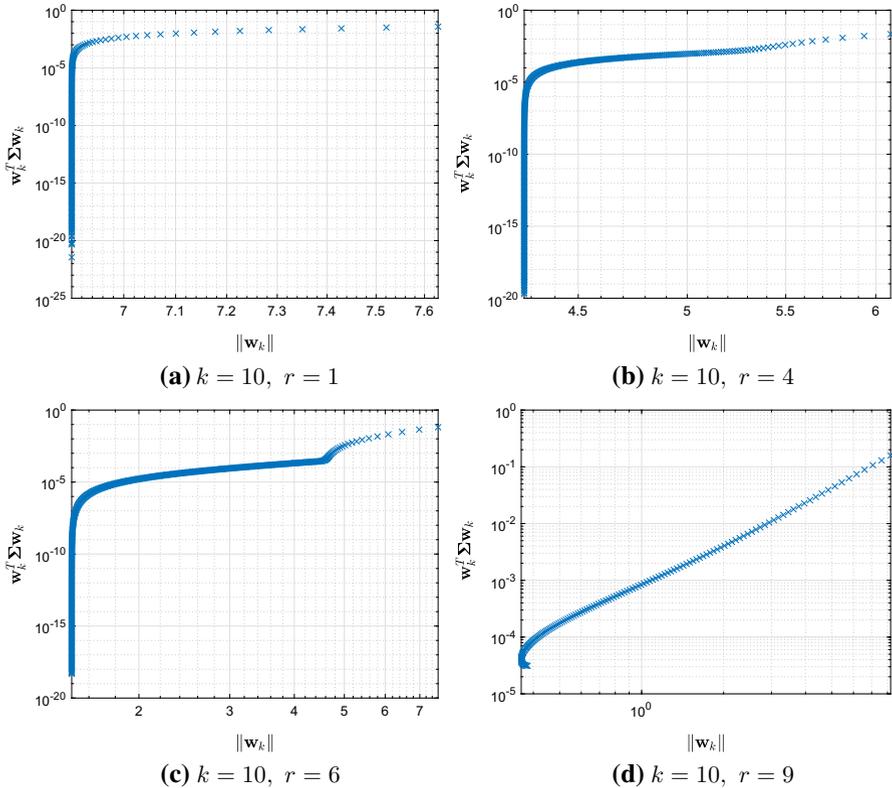


Fig. 1 Plot of the risk as a function of size of portfolio weights (norm of  $\mathbf{w}$ ) for  $k = 10$  and  $r \in \{1, 4, 6, 9\}$

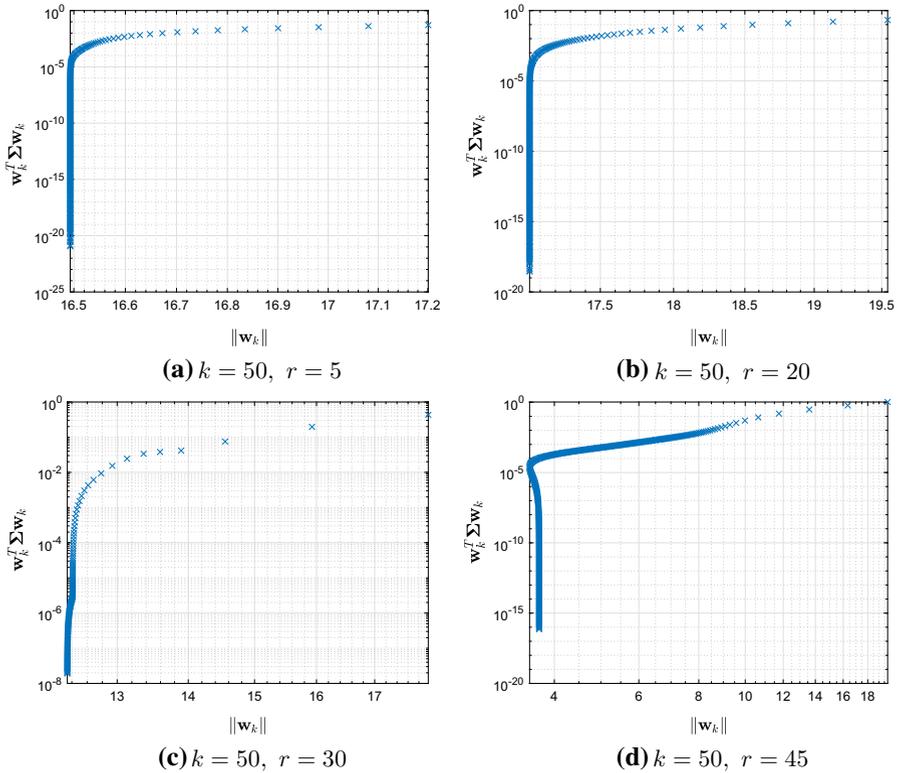
$$\ddot{\mathbf{y}} + \eta \dot{\mathbf{y}} = \mathbf{f} - \Sigma_M \mathbf{y}, \Sigma_M = \text{diag}(s_1, \dots, s_{r_M}, 0, \dots, 0), s_i > 0, \tag{15}$$

where  $r_M$  is the rank of  $\mathbf{M}$ . Since the system now is decoupled we can by partitioning  $\mathbf{f} = [\mathbf{f}_1^T, \mathbf{f}_2^T]^T$  write the solution of (15) as

$$y_i(t) = \frac{1}{s_i} f_i^1 + \alpha_i^{11} e^{-\gamma_i^1 t} + \alpha_i^{12} e^{-\gamma_i^2 t}, \quad i = 1, \dots, r_M,$$

$$y_i(t) = \alpha_i^{21} \alpha_i^{22} e^{-\eta t} + \frac{1}{\eta} f_i^2 t, \quad i = r_M + 1, \dots, n,$$

where  $\gamma_i^j = \eta/2 \pm \sqrt{\eta^2/4 - s_i}$  and  $\alpha_i^{kl}$  are given by the initial conditions. We conclude that the solution is unbounded and grows linearly with  $t$ . However, that does not mean that the dynamical system (15) can not be used for getting one, of infinitely many, solutions to (1). Indeed, by iteratively solving (15) and carefully choosing when to stop the iterations we attain a, hopefully useful, regularized solution. This is called iterative regularization in the literature of ill-posed problems and is an alternative to the least norm solution (Vogel 2002).



**Fig. 2** Plot of the risk as a function of size of portfolio weights (norm of  $\mathbf{w}$ ) for  $k = 50$  and  $r \in \{5, 20, 30, 45\}$

In order to ensure fast convergence of the iterative scheme (13) one must choose the time step and damping such that for convergence  $\|\mathbf{G}\| < 1$  and for efficiency  $\|\mathbf{G}\|$  is as small as possible. The optimal choice of parameters can be summarized in the following theorem, see Gulliksson (2017), Gulliksson et al. (2019) and references therein.

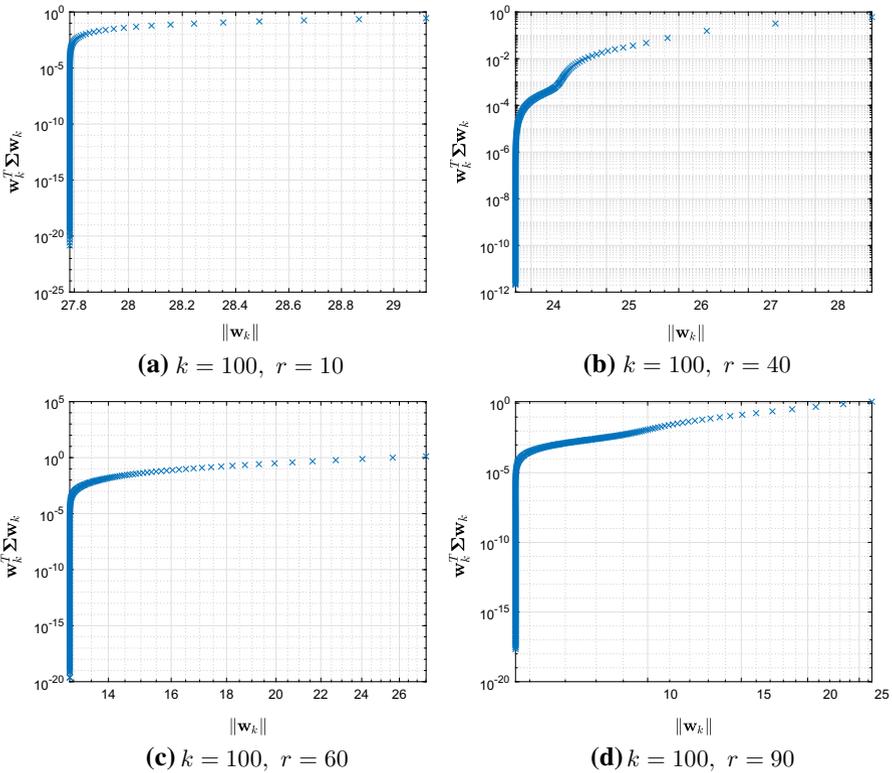
**Theorem 1** Consider symplectic Euler (13) where  $\lambda_i = \lambda_i(\mathbf{M}) > 0$  are the eigenvalues of  $\mathbf{M}$ . Then the parameters

$$\Delta t = \frac{2}{\sqrt{\lambda_{\min}} + \sqrt{\lambda_{\max}}}, \eta = \frac{2\sqrt{\lambda_{\min}}\sqrt{\lambda_{\max}}}{\sqrt{\lambda_{\min}} + \sqrt{\lambda_{\max}}}, \tag{16}$$

where  $\lambda_{\min} = \min_i \lambda_i$  and  $\lambda_{\max} = \max_i \lambda_i$  are the solution to the problem

$$\min_{\Delta t, \eta} \max_{1 \leq i \leq 2n} |\mu_i(\mathbf{G})|,$$

where  $\mu_i(\mathbf{G})$  are the eigenvalues of  $\mathbf{G}$ .



**Fig. 3** Plot of the risk as a function of size of portfolio weights (norm of  $\mathbf{w}$ ) for  $k = 100$  and  $r \in \{10, 40, 60, 90\}$

However, the result in Theorem 1 is not applicable for the case when  $\mathbf{M}$  does not have full rank since the smallest eigenvalue will be zero and the damping will in turn be zero giving an oscillating solution that does not converge to a solution of the underdetermined system. Therefore, we choose the parameters only in the subspace defined by the nonzero eigenvalues of  $\mathbf{M}$ .

**Theorem 2** Assume that  $\lambda_i(\mathbf{M}) > 0, i = 1, \dots, r_M$  are the nonzero eigenvalues of  $\mathbf{M}$  and the rest of the eigenvalues are zero. Then symplectic Euler (13) with parameters

$$\Delta t = \frac{2}{\sqrt{\lambda_{r_M}} + \sqrt{\lambda_{max}}}, \eta = \frac{2\sqrt{\lambda_{r_M}}\sqrt{\lambda_{max}}}{\sqrt{\lambda_{r_M}} + \sqrt{\lambda_{max}}}, \tag{17}$$

where  $\lambda_{r_M} = \min_{\lambda_i > 0} \lambda_i$  and  $\lambda_{max} = \max_i \lambda_i$  is convergent.

**Proof** The time step is smaller and the damping is larger than the parameters given by Theorem 1 and therefore the method is stable and convergent. □

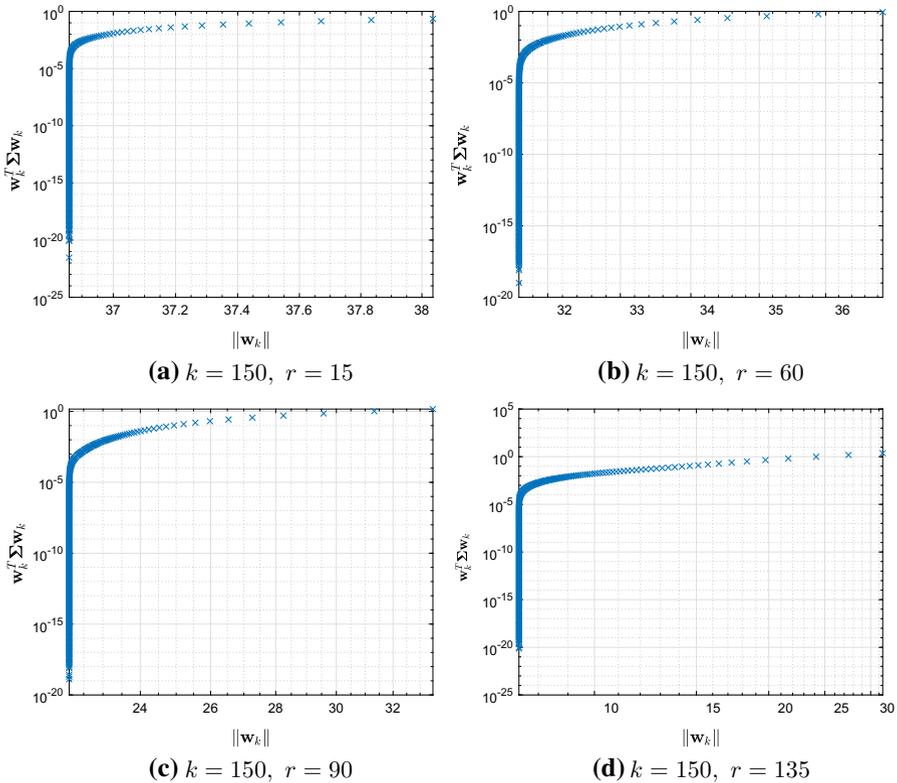


Fig. 4 Plot of the risk as a function of size of portfolio weights (norm of  $\mathbf{w}$ ) for  $k = 150$  and  $r \in \{15, 60, 90, 135\}$

### 3 Numerical and Empirical Studies

#### 3.1 Numerical Study

In this section we examine the iterative approach which is proposed by us. In particular, we evaluate optimal portfolio weights, its norm<sup>4</sup> and variance of the portfolio. All results are compared with the optimal portfolio weights which are obtained by using Moore–Penrose inverse, Lasso-type approach, and naive equal-weighted approach.

In the simulation study, each element of  $\boldsymbol{\mu}$  is uniformly distributed on the interval  $[-0.01, 0.01]$ . The population covariance matrix  $\boldsymbol{\Sigma}$  is generated as follows:

- first  $r$  non-zero eigenvalues of  $\boldsymbol{\Sigma}$  are generated from the uniform distribution on the interval  $(0, 0.01]$ , while the rest  $k - r$  eigenvalues are set to 0;
- the eigenvectors of  $\boldsymbol{\Sigma}$  are generated from the Haar distribution by generating a Wishart matrix with identity covariance and calculating its eigenvectors.

<sup>4</sup> Generally speaking, two measures such as norm and variance (or standard deviation) are equivalent. That’s why, alternatively, one can consider variance of the portfolio weights.

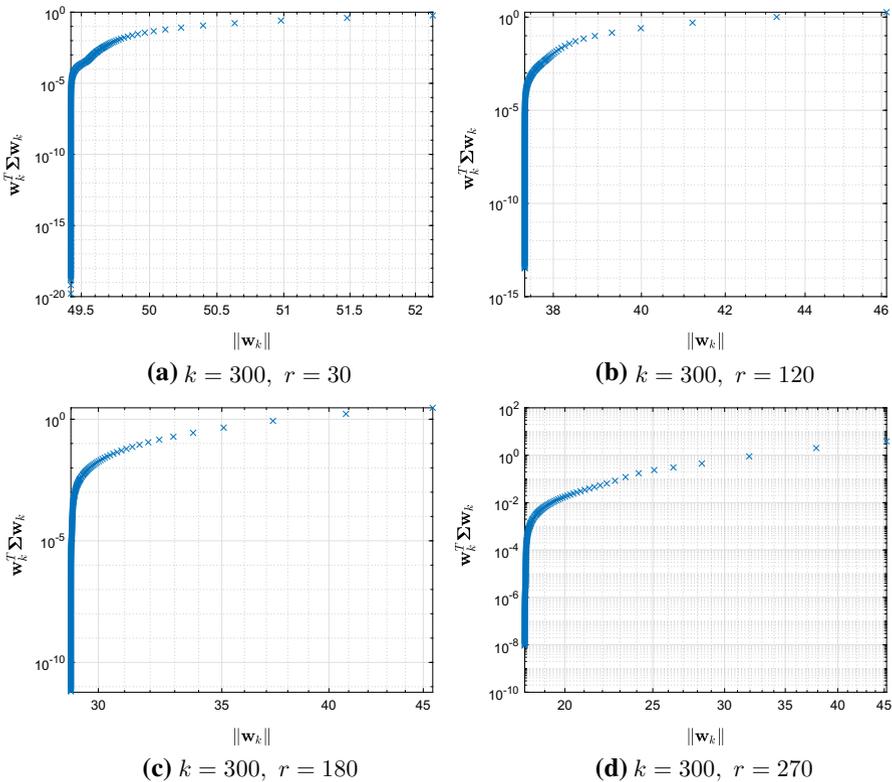


Fig. 5 Plot of the risk as a function of size of portfolio weights (norm of  $\mathbf{w}$ ) for  $k = 300$  and  $r \in \{30, 120, 180, 270\}$

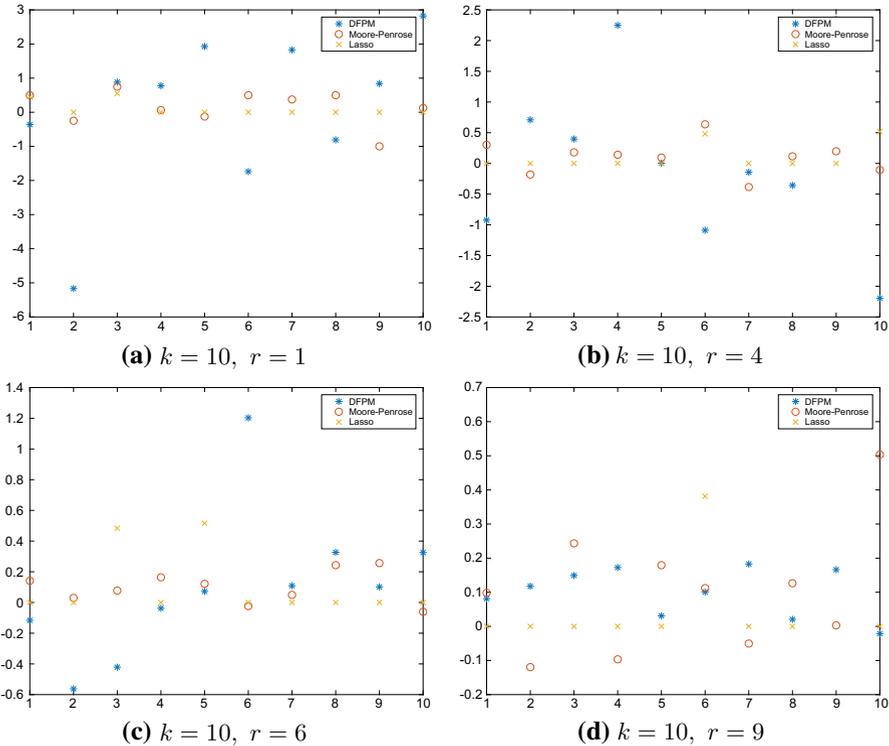
The expected rate of portfolio return  $q$  is taken to be  $\mathbf{1}^T \boldsymbol{\mu} / k$ , i.e.  $q$  is equal to expected rate of naive equal-weighted portfolio. The results are compared for several values of  $k \in \{10, 50, 100, 150, 300\}$  and  $r \in \{0.1k, 0.4k, 0.6k, 0.9k\}$ .<sup>5</sup>

The solution is not unique but one numerical solution can be chosen from the iterates that balances the size of the portfolio and the risk, see Figs. 1, 2, 3, 4 and 5. Therefore, the choice of convergence criteria is important in order to stop the iterations at a satisfactory solution. We have chosen to use a relative convergence criterion based on the projected objective function, i.e., we stop the iterations when

$$\epsilon_k < \text{tolerance}, \quad \epsilon_k = \|\nabla \Phi(\mathbf{u}_k)\| / \Phi(\mathbf{u}_k)$$

where  $\mathbf{u}_k$  is the approximation of the solution of (8) at iteration  $k$  and  $\Phi(\mathbf{u})$  is defined in (8). The tolerance is set to  $10^{-12}$  in order to get a small risk. Again referring to Figs. 1, 2, 3, 4 and 5, we note that a higher tolerance will give a higher risk and size

<sup>5</sup> In practice, if the number of observations  $N$  is less than the number of assets  $k$  in the portfolio, then  $r = N - 1$ . However, one could face the phenomenon of multicollinearity, therefore, additional test on the rank of the covariance matrix could be performed (Nadakuditi and Edelman 2008).



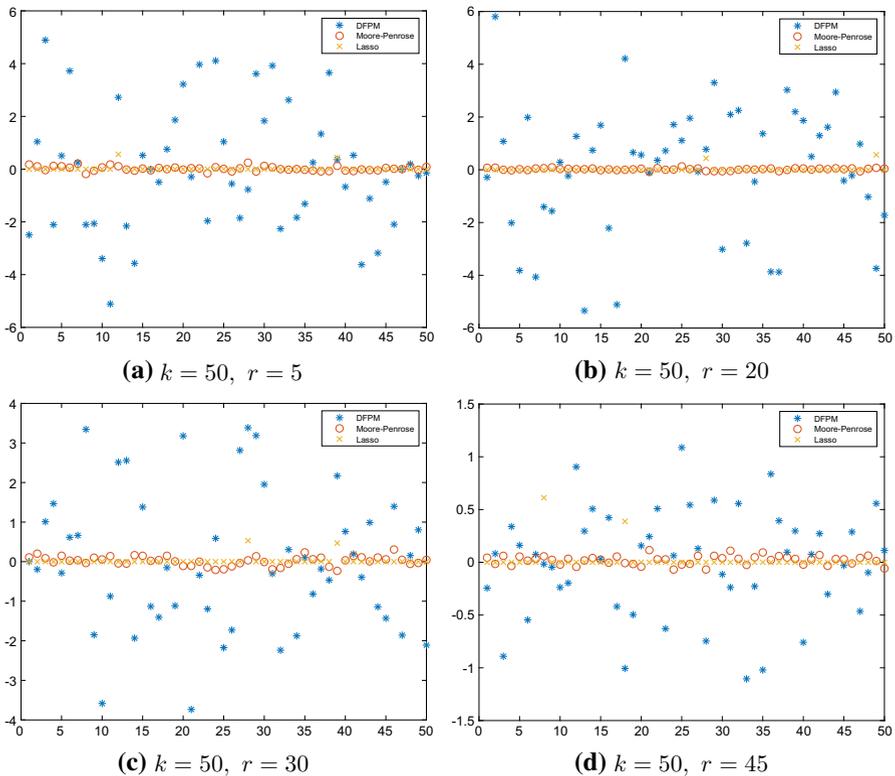
**Fig. 6** Optimal portfolio weights obtained by using DFPM approach and Moore–Penrose inverse for  $k = 10$  and  $r \in \{1, 4, 6, 9\}$

of portfolio (norm of  $\mathbf{w}$ ). Maximum number of iterations is taken to be  $10^4$  but in the presented tables this is not attained.

In Figs. 6, 7, 8, 9 and 10, we present optimal portfolio weights that are obtained by using DFPM, Moore–Penrose inverse, and Lasso approaches. We observe that behaviour of the weights is quite different, i.e., three methods suggest three very different investment strategies. In particular, for  $k \in \{50, 100, 150, 300\}$  we can observe that absolute values of almost all weights obtained by DFPM approach are much larger than the ones obtained by using Moore–Penrose inverse and Lasso methods.

In Table 1, we present the norm of the optimal portfolio weights that is obtained by using different approaches. We can observe that the DFPM method delivers the largest norm in almost all considered cases except the case when  $k = 10$  and  $r = 9$ . In this case, DFPM gives us smaller norm in comparison with Moore–Penrose inverse and Lasso approaches, while the naive equal-weighted method has larger norm than the one which is obtained by using the DFPM approach.

In Table 2, we compare the variance of portfolio for different methods. In all considered cases, we can observe that the DFPM approach shows much better performance than the Moore–Penrose inverse, Lasso, and naive equal-weighted approaches. We would note that the variance of portfolio is much smaller in comparison with vari-

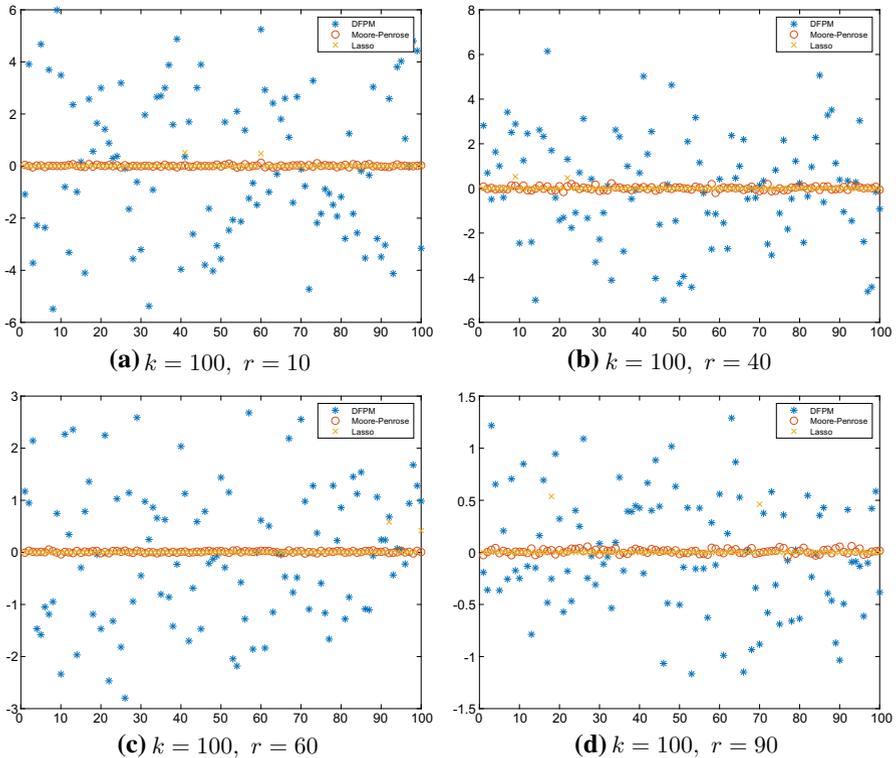


**Fig. 7** Optimal portfolio weights obtained by using DFPM approach and Moore–Penrose inverse for  $k = 50$  and  $r \in \{5, 20, 30, 45\}$

ances that are obtained by using other methods. Since the expected return of portfolio  $q$  is the same for all methods, the smallest variance leads us to the highest Sharpe ratio. Comparing Tables 1 and 2, we can conclude that the DFPM approach delivers smaller variance of portfolio, while the norm of the weights can be larger than the ones obtained by using the Moore–Penrose inverse, Lasso, and naive equal-weighted approaches. So, the smallest norm of the weights doesn’t guarantee us the smallest variance of portfolio.

### 3.2 Empirical Study

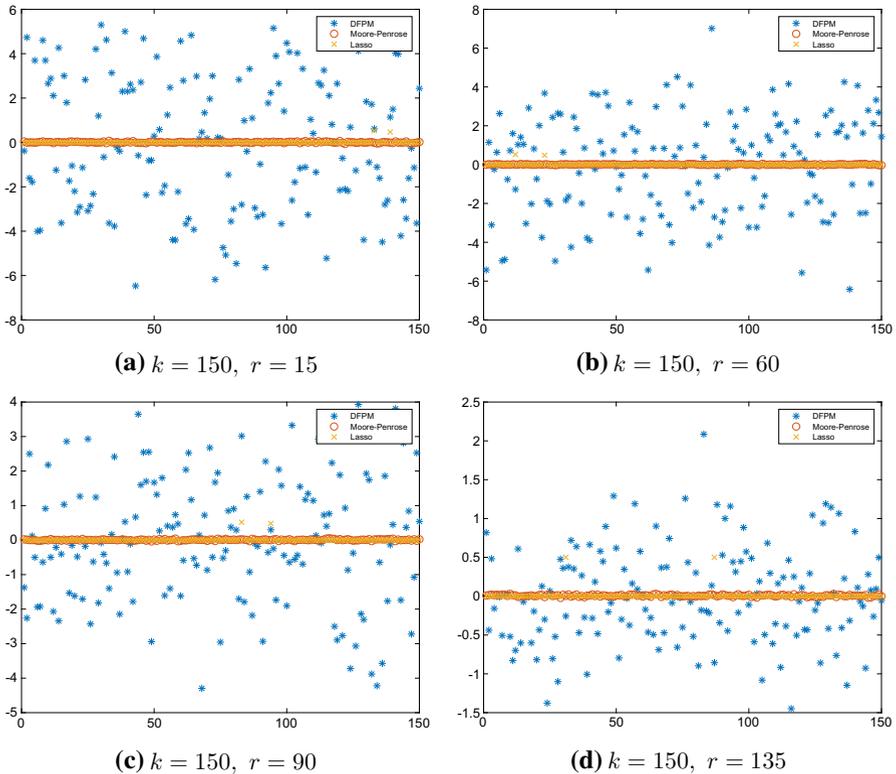
In this section the results of an empirical study are presented. It is shown how one can apply the theory from the previous sections to real data. We use weekly S&P 500 logarithmic returns of 440 stocks for the period from the 4th of May, 2007 to the 25th of January, 2013 resulting in 300 observations. Expected rate of return  $q$  is taken to be equal to 1, i.e.  $q = 1$ . In DFPM approach, the convergence criterion, the tolerance and the maximum number of iterations are taken as in Sect. 3.1.



**Fig. 8** Optimal portfolio weights obtained by using DFPM approach and Moore–Penrose inverse for  $k = 100$  and  $r \in \{10, 40, 60, 90\}$

First of all, we estimate mean vector and covariance matrix by using their empirical counterparts that are defined in (4). Since the number of stocks  $k = 440$  is greater than the sample size  $N = 300$ , sample estimator of the covariance matrix  $\Sigma$  will be singular matrix with  $\text{rank}(\Sigma) = N - 1 = 299$ , i.e.  $r = 299$ . In Fig. 11, we present eigenvalues of the sample covariance matrix. We can observe that the first 299 eigenvalues are much larger than the rest of the eigenvalues. It confirms that the rank of the sample covariance matrix is 299.

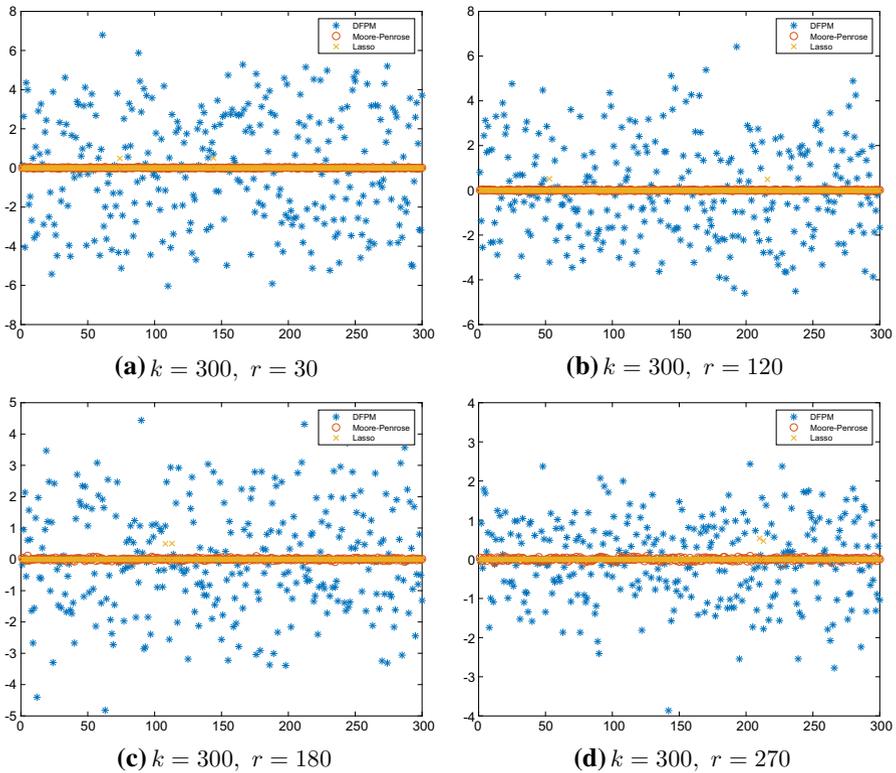
Since we have estimated both mean vector and covariance matrix, we are able to construct optimal portfolio weights by using DFPM, Moore–Penrose inverse, Lasso, and naive equal-weighted methods. In Fig. 12, we deliver the plot with optimal portfolio weights that are obtained by DFPM, Moore–Penrose inverse, and Lasso methods. We can observe quite different behaviour of the weights. Consequently, we can see that these methods suggest completely different investment strategies. One can also note that the behavior of the optimal portfolio weights looks like a white noise process. To check this observation, we should verify several conditions. Here, we focus on the portfolio weights which were obtained by DFPM and Moore–Penrose inverse approaches. First, the expected value of the weights are the same in both methods and equal to  $2.2727e-03$  that is very close to 0. Second, the variance of the weights



**Fig. 9** Optimal portfolio weights obtained by using DFPM approach and Moore–Penrose inverse for  $k = 150$  and  $r \in \{15, 60, 90, 135\}$

obtained by DFPM is  $2.0208e+01$ , while the variance obtained by Moore–Penrose inverse is  $1.8273e+01$ . So, we can observe that both variances are finite and the variance of the weights obtained by using Moore–Penrose inverse is slightly smaller than the variance obtained via DFMP. Third, white noise process is assumed to be uncorrelated. To verify this assumption, we conducted a Ljung–Box Q-test on the optimal portfolio weights. In both cases, this test indicated that there is not enough evidence to reject the null hypothesis of no autocorrelation between the weights. Fourth, we conducted an Engle’s ARCH test on the variance of the weights. This test reported that the null hypothesis of no ARCH effects cannot be rejected in both methods. Thus, we can conclude that optimal portfolio weights that are obtained by DFPM and Moore–Penrose inverse approaches can be modelled by a white noise process.

In Table 3, we present the norm of the optimal weights, variance of portfolio, and Sharpe ratio that are obtained by DFPM, Moore–Penrose inverse, Lasso, and



**Fig. 10** Optimal portfolio weights obtained by using DFPM approach and Moore–Penrose inverse for  $k = 300$  and  $r \in \{30, 120, 180, 270\}$

naive equal-weighted methods. We can observe the highest norm for the weights for the DFPM method, while the lowest norm is obtained for the naive equal-weighted strategy. However, DFPM delivers the smallest variance of the portfolio. Let us note that a quite similar behaviour is observed in the numerical study. In our comparison, we evaluate the Sharpe ratio that is used to help an investor to understand the expected rate of an investment in comparison with its risk (variance of portfolio). According to Table 3, the highest Sharpe ratio is obtained by the DFPM approach. It is remarkable that Sharpe ratio of the portfolio which is obtained via naive equal-weighted strategy is negative, i.e. nobody should invest in this portfolio.

## 4 Summary

Modern portfolio theory plays an important role in economics and suggests an investment strategy that give the lowest risk for a given level of expected return. If covariance matrix of the asset returns is positive definite, then Markowitz' optimization problem has unique solution that depends on the mean vector and covariance matrix of the asset returns. In practice, both mean vector and covariance matrix need to be esti-

**Table 1** Norm of the optimal portfolio weights obtained by using DFPM approach, Moore–Penrose inverse, Lasso-type approach, and naive equal-weighted method

	$r = 0.1k$	$r = 0.4k$	$r = 0.6k$	$r = 0.9k$
$k = 10$				
DFPM	6.9011e+00	4.2709e+00	1.4821e+00	3.8128e-01
Moore–Penrose	1.5971e+00	8.9592e-01	4.4760e-01	6.3991e-01
Lasso	7.1065e-01	7.0761e-01	7.0747e-01	7.2667e-01
Equal-weighted	3.1623e-01	3.1623e-01	3.1623e-01	3.1623e-01
$k = 50$				
DFPM	1.6493e+01	1.7030e+01	1.2342e+01	3.7180e+00
Moore–Penrose	6.4204e-01	3.3156e-01	8.4861e-01	3.3257e-01
Lasso	7.1222e-01	7.1332e-01	7.0844e-01	7.2467e-01
Equal-weighted	1.4142e-01	1.4142e-01	1.4142e-01	1.4142e-01
$k = 100$				
DFPM	2.7783e+01	2.3796e+01	1.2894e+01	5.6451e+00
Moore–Penrose	3.8416e-01	8.2013e-01	1.9755e-01	2.4008e-01
Lasso	7.0753e-01	7.0815e-01	7.1677e-01	7.0909e-01
Equal-weighted	1.0000e-01	1.0000e-01	1.0000e-01	1.0000e-01
$k = 150$				
DFPM	3.6859e+01	3.1612e+01	2.2132e+01	7.4075e+00
Moore–Penrose	3.1669e-01	2.3116e-01	2.2909e-01	1.5941e-01
Lasso	7.0879e-01	7.0792e-01	7.0758e-01	7.0711e-01
Equal-weighted	8.1650e-02	8.1650e-02	8.1650e-02	8.1650e-02
$k = 300$				
DFPM	4.9424e+01	3.7373e+01	2.8989e+01	1.8057e+01
Moore–Penrose	3.1410e-01	3.2062e-01	5.0869e-01	5.8331e-01
Lasso	7.0722e-01	7.0736e-01	7.0713e-01	7.0874e-01
Equal-weighted	5.7735e-02	5.7735e-02	5.7735e-02	5.7735e-02

The comparison is done for  $k \in \{10, 50, 100, 150, 300\}$  and  $r \in \{0.1k, 0.4k, 0.6k, 0.9k\}$

ated. Hence, one can get an estimator of the covariance matrix which is singular and, therefore, optimization problem will not have a unique solution. In our paper, we delivered a new iterative approach (DFPM) that solves the constrained and rank-deficient portfolio selection problem approximately. The method is based on using symplectic solvers for a damped dynamical system that solves the optimization problem and the solution is generally different from the least norm Moore–Penrose, Lasso, and naive equal-weighted solutions. We showed how to determine the optimal time step and damping for symplectic Euler that give fast convergence using only matrix–vector multiplications in each iteration step. In the numerical study we examined DFPM and compared it with solutions that are based on the Moore–Penrose inverse, Lasso, and naive equal-weighted methods. The results are compared for several values of the portfolio size and the rank of the covariance matrix. We observed that iterative and analytical approaches deliver quite different investment strategies. We also found that

**Table 2** Variance of portfolio return obtained by using DFPM approach, Moore–Penrose inverse, Lasso-type approach, and naive equal-weighted method

	$r = 0.1k$	$r = 0.4k$	$r = 0.6k$	$r = 0.9k$
$k = 10$				
DFPM	9.1854e−19	4.6420e−18	1.1801e−18	6.2346e−05
Moore–Penrose	2.7349e−03	4.8452e−04	5.7134e−04	1.0260e−03
Lasso	4.9935e−04	1.8901e−03	2.6884e−03	2.3373e−03
Equal-weighted	1.8953e−05	4.9418e−04	2.4524e−04	2.0013e−04
$k = 50$				
DFPM	2.2432e−18	2.7409e−17	3.7412e−08	1.4133e−16
Moore–Penrose	5.4083e−04	2.9295e−04	3.1993e−06	6.3889e−05
Lasso	3.0026e−04	1.0575e−03	2.1326e−03	2.4936e−03
Equal-weighted	7.7960e−06	2.7411e−05	8.2120e−05	1.1212e−04
$k = 100$				
DFPM	2.8447e−17	4.0131e−12	2.2203e−17	1.3486e−17
Moore–Penrose	6.6776e−04	2.8512e−05	7.1123e−05	1.1062e−05
Lasso	6.2506e−04	9.8181e−04	1.4044e−03	2.3616e−03
Equal-weighted	6.6163e−06	1.8072e−05	3.6201e−05	4.7236e−05
$k = 150$				
DFPM	4.9508e−18	1.0207e−16	7.0666e−17	1.1852e−17
Moore–Penrose	2.0664e−04	3.8681e−05	2.8403e−05	2.0384e−05
Lasso	4.0393e−04	1.3523e−03	1.9754e−03	2.3862e−03
Equal-weighted	4.1166e−06	1.7105e−05	2.1001e−05	2.9336e−05
$k = 300$				
DFPM	1.0500e−17	6.0207e−14	1.1880e−11	1.7420e−08
Moore–Penrose	1.5827e−04	1.2143e−05	7.9741e−06	3.3316e−06
Lasso	3.1657e−04	7.8162e−04	1.5192e−03	2.4490e−03
Equal-weighted	1.3036e−06	5.3795e−06	9.6586e−06	1.4842e−05

The comparison is done for  $k \in \{10, 50, 100, 150, 300\}$  and  $r \in \{0.1k, 0.4k, 0.6k, 0.9k\}$

the norm of the weights in DFPM approach is higher than the ones obtained from the Moore–Penrose inverse, Lasso, and naive equal-weighted methods in almost all considered cases, while the variance of portfolio is always smaller in all considered cases. In the empirical study, we analyzed weekly S&P 500 logarithmic returns of 440 stocks with 300 observations. It is shown that DFPM approach guarantees smaller variance of portfolio return than the solutions based on the Moore–Penrose inverse, Lasso, and naive equal-weighted methods. We also observed that optimal portfolio weights that are obtained by DFPM and Moore–Penrose inverse methods can be modelled by white noise process.

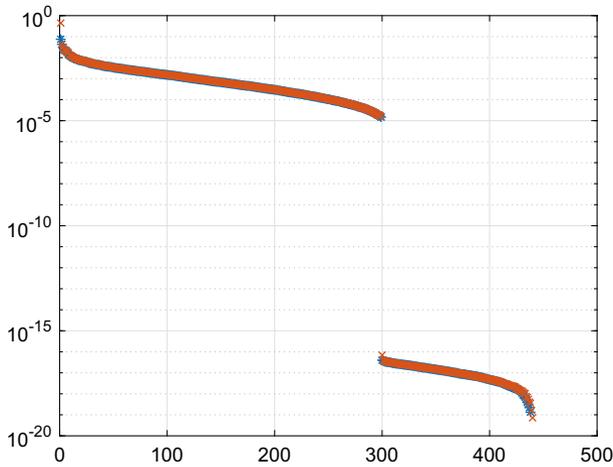


Fig. 11 Eigenvalues of the sample covariance matrix for 440 stocks and 300 observations

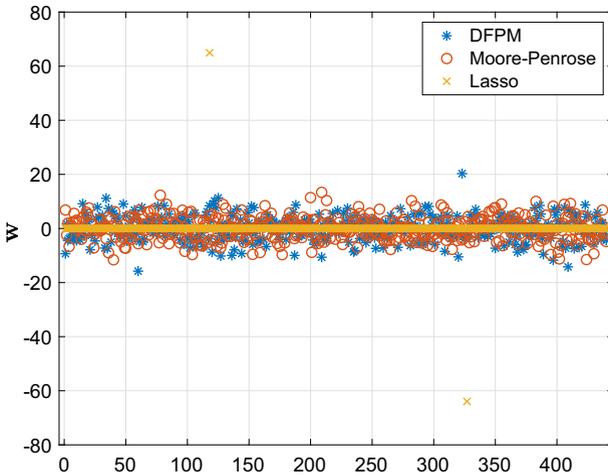


Fig. 12 Optimal portfolio weights obtained by using DFPM, Moore–Penrose inverse, and Lasso approaches, for  $k = 440$ ,  $r = 299$  and  $q = 1$

Table 3 Norm, variance and Sharpe ratio that are obtained by using DFPM, Moore–Penrose, Lasso, and naive equal-weighted methods

	DFPM	Moore–Penrose	Lasso	Equal-weighted
Norm	9.4189e+01	8.9565e+01	9.1146e+01	4.7673e−02
Variance	3.0960e−06	3.2282e−01	5.5834e+01	8.4561e−04
Sharpe ratio	5.6833e+02	1.7600e+00	1.3383e−01	−3.3433e−03

The comparison is done for  $q = 1$

**Acknowledgements** Open access funding provided by Örebro University. The authors are thankful to Prof. Hans Amman and three anonymous Reviewers for careful reading of the manuscript and for their suggestions which have improved an earlier version of this paper. Stepan Mazur acknowledges financial support from the internal research Grants at Örebro University, and from the project “Models for macro and financial economics after the financial crisis” (Dnr: P18-0201) funded by Jan Wallander and Tom Hedelius Foundation.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Adcock, C. J. (2010). Asset pricing and portfolio selection based on the multivariate extended skew-student-t distribution. *Annals of Operation Research*, *176*(1), 221–234.
- Alexander, G. J., & Baptista, M. A. (2004). A comparison of VaR and CVaR constraints on portfolio selection with the mean–variance model. *Management Science*, *50*, 1261–1273.
- Amenguala, D., & Sentana, E. (2010). A comparison of mean–variance efficiency tests. *Journal of Econometrics*, *154*, 16–34.
- Barra, M. (2007). *Barra risk model handbook*. [https://www.academia.edu/4156371/barra\\_risk\\_model\\_handbook](https://www.academia.edu/4156371/barra_risk_model_handbook).
- Bauder, D., Bodnar, T., Mazur, S., & Okhrin, Y. (2018). Bayesian inference for the tangent portfolio. *International Journal of Theoretical and Applied Finance*, *21*(8), 1–27.
- Bégout, P., Bolte, J., & Jendoubi, M. A. (2015). On damped second-order gradient systems. *Journal of Differential Equations*, *259*(7), 3115–3143.
- Björk, T., Murgoci, A., & Zhou, Y. (2014). Mean–variance portfolio optimization with state-dependent risk aversion. *Mathematical Finance*, *24*(1), 1–24.
- Bodnar, O. (2009). Sequential surveillance of the tangency portfolio weights. *International Journal of Theoretical and Applied Finance*, *12*, 797–810.
- Bodnar, T., Dmytriv, S., Parolya, N., & Schmid, W. (2019a). Tests for the weights of the global minimum variance portfolio in a high-dimensional setting. *IEEE Transactions of Signal Processing*, *67*(17), 4479–4493.
- Bodnar, T., Mazur, S., Muhinyuza, S., & Parolya, N. (2018a). On the product of a singular wishart matrix and a singular Gaussian vector in high dimension. *Theory of Probability and Mathematical Statistics*, *99*(2), 37–50.
- Bodnar, T., Mazur, S., & Okhrin, Y. (2013). On the exact and approximate distributions of the product of a wishart matrix with a normal vector. *Journal of Multivariate Analysis*, *122*, 70–81.
- Bodnar, T., Mazur, S., & Okhrin, Y. (2014). Distribution of the product of singular wishart matrix and normal vector. *Theory of Probability and Mathematical Statistics*, *91*, 1–15.
- Bodnar, T., Mazur, S., & Okhrin, Y. (2017a). Bayesian estimation of the global minimum variance portfolio. *European Journal of Operational Research*, *256*(1), 292–307.
- Bodnar, T., Mazur, S., & Parolya, N. (2019b). Central limit theorems for functionals of large dimensional sample covariance matrix and mean vector in matrix-variate skewed model. *Scandinavian Journal of Statistics*, *46*(2), 636–660.
- Bodnar, T., Mazur, S., & Podgórski, K. (2016). Singular inverse Wishart distribution and its application to portfolio theory. *Journal of Multivariate Analysis*, *143*, 314–326.
- Bodnar, T., Mazur, S., & Podgórski, K. (2017b). A test for the global minimum variance portfolio for small sample and singular covariance. *AStA Advances in Statistical Analysis*, *101*, 253–265.
- Bodnar, T., Mazur, S., Podgórski, K., & Tyrcha, J. (2019c). Tangency portfolio weights in small and large dimensions: Estimation and test theory. *Journal of Statistical Planning and Inference*, *201*, 40–57.
- Bodnar, T., & Okhrin, Y. (2008). Properties of the singular, inverse and generalized inverse partitioned Wishart distributions. *Journal of Multivariate Analysis*, *99*, 2389–2405.
- Bodnar, T., & Okhrin, Y. (2011). On the product of inverse Wishart and normal distributions with applications to discriminant analysis and portfolio theory. *Scandinavian Journal of Statistics*, *38*, 311–331.

- Bodnar, T., Parolya, N., & Schmid, W. (2018b). Estimation of the global minimum variance portfolio in high dimensions. *European Journal of Operational Research*, 266(1), 371–390.
- Bodnar, T., & Reiss, M. (2016). Exact and asymptotic tests on a factor model in low and large dimensions with applications. *Journal of Multivariate Analysis*, 150, 125–151.
- Bodnar, T., & Schmid, W. (2008). A test for the weights of the global minimum variance portfolio in an elliptical model. *Metrika*, 67, 179–201.
- Bodnar, T., & Schmid, W. (2009). Econometrical analysis of the sample efficient frontier. *The European Journal of Finance*, 15, 317–327.
- Britten-Jones, M. (1999). The sampling error in estimates of mean–variance efficient portfolio weights. *Journal of Finance*, 54, 655–671.
- Brodie, J., Daubechies, I., De Mol, C., Giannone, D., & Loris, I. (2009). Sparse and stable Markowitz portfolios. In *Proceedings of the National Academy of Sciences of the USA* (Vol. 106, pp. 12267–12272).
- Chernousova, E., & Golubev, Y. (2014). Spectral cut-off regularizations for ill-posed linear models. *Mathematical Methods of Statistics*, 23(2), 116–131.
- Chiarawongse, A., Kiatsupaibul, S., Tirapat, S., & Van Roy, B. (2012). Portfolio selection with qualitative input. *Journal of Banking and Finance*, 36, 489–496.
- Chincarini, L. B., & Kim, D. (2006). *Quantitative equity portfolio management: An active approach to portfolio construction and management*. New York: McGraw-Hill.
- Christopherson, J. A., Carino, D., & Ferson, W. E. (2009). *Portfolio performance measurement and benchmarking*. New York: McGraw-Hill Finance & Investing.
- Connor, G., & Korajczyk, R. A. (2010). Factor models in portfolio and asset pricing theory. In J. B. Guerd (Ed.), *Handbook of portfolio construction* (pp. 401–418). Boston, MA: Springer.
- De Nard, G., Ledoit, O., & Wolf, M. (2019). Factor models for portfolio selection in large dimensions: The good, the better and the ugly. *Journal of Financial Econometrics* <https://doi.org/10.1093/jffinec/nby033>.
- Fan, J., Fan, Y., & Lv, J. (2008). High dimensional covariance matrix estimation using a factor model. *Journal of Econometrics*, 147(1), 186–197.
- Fan, J., Liao, Y., & 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., & Yu, K. (2012). Vast portfolio selection with gross-exposure constraints. *Journal of the American Statistical Association*, 107(498), 592–606.
- Frahm, G., & Memmel, C. (2010). Dominating estimators for minimum–variance portfolios. *Journal of Econometrics*, 159, 289–302.
- Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica*, 57, 1121–1152.
- Golosnoy, V., & Okhrin, Y. (2009). Flexible shrinkage in portfolio selection. *Journal of Economic Dynamics and Control*, 33, 317–328.
- Grinold, R. C., & Kahn, R. (2000). *Active portfolio management: A quantitative approach for producing superior returns and controlling risk*. New York: McGraw-Hill.
- Gulliksson, M. (2017). The discrete dynamical functional particle method for solving constrained optimization problems. *Dolomites Research Notes on Approximation*, 10, 6–12.
- Gulliksson, M., Ögren, M., Oleynik, A., & Zhang, Y. (2019). *Damped dynamical systems for solving equations and optimization problems* (pp. 1–44). Cham: Springer.
- Hairer, E., Lubich, C., & Wanner, G. (2006). *Geometric numerical integration* (2nd ed.). Berlin: Springer.
- Harvey, C. R., Leichty, J. C., Leichty, M. W., & Muller, P. (2010). Portfolio selection with higher moments. *Quantitative Finance*, 10, 469–485.
- Javed, F., Mazur, S., & Ngailo, E. (2017). *Higher order moments of the estimated tangency portfolio weights*. Technical report 10, Örebro University School of Business.
- Jondeau, E., & Rockinger, M. (2006). Optimal portfolio allocation under higher moments. *European Financial Management*, 12, 29–55.
- Jorion, P. (1986). Bayes–Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, 21, 293–305.
- Kan, R., & Smith, D. R. (2008). The distribution of the sample minimum–variance frontier. *Management Science*, 54(7), 1364–1380.

- Kotsiuba, I., & Mazur, S. (2015). On the asymptotic and approximate distributions of the product of an inverse Wishart matrix and a Gaussian random vector. *Theory of Probability and Mathematical Statistics*, *93*, 95–104.
- Kress, R. (1999). *Linear integral equations*. Berlin: Springer.
- Ledito, O., & Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, *10*(5), 603–621.
- Levy, H., & Levy, M. (2014). The benefits of differential variance-based constraints in portfolio optimization. *European Journal of Operational Research*, *234*(2), 372–381.
- Liesiö, J., & Salo, A. (2012). Scenario-based portfolio selection of investment projects with incomplete probability and utility information. *European Journal of Operational Research*, *217*(1), 162–172.
- Markowitz, H. (1952). Mean–variance analysis in portfolio choice and capital markets. *Journal of Finance*, *7*, 77–91.
- Mencia, J., & Sentana, E. (2009). Multivariate location-scale mixtures of normals and mean–variance–skewness portfolio allocation. *Journal of Econometrics*, *153*, 105–121.
- Meucci, A. (2005). *Risk and asset allocation*. Berlin: Springer.
- Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*. New York: Wiley.
- Nadakuditi, R. R., & Edelman, A. (2008). Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples. *IEEE Transactions on Signal Processing*, *56*(7), 2625–2638.
- Neubauer, A. (2000). On Landweber iteration for nonlinear ill-posed problems in hilbert scales. *Numerische Mathematik*, *85*, 309–328.
- Neubauer, A. (2017). On Nesterov acceleration for Landweber iteration of linear ill-posed problems. *Journal of Inverse and Ill-Posed Problems*, *25*, 381–390.
- Okhrin, Y., & Schmid, W. (2006). Distributional properties of portfolio weights. *Journal of Econometrics*, *134*, 235–256.
- Pappas, D., Kiriakopoulos, K., & Kaimakamis, G. (2010). Optimal portfolio selection with singular covariance matrix. *International Mathematical Forum*, *5*(47), 2305–2318.
- Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, *13*(3), 341–360.
- Srivastava, M. S. (2003). Singular Wishart and multivariate beta distributions. *The Annals of Statistics*, *31*(5), 1537–1560.
- Sun, R., Ma, T., Liu, S., & Sathye, M. (2019). Improved covariance matrix estimation for portfolio risk measurement: A review. *Journal of Risk and Financial Management*, *12*(1), 48.
- Tikhonov, A., & Arsenin, V. (1977). *Solutions of ill-posed problems*. New York: Winston.
- Vogel, C. R. (2002). *Computational methods for inverse problems*. Philadelphia, PA: Society for Industrial and Applied Mathematics.
- Wang, Z. (2005). A shrinkage approach to model uncertainty and asset allocation. *Review of Financial Studies*, *18*, 673–705.
- Zhang, Y., & Hofmann, B. (2018). *On the second order asymptotical regularization of linear ill-posed inverse problems*. ArXiv e-prints.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.