IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 15, 2020, accepted October 2, 2020, date of publication October 14, 2020, date of current version October 28, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3031192

Improved Large Dynamic Covariance Matrix Estimation With Graphical Lasso and Its Application in Portfolio Selection

XIN YUAN, WEIQIN YU, ZHIXIANG YIN, AND GUOQIANG WANG, (Member, IEEE)
School of Mathematics, Physics, and Statistics, Shanghai University of Engineering Science, Shanghai 201620, China

Corresponding authors: Xin Yuan (xinyuan1615@gmail.com) and Guoqiang Wang (guoq_wang@hotmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 11971302 and Grant 62072296, and in part by the National Statistical Science Research Project of China under Grant 2020LY067 and Grant 2018LY16.

ABSTRACT The estimation of the large and high-dimensional covariance matrix and precision matrix is a fundamental problem in modern multivariate analysis. It has been widely applied in economics, finance, biology, social networks and health sciences. However, the traditional sample estimators perform poorly for large and high-dimensional data. There are many approaches to improve the covariance matrix estimation. The large dynamic conditional correlation model based on the nonlinear shrinkage and its application in portfolio selection attract increasing attention. In the estimation of the unconditional covariance matrix, the graphical lasso is more robust than the nonlinear shrinkage model, and the leptokurtic and fat tail characteristics of the asset returns are also more obvious. This article proposes improved large dynamic covariance matrix estimation based on the graphical lasso models under the multivariate normal distribution (glasso) and *t* distribution (tlasso), and the corresponding dynamic conditional correlation glasso and tlasso approaches are developed. To verify the effectiveness and robustness of the proposed methods, we conduct simulations and then apply the models to the classic Markowitz portfolio selection problem. Simulations and empirical results show that the combined dynamic conditional correlation glasso and tlasso approaches outperform the current dynamic covariance matrix estimators.

INDEX TERMS Covariance matrix estimation, dynamic conditional correlation, graphical lasso, Tlasso, Markowitz portfolio selection.

I. INTRODUCTION

As an essential input to many financial models, the covariance matrix plays a vital role in asset allocation and risk management. For example, the hedging model must estimate the covariance matrix of asset returns, and the hedging ratio must be adjusted if it changes. The prices of rainbow options and other structured products based on various basic asset designs are sensitive to the covariance matrix of the underlying asset returns [1]. The construction of a minimum variance portfolio also requires an estimate of the covariance matrix of the asset returns [2]. But financial data often has the characteristics of large dimensionality, non-normality, and high positive correlation, which bring significant challenges to the estimation of the covariance matrix [3]. Classical estimation methods are usually based on the assumption that the number of samples is much larger than the variable dimensions. Making use of these approaches to estimate the large covariance matrices, but the changes of the variance and covariance over time are not considered, and they are affected by dimensional disasters and large noise problems [1], [4]–[7]. To overcome these challenges, there are many estimation methods for the large-dimensional covariance matrices from the perspectives of cross-section and time series.

From a cross-sectional perspective, the difficulties lie in large dimensionality and non-normality. The solutions are mainly in two categories: shrinking estimation methods without prior structure information, and sparse estimation methods with prior structure information.

To be without prior structure information means that the covariance matrix to be estimated does not need to satisfy a specific prior structure. Stein [8] pointed out that the sample covariance was overfitting when the dimensionality of the variable was large. To obtain an effective

The associate editor coordinating the review of this manuscript and approving it for publication was Xiao-Jun Yang.

estimator of the covariance matrix, the eigenvectors of the sample covariance matrix should be retained, and the eigenvalues should be shrunk toward the mean of its cross-section. Based on this, Ledoit and Wolf set the target matrix as the single-index covariance matrix [9], the isocorrelation coefficient matrix [10] and the identity matrix [3], respectively. Then they proposed the corresponding linear shrinkage methods. Zhang *et al.* [11] proposed the improved linear shrinkage estimators of the covariance matrix as two types of Toeplitz-structured target matrices are employed in the shrinkage procedure. The essence of the linear shrinkage methods is to use the weighted average of the target matrix and the sample covariance matrix to solve the over-fitting phenomenon of the sample covariance. However, the linear shrinkage estimation assigns the same shrinkage density to all sample eigenvalues. Ledoit and Wolf [4] proposed the classical nonlinear shrinkage estimation method based on random matrix theory, and subsequently introduced a non-random multivariate function, the quantized eigenvalues sampling transform, to the nonlinear contraction to improve the estimation of the large-dimensional covariance matrix [12]. Simulation studies showed that this method is better than previous estimation methods and has ideal finite sample properties. Ledoit and Wolf also conducted a series of detailed studies. For example, they applied the nonlinear shrinkage method to the construction of a minimum variance portfolio [13]. An optimal nonlinear shrinkage estimation method for large-dimensional covariance matrix under Stein's loss is proposed [14]. Recently, they constructed a new estimator for large covariance matrices by drawing a bridge between the classic Stein estimator in finite samples and recent progress under large dimensional asymptotics [15].

The prior information of the sparse estimation method is that many elements of the true covariance matrix are zero or close to zero [16], [17]. The core idea of sparse estimation methods is that when the true covariance matrix is unknown, it is much easier to decide whether to estimate a certain element than to accurately estimate it [18]. Sparse estimation methods can be divided into two categories: 1) to directly estimate the sparse covariance matrix; and 2) to estimate the sparse precision matrix. The threshold method proposed by Bickel and Levina is one of the most convenient methods to estimate the sparse covariance matrix [19]. This method directly sets the smaller elements in the estimation result to zero. In view of the high positive correlation of financial data, scholars have proposed a large class of robust sparse estimation methods, that is, factor-based methods [20]–[22]. The precision matrix captures the conditional correlation between variables [23], which is closely related to the undirected graph in Gaussian graphical models (GGMs). Friedman *et al.* [24] introduced the penalty terms L_1 to the estimation of the precision matrix. The proposed glasso algorithm effectively estimates the precision matrix, and significantly improves the speed of the operation. Because the returns of financial assets often have the characteristics of fat tails, Finegold and Drton [25] replaced the multivariate normal distribution hypothesis in the glasso algorithm with the multivariate t distribution hypothesis, from which they derived the tlasso algorithm. The glasso method based on a multivariate t distribution solves the problems of large dimensionality and non-normality at the same time. Goto and Xu [26] and Torri *et al.* [27] applied the glasso approaches to the construction of minimum variance portfolios. Numerical results showed that portfolios constructed by the glasso method can significantly reduce the risk of out-of-samples compared to the nonlinear shrinkage estimation.

From a time-series perspective, financial data is susceptible to heterogeneity due to policies and financial crises. The ARCH model proposed by Engle [28] and the GARCH model of Bollerslev [29] solve the univariate heteroscedasticity problem. However, due to the influence of dimensionality, these models encounter many challenges when extended from univariate to multivariate. Many breakthrough contributions, including the BEKK model [30], the dynamic conditional correlation (DCC) model [1], and the composite likelihood estimation [31], effectively solve the problem of heteroscedasticity in multivariate situations. The most classic method is the DCC model in multivariate GARCH. Hassanein and Elgohari [32] used the DCC model to effectively analyze the linkage between stock and inter-bank bond markets in China.

Recent research tends to combine methods from two perspectives. For example, Hafner and Reznikova [33] combined the linear shrinkage estimation with the DCC model. Liu *et al.* [34] applied principal components and threshold methods to the estimation of the DCC model. It is easy to find that in these studies, the maximum cross-sectional dimension is less than 200, and the length of the time series is still greater than the number of assets, which does not reach true high dimension. Engle *et al.* [35] combined the nonlinear shrinkage estimation and the DCC model to develop the well-known DCC-NL model, which still performs well when the number of assets is greater than or equal to 1,000.

Considering that the glasso method in [26] and [27] is significantly better than the nonlinear shrinkage in estimating the unconditional covariance matrix, and that asset returns do not obey the normal distribution [5], [11], we apply the glasso and tlasso sparse estimation methods to estimate the unconditional covariance matrix $\bar{\mathbf{Q}}$ in the DCC model. We use Monte Carlo simulation to compare dynamic covariance matrix estimators, and find that the loss of the DCC-glasso and the DCC-tlasso is smaller than other combined DCC types. Empirical results show that both the DCC-glasso and the DCC-tlasso models have smaller standard deviation and higher Sharpe ratio than other combined DCC types.

The rest of this article is organized as follows. Section II introduces five well-known unconditional covariance matrix estimation methods, focusing on the glasso and tlasso approaches. Section III describes the traditional and the improved DCC models based on the glasso and tlasso methods. Section IV discusses our simulations and empirical studies. Some conclusions are drawn in section V.

The following notation is used throughout the paper. $\mathbf{r}$ represents a data matrix of dimension $N \times T$, where N is the number of assets and T is the period. $c = N/T$ is the concentration ratio. $\boldsymbol{\mu}$ is the mean vector of $\mathbf{r}$. Σ and $\widehat{\Sigma}$ are respectively the true covariance matrix and its estimated value. $\mathbf{U}$ and λ_i $(i = 1, 2, \ldots, N)$ are the eigenvector matrix and the eigenvalues of the sample covariance matrix $\mathbf{S}$, respectively. $Tr(\cdot)$ denotes the trace of a matrix. $diag(\cdot)$ represents a diagonal matrix with the elements of the vector on the main diagonal.

II. UNCONDITIONAL COVARIANCE MATRIX ESTIMATION METHODS

A. SAMPLE COVARIANCE MATRIX

The sample covariance matrix,

$$\mathbf{S} = \frac{1}{T-1}\sum_{t=1}^{T}(\mathbf{r}_t - \boldsymbol{\mu})(\mathbf{r}_t - \boldsymbol{\mu})', \tag{1}$$

is the best-known method to estimate the true covariance matrix in traditional multivariate statistical analysis [36].

However, it has some disadvantages. On the one hand, it does not consider changes in variance and covariance over time. On the other hand, it can encounter dimensional disasters. The parameters of the sample covariance matrices is $N \times (N-1)/2$, and the dimension of the data is $N \times T$. When N tends to T, the quantities of the two formulas are similar-sized. It is impossible to estimate $O(N^2)$ parameters with $O(N^2)$ data. Using the sample covariance to construct a minimum variance portfolio will also encounter the problem of large noise.

B. LINEAR SHRINKAGE MODEL

The linear shrinkage model is essentially a weighted average of the sample covariance matrix and the shrinkage target matrix. A tradeoff is made between the setting error of the model and the estimation error, limiting both to a reasonable range. Ledoit and Wolf [3] were the first to propose a shrinkage method based on the identity matrix, which is the most simple and effective. Following is a brief introduction to its basic ideas, the details can be found in [3].

Let the weights given to the identity matrix $\mathbf{I}$ and sample covariance matrix $\mathbf{S}$ be ρ_1 and ρ_2, respectively. Construct a quadratic loss function as follows

$$L(\rho_1, \rho_2) = \|\rho_1\mathbf{I} + \rho_2\mathbf{S} - \Sigma\|^2. \tag{2}$$

The optimal estimates of weights ρ_1 and ρ_2 can be obtained by solving the following optimization problem:

$$\min_{\rho_1,\rho_2} E\left(L\left(\rho_1, \rho_2\right)\right). \tag{3}$$

Let $\rho_1 = \rho v$ and $\rho_2 = 1 - \rho$. Then the optimal solutions v and ρ are given by

$$\hat{v} = \hat{\mu}, \tag{4}$$

$$\hat{\rho} = \frac{\min\left\{\frac{1}{T^2}\sum_{t=1}^{T}\left\|\mathbf{r}_t\mathbf{r}_t' - \mathbf{S}\right\|^2, \left\|\mathbf{S} - \hat{\mu}\mathbf{I}\right\|^2\right\}}{\left\|\mathbf{S} - \hat{\mu}\mathbf{I}\right\|^2}, \tag{5}$$

where $\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N}\lambda_i$. It follows that $\hat{\rho}_1 = \hat{v}\hat{\rho}$ and $\hat{\rho}_2 = 1 - \hat{\rho}$. Finally, the covariance matrix estimate obtained by the linear shrinkage of the identity matrix is

$$\widehat{\Sigma} = \hat{\rho}_1\mathbf{I} + \hat{\rho}_2\mathbf{S}. \tag{6}$$

Other target matrices include the covariance matrix obtained by the Sharpe single index model, and the isocorrelation coefficient matrix. For details, refer to [9], [10].

C. NONLINEAR SHRINKAGE MODEL

It is well-known that the linear shrinkage estimation method gives the same shrinkage density to all sample eigenvalues, while the nonlinear shrinkage estimation assigns the different shrinkage densities. Motivated by the oracle estimator, Ledoit and Wolf [4] extended their linear shrinkage model by applying a nonlinear transformation to the sample eigenvalues. Here is a brief introduction to its basic ideas, the details can be found in [4].

First of all, by shrinking the eigenvalues of the sample covariance matrix while keeping the eigenvectors unchanged, a set of estimators that are "rotationally equivalent" to the sample covariance matrix are constructed. Every rotation-equivariant estimator has the form

$$\mathbf{U}\Delta\mathbf{U}', \tag{7}$$

where $\Delta = diag(d_1, \ldots, d_n)$.

Then, the best estimation of $\widehat{\Sigma}$ can be obtained by minimizing the following loss function:

$$\min_{\Delta}\left\|\mathbf{U}\Delta\mathbf{U}' - \Sigma\right\|. \tag{8}$$

Finally, the nonlinear shrinkage estimate of the covariance matrix is given by

$$\widehat{\Sigma} = \mathbf{U}\Delta^{\mathbf{or}}\mathbf{U}', \tag{9}$$

where $\Delta^{\mathbf{or}} = diag(d_1^{or}, \ldots, d_n^{or})$, with $d_i^{or} = \frac{\lambda_i}{|1-c-c\lambda_i\breve{m}_F(\lambda_i)|^2}$ and $\breve{m}_F(\lambda_i) = \frac{1-c}{c\lambda_i} - \frac{1}{c}\frac{1}{v_{\lambda_i}}$.

The essence of the nonlinear shrinkage estimation is to enlarge the small eigenvalues and shrink the large ones, so that we can compress the range of the sample eigenvalues to the mean. For more the nonlinear shrinkage methods, refer to [12], [13].

D. GRAPHICAL LASSO

Graphical models use the nodes and the edges to describe both the conditional and unconditional dependence structures of a set of variables. Among them, GGMs are the most popular [37]. The minimum variance portfolio model relies on the assumption of normality of asset returns, so GGMs can also be used to estimate the linear dependence between assets [27]. Assume that the return on assets is a normally distributed random variable, $\mathbf{r} \sim \mathcal{N}_N(\boldsymbol{\mu}, \Sigma)$. Then an undirected graph $G(V, E)$ can be defined to be associated with $\mathbf{r}$, where the nodes in the vertex set V correspond to each asset, and the edges E are composed of correlation coefficients with nonzero bias for each group of assets.

The glasso method refers to the introduction of L_1 norms to the maximum likelihood estimation problem, so that most of the non-diagonal elements in the estimated inverse covariance matrix are zero [24]. The corresponding optimization problem is

$$\widehat{\Sigma^{-1}} = \arg\max_{\Sigma^{-1}}(\log|\Sigma^{-1}| - Tr(\Sigma^{-1}\mathbf{S}) - \rho\|(\Sigma^{-1})^{-}\|_1), \tag{10}$$

where ρ is the penalty parameter to control sparsity (larger ρ means more zero elements in the inverse covariance matrix). $(\Sigma^{-1})^{-}$ means that the diagonal elements are set to zero, and the corresponding matrix Σ^{-1} of other elements is unchanged.

The estimated inverse covariance matrix is obtained by solving (10), and the inverse operation is performed to obtain the estimated covariance matrix. For new insights and faster estimates of the glasso methods, we refer the reader to [38].

E. GRAPHICAL MODELING WITH TLASSO

The returns of financial assets often show the characteristics of leptokurtic and fat tails. Therefore, the low-freedom multivariate t-distribution assumption is better for financial models. The tlasso approach is to estimate the inverse covariance matrix sparsely on the premise that the asset returns follow a multivariate t distribution. Finegold and Drton [25] expressed the multivariate t distribution as multivariate normal distribution and gamma distribution, and used the expectation maximization (EM) algorithm to estimate the inverse covariance matrix.

Let $\mathbf{r}$ be a random vector obeying the t distribution of N elements, where the dispersion matrix is Ψ^{-1}, and there are v degrees of freedom, that is, $\mathbf{r} \sim \mathcal{T}_N(\boldsymbol{\mu}, \Psi^{-1}, v)$. The use of Ψ^{-1} to represent the dispersion matrix is to make the subsequent derivation process more convenient. Then the inverse covariance matrix can be expressed as

$$\Sigma^{-1} = \frac{v-2}{v}\Psi. \tag{11}$$

Let $W \sim \mathcal{N}_N(\mathbf{0}, \Psi^{-1})$ and $\tau \sim \Gamma(v/2, v/2)$ be random variables obeying a multivariate normal distribution and a gamma distribution, respectively. Then

$$\mathbf{r} = \boldsymbol{\mu} + \frac{W}{\sqrt{\tau}} \sim \mathcal{T}_N\left(\boldsymbol{\mu}, \Psi^{-1}, v\right). \tag{12}$$

Assuming that $r_1, \ldots, r_T$ are T samples in $\mathbf{r}$, the EM algorithm can be divided into steps E and M, as follows.

- Step E
  Calculate $\hat{\tau}_t^{(j+1)}$ of period $j+1$ according to the estimated values $\widehat{\boldsymbol{\mu}}^{(j)}$ and $\widehat{\Psi}^{(j)}$ of period j, with $t = 1, \ldots, T$, that is,

$$\hat{\tau}_t^{(j+1)} = \frac{v+N}{v + \left((r_t - \widehat{\boldsymbol{\mu}}^{(j)})'\widehat{\Psi}^{(j)}(r_t - \widehat{\boldsymbol{\mu}}^{(j)})\right)}. \tag{13}$$

- Step M
  Calculate $\widehat{\boldsymbol{\mu}}^{(j+1)}$ and $\widehat{\mathbf{S}}^{(j+1)}$ in period $j+1$ as

$$\widehat{\boldsymbol{\mu}}^{(j+1)} = \frac{\sum_{t=1}^{T} \hat{\tau}_t^{(j+1)} r_t}{\sum_{t=1}^{T} \hat{\tau}_t^{(j+1)}}, \tag{14}$$

and

$$\widehat{\mathbf{S}}^{(j+1)} = \frac{1}{T}\sum_{t=1}^{T} \hat{\tau}_t^{(j+1)} \left[r_t - \widehat{\boldsymbol{\mu}}^{(j+1)}\right]\left[r_t - \widehat{\boldsymbol{\mu}}^{(j+1)}\right]'. \tag{15}$$

Then $\widehat{\Psi}^{(j+1)}$ is obtained by solving the following optimization problem

$$\widehat{\Psi}^{(j+1)} = \arg\max_{\Psi}(\log|\Psi| - Tr(\Psi\widehat{\mathbf{S}}^{(j+1)}) - \lambda\|\Psi\|_1). \tag{16}$$

When estimating the inverse covariance matrix, the initial mean vector and dispersion matrix are obtained from the samples. Next, iterate through steps E and M in sequence until the maximum difference between $\widehat{\Psi}^{(j)}$ and $\widehat{\Psi}^{(j+1)}$ is less than a given threshold. Finally, use (11) to calculate the estimated covariance matrix.

III. IMPROVED DYNAMIC CONDITIONAL CORRELATION ESTIMATORS

To solve the conditional heteroscedasticity of the multivariate variables, Engle [1] was the first to consider the classic multivariate DCC-GARCH model,

$$\begin{cases} \mathbf{r}_t = \boldsymbol{\mu}_t + \mathbf{a}_t, \\ \mathbf{a}_t = \mathbf{H}_t^{1/2}\mathbf{z}_t, \\ \mathbf{H}_t = \mathbf{D}_t\mathbf{R}_t\mathbf{D}_t, \end{cases} \tag{17}$$

which is also-called the DCC model.

The symbols $\mathbf{r}_t$, $\mathbf{a}_t$, $\boldsymbol{\mu}_t$, and $\mathbf{z}_t$ denote $N \times 1$ dimensional vectors, which respectively represent the log returns, the returns removing the mean value, the mean, and the independent and identically distributed errors. $\mathbf{H}_t$, $\mathbf{D}_t$, and $\mathbf{R}_t$ denote $N \times N$ dimensional matrices, which are respectively the conditional covariance, the standard deviation, and the correlation coefficient matrixs of $\mathbf{a}_t$ at time t. The vectors $\mathbf{a}_t$ and $\mathbf{z}_t$ satisfy the conditions $E(\mathbf{a}_t) = \mathbf{0}$, $Cov\,[\mathbf{a}_t] = \mathbf{H}_t$ and $E(\mathbf{z}_t) = \mathbf{0}$, $E\left[\mathbf{z}_t\mathbf{z}_t^T\right] = \mathbf{I}$. For more details, we refer the reader to [39].

The elements in the diagonal matrix $\mathbf{D}_t$ are the standard deviations of the univariate GARCH model, that is,

$$\mathbf{D}_t = \begin{bmatrix} \sqrt{h_{1t}} & 0 & \cdots & 0 \\ 0 & \sqrt{h_{2t}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{h_{Nt}} \end{bmatrix}, \tag{18}$$

where

$$h_{it} = \alpha_{i0} + \sum_{q=1}^{Q_i} \alpha_{iq} a_{i,t-q}^2 + \sum_{p=1}^{P_i} \beta_{ip} h_{i,t-p}. \tag{19}$$

Univariate GARCH models can have different orders, but the simplest and most effective model is GARCH (1, 1).

The conditional correlation coefficient matrix is

$$\mathbf{R}_t = \begin{bmatrix} 1 & \rho_{12,t} & \rho_{13,t} & \cdots & \rho_{1N,t} \\ \rho_{12,t} & 1 & \rho_{23,t} & \cdots & \rho_{2N,t} \\ \rho_{13,t} & \rho_{23,t} & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \rho_{N-1,N,t} \\ \rho_{1N,t} & \rho_{2N,t} & \cdots & \rho_{N-1,N,t} & 1 \end{bmatrix}. \tag{20}$$

If follows from (17), (19), and (20) that

$$[\mathbf{H}_t]_{ij} = \sqrt{h_{it} h_{jt}} \rho_{ij}. \tag{21}$$

It should be noted that $\mathbf{R}_t$ must satisfy the following two conditions.

- Condition 1: each element in $\mathbf{R}_t$ must be less than or equal to 1;
- Condition 2: $\mathbf{R}_t$ must be positive definite.

In the DCC model, $\mathbf{R}_t$ is decomposed into the following form

$$\mathbf{R}_t = \mathbf{Q}_t^{*-1} \mathbf{Q}_t \mathbf{Q}_t^{*-1}, \tag{22}$$

where $\epsilon_t = \mathbf{D}_t^{-1} \mathbf{a}_t \sim \mathcal{N}_N(\mathbf{0}, \mathbf{R}_t)$, a and b are the scalar parameters, and $\mathbf{Q}_t^*$ is a diagonal matrix composed of the square roots of the diagonal elements of $\mathbf{Q}_t$ given by

$$\mathbf{Q}_t = (1-a-b)\bar{\mathbf{Q}} + a\epsilon_{t-1}\epsilon_{t-1}^T + b\mathbf{Q}_{t-1}. \tag{23}$$

To ensure that $\mathbf{R}_t$ is positive definite, a and b must meet the following conditions:

$$a \geq 0, \quad b \geq 0, \quad a + b < 1. \tag{24}$$

The DCC model effectively solves the problem of heteroscedasticity in multivariate statistical analysis. However, the sample covariance is used to estimate $\bar{\mathbf{Q}}$ in the traditional DCC model, which performance is poor [6] when the data dimension is much larger than the sample size. Engle *et al.* [35] recently proposed a DCC-NL large-dimensional dynamic covariance matrix estimator, which uses the nonlinear shrinkage model to estimate the unconditional covariance matrix $\bar{\mathbf{Q}}$. They increased the order of magnitude of assets and obtained better numerical results. For minimum variance portfolios, Goto and Xu [26] and Torri *et al.* [27] both pointed out that the performance of the glasso method is better than the nonlinear shrinkage.

To further improve the covariance matrix, we use the glasso and tlasso methods to estimate the unconditional covariance matrix $\bar{\mathbf{Q}}$ in the DCC model. The corresponding DCC-glasso and DCC-tlasso models for large-dimensional dynamic covariance matrix estimators are developed. The estimation steps are given below.

- For each asset return series, fit the GARCH (1, 1) model.
- Use the glasso and tlasso methods to estimate the unconditional covariance matrix $\bar{\mathbf{Q}}$.
- Maximize the composite likelihood function and estimate the dynamic covariance matrix.

IV. SIMULATIONS AND APPLICATIONS

A. DATA DETAILS

The data in this article comes from the China Stock Market & Accounting Research (CSMAR) Database. The target assets of the Shanghai 50 Index (SSE50), Shanghai and Shenzhen 300 Index (HS300), and China Securities 500 Index (CSI500) are selected as the research objects.

We selected the target assets of each index starting in January 1, 2010, and ending in December 31, 2019, and used the sample covariance matrix of 10-year daily data (a total of 2430 trading days) to simulate the true covariance matrix. In practical applications, stock suspensions often increase the risk of a portfolio, so it is not appropriate to select stocks that have been suspended for a long time when constructing a portfolio stock pool. We eliminated stocks that had been suspended for a long time (greater than 20% of the sample period), and studied the remaining stocks. A total of 41 stocks remained in the SSE50, 219 stocks remained in the HS300, and 377 stocks remained in the CSI500.

In the simulations, we standardized the data to force the optimal value of the penalty parameter to appear in the range [0, 1] in the optimization process.

In the empirical research, we divided the period of the data into two parts: within the sample and out of the sample. The sample was from January 1, 2010, to January 31, 2012. Out of the sample was February 1, 2012, to December 31, 2019. The rolling period window method was used to analyze the performance of the out-of-sample model. The window of the estimated covariance matrix was $T = 500$, and the portfolio weights were readjusted by calculating the minimum variance portfolio on a fixed-size window every five trading days (one week).

B. SIMULATIONS

Simulations were divided into two parts. The first part needed to determine the optimal penalty parameters in the glasso and tlasso approaches. The second part compared the advantages and the disadvantages of each estimator through simulations.

For the first part, we used the Bayesian information criterion (BIC) to optimize ρ. We used the maximum likelihood method to estimate that the degree of freedom of the real-life asset returns was close to 4, so the degree of freedom was set to 4 in the tlasso method. The threshold in tlasso was set to 0.01 [25].

For the second part, we simulated the DCC model with $a = 0.008$ and $b = 0.938$ to compare the advantages and disadvantages of different estimators. The parameters were estimated by the actual data of 2,340 days. The univariate GARCH model was fitted with the GARCH (1, 1), where

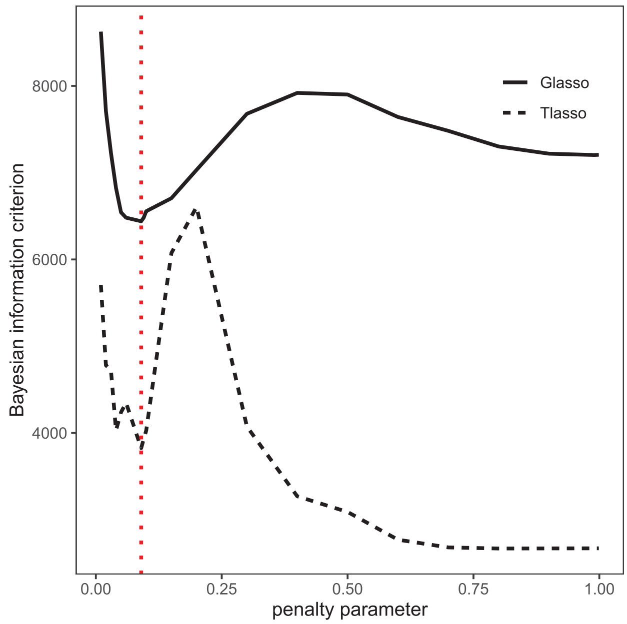

FIGURE 1. **Optimal choice of penalty parameter for glasso and tlasso based on SSE50 with 41 assets.**

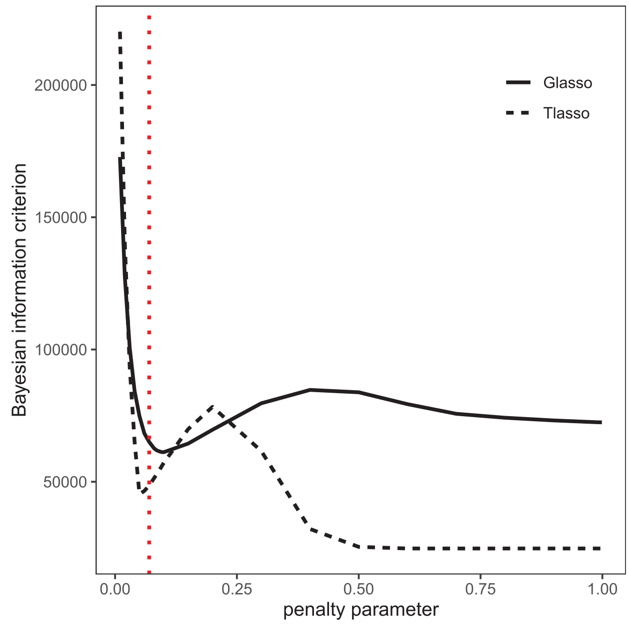

FIGURE 2. **Optimal choice of penalty parameter for glasso and tlasso based on HS300 with 219 assets.**

the parameters $\alpha_i = 0.05$, $\beta_i = 0.90$ [35]. The maximum value of the concentration ratio c is 0.756 (377/500). Both DCC-glasso and DCC-tlasso, as considered in this article, can be applied to the cases where c is greater than 1. Because the traditional DCC model does not work when $c \geq 1$. To compare with the traditional DCC model, this article does not discuss the case where c is greater than 1.

1) OPTIMAL CHOICE OF PENALTY PARAMETER

The estimated inverse covariance matrix in the glasso and tlasso approaches depends largely on the penalty parameter ρ. The larger the value of ρ, the sparser the inverse covariance matrix. In special cases, the result obtained by the model when $\rho = 0$ is the inverse of the sample covariance matrix. We use BIC to optimize ρ, calculated as

$$BIC = -2\log\left(Lik_{\rho_i}\right) + k_{\widehat{\Sigma^{-1}}} \times \log(T), \tag{25}$$

where Lik_{ρ_i} is the likelihood function value corresponding to the i-th ρ, and $k_{\widehat{\Sigma^{-1}}}$ is the number of nonzero elements in the estimated inverse covariance matrix.

Figs. 1 to 3 show the changes of BIC of the glasso and tlasso methods with the penalty parameters for the different asset amounts. It can be seen that BIC in both models tends to decrease first, increase, and then decrease again with the increase of the penalty parameter ρ. The subsequent decrease in BIC in three figures is due to the fact that with the increase of ρ, the inverse covariance matrix has been punished, resulting in the estimated inverse covariance matrix being all zero except for the diagonal elements, which is not the ideal result in this article. Therefore, we choose $\rho = 0.09$

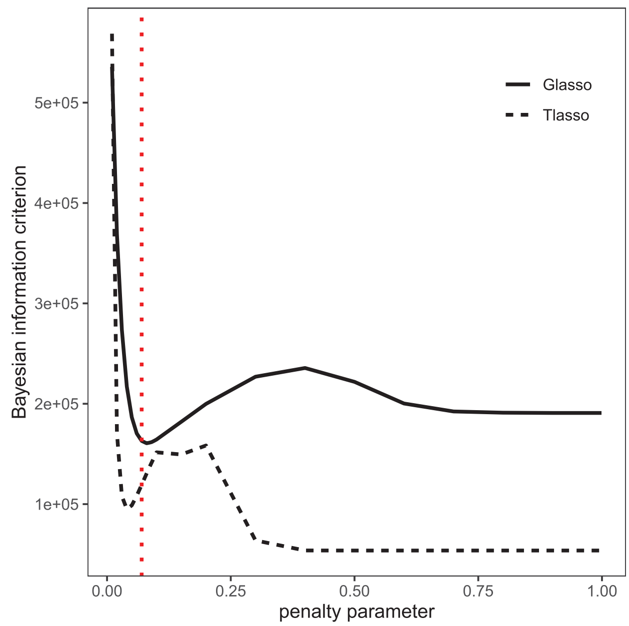

FIGURE 3. **Optimal choice of penalty parameter for glasso and tlasso based on CSI300 with 377 assets.**

for SSE50, $\rho = 0.07$ for HS300, and CSI500 in portfolio selection.

2) LIST OF ESTIMATORS AND THEIR LOSSES

To judge the quality of the estimators, we use the conditional covariance matrix loss function proposed by Engle *et al.* [35], i.e.,

$$\mathcal{L}(\widehat{\Sigma}, \Sigma) = \frac{Tr(\widehat{\Sigma}^{-1}\Sigma\widehat{\Sigma}^{-1})/N}{[Tr(\widehat{\Sigma}^{-1})/N]^2} - \frac{1}{Tr(\Sigma^{-1})/N}. \tag{26}$$

TABLE 1. The loss of each estimator.

N	DCC-S	DCC-L	DCC-NL	DCC-glasso	DCC-tlasso
41	9.058	9.132	8.716	**7.447**	8.363
219	15.23	10.77	9.692	8.714	**8.093**
377	56.57	18.26	10.87	10.12	**9.198**

[a] The unit is 10^{-5}.

The above loss function is extended to the conditional covariance matrix case,

$$L = \frac{1}{T}\sum_{t=1}^{T}\mathcal{L}(\widehat{\mathbf{H}}_t, \mathbf{H}_t). \tag{27}$$

The five methods introduced in II are applied to the DCC model to estimate the unconditional covariance matrix $\bar{\mathbf{Q}}$, and the following five estimators are constructed. We mainly discuss the last two models.

1) DCC-S: The sample covariance matrix is used to estimate the unconditional covariance matrix $\bar{\mathbf{Q}}$ in the DCC model [1].
2) DCC-L: The linear shrinkage estimation method is used to estimate the unconditional covariance matrix $\bar{\mathbf{Q}}$ in the DCC model [35].
3) DCC-NL: The nonlinear shrinkage estimation method is used to estimate the unconditional covariance matrix $\bar{\mathbf{Q}}$ in the DCC model [35].
4) DCC-glasso: The glasso method is used to estimate the unconditional covariance matrix $\bar{\mathbf{Q}}$ in the DCC model.
5) DCC-tlasso: The tlasso method is used to estimate the unconditional covariance matrix $\bar{\mathbf{Q}}$ in the DCC model.

The loss of each estimator is shown in Table 1, where the minimum loss is bolded. It can be seen that the DCC models based on the glasso and tlasso approaches achieve smaller losses. As the number of assets increases, the effect of the glasso and tlasso methods becomes more significant.

C. APPLICATIONS

1) MINIMUM VARIANCE PORTFOLIO

Since Markowitz's pioneering work, the concept of diversifying risks through diversified investment has become the core of modern investment theory [2]. The minimum variance portfolio model is essentially a quadratic optimization problem, i.e.,

$$\begin{aligned} &\min_{\mathbf{w}} \quad \mathbf{w}'\Sigma\mathbf{w} \\ &s.t. \quad \mathbf{1}'\mathbf{w} = 1, \end{aligned} \tag{28}$$

where $\mathbf{w} = (w_1, w_2, \ldots, w_N)'$ is the asset weight vector, with dimensions $N \times 1$. Its analytical solution, the optimal weight vector of the minimum variance portfolio, is given by

$$\widehat{\mathbf{w}}^* = \frac{\widehat{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\widehat{\Sigma}^{-1}\mathbf{1}}. \tag{29}$$

Since the short-selling mechanism of the Chinese stock market is not perfect, this article limits the weight to be greater than or equal to zero in the construction of portfolios. We consider the following minimum variance portfolio problem with nonnegative weights, i.e.,

$$\begin{aligned} &\min_{\mathbf{w}} \quad \mathbf{w}'\Sigma\mathbf{w} \\ &s.t. \ \mathbf{1}'\mathbf{w} = 1, \\ &\qquad w_i \geq 0, \quad i = 1, \ldots, N. \end{aligned} \tag{30}$$

TABLE 2. Descriptive statistics of stock log returns.

Stock	Mean	SD	Skewness	Kurtosis	J-B	p-value
SSE50	0.000	0.015	-0.402	8.289	2899	0.000
HS300	0.000	0.015	-0.665	7.871	2582	0.000
CSI500	0.000	0.017	-0.905	6.651	1682	0.000
CMB	0.000	0.018	0.239	6.824	1504	0.000
Moutai	0.001	0.020	-0.180	6.719	1413	0.000

[a] CMB: China Merchants Bank.
[b] Moutai: Kweichow Moutai Company Limited.
[c] SD: Standard Deviation.
[d] J-B and p-value are the JB statistical test results.

TABLE 3. Performance comparison of five models based on SSE50 with 41 assets.

Model	Standard deviation		Average return		Sharpe ratio	
	In	Out	In	Out	In	Out
DCC-S	0.65	2.29	-0.016	0.012	-3.19	0.32
DCC-L	0.82	1.53	-0.026	0.004	-3.75	-0.05
DCC-NL	0.83	**1.17**	-0.028	**0.014**	-3.95	**0.79**
DCC-glasso	0.85	1.61	-0.028	0.003	-3.85	-0.11
DCC-tlasso	0.88	1.60	-0.027	0.001	-3.61	-0.24

[a] The unit is 10^{-2}.

2) EMPIRICAL RESULTS

We used the five methods in IV-B2 to construct the minimum variance portfolio model, and considered the advantages and disadvantages of each model from the aspects of the standard deviation, the mean return and the Sharpe ratio.

Table 2 shows a descriptive statistical analysis of each index and its main constituent stocks. The results show that the kurtosis coefficient of each stock's returns is greater than 3, and the JB statistical test rejects the assumption of a normal distribution for each return series. Fig. 4 more clearly shows the density curves of the three indices and their main stock returns. Among them, the bold red curve is a normal distribution with a mean value of 0.00025 and a standard deviation of 0.0169 (the mean and the standard deviation are estimated from the sample values of the logarithmic returns). It can be found that the logarithmic returns of stocks have obvious fat tails and leptokurtic characteristics. In this article, the tlasso method is applied to the DCC model, which is more suitable for the distribution of the data itself.

Tables 3-5 shows the performance measures of the five portfolio models with different asset amounts. When constructing a minimum variance portfolio, the standard deviation is the most important performance metric. In addition, we calculate the average return of the portfolio and the

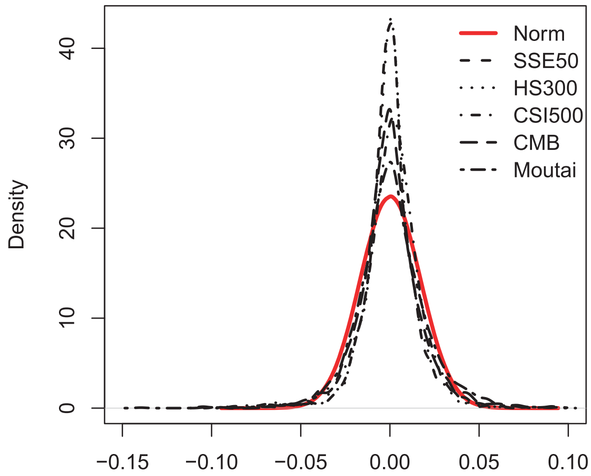

FIGURE 4. Density curve of stock log returns.

TABLE 4. Performance comparison of five models based on HS300 with 219 assets.

Model	Standard deviation		Average return		Sharpe ratio	
	In	Out	In	Out	In	Out
DCC-S	1.65	6.63	0.017	0.002	0.74	-0.04
DCC-L	2.00	4.90	0.024	0.000	0.96	-0.10
DCC-NL	2.09	2.59	0.024	0.000	0.92	-0.18
DCC-glasso	2.08	2.04	0.024	0.000	0.93	-0.23
DCC-tlasso	1.85	**1.91**	0.025	0.000	1.09	-0.25

[a] The unit is 10^{-2}.

TABLE 5. Performance comparison of five models based on CSI500 with 377 assets.

Model	Standard deviation		Average return		Sharpe ratio	
	In	Out	In	Out	In	Out
DCC-S	2.19	6.87	-0.088	0.008	-4.23	0.05
DCC-L	2.60	5.42	-0.068	0.002	-2.80	-0.05
DCC-NL	2.88	3.43	-0.069	0.003	-2.56	-0.05
DCC-glasso	1.92	2.21	-0.070	0.011	-3.88	0.28
DCC-tlasso	1.91	**2.05**	-0.070	**0.015**	-3.91	**0.49**

[a] The unit is 10^{-2}.

Sharpe ratio to analyze the risk-adjusted return. In calculating the Sharpe ratio, the risk-free interest rate is 1.75% of the one-year deposit rate in China.

It can be seen from Table 3 that the standard deviation of the DCC-S model for the in-sample is only 0.0065, which is the smallest of all the models. While the out-of-sample standard deviation is 0.0229, which is significantly higher than the in-sample standard deviation. This phenomenon can also be seen in Tables 4 and 5. The variation of the out-of-sample standard deviation of other models is significantly smaller than that of the DCC-S model. This verifies that the use of sample covariance in the estimation of large-dimensional covariance matrices will have a negative impact on the portfolio selection.

Table 3 compares of the performance results of each model, where the number of assets is 41. The DCC-NL model is optimal in all three aspects.

TABLE 6. Difference in out-of-sample standard deviations between the DCC-tlasso and the alternative methods.

Number of assets	DCC-S	DCC-L	DCC-NL	DCC-glasso
41	-0.404*	0.025	0.293*	-0.019
219	-1.290*	-0.962*	-0.325*	-0.079*
377	-1.255*	-0.992*	-0.535*	-0.088*

[a] * denote 10% significance level.

Table 4 and Table 5 compare the performance results of the models with the numbers of the assets of 219 and 377, respectively. The tables show that the performance of the DCC-glasso and DCC-tlasso models is superior to that of other models from the three out-of-sample performance metrics. The DCC-tlasso obtains the smallest out-of-sample variance and the largest average return, which leads to the highest Sharpe ratio of the models.

In the above empirical analysis, the moving window is $T = 500$. As the number of the assets increases, c will increase. Empirical research results show that the greater the c, the better the DCC-glasso and DCC-tlasso models proposed in this article. At the same time, it can be seen from Table 2 that the assets do not follow a multivariate normal distribution. Therefore, the DCC-tlasso is superior to the DCC-glasso.

The homogeneity of the variance test can determine whether the variances of the two populations are equal, but if the test data do not obey a normal distribution, the robustness of the test will be greatly reduced. In view of this, Ledoit and Wolf [40] proposed an improved test method called bootstrap inference, which we use to test for a significant difference between the DCC-tlasso portfolios and the alternative methods in the out-of-sample standard deviation. Table 6 shows that the out-of-sample standard deviation of the DCC-tlasso is significantly lower than that of other models as the number of assets become larger.

V. CONCLUSION

We applied two sparse estimators, the glasso and tlasso approaches, to estimate the unconditional covariance matrix $\bar{\mathbf{Q}}$ in the DCC model. The corresponding models are called the DCC-glasso and the DCC-tlasso, respectively. Numerical simulations illustrate that the proposed models have smaller losses than other combined DCC types, including the DCC-S, the DCC-L, and the DCC-NL models. We applied them to the real-life stock returns data, and found that the greater the asset dimension, the better DCC-glasso and DCC-tlasso perform.

The anonymous referees pointed out the proposal for the GGM under the long-tailed symmetric distribution by Ağraz and Purutçuoğlu and the contribution to the construction of a robust investment portfolio by Kara *et al.*. The detailed comparisons of models can also be carried out in future research. Another interesting topic about combining approximate factor model (AFM) [41] with the DCC-tlasso to construct the AFM-DCC-tlasso can be considered in future studies.

ACKNOWLEDGMENT

This authors would like to thank the editor and the anonymous referees for their useful comments and suggestions, which helped to improve the presentation of this article.

REFERENCES

[1] R. Engle, "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models," *J. Bus. Econ. Statist.*, vol. 20, no. 3, pp. 339–350, Jul. 2002.
[2] H. Markowitz, "Portfolio selection," *J. Finance*, vol. 7, no. 1, pp. 77–91, 1952.
[3] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *J. Multivariate Anal.*, vol. 88, no. 2, pp. 365–411, Feb. 2004.
[4] O. Ledoit and M. Wolf, "Nonlinear shrinkage estimation of large-dimensional covariance matrices," *Ann. Statist.*, vol. 40, no. 2, pp. 1024–1060, Apr. 2012.
[5] L. Laloux, P. Cizeau, J. P. Bouchaud, and M. Potters, "Noise dressing of financial correlation matrices," *Phys. Rev. Lett.*, vol. 83, no. 7, p. 1467, Aug. 1999.
[6] R. Sun, T. Ma, S. Liu, and M. Sathye, "Improved covariance matrix estimation for portfolio risk measurement: A review," *J. Risk Financial Manage.*, vol. 12, no. 1, p. 48, Mar. 2019.
[7] F. Wen, Y. Yang, P. Liu, and R. C. Qiu, "Positive definite estimation of large covariance matrix using generalized nonconvex penalties," *IEEE Access*, vol. 4, pp. 4168–4182, Jul. 2016.
[8] C. Stein, "Lectures on the theory of estimation of many parameters," *J. Sov. Math.*, vol. 34, no. 1, pp. 1373–1403, Jul. 1986.
[9] O. Ledoit and M. Wolf, "Improved estimation of the covariance matrix of stock returns with an application to portfolio selection," *J. Empirical Finance*, vol. 10, no. 5, pp. 603–621, Dec. 2003.
[10] O. Ledoit and M. Wolf, "Honey, I shrunk the sample covariance matrix," *J. Portfolio Manage.*, vol. 30, no. 4, pp. 110–119, Jul. 2004.
[11] B. Zhang, J. Zhou, and J. Li, "Improved shrinkage estimators of covariance matrices with toeplitz-structured targets in small sample scenarios," *IEEE Access*, vol. 7, pp. 116785–116798, Aug. 2019.
[12] O. Ledoit and M. Wolf, "Spectrum estimation: A unified framework for covariance matrix estimation and PCA in large dimensions," *J. Multivariate Anal.*, vol. 139, pp. 360–384, Jul. 2015.
[13] O. Ledoit and M. Wolf, "Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks," *Rev. Financial Stud.*, vol. 30, no. 12, pp. 4349–4388, Dec. 2017.
[14] O. Ledoit and M. Wolf, "Optimal estimation of a large-dimensional covariance matrix under stein's loss," *Bernoulli*, vol. 24, no. 4B, pp. 3791–3832, Nov. 2018.
[15] O. Ledoit and M. Wolf, "Quadratic shrinkage for large covariance matrices," Univ. Zurich, Zurich, Switzerland, Tech. Rep. 335, Nov. 2019.
[16] V. DeMiguel and F. J. Nogales, "Portfolio selection with robust estimation," *Oper. Res.*, vol. 57, no. 3, pp. 560–577, Jun. 2009.
[17] C. Lam and J. Fan, "Sparsistency and rates of convergence in large covariance matrix estimation," *Ann. Statist.*, vol. 37, no. 6B, pp. 4254–4278, Dec. 2009.
[18] J. Fan, Y. Liao, and H. Liu, "An overview of the estimation of large covariance and precision matrices," *Econometrics J.*, vol. 19, no. 1, pp. C1–C32, Mar. 2016.
[19] P. J. Bickel and E. Levina, "Covariance regularization by thresholding," *Ann. Statist.*, vol. 36, no. 6, pp. 2577–2604, Dec. 2008.
[20] J. Fan, Y. Fan, and J. Lv, "High dimensional covariance matrix estimation using a factor model," *J. Econometrics*, vol. 147, no. 1, pp. 186–197, Nov. 2008.
[21] C. Lam and Q. Yao, "Factor modeling for high-dimensional time series: Inference for the number of factors," *Ann. Statist.*, vol. 40, no. 2, pp. 694–726, Apr. 2012.
[22] J. Fan, Y. Liao, and M. Mincheva, "Large covariance estimation by thresholding principal orthogonal complements," *J. Roy. Stat. Soc., B, Stat. Methodol.*, vol. 75, no. 4, pp. 603–680, Sep. 2013.
[23] G. V. G. Stevens, "On the inverse of the covariance matrix in portfolio analysis," *J. Finance*, vol. 53, no. 5, pp. 1821–1827, Oct. 1998.
[24] J. Friedman, T. Hastie, and R. Tibshirani, "Sparse inverse covariance estimation with the graphical lasso," *Biostatistics*, vol. 9, no. 3, pp. 432–441, Jul. 2008.
[25] M. Finegold and M. Drton, "Robust graphical modeling of gene networks using classical and alternative t-distributions," *Ann. Appl. Statist.*, vol. 5, no. 2A, pp. 1057–1080, Jun. 2011.
[26] S. Goto and Y. Xu, "Improving mean variance optimization through sparse hedging restrictions," *J. Financial Quant. Anal.*, vol. 50, no. 6, pp. 1415–1441, Dec. 2015.
[27] G. Torri and R. Giacometti, "Sparse precision matrices for minimum variance portfolios," *Comput. Manage. Sci.*, vol. 16, no. 3, pp. 375–400, Jul. 2019.
[28] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation," *Econometrica*, vol. 50, no. 4, pp. 987–1007, 1982.
[29] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *J. Econometrics*, vol. 31, no. 3, pp. 307–327, Apr. 1986.
[30] R. F. Engle and K. F. Kroner, "Multivariate simultaneous generalized ARCH," *Econ. Theory*, vol. 11, no. 1, pp. 122–150, Feb. 1995.
[31] C. Pakel, N. Shephard, K. Sheppard, and R. F. Engle, "Fitting vast dimensional time-varying covariance models," *J. Bus. Econ. Statist.*, pp. 1–17, Feb. 2020. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/07350015.2020.1713795
[32] A. Hassanein and H. Elgohari, "The linkage between stock and inter-bank bond markets in China: A dynamic conditional correlation (DCC) analysis," *Int. J. Econ. Bus. Res.*, vol. 20, no. 1, pp. 80–99, Jun. 2020.
[33] C. M. Hafner and O. Reznikova, "On the estimation of dynamic conditional correlation models," *Comput. Statist. Data Anal.*, vol. 56, no. 11, pp. 3533–3545, Nov. 2012.
[34] L. P. Liu, D. Ma, and W. P. Bai, "Estimation and application study on the dynamic conditional covariance of large dimensional data," *Statist. Res.*, vol. 32, no. 6, pp. 105–112, Jun. 2015.
[35] R. F. Engle, O. Ledoit, and M. Wolf, "Large dynamic covariance matrices," *J. Bus. Econ. Statist.*, vol. 37, no. 2, pp. 363–375, Apr. 2019.
[36] R. A. Johnson and D. W. Wichern, "Sample geometry and random sampling," in *Applied Multivariate Statistical Analysis*, 6th ed. London, U.K.: Pearson, 2007, pp. 137–139.
[37] M. Yuan and Y. Lin, "Model selection and estimation in the Gaussian graphical model," *Biometrika*, vol. 94, no. 1, pp. 19–35, Feb. 2007.
[38] D. M. Witten, J. H. Friedman, and N. Simon, "New insights and faster computations for the graphical lasso," *J. Comput. Graph. Statist.*, vol. 20, no. 4, pp. 892–900, Jan. 2011.
[39] E. Orskaug, "Multivariate DCC-GARCH model," M.S. thesis, Dept. Math. Sci., Norwegian Univ. Sci. Technol., Trondheim, Norway, 2009.
[40] O. Ledoit and M. Wolf, "Robust performances hypothesis testing with the variance," *Wilmott*, vol. 2011, no. 55, pp. 86–89, Sep. 2011.
[41] J. Fan, Y. Liao, and M. Mincheva, "High dimensional covariance matrix estimation in approximate factor models," *Ann. Statist.*, vol. 39, no. 6, pp. 3320–3356, Jan. 2011.

XIN YUAN received the B.S. degree in mathematics and applied mathematics from Pingdingshan University, China, in 2015. He is currently pursuing the M.S. degree in statistics with the School of Mathematics, Physics, and Statistics, Shanghai University of Engineering Science, China. His current research interests include high-dimension statistical inference, sparse statistical analysis, and their applications in financial risk management.

WEIQIN YU received the B.S. degree in mathematics and applied mathematics from Yangzhou University, China, in 2004, and the M.S. and Ph.D. degrees in applied mathematics from the Nanjing University of Aeronautics and Astronautics, China, in 2007 and 2010, respectively. She is currently an Associate Professor with the School of Mathematics, Physics, and Statistics, Shanghai University of Engineering Science, China. Her current research interests include high-dimension statistical inference and statistical optimization.

ZHIXIANG YIN received the Ph.D. degree from the Huazhong University of Science and Technology, China, in 2003. He is currently a Professor with the School of Mathematics, Physics, and Statistics, Shanghai University of Engineering Science, China. Furthermore, he is also the Dean of the School of Mathematics, Physics, and Statistics, Shanghai University of Engineering and Technology, China. His current research interests include combinatorial and optimization, intelligent computing, and protein structure prediction.

GUOQIANG WANG (Member, IEEE) received the B.S. degree in mathematics and applied mathematics from Shandong Normal University, China, in 2002, and the M.S. and Ph.D. degrees in operations research and cybernetics from Shanghai University, China, in 2005 and 2009, respectively.

From 2012 to 2013, he was a Visiting Research Associate with the Department of Mathematics and Statistics, Curtin University, Australia. He is currently a Professor with the School of Mathematics, Physics, and Statistics, Shanghai University of Engineering Science, China. His current research interests include high-dimension statistical inference, statistical optimization, optimization theory, methods, and applications.