

Research Article

Improved Shrinkage Estimator of Large-Dimensional Covariance Matrix under the Complex Gaussian Distribution

Bin Zhang ^{1,2}

¹College of Education, Hubei Minzu University, Enshi, Hubei 445000, China

²College of Mathematics, Sichuan University, Chengdu, Sichuan 610064, China

Correspondence should be addressed to Bin Zhang; 2009021@hbmy.edu.cn

Received 29 March 2020; Revised 5 May 2020; Accepted 28 May 2020; Published 7 July 2020

Guest Editor: Jianfeng Li

Copyright © 2020 Bin Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Estimating the covariance matrix of a random vector is essential and challenging in large dimension and small sample size scenarios. The purpose of this paper is to produce an outperformed large-dimensional covariance matrix estimator in the complex domain via the linear shrinkage regularization. Firstly, we develop a necessary moment property of the complex Wishart distribution. Secondly, by minimizing the mean squared error between the real covariance matrix and its shrinkage estimator, we obtain the optimal shrinkage intensity in a closed form for the spherical target matrix under the complex Gaussian distribution. Thirdly, we propose a newly available shrinkage estimator by unbiasedly estimating the unknown scalars involved in the optimal shrinkage intensity. Both the numerical simulations and an example application to array signal processing reveal that the proposed covariance matrix estimator performs well in large dimension and small sample size scenarios.

1. Introduction

The problem of estimating the covariance matrix of a random vector arises in both multivariate statistical theory and various applications [1, 2]. In large sample setting, where the dimension of a random vector is small and the sample size is large enough, the sample covariance matrix (SCM) is a reliable estimator of the real covariance matrix and is widely employed in many scenarios. However, suffering from the curse of dimensionality, the SCM becomes ill-conditioned or even singular in large dimension scenarios [3]. Then, severe consequences may appear if the SCM remained as the covariance matrix estimator [4, 5].

During the last two decades, scientists have proposed many regularization strategies to generate outperformed covariance matrix estimators in large dimension scenarios [6–10]. Among these, the linear shrinkage estimation is an effective strategy to inspire a well-conditioned covariance matrix estimator when the dimension is large compared to the sample size [11, 12]. When the prior information of the covariance structure is available, the linear shrinkage

estimator is modeled as a linear combination between the SCM and a proper target matrix. In the existing literature, the target matrices, which are usually formed through structuring the SCM in line with the prior information, include the spherical target and others, such as the diagonal target, the Toeplitz rectified target, and the tapered SCM [13].

With the aid of prior information, the linear shrinkage estimator can always outperform the SCM when the involved tuning parameter is carefully selected [14]. Therefore, one of the crucial difficulties in linear shrinkage estimation is to determine the optimal tuning parameter which is also called as the shrinkage intensity. By minimizing the mean squared error (MSE) between the shrinkage estimator and the real covariance matrix, the optimal tuning parameter can be expressed in a closed form for an arbitrary target. However, it is comprised of unknown scalars which involve the expectation operator and the real covariance matrix, leading to a chief difficulty in generating an available shrinkage estimator. In particular, when the data follow a specific distribution such as Gaussian distribution or

elliptical distribution, the expectation can be further calculated [11]. What is noteworthy is that the optimal tuning parameter owns different expressions under different distributions, even for the same target [15]. Therefore, it is necessary to discover the specific properties of shrinkage intensity under some typical distributions. In many applications, such as array signal processing, the data come from the complex domain. There have been some related research studies on it. In [16], a linear shrinkage estimator for Toeplitz rectified target is developed under the complex Gaussian distribution, whereas the involved unknown scalars in the optimal tuning parameter are not unbiasedly estimated, resulting in a suboptimal covariance matrix estimator [17]. In [18], a linear shrinkage estimator is proposed via low-complexity crossvalidation under an arbitrary complex distribution. Therefore, when the data follow a specific distribution, the linear shrinkage estimator could be further improved by making full use of the distribution information.

In this paper, we further research the linear shrinkage estimator under the complex Gaussian distribution. The target matrix is chosen as the spherical target which has been widely studied under the real number field [6, 11, 14]. The optimal tuning parameter is obtained by minimizing the MSE. We remind that the above optimal tuning parameter involves both the expectation operator and the real covariance matrix. By developing a novel moment property of the complex Wishart distribution, we can calculate the expectation operator. Then, the optimal tuning parameter turns to be only related to some unknown scalars concerning the real covariance matrix. A popular approach is adopted by replacing these unknown scalars with their estimates to obtain an available tuning parameter. Furthermore, good estimates of unknown scalars can benefit the corresponding available tuning parameter and the corresponding shrinkage estimator [11, 19].

The main contributions of this paper are summarized as three-fold:

- (1) A necessary moment property of the complex Wishart distribution is developed. On this basis, the optimal tuning parameter for the spherical target is analytically expressed under the complex Gaussian distribution.
- (2) All the unknown scalars involved in the optimal tuning parameter are unbiasedly estimated. Then, the corresponding available linear shrinkage estimator under the complex Gaussian distribution is proposed.
- (3) The performance of the proposed covariance matrix estimator is verified with comparison to the existing estimators in numerical simulations and an example application to adaptive beamforming.

The rest of this paper is organized as follows: Section 2 formulates the linear shrinkage estimation under the complex Gaussian distribution as a quadratic programming problem. The optimal solution is analytically obtained. Section 3 unbiasedly estimates the relative unknown scalars and subsequently proposes a new shrinkage estimator for the

spherical target. Section 4 provides some numerical simulations and an example application for verifying the performance of the proposed covariance matrix estimator. Section 5 concludes.

1.1. Notations. The notation \mathbb{C}^m is the set of all m -dimensional complex column vectors, and \mathbb{H}^n is the set of all $n \times n$ Hermitian matrices. The symbol \mathbb{E} denotes the mathematical expectation. The bold symbols $\mathbf{0}$ and $\mathbf{1}$ respectively denote the column vectors having all entries 0 and 1 with an appropriate dimension. The symbol \mathbf{I}_n denotes the $n \times n$ identity matrix. For a matrix \mathbf{A} , \mathbf{A}^H and $\|\mathbf{A}\|$ respectively denote its conjugate transpose and Frobenius matrix norm. For a squared matrix \mathbf{A} , \mathbf{A}^{-1} and $\text{tr}(\mathbf{A})$ respectively denote its inverse and trace. For two real numbers a and b , $a \wedge b$ and $a \vee b$ respectively mean the maximum and minimum of a and b .

2. Formulation and the Optimal Solution

Assume a p -dimensional random vector $\mathbf{x} \in \mathbb{C}^p$ follows the complex Gaussian distribution $\mathcal{CN}(\mathbf{0}, \Sigma)$, where Σ is the unknown covariance matrix. Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n \in \mathbb{C}^p$ be an independent and identically distributed (i.i.d.) sample, then the sample covariance matrix \mathbf{S} is defined by

$$\mathbf{S} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^H. \quad (1)$$

For an arbitrary prespecified target matrix $\mathbf{T} \in \mathbb{H}^p$ which represents an aspect prior information of the real covariance matrix structure [20], the linear shrinkage estimator of covariance matrix Σ is modeled as

$$\hat{\Sigma} = (1 - w)\mathbf{S} + w\mathbf{T}, \quad (2)$$

where $w \in [0, 1]$ is the tuning parameter which is also called shrinkage intensity [21]. Because \mathbf{S} and \mathbf{T} are Hermitian, we have $\hat{\Sigma} \in \mathbb{H}^p$ for an arbitrary $w \in [0, 1]$.

To find the optimal shrinkage intensity, we employ the MSE criterion:

$$\mathcal{M}_{\mathbf{T}}(w) = \mathbb{E}[\|\hat{\Sigma} - \Sigma\|^2] = \mathbb{E}[\|(1 - w)\mathbf{S} + w\mathbf{T} - \Sigma\|^2]. \quad (3)$$

Furthermore, we have

$$\mathcal{M}_{\mathbf{T}}(w) = w^2 \mathbb{E}[\text{tr}(\mathbf{T} - \mathbf{S})^2] - 2w \mathbb{E}[\text{tr}(\mathbf{T} - \mathbf{S})(\Sigma - \mathbf{S})] + c, \quad (4)$$

where $c = \mathbb{E}[\text{tr}(\Sigma - \mathbf{S})(\Sigma - \mathbf{S})^2]$ is a constant. Therefore, the optimal shrinkage intensity can be obtained through solving the following optimization problem:

$$\begin{aligned} \min \quad & w^2 \mathbb{E}[\text{tr}(\mathbf{T} - \mathbf{S})^2] - 2 \mathbb{E}[\text{tr}(\mathbf{T} - \mathbf{S})(\Sigma - \mathbf{S})]w, \\ \text{s.t.} \quad & 0 \leq w \leq 1. \end{aligned} \quad (5)$$

It is worth noticing that the objective function in optimization problem (5) is a convex quadratic function of w , and the optimal shrinkage intensity can be expressed in a closed form as follows:

$$w_0^* = 0 \wedge \frac{\mathbb{E}[\text{tr}(\mathbf{T} - \mathbf{S})(\boldsymbol{\Sigma} - \mathbf{S})]}{\mathbb{E}[\text{tr}(\mathbf{T} - \mathbf{S})^2]} \sqrt{1}. \quad (6)$$

Furthermore, for the spherical target $\mathbf{T} = (\text{tr}(\mathbf{S})/p)\mathbf{I}_p$, the optimal shrinkage intensity w_0^* given by (6) becomes

$$w_1^* = 0 \wedge \frac{(1/p)\text{tr}^2(\boldsymbol{\Sigma}) - \text{tr}(\boldsymbol{\Sigma}^2) + \mathbb{E}[\text{tr}(\mathbf{S}^2)] - (1/p)\mathbb{E}[\text{tr}^2(\mathbf{S})]}{\mathbb{E}[\text{tr}(\mathbf{S}^2)] - (1/p)\mathbb{E}[\text{tr}^2(\mathbf{S})]} \sqrt{1}. \quad (7)$$

Denote the matrix

$$\mathbf{E}_n = \begin{pmatrix} 1 & \frac{1}{n} \\ \frac{1}{n} & 1 \end{pmatrix}. \quad (8)$$

We can obtain the following moment property of the complex Gaussian distribution.

Proposition 1. Assume an i.i.d. sample $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n \in \mathbb{C}^p$ follows the complex Gaussian distribution $\mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma})$ and \mathbf{S} is the sample covariance matrix; we have

$$\mathbb{E}[\text{tr}^2(\mathbf{W})] = \sum_{i=1}^p \sum_{j=1}^p \sum_{k=1}^p \sum_{l=1}^p \mathbb{E}[w_{ij}w_{kl}] = \sum_{i=1}^p \sum_{j=1}^p \sum_{k=1}^p \sum_{l=1}^p (b_1 \delta_{ij} \delta_{kl} + b_2 \delta_{il} \delta_{jk}) = b_1 \text{tr}^2(\mathbf{I}_p) + b_2 \text{tr}(\mathbf{I}_p), \quad (13)$$

where $b_1 = n^2$ and $b_2 = n$. For a random matrix \mathbf{W} which follows complex Wishart distribution $\mathcal{CW}(\boldsymbol{\Sigma}, n)$ with degree of freedom n , let $\boldsymbol{\Sigma} = \mathbf{G}\mathbf{G}^H$; then, we have $\mathbf{G}^{-1}\mathbf{W}\mathbf{G}^{-H} \sim \mathcal{CW}(\mathbf{I}_p, n)$. In the same manner, we can obtain

$$\mathbb{E}[\text{tr}^2(\mathbf{W})] = n^2 \text{tr}^2(\boldsymbol{\Sigma}) + n \text{tr}(\boldsymbol{\Sigma}^2). \quad (14)$$

Noticing that $n\mathbf{S} \sim \mathcal{CW}(\boldsymbol{\Sigma}, n)$, we can obtain

$$\mathbb{E}[\text{tr}^2(n\mathbf{S})] = n^2 \text{tr}^2(\boldsymbol{\Sigma}) + n \text{tr}(\boldsymbol{\Sigma}^2). \quad (15)$$

Therefore, we have

$$\mathbb{E}[\text{tr}^2(\mathbf{S})] = \text{tr}^2(\boldsymbol{\Sigma}) + \frac{1}{n} \text{tr}(\boldsymbol{\Sigma}^2). \quad (16)$$

By (10) and (16), equality (9) holds. \square

Theorem 1. When the target matrix is $\mathbf{T} = (\text{tr}(\mathbf{S})/p)\mathbf{I}_p$, the optimal shrinkage intensity under the complex Gaussian distribution is

$$w^* = \frac{p \text{tr}^2(\boldsymbol{\Sigma}) - \text{tr}(\boldsymbol{\Sigma}^2)}{(p-n)\text{tr}^2(\boldsymbol{\Sigma}) + (np-1)\text{tr}(\boldsymbol{\Sigma}^2)} \in [0, 1]. \quad (17)$$

Proof. By plugging equalities (10) and (16) into (7), we have

$$\mathbb{E} \begin{pmatrix} \text{tr}(\mathbf{S}^2) \\ \text{tr}^2(\mathbf{S}) \end{pmatrix} = \mathbf{E}_n \begin{pmatrix} \text{tr}(\boldsymbol{\Sigma}^2) \\ \text{tr}^2(\boldsymbol{\Sigma}) \end{pmatrix}. \quad (9)$$

Proof. Because the sample $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n \in \mathbb{C}^p$ follows a complex Gaussian distribution with mean $\mathbf{0}$, we have $n\mathbf{S} \sim \mathcal{CW}(\boldsymbol{\Sigma}, n)$. By the moment properties of complex Wishart distribution [22], we can obtain

$$\mathbb{E}[\text{tr}(\mathbf{S}^2)] = \text{tr}(\boldsymbol{\Sigma}^2) + (1/n)\text{tr}^2(\boldsymbol{\Sigma}). \quad (10)$$

Furthermore, when a random matrix $\mathbf{W} = (w_{ij})_{p \times p}$ follows complex Wishart distribution $\mathcal{CW}(\mathbf{I}_p, n)$ with degree of freedom n , we have

$$\text{tr}(\mathbf{W}) = \sum_{i=1}^p \sum_{j=1}^p w_{ij}, \quad (11)$$

$$\text{tr}^2(\mathbf{W}) = \sum_{i=1}^p \sum_{j=1}^p \sum_{k=1}^p \sum_{l=1}^p w_{ij}w_{kl}. \quad (12)$$

By taking expectation on both sides, we have

$$\mathbb{E}[\text{tr}(\mathbf{S}^2)] - \frac{1}{p} \mathbb{E}[\text{tr}^2(\mathbf{S})] = \frac{np-1}{np} \text{tr}(\boldsymbol{\Sigma}^2) + \frac{p-n}{np} \text{tr}^2(\boldsymbol{\Sigma}). \quad (18)$$

Therefore, we can obtain

$$w^* = 0 \wedge \frac{p \text{tr}^2(\boldsymbol{\Sigma}) - \text{tr}(\boldsymbol{\Sigma}^2)}{(p-n)\text{tr}^2(\boldsymbol{\Sigma}) + (np-1)\text{tr}(\boldsymbol{\Sigma}^2)} \sqrt{1}. \quad (19)$$

By Cauchy-Schwarz inequality, we have $w^* \in [0, 1]$. Hence, we have

$$w^* = \frac{p \text{tr}^2(\boldsymbol{\Sigma}) - \text{tr}(\boldsymbol{\Sigma}^2)}{(p-n)\text{tr}^2(\boldsymbol{\Sigma}) + (np-1)\text{tr}(\boldsymbol{\Sigma}^2)} \in [0, 1]. \quad (20)$$

By Theorem 1, the corresponding optimal linear shrinkage estimator is

$$\hat{\boldsymbol{\Sigma}} = (1 - w^*)\mathbf{S} + w^* \frac{\text{tr}(\mathbf{S})}{p} \mathbf{I}_p. \quad (21)$$

We remind that the optimal shrinkage estimator concerns with the real covariance matrix. Thus, it is unavailable in practical applications. Despite this, it provides a theoretical optimal value for evaluating the available ones. \square

3. Available Linear Shrinkage Estimator

Theorem 2. Under the complex Gaussian distribution, the unbiased estimates of $\text{tr}(\Sigma^2)$ and $\text{tr}^2(\Sigma)$ are, respectively, given by

$$\begin{aligned}\alpha &= \frac{1}{n^2 - 1} (n^2 \text{tr}(\mathbf{S}^2) - n \text{tr}^2(\mathbf{S})), \\ \beta &= \frac{1}{n^2 - 1} (n^2 \text{tr}^2(\mathbf{S}) - n \text{tr}(\mathbf{S}^2)).\end{aligned}\quad (22)$$

Proof. Because the inverse matrix of \mathbf{E}_n is

$$\mathbf{E}_n^{-1} = \frac{n^2}{n^2 - 1} \begin{pmatrix} 1 & \frac{1}{n} \\ -\frac{1}{n} & 1 \end{pmatrix}, \quad (23)$$

we have

$$\mathbb{E} \left(\mathbf{E}_n^{-1} \begin{pmatrix} \text{tr}(\mathbf{S}^2) \\ \text{tr}^2(\mathbf{S}) \end{pmatrix} \right) = \begin{pmatrix} \text{tr}(\Sigma^2) \\ \text{tr}^2(\Sigma) \end{pmatrix}. \quad (24)$$

Therefore, we can obtain

$$\begin{aligned}\mathbb{E}[\alpha] &= \text{tr}(\Sigma^2), \\ \mathbb{E}[\beta] &= \text{tr}^2(\Sigma),\end{aligned}\quad (25)$$

revealing that α and β are unbiased estimates of $\text{tr}(\Sigma^2)$ and $\text{tr}^2(\Sigma)$, respectively.

Through plugging the unbiased estimates given by (22) into the optimal shrinkage intensity w^* , we can obtain the available shrinkage intensity:

$$\hat{w} = 0 \wedge \frac{p\beta - \alpha}{(p-n)\beta + (np-1)\alpha} \vee 1. \quad (26)$$

Therefore, the available linear shrinkage estimator is

$$\hat{\Sigma} = (1 - \hat{w})\mathbf{S} + \hat{w} \frac{\text{tr}(\mathbf{S})}{p} \mathbf{I}_p. \quad (27)$$

The linear shrinkage estimator given by (27) is positive definite even when the dimension exceeds the sample size, except that $\hat{\Sigma}$ degenerates into the SCM. \square

4. Numerical Simulations and Adaptive Beamforming

In this section, we provide some numerical simulations and an example application to adaptive beamforming for verifying the performance of the proposed covariance matrix estimator. The proposed linear shrinkage estimator is denoted as T1cg. The linear shrinkage estimator corresponding to the spherical target matrix in [18] is denoted as T1cv.

4.1. Numerical Simulations. As mentioned before, an accurate shrinkage intensity estimate can benefit the linear shrinkage estimator. In this section, we compare the

proposed available shrinkage intensity and the existing one based on crossvalidation in [18] to reveal the advantage of the proposed shrinkage estimator. In our simulations, the real covariance matrix is $\Sigma = (\sigma_{ij})_{p \times p}$ with

$$\sigma_{ij} = t^{|i-j|}. \quad (28)$$

The model parameter t is set to be 0.5, resulting in the real covariance matrix being close to a spherical structured matrix. The data come from the complex Gaussian distribution $\mathcal{CN}(\mathbf{0}, \Sigma)$. The MSE of each available shrinkage intensity relative to the optimal intensity given by (17) is computed by averaging 5×10^4 Marlo runs.

Figure 1 reports the MSEs of available shrinkage intensities versus the sample size and the dimension. We can see that the MSEs of available shrinkage intensities in T1cv and T1cg decrease as the sample size or the dimension gets larger. Because the proposed T1cg by plugging in the unbiased estimates of unknown scalars employs the complex Gaussian distribution information, it outperforms the T1cv based on the nonparameter approach.

4.2. Adaptive Beamforming. In this section, we apply the proposed covariance estimators to array signal processing. Specifically, we consider a uniform linear array (ULA) which consists of p sensors with half-wavelength spacing. At time $t = 1, \dots, n$, the received signal can be modeled as

$$\mathbf{x}(t) = \mathbf{a}(\theta_0)s_0(t) + \sum_{k=1}^K \mathbf{a}(\theta_k)s_k(t) + \mathbf{n}(t) \in \mathbb{C}^p, \quad (29)$$

where θ_0 and θ_k are the directions of desired signal $s_0(t)$ and interference signals $s_k(t)$, respectively, $\mathbf{a}(\theta_0)$ and $\mathbf{a}(\theta_k)$ are the corresponding array responses, and $\mathbf{n}(t)$ is the noise [23]. Then, the minimum variance distortionless response (MVDR) beamformer is expressed as

$$\mathbf{w} = \frac{\Sigma^{-1} \mathbf{a}(\theta_0)}{\mathbf{a}(\theta_0)^H \Sigma^{-1} \mathbf{a}(\theta_0)}. \quad (30)$$

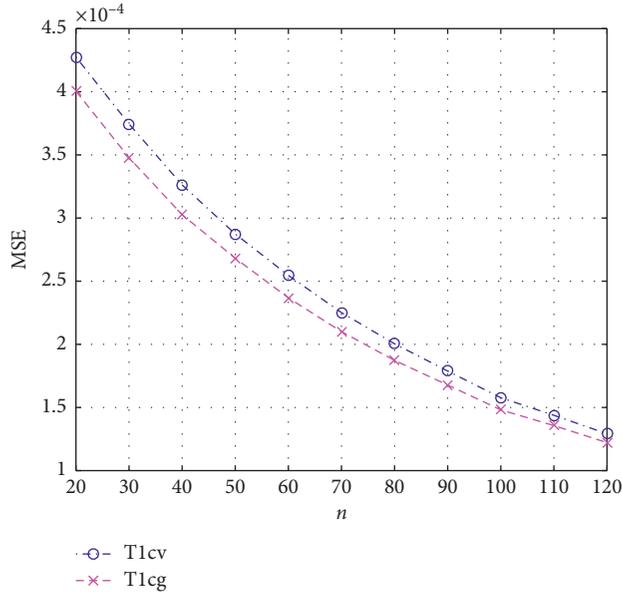
The covariance matrix Σ in (30) is unknown and suggested to be replaced with its estimate $\hat{\Sigma}$ [24]. Then, the corresponding output signal-to-interference-plus-noise ratio (SINR) is

$$\text{SINR} = \frac{\sigma_0^2 |\hat{\mathbf{w}}^H \mathbf{a}(\theta_0)|^2}{\hat{\mathbf{w}}^H (\hat{\Sigma} - \mathbf{a}(\theta_0)\mathbf{a}(\theta_0)^H) \hat{\mathbf{w}}}, \quad (31)$$

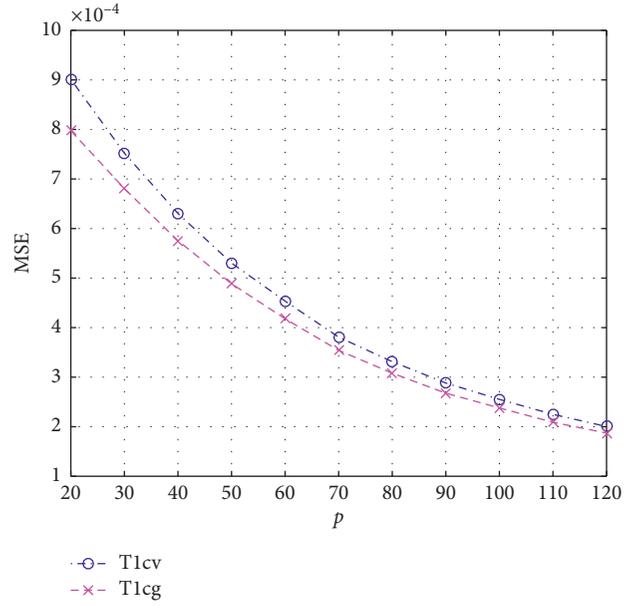
where $\hat{\mathbf{w}}$ is the estimated beamformer corresponding to $\hat{\Sigma}$. The covariance matrix estimator with larger output SINR is preferred in array signal processing.

In our simulations, we assume the desired signal has an angle of arrival of $\theta_0 = 5^\circ$ with power $\sigma_0^2 = 10$ dB, and the interference signals come from the directions $\{-10^\circ, 0^\circ, 10^\circ\}$ with power 8 dB. For each covariance estimator, the corresponding output SINR is approximated by averaging 5×10^6 repetitions.

Figures 2 and 3 report the SINR and corresponding elapsed time of the adaptive beamformers based on different

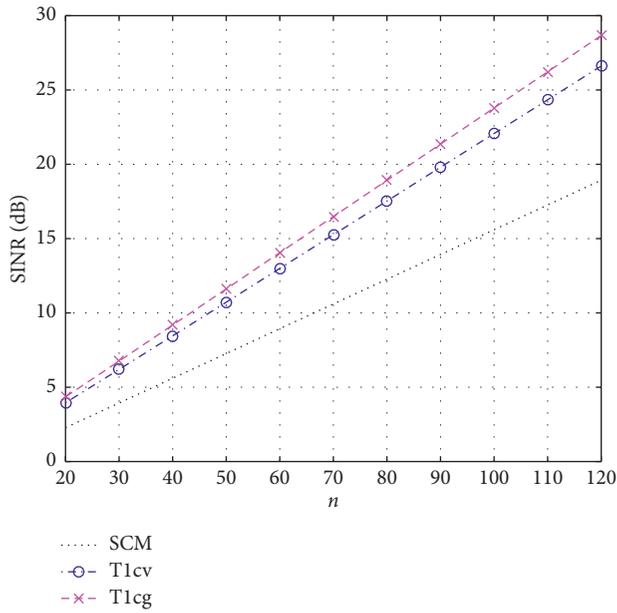


(a)

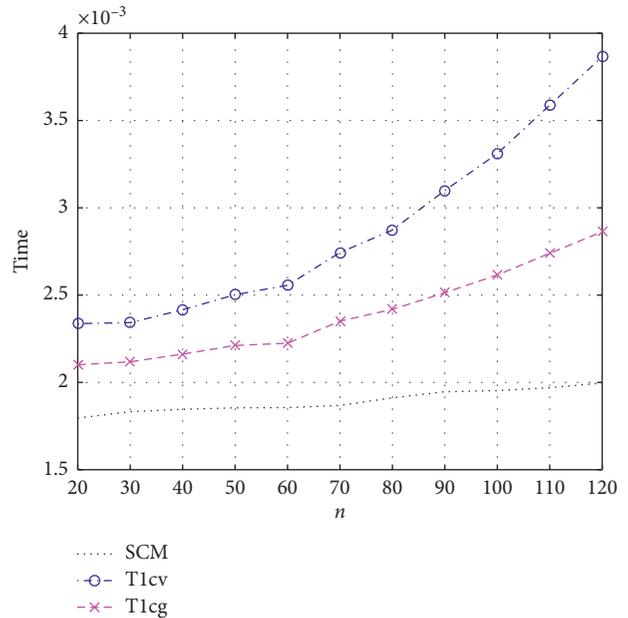


(b)

FIGURE 1: The MSEs of available shrinkage intensities when $p = 60$ (a) and $n = 60$ (b).



(a)



(b)

FIGURE 2: The SINR (a) and corresponding elapsed time (b) of adaptive beamformer based on each covariance matrix estimator versus the sample size under the complex Gaussian distribution.

covariance matrix estimators under the complex Gaussian scenario, where the noise $\mathbf{n}(t)$ follows the complex Gaussian distribution with power 0 dB. In Figure 2, the dimension is $p = 60$ and the sample size ranges from 20 to 120. In Figure 3, the sample size is $n = 60$ and the dimension ranges from 20 to 120. Our observations and analyses are summarized as follows:

- (1) Even though enjoying the lowest computation cost, the classic covariance estimator SCM has an unsatisfactory performance in small sample size scenarios. Therefore, it is not an ideal covariance matrix estimator any more in these scenarios.
- (2) The SINR based on each covariance matrix estimator increases when the sample size gets larger but

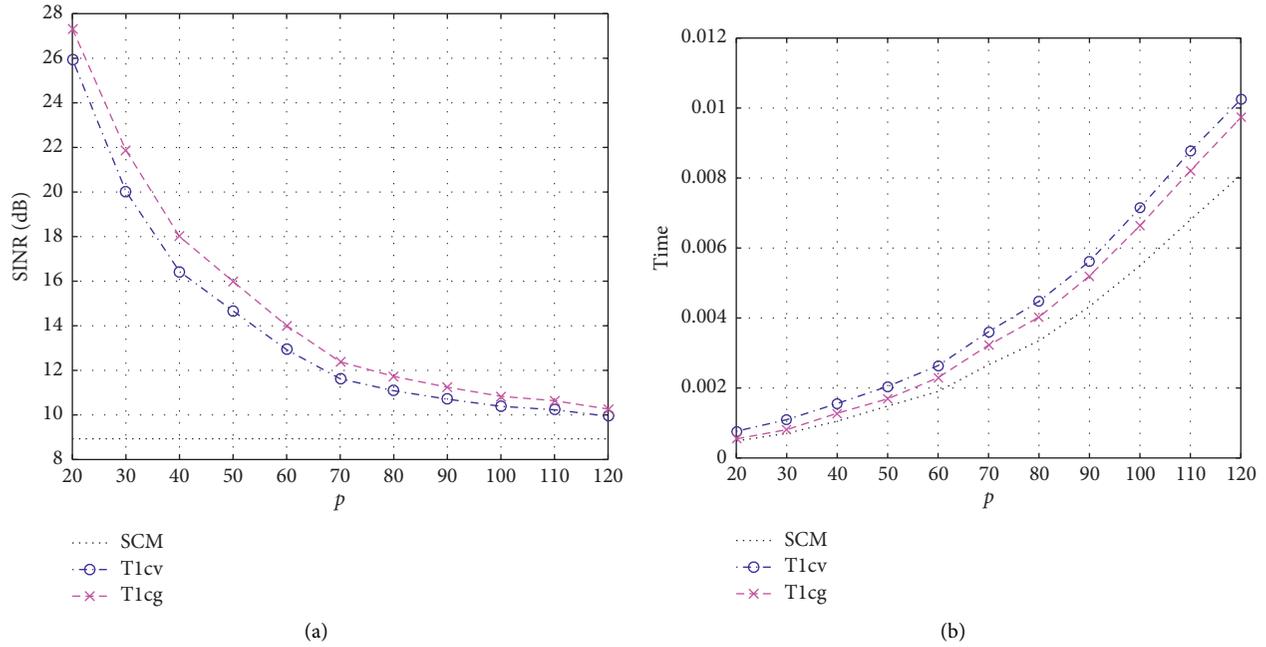


FIGURE 3: The SINR (a) and corresponding elapsed time (b) of adaptive beamformer based on each covariance matrix estimator versus the dimension under the complex Gaussian distribution.

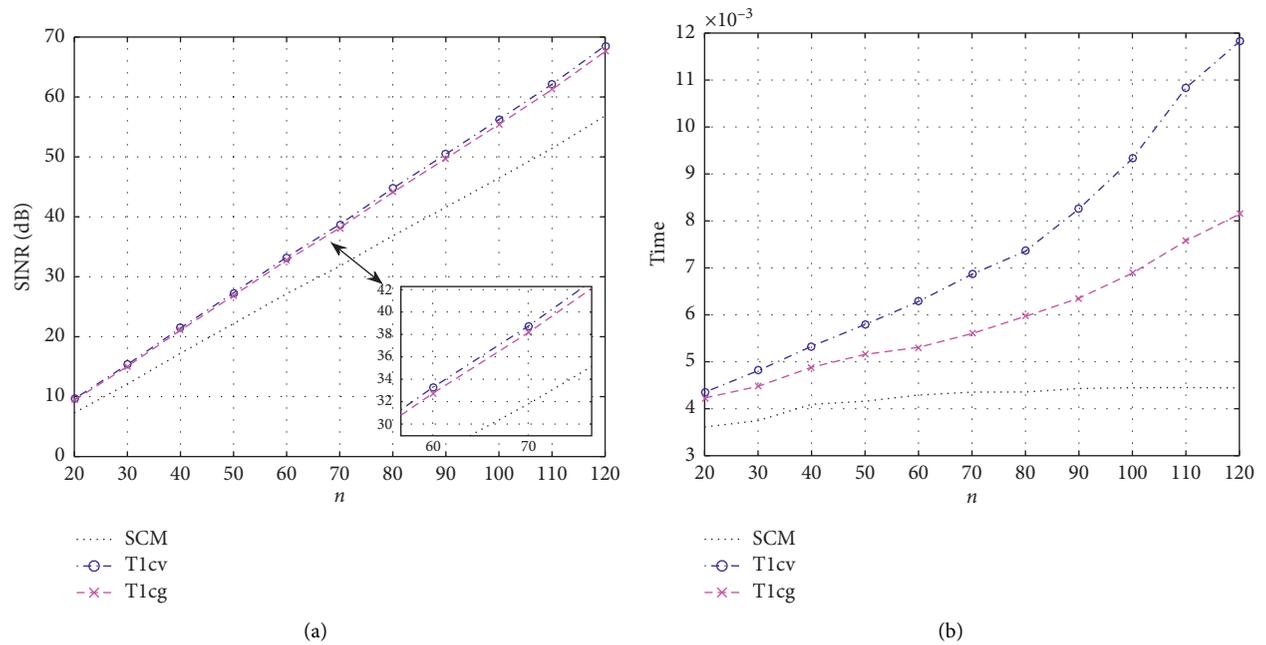


FIGURE 4: The SINR (a) and corresponding elapsed time (b) of adaptive beamformer based on each covariance matrix estimator versus the sample size under the complex non-Gaussian distribution.

decreases when the dimension gets larger. It reveals that the covariance matrix estimators play an important role in adaptive beamforming in large dimension and small sample size scenarios.

- (3) Both the proposed shrinkage estimator T1cg and the existing shrinkage estimator T1cv outperform the SCM with an additional but reasonable

computation cost. Furthermore, the proposed T1cg dominates the T1cv on both SINR and computation cost because the signal comes from the complex Gaussian distribution in the simulation setting, and the proposed covariance matrix estimator has considered the specific distribution information.

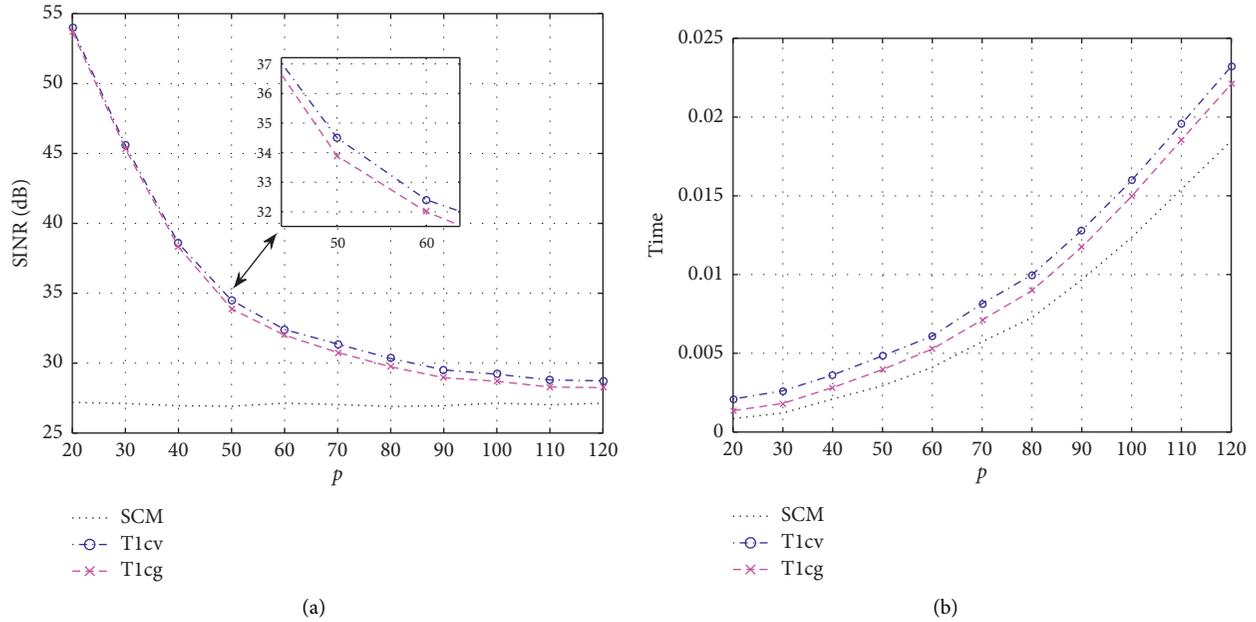


FIGURE 5: The SINR (a) and corresponding elapsed time (b) of adaptive beamformer based on each covariance matrix estimator versus the dimension under the complex non-Gaussian distribution.

Figures 4 and 5 report the SINR and corresponding elapsed time of the adaptive beamformers when each variate in the noise follows the complex Gaussian mixture distribution $0.3\mathcal{CN}(-20, 1) + 0.4\mathcal{CN}(0, 1) + 0.3\mathcal{CN}(20, 1)$ with power 6 dB. We can see that the SINR and elapsed time show the same varying tendencies, as in Figures 2 and 3. When the received signal comes from the complex non-Gaussian distribution, the proposed estimator T1cg performs inferior to the existing T1cv in adaptive beamforming. It is worthy noticing that both T1cg and T1cv have analytical expressions, and there are $(np^2 + np + 3p^2)$ multiplication in the proposed estimators T1cg and $(np^2 + 3np + 5p^2)$ multiplication in T1cv. Therefore, the proposed T1cg can always enjoy a lower computation complexity than T1cv.

On the whole, by employing additional distribution information, the proposed estimator T1cg outperforms T1cv in the complex Gaussian scenario and enjoys a comparable performance with T1cv in the complex non-Gaussian scenario. Moreover, the proposed estimator T1cg always enjoys an advantage over T1cv in computation cost.

5. Conclusion

In this paper, we have proposed a new covariance matrix estimator via linear shrinkage procedure under the complex Gaussian distribution. Through calculating the moment of Wishart distribution, we obtain the optimal shrinkage intensity for the spherical target. Furthermore, the involved unknown scalars are unbiasedly estimated. Subsequently, we propose the corresponding available linear shrinkage estimator. Numerical simulations and application to adaptive beamforming show that the proposed covariance matrix estimator is outperformed compared with the existing estimators. In future work, we will investigate the Cramér–Rao

bound for the linear shrinkage estimation and develop nonlinear shrinkage estimation of the large-dimensional covariance matrix.

Data Availability

All data included in this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interests regarding the publication of this paper.

Acknowledgments

This work was supported in part by the Youth Research Foundation of Hubei Minzu University under Grant MY2017Q021.

References

- [1] J. Bai and S. Shi, “Estimating high dimensional covariance matrices and its applications,” *Annals of Economics and Finance*, vol. 12, no. 2, pp. 199–215, 2011.
- [2] B. D. Van Veen, “Adaptive convergence of linearly constrained beamformers based on the sample covariance matrix,” *IEEE Transactions on Signal Processing*, vol. 39, no. 6, pp. 1470–1473, 1991.
- [3] Z. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, Springer, New York, NY, USA, 2010.
- [4] F. Gini and M. Greco, “Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter,” *Signal Processing*, vol. 82, no. 12, pp. 1847–1859, 2002.

- [5] D. M. Witten and R. Tibshirani, "Covariance-regularized regression and classification for high dimensional problems," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 71, no. 3, pp. 615–636, 2009.
- [6] O. Ledoit and M. Wolf, "Improved estimation of the covariance matrix of stock returns with an application to portfolio selection," *Journal of Empirical Finance*, vol. 10, no. 5, pp. 603–621, 2003.
- [7] T. T. Cai and M. Yuan, "Adaptive covariance matrix estimation through block thresholding," *The Annals of Statistics*, vol. 40, no. 4, pp. 2014–2042, 2012.
- [8] A. J. Rothman, "Positive definite estimators of large covariance matrices," *Biometrika*, vol. 99, no. 3, pp. 733–740, 2012.
- [9] O. Besson and Y. I. Abramovich, "Regularized covariance matrix estimation in complex elliptically symmetric distributions using the expected likelihood approach-Part 2: the under-sampled case," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5819–5829, 2013.
- [10] J. Bien, F. Bunea, and L. Xiao, "Convex banding of the covariance matrix," *Journal of the American Statistical Association*, vol. 111, no. 514, pp. 834–845, 2016.
- [11] T. J. Fisher and X. Sun, "Improved Stein-type shrinkage estimators for the high-dimensional multivariate normal covariance matrix," *Computational Statistics & Data Analysis*, vol. 55, no. 5, pp. 1909–1918, 2011.
- [12] X. Chen, Z. J. Wang, and M. J. McKeown, "Shrinkage-to-tapering estimation of large covariance matrices," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5640–5656, 2012.
- [13] J. Schäfer and K. Strimmer, "A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics," *Statistical Applications in Genetics and Molecular Biology*, vol. 4, no. 1, pp. 1–32, 2009.
- [14] Y. Ikeda, T. Kubokawa, and M. S. Srivastava, "Comparison of linear shrinkage estimators of a large covariance matrix in normal and non-normal distributions," *Computational Statistics & Data Analysis*, vol. 95, pp. 95–108, 2016.
- [15] Y. Konno, "Shrinkage estimators for large covariance matrices in multivariate real and complex normal distributions under an invariant quadratic loss," *Journal of Multivariate Analysis*, vol. 100, no. 10, pp. 2237–2253, 2009.
- [16] Y. Liu, X. Sun, and S. Zhao, "A covariance matrix shrinkage method with Toeplitz rectified target for DOA estimation under the uniform linear array," *AEÜ-International Journal of Electronics and Communications*, vol. 81, pp. 50–55, 2017.
- [17] 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. 116 785–116 798, 2019.
- [18] J. Tong, R. Hu, J. Xi, Z. Xiao, Q. Guo, and Y. Yu, "Linear shrinkage estimation of covariance matrices using low-complexity cross-validation," *Signal Processing*, vol. 148, pp. 223–233, 2018.
- [19] Y. Chen, A. Wiesel, Y. C. Eldar, and A. O. Hero, "Shrinkage algorithms for MMSE covariance estimation," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5016–5029, 2010.
- [20] A. Hannart and P. Naveau, "Estimating high dimensional covariance matrices: a new look at the Gaussian conjugate framework," *Journal of Multivariate Analysis*, vol. 131, no. 131, pp. 149–162, 2014.
- [21] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of Multivariate Analysis*, vol. 88, no. 2, pp. 365–411, 2004.
- [22] J. A. Tague and C. I. Caldwell, "Expectations of useful complex Wishart forms," *Multidimensional Systems and Signal Processing*, vol. 5, no. 3, pp. 263–279, 1994.
- [23] X. Mestre and M. A. Lagunas, "Finite sample size effect on minimum variance beamformers: optimum diagonal loading factor for large arrays," *IEEE Transactions on Signal Processing*, vol. 54, no. 1, pp. 69–82, 2006.
- [24] J. Serra and M. Nájar, "Asymptotically optimal linear shrinkage of sample LMMSE and MVDR filters," *IEEE Transactions on Signal Processing*, vol. 62, no. 14, pp. 3552–3564, 2014.