SCIREA Journal of Mathematics
http://www.scirea.org/journal/Mathematics
February 9, 2023
Volume 8, Issue 1, February 2023
https://doi.org/10.54647/mathematics110384

Unbiased estimators of two second-order moments of the covariance matrix

Xuanci Wang[1], Wei Yi[2], Bin Zhang[1,*]

[1] College of Mathematics and Statistics, Guangxi Normal University, Guilin, Guangxi 541004, China

[2] Xiamen Second Foreign Language School, Xiamen, Fujian 361100, China

wangxc0915@126.com; weiyi2020vv@163.com; binzhang@gxnu.edu.cn

*Corresponding author: B. Zhang (binzhang@gxnu.edu.cn).

Abstract

The covariance matrix, employed for measuring the linear correlation between variables, plays a vital role in data analysis, such as statistical prediction and hypothesis testing. When the data dimension is high, the traditional sample covariance matrix is not an ideal estimator of the population covariance matrix anymore, resulting in degradation or even inaccuracy of the second-order moment estimators' performance based on the sample covariance matrix. This paper studies the unbiased estimators of the second-order moments under complex Gaussian distribution. The proposed unbiased estimators have better statistical properties and numerical performance than the existing estimation methods.

Keywords: Covariance matrix, second-order moments, complex Gaussian distribution,

unbiased estimators.

I. Introduction

With the wide application of large-scale array technology and the popularization of 5G, the need for high-dimensional data analysis in fields including communication and information is increasing[1]–[4]. The following reasons lead to this phenomenon. On the one hand, more array elements can obtain higher resolution, more vital interference suppression ability, and longer detection distance. On the other hand, limited by the hardware level and external environment, it takes a large sample size in the array processing to maintain the corresponding performance of the device. Nevertheless, in fields such as radar signals, it is often difficult to match the number of samples with the number of array elements, leading to a high-dimensional small sample problem in signal statistical analysis. When the sample size is small, the performance of traditional signal estimation algorithms will rapidly decrease as the number of dimensions increases. The main reason is that the signal estimation algorithm often involves the covariance matrix between array elements [5], [6]. In practice, the covariance matrix is often unknown [7], and it needs to be estimated with limited samples [8]. When the dimension is low and the sample size is large, the sample covariance matrix is a good estimator of the population covariance matrix. However, in the case of high-dimensional small samples, the sample covariance matrix becomes ill-conditioned, even singular. Its error becomes non-negligible, causing the corresponding algorithm to be unstable or unusable [9], [10]. A natural processing method uses limited samples to improve the estimation accuracy of the population covariance matrix, which can improve the algorithm's performance. Therefore, estimating the covariance matrix in the case of high-dimensional small samples has received extensive attention [11], [12]. In practice, people often pay more attention to estimating its population moment than the covariance matrix itself. For example, when testing the spherical structure of the covariance matrix, estimating the second-order population moment is the key to constructing the test statistics [13], [14]. Besides, the theoretical optimal shrinkage coefficient needs to be brought into the second-order moment estimator of the covariance matrix to become usable [15].

Since the sample covariance matrix is no longer suitable for high-dimensional small-sample situations, it is no longer reliable to estimate the second-order moment based on the sample covariance matrix. When the dimension tends to infinity, the sample-based second-order moment estimator of the covariance matrix is not a consistent estimate of the second-order population moment. Meanwhile, it is a biased estimator. Therefore, people need to find a better second-order moment estimator. One way is to find a better covariance matrix estimator and then calculate the second-order moment. We then obtain the second-order moment estimation. The estimators obtained by this method are usually consistent, but it takes effort to guarantee unbiasedness. Another method directly estimates the existing second-order moment based on the covariance matrix. The unbiased second-order moment estimation is obtained by correcting the deviation. When the data comes from the real number field, the existing literature has obtained the unbiased estimation of the second-order population moment [16]. However, the data often comes from the complex number field in signal processing. Its distribution is different from the case of the real number field [17]. In this paper, we assume that the data obeys the complex Gaussian distribution and study the unbiased estimator of the second-order moment of the covariance matrix.

II. Estimation of the second-order moment under complex Gaussian distribution

Let the random variable $x \in C^p$ obey the complex Gaussian distribution $CG(\mu, \Sigma)$, where μ , Σ are the expectation and population covariance matrix respectively, $x_1, x_2, \cdots, x_n \in C^p$ is an available sample with a sample size of n . Commonly used second-order moments include $tr^2(\Sigma), tr(\Sigma^2)$, where $tr()$ is the trace of the matrix. They play an important role in multivariate statistical theory. In [18], the ratio in the likelihood of covariance matrix sphericity test is

$$r = \frac{tr^2(\Sigma)}{tr(\Sigma^2)}. \tag{1}$$

In [19], the optimal parameter in shrinkage estimation can be expressed as

$$w = \frac{ptr^2(\Sigma) + (p-2)tr(\Sigma^2)}{(p-n)tr^2(\Sigma) + (np+p-2)tr(\Sigma^2)}. \tag{2}$$

Note that B is a p-order symmetric Boolean matrix, and its elements satisfy $b_{ij} = b_{ji} = 0$ or $b_{ij} = b_{ji} = 1$. This paper studies the estimation of a wider class of second-order moments, namely $tr^2(B \circ \Sigma)$ and $tr\left[(B \circ \Sigma)^2\right]$, where $\circ$ represents the Hadamard product of the matrix. Obviously, when the elements of B are all equal to 1, we have

$$tr^2(B \circ \Sigma) = tr^2(\Sigma),\ tr\left[(B \circ \Sigma)^2\right] = tr(\Sigma^2). \tag{3}$$

We denote d_Σ as the p-dimensional column vector composed of the diagonal elements of the population covariance matrix and d_Σ^T as its conjugate transpose, we can get

$$tr^2(B \circ \Sigma) = d_\Sigma^T B d_\Sigma. \tag{4}$$

Next, we estimate $tr\left[(B \circ \Sigma)^2\right]$ and $d_\Sigma^T B d_\Sigma$.

Denote the sample covariance matrix as $S = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})^T$, where $\bar{x} = \sum_{i=1}^{n} x_i / n$ is the sample mean. Denote d_S as a p-dimensional column vector composed of diagonal elements of the sample covariance matrix. In classical statistical theory, the sample covariance matrix S is an unbiased and consistent estimate of Σ, so the estimator based on the sample covariance matrix $tr\left[(B \circ S)^2\right]$ and $d_S^T B d_S$ are used as the estimates of $tr\left[(B \circ \Sigma)^2\right]$ and $d_\Sigma^T B d_\Sigma$ respectively. Although unbiasedness is not satisfied, the two estimators are still consistent when the sample size n is large enough. However, in the high-dimensional small-sample case, $tr\left[(B \circ S)^2\right]$ and $d_S^T B d_S$ are neither unbiased nor consistent. The estimation error becomes larger with the increasing dimension. Therefore, it is necessary to

correct the deviation of the estimation error and then obtain an estimator with better performance.

Theorem 1: Let $x_1, x_2, \cdots, x_n \in C^p$ be the samples coming from the complex Gaussian distribution $CG(\mu, \Sigma)$. S is the sample covariance matrix. For any $B = \{b_{ij} : b_{ij} = b_{ji} = 1 \text{ or } b_{ij} = b_{ji} = 0,\ i, j = 1, \cdots, p\}$, we have

$$\begin{aligned} \alpha &= \frac{n}{(n+1)(n-1)}[ntr(B \circ S)^2 - d_S^{\mathrm{T}} B d_S] \\ \beta &= \frac{n}{(n+1)(n-1)}[nd_S^{\mathrm{T}} B d_S - tr(B \circ S)^2] \end{aligned} \tag{5}$$

which are the unbiased and consistent estimators of the second-order moment $tr\left[\left(B \circ \Sigma\right)^2\right]$ and $d_\Sigma^T B d_\Sigma$.

Proof 1: Denote $S = \left(s_{ij}\right)_{p\times p}$, $\Sigma = \left(\sigma_{ij}\right)_{p\times p}$ and $B = \left(b_{ij}\right)_{p\times p}$. According to the property of the complex Wishart distribution, we have

$$\begin{aligned} E[tr(B \circ S)^2] &= E(\sum_{i,j=1}^{p} b_{ij}^2 s_{ij} s_{ji}) = \sum_{i,j=1}^{p} \frac{1}{n} b_{ij}^2 \sigma_{ii} \sigma_{jj} + \sum_{i,j=1}^{p} b_{ij}^2 \sigma_{ij} \sigma_{ji} \\ &= \frac{1}{n} d_\Sigma^T B d_\Sigma + tr(B \circ \Sigma)(B \circ \Sigma)^T = \frac{1}{n} d_\Sigma^T B d_\Sigma + tr(B \circ \Sigma)^2, \end{aligned} \tag{6}$$

and

$$\begin{aligned} E[d_S^T B d_S] &= E(\sum_{i,j=1}^{p} b_{ij} s_{ii} s_{jj}) = \sum_{i,j=1}^{p} \frac{1}{n} b_{ij} \sigma_{ij} \sigma_{ji} + \sum_{i,j=1}^{p} b_{ij} \sigma_{ii} \sigma_{jj} \\ &= d_\Sigma^T B d_\Sigma + \frac{1}{n} tr(B \circ \Sigma)(B \circ \Sigma)^T = d_\Sigma^T B d_\Sigma + \frac{1}{n} tr(B \circ \Sigma)^2. \end{aligned} \tag{7}$$

Combining the equation (6) and (7), we have

$$tr[(B \circ \Sigma)^2] = \frac{n}{(n+1)(n-1)} E[ntr(B \circ S)^2 - d_S^T B d_S], \tag{8}$$

$$d_{\Sigma}^{T}Bd_{\Sigma}=\frac{n}{(n+1)(n-1)}E[nd_{S}^{T}Bd_{S}-tr(B\circ S)^{2}]. \tag{9}$$

Therefore, α and β are the unbiased estimators of $tr[(B\circ\Sigma)^{2}]$ and $d_{\Sigma}^{T}Bd_{\Sigma}$. Besides, by the law of large numbers, when $n\to\infty$, we have $E(\alpha)\xrightarrow{p}tr[(B\circ\Sigma)^{2}]$ and $E(\beta)\xrightarrow{p}d_{\Sigma}^{T}Bd_{\Sigma}$, which shows that α, β are the consistent estimators of the second-order moment.

According to Theorem 1, the estimation deviations of estimators $tr[(B\circ S)^{2}]$ and $d_{S}^{T}Bd_{S}$ based on sample covariance matrix are respectively as follows

$$\varepsilon_{1}=\alpha-tr[(B\circ S)^{2}]=\frac{1}{(n+1)(n-1)}[tr(B\circ S)^{2}-nd_{S}^{T}Bd_{S}], \tag{10}$$

$$\varepsilon_{2}=\beta-d_{S}^{T}Bd_{S}=\frac{1}{(n+1)(n-1)}[d_{S}^{T}Bd_{S}-ntr(B\circ S)^{2}]. \tag{11}$$

When $B=\{b_{ij}:b_{ij}=b_{ji}=1,\ i,j=1,\cdots,p\}$, the unbiased estimators of $tr(\Sigma^{2})$ and $tr^{2}(\Sigma)$ are

$$\alpha=\frac{n}{(n+1)(n-1)}[ntr(S^{2})-tr^{2}(S)],\quad \beta=\frac{n}{(n+1)(n-1)}[ntr^{2}(S)-tr(S^{2})] \tag{12}$$

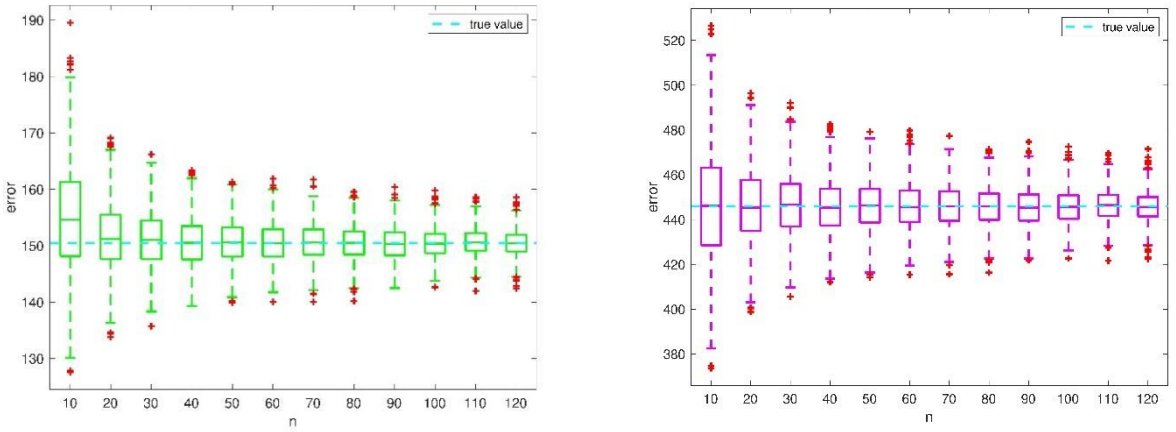
It is consistent with the results in [20]. According to Theorem 1, the estimation errors of the estimators $tr(S^{2})$ and $tr^{2}(S)$ based on the sample covariance matrix are as follows

$$\varepsilon_{1}=\alpha-tr(S^{2})=\frac{1}{(n+1)(n-1)}[tr(S^{2})-ntr^{2}(S)] \tag{13}$$

$$\varepsilon_{2}=\beta-d_{S}^{T}Bd_{S}=\frac{1}{(n+1)(n-1)}[tr^{2}(S)-ntr(S^{2})] \tag{14}$$

III. Numerical Simulation

In this section, we verify the performance of the second-order moment estimator proposed in Theorem1 through numerical simulation. In the experiment, we set the data $x_1, x_2, \cdots, x_n$ from complex Gaussian distribution with mean 0. And the population covariance is $\Sigma = (\sigma_{ij})$, where $\sigma_{ij} = t|i-j|$. The dimension is $p = 150$. The Boolean matrix is a symmetric matrix with a bandwidth of 70, that is, $B = \{b_{ij} : b_{ij} = b_{ji} = 0, if\ |i-j| \geq 70;\ b_{ij} = b_{ji} = 1, if\ |i-j| < 70;\}$. When $t = 0.2$, the real values of the second-oder moments $tr\left[(B \circ \Sigma)^2\right]$ and $d_\Sigma^T B d_\Sigma$ are 150.4736 and 446. When $t = 0.8$, the true values of the second-order moments are 271.2416 and 446, respectively.

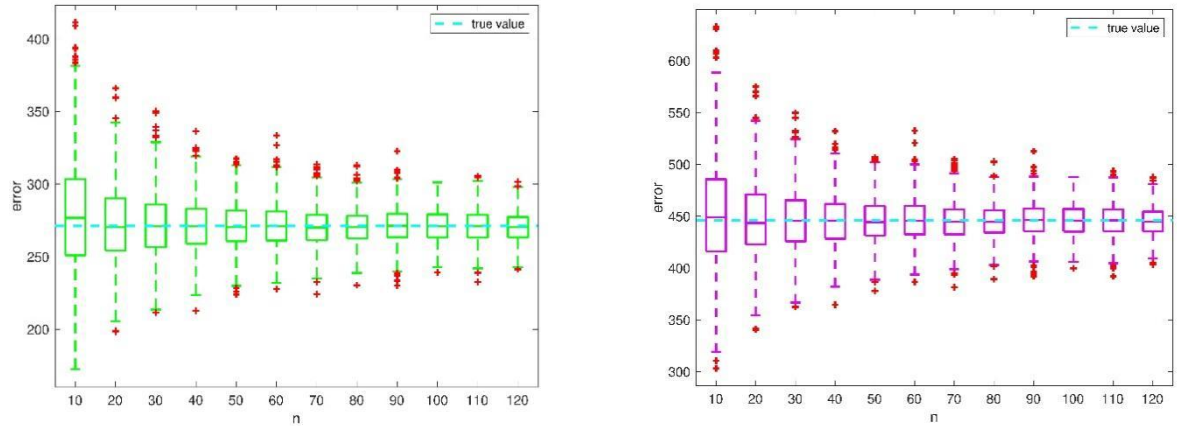


(a) **(b)**

Figure 1. When $t = 0.2$, the box plots of α and β.

Figure 1 describes the values of estimators α and β under different sample sizes. From Figure 1(a), we can see that the average value of $tr\left[(B \circ \Sigma)^2\right]$ is very close to the real value of the second-order moment 150.4736. Figure 1(b) is the estimated result of $d_\Sigma^T B d_\Sigma$. The real value is 446. It can be seen from Figure 1 that although the values of n are different, the median always fluctuates slightly around the real value. Moreover, as the sample size increases,

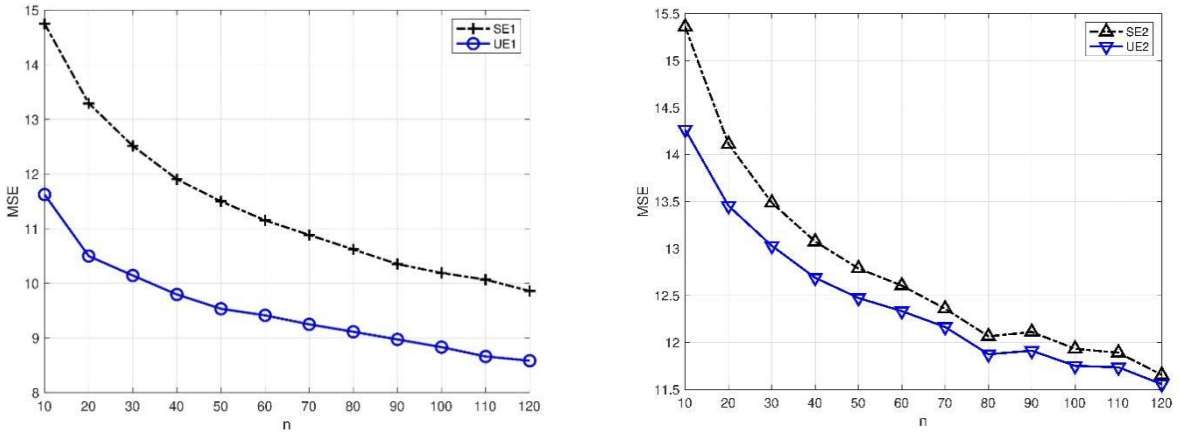
the estimation error decreases significantly.



(a) **(b)**

Figure 2. When $t = 0.8$, the box plots of α and β.

In Figure 2, we increase the correlation coefficient t to 0.8 to verify the effectiveness under a strong correlation. From the results of the box plot in 2(a), we can see that when n takes different values, the medians of $tr\left[\left(B \circ \Sigma\right)^2\right]$ are all around the actual value 271.2416. Moreover, the variances also decrease rapidly as the sample size increases. The results shown in Figure 2(b) are similar to those shown in Figure 1(b). Similarly, the estimated mean of $d_\Sigma^T B d_\Sigma$ is always close to the actual value of 446. From Figure 1(a)–Figure 2(b), it can be concluded that the estimators of the two types of second-order moments proposed in this paper satisfy the unbiasedness. In addition, we compare the proposed unbiased estimator with the existing estimator based on the sample covariance matrix. Figure 3 reflects the changing trend of the mean square error of the two types of estimators as the sample size increases.



(a) **(b)**

Figure 3. The mean square error of the second-order moment estimator.

As shown in Figure 3(a), SE1 is the estimated mean square error based on the sample covariance matrix, and UE1 is the estimated mean square error of α. In Figure 3(b), SE2 represents the estimated mean square error of the existing estimator, while UE2 is the estimated mean square error of β. It is easy to see that the mean square errors of the two types of estimators are relatively large in the case of small high-dimensional samples. As the sample size increases, both mean square errors decrease. It shows that the mean square error of the estimator is greatly affected by the dimension and sample size. Furthermore, the mean square error of the second-order moment estimator proposed in this paper is significantly smaller than that of the existing estimator based on the sample covariance matrix mean square error. In summary, the second-order population moment estimator proposed in this paper conforms to the statistical characteristics of unbiasedness and consistency. Its performance is significantly better than the existing estimators in the case of high-dimensional small samples. In addition, the second-order population estimator proposed in this paper is suitable for any symmetric Boolean matrix and has a wide range of application prospects.

IV. Conclusion

This paper has studied the problem of estimating the second-order moments related to the

covariance matrix in high-dimensional data analysis. The existing estimation methods were directly substituted into the sample covariance matrix. In this paper, using the statistical properties of the complex Wishart distribution, two types of unbiased estimation of the second-order population moment were carried out, and the deviations of the existing estimators were calculated. Numerical simulations showed that, compared with the existing estimators, the estimators proposed in this paper have the advantages of unbiasedness and a minor mean square error. These estimators could be used to construct a new likelihood ratio test statistic and improve the estimation performance of the corresponding algorithm.

Funding

This work was funded by the Guangxi Science and Technology Planning Project (2022AC21276) and the Science and Technology Project of Guangxi Guike (AD21220114).

Reference

[1] A. Jukić, T. van Waterschoot, T. Gerkmann, and S. Doclo, “Multi-channel linear prediction-based speech dereverberation with sparse priors,” IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 23, no. 9, pp. 1509–1520, 2015.

[2] Z. Li, J. Liu, S. Zhang, Y. Yu, H. Liang, Q. Lu, J. Chen, Y. Han, F. Zhang, and J. Li, “Novel potential metabolic biomarker panel for early detection of severe COVID-19 using full-spectrum metabolome and whole-transcriptome analyses,” Signal Transduction and Targeted Therapy, vol. 7, p. 129, 2022.

[3] Y. Xiao, D. Ma, Y. Yang, F. Yang, J. Ding, Y. Gong, L. Jiang, L. Ge, S.-Y. Wu, Q. Yu, Q. Zhang, F. Bertucci, Q. Sun, X. Hu, D. Li, Z. Shao, and Y. Jiang, “Comprehensive metabolomics expands precision medicine for triple-negative breast cancer,” Cell Research, vol. 32, p. 477–490, 2015.

[4] E. Ollila and A. Breloy, “Regularized tapered sample covariance matrix,” IEEE Transactions on Signal Processing, vol. 70, no. 1, pp. 2306–2320, 2022.

[5] A. Moore, S. Hafezi, R. Vos, P. Naylor, and M. Brookes, “A compact noise covariance matrix model for MVDR beamforming,” IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 30, pp. 1–14, 01 2022.

[6] D. Luong, B. Balaji, and S. Rajan, “Structured covariance matrix estimation for noise-type radars,” IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1–13, 2022.

[7] K. W. Chan, “Mean-structure and autocorrelation consistent covariance matrix estimation,” Journal of Business & Economic Statistics,vol. 40, no. 1, pp. 201–215, 2022.

[8] E. Raninen and E. Ollila, “Bias adjusted sign covariance matrix,” IEEE Signal Processing Letters, vol. 29, pp. 339–343, 2022.

[9] W. WU and M. Pourahmadi, “Nonparametric estimation of large covariance matrices of longitudinal data,” Biometrika, vol. 90, no. 4, p. 831–844, 2003.

[10] P. J. Bickel and E. Levina, “Covariance regularization by thresholding,” The annals of statistics, vol. 36, no. 6, pp. 2577–2604, 2008.

[11] R. T. Bengtsson, “Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants,” Journal of Multivariate Analysis, vol. 98, no. 2, pp. 227–255, 2007.

[12] 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.

[13] M. Sumair, T. Aized, M. M. Aslam Bhutta, F. A. Siddiqui, L. Tehreem, and A. Chaudhry, “Method of four moments mixture-a new approach for parametric estimation of Weibull probability distribution for wind potential estimation applications,” Renewable Energy, vol. 191, pp. 291–304, 2022.

[14] X. Ding, “Some sphericity tests for high dimensional data based on ratio of the traces of sample covariance matrices,” Statistics & Probability Letters, vol. 156, p. 108613, 2020.

[15] 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.

[16] B. Zhang, “Improved shrinkage estimator of large-dimensional covariance matrix under the complex Gaussian distribution,” Mathematical Problems in Engineering, vol. 2020, pp. 1–8, 07 2020.

[17] J. A. Tague and C. Caldwell, “Expectations of useful complex Wishart forms,” Multidimensional Systems and Signal Processing, vol. 5, pp. 263–279, 1994.

[18] J. Li, J. Zhou, and B. Zhang, “Estimation of large covariance matrices by shrinking to structured target in normal and non-normal distributions,” IEEE Access, vol. PP, pp. 1–1, 12 2017.

[19] X. Tian, Y. Lu, and W. Li, “A robust test for sphericity of high-dimensional covariance matrices,” Journal of Multivariate Analysis, vol. 141, pp. 217–227, 2015.

[20] M. S. Srivastava, “Some tests concerning the covariance matrix in high dimensional data,” Journal of the Japan Statistical Society. Japanese issue, vol. 35, pp. 251–272, 2005.