Shrinkage estimators of large covariance matrices with Toeplitz targets in array signal processing

Bin Zhang1 & Shoucheng Yuan2*

The problem of estimating a large covariance matrix arises in various statistical applications. This paper develops new covariance matrix estimators based on shrinkage regularization. Individually, we consider two kinds of Toeplitz-structured target matrices as the data come from the complex Gaussian distribution. We derive the optimal tuning parameter under the mean squared error criterion in closed form by discovering the mathematical properties of the two target matrices. We get some vital moment properties of the complex Wishart distribution, then simplify the optimal tuning parameter. By unbiasedly estimating the unknown scalar quantities involved in the optimal tuning parameter, we propose two shrinkage estimators available in the large-dimensional setting. For verifying the performance of the proposed covariance matrix estimators, we provide some numerical simulations and applications to array signal processing compared to some existing estimators.

The covariance matrix, which is a fundamental quantity for quantifying the variance and covariance among the variables, plays an essential role in both multivariate statistics and practical data processing1,2. In general, people can hardly know the exact value of the covariance matrix. An overall experience is to find a well-conditioned covariance matrix estimator by the collected data3–5. In classic statistical theory with a low dimension and large sample size, the sample covariance matrix (SCM) is widely known as one standard estimator of the true covariance matrix. It enjoys lots of favorable statistical properties such as unbiasedness and consistency. Unfortunately, it suffers from either the small sample size or the large-dimension framework and needs to be improved6,7.

During the last two decades, estimating the covariance matrix in large-dimension scenarios has attracted tremendous attention. Numerous regularization strategies have been designed to generate an high-performance covariance matrix estimator8–13. Among these, the linear shrinkage estimation is a widely-adopted regularization strategy to result in a well-conditioned covariance matrix estimator14–17. In detail, the shrinkage estimator consists of the SCM and a structured target matrix that represents prior information on the covariance structure. By carefully choosing the involved tuning parameter, also called the shrinkage intensity, the corresponding shrinkage estimator can significantly improve the SCM18–20. Although owning an analytical expression under the mean squared error (MSE) criterion, the optimal shrinkage estimator is unavailable in the practical data processing. The reason is the loss function in the MSE criterion evaluates the mean Frobenius distance between the true covariance matrix and the shrinkage estimator, resulting in the optimal shrinkage intensity containing some unknown scalar quantities concerning the true covariance matrix and the expectation operator. A direct approach to turning the optimal shrinkage estimator into an available one is the plug-in strategy, which replaces the unknown scalar quantities with their estimates21–23. It is not hard to understand that good estimates of the unknown scalar quantities will benefit both the corresponding available shrinkage intensity and the forthcoming available covariance matrix estimator24,25.

A large proportion of shrinkage estimators is developed for dealing with large-dimensional data from the real number field. In14, the optimal shrinkage intensity is obtained for the spherical target. The concerned unknown scalar quantities are consistently estimated with small biases under the distribution-free setting. In22,24, the unknown scalar quantities are respectively unbiasedly estimated under Gaussian and non-Gaussian distributions to develop shrinkage estimators with better performance. Furthermore, for the diagonal target and the tapered SCM target, the unknown scalar quantities in the corresponding optimal shrinkage intensities are respectively unbiasedly estimated to obtain high-performance covariance matrix estimators. What is worth mentioning is that the unknown scalar quantities corresponding to different target matrices are different. Moreover, estimating the unknown scalar quantities usually becomes difficult for the complicated target matrix. In26, for two kinds of Toeplitz-structured targets, the unknown scalar quantities are respectively unbiasedly estimated under

1College of Mathematics and Statistics, Guangxi Normal University, Guilin 541004, Guangxi, China. 2College of Mathematics and Statistics, Pu’er University, Pu’er 665000, Yunnan, China. *email: ysc-peace@163.com
Gaussian and non-Gaussian distributions. Furthermore, the shrinkage estimation is investigated under elliptically contoured distribution. In signal processing, the data usually comes from the complex number field. The covariance matrix is conjugate symmetric. The primary difference from the real number field is that the SCM with a multiplier follows the complex Wishart distribution. In the large sample scenario, however, the estimates of the involved are biased, leading a poor performance in a small sample size setting. In a non-parametric shrinkage estimator is developed via a proxy MSE under the complex Gaussian distribution. The advantage of this method is applying to any target matrix. However, the non-parametric shrinkage estimator is not optimal in the sense of MSE. In the spherical target is established.

This paper investigates the shrinkage estimators or the two kinds of Toeplitz-structured targets under the complex Gaussian distribution. The main contributions are summarized as four-fold.

1. The optimal shrinkage intensities for the two kinds of Toeplitz-structured targets are obtained in the closed form under the MSE criterion.
2. We provide a useful moment property of the complex Wishart distribution for unbiasedly estimating the unknown scalar quantities involved in the optimal shrinkage intensities.
3. All the involved unknown scalar quantities are unbiasedly estimated under the complex Gaussian distribution. Then the corresponding available shrinkable estimators are immediately analytically expressed.
4. We provide some numerical simulations and applications to array signal processing to reveal the advantages of proposed covariance matrix estimators.

The remainder of this paper is as follows: In "Formulation", we retrospect the shrinkage estimation of the covariance matrix as a quadratic optimization problem. "The optimal shrinkage intensity for Toeplitz-structured targets" provides the optimal shrinkage intensities and corresponding linear shrinkage estimators for the two kinds of Toeplitz-structured targets under complex Gaussian distribution. "Available shrinkage estimator forToeplitz-structured targets" unbiasedly estimates the unknown scalar quantities involved in the optimal shrinkage estimators and proposes newly available shrinkage estimators. "Methods" provides some numerical simulations and two practical applications for exhibiting the advantages of proposed covariance matrix estimators. Section "Conclusion" concludes.

Notations. The notation \( \mathbb{R} \) represents the set of all real numbers. 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 symbol \( \mathbf{x} \) the maximum and minimum of \( b \). For two positive integers \( a, b \), \( a \wedge b \) and \( a \vee b \) respectively denote its inverse and trace. For two real numbers \( a \) and \( b \), \( a \wedge b \) respectively mean the maximum and minimum of \( a \) and \( b \). For two positive integers \( i, j \), \( \delta_{ij} \) denotes the indicator function being 1 as \( i = j \) or 0 otherwise.

Formulation

Let \( x_1, x_2, \ldots, x_n \in \mathbb{C}^p \) be an independent and identically distributed (i.i.d.) data drawn from the complex Gaussian distribution \( \mathcal{CN}(\mathbf{0}, \Sigma) \) with mean \( \mathbf{0} \) and unknown covariance matrix \( \Sigma \). The task is estimating \( \Sigma \) by the available data \( \mathbf{x} \). As is known to all, the SCM \( \Sigma = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^H \) is one of standard covariance matrix estimators in the large sample scenario. However, it suffers from the curse of dimensionality in modern statistics. Furthermore, linear shrinkage regularization is a practical approach to obtaining a covariance matrix estimator with lower variance. Given a target matrix \( \mathbf{T} \in \mathbb{H}^p \) represents an aspect prior information of covariance structure, the corresponding linear shrinkage estimator is expressed as a weighted combination between the SCM and the target, then we have

\[
\hat{\Sigma} = (1 - w) \Sigma + w \mathbf{T} \in \mathbb{H}^p, \tag{1}
\]

where \( w \in \mathbb{R} \), the so-called shrinkage intensity, is an undetermined tuning parameter. For tuning the optimal parameter, we employ the MSE loss function

\[
\mathcal{M}_T(w) = \mathbb{E}[\| \hat{\Sigma} - \Sigma \|^2] = \mathbb{E}[\| (1 - w) \Sigma + w \mathbf{T} - \Sigma \|^2]. \tag{2}
\]

Through a simple transformation, we have

\[
\mathcal{M}_T(w) = \mathbb{E}[(\text{tr}(\Sigma - \mathbf{T}))^2]w^2 - 2\mathbb{E}[\text{tr}(\Sigma - \mathbf{T})(\Sigma - \mathbf{T})]w + c, \tag{3}
\]

where \( c = \mathbb{E}[(\text{tr}(\Sigma - \mathbf{T}))^2] \) is a constant. Hence, we can obtain the optimal shrinkage intensity in (3) by solving the following convex quadratic problem:

\[
\begin{align*}
\min \quad & \mathbb{E}[\text{tr}(\Sigma - \mathbf{T})^2]w^2 - 2\mathbb{E}[\text{tr}(\Sigma - \mathbf{T})(\Sigma - \mathbf{T})]w + c \\
\text{s.t.} \quad & 0 \leq w \leq 1.
\end{align*}
\tag{4}
\]

Then, the optimal shrinkage intensity for an arbitrary target matrix \( \mathbf{T} \) is
\[ w^* = 0 \land \frac{\mathbb{E}[\text{tr}(T - S) (\Sigma - S)]}{\mathbb{E}[\text{tr}(T - S)^2]} \lor 1. \]  

(5)

### The optimal shrinkage intensity for Toeplitz-structured targets

Let \( L_p \) be the \( p \times p \) matrix where the elements below the main diagonal equal 1 and 0 otherwise. For \( q = -(p-1), \ldots, -1, 1, \ldots, (p-1) \), let \( J_q \) be the \( p \times p \) matrix where the elements of \( q \)-th diagonal above (for \( q > 0 \)) or below (for \( q < 0 \)) main diagonal are equal to 1 and 0 otherwise. Then, the two types of Toeplitz-structured targets can be expressed as

\[
T_1 = \frac{\text{tr}(S)}{p} I_p + \frac{2 \text{tr}(S L_p)}{p(p-1)} L_p^{H} S + \frac{2 \text{tr}(S L_p^{H})}{p(p-1)} L_p, \\
T_2 = \sum_{q=-\infty}^{p-1} \frac{\text{tr}(J_q)}{p-|q|} J_q. 
\]

(6)

(7)

Both \( T_1 \) and \( T_2 \) are Hermitian and Toeplitz-structured matrices. Furthermore, they have the following property.

**Proposition 1** For \( i = 1, 2 \), the following equality about the target \( T_i \) and the SCM holds:

\[
\text{tr}(T_i S) = \text{tr}(T_i^2). 
\]

(8)

**Proof** When the target matrix is \( T_1 \), we have

\[
\text{tr}(T_1 S) = \text{tr} \left( \frac{\text{tr}(S)}{p} I_p + \frac{2 \text{tr}(S L_p)}{p(p-1)} L_p^{H} S + \frac{2 \text{tr}(S L_p^{H})}{p(p-1)} L_p \right). 
\]

(9)

and

\[
\text{tr}(T_1^2) = \text{tr} \left( \frac{\text{tr}(S)}{p} I_p + \frac{2 \text{tr}(S L_p)}{p(p-1)} L_p^{H} S + \frac{2 \text{tr}(S L_p^{H})}{p(p-1)} L_p \right)^2. 
\]

(10)

Moreover, for the matrix \( L_p \), we have \( \text{tr}(L_p) = 0 \), \( \text{tr}(L_p^{H}) = 0 \), \( \text{tr}(L_p L_p) = 0 \), \( \text{tr}(L_p^{H} L_p) = 0 \) and \( \text{tr}(L_p^{H} L_p) = \frac{p(p-1)}{2} \).

By substituting these equalities about \( L_p \) into (9) and (10), we can obtain

\[
\text{tr}(T_1^2) = \frac{1}{p} \text{tr}(S) + \frac{4}{p(p-1)} \text{tr}(S L_p) \text{tr}(S L_p^{H}) + \frac{4}{p(p-1)} \text{tr}(S L_p) \text{tr}(S L_p^{H}) \text{tr}(L_p L_p) + \frac{4}{p(p-1)} \text{tr}(S) \text{tr}(S L_p^{H}) \text{tr}(L_p L_p) + \frac{8}{p(p-1)} \text{tr}(S L_p) \text{tr}(S L_p^{H})\text{tr}(L_p L_p). 
\]

(11)

For the target matrix \( T_2 \), we have

\[
\text{tr}(T_2^2) = \text{tr}(T_2 S) = \sum_{q=-\infty}^{p-1} \frac{1}{p-|q|} \text{tr}(J_q) \text{tr}(J_q^{H}). 
\]

(12)

Therefore, the equality (8) holds for \( T_1 \) and \( T_2 \).

**Remark 1** It is not hard to verify that the relationship given by (8) also holds for the spherical target \( T_3 = \frac{\text{tr}(S)}{p} I_p \) and the diagonal target \( T_4 = \text{diag}(s_{11}, \ldots, s_{pp}) \), which are two popular target matrices in linear shrinkage estimation.

By Proposition 1, when the target is \( T = T_1 \) or \( T = T_2 \), the optimal shrinkage intensity in (5) turns to be

\[
w^* = 0 \land \left( 1 - \frac{\text{tr}(\Sigma^2) - \mathbb{E}[\text{tr}(T \Sigma)]}{\mathbb{E}[\text{tr}(S^2)] - \mathbb{E}[\text{tr}(T^2)]} \right) \lor 1. 
\]

(13)
Furthermore, under complex Gaussian distribution, we have the following results.

**Theorem 2** Assume \( x_1, x_2, \ldots, x_n \in \mathbb{C}^p \) is an i.i.d. sample from \( \mathcal{CN}(0, \Sigma) \), \( A \) is an arbitrary nonrandom \( p \times p \) matrix, \( S \) is the sample covariance matrix, the following relationships hold:

\[
E[\text{tr}(SAS^H)] = \text{tr}(\Sigma A \Sigma A^H) + \frac{1}{n} \text{tr}(\Sigma A)\text{tr}(\Sigma A^H),
\]

(14)

\[
E[\text{tr}(SA)\text{tr}(S^AH)] = \frac{1}{n} \text{tr}(\Sigma A \Sigma A^H) + \text{tr}(\Sigma A)\text{tr}(\Sigma A^H).
\]

(15)

**Proof** Because the i.i.d. sample \( x_1, x_2, \ldots, x_n \in \mathbb{C}^p \) comes from the complex Gaussian distribution \( \mathcal{CN}(0, \Sigma) \), then \( nS \) follows the complex Wishart distribution \( \mathcal{CW}(\Sigma, n) \), where \( n \) is the degree of freedom. For an arbitrary nonrandom \( p \times p \) matrix \( A \), we have

\[
E[SAS^H] = \Sigma A \Sigma A^H + \frac{1}{n} \text{tr}(\Sigma A)\Sigma A^H.
\]

(16)

Then, we can obtain

\[
E[\text{tr}(SAS^H)] = \text{tr}(\Sigma A \Sigma A^H) + \frac{1}{n} \text{tr}(\Sigma A)\text{tr}(\Sigma A^H).
\]

(17)

Next, we compute \( E[\text{tr}(SA)\text{tr}(S^AH)] \) under complex Gaussian distribution. Denote \( A = (a_{ij})_{p\times p} \), then we have \( A^H = (a_{ij})_{p\times p} \). For a random matrix \( W = (w_{ij})_{p\times p} \), we have

\[
\text{tr}(WA) = \sum_{i=1}^{p} \sum_{j=1}^{p} w_{ij} a_{ji}, \quad \text{tr}(WA^H) = \sum_{i=1}^{p} \sum_{j=1}^{p} w_{ij} a_{ij}^*.
\]

Therefore, we have

\[
\text{tr}(WA)\text{tr}(WA^H) = \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{k=1}^{p} \sum_{l=1}^{p} a_{ji} a_{kj}^* w_{ij} w_{kl}.
\]

(18)

When \( W \) follows the complex Wishart distribution \( \mathcal{CW}(I_p, n) \), we have

\[
E[\text{tr}(WA)\text{tr}(WA^H)] = \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{k=1}^{p} \sum_{l=1}^{p} a_{ji} a_{kj}^* E[w_{ij} w_{kl}]
\]

(19)

\[
= \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{k=1}^{p} \sum_{l=1}^{p} a_{ji} a_{kj}^* (b_1 \delta_{ij} \delta_{kl} + b_2 \delta_{ij} \delta_{jk})
\]

(20)

\[
= \sum_{i=1}^{p} \sum_{j=1}^{p} b_1 a_{ji} a_{jk}^* + \sum_{i=1}^{p} \sum_{j=1}^{p} b_2 a_{ji} a_{ij}^*
\]

\[
= b_1 \text{tr}(A)\text{tr}(A^H) + b_2 \text{tr}(AA^H),
\]

where \( \delta_{ij} \) is the indicator function which equals 1 when \( i = j \) or 0 otherwise, \( b_1 \) and \( b_2 \) are two undetermined constants. Furthermore, when \( W \) follows the complex Wishart distribution \( \mathcal{CW}(\Sigma, n) \), we have \( G^{-1}WG^{-H} \sim \mathcal{CW}(I_p, n) \) by denoting \( \Sigma = GG^H \). Hence, for an arbitrary nonrandom \( p \times p \) matrix \( A \), we can obtain

\[
E[\text{tr}(G^{-1}WG^{-H}B)\text{tr}(G^{-1}WG^{-H}B^H)]
\]

\[
= E[\text{tr}(WG^{-H}BG^{-1})\text{tr}(WG^{-H}B^HG^{-1})]
\]

\[
= b_1 \text{tr}(B)\text{tr}(B^H) + b_2 \text{tr}(BB^H).
\]

(21)

Denote \( A = G^{-H}BG^{-1} \), then \( B = G^HAG \). Therefore, we have

\[
E[\text{tr}(WA)\text{tr}(WA^H)] = b_1 \text{tr}(G^HAG)\text{tr}(G^HAG) + b_2 \text{tr}(G^HAGG^HAG^HG)
\]

\[
= b_1 \text{tr}(G^HAG)\text{tr}(G^HAG) + b_2 \text{tr}(G^HAGG^HAG^H).
\]

(22)

Next, we turn to determine the constants \( b_1 \) and \( b_2 \). Noticing that (22) holds for arbitrary \( \Sigma \) and \( A \), we can plug some specific matrices into (22) to obtain \( b_1 \) and \( b_2 \). Firstly, let \( \Sigma = I_p, A = (a_{ij})_{p\times p} \) with \( a_{ij} = 1 \) if \( i = j = 1 \) and \( a_{ij} = 0 \) if else, then we have

\[
E[w_{11}^2] = b_1 + b_2.
\]

(23)
Secondly, let $\Sigma = I_p$ and $A = I_p$, we have

$$
\sum_{i=1}^{p} \sum_{j=1}^{p} E[w_{ii}w_{jj}] = b_1 p^2 + b_2 p.
$$

(24)

For $i = 1, \ldots, p$, $w_{ii}$ is the sum of squares of $2n$ i.i.d. Gaussian random variables with variance $\frac{1}{2}$. Hence, we can obtain that $2w_{ii} \sim \chi^2(2n)$. By the moment property of $\chi^2$ distribution, we have $E[w_{ii}] = n$, $E[w_{ii}^2] = n(n + 1)$. Therefore, we have

$$
\sum_{i=1}^{p} \sum_{j=1}^{p} E[w_{ii}w_{jj}] = \sum_{i=1}^{p} E[w_{ii}^2] + \sum_{i \neq j} E[w_{ii}E[w_{jj}]]
$$

(25)

$$
= \sum_{i=1}^{p} E[w_{ii}^2] + \sum_{i \neq j} E[w_{ii}]E[w_{jj}]
$$

$$
= n(n+1)p + p(p-1)n^2.
$$

Then, the equalities (23) and (24) becomes

$$
b_1 + b_2 = n(n+1),
$$

(26)

$$
b_1 p^2 + b_2 p = n(n+1)p + p(p-1)n^2.
$$

(27)

By solving the above system of equations, we can easily obtain $b_1 = n^2$, $b_2 = n$. Therefore, we have

$$
E[\text{tr}(WA)\text{tr}(WA^H)] = n^2\text{tr}(\Sigma A)\text{tr}(\Sigma A^H) + n\text{tr}(\Sigma A\Sigma A^H).
$$

(28)

Due to $nS \sim CVV(\Sigma, n)$, we have

$$
E[\text{tr}(nSA)\text{tr}(nSA^H)] = n^2\text{tr}(\Sigma A)\text{tr}(\Sigma A^H) + n\text{tr}(\Sigma A\Sigma A^H).
$$

(29)

By dividing $n^2$, we have

$$
E[\text{tr}(SA)\text{tr}(SA^H)] = \frac{1}{n}\text{tr}(\Sigma A\Sigma A^H) + \text{tr}(\Sigma A)\text{tr}(\Sigma A^H).
$$

(30)

It is worth noticing that Theorem 2 holds for arbitrary nonrandom complex matrix $A$. As $A = I_p$, we can obtain

$$
E[\text{tr}(S^2)] = \text{tr}(\Sigma^2) + \frac{1}{n}\text{tr}^2(\Sigma),
$$

(31)

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

(32)

Therefore, the expressions of $E[\text{tr}(S^2)]$ under Gaussian distribution and complex Gaussian distribution are different from each other. The same to $E[\text{tr}^2(S)]$. As $A = I_p$, we can obtain

$$
E[\text{tr}(SL_pS_L^H)] = \text{tr}(\Sigma L_p \Sigma L_p^H) + \frac{1}{n}\text{tr}(\Sigma L_p)\text{tr}(\Sigma L_p^H),
$$

(33)

$$
E[\text{tr}(SL_p)\text{tr}(SL_p^H)] = \frac{1}{n}\text{tr}(\Sigma L_p \Sigma L_p^H) + \text{tr}(\Sigma L_p)\text{tr}(\Sigma L_p^H).
$$

(34)

Moreover, for $q = -(p-1), \ldots, p-1$ and $A = I_q$, we can obtain

$$
E[\text{tr}(SL_qS_q^H)] = \text{tr}(\Sigma I_q \Sigma I_q^H) + \frac{1}{n}\text{tr}(\Sigma I_q)\text{tr}(\Sigma I_q^H),
$$

(35)

$$
E[\text{tr}(SL_q)\text{tr}(SL_q^H)] = \frac{1}{n}\text{tr}(\Sigma I_q \Sigma I_q^H) + \text{tr}(\Sigma I_q)\text{tr}(\Sigma I_q^H).
$$

(36)

**Theorem 3**  Assume $x_1, x_2, \ldots, x_n \in C^p$ is an i.i.d. sample from $CN(\mathbf{0}, \Sigma)$, the optimal shrinkage intensity for $T_1$ is

$$
w_{T_1} = 0 \land \frac{\text{tr}(\Sigma) - \frac{4}{p}\text{tr}(\Sigma^2) - \frac{4}{p(p-1)}\text{tr}(\Sigma L_p \Sigma L_p^H)}{\text{tr}(\Sigma^2) - \frac{4}{p}\text{tr}(\Sigma^2) - \frac{4}{p(p-1)}\text{tr}(\Sigma L_p \Sigma L_p^H) + d_{T_1}} \lor 1,
$$

(37)

where
\[ dt_1 = n \text{tr}(\Sigma^2) - \frac{n}{p} \text{tr}^2(\Sigma) - \frac{4n}{p(p-1)} \text{tr}(\Sigma I_p)\text{tr}(\Sigma I_p^H). \]  

The optimal shrinkage intensity for \( T_2 \) is

\[ w^*_T_2 = 0 \wedge \frac{\text{tr}^2(\Sigma) - \sum_{q=-p}^{p-1} \frac{1}{p-|q|} \text{tr}(\Sigma I_q)\Sigma I_q^H}{\text{tr}(\Sigma) - \sum_{q=-p}^{p-1} \frac{1}{p-|q|} \text{tr}(\Sigma I_q)\Sigma I_q^H + dt_2} \wedge 1, \]

where

\[ dt_2 = n \text{tr}(\Sigma^2) - \sum_{q=-p}^{p-1} \frac{n}{p-|q|} \text{tr}(\Sigma I_q)\text{tr}(\Sigma I_q^H). \]

**Proof** By the optimal shrinkage intensity given by (13), we need to compute \( E[\text{tr}(S^2)], E[\text{tr}(T \Sigma)] \) and \( E[\text{tr}(T^2)] \) under complex Gaussian distribution. By Eq. (14), we have

\[ E[\text{tr}(S^2)] = \frac{1}{n} \text{tr}^2(\Sigma) + \text{tr}(\Sigma^2). \]

Next, we just need to compute \( E[\text{tr}(T \Sigma)] \) and \( E[\text{tr}(T^2)]. \) When the target is \( T_1, \) we have

\[ E[\text{tr}(T \Sigma)] = E \left[ \text{tr} \left( \frac{\text{tr}(S)}{p} \Sigma + \frac{2\text{tr}(SL_p)}{p(p-1)} SL_p^H \Sigma + \frac{2\text{tr}(SL_p^H)}{p(p-1)} \Sigma L_p \right) \right] = \frac{1}{p} \text{tr}^2(\Sigma) + \frac{4}{p(p-1)} \text{tr}(\Sigma I_p)\text{tr}(\Sigma I_p^H) \]

and

\[ E[\text{tr}(T^2)] = E[\text{tr}(T_1^2)] = E[\text{tr}(T_1 S)] = \frac{1}{p} E[\text{tr}^2(S)] + \frac{4}{p(p-1)} E[\text{tr}(SL_p)\text{tr}(SL_p^H)]. \]

By Eqs. (33) and (34), we have

\[ E[\text{tr}(T^2)] = \frac{1}{p} \left( \frac{1}{n} \text{tr}^2(\Sigma) + \frac{4}{p(p-1)} \left( \frac{1}{n} \text{tr}(\Sigma I_p \Sigma I_p^H) + \text{tr}(\Sigma I_p)\text{tr}(\Sigma I_p^H) \right) \right) \]

\[ = \frac{1}{np} \text{tr}(\Sigma^2) + \frac{4}{p(p-1)} \text{tr}(\Sigma I_p \Sigma I_p^H) + \frac{4}{p(p-1)} \text{tr}(\Sigma I_p)\text{tr}(\Sigma I_p^H). \]

By substituting (41), (42) and (44) into (13), the optimal shrinkage intensity becomes

\[ w^*_T_1 = 0 \wedge \frac{1}{np} \text{tr}(\Sigma^2) - \frac{1}{np} \text{tr}^2(\Sigma) - \frac{4}{np(p-1)} \text{tr}(\Sigma I_p \Sigma I_p^H) \]

\[ \wedge 1, \]

where \( dt_1 = \text{tr}(\Sigma^2) - \frac{1}{p} \text{tr}^2(\Sigma) - \frac{4}{p(p-1)} \text{tr}(\Sigma I_p)\text{tr}(\Sigma I_p^H). \)

In the same manner, when the target is \( T_2, \) we have

\[ E[\text{tr}(T \Sigma)] = E \left[ \text{tr} \left( \sum_{q=-p}^{p-1} \frac{1}{p-|q|} \text{tr}(\Sigma I_q)\Sigma I_q^H \right) \right] \]

\[ = \sum_{q=-p}^{p-1} \frac{1}{p-|q|} \text{tr}(\Sigma I_q)\text{tr}(\Sigma I_q^H), \]

and

\[ E[\text{tr}(T^2)] = E[\text{tr}(T_2^2)] = \sum_{q=-p}^{p-1} \frac{1}{p-|q|} E[\text{tr}(\Sigma I_q)\text{tr}(\Sigma I_q^H)]. \]

By Eq. (36), we can obtain
\[
\mathbb{E}[\text{tr}(T^2)] = \sum_{q=-(p-1)}^{p-1} \frac{1}{n} \left( \frac{1}{n} \text{tr}(\Sigma I_q \Sigma I_q^H) + \text{tr}(\Sigma I_q)\text{tr}(\Sigma I_q^H) \right).
\]

By substituting (41), (46) and (48) into (13), the optimal shrinkage intensity becomes

\[
w^2_{T^2} = 0 \land \frac{1}{n} \text{tr}^2(\Sigma) - \frac{1}{n^2-1} \sum_{q=-(p-1)}^{p-1} \frac{1}{n} \text{tr}(\Sigma I_q \Sigma I_q^H) + \frac{1}{n^2-1} \frac{1}{n} \text{tr}(\Sigma I_q)\text{tr}(\Sigma I_q^H) + \alpha_{T^2},
\]

where \( \alpha_{T^2} = \text{tr}(\Sigma^2) - \sum_{q=-(p-1)}^{p-1} \frac{1}{n} \text{tr}(\Sigma I_q)\text{tr}(\Sigma I_q^H) \).

We remind that the optimal shrinkage intensities given by (37) and (39) involve the true covariance matrix and are unavailable in real applications. Even so, the theoretical optimal tuning parameters are still crucial in shrinkage estimation and can be employed to evaluate the available ones.

Available shrinkage estimator for Toeplitz-structured targets

In this section, we estimate the unknown scalar quantities in (37) and (39), and obtain available shrinkage intensities by plug-in strategy.

**Corollary 4** Let \( x_1, x_2, \ldots, x_n \in \mathbb{C}^p \) be an i.i.d. sample from \( \mathcal{C}\mathcal{N}(\mathbf{0}, \Sigma) \). For an arbitrary nonrandom matrix \( \mathbf{A} \), the unbiased estimates of \( \text{tr}(\Sigma \mathbf{A}^* \Sigma^H) \) and \( \text{tr}(\Sigma)\text{tr}(\Sigma^H) \) are respectively given by

\[
\alpha_{\mathbf{A}} = \frac{n^2}{n^2-1} \text{tr}(\mathbf{A}^* \mathbf{S} \mathbf{A}^H) - \frac{n}{n^2-1} \text{tr}(\mathbf{S}) \text{tr}(\mathbf{A}^H),
\]

\[
\beta_{\mathbf{A}} = \frac{n^2}{n^2-1} \text{tr}(\mathbf{A}^* \mathbf{S} \mathbf{A}^H) - \frac{n}{n^2-1} \text{tr}(\mathbf{S} \mathbf{A}^H).
\]

**Proof** By Theorem 2, for an arbitrary nonrandom matrix \( \mathbf{A} \), we have

\[
\mathbb{E} \left( \frac{\text{tr}(\mathbf{A}^* \mathbf{S} \mathbf{A}^H)}{\text{tr}(\mathbf{S})\text{tr}(\mathbf{A}^H)} \right) \equiv \left( \frac{1}{n^2} \right) \left( \frac{1}{n} \right) \left( \frac{\text{tr}(\Sigma \mathbf{A}^* \Sigma^H)}{\text{tr}(\Sigma)\text{tr}(\Sigma^H)} \right).
\]

Because

\[
\left( \frac{1}{n^2} \right) \left( \frac{1}{n} \right) = \frac{1}{n^2-1} \left( \frac{n^2 - n}{-n - n^2} \right),
\]

we can obtain

\[
\mathbb{E} \left( \frac{1}{n^2-1} \left( \frac{n^2 - n}{-n - n^2} \right) \left( \frac{\text{tr}(\mathbf{A}^* \mathbf{S} \mathbf{A}^H)}{\text{tr}(\mathbf{S})\text{tr}(\mathbf{A}^H)} \right) \right) = \left( \frac{\text{tr}(\Sigma \mathbf{A}^* \Sigma^H)}{\text{tr}(\Sigma)\text{tr}(\Sigma^H)} \right).
\]

Therefore, we have

\[
\mathbb{E} \left[ \frac{n^2}{n^2-1} \text{tr}(\mathbf{A}^* \mathbf{S} \mathbf{A}^H) - \frac{n}{n^2-1} \text{tr}(\mathbf{S}) \text{tr}(\mathbf{A}^H) \right] = \text{tr}(\Sigma \mathbf{A}^* \Sigma^H),
\]

\[
\mathbb{E} \left[ \frac{n^2}{n^2-1} \text{tr}(\mathbf{S}) \text{tr}(\mathbf{A}^H) - \frac{n}{n^2-1} \text{tr}(\mathbf{A}^* \mathbf{S} \mathbf{A}^H) \right] = \text{tr}(\Sigma)\text{tr}(\Sigma^H).
\]

This concludes the proof.

By Corollary 4, when \( \mathbf{A} = \mathbf{I}_p \), the unknown scalar quantities \( \text{tr}(\Sigma^2) \) and \( \text{tr}^2(\Sigma) \) are unbiasedly estimated by

\[
\alpha_{\mathbf{I}_p} = \frac{n^2}{n^2-1} \text{tr}(\Sigma^2) - \frac{n}{n^2-1} \text{tr}^2(\Sigma),
\]

\[
\beta_{\mathbf{I}_p} = \frac{n^2}{n^2-1} \text{tr}^2(\Sigma) - \frac{n}{n^2-1} \text{tr}(\Sigma^2).
\]

When \( \mathbf{A} = \mathbf{I}_p \), the unbiased estimates of \( \text{tr}(\Sigma \mathbf{I}_p^* \Sigma \mathbf{I}_p^H) \) and \( \text{tr}(\Sigma \mathbf{I}_p)\text{tr}(\Sigma \mathbf{I}_p^H) \) are respectively given by

\[
\alpha_{\mathbf{I}_p} = \frac{n^2}{n^2-1} \text{tr}(\mathbf{I}_p^* \mathbf{S} \mathbf{I}_p^H) - \frac{n}{n^2-1} \text{tr}(\mathbf{I}_p)\text{tr}(\mathbf{I}_p^H),
\]

\[
\beta_{\mathbf{I}_p} = \frac{n^2}{n^2-1} \text{tr}(\mathbf{I}_p)\text{tr}(\mathbf{I}_p^H) - \frac{n}{n^2-1} \text{tr}(\mathbf{I}_p^* \mathbf{S} \mathbf{I}_p^H).
\]
Table 1. Comparison of shrinkage intensities for Toeplitz-structured target.

| References | Target | Intensity expression | Method |
|------------|--------|----------------------|--------|
| Liu \(^{30}\) | \(T_2\) | \((n - 3m)\sigma^2 + (n - 1)\sigma^2 \bigg/ (n - 3m + (n - 1))\) \(\vee 1\) | Complex Gaussian, biased |
| Tong \(^{31}\) | \(T_1, T_2\) | \(0 \wedge \frac{\sum_{q=-p}^{-1}(p-q)\alpha_q}{\sum_{q=-p}^{-1}(p-q)^2\alpha_q} \vee 1\) | Distribution free, cross-validation |
| This paper | \(T_1\) | \(0 \wedge \frac{\alpha_q - \frac{1}{p}\alpha_p - \frac{4}{p(p-1)}\alpha_p}{\frac{1}{p}\alpha_p - \frac{4}{p(p-1)}\alpha_p + \hat{d}_T} \vee 1\) | Complex Gaussian, unbiased |
| This paper | \(T_2\) | \(0 \wedge \frac{\alpha_q - \frac{1}{p}\alpha_p - \frac{4}{p(p-1)}\alpha_p}{\frac{1}{p}\alpha_p - \frac{4}{p(p-1)}\alpha_p + \hat{d}_T} \vee 1\) | Complex Gaussian, unbiased |

Finally, we compare the expressions of the proposed shrinkage intensities and their existing competitors in Table 1.

Methods

This section provides some numerical simulations and apply them to array signal processing to verify the performance of the proposed estimators. For the sake of notation, the proposed shrinkage estimators in (65) and (68) are denoted as LST1 and LST2. The optimal shrinkage estimators for targets \(T_1\) and \(T_2\) are respectively denoted as LSO1 and LSO2. The existing competitor for \(T_2\) in \(^{30}\) is denoted as LSZ. The competitors in \(^{31}\) for \(T_1, T_2\) are respectively denoted as LSCV1 and LSCV2.
Numerical simulations. This subsection verifies the estimation accuracy of the available shrinkage intensities by plugging in the unbiased estimates given in Corollary 4. Then we compare the MSEs of the shrinkage estimators listed in Table 1. In implementation, we generate the sample from the complex Gaussian distribution $CN(0, \Sigma)$, where $\Sigma$ is the true covariance matrix. In numerical simulations, we consider two covariance models:

1. Model 1: $\Sigma = \Sigma_\eta + \nu \Sigma_T$, where $\Sigma_\eta = (\sigma_{ij})_{p \times p}$ and $\sigma_{ij} = |i-j|$, $\eta = 0.8$.
2. Model 2: $\Sigma = \Sigma_\rho + \nu \Sigma_T$, where $\Sigma_\rho = (\sigma_{ij})_{p \times p}$ and $\sigma_{ij} = \frac{1}{2}|i-j|^{2\rho} - |i-j|^{2\rho} + \frac{1}{2}|i-j-1|^{2\rho}$. $\eta = 0.8$, $\rho = 0.9$, and $\Sigma_T = \xi \xi^T$ with $\xi$ being uniformly drawn from the interval $[0, 1]$, to result in a Toeplitz-like covariance matrices. The dimension is $p = 100$, the sample size $n$ varies from 10 to 100.

Figures 1 and 2 depict the variation tendencies of the optimal shrinkage intensities and their estimates under Model 1 and Model 2. We can see that the available shrinkage intensities based on the unbiased estimates of unknown scalar quantities are very close to the corresponding theoretical optimal ones. For the target $T_1$, the proposed LST1 outperforms the existing LSCV1 based on the cross-validation method. Moreover, for the target $T_2$, the proposed LST2 performs better than both LSZ based on the biased estimates of unknown scalar quantities and LSCV2 based on the cross-validation method.

Next, we employ the percentage relative improvement in average losses (PRIAL) over the SCM to compare the MSE performance of the proposed covariance estimators and the existing estimators of the same type\textsuperscript{14}. The PRIAL is defined by

$$PRIAL = \frac{E[\|S - \Sigma\|^2] - E[\|\hat{S} - \Sigma\|^2]}{E[\|S - \Sigma\|^2]}. \quad (69)$$

The covariance matrix estimator with a larger PRIAL is preferred.
Figures 3 and 4 report the PRIAL values of shrinkage estimators for the two kinds of covariance models, where a larger PRIAL means a smaller MSE. We can see that the PRIALs of proposed LST1 and LST2 are very close to the PRIAL of the optimal shrinkage estimators. When $T_1$ is employed as the target in shrinkage estimation, the proposed LST1 outperforms LSCV1. When $T_2$ is employed, the proposed LST2 outperforms the existing LSZ and LSCV2. Therefore, the proposed covariance matrix estimators for $T_1, T_2$ enjoy a satisfactory MSE performance.

Adaptive beamforming. This subsection applies the proposed covariance estimators to array signal processing. In detail, a uniform linear array that consists of sensors with half-wavelength spacing is considered. The received signal at time $t$ is

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

where $\theta_0$ and $\theta_k$ are respectively the directions of desired complex Gaussian signal $s_0(t)$ and interference signals $s_k(t)$, $a(\theta_0)$ and $a(\theta_k)$ are the corresponding responses, and $n(t)$ is a white noise. We can obtain the minimum variance distortionless response (MVDR) beamformer:

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

We usually the covariance matrix $\Sigma$ in (71) and replace it with its estimate $\hat{\Sigma}$. The corresponding output signal-to-interference-plus-noise ratio (SINR) is

$$\text{SINR} = \frac{\sigma_0^2|w^Ha(\theta_0)|^2}{w^H(\hat{\Sigma} - a(\theta_0)a(\theta_0)^H)w},$$
where \( \hat{w} \) is the realized beamformer based on \( \hat{\Sigma} \). The shrinkage estimator with larger output SINR is recommended in array signal processing.

In our experiments, the sensor number is \( p = 50 \) and the desired signal is assumed to have an angle of arrival of \( \theta_0 = 5^\circ \) with power \( \sigma_0^2 = 10 \) dB. There are \( K = 3 \) interference signals on the directions \( -10^\circ, 0^\circ, 10^\circ \) with a same power 8 dB, and an additive complex Gaussian noise with a power 0 dB. For each covariance estimator, we compute the output SINR by averaging \( 10^4 \) repetitions.

Figure 5 reports the SINRs and corresponding elapsed time of adaptive beamformer based on the shrinkage estimators. We can see that the SINR is small when the sample size is small and becomes more significant as the sample size increases. The shrinkage estimators for target \( T_1 \) perform better than the ones for target \( T_2 \) on SINR. We can see that the target matrix is an important factor in the shrinkage estimation of the covariance matrix. Furthermore, the proposed shrinkage estimator LST1 has the largest SINR and the lowest computational cost, showing that the unbiased estimates of the unknown scalar quantities are conducive to improving the performance of the shrinkage estimator.

**DOA estimation.** This subsection investigates the performance of shrinkage estimators in the direction of arrival (DOA) estimation. We assume the linear array consists of \( p = 50 \) uniformly distributed sensors. There are \( K = 5 \) mutually independent complex Gaussian source signals from DOA \( \theta = (-10^\circ, -5^\circ, 0^\circ, 5^\circ, 10^\circ)^T \). The signal-to-noise ratio (SNR) is 6 dB. A DOA estimate \( \hat{\theta} \) can be obtained using the MUSIC algorithm variant by replacing the SCM with shrinkage estimator \(^{30,35}\). The root mean squared error (RMSE) of DOA estimate \( \hat{\theta} \) is defined as

\[
\text{RMSE}(\hat{\theta}) = \sqrt{\frac{1}{R K} \sum_{r=1}^{R} \|\hat{\theta}_r - \theta\|^2},
\]

where \( R \) is the Monte Carlo run and \( \hat{\theta}_r \) is the DOA estimate in \( r \)-th Monte Carlo run.

Figure 6 reports the RMSEs and corresponding elapsed time of DOA based on the shrinkage estimators. The classic SCM is unavailable in this scenario where \( n \leq p \). We can see that the estimated error is disappointing as the dimension is large and the sample size is small. The proposed shrinkage estimator LST1 enjoys the lowest RMSE and elapsed time of the others. With an additional time cost, the proposed LST2 can reach a lower RMSE than the existing competitors such as LSZ, LSCV1, and LSCV2 when \( n \geq 30 \).

**Conclusion**

In this paper, we have studied the linear estimation of the covariance matrix in large-dimension and small-sample size scenarios. Under the complex Gaussian distribution, we obtain the optimal shrinkage intensities for the two kinds of Toeplitz-structured targets. By discovering a crucial moment property of the complex Wishart distribution, we unbiasedly estimate the unknown scalar quantities related to the true covariance matrix. Then we propose the shrinkage estimators for the Toeplitz-structured targets. Some numerical simulations and two example applications to signal processing reveal that the proposed covariance matrix estimators dominate the existing estimators in large-dimension and small-sample size scenarios.

In future work, we will continue in-depth studies on the Cramér–Rao bound of the linear shrinkage estimator and investigate nonlinear shrinkage estimation. Moreover, how to choose the target matrix in practice is still a problem worthy of study in shrinkage estimation.
Figure 6. The RMSE and corresponding elapsed time of DOA based on each shrinkage estimator.

Data availability
All data included in this study are available upon request by contact with the corresponding author.

Received: 15 April 2022; Accepted: 5 October 2022
Published online: 08 November 2022

References
1. Choi, S. et al. Sea clutter covariance matrix estimation and its application to whitening filter. J. Electromagn. Eng. Sci. 21(2), 134–142 (2021).
2. Wang, L., Wang, W.-Q. & Zhou, Y. FDA-MIMO radar covariance matrix estimation via shrinkage processing. Digital Signal Process. 118, 103206 (2021).
3. Hirukawa, M. Robust covariance matrix estimation in time series: A review. Econ. Stat. 12, 1–26 (2021).
4. Liu, C., Wang, M. & Xia, N. Design-free estimation of integrated covariance matrices for high-frequency data. J. Multivar. Anal. 189, 105410 (2022).
5. Jia, S., Zhang, C. & Lu, H. Covariance function versus covariance matrix estimation in efficient semi-parametric regression for longitudinal data analysis. J. Multivar. Anal. 187, 105009 (2022).
6. Fourdrinier, D., Haddouche, A. M. & Mezoued, F. Covariance matrix estimation under data-based loss.
7. Le Thi, H. A. & Ho, V. T. Alternating dca for reduced-rank multitask linear regression with covariance matrix estimation. Ann. Math. Artif. Intell. 1–21 (2021).
8. Won, J.-H., Lim, J., Kim, S.-J. & Rajaratnam, B. Condition-number-regularized covariance estimation. J. R. Stat. Soc. Ser. B (Stat. Methodol.) 75(3), 427–450 (2013).
9. Yang, Y., Zhou, J. & Pan, J. Estimation and optimal structure selection of high-dimensional Toeplitz covariance matrix. J. Multivar. Anal. 184, 104739 (2021).
10. Wang, H., Peng, B., Li, D. & Leng, C. Nonparametric estimation of large covariance matrices with conditional sparsity. J. Econ. 223(1), 53–72 (2021).
11. Raninen, E. & Ollila, E. Coupled regularized sample covariance matrix estimator for multiple classes. Trans. Signal Process. 69, 5681–5692 (2021).
12. Bodnar, T., Drmyriv, S., Okhrin, Y., Parolya, N. & Schmid, W. Statistical inference for the expected utility portfolio in high dimensions. IEEE Trans. Signal Process. 69, 1–14 (2021).
13. Yang, W., & Kang, X. An improved banded estimation for large covariance matrix. Commun. Stat. Theory Methods 1–17 (2021).
14. Ledoit, O. & Wolf, M. Honey, I shrunk the sample covariance matrix. J. Portf. Manag. 30(4), 110–119 (2004).
15. Lopez-Restrepo, S. et al. An efficient ensemble kalman filter implementation via shrinkage covariance matrix estimation: Exploiting prior knowledge. Comput. Geosci. 25(3), 985–1003 (2021).
16. Ledoit, O. & Wolf, M. Shrinkage estimation of large covariance matrices: Keep it simple, statistician?. J. Multivar. Anal. 186, 104796 (2021).
17. Bodnar, O., Bodnar, T. & Parolya, N. Recent advances in shrinkage-based high-dimensional inference. J. Multivar. Anal. 188, 104826 (2021).
18. Tong, J. et al. Shrinkage of covariance matrices for linear signal estimation using cross-validation. IEEE Trans. Signal Process. 64(11), 2965–2975 (2016).
19. Li, J., Zhou, J. & Zhang, B. Estimation of large covariance matrices by shrinking to structured target in normal and non-normal distributions. IEEE Access 6, 2158–2169 (2018).
20. Joo, Y. C. & Park, S. Y. Optimal portfolio selection using a simple double-shrinkage selection rule. Financ. Res. Lett. 43, 102019 (2021).
21. Chen, X., Wang, Z. J. & McKeown, M. J. Shrinkage-to-tapering estimation of large covariance matrices. IEEE Trans. Signal Process. 60(11), 5640–5656 (2012).
22. Ikeda, Y., Kubokawa, T. & Srivastava, M. S. Comparison of linear shrinkage estimators of a large covariance matrix in normal and non-normal distributions. Comput. Stat. Data Anal. 95, 95–108 (2016).
23. Cabana, E., Lillo, E. & Laniado, H. Multivariate outlier detection based on a robust Mahalanobis distance with shrinkage estimators. Stat. Pap. 62(4), 1583–1609 (2021).
24. Fisher, T. J. & Sun, X. Improved Stein-type shrinkage estimators for the high-dimensional multivariate normal covariance matrix. Comput. Stat. Data Anal. 55(5), 1909–1918 (2011).
25. Sun, R., Ma, T., Liu, S. & Sathye, M. Improved covariance matrix estimation for portfolio risk measurement: A review. J. Risk Portf. Manag. 12(1), 48 (2019).
26. Zhang, B. Zhou, J. & Li, J. Improved shrinkage estimators of covariance matrices with Toeplitz-structured targets in small sample scenarios. *IEEE Access* 7, 785–798 (2019).

27. Ollila, E. & Raninen, E. Optimal shrinkage covariance matrix estimation under random sampling from elliptical distributions. *IEEE Trans. Signal Process.* 67(10), 2707–2719 (2019).

28. Hadjouche, A. M., Fourdrinier, D. & Mezoued, F. Scale matrix estimation of an elliptically symmetric distribution in high and low dimensions. *J. Multivar. Anal.* 181, 104680 (2021).

29. Nguyen, V., Kuhn, D. & Esfahani, P. Distributionally robust inverse covariance estimation: The Wasserstein shrinkage estimator. *Oper. Res.* 70(1), 490–515 (2022).

30. Liu, Y., Sun, X. & Zhao, S. A covariance matrix shrinkage method with Toeplitz rectified target for DOA estimation under the uniform linear array. *Int. J. Electron. Commun. (AEÜ)* 81, 50–55 (2017).

31. Tong, J. et al. Linear shrinkage estimation of covariance matrices using low-complexity cross-validation. *Signal Process.* 148, 223–233 (2018).

32. Zhang, B. Improved shrinkage estimator of large-dimensional covariance matrix under the complex Gaussian distribution. *Math. Probl. Eng.* 2020, 1–8 (2020).

33. Mestre, X. & Lagunas, M. Á. Finite sample size effect on minimum variance beamformers: Optimum diagonal loading factor for large arrays. *IEEE Trans. Signal Process.* 54(1), 69–82 (2006).

34. Serra, J. & Nájar, M. Asymptotically optimal linear shrinkage of sample LMMSE and MVDR filters. *IEEE Trans. Signal Process.* 62(14), 3552–3564 (2014).

35. Vallet, P. & Louvaton, P. Toeplitz rectification and DoA estimation with MUSIC. in *IEEE International Conference on Acoustics, Speech and Signal Processing, Florence, Italy* (pp. 2237–2241, 2014).

**Author contributions**

B.Z. conceived the experiments, B.Z. and S.Y. conducted the experiments, B.Z. and S.Y. analysed the results. All authors reviewed the manuscript.

**Funding**

This work was supported in part by the Guangxi Science and Technology Planning Project (2021AC19061), the Science and Technology Project of Guangxi(2021AC06001), the Innovation team of Pu’er University (CXTD019), and the Scientific Research Fund Project of Yunnan Education Department (2022J0985).

**Competing interests**

The authors declare no competing interests.

**Additional information**

**Correspondence** and requests for materials should be addressed to S.Y.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit [http://creativecommons.org/licenses/by/4.0/](http://creativecommons.org/licenses/by/4.0/).

© The Author(s) 2022