

Statistical Performance of a Data-Based Covariance Estimator

Zhengyuan Xu

Abstract—Data covariance is typically estimated by a sample average method. The second-order statistics of this sample covariance are encountered in evaluating the performance of many covariance-based processing methods. This correspondence derives closed-form expressions for the second-order statistics of a vectored sample covariance. Given a baseband communication system model, the results explicitly rely on the channel parameters and the second- and fourth-order statistics of the channel noise and inputs. They can be directly applied to study statistical properties of covariance-based methods such as a correlation-matching-based channel-estimation method in a code-division multiple-access system. Numerical examples are provided to further verify the derived results.

Index Terms—Channel estimation, code-division multiple access (CDMA), covariance estimation, performance analysis.

I. INTRODUCTION

The covariance-matching technique has been successfully applied to blind channel estimation for either time-division multiple-access (TDMA) systems [1], [2], periodic (short-code) code-division multiple-access (CDMA) systems [3], [4], or aperiodic (long-code) CDMA systems [5], [6]. This technique compares the ideal covariance of a model-based channel output with its estimate and minimizes certain matching errors to obtain the corresponding channel estimators based on either the matrix Frobenius norm or the vector norm [7].

Data-covariance estimation is necessary for the implementation of covariance-based algorithms. Many algorithms have been developed to estimate covariance, most of which are contributed to array processing applications such as direction-of-arrival (DOA) estimation. In [8], a maximum-likelihood (ML) method is presented to estimate structured real covariance matrices of Gaussian processes. Under low rank modeling and a Gaussian assumption of source signals and noise, [9] reveals a relationship between ML covariance estimation and multiple-window spectrum analysis. A robust ML estimation method is proposed in [10] by minimizing the worst case asymptotic estimated covariance even in the presence of non-Gaussian noise. A computationally efficient asymptotic ML estimate is derived for a structured covariance matrix using the extended invariance principle (including reparameterization of the ML criterion and solution refinement based on a weighted least-squares technique) in [11]. Asymptotic performance for direction estimation is studied in [12] when the number of snapshots is sufficiently large in estimating the data covariance. A Toeplitz covariance matrix can be parameterized by only common unknown elements in each of the diagonals [13]. Those parameters are then estimated with substantially reduced computational complexity. Incorporating the Toeplitz structure of the covariance matrix also can significantly improve the adaptive beamformer and detector performance [14].

Certainly, the performance of a covariance-estimation method characterizes each covariance-based algorithm. In practice, estimation of the data covariance can easily be performed by a sample averaging method based on received data. If Gaussian sources are assumed, the second-order statistics of the estimated covariance can be found in [15].

They have been applied to analyze asymptotic properties of eigenvectors of a covariance matrix [16]. Performance of such a covariance estimator has also been analyzed in [17] and [18] for further analysis of DOA estimators. Statistical analysis of a subspace channel-estimation method is carried out from the statistics of the sample covariance in [19]. Those results are based on an assumption that a sufficiently large number of snapshots are available at the receiver, causing the sample covariance to converge in distribution to a Gaussian process according to the central limit theorem [20].

The previously mentioned situation motivates us to derive an easy-to-use expression for the statistics of the vectored sample covariance with a finite number of snapshots and not necessarily Gaussian channel inputs. Although statistics of a weighted covariance matrix have been derived in [21], they do not provide pairwise statistics of all elements in the covariance, such as vectorized covariance used in covariance matching techniques. More general results suitable for study of covariance-based algorithms are desirable and, thus, are derived in this correspondence.

Starting from a typical linear channel input–output model, the covariance of a vectored sample covariance is simplified according to a powerful mathematical tool—the Kronecker product (“ \otimes ”) [7]. It is derived in a closed form as a function of statistics of channel inputs, channel parameters, and statistics of Gaussian noise, for either a real system or a system involving complex quantities (complex data modulation, complex channel, etc.). Extension to a system with colored noise is made possible by whitening the received data first and then obtaining the statistics of the covariance estimated from the resulting data. Using the transformation relation, the desired covariance then follows. The derived results are shown to be applicable to evaluate performance of a correlation-matching-based channel estimator [4]. They may also be useful for the design of constraints in robust multiuser detection techniques such as [22] and [23].

This correspondence has the following structure. Section II describes a data model to be used and Section III provides closed-form results for the covariance of a vectored sample covariance together with detailed derivation. Section IV shows an application example to evaluate a correlation-matching based channel estimator in a CDMA system. Simulation examples are provided for verification of analytical results in Section V. Finally, some conclusions are drawn in Section VI.

II. DATA MODEL

Consider a vector form channel input/output data model

$$\mathbf{y}_n = \mathbf{H}\mathbf{b}_n + \mathbf{v}_n \quad (1)$$

where \mathbf{y}_n is a $\nu \times 1$ received data vector at time n , \mathbf{H} is a $\nu \times L$ channel matrix, \mathbf{b}_n is an $L \times 1$ input vector, and \mathbf{v}_n represents the additive Gaussian noise independent of inputs. This model can describe a variety of communication systems. For example, in a single-input–multiple-output (SIMO) system, \mathbf{b}_n includes inputs from time n to $n+L-1$ emitted from the same source and \mathbf{H} has a block Toeplitz structure. When it is used to represent a multiple-input–multiple-output (MIMO) system, \mathbf{y}_n has samples from multiple sources. Accordingly, \mathbf{b}_n contains symbols from different sources. In particular, the received signal in a CDMA system can also fit into this model where \mathbf{H} becomes a signature waveform matrix that captures the overall effects of spreading codes, propagation channels, and front-end pulse shaping and matched filtering.

In this correspondence, we assume all entries in \mathbf{b}_n are mutually independent identically distributed (i.i.d.) sequences. They are drawn from the same symmetric constellation with zero odd moments, equal

Manuscript received May 5, 2003; revised October 30, 2003 and January 6, 2004. This work was supported in part by the National Science Foundation under Grant NSF-CCR 0207931.

The author is with the Department of Electrical Engineering, University of California, Riverside, CA 92521 USA (e-mail: dxu@ee.ucr.edu).

Digital Object Identifier 10.1109/TVT.2004.825698

variance σ_b^2 and fourth-order absolute moment m_{4b} , and fourth-order cumulant κ_{4b} ($\kappa_{4b} = m_{4b} - 3\sigma_b^4$ for real sequences and $\kappa_{4b} = m_{4b} - 2\sigma_b^4$ for complex sequences [20]). In a case when multiple sources have different transmission powers, they can be easily absorbed into the columns of channel matrix \mathbf{H} by proper scaling. Additionally, we temporarily assume that elements in \mathbf{v}_n represent i.i.d. additive white Gaussian noise (AWGN) with zero mean and variance σ_v^2 . An extension to a system with colored noise will be made later. If the inputs and noise are complex quantities, their real and imaginary parts are assumed to be i.i.d. sequences.

For notational convenience, denote Hermitian complex conjugate (*) transpose (T) by H , the l th ($l = 1, 2, \dots$) column of a matrix \mathbf{U} as \mathbf{u}_l , \mathbf{I}_ν as an identity matrix of degree ν , the Khatri–Rao product “ \square ” to represent column-by-column Kronecker product of two matrices with the same number of columns [24] $\mathbf{U}\square\mathbf{W} = [\mathbf{u}_1 \otimes \mathbf{w}_1, \dots, \mathbf{u}_m \otimes \mathbf{w}_m]$, and a “vec” operation to successively stack columns of a matrix into a big column vector [7]. A diagonal or block diagonal matrix with main diagonal entries \mathbf{x}_i is denoted as $\text{diag}\{\mathbf{x}_1, \mathbf{x}_2, \dots\}$. We also define $\mathbf{U} \odot_r \mathbf{W} \triangleq [\mathbf{U} \otimes \mathbf{w}_1, \dots, \mathbf{U} \otimes \mathbf{w}_m]$, which successively applies Kronecker product to each column of the right matrix and, correspondingly, $\mathbf{U} \odot_l \mathbf{W} \triangleq [\mathbf{u}_1 \otimes \mathbf{W}, \dots, \mathbf{u}_m \otimes \mathbf{W}]$ to the left matrix.

III. PERFORMANCE EVALUATION

In the covariance-matching context, the covariance of \mathbf{y}_n is first computed based on model (1) as

$$\mathbf{R} = E\{\mathbf{y}_n \mathbf{y}_n^H\} = \sigma_b^2 \mathbf{H} \mathbf{H}^H + \sigma_v^2 \mathbf{I}_\nu. \quad (2)$$

Define $\mathbf{r} = \text{vec}(\mathbf{R})$. If \mathbf{R} is estimated by sample average $\hat{\mathbf{R}} = (1/N) \sum_{n=1}^N \mathbf{y}_n \mathbf{y}_n^H$, then it has been observed that the covariance of $\hat{\mathbf{r}} = \text{vec}(\hat{\mathbf{R}})$ characterizes the asymptotic property of each channel estimator [2]–[4]. For simplicity of derivation, assume all N data vectors are independent and define $\hat{\mathbf{r}}_n = \text{vec}(\mathbf{y}_n \mathbf{y}_n^H)$ as the instantaneous vectored covariance. Then, it can be easily verified that the covariance of $\hat{\mathbf{r}}$ has a form

$$\Phi(\hat{\mathbf{r}}) = E\{(\hat{\mathbf{r}} - \mathbf{r})(\hat{\mathbf{r}} - \mathbf{r})^H\} = \frac{1}{N^2} \sum_{n_1, n_2=1}^N E\{\hat{\mathbf{r}}_{n_1} \hat{\mathbf{r}}_{n_2}^H\} - \mathbf{r} \mathbf{r}^H. \quad (3)$$

Considering all possible values that n_1 and n_2 take and invoking the independence assumption on \mathbf{y}_n at different time, (3) can be simplified as

$$\Phi(\hat{\mathbf{r}}) = \frac{1}{N} \Psi, \quad \Psi = E\{\hat{\mathbf{r}}_n \hat{\mathbf{r}}_n^H\} - \mathbf{r} \mathbf{r}^H. \quad (4)$$

For clarity, we provide analytical expressions for $\Phi(\hat{\mathbf{r}})$ in the following proposition, whose proof is given in the Appendix.

Proposition: If the channel model follows (1), with L inputs and ν outputs, and data covariance is estimated by $\hat{\mathbf{R}} = (1/N) \sum_{n=1}^N \mathbf{y}_n \mathbf{y}_n^H$ from N independent data vectors, then, for a real system, the covariance of $\hat{\mathbf{r}} = \text{vec}(\hat{\mathbf{R}})$ satisfies

$$\begin{aligned} N\Phi(\hat{\mathbf{r}}) &= \kappa_{4b}(\mathbf{H}\square\mathbf{H})(\mathbf{H}\square\mathbf{H})^T \\ &\quad + \sigma_b^4(\mathbf{H} \odot_r \mathbf{H})(\mathbf{H} \odot_l \mathbf{H})^T \\ &\quad + \mathbf{R} \otimes \mathbf{R} + \sigma_v^4(\mathbf{I}_\nu \odot_r \mathbf{I}_\nu)(\mathbf{I}_\nu \odot_l \mathbf{I}_\nu)^T \\ &\quad + \sigma_b^2 \sigma_v^2(\mathbf{H} \odot_r \mathbf{I}_\nu)(\mathbf{I}_\nu \otimes \mathbf{H}^T) \\ &\quad + \sigma_b^2 \sigma_v^2(\mathbf{I}_\nu \otimes \mathbf{H})(\mathbf{H} \odot_r \mathbf{I}_\nu)^T \end{aligned} \quad (5)$$

while for a complex system it satisfies

$$N\Phi(\hat{\mathbf{r}}) = \kappa_{4b}(\mathbf{H}^* \square \mathbf{H})(\mathbf{H}^* \square \mathbf{H})^H + \mathbf{R}^* \otimes \mathbf{R}. \quad (6)$$

This proposition provides a direct method to compute $\Phi(\hat{\mathbf{r}})$ for given system parameters. It can be observed that $\Phi(\hat{\mathbf{r}})$ is inversely proportional to N .

A. Discussion

Our model is based on the assumption that all inputs have the same power. If their powers are different, then proper scaling of the corresponding column vector in the channel matrix can be performed. Another assumption is that the noise is AWGN. If the noise is colored Gaussian, then some preprocessing is necessary in order to directly apply our results. Suppose that $E\{\mathbf{v}_n \mathbf{v}_n^H\} = \mathbf{R}_v$. Prewhitening \mathbf{y}_n by $\mathbf{R}_v^{-1/2}$ yields a new data vector

$$\tilde{\mathbf{y}}_n \triangleq \mathbf{R}_v^{-1/2} \mathbf{y}_n = \tilde{\mathbf{H}} \mathbf{b}_n + \tilde{\mathbf{v}}_n, \quad \tilde{\mathbf{H}} = \mathbf{R}_v^{-1/2} \mathbf{H} \quad (7)$$

where $\tilde{\mathbf{v}}_n$ becomes a white Gaussian noise vector with $E\{\tilde{\mathbf{v}}_n \tilde{\mathbf{v}}_n^H\} = \mathbf{I}_\nu$. Then, each of its elements have unit covariance. After comparing (7) with (1), replacing \mathbf{H} by $\tilde{\mathbf{H}}$, and setting σ_v^2 to be one, it is clear that either (5) or (6) can be applied to $\tilde{\mathbf{y}}_n$ to obtain the corresponding statistics of the covariance termed as $\tilde{\Psi} = N\Phi(\tilde{\mathbf{r}})$. Once it is derived, we can obtain Ψ . According to (7) and (4), we have

$$\Psi = E\left\{\text{vec}\left(\mathbf{R}_v^{1/2} \tilde{\mathbf{y}}_n \tilde{\mathbf{y}}_n^H \mathbf{R}_v^{1/2}\right) \text{vec}^H\left(\mathbf{R}_v^{1/2} \tilde{\mathbf{y}}_n \tilde{\mathbf{y}}_n^H \mathbf{R}_v^{1/2}\right)\right\} - \mathbf{r} \mathbf{r}^H.$$

Using the property of the Kronecker product [7] $\text{vec}(\mathbf{Z}_1 \mathbf{Z}_2 \mathbf{Z}_3) = (\mathbf{Z}_3^T \otimes \mathbf{Z}_1) \text{vec}(\mathbf{Z}_2)$, it becomes

$$\Psi = \left[(\mathbf{R}_v^*)^{1/2} \otimes \mathbf{R}_v^{1/2} \right] \tilde{\Psi} \left[(\mathbf{R}_v^*)^{1/2} \otimes \mathbf{R}_v^{1/2} \right]^H - \mathbf{r} \mathbf{r}^H.$$

If we choose to replace \mathbf{r} by $\text{vec}(E\{\mathbf{R}_v^{1/2} \tilde{\mathbf{y}}_n \tilde{\mathbf{y}}_n^H \mathbf{R}_v^{1/2}\}) = [(\mathbf{R}_v^*)^{1/2} \otimes \mathbf{R}_v^{1/2}] \tilde{\mathbf{r}}$, it becomes

$$\Psi = \left[(\mathbf{R}_v^*)^{1/2} \otimes \mathbf{R}_v^{1/2} \right] (\tilde{\Psi} - \tilde{\mathbf{r}} \tilde{\mathbf{r}}^H) \left[(\mathbf{R}_v^*)^{1/2} \otimes \mathbf{R}_v^{1/2} \right]^H - \mathbf{r} \mathbf{r}^H.$$

Next, we will show an example of how to apply the *proposition* to analyze a channel estimator.

IV. APPLICATION EXAMPLE

In this section, we will show how to apply our analytical results to analyze a correlation-matching-based channel estimator proposed for a CDMA system in [4]. For simplicity, consider a K -user synchronous CDMA system. User k is assigned spreading codes $c_k(0), \dots, c_k(P-1)$. Its channel is assumed to have length q_k and corresponding coefficients are stacked in a vector \mathbf{g}_k . Then, in one symbol interval, the received data vector of length $\nu = P$ has a form [4]

$$\mathbf{y}_n = \sum_{k=1}^K [\mathbf{C}_{k,0} \mathbf{g}_k b_k(n) + \mathbf{C}_{k,-1} \mathbf{g}_k b_k(n-1)] + \mathbf{v}_n \quad (8)$$

where $b_k(n)$ and $b_k(n-1)$ are the k th user's current and previous symbols, respectively, $\mathbf{C}_{k,0}$ and $\mathbf{C}_{k,-1}$ are corresponding code matrices obtained from the spreading codes of user k , \mathbf{g}_k is the channel vector, and \mathbf{v}_n is the noise vector. Without loss of generality, assume user 1 is the desired user and the maximum channel length is q . The channel-estimation method [4] is based on minimizing the error between the correlation of (8), which reads

$$\mathbf{R} = \sum_{k=1}^K (\mathbf{C}_{k,0} \mathbf{G}_k \mathbf{C}_{k,0}^H + \mathbf{C}_{k,-1} \mathbf{G}_k \mathbf{C}_{k,-1}^H) + \sigma_v^2 \mathbf{I}_\nu \quad (9)$$

where $\mathbf{G}_k = \sigma_b^2 \mathbf{g}_k \mathbf{g}_k^H$ and its estimate. For convenience, define $\mathbf{x}_k = \text{vec}(\mathbf{G}_k)$. Collect all \mathbf{x}_k and σ_b^2 in an unknown vector \mathbf{x} . Then, after vec operation on \mathbf{R} : $\mathbf{r} = \text{vec}(\mathbf{R})$, it can be easily found that \mathbf{x} satisfies [4]

$$\mathbf{x} = (\mathbf{S}^H \mathbf{S})^{-1} \mathbf{S}^H \mathbf{r} \quad (10)$$

with \mathbf{S} being defined as

$$\mathbf{S} = [\mathbf{C}_{1,0}^* \otimes \mathbf{C}_{1,0} + \mathbf{C}_{1,-1}^* \otimes \mathbf{C}_{1,-1}, \dots, \mathbf{C}_{K,0}^* \otimes \mathbf{C}_{K,0} + \mathbf{C}_{K,-1}^* \otimes \mathbf{C}_{K,-1}, \text{vec}(\mathbf{I}_\nu)].$$

If \mathbf{r} is estimated from received data vectors, then \mathbf{x} can be estimated according to (10) under some identifiability conditions [4]. For the desired user, \mathbf{x}_1 can be extracted from \mathbf{x} . Then, \mathbf{G}_1 is reconstructed by the reverse vec operation. These operations can be described by

$$\hat{\mathbf{G}}_1 = [\mathbf{A}_1 \hat{\mathbf{r}}, \dots, \mathbf{A}_q \hat{\mathbf{r}}] \quad (11)$$

where q is the length of \mathbf{g}_1

$$\mathbf{A}_i = (\mathbf{e}_i^T \otimes \mathbf{I}_q) [\mathbf{I}_{q^2}, \mathbf{0}, \dots, \mathbf{0}]_{q^2 \times (Kq^2+1)} (\mathbf{S}^H \mathbf{S})^{-1} \mathbf{S}^H.$$

For $i = 1, \dots, q$, \mathbf{e}_i is a unitary vector of length q with only the i th element as 1 but zeros elsewhere. From our definition of \mathbf{G}_1 , \mathbf{g}_1 is an eigenvector corresponding to its unique nonzero eigenvalue. If $\hat{\mathbf{r}}$ has an estimation error $\delta \mathbf{r} = \hat{\mathbf{r}} - \mathbf{r}$ due to finite N , then an error is introduced to $\hat{\mathbf{G}}_1$. From (11), $\hat{\mathbf{G}}_1$ is perturbed by $\delta \mathbf{G}_1$ as

$$\delta \mathbf{G}_1 = [\mathbf{A}_1 \delta \mathbf{r}, \dots, \mathbf{A}_q \delta \mathbf{r}]. \quad (12)$$

Then, the first-order perturbation in this eigenvector becomes [25]

$$\delta \mathbf{g}_1 \approx \frac{1}{\sigma_b^2 \|\mathbf{g}_1\|^2} \mathbf{\Pi} \delta \mathbf{G}_1 \mathbf{g}_1, \quad \mathbf{\Pi} = \mathbf{\Sigma}_n \mathbf{\Sigma}_n^H \quad (13)$$

where $\mathbf{\Sigma}_n$ spans a $(q-1)$ -dimensional subspace orthogonal to \mathbf{g}_1 . Using (12) and (13), we obtain the covariance of the channel estimate

$$\text{Cov}(\delta \mathbf{g}_1) \approx \frac{1}{\sigma_b^2 \|\mathbf{g}_1\|^4} \mathbf{\Pi} \sum_{i,j=1}^q g_1(i) g_1^*(j) \mathbf{A}_i \mathbf{\Phi}(\hat{\mathbf{r}}) \mathbf{A}_j^H \mathbf{\Pi} \quad (14)$$

where $\mathbf{\Phi}(\hat{\mathbf{r}})$ is the covariance of $\hat{\mathbf{r}}$ given by *proposition*. Then, the channel mean square error (mse) is the trace of $\text{Cov}(\delta \mathbf{g}_1)$. This result may be useful for the design of constraints in robust multiuser detection techniques [22], [23]. Together with both (5) and (6), it will next be verified by computer simulations.

V. SIMULATIONS

First, we show how satisfactory the results from our *proposition* are when they are applied to predict the performance of a covariance estimator in practical conditions. We adopt the mse of estimated $\mathbf{\Psi}$ (instead of $\mathbf{\Phi}(\hat{\mathbf{r}})$ for clear observation of the error) to quantify the covariance estimation error. It is defined as the following normalized Frobenius squared norm of a difference matrix between our experimental result and its analytical result based on 100 independent realizations:

$$\frac{\|\hat{\mathbf{\Psi}} - \mathbf{\Psi}\|_F^2}{\|\mathbf{\Psi}\|_F^2}.$$

Parameters are chosen as $\nu = 5$, $L = 3$, and $\text{SNR} = 10$ dB. Input signals are taken from different constellations with statistics $\sigma_b^2 = 1$. Channel matrix \mathbf{H} in (1) is randomly and independently generated in each realization according to Gaussian distribution with zero mean

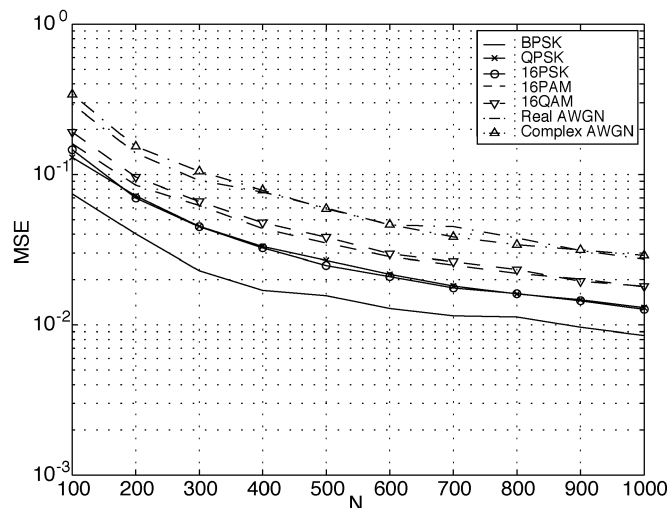


Fig. 1. Mean-square error (mse) of covariance estimation versus N with different inputs.

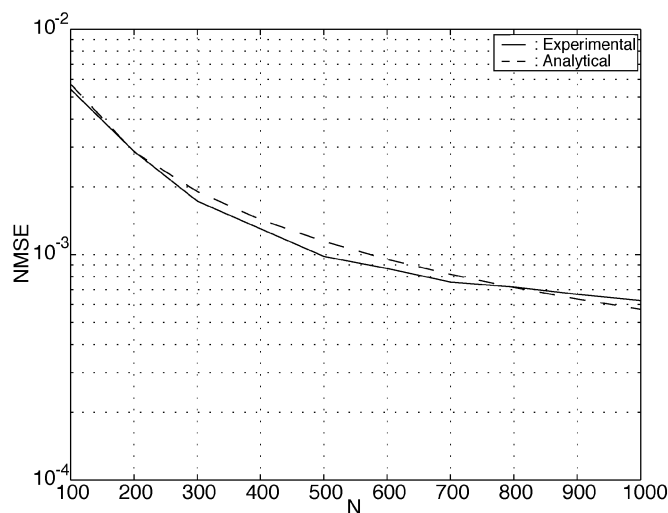


Fig. 2. Mean-square channel-estimation error versus N for a CDMA system.

and unit variance. Both real and complex systems are considered. In particular, effects of data modulations of binary phase-shift keying (BPSK), quaternary PSK (QPSK), 16 PSK, 16 pulse-amplitude modulation (PAM), 16 quadrature amplitude modulation (QAM) [26] and Gaussian sources are investigated. The mse versus N is plotted in Fig. 1. It can be observed that when N increases, the mse decreases, indicating that the covariance estimate becomes more accurate. It is also seen that PSK modulation gives the smallest error compared with amplitude modulation (AM) and Gaussian sources. Second, we apply our analytical results to a CDMA channel-estimation problem. Gold sequences of length 31 are used for a CDMA system with eight users of equal power, binary inputs, and 10-dB noise. Entries of users' channel vectors are Gaussian random variables with zero mean and unit variance. Each channel has three taps. Fig. 2 shows the normalized channel mse of the desired user. It is clear that the error decreases to a level of 10^{-3} even with only 500 data vectors. Again, the experimental result is highly consistent with the analytical result.

VI. CONCLUSION

In this correspondence, closed-form expressions for the covariance of a vectored sample covariance were derived for both real and complex

systems when a typical communication system model involves independent sources and AWGN. The results can be generalized to systems with colored Gaussian noise. Those analytical results can be applied to covariance-based algorithms. An application example and numerical examples were provided for explanation and verification.

APPENDIX I PROOF OF PROPOSITION

We will simplify Ψ based on (1). If we write $\hat{\mathbf{r}}_n$ as $\mathbf{y}_n^* \otimes \mathbf{y}_n$, then we obtain

$$\begin{aligned} E \left\{ \hat{\mathbf{r}}_n \hat{\mathbf{r}}_n^H \right\} &= (\mathbf{H}^* \otimes \mathbf{H}) E \left\{ \left(\mathbf{b}_n \mathbf{b}_n^H \right)^* \otimes \left(\mathbf{b}_n \mathbf{b}_n^H \right) \right\} \\ &\quad \times (\mathbf{H}^T \otimes \mathbf{H}^H) \\ &+ E \left\{ \left(\mathbf{v}_n \mathbf{v}_n^H \right)^* \otimes \left(\mathbf{v}_n \mathbf{v}_n^H \right) \right\} \\ &+ E \left\{ \left(\mathbf{H} \mathbf{b}_n \mathbf{v}_n^H \right)^* \otimes \left(\mathbf{H} \mathbf{b}_n \mathbf{v}_n^H \right) \right\} \\ &+ E \left\{ \left(\mathbf{v}_n \mathbf{b}_n^H \mathbf{H}^H \right)^* \otimes \left(\mathbf{v}_n \mathbf{b}_n^H \mathbf{H}^H \right) \right\} \\ &+ E \left\{ \left(\mathbf{H} \mathbf{b}_n \mathbf{v}_n^H \right)^* \otimes \left(\mathbf{v}_n \mathbf{b}_n^H \mathbf{H}^H \right) \right\} \\ &+ E \left\{ \left(\mathbf{v}_n \mathbf{b}_n^H \mathbf{H}^H \right)^* \otimes \left(\mathbf{H} \mathbf{b}_n \mathbf{v}_n^H \right) \right\} \\ &+ \sigma_b^2 (\mathbf{H}^* \mathbf{H}^T) \otimes \mathbf{I}_L \\ &+ \sigma_b^2 \mathbf{I}_L \otimes (\mathbf{H} \mathbf{H}^H). \end{aligned} \quad (15)$$

Further simplification of (15) requires differentiation between real and complex systems.

A. Case 1—Real Systems

The underlined term denoted as \mathbf{A} can be simplified as [21]

$$\mathbf{A} = \kappa_{4b} \mathbf{X}_1 + \sigma_b^4 \mathbf{X}_2 + \sigma_b^4 \text{vec}(\mathbf{I}_L) \text{vec}^T(\mathbf{I}_L) + \sigma_b^4 \mathbf{I}_{L^2} \quad (16)$$

where \mathbf{X}_1 and \mathbf{X}_2 are block matrices, as in [21]

$$\begin{aligned} \mathbf{X}_1 &= \text{diag} \left\{ \mathbf{a}_1 \mathbf{a}_1^T, \dots, \mathbf{a}_L \mathbf{a}_L^T \right\} \\ \mathbf{a}_l^T &= \left[0, \dots, 0, \underbrace{1}_{l\text{-th}}, 0, \dots, 0 \right]_{1 \times L} \\ \mathbf{X}_2 &= \left[\mathbf{a}_j \mathbf{a}_i^T \right]_{L \times L}. \end{aligned}$$

The first term of (15) then becomes

$$\begin{aligned} (\mathbf{H} \otimes \mathbf{H}) \mathbf{A} (\mathbf{H} \otimes \mathbf{H})^T &= \kappa_{4b} (\mathbf{H} \otimes \mathbf{H}) \mathbf{X}_1 (\mathbf{H} \otimes \mathbf{H})^T \\ &\quad + \sigma_b^4 (\mathbf{H} \otimes \mathbf{H}) \mathbf{X}_2 (\mathbf{H} \otimes \mathbf{H})^T \\ &\quad + \sigma_b^4 \text{vec}(\mathbf{H} \mathbf{H}^T) \text{vec}^T(\mathbf{H} \mathbf{H}^T) \\ &\quad + \sigma_b^4 (\mathbf{H} \mathbf{H}^T) \otimes (\mathbf{H} \mathbf{H}^T). \end{aligned} \quad (17)$$

Expressing \mathbf{X}_1 as $\text{diag} \{ \mathbf{a}_1, \dots, \mathbf{a}_L \} (\text{diag} \{ \mathbf{a}_1, \dots, \mathbf{a}_L \})^T$ and noticing that the l th column of $\text{diag} \{ \mathbf{a}_1, \dots, \mathbf{a}_L \}$ equals $\mathbf{a}_l \otimes \mathbf{a}_l$, we obtain $(\mathbf{H} \otimes \mathbf{H}) \mathbf{X}_1 (\mathbf{H} \otimes \mathbf{H})^T = (\mathbf{H} \square \mathbf{H}) (\mathbf{H} \square \mathbf{H})^T$. Noticing that $\mathbf{X}_2 = [\mathbf{I}_L \otimes \mathbf{a}_1, \dots, \mathbf{I}_L \otimes \mathbf{a}_L]$, we have $(\mathbf{H} \otimes \mathbf{H}) \mathbf{X}_2 = \mathbf{H} \odot_r \mathbf{H}$. After rewriting $\mathbf{H} \otimes \mathbf{H}$ as $\mathbf{H} \odot_l \mathbf{H}$, we obtain $(\mathbf{H} \otimes \mathbf{H}) \mathbf{X}_2 (\mathbf{H} \otimes \mathbf{H})^T = (\mathbf{H} \odot_r \mathbf{H}) (\mathbf{H} \odot_l \mathbf{H})^T$. Then (17) becomes

$$\begin{aligned} (\mathbf{H} \otimes \mathbf{H}) \mathbf{A} (\mathbf{H} \otimes \mathbf{H})^T &= \kappa_{4b} (\mathbf{H} \square \mathbf{H}) (\mathbf{H} \square \mathbf{H})^T \\ &\quad + \sigma_b^4 (\mathbf{H} \odot_r \mathbf{H}) (\mathbf{H} \odot_l \mathbf{H})^T \\ &\quad + \sigma_b^4 \text{vec}(\mathbf{H} \mathbf{H}^T) \text{vec}^T(\mathbf{H} \mathbf{H}^T) \\ &\quad + \sigma_b^4 (\mathbf{H} \mathbf{H}^T) \otimes (\mathbf{H} \mathbf{H}^T). \end{aligned} \quad (18)$$

Following similar procedures in simplifying \mathbf{A} and noticing the zero fourth-order cumulant for a Gaussian random variable, we obtain the second term of (15) [21]

$$\begin{aligned} E \left\{ \left(\mathbf{v}_n \mathbf{v}_n^T \right) \otimes \left(\mathbf{v}_n \mathbf{v}_n^T \right) \right\} &= \sigma_v^4 (\mathbf{I}_\nu \odot_r \mathbf{I}_\nu) (\mathbf{I}_\nu \odot_l \mathbf{I}_\nu)^T \\ &\quad + \sigma_v^4 \text{vec}(\mathbf{I}_\nu) \text{vec}^T(\mathbf{I}_\nu) \\ &\quad + \sigma_v^4 \mathbf{I}_{\nu^2}. \end{aligned} \quad (19)$$

Using $\mathbf{b}_n \otimes \mathbf{b}_n = \text{vec}(\mathbf{b}_n \mathbf{b}_n^T)$ and $\mathbf{v}_n \otimes \mathbf{v}_n = \text{vec}(\mathbf{v}_n \mathbf{v}_n^T)$, the third term in (15) can be easily found to be $E \{ (\mathbf{H} \mathbf{b}_n \mathbf{v}_n^T) \otimes (\mathbf{H} \mathbf{b}_n \mathbf{v}_n^T) \} = \sigma_b^2 \sigma_v^2 \text{vec}(\mathbf{H} \mathbf{H}^T) \text{vec}^T(\mathbf{I}_\nu)$. Its transpose becomes the fourth term. For the fifth term, noticing that $E \{ \mathbf{b}_n \otimes \mathbf{b}_n^T \} = \sigma_b^2 \mathbf{I}_L$, it can be easily verified that

$$E \left\{ \left(\mathbf{H} \mathbf{b}_n \mathbf{v}_n^T \right) \otimes \left(\mathbf{v}_n \mathbf{b}_n^T \mathbf{H}^T \right) \right\} = \sigma_b^2 E \left\{ \mathbf{H} \otimes \mathbf{v}_n \right\} \left(\mathbf{v}_n^T \otimes \mathbf{H}^T \right). \quad (20)$$

Explicitly expanding $\mathbf{v}_n^T \otimes \mathbf{H}^T$ by the definition of \otimes and combining the corresponding random variables to evaluate the expectation, (20) becomes

$$E \left\{ \left(\mathbf{H} \mathbf{b}_n \mathbf{v}_n^T \right) \otimes \left(\mathbf{v}_n \mathbf{b}_n^T \mathbf{H}^T \right) \right\} = \sigma_b^2 \sigma_v^2 (\mathbf{H} \odot_r \mathbf{I}_\nu) (\mathbf{I}_\nu \otimes \mathbf{H}^T). \quad (21)$$

The sixth term is the transpose of (21).

Considering the results above for all terms in (15), we obtain (5).

B. Case 2—Complex Systems

We will take similar steps in the following. First, \mathbf{A} is simplified as

$$\mathbf{A} = \kappa_{4b} \mathbf{X}_1 + \sigma_b^4 \text{vec}(\mathbf{I}_L) \text{vec}^T(\mathbf{I}_L) + \sigma_b^4 \mathbf{I}_{L^2}. \quad (22)$$

Similar to (18) and observing that no term involving \mathbf{X}_2 exists in (22) when compared with (16), we obtain

$$\begin{aligned} (\mathbf{H}^* \otimes \mathbf{H}) \mathbf{A} (\mathbf{H}^T \otimes \mathbf{H}^H) &= \kappa_{4b} (\mathbf{H}^* \square \mathbf{H}) (\mathbf{H}^* \square \mathbf{H})^H \\ &\quad + \sigma_b^4 \text{vec}(\mathbf{H} \mathbf{H}^H) \text{vec}^H(\mathbf{H} \mathbf{H}^H) \\ &\quad + \sigma_b^4 (\mathbf{H}^* \mathbf{H}^T) \otimes (\mathbf{H} \mathbf{H}^H). \end{aligned} \quad (23)$$

For the second term, we have [21]

$$E \left\{ \left(\mathbf{v}_n \mathbf{v}_n^H \right)^* \otimes \left(\mathbf{v}_n \mathbf{v}_n^H \right) \right\} = \sigma_v^4 \text{vec}(\mathbf{I}_\nu) \text{vec}^T(\mathbf{I}_\nu) + \sigma_v^4 \mathbf{I}_{\nu^2}. \quad (24)$$

The third term can be simplified as before. Its Hermitian equals the fourth term. Due to an independence assumption between the real and imaginary parts of a complex random variable, the fifth and sixth terms are both zeros. Considering all above, we obtain (6). \square

REFERENCES

- [1] H. Gazzah, P. Regalia, J. Delmas, and K. Abed-Meraim, "A blind multichannel identification algorithm robust to order overestimation," *IEEE Trans. Signal Processing*, vol. 50, pp. 1449–1458, June 2002.
- [2] G. Giannakis and S. Halford, "Asymptotically optimal blind fractionally spaced channel estimation and performance analysis," *IEEE Trans. Signal Processing*, vol. 45, pp. 1815–1830, July 1997.
- [3] J. Choi, "A moment-based blind method for identification of FIR down-link DS-SS channels," *IEEE Trans. Signal Processing*, vol. 48, pp. 966–981, Apr. 2000.
- [4] Z. Xu, "Asymptotically near-optimal blind estimation of multipath CDMA channels," *IEEE Trans. Signal Processing*, vol. 49, pp. 2003–2017, Sept. 2001.
- [5] Z. Xu and M. Tsatsanis, "Blind channel estimation for long code multiuser CDMA systems," *IEEE Trans. Signal Processing*, vol. 48, pp. 988–1001, Apr. 2000.
- [6] Z. Xu, "Low complexity multiuser channel estimation with aperiodic spreading codes," *IEEE Trans. Signal Processing*, vol. 49, pp. 2813–2822, Nov. 2001.

- [7] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd ed. San Diego, CA: Academic, 1985.
- [8] J. P. Burg, D. G. Luenberger, and D. L. Wenger, "Estimation of structured covariance matrices," *Proc. IEEE*, vol. 70, pp. 963–974, Sept. 1982.
- [9] B. D. van Veen and L. L. Scharf, "Estimation of structured covariance matrices and multiple window spectrum analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1467–1472, Aug. 1990.
- [10] D. B. Williams and D. H. Johnson, "Robust estimation of structured covariance matrices," *IEEE Trans. Signal Processing*, vol. 41, pp. 2891–2906, Sept. 1993.
- [11] H. Li, P. Stoica, and J. Li, "Computationally efficient maximum likelihood estimation of structured covariance matrices," *IEEE Trans. Signal Processing*, vol. 47, pp. 1314–1323, May 1999.
- [12] P. Stoica, E. G. Larsson, and A. B. Gershman, "The stochastic CRB for array processing: A textbook derivation," *IEEE Signal Processing Lett.*, vol. 8, pp. 148–150, May 2001.
- [13] M. Jansson and B. Ottersten, "Structured covariance matrix estimation: a parametric approach," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 5, Istanbul, Turkey, June 5–9, 2000, pp. 3172–3175.
- [14] D. R. Fuhrmann, "Application of Toeplitz covariance estimation to adaptive beamforming and detection," *IEEE Trans. Signal Processing*, vol. 39, pp. 2194–2198, Oct. 1991.
- [15] D. R. Brillinger, *Time Series Data Analysis and Theory*. San Francisco, CA: Holden-Day, 1981.
- [16] B. Friedlander and A. Weiss, "On the second-order statistics of the eigenvectors of sample covariance matrices," *IEEE Trans. Signal Processing*, vol. 46, pp. 3136–3139, Nov. 1998.
- [17] J.-F. Cardoso and E. Moulines, "A robustness property of DOA estimators based on covariance," *IEEE Trans. Signal Processing*, vol. 42, pp. 3285–3287, Nov. 1994.
- [18] J.-P. Delmas and Y. Meurisse, "Asymptotic performance of DOA finding algorithms with temporally correlated narrowband signals," *IEEE Trans. Signal Processing*, vol. 48, pp. 2669–2674, Sept. 2000.
- [19] M. Kristensson and B. Ottersten, "Statistical analysis of a subspace method for blind channel identification," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 5, Atlanta, GA, May 7–10, 1996, pp. 2435–2438.
- [20] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*, 2nd ed. Reading, MA: Addison-Wesley, 1994.
- [21] Z. Xu, "On the second-order statistics of the weighted sample covariance matrix," *IEEE Trans. Signal Processing*, vol. 51, pp. 527–534, Feb. 2003.
- [22] S. Cui, Z.-Q. Luo, and Z. Ding, "Robust blind multiuser detection against CDMA signature mismatch," in *Proc. ICASSP '01*, Salt Lake City, UT, May 2001, pp. 2297–2300.
- [23] A. B. Gershman and S. Shahbazpanahi, "Robust blind multiuser detection for synchronous CDMA systems," in *Proc. ICASSP '03*, vol. 4, Hong Kong, Apr. 2003, pp. 53–56.
- [24] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation and Modulation Theory*. New York: Wiley, 2002.
- [25] Z. Xu, "Perturbation analysis for subspace decomposition with applications in subspace-based algorithms," *IEEE Trans. Signal Processing*, vol. 50, pp. 2820–2830, Nov. 2002.
- [26] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.