Estimating Statistical Properties of MIMO Fading Channels

Aleksandar Dogandžić, *Member, IEEE,* and Jinghua Jin

***Abstract*—We propose maximum likelihood (ML) and restricted maximum likelihood (REML) methods for estimating the mean and covariance parameters of multi-input multi-output (MIMO) Ricean and Rayleigh block-fading channels using measurements from multiple coherent intervals containing both amplitudes and phases of the received signal. Correlated and independent fading scenarios with structured and unstructured line-of-sight (LOS) array response models are considered. Computationally efficient ML and approximate ML (AML) estimators are proposed for unitary space-time modulation schemes and orthogonal designs in correlated fading. We also derive Cramér–Rao bounds (CRBs) for the unknown parameters, discuss initialization of the proposed algorithms, and evaluate their performance via numerical simulations under the block- and continuous-fading scenarios.**

***Index Terms*—Cramér–Rao bound, maximum likelihood estimation, multi-input multi-output (MIMO) fading channels, Rayleigh fading, restricted maximum likelihood estimation, Ricean fading.**

I. Introduction

STATISTICAL modeling of multi-input multi-output (MIMO) Ricean and Rayleigh fading channels has recently attracted considerable attention; see [1]–[3] and references therein. The effects of correlated MIMO Rayleigh and Ricean fading on capacity and error-probability performance are discussed in [1]–[9]. However, it is assumed in [1]–[9] that statistical properties of the fading process are *known*. In [10], a method is proposed for consistent estimation of fading channel correlations in a single-input single-output frequency-selective Rayleigh fading scenario. In [11], an expectation–maximization (EM) algorithm is derived for estimating the mean and covariance parameters of a multivariate complex Ricean density from noiseless measurements and applied to polarimetric synthetic aperture radar (SAR). Most existing approaches to estimating fading-channel statistical properties do not account for noise effects and are based on signal-power measurements only; see, e.g., [12], [13], and references therein. In this paper (see also [14]), we present maximum likelihood (ML) and restricted maximum likelihood (REML) methods for estimating statistical properties of MIMO Ricean and Rayleigh block-fading channels using *complex noisy measurements* (containing *both* the phases and amplitudes of the received signals) from multiple coherent intervals. Knowing these properties is beneficial for (i) performance analysis [1]–[9] and design of wireless communication systems [6], [15], [16], (ii) implementation of space-time transmit precoding schemes that utilize mean and covariance feedback (see, e.g., [17] and references therein), (iii) antenna and constellation selection in spatial-multiplexing MIMO systems [18], (iv) implementation of noncoherent ML space-time receivers [5], [6], (v) mobile positioning [19], and (vi) channel estimation and sounding[1] [10], [22], [23, ch. 5.3.7]. Furthermore, the estimation methods developed herein are applicable to sensor array processing for moving arrays (which shares a similar measurement model; compare, e.g., the models in [24] and Section II).

We introduce the measurement model in Section II. In Section III, *expectation-conditional maximization either (ECME) algorithms*[2] are developed for computing the ML and REML estimates of the mean and covariance parameters of MIMO channels under correlated and independent block-fading scenarios (Sections III-A and B, respectively). In Section III-A3, we derive *closed-form* ML and approximate ML (AML) estimators for unitary space-time modulation schemes and orthogonal designs in correlated fading. We also derive Cramér–Rao bound (CRB) expressions for the unknown parameters (Appendix E and Section III-B1), discuss initialization of the proposed algorithms (Section III-A1), and evaluate their performance via numerical simulations under both block- and continuous-fading scenarios (Section IV). Concluding remarks are given in Section V.

II. Measurement Model

We adopt a *block-fading model* where the fading coefficients are constant within a coherent interval but vary randomly from one coherent interval to another. Assume that spatiotemporal measurements from K coherent intervals are available. Denote by $\boldsymbol{y}_k(t)$ an $n_{\mathrm{R}} \times 1$ data vector received by an array of n_{R} antennas at time $t \in \{1, 2, \ldots, N\}$ in the kth coherent interval, where $k \in \{1, 2, \ldots, K\}$. We consider the following measurement model:

$$\boldsymbol{y}_k(t) = H_k \boldsymbol{\phi}_k(t) + \boldsymbol{e}_k(t) \tag{2.1}$$

where $t = 1, \ldots, N,\ k = 1, \ldots, K$, and

- H_k is the $n_{\mathrm{R}} \times n_{\mathrm{T}}$ *channel response matrix*;

Manuscript received November 25, 2003; revised October 21, 2004. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Franz Hlawatsch.

The authors are with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011 USA (e-mail: ald@iastate.edu; jinjh@iastate.edu).

Digital Object Identifier 10.1109/TSP.2005.850877

[1]In particular, the channel mean and covariance parameters can be incorporated into channel estimation by utilizing the Bayesian linear-model minimum mean-square error (MMSE) estimator [20, Th. 11.1], [21], which outperforms the classical least-squares channel estimator [10].

[2]The ECME algorithms belong to the general class of EM algorithms; see [25]. The EM algorithms converge monotonically to a local or the global maximum of the likelihood function; see, e.g., [25, ch. 3] and [26, ch. 12.4].

- $\boldsymbol{\phi}_k(t)$ is an $n_{\mathrm{T}} \times 1$ vector of signals transmitted by n_{T} transmitter antennas and received by the receiver array at time t;
- $\boldsymbol{e}_k(t)$ is additive white complex Gaussian noise with

$$\mathrm{E}\big[\boldsymbol{e}_{k_1}(t_1)\boldsymbol{e}_{k_2}(t_2)^H\big] = \delta_{k_1,k_2}\delta_{t_1,t_2}\cdot\sigma^2 I_{n_{\mathrm{R}}}.$$

Here, δ_{ij} denotes the Kronecker delta symbol, I_n the identity matrix of size n, and "^{H}" the Hermitian (conjugate) transpose. We assume that the transmitted symbols are *known*, i.e., the coherent intervals contain training or previously detected data. Stacking all N time samples from the kth coherent interval into a single vector and using ([27, (2.11) in ch. 16]), we write (2.1) as

$$\boldsymbol{y}_k = Z_k\boldsymbol{h}_k + \boldsymbol{e}_k \tag{2.2a}$$

where $\boldsymbol{h}_k = \mathrm{vec}\{H_k\}$ is the $n_{\mathrm{R}}n_{\mathrm{T}}\times 1$ *channel response vector*, $\boldsymbol{y}_k = [\boldsymbol{y}_k(1)^T, \boldsymbol{y}_k(2)^T, \ldots, \boldsymbol{y}_k(N)^T]^T$ is the $n_{\mathrm{R}}N \times 1$ *spatiotemporal data vector*, $\boldsymbol{e}_k = [\boldsymbol{e}_k(1)^T, \boldsymbol{e}_k(2)^T, \ldots, \boldsymbol{e}_k(N)^T]^T$, and

$$Z_k = {\Phi_k}^T \otimes I_{n_{\mathrm{R}}}, \quad \Phi_k = [\boldsymbol{\phi}_k(1)\ldots\boldsymbol{\phi}_k(N)]. \tag{2.2b}$$

Here, "^{T}" denotes a transpose, $\otimes$ the Kronecker product, and the vec operator stacks the columns of a matrix one below another into a single column vector[3]. The $n_{\mathrm{T}} \times N$ matrix Φ_k in (2.2b) is the *signal matrix* in the kth coherent interval; we also define the "augmented" signal matrix Z_k of size $n_{\mathrm{R}}N \times n_{\mathrm{R}}n_{\mathrm{T}}$. The "vectorized" model (2.2) is used in [5] and [6] to describe MIMO measurements from a single coherent interval.

We now decompose the channel response vector $\boldsymbol{h}_k$ into a sum of the (deterministic) *line-of-sight (LOS) component* $\boldsymbol{h}_{\mathrm{LOS},k}$ and (random) *scattering component* $\boldsymbol{h}_{\mathrm{SC},k}$

$$\boldsymbol{h}_k = \boldsymbol{h}_{\mathrm{LOS},k} + \boldsymbol{h}_{\mathrm{SC},k}. \tag{2.3}$$

Let us adopt the following model for the LOS component:

$$\boldsymbol{h}_{\mathrm{LOS},k} = A_{\mathrm{LOS},k}\,\boldsymbol{x} \tag{2.4a}$$

where

- $A_{\mathrm{LOS},k}, k = 1, 2, \ldots, K$ are $n_{\mathrm{R}}n_{\mathrm{T}} \times r$ matrices;
- $\boldsymbol{x}$ is an $r \times 1$ vector of unknown complex coefficients.

The model (2.4a) is fairly general and can be used to describe the LOS component when dual-polarized antenna elements are employed.

When the transmitter and receiver LOS array responses are not known and the variation of the LOS component from one coherent interval to another can be described with a simple Doppler-shift model, we select $A_{\mathrm{LOS},k}$ as

$$A_{\mathrm{LOS},k} = \exp(j\omega_{\mathrm{D,LOS}}Nk)\cdot I_{n_{\mathrm{R}}n_{\mathrm{T}}} \tag{2.4b}$$

where $\omega_{\mathrm{D,LOS}}$ is the *LOS Doppler shift* (in radians) due to the relative movement between receiver and transmitter.[4] We assume that $\omega_{\mathrm{D,LOS}}$ is *known*, unless specified otherwise (see, e.g., Figs. 4, 9, and 10 in Section IV). Here, (2.4a) simplifies to $\boldsymbol{h}_{\mathrm{LOS},k} = \exp(j\omega_{\mathrm{D,LOS}}Nk)\cdot\boldsymbol{x}$, where $\boldsymbol{x}$ is the $n_{\mathrm{R}}n_{\mathrm{T}} \times 1$ *unstructured LOS array response vector*; hence, $r = n_{\mathrm{R}}n_{\mathrm{T}}$ in this case.

If the transmitter and receiver LOS array responses are known, we utilize the *structured LOS array response model* (see [6], [14], [28], and [29]):

$$A_{\mathrm{LOS},k} = \exp(j\omega_{\mathrm{D,LOS}}Nk)\cdot\boldsymbol{a}_{\mathrm{T,LOS}}\otimes\boldsymbol{a}_{\mathrm{R,LOS}} \tag{2.4c}$$

where $\boldsymbol{a}_{\mathrm{T,LOS}}$ and $\boldsymbol{a}_{\mathrm{R,LOS}}$ are the transmitter and receiver LOS array response vectors of dimensions $n_{\mathrm{T}} \times 1$ and $n_{\mathrm{R}} \times 1$, respectively. Now, (2.4a) becomes $\boldsymbol{h}_{\mathrm{LOS},k} = \boldsymbol{a}_{\mathrm{T,LOS}} \otimes \boldsymbol{a}_{\mathrm{R,LOS}}\exp(j\omega_{\mathrm{D,LOS}}Nk)\cdot x$, where $\boldsymbol{x} = x$ is the scalar LOS complex amplitude (implying $r = 1$), and $A_{\mathrm{LOS},k}$ is an $n_{\mathrm{R}}n_{\mathrm{T}} \times 1$ vector.

To describe the channel variation from one coherent interval to another, we assume that the scattering channel vectors $\boldsymbol{h}_{\mathrm{SC},k}$ are zero-mean independent, identically distributed (i.i.d.) complex Gaussian, with an $n_{\mathrm{R}}n_{\mathrm{T}} \times n_{\mathrm{R}}n_{\mathrm{T}}$ covariance matrix:

$$\Psi = \mathrm{E}[\boldsymbol{h}_{\mathrm{SC},k}\boldsymbol{h}_{\mathrm{SC},k}^H], \quad k = 1, 2, \ldots, K. \tag{2.5}$$

In addition, $\boldsymbol{h}_{\mathrm{SC},k}$ and noise vectors $\boldsymbol{e}_k$ are assumed to be independent, i.e., $\mathrm{E}[\boldsymbol{e}_{k_1}\boldsymbol{h}_{\mathrm{SC},k_2}^H] = 0$, where $k_1, k_2 \in \{1, 2, \ldots, K\}$.

Our goal is to estimate the unknown parameters in the above model:

- the LOS coefficient vector $\boldsymbol{x}$;
- spatial fading covariance matrix Ψ;
- noise variance σ^2;

which are assumed to be *constant* over the K coherent intervals. This assumption is justified by the fact that the channel mean and covariance parameters depend on large-scale variations in the scattering environment, which are typically slow (see also the discussion in, e.g., [15] and [22]). Define the vector of unknown parameters:

$$\boldsymbol{\rho} = [\mathrm{Re}\{\boldsymbol{x}\}^T, \mathrm{Im}\{\boldsymbol{x}\}^T, \boldsymbol{\gamma}^T]^T \tag{2.6}$$

where

$$\boldsymbol{\gamma} = [\sigma^2, \boldsymbol{\psi}^T]^T \tag{2.7}$$

is the vector of *variance components*, and $\boldsymbol{\psi}$ describes a parametrization of the fading covariance matrix Ψ. We consider two models for Ψ:

i) **unstructured (correlated fading)**

$$\boldsymbol{\psi} = [\mathrm{Re}\{\mathrm{vech}(\Psi)\}^T, \mathrm{Im}\{\underline{\mathrm{vech}}(\Psi)\}^T]^T$$

where the correlation structure of the fading channel is completely unknown;

ii) **diagonal (independent fading)**

$$\Psi = \mathrm{diag}(\psi_1, \psi_2, \ldots, \psi_{n_{\mathrm{R}}n_{\mathrm{T}}})$$

and

$$\boldsymbol{\psi} = [\psi_1, \psi_2, \ldots, \psi_{n_{\mathrm{R}}n_{\mathrm{T}}}]^T$$

where the fading-channel coefficients are independent with nonequal variances.

Here, the vech and $\underline{\mathrm{vech}}$ operators create a single column vector by stacking elements below the main diagonal columnwise; vech includes the main diagonal, whereas $\underline{\mathrm{vech}}$ omits it. Note

[3]For the definition and properties of the Kronecker product and vec operator, see [27, ch. 16].

[4]Note that $\omega_{\mathrm{D,LOS}}$ corresponds to the continuous-time LOS Doppler shift $\Omega_{\mathrm{D,LOS}} = \omega_{\mathrm{D,LOS}}/\Delta t$, where Δt is the symbol duration.

that $\boldsymbol{\psi}$ is a valid parametrization *only if* Ψ is a positive semidefinite Hermitian matrix.

In the following, we derive ML and REML algorithms for estimating the unknown parameter vector $\boldsymbol{\rho}$ under the two fading scenarios described above. We also derive efficient algorithms for estimating $\boldsymbol{\rho}$ when $\Phi_k\Phi_k^H$ is a constant or an identity matrix.

III. ML and REML Estimation

We first outline the ML and REML approaches to estimating $\boldsymbol{\rho}$ and then present the proposed algorithms.

Under the measurement model in Section II, the spatiotemporal data vectors $\boldsymbol{y}_k$ are independent, complex Gaussian with means and covariances

$$\mathrm{E}[\boldsymbol{y}_k] = \Upsilon_k\boldsymbol{x} = Z_k A_{\mathrm{LOS},k}\boldsymbol{x} \tag{3.1a}$$

$$\begin{aligned}\Sigma_k(\boldsymbol{\gamma}) &= \mathrm{cov}(\boldsymbol{y}_k) = \mathrm{E}[(\boldsymbol{y}_k - \mathrm{E}[\boldsymbol{y}_k])\,(\boldsymbol{y}_k - \mathrm{E}[\boldsymbol{y}_k])^H] \\ &= Z_k\Psi Z_k^H + \sigma^2 I_{n_{\mathrm{R}}N}.\end{aligned} \tag{3.1b}$$

Thus, the log-likelihood function to be maximized is the logarithm of the joint probability density function (pdf) of $\boldsymbol{y}_k, k = 1, 2, \ldots, K$:

$$L(\boldsymbol{\rho}) = -(\boldsymbol{y} - \Upsilon\boldsymbol{x})^H\Sigma(\boldsymbol{\gamma})^{-1}(\boldsymbol{y} - \Upsilon\boldsymbol{x}) - \ln|\pi\Sigma(\boldsymbol{\gamma})| \tag{3.2}$$

where $|\cdot|$ denotes the determinant, $\boldsymbol{y} = [\boldsymbol{y}_1^T, \boldsymbol{y}_2^T \ldots \boldsymbol{y}_K^T]^T$, $\Upsilon = [\Upsilon_1^T, \Upsilon_2^T \ldots \Upsilon_K^T]^T$ is an $Kn_{\mathrm{R}}N \times r$ matrix of rank r, and $\Sigma(\boldsymbol{\gamma})$ is an $Kn_{\mathrm{R}}N \times Kn_{\mathrm{R}}N$ block-diagonal matrix:

$$\Sigma(\boldsymbol{\gamma}) = \mathrm{bdiag}\{\Sigma_1(\boldsymbol{\gamma}), \Sigma_2(\boldsymbol{\gamma}) \ldots \Sigma_K(\boldsymbol{\gamma})\}.$$

The estimate of the LOS coefficient vector $\boldsymbol{x}$ that maximizes (3.2) for any fixed $\boldsymbol{\gamma}$ is given by

$$\begin{aligned}\hat{\boldsymbol{x}}(\boldsymbol{\gamma}) &= [\Upsilon^H\Sigma(\boldsymbol{\gamma})^{-1}\Upsilon]^{-1}\Upsilon^H\Sigma(\boldsymbol{\gamma})^{-1}\boldsymbol{y} \\ &= \left[\sum_{k=1}^{K}\mathcal{A}_{\mathrm{LOS},k}^H W_k(\boldsymbol{\gamma})\mathcal{A}_{\mathrm{LOS},k}\right]^{-1} \\ &\quad\times \sum_{k=1}^{K}\mathcal{A}_{\mathrm{LOS},k}^H W_k(\boldsymbol{\gamma})\boldsymbol{z}_k\end{aligned} \tag{3.3}$$

where

$$W_k(\boldsymbol{\gamma}) = \left(\sigma^2 Z_k^H Z_k + Z_k^H Z_k \Psi Z_k^H Z_k\right)^{-1} \tag{3.4a}$$

$$\mathcal{A}_{\mathrm{LOS},k} = Z_k^H\Upsilon_k = Z_k^H Z_k A_{\mathrm{LOS},k} \tag{3.4b}$$

$$\boldsymbol{z}_k = Z_k^H\boldsymbol{y}_k = \mathrm{vec}\left(Y_k\Phi_k^H\right) \tag{3.4c}$$

and

$$Y_k = [\boldsymbol{y}_k(1) \ldots \boldsymbol{y}_k(N)]$$

is the *spatiotemporal data matrix* in the kth coherent interval. The second equalities in (3.3) and (3.4c) follow by using (A.2) in Appendix A and [27, (16.2.11)], respectively. Replacing $\boldsymbol{x}$ in (3.2) with its ML estimate in (3.3) yields the concentrated log-likelihood function:

$$\begin{aligned}&L(\boldsymbol{\gamma}\,|\,\hat{\boldsymbol{x}}(\boldsymbol{\gamma})) \\ &\quad= -[\boldsymbol{y} - \Upsilon\hat{\boldsymbol{x}}(\boldsymbol{\gamma})]^H\Sigma(\boldsymbol{\gamma})^{-1}[\boldsymbol{y} - \Upsilon\hat{\boldsymbol{x}}(\boldsymbol{\gamma})] - \ln|\pi\Sigma(\boldsymbol{\gamma})| \\ &\quad= -\boldsymbol{y}^H\Pi(\boldsymbol{\gamma})\boldsymbol{y} - \ln|\pi\Sigma(\boldsymbol{\gamma})|\end{aligned} \tag{3.5}$$

where

$$\Pi(\boldsymbol{\gamma}) = \Sigma(\boldsymbol{\gamma})^{-1} - \Sigma(\boldsymbol{\gamma})^{-1}\Upsilon[\Upsilon^H\Sigma(\boldsymbol{\gamma})^{-1}\Upsilon]^{-1}\Upsilon^H\Sigma(\boldsymbol{\gamma})^{-1}.$$

Note that (3.5) is a nonlinear function of the variance-component parameters $\boldsymbol{\gamma}$ that generally needs to be maximized using iterative algorithms. Once the ML estimate $\hat{\boldsymbol{\gamma}}$ is computed by maximizing (3.5), the ML estimate of $\boldsymbol{x}$ is obtained by substituting $\hat{\boldsymbol{\gamma}}$ into (3.3). Interestingly, closed-form solutions for the ML estimates of $\boldsymbol{\gamma}$ exist under the correlated fading scenario with constant $\Phi_k\Phi_k^H$; see Section III-A3.

We now introduce the REML method for estimating the unknown variance components. The REML estimate of $\boldsymbol{\gamma}$ is obtained by filtering out the deterministic Ricean component from the received data and applying ML estimation to the *error contrasts* (i.e., filtered data), which corresponds to maximizing the REML log-likelihood function (see Appendix B)

$$\begin{aligned}L_{\mathrm{REML}}(\boldsymbol{\gamma}) &= L(\boldsymbol{\gamma}\,|\,\hat{\boldsymbol{x}}(\boldsymbol{\gamma})) - \ln|\Upsilon^H\Sigma(\boldsymbol{\gamma})^{-1}\Upsilon| \\ &= L(\boldsymbol{\gamma}\,|\,\hat{\boldsymbol{x}}(\boldsymbol{\gamma})) - \ln\left|\sum_{k=1}^{K}\Upsilon_k^H\Sigma_k(\boldsymbol{\gamma})^{-1}\Upsilon_k\right|\end{aligned} \tag{3.6}$$

with respect to $\boldsymbol{\gamma}$, where $L(\boldsymbol{\gamma}\,|\,\hat{\boldsymbol{x}}(\boldsymbol{\gamma}))$ is the concentrated log likelihood in (3.5). We can also derive (3.6) by using an *integrated-likelihood approach* for eliminating nuisance parameters [30]. Here, we treat $\boldsymbol{x}$ as a nuisance parameter vector and integrate it out using a noninformative prior[5]; see [31, ch. 3.3.3]. The REML method provides only estimates of the variance components $\boldsymbol{\gamma}$; however, a good estimate of $\boldsymbol{x}$ is obtained by substituting the REML estimate of $\boldsymbol{\gamma}$ into (3.3), which we call the "REML" estimate of $\boldsymbol{x}$ (with a slight abuse of terminology). Since the Ricean component has been filtered out, the REML estimate of $\boldsymbol{\gamma}$ is invariant to the value of $\boldsymbol{x}$, i.e., changing $\boldsymbol{x}$ does not alter the REML estimate of $\boldsymbol{\gamma}$. Finally, the REML estimates of the variance components have *smaller bias* than the corresponding ML estimates; see [31] and [32].

In Appendix E, we compute general CRB expressions for the unknown parameters $\boldsymbol{\rho}$, assuming an arbitrary parametrization $\Psi = \Psi(\boldsymbol{\psi})$ of the fading covariance matrix. We then specialize these general results to the independent and correlated fading scenarios with constant $\Phi_k\Phi_k^H$ (e.g., equal to the identity matrix); see (3.23) in Section III-B1 and Appendices E-A and E-B.

A. Correlated Fading

We compute the ML and REML estimates of the unknown parameters under the correlated fading scenario. ECME algorithms for arbitrary $\Phi_k\Phi_k^H$ are presented in Sections III-A1 and 2. In Section III-A3, we derive an alternating-projection ML algorithm for the case where $\Phi_k\Phi_k^H$ are independent of k.

1) ECME Algorithm for ML Estimation: In Appendix C.A, we derive an ML ECME algorithm for estimating $\boldsymbol{\rho}$ under the correlated fading scenario. An ECME algorithm maximizes *either* the expected complete-data log-likelihood function (where

[5]Choices of a noninformative prior pdf for $\boldsymbol{x}$ could be a complex Gaussian with an arbitrary mean and a covariance matrix whose inverse is a zero matrix, or $\mathrm{Re}\{\boldsymbol{x}\}, \mathrm{Im}\{\boldsymbol{x}\} \in \mathrm{uniform}(-\infty, \infty)$; see [31, p. 84].

the expectation is computed with respect to the conditional distribution of the unobserved data given the observed measurements) *or* the actual observed-data log-likelihood; see [25, ch. 5.7], [33], and [34]. Here, we treat the scattering channel vectors $\boldsymbol{h}_{\mathrm{SC},k}, k = 1, 2, \ldots, K$ as the *unobserved* (or missing) data and derive the following ECME algorithm: Iterate between

$$W_k^{(i)} = \left[(\sigma^2)^{(i)} Z_k^H Z_k + Z_k^H Z_k \Psi^{(i)} Z_k^H Z_k\right]^{-1} \quad (3.7a)$$

$$\boldsymbol{x}^{(i)} = \left(\sum_{k=1}^{K} \mathcal{A}_{\mathrm{LOS},k}^H W_k^{(i)} \mathcal{A}_{\mathrm{LOS},k}\right)^{-1} \times \sum_{k=1}^{K} \mathcal{A}_{\mathrm{LOS},k}^H W_k^{(i)} \boldsymbol{z}_k \quad (3.7b)$$

$$\boldsymbol{h}_{\mathrm{SC},k}^{(i)} = \Psi^{(i)} Z_k^H Z_k W_k^{(i)} \left(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x}^{(i)}\right) \quad (3.7c)$$

for $k = 1, 2, \ldots, K$, and

$$(\sigma^2)^{(i+1)} = \tilde{\sigma}^2 + \frac{(\sigma^2)^{(i)}}{K n_{\mathrm{R}} N} \sum_{k=1}^{K} \left(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x}^{(i)}\right)^H \times W_k^{(i)} \left(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x}^{(i)}\right) \quad (3.8a)$$

$$\Psi^{(i+1)} = \Psi^{(i)} + \frac{1}{K} \sum_{k=1}^{K} \left\{ \boldsymbol{h}_{\mathrm{SC},k}^{(i)} \left(\boldsymbol{h}_{\mathrm{SC},k}^{(i)}\right)^H - \Psi^{(i)} Z_k^H Z_k W_k^{(i)} Z_k^H Z_k \Psi^{(i)} \right\} \quad (3.8b)$$

where

$$\tilde{\sigma}^2 = \frac{1}{K n_{\mathrm{R}} N} \sum_{k=1}^{K} \left[\boldsymbol{y}_k^H \boldsymbol{y}_k - \boldsymbol{z}_k^H \left(Z_k^H Z_k\right)^{-1} \boldsymbol{z}_k\right] = \frac{1}{K n_{\mathrm{R}} N} \sum_{k=1}^{K} \mathrm{tr}\Big\{ Y_k \Big[I_N - \Phi_k^H \left(\Phi_k \Phi_k^H\right)^{-1} \Phi_k \Big] {Y_k}^H \Big\}. \quad (3.8c)$$

The matrices $\mathcal{A}_{\mathrm{LOS},k}$ and vectors $\boldsymbol{z}_k, k = 1, 2, \ldots, K$ [defined in (3.4b) and (3.4c)] and $\tilde{\sigma}^2$ (above) do not depend on the unknown parameters and can therefore be computed beforehand. In addition, $\boldsymbol{h}_{\mathrm{SC},k}^{(i)}$ in (3.7c) are the (estimated) Bayesian linear-model MMSE estimators of the scattering channel vectors $\boldsymbol{h}_{\mathrm{SC},k}$, with $\boldsymbol{x}$, Ψ, and σ^2 replaced by their ith-iteration estimates (see also Footnote 1). To derive the second equality in (3.8c), we have used the property of the vec operator in [27, Th. 16.2.2].

Initialization: The above iteration can be initialized with $\Psi^{(-1)} = 0$, implying that the initial estimate $\boldsymbol{x}^{(-1)}$ of the LOS coefficient vector is simply its linear *least-squares (LS) estimate* [see also (3.3)]:

$$\boldsymbol{x}^{(-1)} = \left(\sum_{k=1}^{K} \Upsilon_k^H \Upsilon_k\right)^{-1} \sum_{k=1}^{K} \Upsilon_k^H \boldsymbol{y}_k = \left(\sum_{k=1}^{K} A_{\mathrm{LOS},k}^H Z_k^H Z_k A_{\mathrm{LOS},k}\right)^{-1} \sum_{k=1}^{K} A_{\mathrm{LOS},k}^H \boldsymbol{z}_k. \quad (3.9)$$

After computing $\boldsymbol{x}^{(-1)}$, a good initial estimate of Ψ is its *modified method-of-moments estimate* (similar to [35, p. 244]):

$$\Psi^{(0)} = \left[\frac{1}{K} \sum_{k=1}^{K} \left(Z_k^H Z_k\right)^{-1} \left(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x}^{(-1)}\right) \times \left(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x}^{(-1)}\right)^H \left(Z_k^H Z_k\right)^{-1}\right] - \min\left\{(\sigma^2)^{(0)}, \hat{\lambda}\right\} \cdot \frac{1}{K} \sum_{k=1}^{K} \left(Z_k^H Z_k\right)^{-1} \quad (3.10a)$$

where $(\sigma^2)^{(0)}$ is a method-of-moments estimate of σ^2:

$$(\sigma^2)^{(0)} = \frac{1}{K n_{\mathrm{R}} (N - n_{\mathrm{T}})} \times \sum_{k=1}^{K} \left[\boldsymbol{y}_k^H \boldsymbol{y}_k - \boldsymbol{z}_k^H \left(Z_k^H Z_k\right)^{-1} \boldsymbol{z}_k\right] = \frac{K n_{\mathrm{R}} N}{K n_{\mathrm{R}} (N - n_{\mathrm{T}})} \cdot \tilde{\sigma}^2 \quad (3.10b)$$

which is a good initial estimate of σ^2, and $\hat{\lambda}$ is the smallest generalized eigenvalue of the matrices $(1/K) \cdot \sum_{k=1}^{K} (Z_k^H Z_k)^{-1} (\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x}^{(-1)}) (\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x}^{(-1)})^H (Z_k^H Z_k)^{-1}$ and $(1/K) \cdot \sum_{k=1}^{K} (Z_k^H Z_k)^{-1}$. Note that $(Z_k^H Z_k)^{-1}$ can be efficiently computed as

$$\left(Z_k^H Z_k\right)^{-1} = \left(\Phi_k {}^* \Phi_k {}^T\right)^{-1} \otimes I_{n_{\mathrm{R}}}$$

where * denotes complex conjugation. The moment estimator in (3.10a) follows by pre- and post-multiplying $Z_k^H (\boldsymbol{y}_k - \mathrm{E}[\boldsymbol{y}_k])(\boldsymbol{y}_k - \mathrm{E}[\boldsymbol{y}_k])^H Z_k$ by $(Z_k^H Z_k)^{-1}$, summing over $k = 1, 2 \ldots, K$, taking the expectation of the resulting expression, and solving for Ψ; the moment estimator in (3.10b) follows by verifying that its expectation is σ^2. To derive (3.10a), we applied a modification similar to [35, p. 244] to ensure that $\Psi^{(0)}$ is always a valid (i.e., positive semidefinite) covariance matrix.

2) ECME Algorithm for REML Estimation: In Appendix C-B, we derive an ECME algorithm for REML estimation of $\boldsymbol{\rho}$, which follows by replacing (3.8a) and (3.8b) with

$$(\sigma^2)^{(i+1)} = \frac{K n_{\mathrm{R}} N}{K n_{\mathrm{R}} N - r} \cdot \tilde{\sigma}^2 + \frac{(\sigma^2)^{(i)}}{K n_{\mathrm{R}} N - r} \cdot \sum_{k=1}^{K} \left(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x}^{(i)}\right)^H \times W_k^{(i)} \left(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x}^{(i)}\right) \quad (3.11a)$$

$$\Psi^{(i+1)} = \Psi^{(i)} + \frac{1}{K} \sum_{k=1}^{K} \left\{ \boldsymbol{h}_{\mathrm{SC},k}^{(i)} \left(\boldsymbol{h}_{\mathrm{SC},k}^{(i)}\right)^H - \Psi^{(i)} Z_k^H Z_k \cdot \left[W_k^{(i)} - W_k^{(i)} \mathcal{A}_{\mathrm{LOS},k} \left(\sum_{l=1}^{K} \mathcal{A}_{\mathrm{LOS},l}^H W_l^{(i)} \mathcal{A}_{\mathrm{LOS},l} \right)^{-1} \times \mathcal{A}_{\mathrm{LOS},k}^H W_k^{(i)} \right] Z_k^H Z_k \Psi^{(i)} \right\} \quad (3.11b)$$

in the iteration (3.7) and (3.8) and keeping the other steps intact.

The above ML and REML ECME algorithms *always* converge to estimates that are in the parameter space: $(\sigma^2)^{(i)} \geq 0$ and $\Psi^{(i)} \geq 0$ (i.e., $\Psi^{(i)}$ is positive semidefinite) at each iteration step i, provided that the initial values are in the parameter space. This is an important general property of the EM and related algorithms (such as ECME), see [26, ch. 12.4]. The fact that $\Psi^{(i+1)} \geq 0$ in (3.8b) follows from the derivation in Appendix C; see (C.4) and (C.7), and observe that a sample covariance matrix is always positive semidefinite. A similar argument applies to the REML case since the ECME REML algorithm is simply the ECME ML algorithm applied to the error contrasts. Rayleigh-fading versions of the above algorithms are obtained by removing the step (3.7b) and setting $\mathcal{A}_{\mathrm{LOS},k} = 0$ in (3.7c) and (3.8a) or (3.11a).

3) ML Estimation for Constant $\Phi_k \Phi_k^H$: We now present computationally efficient estimators for the scenario where

$$\Phi_k \Phi_k^H = \Gamma_\Phi = \text{constant} \tag{3.12a}$$

(independent of $k \in \{1, 2, \ldots, K\}$), which holds for many practically important signaling schemes, e.g., unitary space-time codes [36] and space-time block codes based on orthogonal designs [37], [38]. The above condition implies that $Z_k^H Z_k$ also does not depend on k; hence, we define

$$C = Z_k^H Z_k = {\Gamma_\Phi}^T \otimes I_{n_\mathrm{R}}. \tag{3.12b}$$

In this case, there exists a closed-form expression for the ML estimate of σ^2 (see Appendix D):

$$\begin{aligned}\hat{\sigma}^2_{\mathrm{ML}} &= \frac{1}{K n_\mathrm{R}(N - n_\mathrm{T})} \sum_{k=1}^{K} \left(\boldsymbol{y}_k^H \boldsymbol{y}_k - \boldsymbol{z}_k^H C^{-1} \boldsymbol{z}_k \right) \\ &= \frac{1}{K n_\mathrm{R}(N - n_\mathrm{T})} \\ &\quad \cdot \sum_{k=1}^{K} \mathrm{tr}\left[Y_k (I_N - \Phi_k^H {\Gamma_\Phi}^{-1} \Phi_k) Y_k^H \right]\end{aligned} \tag{3.13}$$

which coincides with the moment estimator in (3.10b). The exact ML estimates of $\boldsymbol{x}$ and Ψ can be computed by iterating between the following two steps (see Appendix D):

$$W^{(i)} = \left(\hat{\sigma}^2_{\mathrm{ML}} C + C \Psi^{(i)} C \right)^{-1} \tag{3.14a}$$

$$\begin{aligned}\boldsymbol{x}^{(i)} &= \left[\sum_{k=1}^{K} \mathcal{A}_{\mathrm{LOS},k}^H W^{(i)} \mathcal{A}_{\mathrm{LOS},k} \right]^{-1} \\ &\quad \times \sum_{k=1}^{K} \mathcal{A}_{\mathrm{LOS},k}^H W^{(i)} \boldsymbol{z}_k\end{aligned} \tag{3.14b}$$

$$\begin{aligned}\boldsymbol{\xi}_k^{(i)} &= C^{-1} \left(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x}^{(i)} \right) \\ &= \mathrm{vec}\left(Y_k \Phi_k^H {\Gamma_\Phi}^{-1} \right) - A_{\mathrm{LOS},k} \boldsymbol{x}^{(i)}\end{aligned} \tag{3.14c}$$

for $k = 1, 2, \ldots, K$, and

$$\Psi^{(i+1)} = \left[\frac{1}{K} \sum_{k=1}^{K} \boldsymbol{\xi}_k^{(i)} (\boldsymbol{\xi}_k^{(i)})^H \right] - \hat{\sigma}^2_{\mathrm{ML}} \cdot \left({\Gamma_\Phi}^T \right)^{-1} \otimes I_{n_\mathrm{R}}. \tag{3.15}$$

The above iteration increases the log-likelihood function (3.2) at each cycle but may converge to solutions that are not in the parameter space; see also the discussion below. It can be shown that if (3.12) holds, the estimators (3.13) and (3.15) are fixed points of the ECME iterations (3.8a) and (3.8b).

In Appendix E-A, we derive the CRB expressions for this scenario.

ML Estimation for Unstructured LOS Array Response Model: Under the unstructured LOS array response model

$$\begin{aligned}\hat{\boldsymbol{x}}_{\mathrm{UML}} &= C^{-1} \cdot \frac{1}{K} \sum_{k=1}^{K} \exp(-j \omega_{\mathrm{D,LOS}} N k) \boldsymbol{z}_k \\ &= \frac{1}{K} \sum_{k=1}^{K} \mathrm{vec}\left(Y_k \Phi_k^H {\Gamma_\Phi}^{-1} \right) \exp(-j \omega_{\mathrm{D,LOS}} N k)\end{aligned} \tag{3.16a}$$

$$\hat{\Psi}_{\mathrm{UML}} = \left[\frac{1}{K} \sum_{k=1}^{K} \hat{\boldsymbol{\xi}}_k \hat{\boldsymbol{\xi}}_k^H \right] - \hat{\sigma}^2_{\mathrm{ML}} \cdot \left({\Gamma_\Phi}^T \right)^{-1} \otimes I_{n_\mathrm{R}} \tag{3.16b}$$

are the *closed-form expressions* for the ML estimates of $\boldsymbol{x}$ and Ψ, where $\hat{\sigma}^2_{\mathrm{ML}}$ is the ML estimate of σ^2 in (3.13), and

$$\hat{\boldsymbol{\xi}}_k = \mathrm{vec}\left(Y_k \Phi_k^H {\Gamma_\Phi}^{-1} \right) - \exp(j \omega_{\mathrm{D,LOS}} N k) \cdot \hat{\boldsymbol{x}}_{\mathrm{UML}} \tag{3.16c}$$

for $k = 1, 2, \ldots, K$. Here, (3.16a) follows by substituting (2.4b) into (3.3) [see also (3.4b) and (3.12b)], and (3.16b) is obtained by substituting $\hat{\boldsymbol{x}}_{\mathrm{UML}}$ into (D.2c); see Appendix D.

If the LOS Doppler shift $\omega_{\mathrm{D,LOS}}$ is *unknown*, its ML estimate can be computed by maximizing the following concentrated log-likelihood function:

$$\begin{aligned}\hat{\omega}_{\mathrm{D,LOS}} = \arg \max_{\omega_{\mathrm{D,LOS}}} &\boldsymbol{z}_{\mathrm{DTFT}}(\omega_{\mathrm{D,LOS}} N)^H \\ &\times \left(\sum_{k=1}^{K} \boldsymbol{z}_k \boldsymbol{z}_k^H \right)^{-1} \boldsymbol{z}_{\mathrm{DTFT}}(\omega_{\mathrm{D,LOS}} N)\end{aligned} \tag{3.17}$$

where

$$\begin{aligned}\boldsymbol{z}_{\mathrm{DTFT}}(\omega) &= \frac{1}{K} \cdot \sum_{k=1}^{K} \exp(-j \omega k) \boldsymbol{z}_k \\ &= \frac{1}{K} \cdot \sum_{k=1}^{K} \exp(-j \omega k) \mathrm{vec}\left(Y_k \Phi_k^H \right)\end{aligned} \tag{3.18}$$

is proportional to the discrete-time Fourier transform (DTFT) of $\boldsymbol{z}_k$, $k = 1, 2, \ldots, K$. The above concentrated log-likelihood is obtained by replacing $\boldsymbol{x}, \Psi, \sigma^2$, and $Z_k^H Z_k$ in (A.3) (see Appendix A) with $\hat{\boldsymbol{x}}_{\mathrm{UML}}, \hat{\Psi}_{\mathrm{UML}}, \hat{\sigma}^2_{\mathrm{ML}}$, and C (respectively), neglecting constant terms, using ([27, Th. 18.1.1, p. 416]), and applying a monotonic transformation. The classical algorithm for DTFT-based frequency estimation in [39, ch. 6.4.4] can be easily extended and applied to maximizing (3.17).

AML Estimation for Structured LOS Array Response Model: Due to the CRB decoupling between the mean and variance-component parameters [see (E.1a) in Appendix E], the ML estimate of Ψ for the unstructured LOS array response model in (3.16b) is asymptotically efficient under the structured

LOS array response model. Hence, $\hat{\Psi}_{\mathrm{UML}}$ is an AML estimate of Ψ, provided that it is positive semidefinite. Now, we obtain a closed-form AML estimate of the structured-array complex LOS amplitude by substituting (2.4c) and $\hat{\sigma}^2_{\mathrm{ML}}$ and $\hat{\Psi}_{\mathrm{UML}}$ into (3.3)

$$\hat{x}_{\mathrm{AML}} = \frac{\left(\boldsymbol{a}^H_{\mathrm{T,LOS}} \otimes \boldsymbol{a}^H_{\mathrm{R,LOS}}\right) C \Xi^{-1} \boldsymbol{z}_{\mathrm{DTFT}}}{\left(\boldsymbol{a}^H_{\mathrm{T,LOS}} \otimes \boldsymbol{a}^H_{\mathrm{R,LOS}}\right) C \Xi^{-1} C (\boldsymbol{a}_{\mathrm{T,LOS}} \otimes \boldsymbol{a}_{\mathrm{R,LOS}})} \tag{3.19a}$$

where

$$\Xi = \left[\frac{1}{K}\sum_{k=1}^{K} \boldsymbol{z}_k \boldsymbol{z}_k{}^H\right] - \boldsymbol{z}_{\mathrm{DTFT}} \boldsymbol{z}_{\mathrm{DTFT}}{}^H. \tag{3.19b}$$

To simplify the notation, we have omitted the dependence of $\boldsymbol{z}_{\mathrm{DTFT}}(\omega_{\mathrm{D,LOS}} N)$ on $\omega_{\mathrm{D,LOS}} N$ in (3.19). The above estimator is asymptotically efficient. However, it is based on the assumption that $\hat{\Psi}_{\mathrm{UML}} \geq 0$ and performs poorly when this assumption does not hold (e.g., when K is small, see Figs. 2, 6, and 8 in Section IV).

ML Estimation for Rayleigh Fading: Under the Rayleigh-fading scenario, the closed-form expressions for the ML estimates of σ^2 and Ψ are given by (3.13) and

$$\hat{\Psi}_{\mathrm{ML}} = \left[\frac{1}{K}\sum_{k=1}^{K} \mathrm{vec}\left(Y_k \Phi_k^H \Gamma_\Phi{}^{-1}\right) \cdot \mathrm{vec}\left(Y_k \Phi_k^H \Gamma_\Phi{}^{-1}\right)^H\right] - \hat{\sigma}^2_{\mathrm{ML}} \cdot \left(\Gamma_\Phi{}^T\right)^{-1} \otimes I_{n_{\mathrm{R}}}. \tag{3.20}$$

The closed-form ML estimates in (3.13), (3.16), and (3.20) can be used to implement noncoherent ML space-time receivers, which require fast estimation of the fading parameters.

The estimates of Ψ obtained using (3.14) and (3.15) and closed-form expressions (3.16b) and (3.20) are maximum likelihood only if they are positive semidefinite; otherwise, we can apply the ECME algorithm in Section III-A1, which always converges to solutions within the parameter space. Clearly, a necessary condition for (3.15), (3.16b), and (3.20) to be positive semidefinite is $K \geq n_{\mathrm{R}} n_{\mathrm{T}}$. The probability that (3.14)–(3.15), (3.16b), and (3.20) yield nonpositive semidefinite estimates of Ψ is asymptotically zero as either $K \to \infty$ or $N n_{\mathrm{R}} \to \infty$. For the unstructured LOS array response model, we can remove (3.7b) from the ML and REML ECME iterations and use the closed-form expression for the ML estimate of $\boldsymbol{x}$ in (3.16a).

B. Independent Fading

We develop ECME ML and REML algorithms for estimating $\boldsymbol{\rho}$ under the independent fading scenario and simplify them in the case where $\Phi_k \Phi_k^H$ is an identity matrix. Approximately independent fading occurs, for example, in virtual channel representations; see [22] and references therein. (In [14], we derived Henderson's methods [40] for this scenario, which performed similarly to the algorithms proposed here.)

1) ECME Algorithm for ML Estimation: The ECME ML algorithm for independent fading follows using arguments similar to those in Appendix C-A (where ECME algorithms were derived for the correlated fading scenario): Iterate between (3.7a)–(3.7c) for $k = 1, 2, \ldots, K$ and

$$\begin{aligned}(\sigma^2)^{(i+1)} &= \tilde{\sigma}^2 + \frac{(\sigma^2)^{(i)}}{K n_{\mathrm{R}} N} \cdot \sum_{k=1}^{K} \left(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x}^{(i)}\right)^H \\ &\quad \times W_k^{(i)} \left(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k} \boldsymbol{x}^{(i)}\right) \end{aligned} \tag{3.21a}$$

$$\begin{aligned}(\psi_n)^{(i+1)} &= (\psi_n)^{(i)} + \frac{1}{K} \cdot \sum_{k=1}^{K} \left\{ \left| \left[\boldsymbol{h}^{(i)}_{\mathrm{SC},k}\right]_n \right|^2 - \left[(\psi_n)^{(i)}\right]^2 \right. \\ &\quad \left. \cdot \left[Z_k^H Z_k W_k^{(i)} Z_k^H Z_k\right]_{n,n} \right\} \end{aligned} \tag{3.21b}$$

for $n = 1, 2, \ldots, n_{\mathrm{R}} n_{\mathrm{T}}$, where $\tilde{\sigma}^2$ has been defined in (3.8c), and

$$\Psi^{(i)} = \mathrm{diag}\left\{(\psi_1)^{(i)}, (\psi_2)^{(i)}, \ldots, (\psi_{n_{\mathrm{R}} n_{\mathrm{T}}})^{(i)}\right\}.$$

Here, $|\cdot|$ denotes the absolute value, $[Z_k^H Z_k W_k^{(i)} Z_k^H Z_k]_{n,n}$ the (n,n) element of $Z_k^H Z_k W_k^{(i)} Z_k^H Z_k$, and $[\boldsymbol{h}^{(i)}_{\mathrm{SC},k}]_n$ the nth element of $\boldsymbol{h}^{(i)}_{\mathrm{SC},k}$.

$\boldsymbol{\Phi_k \Phi_k^H = I_{n_{\mathrm{T}}}}$: When $\Phi_k \Phi_k^H = I_{n_{\mathrm{T}}}$, (3.7a) and (3.7c) simplify to

$$W^{(i)} = W_k^{(i)} = \mathrm{diag}\left\{\left[(\sigma^2)^{(i)} + (\psi_1)^{(i)}\right]^{-1}, \ldots \left[(\sigma^2)^{(i)} + (\psi_{n_{\mathrm{R}} n_{\mathrm{T}}})^{(i)}\right]^{-1}\right\} \tag{3.22a}$$

$$\left[\boldsymbol{h}^{(i)}_{\mathrm{SC},k}\right]_n = \frac{(\psi_n)^{(i)}}{(\sigma^2)^{(i)} + (\psi_n)^{(i)}} \cdot \left[\boldsymbol{z}_k - A_{\mathrm{LOS},k} \boldsymbol{x}^{(i)}\right]_n \tag{3.22b}$$

for $n = 1, 2, \ldots, n_{\mathrm{R}} n_{\mathrm{T}}$, and the conditional maximization (CM) steps in (3.21a) and (3.21b) simplify accordingly. In this case, the CRB expressions also simplify (see Appendix E-B):

$$\mathrm{CRB}_{\sigma^2,\sigma^2} = \frac{\sigma^4}{n_{\mathrm{R}}(N - n_{\mathrm{T}})K} \tag{3.23a}$$

$$\begin{aligned}[\mathrm{CRB}_{\psi,\psi}]_{n,n} &= \mathrm{CRB}_{\psi_n \psi_n} \\ &= \frac{(\sigma^2 + \psi_n)^2}{K} + \frac{\sigma^4}{n_{\mathrm{R}}(N - n_{\mathrm{T}})K} \end{aligned} \tag{3.23b}$$

for $n = 1, 2, \ldots, n_{\mathrm{R}} n_{\mathrm{T}}$. Here, $\mathrm{CRB}_{\psi_n,\psi_n}$ is an increasing function of *both* σ^2 and ψ_n. As expected, both $\mathrm{CRB}_{\sigma^2,\sigma^2}$ and $\mathrm{CRB}_{\psi_n,\psi_n}$ decrease proportionally to $1/K$ as the number of coherent intervals K grows.

TABLE I
PERCENTAGES OF TRIALS IN WHICH THE ALTERNATING-PROJECTION AND CLOSED-FORM ESTIMATORS FOR THE STRUCTURED AND UNSTRUCTURED LOS ARRAY RESPONSE MODELS IN (3.15) AND (3.16b) WERE NOT POSITIVE SEMIDEFINITE, AS FUNCTIONS OF K

	$K=10$	$K=20$	$K=30$	$K=40$	$K=50$	$K=60$
% invalid estimates of Ψ, structured LOS	63.2%	13.7%	3.7%	0.5%	0.1%	0.05%
% invalid estimates of Ψ, unstructured LOS	65.3%	18.3%	3.9%	0.8%	0.1%	0.04%

2) ECME Algorithm for REML Estimation: The ECME REML algorithm for independent fading follows using arguments similar to those in Appendix C-B: Iterate between (3.7a)–(3.7c) for $k=1,2,\ldots,K$, and

$$(\sigma^2)^{(i+1)} = \frac{Kn_{\mathrm{R}}N}{Kn_{\mathrm{R}}N-r}\cdot\tilde{\sigma}^2 + \frac{(\sigma^2)^{(i)}}{Kn_{\mathrm{R}}N-r} \cdot\sum_{k=1}^{K}\left(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\boldsymbol{x}^{(i)}\right)^H \times W_k^{(i)}\left(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\boldsymbol{x}^{(i)}\right) \tag{3.24a}$$

$$(\psi_n)^{(i+1)} = (\psi_n)^{(i)} + \frac{1}{K}\cdot\sum_{k=1}^{K}\left\{\left|\left[\boldsymbol{h}_{\mathrm{SC},k}^{(i)}\right]_n\right|^2 - \left[(\psi_n)^{(i)}\right]^2 \cdot\left[Z_k^H Z_k\left(W_k^{(i)} - W_k^{(i)}\mathcal{A}_{\mathrm{LOS},k} \cdot\left(\sum_{l=1}^{K}\mathcal{A}_{\mathrm{LOS},l}^H W_l^{(i)}\mathcal{A}_{\mathrm{LOS},l}\right)^{-1} \times\mathcal{A}_{\mathrm{LOS},k}^H W_k^{(i)}\right)Z_k^H Z_k\right]_{n,n}\right\} \tag{3.24b}$$

where $n=1,2,\ldots,n_{\mathrm{R}}n_{\mathrm{T}}$. For $\Phi_k\Phi_k^H = I_{n_{\mathrm{T}}}$, we can use (3.22) to simplify (3.24a) and (3.24b).

As in the correlated fading case, the above ECME algorithms converge to variance estimates that are *always* in the parameter space (i.e., non-negative). They can be initialized using the moment estimators in (3.10b) and diagonal elements of (3.10a). For the unstructured LOS array response model, we can remove (3.7b) from the ML and REML ECME iterations and use the closed-form expression for the ML estimate of $\boldsymbol{x}$ in (3.16a). Rayleigh-fading versions of these algorithms follow by removing (3.7b) and setting $\mathcal{A}_{\mathrm{LOS},k}=0$ in (3.7c) and (3.21a) or (3.24a).

IV. NUMERICAL EXAMPLES

We evaluate the estimation accuracy and computational efficiency of the ML and REML methods in Section III. Our performance metric is the mean-square error (MSE) of an estimator, calculated using 5000 independent trials. Numerical simulations were performed using both block- and continuous-fading scenarios. Throughout this section, we employed the Alamouti transmission scheme for a $n_{\mathrm{R}}\times n_{\mathrm{T}}=2\times 2$ MIMO system with $N=30$ quadrature phase-shift keying (QPSK) symbols per coherent interval (normalized so that $\Phi_k\Phi_k^H = I_2$) and generated additive white complex Gaussian noise $\boldsymbol{e}_k(t)$ with variance $\sigma^2=0.01$.

A. Block-Fading Scenario

In the block-fading case, we generated the simulated data using the measurement model in Section II. The LOS component was generated using (2.4c) with $x=1, \boldsymbol{a}_{\mathrm{T,LOS}} = [1,\exp(-j\pi/6)]^T, \boldsymbol{a}_{\mathrm{R,LOS}} = [1,\exp(-j\pi/3)]^T$, and $\omega_{\mathrm{D,LOS}} = \pi/2\cdot 10^{-2}$.

1) Correlated Fading: In the first set of simulations, we consider the *correlated block-fading scenario* and apply the ML and REML algorithms in Section III-A using the unstructured and structured LOS array response models in (2.4b) and (2.4c). The spatial fading covariance matrix was

$$\Psi = \sigma^2\cdot\begin{bmatrix} 1 & 0.2+0.1j & 0.4-0.5j & 0 \\ 0.2-0.1j & 2 & 0 & 0.1+0.1j \\ 0.4+0.5j & 0 & 4 & 0.3-0.3j \\ 0 & 0.1-0.1j & 0.3+0.3j & 8 \end{bmatrix}. \tag{4.1}$$

In Figs. 1 and 2, we present the MSEs (and corresponding CRBs) for the ML and REML estimates of selected parameters as functions of the number of coherent intervals K. The ML estimates of $\boldsymbol{\rho}$ were computed using the closed-form expressions in (3.13) and (3.16) for the unstructured LOS model and the alternating-projection ML algorithm (3.14) and (3.15) for the structured LOS model. For $K\geq 10$, the alternating-projection algorithm converged in less than five iterations. In the cases where (3.16b) and (3.15) were not positive semidefinite, we ran the ECME ML algorithm described in Section III-A1. In terms of CPU time, the alternating-projection ML algorithm was five to seven times faster than the ECME ML algorithm. The REML estimation was performed using the ECME algorithm in Section III-A2, which converged in less than seven iterations.

In Table I, we show the percentages of trials in which the estimates of Ψ in (3.15) and (3.16b) were not positive semidefinite as functions of K. These percentages decay rapidly with K; however, they are high for small K, underlining the importance of the ECME approach, which handles parameter constraints automatically.

In Fig. 1, the MSEs and CRBs for the ML and REML estimates of $\mathrm{Re}\{\Psi_{3,2}\}, \Psi_{3,3}, \Psi_{4,4}$, and sum of all elements of $\boldsymbol{\psi}$ are shown as functions of K for the (left) unstructured and (right) structured LOS array response models. Due to the CRB decoupling between the mean and variance-component parameters [see (E.1a) in Appendix E], the CRBs for the variance components are the same regardless of the LOS array response parametrization; the corresponding MSEs are also approximately equal.

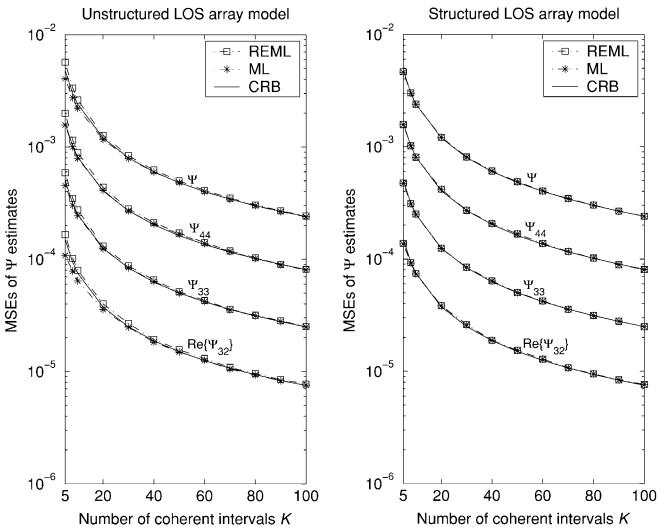

Fig. 1. Mean-square errors and Cramér–Rao bounds for the ML and REML estimates of $\mathrm{Re}\{\Psi_{3,2}\}, \Psi_{3,3}, \Psi_{4,4}$ and sum of all elements of $\boldsymbol{\psi}$ under the correlated block-fading scenario and (left) unstructured and (right) structured LOS array response models, as functions of K.

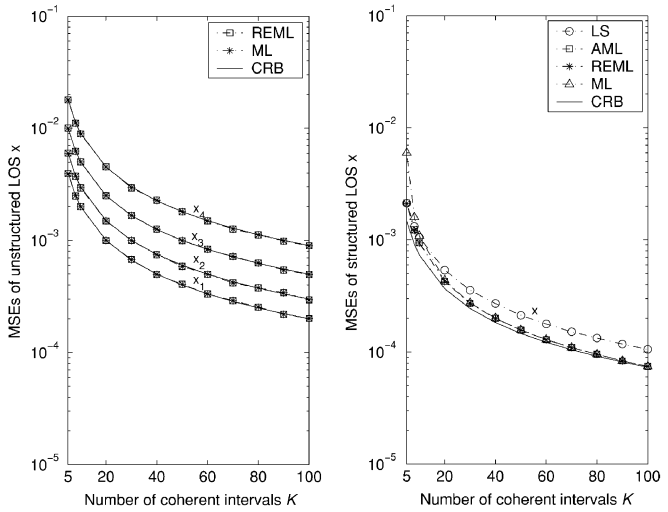

Fig. 2. MSEs and CRBs for the ML and REML estimates of the LOS coefficients under the correlated block-fading scenario and (left) unstructured and (right) structured LOS array response models, as functions of K.

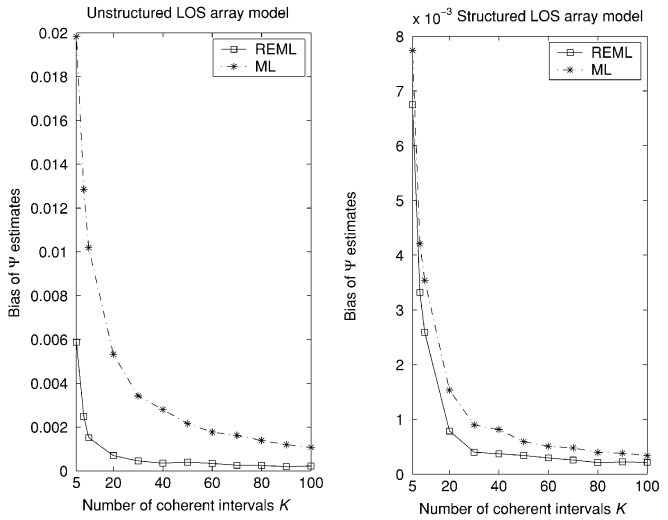

Fig. 3. Absolute biases for the ML and REML estimates of the sum of all elements of $\boldsymbol{\psi}$ under the correlated block-fading scenario and (left) unstructured and (right) structured LOS array response models, as functions of K.

In Fig. 2, we present the MSEs and CRBs for the (left) ML and REML estimates of the unstructured LOS array response vector and (right) ML, REML, AML, and LS estimates of the structured-array LOS complex amplitude x, as functions of K. Here, the linear LS estimate of x is computed by substituting (2.4c) into (3.9). As expected, the MSEs and CRBs are smaller for the structured LOS model. For larger K, the (closed-form) AML and (iterative) ML and REML estimates of x achieve similar MSE performances. However, the AML estimator performs poorly when K is small; see also the discussion in Section III-A3. An analytical expression for the MSE of the linear LS estimate of x is given below (for the special case of $\Phi_k \Phi_k^H = I_{n_{\mathrm{T}}}$ and assuming the block-fading scenario):

$$\begin{aligned}\mathrm{MSE}_x &= \mathrm{MSE}_{\mathrm{Re}\{x\}} + \mathrm{MSE}_{\mathrm{Im}\{x\}} \\ &= \frac{\sigma^2}{K \cdot \boldsymbol{a}_{\mathrm{T}}^H \boldsymbol{a}_{\mathrm{T}} \cdot \boldsymbol{a}_{\mathrm{R}}^H \boldsymbol{a}_{\mathrm{R}}} \\ &\quad + \frac{\left(\boldsymbol{a}_{\mathrm{T,LOS}}^H \otimes \boldsymbol{a}_{\mathrm{R,LOS}}^H\right) \Psi \left(\boldsymbol{a}_{\mathrm{T,LOS}} \otimes \boldsymbol{a}_{\mathrm{R,LOS}}\right)}{K \cdot \left(\boldsymbol{a}_{\mathrm{T}}^H \boldsymbol{a}_{\mathrm{T}}\right)^2 \cdot \left(\boldsymbol{a}_{\mathrm{R}}^H \boldsymbol{a}_{\mathrm{R}}\right)^2}. \end{aligned} \quad (4.2)$$

The CRBs were computed using the results in Appendix E-A.

Since the MSE performances of the ML and REML estimators are similar, it is of interest to compare their biases as well. Fig. 3 compares the absolute biases for the ML and REML estimates of the variance components $\boldsymbol{\psi}$ (computed using 5000 independent trials) under the (left) unstructured and (right) structured LOS array response models; the biases are shown as functions of K. The obtained results confirm the *bias-correction property* of REML variance-component estimation; see also the discussion in Section III. Compared with ML, the REML approach yields significant bias improvements when the rank r of the deterministic component is large (e.g., unstructured LOS array model) and for small sample sizes K.

Unknown $\boldsymbol{\omega}_{\mathrm{D,LOS}}$: If the LOS Doppler shift $\omega_{\mathrm{D,LOS}}$ is *unknown*, we estimate it using the ML estimator in (3.17) for the unstructured LOS model. Fig. 4 shows the MSE performance of this estimator as a function of K.

2) Independent Fading: Consider the *independent block-fading scenario* with the following fading covariance matrix:

$$\Psi = \mathrm{diag}\{\psi_1, \psi_2, \psi_3, \psi_4\} = \sigma^2 \cdot \mathrm{diag}\{1, 2, 4, 8\}. \quad (4.3)$$

We have applied the ECME, ML, and REML algorithms in Section III-B; their MSE performances are shown in Figs. 5 and 6. For $K \geq 20$, the ECME algorithms converged in less than ten iterations. Fig. 5 shows the MSEs and CRBs for the ML and REML estimates of the fading variances ψ_1, ψ_2, ψ_3, and ψ_4 as functions of K for the (left) unstructured and (right) structured LOS array response models. As expected, the CRBs for these variances are the same regardless of the LOS array response parametrization, and the corresponding MSEs are approximately equal as well. In Fig. 6, we compare the MSEs and

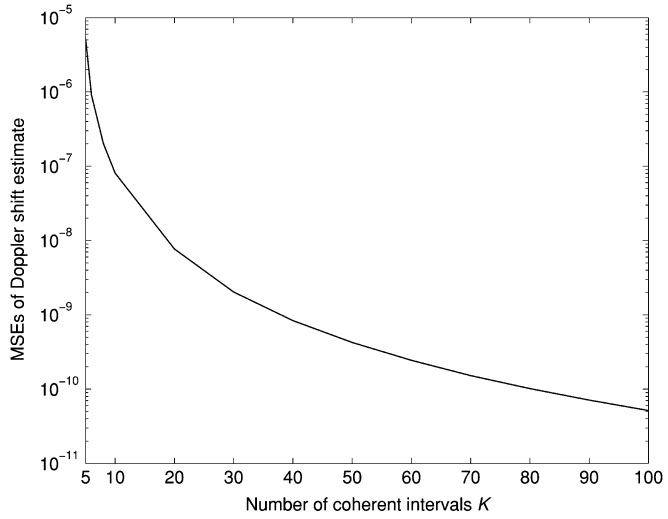

Fig. 4. MSE for the ML estimate of the LOS Doppler shift in (3.17), as a function of K.

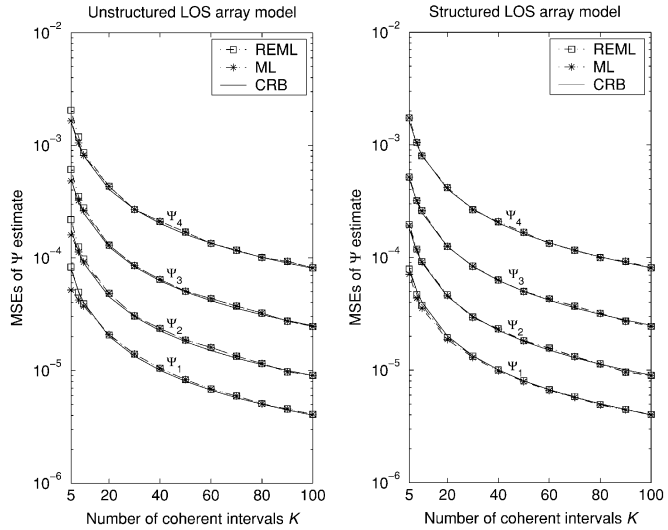

Fig. 5. MSEs and CRBs for the ML and REML estimates of ψ_4, ψ_3, ψ_2, and ψ_1 under the independent block-fading scenario and (left) unstructured and (right) structured LOS array response models, as functions of K.

CRBs for the ML and REML estimates of the (left) unstructured LOS array response vector and ML, REML, AML, and LS estimates of the (right) structured-array LOS complex amplitude x. To compute the CRBs, we used (3.23) and the results in Appendix E-B.

B. Continuous-Fading Scenario

We now study the performance of the proposed methods in *continuous fading*, where the scattering channel coefficients are temporally correlated according to the Jakes' model; see, e.g., [12], [41], and references therein. First, denote by $\boldsymbol{h}_{\mathrm{SC},k}(t)$ the scattering channel vector at time t in the kth coherent interval. Assuming adjacent coherent intervals, we model the $Kn_{\mathrm{R}}n_{\mathrm{T}}N \times 1$ vector of all scattering coefficients: $\boldsymbol{h}_{\mathrm{SC}} = [\boldsymbol{h}_{\mathrm{SC},1}(1)^T \ldots \boldsymbol{h}_{\mathrm{SC},1}(N)^T,\ \boldsymbol{h}_{\mathrm{SC},2}(1)^T \ldots$

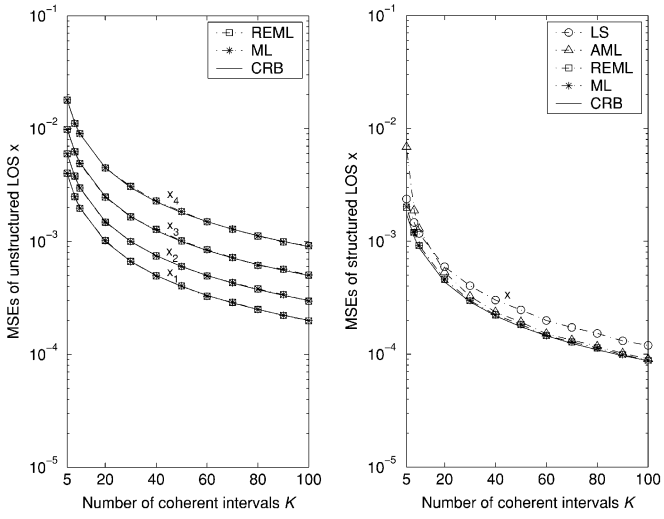

Fig. 6. MSEs and CRBs for the ML, REML, AML, and LS estimates of the LOS coefficients under the independent block-fading scenario and (left) unstructured and (right) structured LOS array response models, as functions of K.

$\boldsymbol{h}_{\mathrm{SC},2}(N)^T \ldots \boldsymbol{h}_{\mathrm{SC},K}(1)^T, \ldots \boldsymbol{h}_{\mathrm{SC},K}(N)^T]^T$ as a zero-mean complex Gaussian vector with covariance matrix (see [41]):

$$\mathrm{E}\Big[\boldsymbol{h}_{\mathrm{SC}}\boldsymbol{h}_{\mathrm{SC}}{}^H\Big] = J(\omega_{\mathrm{D}}) \otimes \Psi \tag{4.4a}$$

where $\omega_{\mathrm{D}} \in (0, \pi)$ is the *maximum angular Doppler frequency* (corresponding to the Doppler spread of $2\omega_{\mathrm{D}}$), the (p, q) element of the $KN \times KN$ matrix $J(\omega_{\mathrm{D}})$ is

$$[J(\omega_{\mathrm{D}})]_{p,q} = J_0(\omega_{\mathrm{D}}(p - q)) \tag{4.4b}$$

and $J_0(\cdot)$ denotes the zeroth-order Bessel function of the first kind [12], [41]. We further assume that the LOS component of the channel response matrix changes with t according to the following model:

$$H_{\mathrm{LOS},k}(t) = \boldsymbol{a}_{\mathrm{R,LOS}}\boldsymbol{a}_{\mathrm{T,LOS}}^T \times \exp\{j\omega_{\mathrm{D,LOS}} \cdot [(k - 0.5)N + t]\} \cdot x \tag{4.4c}$$

where the LOS angular Doppler shift $\omega_{\mathrm{D,LOS}}$ should be bounded by the maximum Doppler frequency, i.e., $|\omega_{\mathrm{D,LOS}}| \leq \omega_{\mathrm{D}}$. Combining the scattering and LOS channel components, we obtain the following *continuous-fading measurement model*:

$$\boldsymbol{y}_k(t) = [H_{\mathrm{SC},k}(t) + H_{\mathrm{LOS},k}(t)]\boldsymbol{\phi}_k(t) + \boldsymbol{e}_k(t) \tag{4.4d}$$

for $t = 1, \ldots, N, k = 1, \ldots, K$, where $\boldsymbol{\phi}_k(t)$ and $\boldsymbol{e}_k(t)$ have been defined in Section II, and $\boldsymbol{h}_{\mathrm{SC},k}(t) = \mathrm{vec}\{H_{\mathrm{SC},k}(t)\}$. The above model accounts for correlations among the coherent intervals and time variations of the scattering and LOS channel coefficients within a coherent interval.

In the following examples, we consider the correlated fading scenario with Ψ given in (4.1), maximum Doppler frequency $\omega_{\mathrm{D}} = 2\pi \cdot 10^{-2}$ (consistent with the mobile speed of 100 mi/h for the carrier frequency 1.9 GHz and a symbol rate

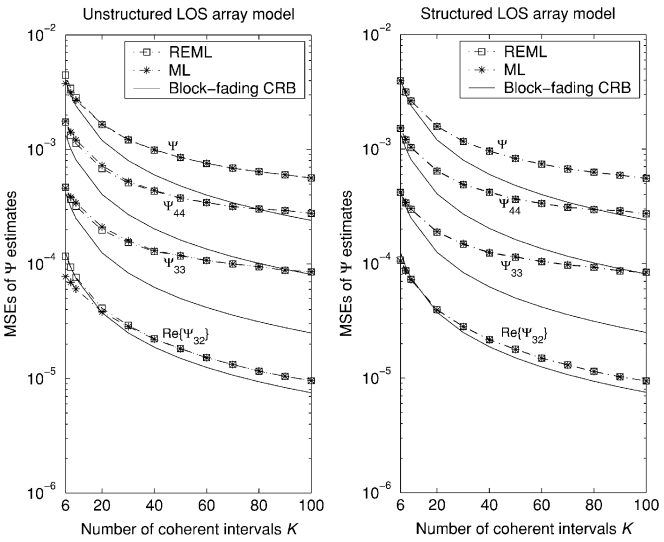

Fig. 7. MSEs and block-fading CRBs for the ML and REML estimates of $\mathrm{Re}\{\Psi_{3,2}\}, \Psi_{3,3}, \Psi_{4,4}$, and sum of all elements of $\boldsymbol{\psi}$ under the correlated Ricean continuous-fading scenario and (left) unstructured and (right) structured LOS array response models, as functions of K.

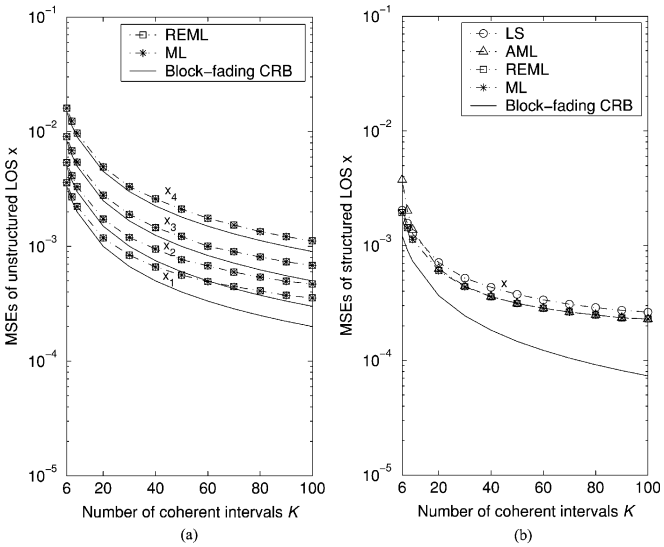

Fig. 8. MSEs and block-fading CRBs for the ML, REML, AML, and LS estimates of the LOS coefficients under the correlated continuous-fading scenario and (left) unstructured and (right) structured LOS array response models, as functions of K.

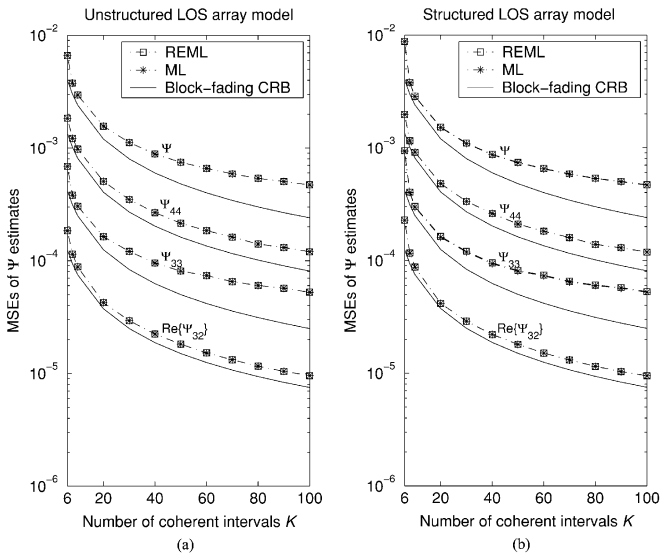

Fig. 9. MSEs and block-fading CRBs for the ML and REML estimates of $\mathrm{Re}\{\Psi_{3,2}\}, \Psi_{3,3}, \Psi_{4,4}$, and sum of all elements of $\boldsymbol{\psi}$ under the correlated Ricean continuous-fading scenario with unknown $\omega_{\mathrm{D,LOS}}$ and (left) unstructured and (right) structured LOS array response models, as functions of K.

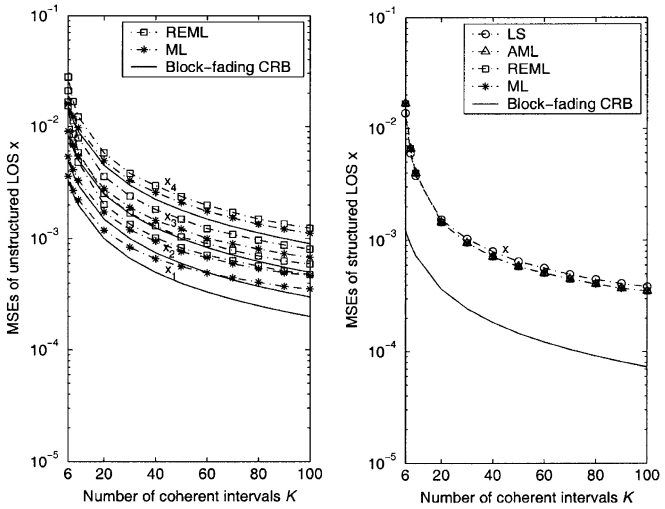

Fig. 10. MSEs and block-fading CRBs for the ML, REML, AML, and LS estimates of the LOS coefficients under the correlated continuous-fading scenario with unknown $\omega_{\mathrm{D,LOS}}$ and (left) unstructured and (right) structured LOS array response models, as functions of K.

of 30 kHz; see also [36]), and the LOS parameters $x = 1, \boldsymbol{a}_{\mathrm{T,LOS}} = [1, \exp(-j\pi/6)]^T, \boldsymbol{a}_{\mathrm{R,LOS}} = [1, \exp(-j\pi/3)]^T$, and $\omega_{\mathrm{D,LOS}} = \pi/2 \cdot 10^{-2}$. We first assume that the LOS Doppler shift $\omega_{\mathrm{D,LOS}}$ is known, and then consider the case where $\omega_{\mathrm{D,LOS}}$ is unknown.

Known $\boldsymbol{\omega}_{\mathrm{D,LOS}}$: We computed the ML and REML estimates of $\boldsymbol{\rho}$ using the methods in Section III-A, where the coherent-interval length was chosen as $N = 30$. Figs. 7 and 8 show the MSEs for the ML and REML estimates of $\mathrm{Re}\{\Psi_{3,2}\}, \Psi_{3,3}, \Psi_{4,4}$, sum of all elements of $\boldsymbol{\psi}$, and the LOS coefficients, as functions of K; Fig. 8 (right) shows the MSEs for the AML and LS estimates of the structured-array LOS complex amplitude x. We also compare these MSEs with the corresponding block-fading CRBs, thus quantifying the *performance loss* that each method incurs due to continuous fading. Interestingly, for small K, the MSEs of the ML and REML estimates are close to the block-fading CRBs. Under the continuous-fading scenario, the proposed variance-component estimates are mostly affected by correlations among the coherent intervals and time variations of the scattering component within a coherent interval, whereas the LOS-coefficient estimates are mostly affected by time variations of the LOS component.

Unknown $\boldsymbol{\omega}_{\mathrm{D,LOS}}$: We now consider the scenario where the LOS Doppler shift $\omega_{\mathrm{D,LOS}}$ is *unknown* and *estimated* by maximizing (3.17). Here, we selected the coherent interval length $N = 30$, as in the previous example. Following the *estimated likelihood approach* in, e.g., [26, ch. 10.7], we treat the obtained

estimate of $\omega_{\mathrm{D,LOS}}$ as a known constant and apply the ML and REML methods in Section III-A. Figs. 9 and 10 show the MSEs and block-fading CRBs for the (estimated) ML, REML, AML, and LS estimates of $\mathrm{Re}\{\Psi_{3,2}\}$, $\Psi_{3,3}$, $\Psi_{4,4}$, sum of all elements of $\boldsymbol{\psi}$, and LOS coefficients, as functions of K. For the unstructured LOS array response model, the ML estimates of $\boldsymbol{x}$ outperform the corresponding REML estimates. Interestingly, estimating the LOS Doppler shift does not significantly affect the performances of the ML and REML estimates of the variance components (compare Figs. 7 and 9).

V. Concluding Remarks

We derived ML and REML methods for estimating the mean and covariance parameters of MIMO fading channels under correlated and independent block-fading scenarios. CRBs were computed for the unknown fading-covariance and line-of-sight parameters. For unitary space-time codes and orthogonal designs in correlated fading, we obtained closed-form exact and approximate ML estimates of the unknown parameters. We evaluated the performance of the proposed methods via numerical simulations under the block- and continuous-fading scenarios. Simulation results show that the proposed estimators are almost efficient under the block-fading scenario, having mean-square errors close to the corresponding CRBs.

Further research will include incorporating the proposed estimators into the design of space-time receivers and deriving efficient estimation algorithms and CRBs for the continuous-fading scenario. It is also of interest to develop estimators for the case where the LOS channel response vector follows the multivariate complex Ricean model in [11].

Appendix A
Expressions for $\Sigma_k(\boldsymbol{\gamma})^{-1}$ and the Log-Likelihood Function

We derive expressions for $\Sigma_k(\boldsymbol{\gamma})^{-1}$ and simplify the log-likelihood expression in (3.2). Using the matrix inversion lemma (in e.g., [27, Th. 18.2.8, p. 424]), we get

$$\Sigma_k(\boldsymbol{\gamma})^{-1} = \sigma^{-2}\left[I_{n_{\mathrm{R}}N} - Z_k\left(Z_k^H Z_k\right)^{-1} Z_k^H\right] + Z_k W_k(\boldsymbol{\gamma}) Z_k^H \quad \text{(A.1)}$$

which further implies

$$\Upsilon_k^H \Sigma_k(\boldsymbol{\gamma})^{-1} = \mathcal{A}_{\mathrm{LOS},k}^H W_k(\boldsymbol{\gamma}) Z_k^H. \quad \text{(A.2)}$$

Now, the log-likelihood (3.2) can be rewritten as

$$\begin{aligned} L(\boldsymbol{\rho}) = &-Kn_{\mathrm{R}}N\ln(\pi\sigma^2) \\ &- \frac{1}{\sigma^2}\sum_{k=1}^{K} \boldsymbol{y}_k^H\left[I_{n_{\mathrm{R}}N} - Z_k\left(Z_k^H Z_k\right)^{-1} Z_k^H\right]\boldsymbol{y}_k \\ &- \sum_{k=1}^{K}(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\boldsymbol{x})^H W_k(\boldsymbol{\gamma})(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\boldsymbol{x}) \\ &- \sum_{k=1}^{K}\ln\left|I_{n_{\mathrm{R}}n_{\mathrm{T}}} + (1/\sigma^2)\cdot Z_k^H Z_k \Psi\right| \end{aligned} \quad \text{(A.3)}$$

which follows by using (A.1) and the fact that

$$|\Sigma_k(\boldsymbol{\gamma})| = (\sigma^2)^{n_{\mathrm{R}}N}\cdot\left|I_{n_{\mathrm{R}}n_{\mathrm{T}}} + (1/\sigma^2)\cdot Z_k^H Z_k \Psi\right| \quad \text{(A.4)}$$

see (3.1b) and [27, Th. 18.1.1, p. 416].

Appendix B
Restricted Maximum Likelihood

We derive a REML log-likelihood expression for the measurement model in Section II. Define a vector of error contrasts $\boldsymbol{u} = B^H\boldsymbol{y}$, where B is a matrix whose columns span the space orthogonal to the column space of Υ. Then, the REML log-likelihood is obtained as the log-likelihood function for the error contrasts. Without loss of generality, we choose $B^H B = I_{Kn_{\mathrm{R}}N-1}$; see [32]. Then, (3.6) follows by using the identities (see [32], and [42, p. 77])

$$\begin{aligned} &|B^H\Sigma(\boldsymbol{\gamma})B| = |\Upsilon^H\Sigma(\boldsymbol{\gamma})^{-1}\Upsilon|\cdot|\Sigma(\boldsymbol{\gamma})|/|\Upsilon^H\Upsilon| \\ &B[B^H\Sigma(\boldsymbol{\gamma})B]^{-1}B^H = \Pi(\boldsymbol{\gamma}) \end{aligned}$$

and neglecting terms that do not depend on $\boldsymbol{\gamma}$. Note that (3.6) can be further simplified by using (A.3):

$$\begin{aligned} L_{\mathrm{REML}}(\boldsymbol{\gamma}) = &-Kn_{\mathrm{R}}N\ln\sigma^2 \\ &- \sum_{k=1}^{K}\ln\left|I_{n_{\mathrm{R}}n_{\mathrm{T}}} + Z_k^H Z_k\Psi/\sigma^2\right| \\ &- \frac{1}{\sigma^2}\sum_{k=1}^{K}\boldsymbol{y}_k^H\left[I_{n_{\mathrm{R}}N} - Z_k\left(Z_k^H Z_k\right)^{-1} Z_k^H\right]\boldsymbol{y}_k \\ &- \ln\left|\sum_{k=1}^{K}\mathcal{A}_{\mathrm{LOS},k}^H W_k(\boldsymbol{\gamma})\mathcal{A}_{\mathrm{LOS},k}\right| \\ &- \sum_{k=1}^{K}[\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\hat{\boldsymbol{x}}(\boldsymbol{\gamma})]^H W_k(\boldsymbol{\gamma}) \\ &\times [\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\hat{\boldsymbol{x}}(\boldsymbol{\gamma})] \end{aligned} \quad \text{(B.1)}$$

where $\hat{\boldsymbol{x}}(\boldsymbol{\gamma})$ has been defined in (3.3).

Appendix C
ECME Algorithms for Correlated Fading

A. ML Estimation

We derive the ECME ML algorithm in Section III-A1. The first conditional maximization (CM) step in (3.7b) follows by substituting the most recent estimate of $\boldsymbol{\gamma}$ into (3.3), where (3.3) is the ML estimate of $\boldsymbol{x}$ that maximizes the observed-data likelihood function for fixed $\boldsymbol{\gamma}$. The second CM step in (3.8a) follows by reparametrizing the variance components using $\Psi' = \Psi/\sigma^2$ and σ^2 (rather than Ψ and σ^2), which allows us to find the closed-form solution for the ML estimate of σ^2 that maximizes the observed-data likelihood function for fixed Ψ' and $\boldsymbol{x}$:

$$\begin{aligned} &\hat{\sigma}^2(\boldsymbol{x}, \Psi') \\ &\quad = \frac{1}{Kn_{\mathrm{R}}N}\sum_{k=1}^{K}\Big\{\boldsymbol{y}_k^H\left[I_{n_{\mathrm{R}}N} - Z_k\left(Z_k^H Z_k\right)^{-1} Z_k^H\right]\boldsymbol{y}_k \\ &\qquad + (\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\boldsymbol{x})^H W_k'(\Psi')(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\boldsymbol{x})\Big\} \end{aligned} \quad \text{(C.1)}$$

where

$$W_k'(\Psi') = \left(Z_k^H Z_k + Z_k^H Z_k \Psi' Z_k^H Z_k\right)^{-1}. \tag{C.2}$$

Equation (C.1) follows by maximizing the (reparametrized) log-likelihood function in (A.3) with $W_k(\boldsymbol{\gamma})$ replaced by $(1/\sigma^2)\cdot W_k'(\Psi')$. Then, the second CM step for updating the estimate of σ^2 in (3.8a) is obtained by substituting the most recent estimates of $\boldsymbol{x}$ and $\Psi' = \Psi/\sigma^2$ into (C.1). Finally, we apply the standard EM algorithm to update Ψ by treating $\boldsymbol{h}_{\mathrm{SC},k}$ as the missing data, where $\boldsymbol{x}$ and σ^2 are *fixed.* Consequently

$$\boldsymbol{r}_k = \boldsymbol{y}_k - \Upsilon_k \boldsymbol{x}, \quad k = 1, 2, \ldots, K \tag{C.3}$$

are the *observed data*. If $\boldsymbol{h}_{\mathrm{SC},k}, k = 1, 2, \ldots, K$ were known (forming the *complete data* together with $\boldsymbol{r}_k$, $k = 1, 2, \ldots, K$), we could easily find the complete-data ML estimate of Ψ as follows:

$$\hat{\Psi} = \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{h}_{\mathrm{SC},k}\boldsymbol{h}_{\mathrm{SC},k}^H \tag{C.4}$$

where $\hat{\Psi}$ is also the *natural complete-data sufficient statistic* for estimating Ψ. Then, the third CM step for updating the estimate of Ψ in (3.8b) follows by computing the conditional expectation of (C.4) given the observed data $\boldsymbol{r}_k, k = 1, 2, \ldots, K$. We first find the distribution of the missing data $\boldsymbol{h}_{\mathrm{SC},k}, k = 1, 2, \ldots, K$ conditional on the observed data $\boldsymbol{r}_k, k = 1, 2, \ldots, K$. The joint distribution of $\boldsymbol{r}_k$ and $\boldsymbol{h}_{\mathrm{SC},k}$ is complex Gaussian with mean and covariance

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{r}_k \\ \boldsymbol{h}_{\mathrm{SC},k}\end{bmatrix}\right\} = \boldsymbol{0} \tag{C.5a}$$

$$\begin{aligned}\mathrm{cov}\left\{\begin{bmatrix}\boldsymbol{r}_k \\ \boldsymbol{h}_{\mathrm{SC},k}\end{bmatrix}\right\} &= \mathrm{E}\left\{\begin{bmatrix}\boldsymbol{r}_k \\ \boldsymbol{h}_{\mathrm{SC},k}\end{bmatrix}\begin{bmatrix}\boldsymbol{r}_k \\ \boldsymbol{h}_{\mathrm{SC},k}\end{bmatrix}^H\right\} \\ &= \begin{bmatrix}\Sigma_k(\boldsymbol{\gamma}) & Z_k\Psi \\ \Psi^H Z_k^H & \Psi\end{bmatrix}\end{aligned} \tag{C.5b}$$

and then, [20, result 7, pp. 508 and 509] implies that $\boldsymbol{h}_{\mathrm{SC},k}$ conditional on $\boldsymbol{r}_k$ are complex Gaussian vectors with means and covariances equal to

$$\begin{aligned}\mathrm{E}[\boldsymbol{h}_{\mathrm{SC},k}\,|\,\boldsymbol{r}_k] &= \Psi Z_k^H \Sigma_k(\boldsymbol{\gamma})^{-1}\boldsymbol{r}_k \\ &= \Psi Z_k^H Z_k W_k(\boldsymbol{\gamma}) Z_k^H \boldsymbol{r}_k \\ &= \Psi Z_k^H Z_k W_k(\boldsymbol{\gamma})(\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\boldsymbol{x})\end{aligned} \tag{C.6a}$$

$$\begin{aligned}\mathrm{cov}(\boldsymbol{h}_{\mathrm{SC},k}\,|\,\boldsymbol{r}_k) &= \Psi - \Psi Z_k^H \Sigma_k(\boldsymbol{\gamma})^{-1} Z_k \Psi \\ &= \Psi - \Psi Z_k^H Z_k W_k(\boldsymbol{\gamma}) Z_k^H Z_k \Psi\end{aligned} \tag{C.6b}$$

where (C.6a) and (C.6b) follow by using (A.1) and (3.4). Now, (3.7c) follows from (C.6a), and the third CM step in (3.8b) is obtained by substituting the most recent estimates of $\mathrm{E}[\boldsymbol{h}_{\mathrm{SC},k}\,|\,\boldsymbol{r}_k]$ and Ψ into

$$\begin{aligned}\mathrm{E}[\hat{\Psi}\,|\,\boldsymbol{r}_k] = \frac{1}{K}\sum_{k=1}^{K}\Big\{&\mathrm{E}[\boldsymbol{h}_{\mathrm{SC},k}\,|\,\boldsymbol{r}_k]\mathrm{E}[\boldsymbol{h}_{\mathrm{SC},k}\,|\,\boldsymbol{r}_k]^H \\ &+ \Psi\left[I_{n_{\mathrm{R}}n_{\mathrm{T}}} - Z_k^H Z_k W_k(\boldsymbol{\gamma}) Z_k^H Z_k \Psi\right]\Big\}.\end{aligned} \tag{C.7}$$

B. REML Estimation

We derive the ECME REML algorithm in Section III-A2. The CM step for updating the fading covariance matrix in (3.11b) follows by replacing $\Sigma_k(\boldsymbol{\gamma})^{-1}$ in (C.7) with (see [43, (3.10)] and [34, Sec. 2.3])

$$\begin{aligned}&\Sigma_k(\boldsymbol{\gamma})^{-1} - \Sigma_k(\boldsymbol{\gamma})^{-1}\Upsilon_k\left[\sum_{l=1}^{K}\Upsilon_l^H\Sigma_l(\boldsymbol{\gamma})^{-1}\Upsilon_l\right]^{-1}\Upsilon_k^H\Sigma_k(\boldsymbol{\gamma})^{-1} \\ &\quad= \left[(\sigma^2)^{(i)}\right]^{-1}\left[I_{n_{\mathrm{R}}N} - Z_k\left(Z_k^H Z_k\right)^{-1}Z_k^H\right] \\ &\qquad + Z_k W_k(\boldsymbol{\gamma}) Z_k^H - Z_k W_k(\boldsymbol{\gamma})\mathcal{A}_{\mathrm{LOS},k} \\ &\qquad \times\left[\sum_{l=1}^{K}\mathcal{A}_{\mathrm{LOS},l}^H W_l(\boldsymbol{\gamma})\mathcal{A}_{\mathrm{LOS},l}\right]^{-1}\mathcal{A}_{\mathrm{LOS},k}^H W_k(\boldsymbol{\gamma}) Z_k^H\end{aligned} \tag{C.8}$$

and substituting the most recent estimates of $\boldsymbol{x}$ and Ψ. To obtain the right-hand side of (C.8), we used (A.1) and (A.2). The CM step in (3.11a) follows by reparametrizing the variance components using $\Psi' = \Psi/\sigma^2$ and σ^2, which allows us to find the closed-form solution for the REML estimate of σ^2 that maximizes the observed-data restricted likelihood function for fixed Ψ' and $\boldsymbol{x}$

$$\begin{aligned}\hat{\sigma}^2(\Psi') = \frac{Kn_{\mathrm{R}}N}{Kn_{\mathrm{R}}N - r}\cdot\tilde{\sigma}^2 + \frac{1}{Kn_{\mathrm{R}}N - r} \\ \cdot\sum_{k=1}^{K}[\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\hat{\boldsymbol{x}}(\Psi')]^H W_k'(\Psi') \\ \times[\boldsymbol{z}_k - \mathcal{A}_{\mathrm{LOS},k}\hat{\boldsymbol{x}}(\Psi')]\end{aligned} \tag{C.9}$$

where

$$\begin{aligned}\hat{\boldsymbol{x}}(\Psi') = \left[\sum_{k=1}^{K}\mathcal{A}_{\mathrm{LOS},k}^H W_k'(\Psi')\mathcal{A}_{\mathrm{LOS},k}\right]^{-1} \\ \times\sum_{k=1}^{K}\mathcal{A}_{\mathrm{LOS},k}^H W_k'(\Psi')\boldsymbol{z}_k\end{aligned} \tag{C.10}$$

and $W_k'(\Psi')$ has been defined in (C.2). Then, the CM step for updating the estimate of σ^2 in (3.11a) is obtained by substituting the most recent estimates of $\hat{\boldsymbol{x}}(\Psi')$ and $\Psi' = \Psi/\sigma^2$ into (C.9).

Appendix D
ML Estimation for Correlated Fading and Constant $\Phi_k\Phi_k^H$

We derive the ML estimation algorithm in Section III-A3 that is described by (3.13) and iteration (3.14) and (3.15). We estimate $\boldsymbol{\rho}$ by iterating between the following two steps:

i) **(LOS component estimation)** Fix $\boldsymbol{\gamma}$, and estimate $\boldsymbol{x}$ using (3.3).

ii) **(variance-component estimation)** Fix $\boldsymbol{x}$, compute $\boldsymbol{r}_k, k = 1, 2, \ldots, K$ using (C.3), and estimate the variance-component vector $\boldsymbol{\gamma}$ by maximizing the log-likelihood function [see (3.2)]:

$$L(\boldsymbol{\gamma};\boldsymbol{x}) = -\sum_{k=1}^{K}\left[\boldsymbol{r}_k^H\Sigma_k(\boldsymbol{\gamma})^{-1}\boldsymbol{r}_k + \ln|\pi\Sigma_k(\boldsymbol{\gamma})|\right]. \tag{D.1}$$

Note that the above *alternating-projection* approach is embedded in all the algorithms discussed in this paper. We show that for constant $\Phi_k\Phi_k^H = \Gamma_\Phi$ and fixed $\boldsymbol{x}$, the ML estimates of

Ψ and σ^2 [that maximize (D.1)] are their method-of-moments estimates:

$$\hat{\sigma}^2_{\mathrm{ML}} = \frac{1}{Kn_{\mathrm{R}}(N-n_{\mathrm{T}})} \times \sum_{k=1}^{K} \boldsymbol{r}_k^H \left[I_{n_{\mathrm{R}}N} - Z_k \left(Z_k^H Z_k \right)^{-1} Z_k^H \right] \boldsymbol{r}_k \tag{D.2a}$$

$$= \frac{1}{Kn_{\mathrm{R}}(N-n_{\mathrm{T}})} \times \sum_{k=1}^{K} \boldsymbol{y}_k^H \left[I_{n_{\mathrm{R}}N} - Z_k C^{-1} Z_k^H \right] \boldsymbol{y}_k \tag{D.2b}$$

$$\hat{\Psi}_{\mathrm{ML}}(\boldsymbol{x}) = \frac{1}{K} \sum_{k=1}^{K} \left[C^{-1} Z_k^H \boldsymbol{r}_k \boldsymbol{r}_k^H Z_k C^{-1} \right] - \hat{\sigma}^2_{\mathrm{ML}} \cdot C^{-1} \tag{D.2c}$$

where (D.2b) follows from (D.2a) by using (3.12b) and (A.1). We now prove that the expressions (D.2) indeed maximize (D.1). First, simplify (D.1) using $C = Z_k^H Z_k$ and (A.3):

$$L(\boldsymbol{\gamma};\boldsymbol{x})|_{Z_k^H Z_k = C} = -\sum_{k=1}^{K} \ln|\pi\Sigma_k(\boldsymbol{\gamma})| - \frac{1}{\sigma^2} \cdot Kn_{\mathrm{R}}(N-n_{\mathrm{T}}) \cdot \hat{\sigma}^2_{\mathrm{ML}} - \mathrm{tr}\left[(\sigma^2 C^{-1} + \Psi)^{-1} \sum_{k=1}^{K} C^{-1} Z_k^H \boldsymbol{r}_k \boldsymbol{r}_k^H Z_k C^{-1} \right]. \tag{D.3}$$

The above log-likelihood function is the logarithm of a multivariate complex Gaussian pdf that belongs to the multiparameter exponential family of distributions; see [44, ch. 1.6.2]. This fact can be directly verified by inspecting (D.3), which is a linear function of the following *natural sufficient statistics*: $\hat{\sigma}^2_{\mathrm{ML}}$ and $(1/K) \cdot \sum_{k=1}^{K} C^{-1} Z_k^H \boldsymbol{r}_k \boldsymbol{r}_k^H Z_k C^{-1}$. Then, [44, Th. 2.3.1] implies that the moment estimates of σ^2 and Ψ in (D.2) are also their ML estimates, provided that $\hat{\Psi}_{\mathrm{ML}}$ in (D.2c) is positive semidefinite.

Appendix E
Cramér–Rao Bound

We derive the CRB for the vector of unknown parameters $\boldsymbol{\rho}$. Using the well-known expression for the Fisher information matrix in e.g., [20, p. 525], the CRB for $\boldsymbol{\rho}$ is

$$\mathrm{CRB} = \begin{bmatrix} \mathrm{CRB}_{x,x} & 0 \\ 0 & \mathrm{CRB}_{\gamma,\gamma} \end{bmatrix} = \begin{bmatrix} \mathcal{I}_{x,x}^{-1} & 0 \\ 0 & \mathcal{I}_{\gamma,\gamma}^{-1} \end{bmatrix} \tag{E.1a}$$

where $\mathrm{CRB}_{x,x} = \mathrm{CRB}_{[\mathrm{Re}\{\boldsymbol{x}\}^T, \mathrm{Im}\{\boldsymbol{x}\}^T]^T, [\mathrm{Re}\{\boldsymbol{x}\}^T, \mathrm{Im}\{\boldsymbol{x}\}^T]^T}$

$$\mathcal{I}_{x,x} = \mathcal{I}_{[\mathrm{Re}\{\boldsymbol{x}\}^T, \mathrm{Im}\{\boldsymbol{x}\}^T]^T, [\mathrm{Re}\{\boldsymbol{x}\}^T, \mathrm{Im}\{\boldsymbol{x}\}^T]^T} = 2\mathrm{Re}\left\{ \sum_{k=1}^{K} D_{x,k}^H \Sigma_k(\boldsymbol{\gamma})^{-1} D_{x,k} \right\} \tag{E.1b}$$

$$[\mathcal{I}_{\gamma,\gamma}]_{p,q} = \sum_{k=1}^{K} \mathrm{tr}\left\{ \Sigma_k(\boldsymbol{\gamma})^{-1} \frac{\partial \Sigma_k(\boldsymbol{\gamma})}{\partial \gamma_p} \Sigma_k(\boldsymbol{\gamma})^{-1} \frac{\partial \Sigma_k(\boldsymbol{\gamma})}{\partial \gamma_q} \right\} \tag{E.1c}$$

for $p, q = 1, 2, \ldots, \dim(\boldsymbol{\gamma})$, and

$$D_{x,k} \left[\frac{\partial(\Upsilon_k \boldsymbol{x})}{\partial(\mathrm{Re}\{\boldsymbol{x}\})^T}, \frac{\partial(\Upsilon_k \boldsymbol{x})}{\partial(\mathrm{Im}\{\boldsymbol{x}\})^T} \right] = [1, j] \otimes \Upsilon_k. \tag{E.1d}$$

Substituting (E.1d) into (E.1b) yields

$$\mathcal{I}_{x,x} = \mathrm{Re}\left\{ \begin{bmatrix} 1 & j \\ -j & 1 \end{bmatrix} \otimes P_{x,x}(\boldsymbol{\gamma}) \right\} \tag{E.2}$$

where

$$P_{x,x}(\boldsymbol{\gamma}) = 2 \cdot \sum_{k=1}^{K} \Upsilon_k^H \Sigma_k(\boldsymbol{\gamma})^{-1} \Upsilon_k = 2 \cdot \sum_{k=1}^{K} \mathcal{A}_{\mathrm{LOS},k}^H W_k(\boldsymbol{\gamma}) \mathcal{A}_{\mathrm{LOS},k}. \tag{E.3}$$

It is easy to show that

$$\mathrm{CRB}_{x,x} = \mathcal{I}_{x,x}^{-1} = \begin{bmatrix} \mathrm{Re}\{P_{x,x}(\boldsymbol{\gamma})^{-1}\} & -\mathrm{Im}\{P_{x,x}(\boldsymbol{\gamma})^{-1}\} \\ \mathrm{Im}\{P_{x,x}(\boldsymbol{\gamma})^{-1}\} & \mathrm{Re}\{P_{x,x}(\boldsymbol{\gamma})^{-1}\} \end{bmatrix}. \tag{E.4}$$

We now use (A.1) to simplify (E.1c). For $p = 1$, we have (E.5), shown at the bottom of the page, whereas for $p, q > 1$, we have

$$[\mathcal{I}_{\gamma,\gamma}]_{p,q} = [\mathcal{I}_{\gamma,\gamma}]_{\psi_{p-1},\psi_{q-1}} = [\mathcal{I}_{\psi\psi}]_{p-1,q-1} = \sum_{k=1}^{K} \mathrm{tr}\left[\Sigma_k(\boldsymbol{\gamma})^{-1} \frac{\partial \Sigma_k(\boldsymbol{\gamma})}{\partial \psi_{p-1}} \Sigma_k(\boldsymbol{\gamma})^{-1} \frac{\partial \Sigma_k(\boldsymbol{\gamma})}{\partial \psi_{q-1}} \right] = \sum_{k=1}^{K} \mathrm{tr}\left[Z_k^H Z_k W_k(\boldsymbol{\gamma}) Z_k^H Z_k \frac{\partial \Psi}{\partial \psi_{p-1}} \times Z_k^H Z_k W_k(\boldsymbol{\gamma}) Z_k^H Z_k \frac{\partial \Psi}{\partial \psi_{q-1}} \right]. \tag{E.6}$$

We partition $\mathcal{I}_{\gamma,\gamma}$ as

$$\mathcal{I}_{\gamma,\gamma} = \begin{bmatrix} i_{\sigma^2\sigma^2} & \boldsymbol{i}_{\psi,\sigma^2}^T \\ \boldsymbol{i}_{\psi,\sigma^2} & \mathcal{I}_{\psi,\psi} \end{bmatrix} \tag{E.7}$$

$$[\mathcal{I}_{\gamma,\gamma}]_{1,q} = [\mathcal{I}_{\gamma,\gamma}]_{\sigma^2,\gamma_q} = \sum_{k=1}^{K} \mathrm{tr}\left[\Sigma_k(\boldsymbol{\gamma})^{-1} \frac{\partial \Sigma_k(\boldsymbol{\gamma})}{\partial \sigma^2} \Sigma_k(\boldsymbol{\gamma})^{-1} \frac{\partial \Sigma_k(\boldsymbol{\gamma})}{\partial \gamma_q} \right] = \sum_{k=1}^{K} \mathrm{tr}\left[\Sigma_k(\boldsymbol{\gamma})^{-2} \frac{\partial \Sigma_k(\boldsymbol{\gamma})}{\partial \gamma_q} \right] = \begin{cases} n_{\mathrm{R}}(N-n_{\mathrm{T}})K\sigma^{-4} + \sum_{k=1}^{K} \mathrm{tr}\left[Z_k^H Z_k W_k(\boldsymbol{\gamma}) Z_k^H Z_k W_k(\boldsymbol{\gamma}) \right], & q = 1 \\ \sum_{k=1}^{K} \mathrm{tr}\left[Z_k^H Z_k W_k(\boldsymbol{\gamma}) Z_k^H Z_k W_k(\boldsymbol{\gamma}) Z_k^H Z_k \cdot (\partial\Psi/\partial\psi_{q-1}) \right], & q = 2, 3, \ldots, \dim(\boldsymbol{\gamma}) \end{cases} \tag{E.5}$$

where i_{σ^2,σ^2} and $\boldsymbol{i}_{\psi,\sigma^2}$ are computed using (E.5), and $\mathcal{I}_{\psi,\psi}$ is computed using (E.6). We adopt the same block partitioning of $\mathrm{CRB}_{\gamma\gamma} = \mathcal{I}_{\gamma\gamma}^{-1}$. Then

$$\mathrm{CRB}_{\psi,\psi} = \mathcal{I}_{\psi,\psi}^{-1} + \frac{\mathcal{I}_{\psi,\psi}^{-1}\boldsymbol{i}_{\psi,\sigma^2}\boldsymbol{i}_{\psi,\sigma^2}^T\mathcal{I}_{\psi,\psi}^{-1}}{i_{\sigma^2,\sigma^2} - \boldsymbol{i}_{\psi,\sigma^2}^T\mathcal{I}_{\psi,\psi}^{-1}\boldsymbol{i}_{\psi,\sigma^2}} \quad \text{(E.8a)}$$

$$\mathrm{CRB}_{\sigma^2,\sigma^2} = \frac{1}{i_{\sigma^2,\sigma^2} - \boldsymbol{i}_{\psi,\sigma^2}^T\mathcal{I}_{\psi,\psi}^{-1}\boldsymbol{i}_{\psi,\sigma^2}} \quad \text{(E.8b)}$$

which follow by using the formula for the inverse of a partitioned matrix; see, e.g., [27, Th. 8.5.11].

A. CRB for Correlated Fading and Constant $\Phi_k\Phi_k^H$

We simplify the above CRB expressions for correlated fading and constant $\Phi_k\Phi_k^H = \Gamma_\Phi$. In this case

$$W_k(\boldsymbol{\gamma}) = W(\boldsymbol{\gamma}) = (\sigma^2 C + C\Psi C)^{-1}$$

where C has been defined in (3.12b). For the unstructured LOS array response model in (2.4b), (E.4) simplifies to

$$\mathrm{CRB}_{x,x} = \frac{1}{2K} \cdot \begin{bmatrix} \mathrm{Re}\{\sigma^2 C + C\Psi C\} & -\mathrm{Im}\{\sigma^2 C + C\Psi C\} \\ \mathrm{Im}\{\sigma^2 C + C\Psi C\} & \mathrm{Re}\{\sigma^2 C + C\Psi C\} \end{bmatrix}. \quad \text{(E.9a)}$$

Under the structured LOS array response model in (2.4c), (E.4) becomes

$$\mathrm{CRB}_{x,x} = \frac{1}{2K \cdot \left(\boldsymbol{a}_{\mathrm{T,LOS}}^H \otimes \boldsymbol{a}_{\mathrm{R,LOS}}^H\right) W(\boldsymbol{\gamma})(\boldsymbol{a}_{\mathrm{T,LOS}} \otimes \boldsymbol{a}_{\mathrm{R,LOS}})} \cdot I_2. \quad \text{(E.9b)}$$

The equations in (E.5) simplify to

$$i_{\sigma^2,\sigma^2} = K \cdot \{n_{\mathrm{R}}(N - n_{\mathrm{T}})\sigma^{-4} + \mathrm{tr}[CW(\boldsymbol{\gamma})CW(\boldsymbol{\gamma})]\} \quad \text{(E.10a)}$$

$$[\mathcal{I}_{\gamma,\gamma}]_{\sigma^2,\Psi_{p,p}} = K \cdot \mathrm{tr}\left[CW(\boldsymbol{\gamma})CW(\boldsymbol{\gamma})C\frac{\partial\Psi}{\partial\Psi_{p,p}}\right] = K \cdot [CW(\boldsymbol{\gamma})CW(\boldsymbol{\gamma})C]_{p,p} \quad \text{(E.10b)}$$

and, for $p > q$

$$[\mathcal{I}_{\gamma,\gamma}]_{\sigma^2,\mathrm{Re}\{\Psi_{p,q}\}} = K \cdot \mathrm{tr}\left[CW(\boldsymbol{\gamma})CW(\boldsymbol{\gamma})C \cdot \frac{\partial\Psi}{\partial\mathrm{Re}\{\Psi_{p,q}\}}\right] = 2K \cdot \mathrm{Re}\{[CW(\boldsymbol{\gamma})CW(\boldsymbol{\gamma})C]_{p,q}\} \quad \text{(E.11a)}$$

$$[\mathcal{I}_{\gamma,\gamma}]_{\sigma^2,\mathrm{Im}\{\Psi_{p,q}\}} = K \cdot \mathrm{tr}\left[CW(\boldsymbol{\gamma})CW(\boldsymbol{\gamma})C \cdot \frac{\partial\Psi}{\partial\mathrm{Im}\{\Psi_{p,q}\}}\right] = 2K \cdot \mathrm{Im}\{[CW(\boldsymbol{\gamma})CW(\boldsymbol{\gamma})C]_{p,q}\} \quad \text{(E.11b)}$$

where $\Psi_{p,q}$ denotes the (p,q) element of Ψ, for $p,q = 1,2,\ldots,n_{\mathrm{R}}n_{\mathrm{T}}$. In addition, for $p_1 > q_1$ and $p_2 > q_2$, (E.6) becomes

$$\begin{aligned}
[\mathcal{I}_{\gamma,\gamma}]_{\mathrm{Re}\{\Psi_{p_1,q_1}\},\mathrm{Re}\{\Psi_{p_2,q_2}\}} &= [\mathcal{I}_{\gamma,\gamma}]_{\mathrm{Re}\{\Psi_{p_2,q_2}\},\mathrm{Re}\{\Psi_{p_1,q_1}\}} \\
&= K \cdot \mathrm{tr}\left[CW(\boldsymbol{\gamma})C\frac{\partial\Psi}{\partial\mathrm{Re}\{\Psi_{p_1,q_1}\}} \times CW(\boldsymbol{\gamma})C\frac{\partial\Psi}{\partial\mathrm{Re}\{\Psi_{p_2,q_2}\}}\right] \\
&= 2K \cdot \mathrm{Re}\{[CW(\boldsymbol{\gamma})C]_{q_2,p_1} \cdot [CW(\boldsymbol{\gamma})C]_{q_1,p_2} + [CW(\boldsymbol{\gamma})C]_{q_2,q_1} \cdot [CW(\boldsymbol{\gamma})C]_{p_1,p_2}\} \quad \text{(E.12a)} \\
[\mathcal{I}_{\gamma,\gamma}]_{\mathrm{Re}\{\Psi_{p_1,q_1}\},\mathrm{Im}\{\Psi_{p_2,q_2}\}} &= [\mathcal{I}_{\gamma,\gamma}]_{\mathrm{Im}\{\Psi_{p_2,q_2}\},\mathrm{Re}\{\Psi_{p_1,q_1}\}} \\
&= K \cdot \mathrm{tr}\left[CW(\boldsymbol{\gamma})C\frac{\partial\Psi}{\partial\mathrm{Re}\{\Psi_{p_1,q_1}\}} \times CW(\boldsymbol{\gamma})C\frac{\partial\Psi}{\partial\mathrm{Im}\{\Psi_{p_2,q_2}\}}\right] \\
&= -2K \cdot \mathrm{Im}\{[CW(\boldsymbol{\gamma})C]_{q_2,p_1} \cdot [CW(\boldsymbol{\gamma})C]_{q_1,p_2} + [CW(\boldsymbol{\gamma})C]_{q_2,q_1} \cdot [CW(\boldsymbol{\gamma})C]_{p_1,p_2}\} \quad \text{(E.12b)} \\
[\mathcal{I}_{\gamma,\gamma}]_{\mathrm{Im}\{\Psi_{p_1,q_1}\},\mathrm{Im}\{\Psi_{p_2,q_2}\}} &= [\mathcal{I}_{\gamma,\gamma}]_{\mathrm{Im}\{\Psi_{p_2,q_2}\},\mathrm{Im}\{\Psi_{p_1,q_1}\}} \\
&= K \cdot \mathrm{tr}\left[CW(\boldsymbol{\gamma})C\frac{\partial\Psi}{\partial\mathrm{Im}\{\Psi_{p_1,q_1}\}} \times CW(\boldsymbol{\gamma})C\frac{\partial\Psi}{\partial\mathrm{Im}\{\Psi_{p_2,q_2}\}}\right] \\
&= 2K \cdot \mathrm{Re}\{-[CW(\boldsymbol{\gamma})C]_{q_2,p_1} \cdot [CW(\boldsymbol{\gamma})C]_{q_1,p_2} + [CW(\boldsymbol{\gamma})C]_{q_2,q_1} \cdot [CW(\boldsymbol{\gamma})C]_{p_1,p_2}\} \quad \text{(E.12c)}
\end{aligned}$$

and, for $p_1 = q_1$ and $p_2 > q_2$

$$\begin{aligned}
[\mathcal{I}_{\gamma,\gamma}]_{\Psi_{p_1,p_1},\mathrm{Re}\{\Psi_{p_2,q_2}\}} &= [\mathcal{I}_{\gamma,\gamma}]_{\mathrm{Re}\{\Psi_{p_2,q_2}\},\Psi_{p_1,p_1}} \\
&= K \cdot \mathrm{tr}\left[CW(\boldsymbol{\gamma})C\frac{\partial\Psi}{\partial\Psi_{p_1,p_1}}CW(\boldsymbol{\gamma})C\frac{\partial\Psi}{\partial\mathrm{Re}\{\Psi_{p_2,q_2}\}}\right] \\
&= 2K \cdot \mathrm{Re}\{[CW(\boldsymbol{\gamma})C]_{q_2,p_1} \cdot [CW(\boldsymbol{\gamma})C]_{p_1,p_2}\} \quad \text{(E.13a)} \\
[\mathcal{I}_{\gamma,\gamma}]_{\Psi_{p_1,p_1},\mathrm{Im}\{\Psi_{p_2,q_2}\}} &= [\mathcal{I}_{\gamma,\gamma}]_{\mathrm{Im}\{\Psi_{p_2,q_2}\},\Psi_{p_1,p_1}} \\
&= K \cdot \mathrm{tr}\left[CW(\boldsymbol{\gamma})C\frac{\partial\Psi}{\partial\Psi_{p_1,p_1}}CW(\boldsymbol{\gamma})C\frac{\partial\Psi}{\partial\mathrm{Im}\{\Psi_{p_2,q_2}\}}\right] \\
&= -2K \cdot \mathrm{Im}\{[CW(\boldsymbol{\gamma})C]_{q_2,p_1} \cdot [CW(\boldsymbol{\gamma})C]_{p_1,p_2}\} \quad \text{(E.13b)}
\end{aligned}$$

and, for $p_1 = q_1$ and $p_2 = q_2$

$$[\mathcal{I}_{\gamma,\gamma}]_{\Psi_{p_1,p_1},\Psi_{p_2,p_2}} = K \cdot \mathrm{tr}\left[CW(\boldsymbol{\gamma})C\frac{\partial\Psi}{\partial\Psi_{p_1,p_1}}CW(\boldsymbol{\gamma})C\frac{\partial\Psi}{\partial\Psi_{p_2,p_2}}\right] = K \cdot |[CW(\boldsymbol{\gamma})C]_{p_1,p_2}|^2. \quad \text{(E.14)}$$

B. CRB for Independent Fading and $\Phi_k \Phi_k^H = I_{n_\mathrm{T}}$

For independent fading and unitary space-time codes ($Z_k^H Z_k = I_{n_\mathrm{R} n_\mathrm{T}}$), we have

$$W_k(\boldsymbol{\gamma}) = W(\boldsymbol{\gamma}) = \mathrm{diag}\Big\{(\sigma^2+\psi_1)^{-1}, \ldots, \big(\sigma^2+\psi_{n_\mathrm{R} n_\mathrm{T}}\big)^{-1}\Big\} \tag{E.15}$$

and the CRB expressions (E.5) and (E.6) simplify to

$$i_{\sigma^2,\sigma^2} = n_\mathrm{R}(N-n_\mathrm{T})K\cdot\sigma^{-4} + K\cdot\sum_{n=1}^{n_\mathrm{R} n_\mathrm{T}}(\sigma^2+\psi_n)^{-2} \tag{E.16a}$$

$$\boldsymbol{i}_{\psi,\sigma^2} = K\cdot\Big[(\sigma^2+\psi_1)^{-2}, (\sigma^2+\psi_2)^{-2}, \ldots \big(\sigma^2+\psi_{n_\mathrm{R} n_\mathrm{T}}\big)^{-2}\Big]^T \tag{E.16b}$$

$$\mathcal{I}_{\psi,\psi} = K\cdot\mathrm{diag}\Big\{(\sigma^2+\psi_1)^{-2}, (\sigma^2+\psi_2)^{-2}, \ldots \big(\sigma^2+\psi_{n_\mathrm{R} n_\mathrm{T}}\big)^{-2}\Big\} \tag{E.16c}$$

which yields (3.23) after applying (E.8). The CRB expressions for the LOS coefficients under the unstructured and structured LOS array response models follow by substituting $C = I_{n_\mathrm{R} n_\mathrm{T}}, \Psi = \mathrm{diag}(\psi_1, \psi_2, \ldots, \psi_{n_\mathrm{R} n_\mathrm{T}})$ and (E.15) into (E.9a) and (E.9b).

ACKNOWLEDGMENT

The authors are grateful to the anonymous reviewers for their helpful comments.

REFERENCES

[1] A. L. Moustakas, H. U. Brainier, L. Balents, A. M. Sengupta, and S. H. Simon, "Communication through a diffusive medium: Coherence and capacity," *Science*, vol. 287, pp. 287–290, Jan. 14, 2000.

[2] C.-N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 637–650, Mar. 2002.

[3] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge, U.K.: Cambridge Univ. Press, 2003.

[4] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[5] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of quadratic receivers in Rayleigh-fading channels with applications to a unified analysis of coherent and noncoherent space-time receivers," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 2383–2399, Sep. 2001.

[6] A. Dogandžić, "Chernoff bounds on pairwise error probabilities of space-time codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 1327–1336, May 2003.

[7] *IEEE J. Sel. Areas Commun.*, Apr./Jun. 2003. Special Issue on MIMO Systems and Applications.

[8] *IEEE Trans. Inf. Theory*, Oct. 2003. Special Issue on Space-time Transmission, Reception, Coding, and Signal Processing.

[9] *IEEE Trans. Signal Process.*, Nov. 2003. Special Issue on MIMO Wireless Communications.

[10] J.-M. Chaufray, P. Loubaton, and P. Chevalier, "Consistent estimation of Rayleigh fading channel second-order statistics in the context of the wideband CDMA mode of the UMTS," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3055–3064, Dec. 2001.

[11] T. L. Marzetta, "EM algorithm for estimating the parameters of a multivariate complex Rician density for polarimetric SAR," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, Detroit, MI, May 1995, pp. 3651–3654.

[12] G. Stüber, *Principles of Mobile Communication*, Second ed. Norwell, MA: Kluwer, 2001.

[13] C. Tepedelenlioğlu, A. Abdi, G. B. Giannakis, and M. Kaveh, "Estimation of Doppler spread and signal strength in mobile communications with applications to handoff and adaptive transmission," *Wireless Commun. Mobile Comput.*, vol. 1, pp. 221–242, Apr.–Jun. 2001.

[14] A. Dogandžić and J. Jin, "Estimating statistical properties of MIMO Ricean fading channels," in *Proc. 2nd IEEE Sensor Array Multichannel Signal Process. Workshop*, Rosslyn, VA, Aug. 2002, pp. 149–153.

[15] M. K. Özdemir, E. Arvas, and H. Arslan, "Dynamics of spatial correlation and implications on MIMO systems," *IEEE Commun. Mag.*, vol. 42, no. 6, pp. 514–519, Jun. 2004.

[16] H. Bölcskei, A. J. Paulraj, K.V.S. Hari, R. U. Nabar, and W. W. Lu, "Fixed broadband wireless access: State of the art, challenges, and future directions," *IEEE Commun. Mag.*, vol. 39, no. 1, pp. 100–108, Jan. 2001.

[17] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 684–702, Jun. 2003.

[18] R. Narasimhan, "Spatial multiplexing with transmit antenna and constellation selection for correlated MIMO fading channels," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2829–2838, Nov. 2003.

[19] E. Grosicki, K. Abed-Meraim, P. Loubaton, and J.-M. Chaufray, "Comparison of downlink mobile positioning methods for the UMTS FDD mode without using IPDL periods," in *Proc. Int. Symp. Signal Process. Applications*, Paris, France, Jul. 2003, pp. 347–350.

[20] S. M. Kay, *Fundamentals of Statistical Signal Processing—Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[21] G. K. Robinson, "That BLUP is a good thing: The estimation of random effects (with discussion)," *Statist. Sci.*, vol. 6, no. 1, pp. 15–50, 1991.

[22] J. H. Kotecha and A. M. Sayeed, "Transmit signal design for optimal estimation of correlated MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 546–557, Feb. 2004.

[23] M. Haardt, R. S. Thomä, and A. Richter, "Multidimensional high-resolution parameter estimation with applications to channel sounding," in *High-Resolution and Robust Signal Processing*, Y. Hua, A. B. Gershman, and Q. Cheng, Eds. New York: Marcel Dekker, 2003, ch. 5, pp. 253–337.

[24] A. Zeira and B. Friedlander, "Direction finding with time-varying arrays," *IEEE Trans. Signal Process.*, vol. 43, no. 4, pp. 927–937, Apr. 1995.

[25] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. New York: Wiley, 1997.

[26] Y. Pawitan, *In All Likelihood: Statistical Modeling and Inference Using Likelihood*. New York: Oxford Univ. Press, 2001.

[27] D. A. Harville, *Matrix Algebra From a Statistician's Perspective*. New York: Springer-Verlag, 1997.

[28] F. R. Farrokhi, G. J. Foschini, A. Lozano, and R. A. Valenzuela, "Link-optimal space-time processing with multiple transmit and receive antennas," *IEEE Commun. Lett.*, vol. 5, no. 3, pp. 85–87, Mar. 2001.

[29] G. Fuks, J. Goldberg, and H. Messer, "Bearing estimation in a Ricean channel—Part I: Inherent accuracy limitations," *IEEE Trans. Signal Process.*, vol. 49, no. 5, pp. 925–937, May 2001.

[30] J. O. Berger, B. Liseo, and R. L. Wolpert, "Integrated likelihood methods for eliminating nuisance parameters," *Statist. Sci.*, vol. 14, no. 1, pp. 1–29, 1999.

[31] M. Davidian and D. M. Giltinan, *Nonlinear Models for Repeated Measurement Data*. London, U.K.: Chapman & Hall, 1995.

[32] D. A. Harville, "Bayesian inference for variance components using only error contrasts," *Biometrika*, vol. 61, pp. 383–385, Aug. 1974.

[33] C. H. Liu and D. B. Rubin, "The ECME algorithm: A simple extension of EM and ECM with fast monotone convergence," *Biometrika*, vol. 81, pp. 633–648, Dec. 1994.

[34] X.-L. Meng and D. van Dyk, "Fast EM-type implementations for mixed effects models," *J. R. Stat. Soc.*, ser. B, vol. 60, pp. 559–578, 1998.

[35] E. F. Vonesh and V. M. Chinchilli, *Linear and Nonlinear Models for the Analysis of Repeated Measurements*. New York: Marcel Dekker, 1997.

[36] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 543–564, Mar. 2000.

[37] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1456–1467, Jul. 1999.

[38] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 10, pp. 1451–1468, Oct. 1998.

[39] B. Porat, *A Course in Digital Signal Processing*. New York: Wiley, 1997.

[40] C. R. Henderson, "Maximum Likelihood Estimation of Variance Components," Dept. Animal Sci., Cornell Univ., Ithaca, NY, 1973. unpublished manuscript.
[41] A. Dogandžić and B. Zhang, "Estimating Jakes' Doppler power spectrum parameters using the Whittle approximation," *IEEE Trans. Signal Processing*, vol. 53, no. 3, pp. 987–1005, Mar. 2005.
[42] C. R. Rao, *Linear Statistical Inference and Its Applications*, Second ed. New York: Wiley, 1973.
[43] N. Laird, N. Lange, and D. Stram, "Maximum likelihood computations with repeated measures: Applications of the EM algorithm," *J. Amer. Stat. Assoc.*, vol. 82, pp. 97–105, 1987.
[44] P. J. Bickel and K. A. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics*, Second ed. Upper Saddle River, NJ: Prentice-Hall, 2000.

Aleksandar Dogandžić (S'96–M'01) received the Dipl. Ing. degree (summa cum laude) in electrical engineering from the University of Belgrade, Belgrade, Yugoslavia, in 1995 and the M.S. and Ph.D. degrees in electrical engineering and computer science from the University of Illinois at Chicago (UIC) in 1997 and 2001, respectively.

In August 2001, he joined the Department of Electrical and Computer Engineering, Iowa State University, Ames, as an Assistant Professor. His research interests are in statistical signal processing theory and applications.

Dr. Dogandžić received the Distinguished Electrical Engineering M.S. Student Award by the Chicago Chapter of the IEEE Communications Society in 1996. He was awarded the Aileen S. Andrew Foundation Graduate Fellowship in 1997, the UIC University Fellowship in 2000, and the 2001 Outstanding Thesis Award in the Division of Engineering, Mathematics, and Physical Sciences, UIC. He is a recipient of the 2003 Young Author Best Paper Award and the 2004 Signal Processing Magazine Award by the IEEE Signal Processing Society.

Jinghua Jin received the B.S. and M.S. degrees in electrical engineering from Northern Jiaotong University, Beijing, China, in 1997 and 2000, respectively. She is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Iowa State University, Ames.

Her research interests are in statistical signal processing and its applications in wireless communications.