C. B. Ribeiro, E. Ollila, and V. Koivunen, “Stochastic maximum-likelihood method for MIMO propagation parameter estimation,” *IEEE Transactions on Signal Processing*, vol. 55, no. 1, pp. 46–55, 2007.

Stochastic Maximum-Likelihood Method for MIMO Propagation Parameter Estimation

Cássio B. Ribeiro, *Associate Member, IEEE*, Esa Ollila, *Member, IEEE*, and Visa Koivunen, *Senior Member, IEEE*

***Abstract*—In this paper, we derive a stochastic maximum-likelihood (ML) method for estimating spatio–temporal parameters for multiple-input multiple-output (MIMO) channels. Such estimators are needed in propagation studies where extensive channel measurements and sounding are required. These are seminal tasks in the process of developing advanced channel models. The proposed method employs an angular von Mises distribution model which is appropriate for angular data observed in channel measurement campaigns. The signal model is stochastic, and consequentially the method is particularly useful for estimation of the diffuse scattering component. This approach leads to lower complexity and faster convergence in comparison to deterministic models. These benefits are due to lower dimensionality of the model, leading to a simpler optimization problem. The statistical performance of the estimator is studied by establishing the Cramér–Rao lower bound (CRLB) and comparing the variances. The simulations show that the variance of the proposed estimation technique reaches the CRLB for relatively small sample size. The estimator is robust in the sense that meaningful results are obtained when applied to data generated by channel models other than the one used in its derivation.**

***Index Terms*—Channel sounding, parameter estimation.**

I. Introduction

CHANNEL sounding and extensive channel measurement campaigns are needed in developing powerful multidimensional channel models. Such models are an important tool in developing transceiver structures and designing networks for future wireless systems with high spectral efficiency. An important application is the development of multiple-input multiple-output (MIMO) communication systems, where channel sounding using multiple transmit and receive antennas and subsequent estimation of the channel propagation parameters are needed.

In radio propagation, it is common to classify the signals that reach the receiver as originated by specular reflections or diffuse scattering. The specular components are usually believed to carry most of the power and are often modeled by a relatively large number of deterministic signals with unknown parameters. Diffuse scattering is frequently regarded as noise and neglected in the models. However, it has been observed in measurement campaigns such as [1]–[3] that the diffuse scattering can be significant, and even dominant, especially in non-line-of-sight (NLOS) situations. Since rich scattering environments are of great interest to the design of MIMO systems, understanding the diffuse component is also crucial. A comprehensive description of analytical and computational models for wave propagation and multiple scattering by randomly distributed obstacles and rough surfaces can be found in [4] and [5].

Deterministic techniques for propagation parameter estimation [6], [7] commonly employ models with a large number of discrete waves. This approach leads to the maximization of highly nonlinear likelihood functions with many local maxima, which may cause convergence problems. The computational complexity and variance of estimates are increased, since parameters of a large number of waves need to be estimated. In addition, deterministic techniques using the discrete ray model such as [6] may lead to undesired artifacts, e.g., when a group of waves is estimated from a diffusely scattered cluster. In this case, the parameters of each wave will be randomly distributed with a distribution that do not correspond to any physical phenomena, being only an artifact of the estimation procedure, as observed in [8].

In this paper, we derive an estimation method for the parameters of the angular distribution of the scattering component in MIMO systems. The proposed approach is able to avoid the difficult problem of optimizing very high dimensional and highly nonlinear function. This benefit is due to the lower dimensionality of the model, leading to a simpler optimization problem. In [9], a scheme has been also proposed for the estimation of time-domain behavior of the scattering component. However, the parameters of the angular distribution are not estimated in [9]. In [10], a similar angular distribution model is used for estimation of the channel power spectrum in mobile stations.

A ML estimator for scattered sources is presented in [11], where the angular distribution is assumed Gaussian. The angular spread is assumed small, so that the correlation between the antenna elements can be approximated by a Taylor series expansion. In [12], for small angular spreads, the authors propose a first order Taylor expansion of the spatial signature of each source, leading to the generalized array manifold (GAM) model, which can be used in conjunction with well-known subspace-based methods, such as MUSIC. Again, by assuming small angular spreads, the authors in [13] show that it is possible to approximate a scattered source as a combination of two rays symmetrically located around the nominal direction. This approximation allows the use of efficient algorithms, such as ESPRIT and root-MUSIC. The resulting algorithms are called Spread

Manuscript received November 22, 2004; revised December 23, 2005. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Martin Haardt.

C. B. Ribeiro is with the Signal Processing Laboratory, Federal University of Rio de Janeiro, RJ, 21941-972, Brazil (e-mail: cassio.ribeiro@hut.fi).

E. Ollila and V. Koivunen are with the Signal Processing Laboratory, Helsinki University of Technology, FIN-02015, Finland (e-mail: esollila@wooster.hut.fi; visa.koivunen@hut.fi).

Color versions of Figs. 4–8 are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2006.882057

ESPRIT, Spread root-MUSIC, and so on. Other methods stemming from these can be found, e.g., in [14] and [15].

In this paper, we want to estimate the diffuse scattering resulting from the scattering environment surrounding the receiver. Hence, the assumption of small angular spreads does not necessarily hold, and the methods in [11]–[13] may not be applicable. In particular, we adopt the MIMO channel correlation model presented in [16], where it is assumed that the receiver is surrounded by a large number of scatterers. For the angular distribution model, we employ the von Mises distribution (see [17]), because it has finite support, which is suitable for angular data. It has been reported based on measured data that it is a good model for the angular probability density function (PDF) [18]. Another advantage of using the von Mises PDF is that it provides a closed-form expression for the MIMO channel correlation matrix. Hence, we do not need to assume that the angular spread is small. This matrix may be described analytically as a function of the parameters of the underlying random processes. Hence, it can be used to derive estimators for the channel parameters. A technique for estimating the parameters of the von Mises distribution in channel sounding applications is proposed. The estimated parameters include power azimuth spectrum, power delay spectrum (PDS), angle spread, and angular position of the transmitter array. The Cramér–Rao lower bound (CRLB) for the model is derived, and simulations are conducted to compare the mean-squared error (MSE) of the estimates to the CRLB. The estimator attains the bound with relatively small sample sizes, i.e., it is asymptotically optimal.

This paper is organized as follows. In Section II, we describe the MIMO signal model used in this paper. In Section III, the technique for parameter estimation is described. In Section IV, the CRLB is established. In Section V, we extend the model to also estimate time-domain propagation parameters in frequency-selective channels. Finally, in Section VI, we present some simulation results and compare the large sample performance of the estimation technique to the CRLB.

II. Signal Model

The transmitter is assumed to be elevated and therefore not obstructed by local scatterers, while the receiver is surrounded by a large number of local scatterers. No line-of-sight is assumed between the transmitter and the receiver. We consider that the waves are planar (far-field) and only single scattering occurs. This is what is called the "one-ring" model [19], [20], and has been used, e.g., in [16] and [19], to study the effect of fading correlation on the capacity of MIMO fading channels, and in [21] to study the effect of antenna separation on capacity. Fig. 1 illustrates the propagation environment. It should be noted that the correlation matrix derived using this model also arises in the derivation of other models, like the Saleh–Valenzuela model (SVA) in [22] and [23]. Hence, the estimation methods developed here and analysis of their performance extend to the SVA model in a straightforward manner.

Let N and M be the number of receive and transmit antennas, respectively. We then define the output of the antennas at the receiver by the $N \times 1$ vector $\mathbf{y}(k)$, the known transmitted sequences by the $M \times 1$ vector $\mathbf{u}(k)$, the $N \times M$ random MIMO channel matrix $\mathbf{H}(k)$, and the random noise $N \times 1$ vector $\mathbf{n}(k)$.

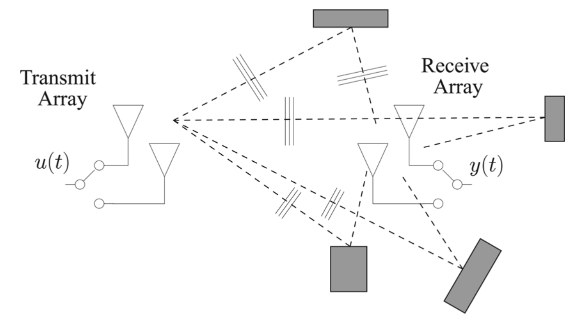

Fig. 1. Channel sounding environment.

The signals defined above are discrete-time versions of continuous-time signals sampled at time instants kT_s, where T_s is the sampling period. Using these definitions, we model the received signal as

$$\mathbf{y}(k) = \mathbf{H}(k)\mathbf{u}(k) + \mathbf{n}(k). \tag{1}$$

In this model, it is assumed that the signal bandwidth is sufficiently narrow so that the channel can be considered frequency nonselective. The statistical properties of $\mathbf{n}(k)$ will be defined in Section III.

For channel sounding applications, it is useful that information about each MIMO subchannel is available separately, and then the received signal can be represented by an $N \times M$ matrix $\mathbf{Y}(k)$, where each row corresponds to a receive antenna and each column corresponds to a transmit antenna. In Fig. 1, this is illustrated by means of switches at both ends, which is a common sounding arrangement, but other techniques can be used. The $\{i,j\}$th element of $\mathbf{Y}(k)$ represents the signal transmitted from antenna j and received by antenna i, i.e.,

$$\mathbf{Y}(k) = \begin{bmatrix} y_{0,0}(k) & \dots & y_{0,M-1}(k) \\ \vdots & \ddots & \vdots \\ y_{N-1,0}(k) & \dots & y_{N-1,M-1}(k) \end{bmatrix}. \tag{2}$$

Hence, we can modify the signal model in (1) to

$$\mathbf{Y}(k) = \mathbf{H}(k)u(k) + \mathbf{N}(k) \tag{3}$$

where $\mathbf{Y}(k)$ and $\mathbf{N}(k)$ are $N \times M$, and the scalar $u(k)$ is the transmitted sequence. The sounding sequence $u(k)$ is composed of a periodic M-sequence [24] with power P_u. It is not necessary to send different sounding sequences, since each MIMO subchannel can be measured individually. This is true for channel sounders based on switched multiple element transmit and receive antenna, like in [6], [7], [25], and [26].

The channel matrix in (3) represents the combination of all waves that impinge the receiver array after being reflected from the surrounding scatterers. Since it is difficult to explicitly describe the channel matrix as a function of the underlying waves, a different signal model is considered in [6], [7], [25], and [26], where the received signal is described as a combination of several waves with deterministic but unknown parameters and additive noise. Once the received signal is written as a function of the underlying wave parameters, ML estimation techniques, such as the space-alternating generalized expectation-maximization

(SAGE) method in [25], can be used to yield high-precision estimates of the parameters. However, since the parameters from all waves must be estimated, the dimension of the model becomes very high, and the algorithms may experience convergence problems due to local maxima in the likelihood function.

In [16], the authors present a MIMO channel model based on the "one-ring" model described here. Its correlation matrix can be described analytically and used to derive estimators for the channel parameters. In this description, the diffuse scattering is considered as realization of an underlying random process and not as a combination of deterministic signals. The authors derived a closed-form solution for the cross correlation between any two subchannels assuming that the incident angles follow a von Mises distribution [17], but other distributions can also be used. A similar model is considered in [19], [21], and [27], but the angular distribution is assumed to be uniform.

The added flexibility in the angular distribution allows for an interesting interpretation of the model as a warped ring, such that those parts of the ring corresponding to stronger reflections are closer to the receiver than those parts of ring corresponding to weaker reflections. In addition, by employing an angular distribution that is a finite mixture of angular von Mises PDFs, the model can be interpreted as a combination of many "warped rings" with different radii. Consequently, arbitrary scattering environments may be covered. This added flexibility can also be seen by the fact that both the proposed model and the SVA model [22], [23] for a single-input multiple-output (SIMO) configuration correspond to the same covariance matrix.

In case scattering occurs at both the transmitter and receiver, the "two-ring" model derived in [28] can be more appropriate. The model assumes two rings of scatterers: one around the transmitter and one around the receiver. In this model, each ray is reflected twice. The drawback in this model is that, even if the number of scatterers in both rings go to infinity, the central limit theorem does not hold, and the channel coefficients do not converge to a Gaussian random variable. Hence, it is not possible to characterize the channel based only on first- and second-order statistics. Ray-tracing methods combined with Monte Carlo simulations have been used to evaluate the model.

An alternative model for scattering at both ends of the link is the extended SVA model. The model assumes rich scattering at both ends, such that the transmit and receive scattering are independent. Under mild assumptions, the channel coefficients converge to a Gaussian variable for an infinite number of rays. Therefore, the channel can be characterized by its first- and second-order statistics.

In this paper, we will restrict the discussion to local scattering around the receiver, but the methods can be straightforwardly extended to SVA model as well.

III. Angular Propagation Parameter Estimation

The following assumptions will be used throughout this section:

a) no line of sight (NLOS) exists and no specular components are present (Rayleigh-fading channel);
b) the receiver is surrounded by a large number of local scatterers;
c) the channel is constant during one measurement cycle, and we assume $E[\mathbf{H}(k)] = 0$;
d) the additive noise $\mathrm{vec}(\mathbf{N}(k))$ is a zero-mean complex circularly symmetric Gaussian process with known covariance matrix $\mathbf{C}_n = E[\mathrm{vec}(\mathbf{N}(k))\mathrm{vec}^H(\mathbf{N}(k))]$ and independent of $\mathbf{H}(k)$;
e) the received signal $\mathrm{vec}(\mathbf{Y}(k))$ is a zero-mean complex temporally white circular Gaussian process.

Assumption e) is justified by the central limit theorem, since $\mathbf{Y}(k)$ results from a sum of an infinite number of waves, whose complex amplitudes are independent and identically distributed (i.i.d.) with finite variances [16].

In assumption d), no particular structure is assumed for the noise covariance matrix. However, in the simulations, we assume that the noise is spatially white. Assumptions d) and e) imply that $\mathrm{vec}(\mathbf{H})$ is circular Gaussian. The time index in $\mathbf{H}(k)$ has been dropped to simplify the notation, in view of assumption c).

We vectorize the model by stacking the matrices into column vectors, so $\mathrm{vec}(\mathbf{Y}(k)) = \mathrm{vec}(\mathbf{H})u(k) + \mathrm{vec}(\mathbf{N}(k))$. Since $\mathrm{vec}(\mathbf{Y}(k))$ is zero-mean complex temporally white circular Gaussian, its PDF can be written as

$$f = \frac{1}{(\pi^{NM}|\mathbf{C}_y|)^{N_s}} \exp(-N_s \mathrm{tr}\{\mathbf{C}_y^{-1}\widehat{\mathbf{C}}_y\})$$

where $\mathrm{tr}\{\cdot\}$ denotes the trace, $\mathbf{C}_y$ is the $NM \times NM$ covariance matrix of the received signal, and $\widehat{\mathbf{C}}_y$ is the $NM \times NM$ sample covariance matrix, defined as

$$\mathbf{C}_y = E[\mathrm{vec}(\mathbf{Y}(k))\mathrm{vec}^H(\mathbf{Y}(k))] \quad (4)$$

$$\widehat{\mathbf{C}}_y = \frac{1}{N_s}\sum_{k=0}^{N_s-1} \mathrm{vec}(\mathbf{Y}(k))\mathrm{vec}^H(\mathbf{Y}(k)). \quad (5)$$

Hence, after removing the constant terms not dependent on the signal, the log-likelihood function can be written as

$$\mathcal{L} \propto -\log|\mathbf{C}_y| - \mathrm{tr}\{\mathbf{C}_y^{-1}\widehat{\mathbf{C}}_y\}. \quad (6)$$

From (3) and using assumptions a)–e)

$$\begin{aligned}\mathbf{C}_y &= E[(\mathrm{vec}(\mathbf{H})u(k) + \mathrm{vec}(\mathbf{N}(k))) \\ &\quad \cdot (\mathrm{vec}(\mathbf{H})u(k) + \mathrm{vec}(\mathbf{N}(k)))^H] \\ &= E[\mathrm{vec}(\mathbf{H})\mathrm{vec}^H(\mathbf{H})]|u(k)|^2 + \mathbf{C}_n \\ &= P_u\mathbf{C}_H + \mathbf{C}_n \end{aligned} \quad (7)$$

where P_u is the power of the transmitted sequence $u(k)$, $\mathbf{C}_H$ is the $NM \times NM$ MIMO channel covariance matrix defined as

$$\mathbf{C}_H = E[\mathrm{vec}(\mathbf{H})\mathrm{vec}^H(\mathbf{H})] \quad (8)$$

and $\mathbf{C}_n$ is the $NM \times NM$ known noise covariance matrix defined as $\mathbf{C}_n = E[\mathrm{vec}(\mathbf{N}(k))\mathrm{vec}^H(\mathbf{N}(k))]$.

For a description of the MIMO channel covariance matrix as a function of propagation parameters, we will use the channel model in [16] and [20]. This channel model assumes a ring of

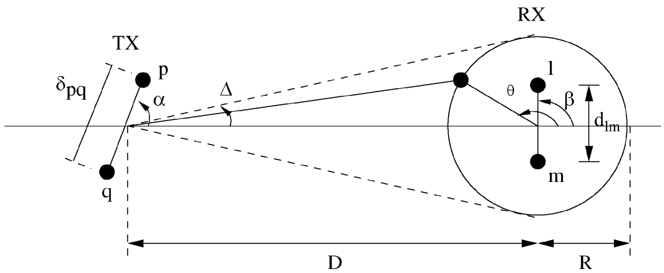

Fig. 2. Illustration of the geometrical configuration of a 2×2 channel with local scatterers at the receiver, where D is the distance between the transmitter and receiver arrays, R is the radius of the ring of scatterers around the receiver, and d_{lm} is the distance between elements l and m in the receive array.

scatterers around the receiver, as depicted in Fig. 2 for any two antennas at the transmitter and receiver.

In addition to assumptions a)–e), the following assumptions will be considered:

f) the angle spread, Δ, at the transmitter is small;
g) $D \gg R \gg \max(\delta_{pq}, d_{lm})$;
h) the position of the receiver array is known, i.e., β is known;
i) the variances $E[|h_{lp}|^2]$ for each MIMO subchannel h_{lp} connecting transmit antenna p to receive antenna l are equal and assumed to be known; we denote this variance by Ω and call it the path loss.

The transmitter spread Δ is defined as the maximum angle with respect to the line connecting the transmitter and receiver where the waves departing from the transmitter illuminate the scatterers around the receiver. For the diffuse scattering component of the radio channel, the support of the angular distribution is always 2π, even for a concentrated distribution around the mean, i.e., large κ. This implies that the angular spread is related to the radius of the ring of scatterers around the receiver and the distance between the transmitter and receiver as $\Delta \approx \arcsin(R/D)$. This relation is also observed in [20].

Assumption i) is a consequence of assumption g), i.e., the interelement distances are small compared with the distance between the transmitter and the receiver. This implies that all elements experience approximately the same path loss.

Using assumptions f)–i), it is possible to show that the cross correlation between any two subchannels lp and mq is given by [16]

$$\rho_{lp,mq} = \frac{1}{\Omega} E[h_{lp} h_{mq}^*] = e^{c_{pq} \cos\alpha} \cdot \int_{-\pi}^{\pi} \exp(c_{pq} \Delta \sin(\alpha) \sin(\theta) + b_{lm} \cos(\theta - \beta)) f(\theta) d\theta \quad (9)$$

where Ω is the known path loss, $f(\theta)$ is any angular PDF of θ, $c_{pq} = j2\pi\delta_{pq}/\lambda$, $b_{lm} = j2\pi d_{lm}/\lambda$, λ is the transmitted signal wavelength, and h_{lp} is the l,pth element of $\mathbf{H}$. Parameters Δ and α denote the angle spread at the transmitter and the angle of the transmit array relative to the line connecting the transmit and receive arrays, respectively.

For simplicity, we will assume in the sequel without loss of generality that the transmitter and receiver are equipped with uniform linear arrays (ULAs). Since the correlation is defined for each antenna pair, any other antenna configuration can be used as well. For antenna configurations other than ULA, an antenna pair must be chosen as a reference pair in the transmitter and in the receiver. Then, for the other antenna pairs, α and β must be redefined as $\alpha + \alpha_{pq}$ and $\beta + \beta_{lm}$, where α_{pq} and β_{lm} are the relative angle of the transmit and receive antenna pairs, respectively, with respect to the chosen reference antenna pair. We assume isotropic antennas, hence the effect of the field patterns are not included in the model. For nonisotropic antennas in the receiver side, the effect is equivalent to factorizing $f(\theta)$ as

$$f(\theta) = f'(\theta) g_l(\theta) g_m^*(\theta),$$

where $g_l(\theta)$ is the field pattern for the lth antenna, and $f'(\theta)$ is the actual angular PDF. This is result is similar to the SVA model in [22], [23].

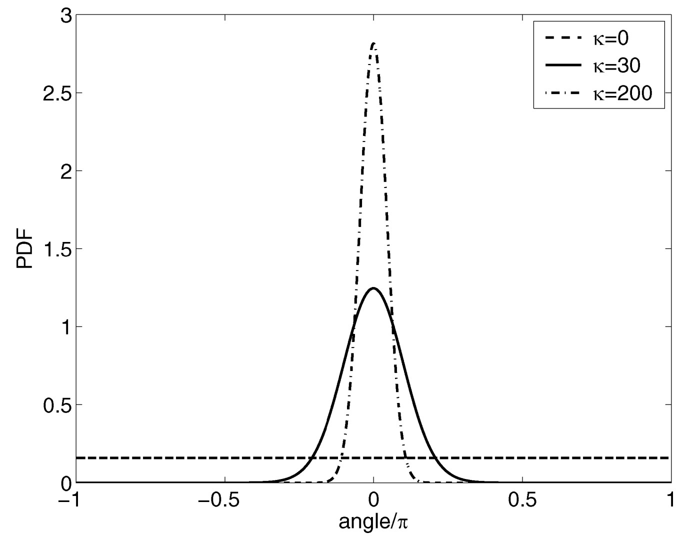

Fig. 3. von Mises PDF for different values of κ, with $\mu = 0$.

An angular PDF must satisfy $f(\theta) = f(\theta + 2\pi k)$ for any integer k. Hence, a Gaussian PDF can not be used. A suitable angular PDF is the von Mises [17], defined as

$$f(\theta) = \frac{1}{2\pi I_0(\kappa)} \exp(\kappa \cos(\theta - \mu)) \quad (10)$$

where μ is the symmetry center ("mean direction"), κ can be chosen between 0 (isotropic scattering) and ∞ (extremely concentrated), and $I_0(\cdot)$ is the modified Bessel function of the first kind of order zero. Fig. 3 illustrates the von Mises PDF for different values of κ. Using the von Mises PDF, the cross correlation in (9) may be written as [16]

$$\rho_{lp,mq} = \frac{\exp(c_{pq}\cos(\alpha))}{I_0(\kappa)} I_0(\{c_{pq}^2 \Delta^2 \sin^2(\alpha) + 2c_{pq}\Delta \sin(\alpha) \cdot (b_{lm}\sin(\beta) + \kappa \sin(\mu)) + 2\kappa b_{lm} \cos(\mu - \beta) + \kappa^2 + b_{lm}^2\}^{1/2}). \quad (11)$$

We can now find the values for μ, κ, α, and Δ that maximize the log-likelihood function in (6), i.e.,

$$\{\widehat{\mu}, \widehat{\kappa}, \widehat{\alpha}, \widehat{\Delta}\} = \operatorname{argmax}_{\mu,\kappa,\alpha,\Delta} \left\{ -\log|\mathbf{C}_y| - \operatorname{tr}\{\mathbf{C}_y^{-1} \widehat{\mathbf{C}}_y\} \right\} \quad (12)$$

where $\mathrm{tr}\{\cdot\}$ denotes the trace. In order to optimize (12), we will use a sequential quadratic programming algorithm [29], [30], implemented as the fmincon function in Matlab.

An extension of this model to multiple scatterer clusters is found in [31], where the angular distribution is a mixture of von Mises PDFs. This extension is not considered here due to lack of space. A similar approach is used in [32], where Laplace and Gaussian PDFs are used. In case specular or line-of-sight (LOS) components are present, they can be modeled by one or some mixture components. Each mixture component may have very different shape (see Fig. 3). This approach has the benefit of already estimating the local scattering around the specular reflections [31].

Another option is to write the model as a sum of a stochastic part, due to scattering, and a deterministic part, which models the specular and LOS components. This extended model can be found in [16] for the LOS component. For the estimation of these specular components, one can maximize the likelihood function directly or use an iterative procedure, such as the RIMAX algorithm [33], where the parameters of deterministic and stochastic components are estimated in an alternating manner (see [33] for detailed description).

IV. Cramér–Rao Bound

In this section, we derive the CRLB, which may be used to compare the large sample performance of estimators to the theoretical lower bound. It is also necessary for establishing asymptotic optimality of an estimator. A simulation result demonstrating the performance of the proposed estimator in comparison to the CRLB is given in Section VI. A detailed derivation can be found in Appendix I.

Define $\boldsymbol{\theta} = \{\mu, \kappa, \alpha, \Delta\}$. The $\{i,j\}$ element of the Fisher information matrix can be shown to be

$$\{\mathbf{I}\}_{i,j} = N_s \,\mathrm{tr}\{\mathbf{C}_y^{-1}\mathbf{D}_j\mathbf{C}_y^{-1}\mathbf{D}_i\} \tag{13}$$

where $\mathbf{D}_i = \mathbf{D}_{\boldsymbol{\theta}_i}$ is the derivative of $\mathbf{C}_y = \mathbf{C}_y(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}_i$.

From (7) and (11), the derivative of the (l,m)th element of the (p,q)th block of $\mathbf{C}_y(\boldsymbol{\theta})$ with respect to μ, κ, α and Δ, is given by

$$\begin{aligned}\{\mathrm{block}_{p,q}(\mathbf{D}_\mu)\}_{l,m} &= \Omega P_u \exp(c_{pq}\cos(\alpha)) I_0^{-1}(\kappa) I_1(\gamma^{1/2})\frac{1}{2}\gamma^{-1/2}\frac{\partial\gamma}{\partial\mu}\\ &= \Omega P_u \exp(c_{pq}\cos(\alpha)) I_0^{-1}(\kappa) I_1(\gamma^{1/2})\gamma^{-1/2}\\ &\quad\cdot[c_{pq}\Delta\sin(\alpha)\kappa\cos(\mu) - \kappa b_{lm}\sin(\mu-\beta)]\end{aligned} \tag{14}$$

$$\begin{aligned}\{\mathrm{block}_{p,q}(\mathbf{D}_\kappa)\}_{l,m} &= \Omega P_u \exp(c_{pq}\cos(\alpha))\Big[-I_0^{-2}(\kappa)I_1(\kappa)I_0(\gamma^{1/2})\\ &\qquad + I_0^{-1}(\kappa)I_1(\gamma^{1/2})\frac{1}{2}\gamma^{-1/2}\frac{\partial\gamma}{\partial\kappa}\Big]\\ &= \Omega P_u \exp(c_{pq}\cos(\alpha))\big[-I_0^{-2}(\kappa)I_1(\kappa)I_0(\gamma^{1/2}) +\\ &\qquad + I_0^{-1}(\kappa)I_1(\gamma^{1/2})\gamma^{-1/2}\\ &\qquad (c_{pq}\Delta\sin(\alpha)\sin(\mu)\\ &\qquad + \kappa + b_{lm}\cos(\mu-\beta))\big]\end{aligned} \tag{15}$$

$$\begin{aligned}\{\mathrm{block}_{p,q}(\mathbf{D}_\alpha)\}_{l,m} &= \Omega P_u \Big[\exp(c_{pq}\cos(\alpha)) I_0^{-1}(\kappa) I_1(\gamma^{1/2})\frac{1}{2}\gamma^{-1/2}\frac{\partial\gamma}{\partial\alpha}\\ &\qquad - \exp(c_{pq}\cos(\alpha)) I_0(\gamma^{1/2}) c_{pq}\sin(\alpha)\Big]\\ &= \Omega P_u \exp(c_{pq}\cos(\alpha))\\ &\quad I_0^{-1}(\kappa)[I_1(\gamma^{1/2})\gamma^{-1/2}(c_{pq}^2\Delta^2\sin(\alpha)\cos(\alpha)\\ &\qquad + c_{pq}\Delta\cos(\alpha)(b_{lm}\sin(\beta)+\kappa\sin(\mu)))\\ &\qquad - I_0(\gamma^{1/2})c_{pq}\sin(\alpha)]\end{aligned} \tag{16}$$

$$\begin{aligned}\{\mathrm{block}_{p,q}(\mathbf{D}_\Delta)\}_{l,m} &= \Omega P_u \exp(c_{pq}\cos(\alpha)) I_0^{-1}(\kappa) I_1(\gamma^{1/2})\frac{1}{2}\gamma^{-1/2}\frac{\partial\gamma}{\partial\Delta}\\ &= \Omega P_u \exp(c_{pq}\cos(\alpha)) I_0^{-1}(\kappa) I_1(\gamma^{1/2})\gamma^{-1/2}\\ &\quad\cdot[c_{pq}^2\Delta\sin^2(\alpha) + c_{pq}\sin(\alpha)(b_{lm}\sin(\beta)+\kappa\sin(\mu))]\end{aligned} \tag{17}$$

where

$$\gamma = (c_{pq}^2\Delta^2\sin^2(\alpha) + 2c_{pq}\Delta\sin(\alpha)(b_{lm}\sin(\beta)+\kappa\sin(\mu)) + 2\kappa b_{lm}\cos(\mu-\beta) + \kappa^2 + b_{lm}^2),$$

Ω is the path loss, P_u is the power of the transmitted sequence, $I_0(\cdot)$ and $I_1(\cdot)$ are the modified Bessel function of the first kind of order zero and one, respectively.

The Cramér–Rao bounds can then be evaluated for each parameter as the diagonal elements of the inverse of the Fisher information matrix in (13).

V. Time-Domain Propagation Parameter Estimation

We can extend the model in (3) to frequency-selective channels, if we note that even though the mechanisms leading to angle and time dispersion are highly related, the power azimuth-delay spectrum can be obtained as a function of the individual spectra [6]. Based on these insights, we generalize the model in (2) to

$$\mathbf{Y}(k) = \mathbf{H}\sum_{l=0}^{L-1} w(l)u(k-l) + \mathbf{N}(k) \tag{18}$$

where L is the maximum number of nonzero elements in the channel impulse response. Since $w(k)$ is complex valued, it can be written as $w(k) = |w(k)|e^{j\varphi(k)}$, where we assume $\varphi(k)$, $k = 0, \ldots, L-1$ are independent and uniformly distributed in the interval $[0, 2\pi)$, and $|w(k)|$ is deterministic.

In (6), it is assumed that the observations are i.i.d. This assumption does not necessarily hold if the channel is frequency selective, hence some loss of performance compared to optimal procedure may be experienced. However, this loss is not expected to be significant since all spatial information is concentrated on matrix $\mathbf{H}$. Hence, the estimator can be derived with low complexity in comparison to more optimal procedures. As shown in Section VI, the proposed method produces reliable estimates even for frequency-selective channels.

As shown in Appendix II, using assumptions a)–d) and the assumptions for $w(k)$ above, we can write

$$\mathbf{C}_y \approx P_u\mathbf{C}_H + \mathbf{C}_n \tag{19}$$

where P_u is defined in (7). Hence, the angular parameters can be estimated with the same techniques used for the narrowband model. In this section we focus on the time-domain parameters, which can be estimated using standard estimation techniques.

The PDS have been investigated in several measurement campaigns, and we will only discuss it briefly here. Experimental evidence indicates that the PDS may be accurately modeled by an exponential decaying function [6], i.e.,

$$|w(k)|^2 \propto e^{-(1/\sigma_d)(k-k_d)}, k \geq 0 \tag{20}$$

where k_d is the discretized group delay, or time of arrival (TOA). Thus, we only have to estimate the TOA of the cluster and the decay factor σ_d of the exponential.

Here, we assume the transmitted signal is a length-N_s M-sequence, which has a cyclic autocorrelation function $r(i)$ with the property that $r(0) = N_s$ and $r(i) = -1, 0 < i \leq N_s - 1$ [34]. The cyclic cross correlation between $u(k)$ and the $\{n,m\}$th element of $\mathbf{Y}(k)$ is given by

$$\begin{aligned} r_{nm}^{yu}(i) &= \sum_{k=0}^{N_s-1} y_{nm}(k)u^*(k-i) \\ &= \sum_{k=0}^{N_s-1}\left[h_{nm}\sum_{l=0}^{L-1} w(l)u(k-l) + n_{nm}(k)\right]u^*(k-i) \end{aligned} \tag{21}$$

where h_{nm} is the $\{n,m\}$th element of $\mathbf{H}$, and $n_{nm}(k)$ is the $\{n,m\}$th element of $\mathbf{N}(k)$. Using the autocorrelation properties of the transmitted M-sequence, and since it is assumed that the cyclic cross correlation between $u(k)$ and the measurement noise $n_{nm}(k)$ is negligible, we obtain

$$r_{nm}^{yu}(i) \approx N_s h_{nm} w(i)|u(k-i)|^2. \tag{22}$$

From (11), $E[|h_{nm}|^2] = 1$, and it is assumed that h_{nm} and $w(k)$ are independent. By averaging $|r_{nm}^{yu}(i)|^2$ over all antennas, we get

$$r^{yu}(i) = \frac{1}{NM}\sum_{n=1}^{N}\sum_{m=1}^{M}|r_{nm}^{yu}(i)|^2 \approx N_s^2 P_u |w(i)|^2 \tag{23}$$

where $P_u = |u(k)|^2$.

We can then derive the following algorithm for estimation of the time-dependent parameters:

a) Calculate

$$\begin{aligned} \xi(l) = (\widehat{|w(l)|^2}) &= \frac{1}{NMIN_s^2 P_u} \\ &\times \sum_{i=1}^{I}\sum_{n=1}^{N}\sum_{m=1}^{M}\left|\sum_{k=0}^{N_s-1} u(k-l)^* y_{nm}(k-(i-1)N_s)\right|^2 \end{aligned} \tag{24}$$

where I is the number of snapshots and N_s is the number of samples in each snapshot.

b) Due to the noisy nature of $\xi(l)$, it is not convenient in practice to set the estimate of the TOA as the sample corresponding to the highest peak in $\xi(l)$. Instead, we take the first sample $\hat{k}$ such that $\xi(\hat{k}) \geq \epsilon$, where ϵ is a fixed threshold. Since in channel sounding applications the signal to noise ratio is typically high, the choice of ϵ is not critical. In the simulations in this paper, we use $\epsilon = 0.02$. For other applications, ϵ can be determined based on standard detection theory [35].

c) Calculate the decay factor σ_d using the sequence $\xi(l)$, $l = \hat{k}, \ldots, N_s - 1$. From (20), $\log|w(k)|^2 \propto -k/\sigma_d$, and hence σ_d can be obtained by least-squares polynomial fitting [36].

Similar techniques have been derived for CDMA receivers, where sequences with properties similar to M-sequences are used [34], [37].

VI. Simulation Results

In this section, we present some results obtained in simulation in order to verify the theoretical results and assess the performance of the estimator.

A. Simulation 1: Performance Comparison to the CRLB

In all simulations both transmitter and receiver are composed of an uniform linear array (ULA) with $M = N = 4$ antennas. The received signal is generated as [38], [39]

$$\text{vec}(\mathbf{Y}(k)) = \mathbf{R}^{1/2}\mathbf{w}(k) + \text{vec}(\mathbf{N}(k)) \tag{25}$$

where $\mathbf{R}^{1/2}$ is obtained as by the Cholesky decomposition [40] of $E[\text{vec}(\mathbf{H})\text{vec}^H(\mathbf{H})]$, $\mathbf{w}(k)$ and $\text{vec}(\mathbf{N}(k))$ are circular complex white Gaussian processes with known variance, and $E[\mathbf{w}(k)\mathbf{w}^H(k)]$ is the identity matrix. The signal-to-noise ratio (SNR) per antenna is fixed to 20 dB and defined as $\text{SNR} = 10\log_{10}(E[|y(k)|^2]/E[|n(k)|^2])$. The path loss is fixed as $\Omega = 1$. It should be noted that even though $\mathbf{R}^{1/2}$ and $\mathbf{w}(k)$ in (25) do not correspond directly to $\mathbf{H}(k)$ and $u(k)$ in (3), the vector $\text{vec}(\mathbf{Y}(k))$ in (25) and (3) have the same statistics, since

$$\begin{aligned} E[\mathbf{R}^{1/2}\mathbf{w}(k)\mathbf{w}^H(k)(\mathbf{R}^{1/2})^H] &= \mathbf{R}^{1/2}(\mathbf{R}^{1/2})^H \\ &= E[\text{vec}(\mathbf{H})\text{vec}^H(\mathbf{H})]. \end{aligned}$$

In Fig. 4, we compare the MSE of the estimates to the CRLB as a function of the number of samples. The MSE is estimated over 200 runs. The angle of the receiver array is $\beta = 90°$, the symmetry center is $\mu = 70° + \beta$, the dispersion $\kappa = \{5,30,200\}$, the angle of the transmitter array is $\alpha = 100°$, and the angle spread at the transmitter is $\Delta = 10°$. It can be seen that the MSE reaches the CRLB already for small number of samples for all values of κ. In Fig. 5, we compare the MSE of the estimates to the CRLB as a function of SNR. Again, it can be seen that the MSE reaches the CRLB for all parameters.

B. Simulation 2: Channel With Time Dispersion

In the next simulations, we will verify the behavior of the proposed estimation method for channels with time dispersion. The transmitter has $M = 1$ antenna and the receiver array has

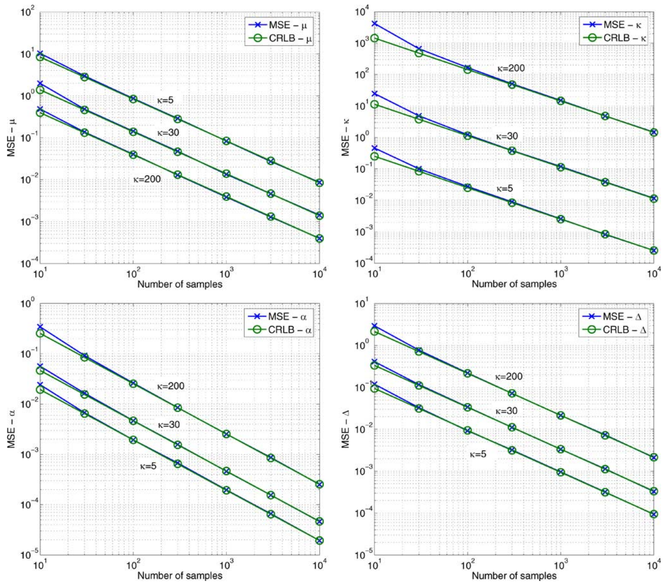

Fig. 4. MSE of the proposed estimator and the CRLB for each unknown parameter as a function of the number of samples for $\kappa = \{5{,}30{,}200\}$. The receive and transmit array are ULA with four antennas. The MSE reaches the CRLB for relatively small number of samples for all values of κ.

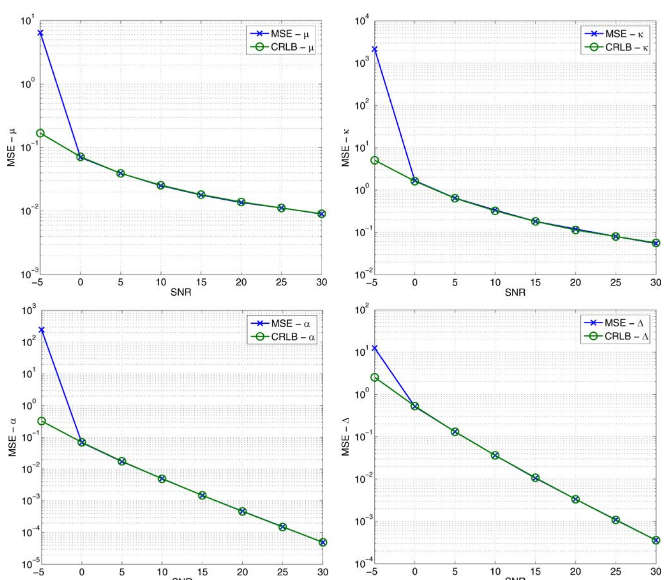

Fig. 5. MSE of the proposed estimator and the CRLB for each unknown parameter as a function of SNR. The receive and transmit array are ULA with four antennas. The MSE reaches the CRLB for relatively small number of samples.

an uniform linear array (ULA) with $N = 11$ antennas. We simulate the channel as a superposition of a large number of waves, defined as

$$\text{vec}(\mathbf{Y}(k)) = \sum_{n=1}^{N_w} \mathbf{x}_n(k) + \text{vec}(\mathbf{N}(k)) \tag{26}$$

where

$$\mathbf{x}_n(k) = \frac{1}{\sqrt{N_w}} e^{j\psi} \mathbf{c}(\phi_n) u(k - \tau_n),\ n = 1, \ldots, N_w, \tag{27}$$

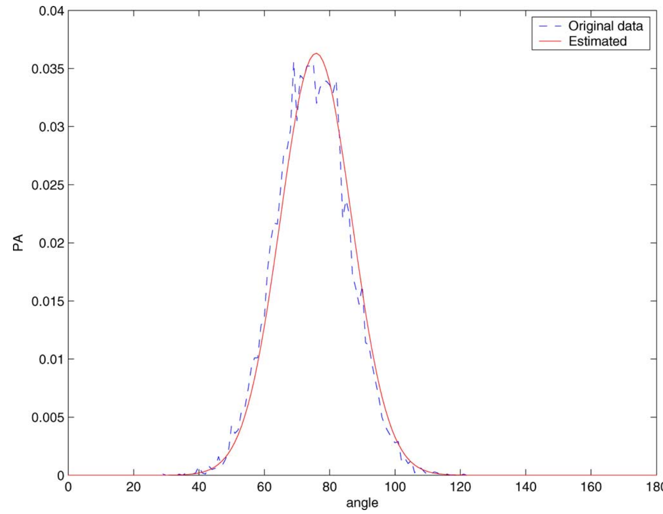

Fig. 6. Comparison between the true and estimated power azimuth spectrum. The estimated spectrum is given by the von Mises PDF with estimated parameters $\hat{\mu} = 75.78^\circ + \beta$ and $\hat{\kappa} = 27.08$. The receive array is ULA with 11 antennas.

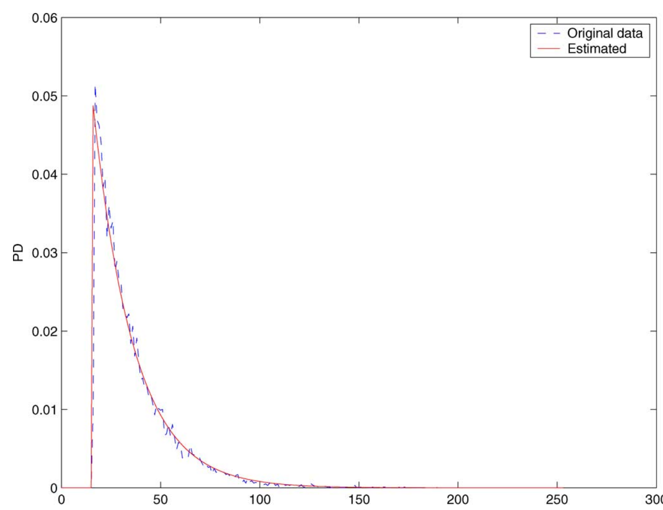

Fig. 7. Comparison between the true and estimated power delay spectrum. The estimated spectrum is given by an exponential decaying function with $\sigma_d = 2.51\ \mu$s and a delay of 16 samples. The receive array is ULA with 11 antennas.

where ψ is a random phase, $\mathbf{c}(\phi_n)$ is the steering vector, and τ_n is the delay of the nth wave. In order to model the scattered signal, we would need an infinite number of waves, but we will use $N_w = 10\,000$ to approximate the channel. The random phase ψ is uniformly distributed in $[0, 2\pi)$. The angle of arrival ϕ_n follows a von Mises distribution with parameters $\mu = 75^\circ + \beta$ and $\kappa = 26.93$. The delays were chosen from an exponential distribution with decay factor $\sigma_d = 2.34\ \mu$s.

The sounding sequence $u(k)$ is a complex length-127 M-sequence, with symbol duration $T_p = 1/(4.096 \times 10^6)$ s and sampling period $T_s = T_p/2$. The random noise $\text{vec}(\mathbf{N}(k))$ is generated from a complex white Gaussian process. The results are obtained during one snapshot, and are averaged over 50 runs.

In Figs. 6 and 7, we compare the true power-azimuth and power-delay spectra with the estimates obtained using the method described in this paper. The estimated power-azimuth spectrum is given by the von Mises PDF with parameters

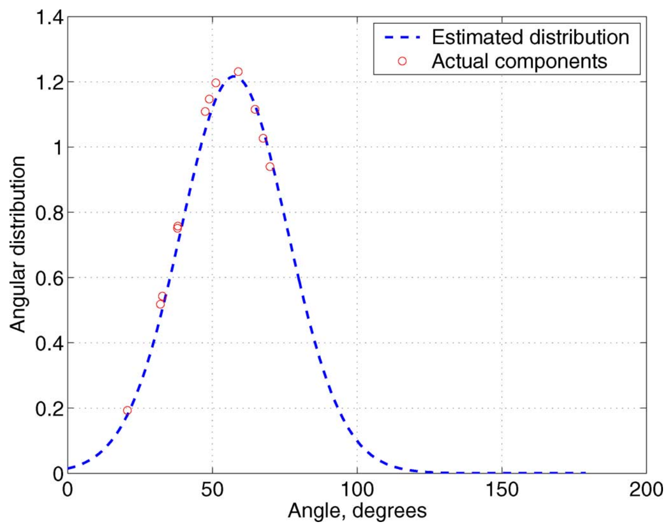

Fig. 8. Comparison between the true and estimated power azimuth spectrum. The channel is simulated using experimentally validated I-METRA's MIMO channel model [38]. Since the model is deterministic, only 12 specular paths are present. The estimated angular distribution provides a good fit for the envelope of the power azimuth spectrum.

$\widehat{\mu} = 75.78^\circ + \beta$ and $\widehat{\kappa} = 27.08$ calculated using (12). The estimated power-delay spectrum is given by (20) with $\widehat{\sigma}_d = 2.51$ μs and a delay of 16 samples, calculated as in Section V. The actual power-azimuth and power-delay spectra are calculated as the histogram of delays τ_n and angles ϕ_n with bin width $\Delta\tau = T_p$ and $\Delta\phi = 1^\circ$, respectively. The word spectrum is used since these histograms are closely related to the power azimuth-delay spectrum in [6]. The estimated spectra closely match the true ones.

C. Simulation 3: I-METRA MIMO Channel

For the last simulation, we simulate the channel using experimentally validated I-METRA's MIMO channel model [38]. The channel model is double-directional and has been validated by measurement data in picocell and microcell environments. It is a widely used realistic channel whose parameters are derived based on measurement results.

The chip rate is 3.84 microchips per second (Mcps), the carrier frequency is equal to 2.15 GHz, and the velocity is 0.01 km/h. The transmitter has one antenna and the receiver has an ULA with four antennas. Twelve paths are generated arriving at time instant $k = 0$ with angles of arrival generated randomly from a von Mises distribution. The power for each path is given by the von Mises PDF calculated at the corresponding angle. This setting is not exactly the same one considered in [38], but it is applied here to simulate a source with local spreading around the nominal direction. The sounding sequence is a length-1023 M-sequence. The signal obtained with the METRA model is normalized such that SNR $\geq$ 20 dB per antenna.

In Fig. 8, we compare the estimated angular distribution obtained by the proposed method and the actual power azimuth spectrum for one realization of the channel. Even though the channel considered in this simulation does not follow exactly the model in Sections II and III, the estimated distribution provides a good fit to the envelope of the actual power azimuth spectrum. This result is not surprising, since stochastic ML estimation can be applied for deterministic signals, usually achieving better performance than deterministic ML for small number of sensors and low SNR [41].

VII. Conclusion

In this paper, we derived a stochastic ML technique to estimate the angular distribution of signals in a MIMO channel. Since MIMO systems rely on rich scattering, modeling and estimation of the diffuse scattering is of high importance. We used the channel model derived in [16] in order to compute the channel correlation matrix that depends on the unknown parameters. The underlying distribution model is angular von Mises distribution which is well suited for directional data. The methods extends to different scattering models and may be extended to arbitrary scattering environments using mixture model [31].

This stochastic model leads to more compact representation of the propagation model, in particular for channels with diffuse scattering. Moreover, the computational complexity is reduced comparing to deterministic techniques, because fewer parameters are needed, simplifying the optimization problem.

We also derived the CRLB for the model and evaluated the large sample performance of the estimator by comparing the variances to the CRLB. The results show that the variance for the proposed estimation technique reaches the CRLB for all parameters with relatively small sample sizes.

We used two different channel models in order to assess the quality of the estimates provided by the proposed method when some of the assumptions used during the development do not hold. The results show that the method is able to provide meaningful and precise estimates of the power-azimuth and power-delay spectra.

Appendix I

The log-likelihood equation is given by

$$\mathcal{L}\{\boldsymbol{\theta}\} = -N_s \log \pi^N - N_s \log|\mathbf{C}_y(\boldsymbol{\theta})| - N_s \mathrm{tr}\{\mathbf{C}_y^{-1}(\boldsymbol{\theta})\widehat{\mathbf{C}}_y\}. \tag{28}$$

In order to compute the derivatives of $\mathcal{L}\{\boldsymbol{\theta}\}$ with respect to each parameter we will use the relations [42]

$$\mathbf{C}(\mathbf{X}) = \mathrm{tr}\{\mathbf{AX}\} \Rightarrow d\mathbf{C}(\mathbf{X}) = \mathrm{tr}\{\mathbf{A}d\mathbf{X}\} \tag{29}$$

$$\mathbf{C}(\mathbf{X}) = \log|\mathbf{X}| \Rightarrow d\mathbf{C}(\mathbf{X}) = \mathrm{tr}\{\mathbf{X}^{-1}d\mathbf{X}\} \tag{30}$$

$$\mathbf{C}(\mathbf{X}) = \mathbf{X}^{-1} \Rightarrow d\mathbf{C}(\mathbf{X}) = -\mathbf{X}^{-1}(d\mathbf{X})\mathbf{X}^{-1}. \tag{31}$$

Using the expressions above, we can calculate

$$\begin{aligned} d\mathcal{L}\{\boldsymbol{\theta}\} &= -N_s d\log|\mathbf{C}_y(\boldsymbol{\theta})| - N_s \mathrm{tr}\{d\mathbf{C}_y^{-1}(\boldsymbol{\theta})\widehat{\mathbf{C}}_y\} \\ &= -N_s \mathrm{tr}\{\mathbf{C}_y^{-1}(\boldsymbol{\theta})d\mathbf{C}_y(\boldsymbol{\theta})\} \\ &\quad + N_s \mathrm{tr}\{\mathbf{C}_y^{-1}(\boldsymbol{\theta})(d\mathbf{C}_y(\boldsymbol{\theta}))\mathbf{C}_y^{-1}(\boldsymbol{\theta})\widehat{\mathbf{C}}_y\} \\ &= -N_s \mathrm{tr}\{(\mathbf{I} - \mathbf{C}_y^{-1}(\boldsymbol{\theta})\widehat{\mathbf{C}}_y)\mathbf{C}_y^{-1}(\boldsymbol{\theta})(d\mathbf{C}_y(\boldsymbol{\theta}))\}. \end{aligned} \tag{32}$$

$$\mathbf{C}_y'' = \mathbf{C}_H \sum_{l=0}^{L-1} \sum_{l'=0, l'\neq l}^{L-1} E[e^{j\varphi(l)}]E[e^{-j\varphi(l')}]|w(l)||w(l')|u(k-l)u^*(k-l') + \mathbf{C}_n. \tag{41}$$

Hence, using (32), we can write

$$\frac{\partial \mathcal{L}\{\boldsymbol{\theta}\}}{\partial \boldsymbol{\theta}_i} = -N_s \mathrm{tr}\{(\mathbf{I}_N - \mathbf{C}_y^{-1}\widehat{\mathbf{C}}_y)\mathbf{C}_y^{-1}\mathbf{D}_i\} \tag{33}$$

where $\mathbf{D}_i = \mathbf{D}_{\boldsymbol{\theta}_i}$.

We obtain the second derivatives by differentiating (33) with respect to $\boldsymbol{\theta}_j$ as

$$\begin{aligned}\frac{\partial^2 \mathcal{L}\{\boldsymbol{\theta}\}}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_j} =& -N_s \mathrm{tr}\Bigg\{\left[\frac{\partial(\mathbf{I}_N - \mathbf{C}_y^{-1}\widehat{\mathbf{C}}_y)}{\partial \boldsymbol{\theta}_j}\right]\mathbf{C}_y^{-1}\mathbf{D}_i \\ &+ (\mathbf{I}_N - \mathbf{C}_y^{-1}\widehat{\mathbf{C}}_y)\left[\frac{\partial(\mathbf{C}_y^{-1}\mathbf{D}_i)}{\partial \boldsymbol{\theta}_j}\right]\Bigg\} \\ =& -N_s \mathrm{tr}\Bigg\{\mathbf{C}_y^{-1}\mathbf{D}_j\mathbf{C}_y^{-1}\widehat{\mathbf{C}}_y\mathbf{C}_y^{-1}\mathbf{D}_i \\ &+ (\mathbf{I}_N - \mathbf{C}_y^{-1}\widehat{\mathbf{C}}_y)\left[\frac{\partial(\mathbf{C}_y^{-1}\mathbf{D}_i)}{\partial \boldsymbol{\theta}_j}\right]\Bigg\}. \end{aligned} \tag{34}$$

We can then write

$$E\left[\frac{\partial^2 \mathcal{L}\{\boldsymbol{\theta}\}}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_j}\right] = -N_s \mathrm{tr}\{\mathbf{C}_y^{-1}\mathbf{D}_j\mathbf{C}_y^{-1}\mathbf{D}_i\} \tag{35}$$

where the fact that $E[\widehat{\mathbf{C}}_y] = (1/N_s)\sum_{k=0}^{N_s-1} E[\mathbf{y}(k)\mathbf{y}(k)^H] = \mathbf{C}_y$ is used. The $\{i,j\}$ element of the Fisher information matrix is then given by

$$\{\mathbf{I}\}_{i,j} = N_s \mathrm{tr}\{\mathbf{C}_y^{-1}\mathbf{D}_j\mathbf{C}_y^{-1}\mathbf{D}_i\} \tag{36}$$

where $\mathbf{D}_i$ is the derivative of $\mathbf{C}_y(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}_i$.

Appendix II

From (18), the covariance matrix of the received signal $\mathbf{Y}(k)$ is given by

$$\begin{aligned}&\mathbf{C}_y \\ &= E[\mathrm{vec}(\mathbf{Y}(k))\mathrm{vec}^H(\mathbf{Y}(k))] \\ &= E\left[\left(\sum_{l=0}^{L-1}\mathrm{vec}(\mathbf{H})e^{j\varphi(l)}|w(l)|u(k-l) + \mathrm{vec}(\mathbf{N}(k))\right)\right. \\ &\quad \left.\times\left(\sum_{l'=0}^{L-1}\mathrm{vec}(\mathbf{H})^H e^{-j\varphi(l')}|w(l')|u^*(k-l') + \mathrm{vec}(\mathbf{N}(k))\right)\right] \\ &= E[\mathrm{vec}(\mathbf{H})\mathrm{vec}^H(\mathbf{H})] \\ &\quad \times\sum_{l=0}^{L-1}\sum_{l'=0}^{L-1}E[e^{j(\varphi(l)-\varphi(l'))}]|w(l)||w(l')|u(k-l)u^*(k-l') + \mathbf{C}_n\end{aligned} \tag{37}$$

where $\mathbf{C}_n$ is the noise covariance matrix, and it was used the assumptions that $\mathbf{H}$ is constant, $\mathbf{H}$, $\mathbf{N}(k)$, and $\varphi(k)$ are independent.

For $l = l'$, (37) simplifies to

$$\mathbf{C}_y' = \mathbf{C}_H \sum_{l=0}^{L-1} |w(l)|^2|u(k-l)|^2 + \mathbf{C}_n \tag{38}$$

where $\mathbf{C}_H = E[\mathrm{vec}(\mathbf{H})\mathrm{vec}^H(\mathbf{H})]$. For the channel sounding technique based on time-division multiplex that is assumed in this paper, a guard time is defined that is longer than the channel impulse response. Hence, for $k \geq 0$, the convolution term in (38) converges to

$$\sum_{l=0}^{L-1} |w(l)|^2|u(k-l)|^2 \approx P_u \sum_{l=0}^{L-1} |w(l)|^2 \tag{39}$$

where $P_u = |u(k-l)|^2$. Since the path-loss effects are already taken into account in Ω, defined in assumption i) and (9), we have $\sum_{l=0}^{L-1} |w(l)|^2 = 1$. Hence, for $k \geq 0$

$$\mathbf{C}_y' \approx P_u \mathbf{C}_H + \mathbf{C}_n. \tag{40}$$

For $l \neq l'$, (37), we obtain (41), shown at the top of the page. The expected value of $e^{j\varphi}$ in (41) is easily calculated as

$$E[e^{j\varphi}] = \int_0^{2\pi} \frac{1}{2\pi} e^{j\varphi} d\varphi = 0. \tag{42}$$

Hence,

$$\mathbf{C}_y = \mathbf{C}_y' + \mathbf{C}_y'' - \mathbf{C}_n \approx P_u \mathbf{C}_H + \mathbf{C}_n. \tag{43}$$

References

[1] Q. H. Spencer, B. D. Jeffs, M. A. Jensen, and A. L. Swindlehurst, "Modeling the statistical time and angle of arrival characteristics of an indoor multipath channel," *IEEE J. Sel. Areas Commun.*, vol. 18, pp. 347–360, Mar. 2000.

[2] W. Newhall, R. Mostafa, K. Dietze, J. Reed, and W. Stutzmad, "Measurement of multipath signal component amplitude correlation coefficients versus propagation delay," in *Proc. IEEE Radio Wireless Conf., 2002*, Aug. 2002, pp. 133–136.

[3] T. Rappaport, S. Seidel, and R. Singh, "900-mhz multipath propagation measurements for us digital cellular radiotelephone," *IEEE Trans. Veh. Technol.*, vol. 39, pp. 132–139, May 1990.

[4] A. Ishimaru, "Wave propagation and scattering in random media and rough surfaces," *Proc. IEEE*, vol. 79, no. 10, pp. 1359–1366, Oct. 1991.

[5] C. M. Chu and S. W. Churchill, "Multiple scattering by randomly distributed obstacles—Methods of solution," *IEEE Trans. Antennas Propag.*, vol. 4, pp. 142–148, Apr. 1956.

[6] K. I. Pedersen, P. E. Mogensen, and B. H. Fleury, "A stochastic model of the temporal and azimuthal dispersion seen at the base station in outdoor propagation environments," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 437–447, Mar. 2000.

[7] B. H. Fleury, P. Jourdan, and A. Stucki, "High-resolution channel parameter estimation for MIMO applications using the SAGE algorithm," in *Proc. 2002 Int. Zurich Seminar BroadBand Communications. Access, Transmission, Networking*, Feb. 2002, pp. 30-1–30-9.

[8] M. Bengtsson and B. Völcker, "On the estimation of azimuth distributions and azimuth spectra," in *Proc. IEEE VTS 54th Vehicular Technology Conf., 2001*, Oct. 2001, vol. 3, pp. 1612–1615.

[9] A. Richter and R. Thoma, "Parametric modelling and estimation of distributed diffuse scattering components of radio channels," in *COST 273 Temporary Document,*, Sep. 2003, TD(03) 198.

[10] A. Abdi, J. A. Barger, and M. Kaveh, "A parametric model for the distribution of the angle of arrival and the associated correlation function and power spectrum at the mobile station," *IEEE Trans. Veh. Technol.*, vol. 51, pp. 425–434, May 2002.

[11] T. Trump and B. Ottersten, "Estimation of nominal direction of arrival and angular spread using an array of sensors," *Signal Process.*, vol. 50, no. 1–2, pp. 57–69, Apr. 1996.

[12] D. Asztély, B. Ottersten, and A. L. Swindlehurst, "A generalized array manifold model for local scattering in wireless communications," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, 1997, vol. 5, pp. 4021–4024.

[13] M. Bengtsson and B. Ottersten, "Low-complexity estimators for distributed sources," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2185–2194, Aug. 2000.

[14] O. Besson and P. Stoica, "Decoupled estimation of DOA and angular spread for a spatially distributed source," *IEEE Trans. Signal Process.*, vol. 48, no. 7, pp. 1872–1882, Jul. 2000.

[15] Q. Wan, J. Yuan, and Y. N. Peng, "Estimation of nominal direction of arrival and angular spread using the determinant of the data matrix," in *Proc. 4th Int. Workshop Mobile Wireless Communications Network*, 2002, pp. 76–79.

[16] A. Abdi and M. Kaveh, "A space–time correlation model for multielement antenna systems in mobile fading channels," *IEEE J. Sel. Areas Commun.*, vol. 20, pp. 550–560, Apr. 2002.

[17] K. V. Mardia*, Statistics of Directional Data*. New York: Academic, 1972.

[18] A. Abdi and M. Kaveh, "A versatile spatio–temporal correlation function for mobile fading channels with non-isotropic scattering," in *Proc. IEEE Workshop Statistical Signal Array Processing*, 2000, pp. 58–62.

[19] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Mar. 2000.

[20] K. Yu and B. Ottersten, "Models for MIMO propagation channels, A review," *Wiley J. Wireless Commun. Mobile Comput. (Special Issue on Adaptive Antennas and MIMO Systems)*, vol. 2, no. 7, pp. 653–666, Nov. 2002.

[21] D. Chizhik, F. Rashid-Farrokhi, J. Ling, and A. Lozano, "Effect of antenna separation on the capacity of BLAST in correlated channels," *IEEE Commun. Lett.*, vol. 4, pp. 337–339, Nov. 2000.

[22] J. Wallace and M. Jensen, "Statistical characteristics of measured MIMO wireless channel data and comparison to conventional models," in *Proc. IEEE Vehicular Technology Conf. (VTC)—Fall*, 2001, vol. 2, pp. 1078–1082.

[23] J. W. Wallace and M. Jensen, "Modeling the indoor MIMO wireless channel," *IEEE Trans. Antennas Propag.*, vol. 50, no. 5, pp. 591–599, May 2002.

[24] J. S. Lee and L. E. Miller*, CDMA Systems Engineering Handbook* Boston, MA: Artech House, 1998.

[25] B. H. Fleury, X. Yin, K. G. Rohbrandt, P. Jourdan, and A. Stucki, "Performance of a high-resolution scheme for joint estimation of delay and bidirection dispersion in the radio channel," in *Proc. IEEE 55th Vehicular Technology Conf. (VTC)*, May 2002, vol. 1, pp. 522–526.

[26] X. Yin, B. H. Fleury, P. Jourdan, and A. Stucki, "Doppler frequency estimation for channel sounding using switched multiple-element transmit and receive antennas," in *Proc. IEEE GLOBECOM*, Dec. 2003, vol. 4, pp. 2177–2181.

[27] W. C. Jakes, Ed., Microwave Mobile Communications New York, IEEE Press, 1994.

[28] D. S. Shiu*, Wireless Communications Using Dual Antenna Arrays* Hingham, MA: Kluwer Academic, 1999.

[29] R. Fletcher*, Practical Methods of Optimization—Unconstrained Optimization*. Chichester, U.K.: Wiley, 1981, vol. 1.

[30] ——*, Practical Methods of Optimization—Constrained Optimization*. Chichester, U.K.: Wiley, 1981, vol. 2.

[31] C. B. Ribeiro, E. Ollila, and V. Koivunen, "Propagation parameter estimation in MIMO systems using mixture of angular distributions model," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Mar. 2005, vol. 4, pp. 885–888.

[32] J. Wallace, H.Özcelik, M. Herdin, E. Bonek, and M. Jensen, "A diffuse multipath spectrum estimation technique for directional channel modeling," in *Proc. IEEE Int. Conf. Communications*, Jun. 2004, vol. 6, pp. 3183–3187.

[33] A. Richter, M. Landmann, and R. S. Thomä, "RIMAX—A flexible algorithm for channel parameter estimation from channel sounding measurements," in *COST 273*, Athens, Greece, 2004.

[34] J. G. Proakis*, Digital Communications*, 3rd ed. New York: McGraw-Hill Int., 1995.

[35] H. V. Poor*, An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer, 1994.

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

[37] S. Glisic and B. Vucetic*, Spread Spectrum CDMA Systems for Wireless Communications*. Boston, MA: Artech House, 1997.

[38] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Sel. Areas Commun.*, vol. 20, pp. 1211–1226, Aug. 2002.

[39] W. Weichselberger, "Spatial structure of multiple antenna radio channels—A signal processing viewpoint," Ph.D. dissertation, Fakultät für Elektrotechnik und Informationstechnik, Technische Universtat Wien, Vienna, Austria, 2003.

[40] G. H. Golub and C. F. V. Loan*, Matrix Computations*. Baltimore, MD: The Johns Hopkins Univ. Press, 1989.

[41] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul. 1996.

[42] J. R. Magnus and H. Neudecker*, Matrix Differential Calculus with Applications in Statistics and Econometrics*. New York: Wiley, 1988.

Cássio B. Ribeiro (S'00–A'01) was born in Brazil in 1977. He received the Electronics Engineering degree from the Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil, in 2000 and the M.Sc. degree in electrical engineering from COPPE/UFRJ in 2002. He is currently working towards the D.Sc. degree at UFRJ and Helsinki University of Technology, Helsinki, Finland.

His current research interests are in statistical signal processing and parameter estimation methods.

Esa Ollila (M'03) was born in Finland in 1974. He received the M.Sc. degree in mathematics from the University of Oulu, Finland, in 1998 and the Ph.D. degree in statistics with honors from the University of Jyväskylä, Finland, in 2002.

He joined the Signal Processing Laboratory at Helsinki University of Technology, Finland, in 2002, where he is currently a Senior Researcher. His current research interests focus on statistical signal processing and robust and nonparametric statistical methods.

Visa Koivunen (S'87–M'87–SM'98) received the D.Sc. (Tech.) degree with honors from the Department of Electrical Engineering, University of Oulu, Finland.

From 1992 to 1995, he was a visiting Researcher at the University of Pennsylvania, Philadelphia. In 1996, he held a faculty position at the Department of Electrical Engineering, University of Oulu. From August 1997 to August 1999, he was an Associate Professor at the Signal Processing Laboratory, Tampere University of Technology, Finland. Since 1999, he has been a Professor of Signal Processing at the Department of Electrical and Communications Engineering, Helsinki University of Technology (HUT), Finland. He is one of the Principal Investigators in SMARAD Center of Excellence in Radio and Communications Engineering nominated by the Academy of Finland. Since 2003, he has been also Adjunct Professor at the University of Pennsylvania. His research interests include statistical, communications, and sensor array signal processing. He has published more than 180 papers in international scientific conferences and journals.

Dr. Koivunen has served as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS. He is a member of the editorial board for the *Signal Processing* journal. He is also a member of the IEEE Signal Processing for Communication Technical Committee (SPCOM-TC).