

Copyright©2005 IEEE. Reprinted from:

C. B. Ribeiro, A. Richter, and V. Koivunen, "Stochastic maximum likelihood estimation of angle- and delay-domain propagation parameters," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Sep. 2005.

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the Helsinki University of Technology's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to [pubs-permissions@ieee.org](mailto:pubs-permissions@ieee.org).

By choosing to view this material, you agree to all provisions of the copyright laws protecting it.

# Stochastic Maximum Likelihood Estimation of Angle- and Delay-Domain Propagation Parameters

Cássio B. Ribeiro, Andreas Richter, Visa Koivunen  
Helsinki University of Technology (TKK)  
Signal Processing Laboratory - SMARAD Center of Excellence

**Abstract**—In this paper we derive an estimator for both time-delay and angular channel propagation parameters of the diffuse scattering component that is frequently observed in channel sounding measurements. The joint angular-delay model leads to correlation matrix with high dimensionality, which prevents direct implementation of a maximum-likelihood (ML) estimator using finite precision arithmetics and finite memory resources. We derive low complexity methods for computing the ML estimates that exploit the structure of the covariance matrices. The estimator is based on a two step procedure: first, the parameters of the power delay profile are estimated, as well as measurement noise power. Then, using the estimated time-delay parameters, the parameters of the angular distributions are estimated. We present simulation results and compare the estimated time-delay and angular distributions to the actual distributions, showing that high precision estimates are obtained.

## I. INTRODUCTION

In radio propagation it is usual to classify the signals that reach the receiver as been originated by specular reflections or scattering. The specular components usually carry most of the power, and are modeled by a relatively large number of deterministic signals with unknown parameters [1]. Scattering is frequently regarded as noise and neglected. However, even though each scattered wave arrives with low power, the overall sum of scattering components can be significant, and even dominant, especially in non line-of-sight (NLOS) situations. This behavior has been observed in measurement campaigns such as [2]. Also, in capacity studies of MIMO systems this NLOS component is of high importance, and hence it is necessary to derive estimation methods that are suitable for the stochastic nature of this component.

Deterministic techniques for propagation parameter estimation [1] commonly employ models with large number of discrete waves. This approach leads to maximization of highly non-linear likelihood functions with many local optima, which causes 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, it has been observed in [3] that, in case of diffuse scattering, deterministic estimation techniques using the discrete ray model may lead to some undesired artifacts. In [4], [5], a scheme has been proposed for estimation of time-delay domain behavior of this diffuse scattering component.

Typically isotropic scattering is employed in channel models [6]. The channel model considered in this work is suitable for non-isotropic scattering model, i.e., nonuniform distribution of angles of arrival. The model stems from the MIMO channel correlation model presented in [7], but it is also equivalent to the channel model obtained in [8] for the system setup

being considered. The MIMO channel matrix may be described analytically as a function of the parameters of the underlying random processes. In [9], this model has been used for estimation of angular channel propagation parameters, but no information from time-delay domain channel propagation parameters is used by the estimator.

In this paper we derive an estimator for the diffuse scattering component that estimates both time-delay and angular channel propagation parameters. The power-delay profile is modeled using an exponential distribution, which is typically observed in measurement campaigns. The power angular profile is modeled using a mixture of Von Mises distributions. The mixture distributions are employed in order to estimate the propagation parameters in scattering environments with multiple clusters of scatterers with high fidelity [9]. The estimation procedure is divided in two steps. We propose a two step procedure to estimate the power-delay profile and power-angular profile. First, the parameters of the exponential distribution in time-delay domain are estimated, as well as measurement noise power. Then, using the estimated time-delay parameters, the parameters of the angular distributions in angular domain are estimated.

The joint angular-delay model leads to correlation matrix with high dimensionality, which prevents direct implementation of a maximum-likelihood (ML) estimator using finite precision arithmetics and finite memory resources. We derive computationally efficient structures for computation of the ML estimates that take full advantage of the structured covariance matrices.

This paper is organized as follows: in Section II we describe the signal model used in this article. In Section III the technique for parameter estimation is described. Finally, in Section IV we present simulation results and compare the estimated time-delay and angular distributions to the actual distributions.

## II. SIGNAL MODEL

Assuming  $M_r$  antennas at the receiver and  $M_t$  antennas at the transmitter, the signal at the receiver is given by

$$\mathbf{Y}(t) = \mathbf{H}w(t) * u(t) + \mathbf{N}(t), \quad (1)$$

where  $\mathbf{H}$  is the  $M_r \times M_t$  spatial channel matrix,  $w(t)$  is the channel impulse response,  $u(t)$  is the transmitted signal and  $\mathbf{N}(t)$  is zero-mean complex Gaussian noise. We can write

$$\mathbf{y}(t) = \mathbf{h}w(t) * u(t) + \mathbf{n}(t), \quad (2)$$

where  $\mathbf{y}(t) = \text{vec}(\mathbf{Y}(t))$ ,  $\mathbf{h} = \text{vec}(\mathbf{H})$ , and  $\mathbf{n}(t) = \text{vec}(\mathbf{N}(t))$ .

We can avoid the complexity due to the convolution between the channel impulse response and the transmitted signal by expressing  $\text{vec}(\mathbf{Y}(t))$  in the frequency domain. In frequency domain the signal at the receiver can be expressed as

$$\mathbf{y}(f) = \mathbf{h}w(f)\mathbf{u}(f) + \mathbf{n}(f). \quad (3)$$

We assume that the excitation signal  $\mathbf{u}(f)$  is a multi-carrier spread spectrum signal (MCSST) [5], which is designed such that  $\mathbf{u}(f)$  is constant in the bandwidth of interest. Hence, the expression above can be further simplified to

$$\mathbf{y}(f) = \mathbf{h}w(f) + \mathbf{n}(f). \quad (4)$$

Let  $M_f$  be the number of observed frequency samples. We then define the  $M_o \times 1$  vector  $\mathcal{Y}$  as

$$\begin{aligned} \mathcal{Y} &= \begin{bmatrix} \mathbf{y}(0) \\ \vdots \\ \mathbf{y}(M_f - 1) \end{bmatrix} = \begin{bmatrix} \mathbf{h}w(0) + \mathbf{n}(0) \\ \vdots \\ \mathbf{h}w(M_f - 1) + \mathbf{n}(M_f - 1) \end{bmatrix} \\ &= \mathbf{w} \otimes \mathbf{h} + \bar{\mathbf{n}}, \end{aligned} \quad (5)$$

where  $\mathbf{w} = [w(0) \dots w(M_f - 1)]^T$ ,  $\bar{\mathbf{n}} = [\mathbf{n}^T(0) \dots \mathbf{n}^T(M_f - 1)]^T$ ,  $M_o = M_r M_t M_f$ , and  $\otimes$  denotes the tensor (Kronecker) product.

The channel sounding technique assumed in this work is based on time division multiplexing of each transmitter and receiver antenna, like in PropSound channel sounder [1]. This particular structure makes it possible to separate the contribution from each transmitter and receiver antenna.

The defined channel matrix represents the combination of all waves that impinge on the receiver array after being reflected by the surrounding scatterers. Deterministic maximum likelihood estimation techniques such as SAGE based [1] represent the received signal as a combination of several discrete waves. Consequently parameters from a large number of waves must be estimated. Hence, the algorithms often have convergence problems and the estimates contain artifacts due to local minima in likelihood function.

The following assumptions are employed throughout this article:

- if present, specular and line of sight components are estimated separately and removed from the signal (Rayleigh fading channel);
- the received signal  $\mathcal{Y}$  is a zero-mean complex temporally white circular Gaussian process;
- the channel is constant during one measurement cycle, and we assume  $E[\mathbf{h}] = 0$  and  $E[\mathbf{w}] = 0$ .
- the additive noise  $\bar{\mathbf{n}}$  is a white zero-mean circular complex Gaussian process with known covariance matrix,  $\mathbf{C}_n = E[\bar{\mathbf{n}}\bar{\mathbf{n}}^H]$ , and independent of  $\mathbf{w} \otimes \mathbf{h}$ .

From assumption (b), the PDF of the received signal  $\mathcal{Y}$  is completely characterized by its  $M_o \times M_o$  covariance matrix

$$\begin{aligned} \mathbf{C}_y &= E[\mathcal{Y}\mathcal{Y}^H] = E[(\mathbf{w} \otimes \mathbf{h} + \bar{\mathbf{n}})(\mathbf{w} \otimes \mathbf{h} + \bar{\mathbf{n}})^H] \\ &= E[\mathbf{w}\mathbf{w}^H] \otimes E[\mathbf{h}\mathbf{h}^H] + E[\bar{\mathbf{n}}\bar{\mathbf{n}}^H] \\ &= \mathbf{C}_w \otimes \mathbf{C}_h + \sigma_n^2 \mathbf{I}, \end{aligned} \quad (6)$$

where  $\mathbf{I}$  is the  $M_o \times M_o$  identity matrix.

### A. Delay and Frequency Domain Characterization

For the delay domain we use the model in [5], which is based on the observation that the power delay profile has an exponential decay over time and a base delay which is related to the distance between the transmitter and receiver. The power delay profile for infinite bandwidth is given by

$$\psi(\tau) = E[|w(\tau)|^2] = \begin{cases} 0, & \tau < \tau'_d \\ \alpha_1/2, & \tau = \tau'_d \\ \alpha_1 e^{-B_d(\tau - \tau'_d)}, & \tau > \tau'_d \end{cases} \quad (7)$$

where  $B_d$  is the coherence bandwidth,  $\alpha_1$  denotes the maximum power, and  $\tau'_d$  is the base delay.

The related power spectrum density is given by the Fourier transform of (7) as

$$\psi(\Delta f) = \frac{\alpha_1}{\beta_d + j2\pi\Delta f} e^{-j2\pi\Delta f\tau'_d}, \quad (8)$$

where  $\beta_d = B_d/(M_f f_0)$  is the normalized coherence bandwidth, and  $f_0$  is the measurement center frequency. Let us define the sampled version of the correlation function  $\mathbf{v}(\boldsymbol{\theta}_w)$ ,  $\boldsymbol{\theta}_w = \{\alpha_1, \beta_d, \tau_d\}$ , in frequency-domain as

$$\mathbf{v}(\boldsymbol{\theta}_w) = \frac{\alpha_1}{M_f} \begin{bmatrix} 1 & \frac{e^{-j2\pi\tau_d}}{\beta_d + j\frac{2\pi}{M_f}} & \dots & \frac{e^{-j2\pi(M_f-1)\tau_d}}{\beta_d + j2\pi\frac{M_f-1}{M_f}} \end{bmatrix}, \quad (9)$$

where  $\tau_d$  is the normalized base delay.

Since the process is stationary in frequency domain, the correlation between components at different frequencies is given by

$$\Psi(f_1, f_2) = \psi(f_1 - f_2), \quad (10)$$

and hence the covariance matrix of the diffuse scattering is modeled as a Toeplitz matrix

$$\mathbf{C}_w = \text{toep}(\mathbf{v}(\boldsymbol{\theta}_w), \mathbf{v}(\boldsymbol{\theta}_w)^H). \quad (11)$$

### B. Angular Domain Characterization

Using the channel model in [7], [9] we can write  $E[\mathbf{h}\mathbf{h}^H]$  as a function of the angular parameters. It is assumed that the receiver is surrounded by a large number of local scatterers, and that the waves reflected by different scatterers arrive at the array with the same power. A similar model is obtained in [8] following a different approach, based only on the statistical properties of the received waves, and not constrained to a specific geometry of the distribution of scatterers.

For simplicity, we will consider the correlation at the receiver only, and thus we assume  $M_t = 1$ . However, the techniques can be naturally extended to double-directional estimation. The cross-correlation between any two MIMO sub-channels is given by

$$\rho_{lm} = \frac{1}{\Omega} E[h_l h_m^*] = \int_{-\pi}^{\pi} \exp(jb_{lm} \cos(\phi)) f(\phi) d\phi, \quad (12)$$

where  $\Omega$  is the path loss,  $f(\phi)$  is any angular PDF of  $\phi$ ,  $b_{lm} = j2\pi d_{lm}$ ,  $d_{lm}$  is the distance between elements  $l$  and  $m$  in the receive array, and  $h_l$  is the  $l$ -th element of  $\mathbf{h}$ .

In channel measurements it is often found that the signal is received as coming from a number of different clusters, specially in bad urban scenarios. This is equivalent to a

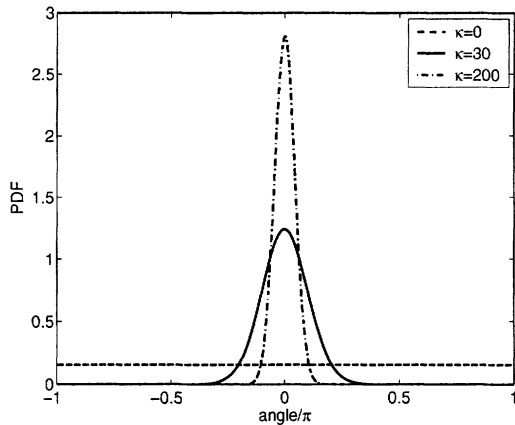


Fig. 1. Von Mises PDF for different values of  $\kappa$ , with  $\mu = 0$ .

situation where the angular PDF  $f(\phi)$  in equation (12) is a mixture of distributions,

$$f(\phi) = \sum_{p=1}^P \epsilon_p f_p(\phi), \quad (13)$$

where  $P$  is the number of clusters,  $\sum_{p=1}^P \epsilon_p = 1$ ,  $\epsilon_p$  are unknown mixture proportions, and  $f_p(\phi)$  is any valid angular PDF. An angular PDF must satisfy  $f(\theta) = f(\theta + 2\pi k) \forall k \in \mathbb{Z}$ . Hence, a Gaussian PDF, which has infinite support, is improper. The von Mises distribution [10] defined in angular domain is more appropriate. It is defined as follows:

$$f_p(\phi) = \frac{1}{2\pi I_0(\kappa)} \exp(\kappa \cos(\phi - \mu)), \quad (14)$$

where  $\mu$  is the symmetry center or “mean angle”,  $\kappa$  can be chosen between 0 (isotropic scattering) and  $\infty$  (extremely concentrated), and  $I_0(\cdot)$  is the modified Bessel function of the first kind of order zero. The representation power of the Von Mises PDF for different scattering environments is illustrated in Figure 1 for several values of  $\kappa$ . This representation assumes the antenna elements are either omni-directional or that the product of the antenna pattern and the actual angular distribution can be modeled by the mixture von Mises distributions in (13).

Using (13) and (14), the cross correlation in (12) may be written analytically as [7], [9]

$$E[h_l h_m^*] = \sum_{p=1}^P \epsilon_p \frac{I_0(\{\kappa_p^2 + b_{lm}^2 + 2\kappa_p b_{lm} \cos(\mu_p)\}^{\frac{1}{2}})}{I_0(\kappa_p)}. \quad (15)$$

### III. PARAMETER ESTIMATION

Let us denote by  $\mathcal{Y}_m$  the  $m$ -th observation of  $\mathcal{Y}$ ,  $m = 1, \dots, M_s$ . Assuming  $\mathcal{Y}$  is circular complex Gaussian and that the realizations  $\mathcal{Y}_m$  are i.i.d., we can write the likelihood function as

$$L(\mathcal{Y}_1, \dots, \mathcal{Y}_{M_s}) \propto -\log |\mathbf{C}_y| - \frac{1}{M_s} \sum_{m=1}^{M_s} \mathcal{Y}_m^H \mathbf{C}_y^{-1} \mathcal{Y}_m, \quad (16)$$

where  $M_s$  is the number of realizations or snapshots. We will also assume the noise is circular complex white Gaussian with variance  $\sigma_n^2$ .

In [5] a method is proposed that estimates  $\mathbf{C}_w$  and the noise variance assuming the input signal to be spatially white,

i.e.,  $\mathbf{C}_h = \mathbf{I}$ . The method exploits the Toeplitz structure of  $\mathbf{C}_w$  for the computation of the ML estimates, reducing the computation complexity by avoiding the direct computation of determinants and matrix inversions. Unfortunately it is not possible to directly extend the method for joint estimation of  $\mathbf{C}_w$ ,  $\mathbf{C}_h$ , and  $\mathbf{C}_n$ , since  $\mathbf{C}_y$  in (6) is not Toeplitz in general. Also, direct optimization of the likelihood function using (6) is not feasible due to the high dimensionality of the matrices involved. Typical values for  $M_f$  and  $M_r M_t$  are in the range  $M_f = [100, 2000]$ , and  $M_r M_t = [4, 64]$ , but higher values can be used. This leads to  $\mathbf{C}_y$  ranging from  $400 \times 400$  to  $128000 \times 128000$ , or even higher.

We propose an estimation method that reduces the computational complexity by calculating the estimates in two steps:

- Optimize for the frequency-domain parameters and noise variance using the algorithm in Section III-A.
- Optimize for the angular-domain parameters using the procedure in Section III-B, with  $\mathbf{C}_w$  as calculated in the previous item.

With this two step procedure it is possible to exploit the Toeplitz structure of  $\mathbf{C}_w$  for the computation of the ML estimates. The covariance matrix to be manipulated in step (b) is  $\mathbf{C}_h$  only, which is typically much smaller than  $\mathbf{C}_w$ .

#### A. Frequency-Domain Parameters

An estimator for the frequency-domain parameters defined in Section II-A is derived in [5]. In [5] it is assumed that the channel covariance matrix has the structure  $\mathbf{C}_y = (\mathbf{C}_w + \sigma_n^2 \mathbf{I}) \otimes \mathbf{I}$ , i.e., the channel is assumed to be spatially white. This is a special case of the situation depicted in equation (6).

We will use the estimator in [5] to provide an estimate of the frequency-domain parameters prior to estimation of angular-domain parameters. Even though we expect a loss in performance due to correlation in angular domain, this loss should not be significant, since the estimator has been successfully applied to data originated from measurements, providing reasonable estimates even in (possibly) correlated environments.

Because of the whiteness assumption, the signal received from each antenna is considered as an independent realization of the diffuse scattering process. Let us define the  $M_f \times M_r M_t$  matrix  $\mathbf{X}'$  whose columns are the signals received by each antenna  $m_r$ ,  $m_r = 1, \dots, M_r$ . Since for the frequency-domain estimator the diffuse scattering is assumed spatially white, each column of  $\mathbf{X}'$  is considered as an independent realization of the random process. Assuming we have  $M_s$  independent measurements, we define the extended  $M_f \times M_r M_t M_s$  matrix  $\mathbf{X}$  whose columns are independent realizations of the diffuse scattering.

Based on the definitions above, the estimator in [5] is a maximum likelihood estimator that maximizes the log-likelihood function

$$L_w \propto -M_r M_t M_s \log(\det(\mathbf{C}_w(\theta_w) + \sigma_n^2 \mathbf{I})) - \text{tr}(\mathbf{X}^H (\mathbf{C}_w(\theta_w) + \sigma_n^2 \mathbf{I})^{-1} \mathbf{X}) \quad (17)$$

#### B. Angular-Domain Parameters

Assuming the number of mixture components in angular domain,  $P$ , is known, the angular parameters are the parameters of the mixture of von Mises distributions:  $\theta_h =$

$\{\phi_1, \kappa_1, \epsilon_1, \dots, \phi_P, \kappa_P, \epsilon_P\}$ ,  $p = 1, \dots, P$ , with  $\sum_{p=1}^P \epsilon_P = 1$ . Due to the model in (6), the path loss  $\Omega$  is already estimated as part of the delay-domain parameters, and hence we use the constraint  $\Omega = 1$ .

For the angular-domain parameters we follow an approach similar to [9], but using the frequency-domain parameters calculated previously. This division in two steps simplifies the optimization procedure, since the parameter space is reduced, but we still have not solved the problem of calculating the determinant and inverse of  $\mathbf{C}_y$  at every iteration.

These computations can be simplified by writing  $\mathbf{C}_y$  as a function of its eigenvalues and eigenvectors. Let us define the matrix  $\mathbf{V}_y$  whose columns are the eigenvectors of  $\mathbf{C}_y$ , and the matrix  $\Lambda_y$  containing the eigenvalues of  $\mathbf{C}_y$  in its main diagonal. We can write

$$\mathbf{C}_y = \mathbf{V}_y \Lambda_y \mathbf{V}_y^H, \quad (18)$$

where we have used the fact that the eigenvectors of a Hermitian matrix are orthogonal. Similarly, we can also define  $\mathbf{V}_w$ ,  $\Lambda_w$ ,  $\mathbf{V}_h$ , and  $\Lambda_h$ , such that

$$\mathbf{C}_w = \mathbf{V}_w \Lambda_w \mathbf{V}_w^H \quad (19)$$

$$\mathbf{C}_h = \mathbf{V}_h \Lambda_h \mathbf{V}_h^H. \quad (20)$$

Substituting (19) and (20) in (6) yields

$$\begin{aligned} \mathbf{C}_y &= (\mathbf{V}_w \Lambda_w \mathbf{V}_w^H) \otimes (\mathbf{V}_h \Lambda_h \mathbf{V}_h^H) + \sigma_n^2 \mathbf{I} \\ &= (\mathbf{V}_w \otimes \mathbf{V}_h) (\Lambda_w \otimes \Lambda_h + \sigma_n^2 \mathbf{I}) (\mathbf{V}_w^H \otimes \mathbf{V}_h^H). \end{aligned} \quad (21)$$

Comparing (21) with (18) we conclude that

$$\mathbf{V}_y = \mathbf{V}_w \otimes \mathbf{V}_h \quad (22)$$

$$\Lambda_y = \Lambda_w \otimes \Lambda_h + \sigma_n^2 \mathbf{I}. \quad (23)$$

We can exploit the Kronecker structure of the eigenvalues and eigenvectors of  $\mathbf{C}_y$  to simplify the optimization procedure. This allows us to compute only the eigenvalues and eigenvectors of  $\mathbf{C}_w$  and  $\mathbf{C}_h$ , and then obtain  $\mathbf{V}_y$  and  $\Lambda_y$ . The logarithm of the determinant of  $\mathbf{C}_y$  can now be calculated as

$$\begin{aligned} \log |\mathbf{C}_y| &= \log (|\mathbf{V}_w \otimes \mathbf{V}_h| |\Lambda_w \otimes \Lambda_h + \sigma_n^2 \mathbf{I}| |\mathbf{V}_w^H \otimes \mathbf{V}_h^H|) \\ &= \log \prod_{j=1}^{M_o} (\lambda_w \otimes \lambda_h + \sigma_n^2 \mathbf{1}_{M_o})_{\{j\}} \\ &= \sum_{j=1}^{M_o} \log ((\lambda_w \otimes \lambda_h + \sigma_n^2 \mathbf{1}_{M_o})_{\{j\}}), \end{aligned} \quad (24)$$

where  $\lambda_w$  and  $\lambda_h$  are vectors containing the eigenvalues of  $\mathbf{C}_w$  and  $\mathbf{C}_h$ , respectively,  $\mathbf{1}_{M_o}$  is a  $M_o \times 1$  vector whose entries are equal to 1, and  $(\cdot)_{\{j\}}$  denotes the  $j$ -th element of  $(\cdot)$ . It is clear that the computational complexity of calculating the determinant is reduced, but another important observation is that the exchange of the order in which the log is computed allows for easier implementation with finite precision, since the eigenvalues can assume very low values, specially for  $\lambda_w$  due to the large number of frequency samples.

The computation of  $\mathbf{C}_y^{-1}$  can also be simplified using

$$\begin{aligned} \mathbf{C}_y^{-1} &= [(\mathbf{V}_w \otimes \mathbf{V}_h) (\Lambda_w \otimes \Lambda_h + \sigma_n^2 \mathbf{I}) (\mathbf{V}_w^H \otimes \mathbf{V}_h^H)]^{-1} \\ &= (\mathbf{V}_w \otimes \mathbf{V}_h) (\Lambda_w \otimes \Lambda_h + \sigma_n^2 \mathbf{I})^{-1} (\mathbf{V}_w^H \otimes \mathbf{V}_h^H). \end{aligned} \quad (25)$$

Further simplifications are possible if we take into account that  $\mathbf{C}_w$  is fixed while optimizing for  $\mathbf{C}_h$ . Clearly, it is not necessary to calculate  $\mathbf{V}_w^H \otimes \mathbf{V}_h^H$  and the multiplication of the resulting  $M_o \times M_o$  matrix by  $\mathcal{Y}_m$  at every iteration of (16), since  $\mathbf{V}_w$  and  $\mathcal{Y}_m$  are fixed. In order to simplify the problem, we will define the transformed signal

$$\bar{\mathcal{Y}}_m = (\Lambda_w^{-1/2} \mathbf{V}_w^H \otimes \mathbf{I}_{M_r M_t}) \mathcal{Y}_m. \quad (26)$$

The covariance matrix of  $\bar{\mathcal{Y}}_m$  is given by

$$\begin{aligned} \mathbf{C}_{\bar{y}} &= (\Lambda_w^{-1/2} \mathbf{V}_w^H \otimes \mathbf{I}_{M_r M_t}) E[\mathcal{Y}_m \mathcal{Y}_m^H] (\mathbf{V}_w \Lambda_w^{-1/2} \otimes \mathbf{I}_{M_r M_t}) \\ &= (\mathbf{I}_{M_f} \otimes \mathbf{V}_h) (\mathbf{I}_{M_f} \otimes \Lambda_h + \sigma_n^2 \Lambda_w^{-1} \otimes \mathbf{I}_{M_r M_t}) \\ &\quad \cdot (\mathbf{I}_{M_f} \otimes \mathbf{V}_h^H). \end{aligned} \quad (27)$$

The covariance matrix  $\mathbf{C}_{\bar{y}}$  is block diagonal, implying that blocks of  $M_r$  elements of  $\bar{\mathcal{Y}}_m$  are uncorrelated. Also, there is no Kronecker product between the eigenvalues of  $\mathbf{C}_w$  and  $\mathbf{C}_h$ , what simplifies numerical implementations. Consequently, we will estimate the angular parameters using  $\bar{\mathcal{Y}}_m$  instead of  $\mathcal{Y}_m$ . The likelihood function for  $\bar{\mathcal{Y}}_m$  is given by

$$L(\bar{\mathcal{Y}}_1, \dots, \bar{\mathcal{Y}}_{M_s}) \propto -\log |\mathbf{C}_{\bar{y}}| - \frac{1}{M_s} \sum_{m=1}^{M_s} \bar{\mathcal{Y}}_m^H \mathbf{C}_{\bar{y}}^{-1} \bar{\mathcal{Y}}_m. \quad (28)$$

Let us define the  $M_o \times M_o$  diagonal matrix  $\Lambda$  as

$$\Lambda = (\mathbf{I}_{M_f} \otimes \Lambda_h + \sigma_n^2 \Lambda_w^{-1} \otimes \mathbf{I}_{M_r M_t}), \quad (29)$$

and the  $M_o \times 1$  vector  $\lambda = \text{diag}(\Lambda)$ . Now we can write the likelihood function in (28) as

$$\begin{aligned} L(\bar{\mathcal{Y}}_1, \dots, \bar{\mathcal{Y}}_{M_s}) &\propto -\sum_{j=1}^{M_o} \log \lambda_j \\ &\quad - \frac{1}{M_s} \sum_{m=1}^{M_s} \bar{\mathcal{Y}}_m^H (\mathbf{I}_{M_f} \otimes \mathbf{V}_h) \Lambda^{-1} \\ &\quad \cdot (\mathbf{I}_{M_f} \otimes \mathbf{V}_h^H) \bar{\mathcal{Y}}_m. \end{aligned} \quad (30)$$

Let us define the  $M_o \times 1$  vector  $\mathcal{X}_m$  as

$$\mathcal{X}_m = (\mathbf{I}_{M_f} \otimes \mathbf{V}_h^H) \bar{\mathcal{Y}}_m. \quad (31)$$

A computationally efficient form for computation of  $\mathcal{X}_m$  is given by

$$\mathcal{X}_m = \text{vec}(\mathbf{V}_h^H \text{mat}(\bar{\mathcal{Y}}_m, M_r M_t, M_f)), \quad (32)$$

where the mat operator reshapes a vector into a matrix:

$$\text{mat} \left( \begin{bmatrix} \mathbf{a}_1^{m \times 1} \\ \vdots \\ \mathbf{a}_N^{m \times 1} \end{bmatrix}, M, N \right) = [\mathbf{a}_1^{m \times 1} \quad \dots \quad \mathbf{a}_N^{m \times 1}]. \quad (33)$$

Finally, we can write  $L(\bar{\mathcal{Y}}_1, \dots, \bar{\mathcal{Y}}_{M_s})$  in a computationally efficient form as

$$L(\bar{\mathcal{Y}}_1, \dots, \bar{\mathcal{Y}}_{M_s}) \propto -\sum_{j=1}^{M_o} \log \lambda_j - \frac{1}{M_s} \sum_{m=1}^{M_s} \mathcal{X}_m^H \Lambda^{-1} \mathcal{X}_m, \quad (34)$$

where it should be noted that the multiplication between a diagonal matrix and a vector is element-wise multiplication of

TABLE I

REDUCED COMPLEXITY ALGORITHM FOR ESTIMATION OF FREQUENCY- AND ANGULAR-DOMAIN PARAMETERS

1)	Estimate frequency-domain parameters, $\theta_w$ , using method in Section III-A.
2)	Compute $\Lambda_w(\theta_w) = \mathbf{F}^H \mathbf{R}_w \mathbf{F}$ using FFT.
3)	Compute $\bar{\mathcal{Y}}_m$ using (26) and FFT.
4)	Optimize for angular-domain parameters, $(\theta_h)$ , using (34). For each iteration compute:
a)	$\mathbf{V}_h$ and $\Lambda_h$ .
b)	$\mathcal{X}$ using (32).

the main diagonal and the vector. The ML estimates of  $\theta_p = \{\kappa_p, \mu_p, \epsilon_p\}$ ,  $p = 1, \dots, P$ , are those values that maximize the likelihood function in (34).

A further reduction of complexity is possible in the calculation of the eigenvalues and eigenvectors of  $\mathbf{C}_w$ , since it is a Toeplitz matrix, and hence it can be approximated by a circulant matrix. A circulant matrix can be decomposed as

$$\mathbf{R}_w = \mathbf{F} \mathbf{D}_w \mathbf{F}^H, \quad (35)$$

where  $\mathbf{F}$  is the DFT matrix and  $\mathbf{D}_w$  is a diagonal matrix with the eigenvalues of  $\mathbf{R}_w$ . Hence,  $\mathbf{V}_w = \mathbf{F}$ , and  $\Lambda_w = \mathbf{F}^H \mathbf{R}_w \mathbf{F}$ . The use of FFT reduces even further the computational complexity of  $\Lambda_w$  and  $\bar{\mathcal{Y}}_m$  in (26).

Table I summarizes the reduced complexity algorithm for estimation of frequency- and angular-domain parameters.

#### IV. SIMULATION RESULTS

In this Section some simulations are presented in order to illustrate the performance of the described optimization procedure. The receiver has an ULA with  $M_r = 8$  antennas and the transmitter uses  $M_t = 1$  antenna. The number of frequency points is  $M_f = 128$ , and  $M_s = 50$ . The received signal is generated as

$$\mathbf{y}(k) = \mathbf{R}^{1/2} \mathbf{n}_2(k) + \mathbf{n}(k), \quad (36)$$

where  $\mathbf{R}^{1/2}$  is obtained by the Cholesky decomposition of  $\mathbf{C}_y$ , and  $\mathbf{n}_2(k)$  is a circular complex white Gaussian process. For the frequency-domain parameters, typical values often observed in channel sounding experiments are used:  $\sigma_n^2 = 0.1$ ,  $\alpha_1 = 1$ ,  $\beta = 0.07$ , and  $\tau = 0.1$ . The angular-domain parameters are defined as  $\phi = \{65^\circ, 90^\circ\}$ ,  $\kappa = \{10, 150\}$ ,  $\epsilon = \{0.5, 0.5\}$ .

In Figures 2 and 3 we compare the delay and angular distributions plus noise obtained using the estimation procedure described in this article with the actual distributions, respectively. The estimator provides high-precision estimates for both the time-delay distribution and the angular distribution. The estimate of the angular distribution is compared to the output of the Bartlett beamformer, showing the gain in using a parametric approach for angular power spectrum estimation.

#### V. CONCLUSION

We propose a channel model for the diffuse scattering component incorporating both frequency- and angular-domain parameters, and derive an optimization procedure to compute the stochastic maximum likelihood estimates with low complexity. Since the estimation is done in a two-step procedure, a performance loss is expected compared to the joint optimization procedure. However, the simulation results show that high precision estimates are obtained.

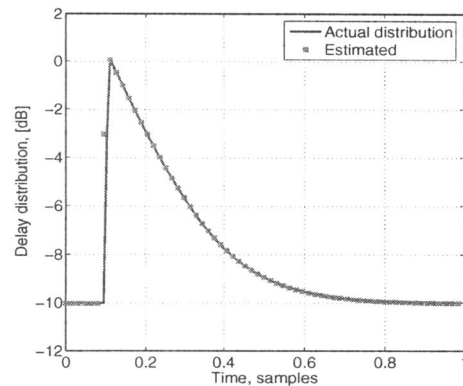


Fig. 2. Comparison of estimated delay distribution and actual distribution. The curves overlap almost perfectly.

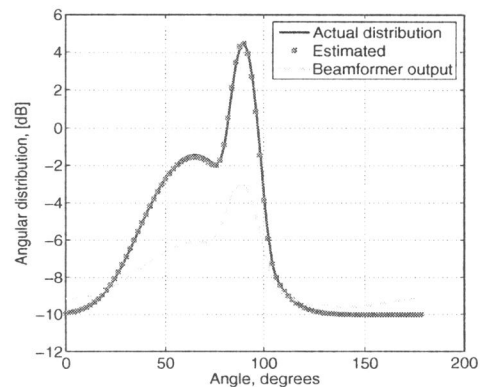


Fig. 3. Comparison of estimated angular distribution and actual distribution. The curves overlap almost perfectly. Also shown is the output of the Bartlett beamformer.

#### REFERENCES

- [1] B. H. Fleury, P. Jourdan, and A. Stucki, "High-resolution channel parameter estimation for MIMO applications using the SAGE algorithm," in *2002 International Zurich Seminar on BroadBand Communications. Access, Transmission, Networking*, Feb 2002, pp. 30-1 – 30-9.
- [2] 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 Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 347-360, Mar. 2000.
- [3] M. Bengtsson and B. Völcker, "On the estimation of azimuth distributions and azimuth spectra," in *IEEE VTS 54th Vehicular Technology Conference, 2001*, Oct. 2001, vol. 3, pp. 1612-1615.
- [4] A. Richter and R. Thoma, "Parametric modelling and estimation of distributed diffuse scattering components of radio channels," *COST 273 Temporary Document*, vol. 3, no. 198, Sep 2003.
- [5] A. Richter, *Estimation of Radio Channel Parameters: Models and Algorithms*, Ph.D. thesis, Technische Universität Ilmenau, 2005.
- [6] 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 Transactions on Communications*, vol. 48, no. 3, pp. 502-513, Mar. 2000.
- [7] A. Abdi and M. Kaveh, "A space-time correlation model for multielement antenna systems in mobile fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 3, pp. 550-560, April 2002.
- [8] J. W. Wallace and M. Jensen, "Modeling the indoor mimo wireless channel," *IEEE Transactions on Antennas and Propagation*, vol. 50, no. 5, pp. 591-599, May 2002.
- [9] C. B. Ribeiro, E. Ollila, and V. Koivunen, "Propagation parameter estimation in mimo systems using mixture of angular distributions model," in *Proc. IEEE Conference on Acoustics, Speech and Signal Processing*, Mar. 2005.
- [10] K. V. Mardia, *Statistics of Directional Data*, Academic Press, London and New York, 1972.