

Copyright©2004 IEEE. Reprinted from:

C. B. Ribeiro, E. Ollila, and V. Koivunen. “Stochastic maximum likelihood method for propagation parameter estimation,” in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Sep. 2004.

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.

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

STOCHASTIC MAXIMUM LIKELIHOOD METHOD FOR PROPAGATION PARAMETER ESTIMATION

Cássio B. Ribeiro^{1,2}, Esa Ollila¹, Visa Koivunen¹

¹Helsinki University of Technology, P.O. Box 3000, FIN-02015 HUT, Finland.

²Federal University of Rio de Janeiro, P.O. Box 68504, 21945-970, Rio de Janeiro, RJ, Brazil.

Abstract - In this paper we will derive a stochastic maximum likelihood method for estimating spatio-temporal channel parameters. 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 angular Von Mises distribution model which is appropriate for directional data typically observed in channel measurement campaigns. The signal model is stochastic. The performance of the proposed method is compared to SAGE algorithm where the signal model is deterministic. The computational complexity of the proposed method is lower and channel parameters are estimated with higher fidelity because the underlying distribution model is well-suited for directional data.

Keywords - Channel sounding, maximum likelihood, directional data.

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 transceivers structures and designing networks for future wireless systems with high spectral efficiency. For example, in the process of developing Multiple-Input Multiple-output (MIMO) communication systems, channel sounding using multiple transmit and receive antennas and subsequent estimation of the channel propagation parameters are needed.

In this paper we are interested in estimating the parameters of MIMO channels via channel sounding. The channel sounder model used in this study is based on [1–3]. It uses a time division approach to scan the outputs of the antennas.

We adopt the MIMO channel covariance matrix model presented in [4]. This matrix may be described analytically as a function of the parameters of the underlying random processes, and thus can be used to derive estimators for the channel parameters. In this work we will derive a stochastic maximum likelihood algorithm for estimation of the spatio-temporal behavior of MIMO channels. We will model the impinging waves as realizations of an underlying random process that characterizes the MIMO channel matrix. Hence, our estimation problem becomes the problem of estimating only the parameters of the distribution of this random process. Von Mises distribution model (see, [5]) is used because it is well suited for angular data. This approach leads to lower complexity and faster convergence in comparison to deterministic

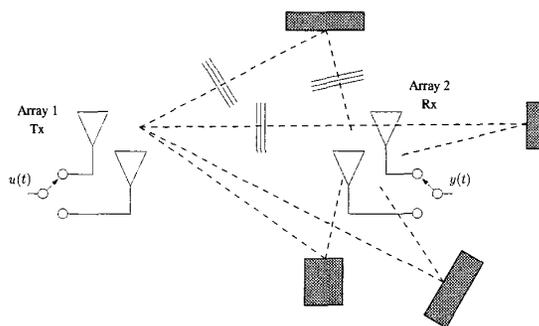


Fig. 1. Channel sounding environment.

model based on Space Alternating Generalized Expectation-maximization (SAGE) algorithm [1,6]. These benefits are due to the smoother likelihood function and the reduced number of unknowns.

This paper is organized as follows. The signal model is presented in Section II. The channel sounding principle is explained in Section III. The proposed channel parameter estimation method is developed in Section IV. Simulation results showing that the proposed method estimates the channel parameters with high fidelity are presented in Section V. The results are compared to the SAGE algorithm in [1–3], and the proposed method is shown to outperform SAGE both in the root mean squared error of the estimates and in computational complexity.

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 [7], and has been used, e.g., in [4] to study the effect of fading correlation on the capacity of MIMO fading channels. Figure 1 illustrates the propagation environment.

Let us define the $N \times 1$ vector $\mathbf{y}(k)$ with the output of the each antenna at the receiver, the $M \times 1$ vector $\mathbf{u}(k)$ of the transmitted sequences, the $N \times M$ MIMO channel matrix $\mathbf{H}_W(k)$, and the $N \times 1$ vector $\mathbf{n}(k)$ with random noise. The signals thus defined 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 can model the received signal as

$$\mathbf{y}(k) = \mathbf{H}_W(k) * \mathbf{u}(k) + \mathbf{n}(k), \quad (1)$$

where $*$ denotes convolution. This corresponds to a frequency selective, or wideband, channel [7], as observed in previous measurement campaigns for typical urban scenarios [6]. We assume that $\mathbf{H}_W(k)$ can be represented as a combination of MIMO channel matrix that depends on the spatial structure of the channel (and possibly on time) and a complex scalar that is dependent on time only, i.e., $\mathbf{H}_W(k) = \mathbf{H}(k)\alpha(k)$. If we also assume the delay spread of the channel to be finite, we can rewrite (1) as

$$\mathbf{y}(k) = \sum_{l=0}^{L-1} \mathbf{H}(l)\alpha(l)\mathbf{u}(k-l) + \mathbf{n}(k), \quad (2)$$

where L is the duration of the channel impulse response.

The defined channel matrix represents the combination of all waves that impinge the receiver array after being reflected by the surrounding scatterers. In [1–3, 6] the channel is modeled using a different approach, where the received signal is described as a combination of several waves with unknown parameters together with random noise. The received signal is then written as a function of the wave parameters, and estimation techniques such as SAGE can be used to yield estimates of the parameters. However, since the parameters from several waves must be estimated, a search in a very high dimensional space must be performed, and thus the algorithms often present convergence problems due to large number of local minima in the likelihood function. Consequently, we choose to use a more compact stochastic model for the parameters.

III. CHANNEL SOUNDING TECHNIQUE

We consider the same channel sounding technique used in [1–3, 6], where the radio channel sounder is assumed to be equipped with switches at both transmit and receive endings, as illustrated in Figure 2.

A. Sounding signal

The sounding signal is applied by means of Switch 1 during a period T_t at the input of each element of the transmit array. The time Switch 1 takes to switch all transmit antennas is called the sounding cycle and denoted by $T_{cy} \geq MT_t$. The total duration of the sounding signal is IT_{cy} , where I is the number of times the sounding cycle is repeated during one measurement shot.

The sounding signal $u(k)$ with power P_u is periodic with period T_t . Each period of the sounding signal can be described as a periodically repeated burst $u_s(k)$ of duration T_u and a cyclic prefix of duration T_g , i.e.,

$$u(k) = \sum_{p=-\infty}^{\infty} u_s(k - pN_u), \quad -N_g \leq k < N_t - N_g, \quad (3)$$

where $N_u = T_u/T_s$, $N_g = T_g/T_s$, and $N_t = T_t/T_s$. The cyclic prefix is transmitted during the guard time T_g , which must be greater than the maximum propagation delay in order to ensure that the signals from different antennas in Array 1 do not interfere with each other in the receiver.

B. Scanning of the receiver array outputs

At the receiver side the switch is activated as depicted in Figure 2. The outputs of Array 2 are successively scanned during a period T_{sc} . For each period T_t , corresponding to the transmission of the sounding signal by one element of Array 1, one sensing cycle is performed, during which all elements in Array 2 are scanned once. In a practical implementation, a guard interval between two consecutive scans should be considered to embody switching time and all transient effects, but these effects are not considered in this work.

If the channel can be considered to be constant during one measurement cycle, the outputs of the channel sounder can be rearranged as in (1) by defining the n -th element of $\mathbf{y}(k)$ as the output of the n -th antenna for the time during which it is active. The sensing windows also have to be synchronized with the transmit windows as in Figure 2, and the guard time has to be properly defined, i.e., $T_g \geq LT_s$. If the channel is assumed to be time varying, then it is still possible to make a description similar to (1), but the elements of the channel matrix have to be properly defined in order to compensate for the time differences between the scanning windows.

IV. PARAMETER ESTIMATION

For the derivation of the estimation method we will consider the transmission from one transmitter antenna at a time, as implemented in the channel sounder in Figure 2. However, the estimation method applies to MIMO systems in general, since after one cycle all transmit antennas have been switched, leading to an M -transmit, N -receive MIMO system. If one transmit antenna is considered at the time, the channel matrix $\mathbf{H}(k)$ in (2) is a $N \times 1$ vector, henceforth denoted by lowercase $\mathbf{h}(k)$, and the transmitted signal $u(k)$ is a scalar. We will assume that the channel is constant during one measurement cycle, in which case the signal model can be simplified to

$$\mathbf{y}(k) = \mathbf{h} \sum_{l=0}^{L-1} \alpha(l)u(k-l) + \mathbf{n}(k). \quad (4)$$

We will assume that $\mathbf{y}(k)$ is a zero-mean complex Gaussian process. After removing the constant terms not dependent on the signal, the log-likelihood function can be rewritten as

$$\lambda\{\mathbf{Y}(k)\} \propto -\log |\mathbf{C}_y| - \text{tr}\{\mathbf{C}_y^{-1}\widehat{\mathbf{C}}_y\}, \quad (5)$$

where \mathbf{C}_y is the covariance matrix of the received signal, $\mathbf{C}_y = E[\mathbf{y}(k)\mathbf{y}^H(k)]$, $\widehat{\mathbf{C}}_y$ is the sample covariance matrix, defined as $\widehat{\mathbf{C}}_y = \frac{1}{N_u} \sum_{k=0}^{N_u-1} \mathbf{y}(k)\mathbf{y}^H(k)$, and $\text{tr}\{\cdot\}$ denotes the trace.

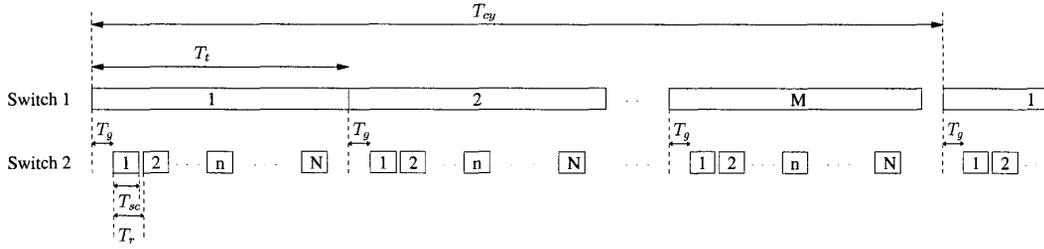


Fig. 2. Timing structure for channel sounding technique.

From (4) we can write \mathbf{C}_y as

$$\mathbf{C}_y = \sum_{l=0}^{L-1} \sum_{l'=0}^{L-1} E[\mathbf{h}\mathbf{h}^H \alpha(l)\alpha^*(l')] u(k-l)u^*(k-l') + \mathbf{C}_n, \quad (6)$$

where \mathbf{C}_n is the noise covariance matrix. Noise is assumed to be zero mean and independent of \mathbf{h} . The expression in (6) can be further simplified using the assumption that the complex gains $\alpha(l)$ and $\alpha^*(l')$ are uncorrelated for $l \neq l'$, and independent of \mathbf{h} . Thus,

$$\mathbf{C}_y = E[\mathbf{h}\mathbf{h}^H] E\left[\sum_{l=0}^{L-1} |\alpha(l)|^2 |u(k-l)|^2\right] + \mathbf{C}_n \quad (7)$$

As described in Section III, the sounding sequence is preceded by a time-guard interval whose duration is greater than or equal to the channel delay spread. Hence, for $k \geq 0$ we can approximate the convolution term in (7) as $E[\sum_{l=0}^{L-1} |\alpha(l)|^2 |u(k-l)|^2] \approx P_u$, and finally rewrite (7) as

$$\mathbf{C}_y = P_u E[\mathbf{h}\mathbf{h}^H] + \mathbf{C}_n \quad (8)$$

Using the channel model in [4] we can write $E[\mathbf{h}\mathbf{h}^H]$ as a function of the channel 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. This situation can be represented as a ring of scatterers around the receiver, as depicted in Figure 3 for any 2 antennas at the transmitter and receiver.

Assuming that the angle spread at the transmitter is small and $D \gg R \gg d_{lm}$, it is possible to show that the cross-correlation between any two sub-channels l and m is given by [4]

$$E[h_l h_m^*] = \int_{-\pi}^{\pi} \exp(b_{lm} \cos(\phi)) f(\phi) d\phi, \quad (9)$$

where $f(\phi)$ is any angular PDF of ϕ , $b_{lm} = j2\pi d_{lm}$, and h_l is the l -th element of \mathbf{h} . Equation (9) is a simplification from the more general expression derived in [4], obtained with the assumption that only one antenna is transmitting at a time. A suitable angular PDF is the Von Mises [5], defined as

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

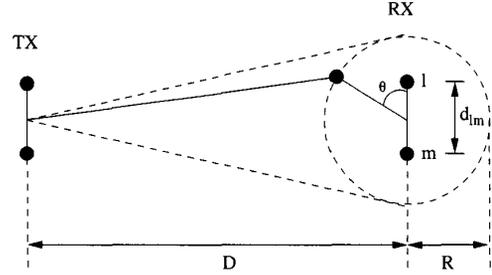


Fig. 3. Illustration of the geometrical configuration of a 2x2 channel with local scatterers at the receiver, where D is the distance between the transmitter and receiver arrays, R is the radius of the ring of scatterers around the receiver, and d_{lm} is the distance between elements l and m in the receive array.

where μ is the mean angle and κ can be chosen between 0 (isotropic scattering) and ∞ (extremely concentrated). Using the Von Mises PDF the cross correlation in (9) may be written as [4]

$$E[h_l h_m^*] = \frac{1}{I_0(\kappa)} I_0(\{\kappa^2 + b_{lm}^2 + 2\kappa[b_{lm} \cos(\mu)]\}^{\frac{1}{2}}). \quad (11)$$

We can now find the values for μ and κ that maximize the log-likelihood function in (5), i.e.,

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

where $\operatorname{tr}\{\cdot\}$ denotes the trace. In order to optimize (12), we will use the Nelder-Mead simplex algorithm [8], as implemented by the `fminsearch` function of Matlab, but any maximization procedure can be used.

We are also interested in the characterization of the behavior of the channel with respect to time. According to [6], the sequence $|\alpha(l)|^2$, $l = 0, \dots, L$, can be approximated by an exponential distribution, i.e.,

$$|\alpha(l)|^2 \propto \exp(-l/\sigma_d). \quad (13)$$

Thus, we only have to estimate the time of arrival (TOA) of the cluster and the decay factor σ_d of the exponential. Using the properties of the M-sequence used for transmission we can derive the following algorithm for estimation of the time-dependent parameters:

1) Calculate

$$\beta(l) = \sum_{i=1}^I \sum_{n=1}^N \left| \sum_{k=0}^{Nu-1} u(k-l)^* y_n(k+(i-1)T_u) \right|^2,$$

where $y_n(k)$ is the n -th element of $\mathbf{y}(k)$.

2) Due to the noisy nature of $\beta(l)$, it is not convenient in practice to set the estimate of the TOA as the sample corresponding to the highest peak in $\beta(l)$. Instead, we take the first sample \hat{k} such that $\beta(\hat{k}) \geq \epsilon$, where ϵ is a fixed threshold.

3) Calculate the decay factor using the sequence $\beta(l)$, $l = \hat{k}, \dots, Nu - 1$.

Although the optimization of (12) involves the computation of determinants and matrix inversions, in practice it is less complex than methods like the SAGE algorithm in [1–3]. This happens because in [1–3] a big number of discrete waves must be estimated, each wave described by 3 to 6 parameters (depending on the model), leading to several optimizations. In the proposed method, however, only one search over a 2-dimensional parameter space has to be carried out. The estimation of (13) is similar to one step of the initialization procedure used in the SAGE, and thus do not increase complexity.

V. SIMULATION RESULTS

The propagation channel is simulated as a superposition of $N_w = 10000$ discrete waves defined as

$$\mathbf{x}_n(k) = \frac{1}{\sqrt{N_w}} e^{j\psi} \mathbf{c}(\phi_n) u(k - \tau_n), \quad n = 1, \dots, N_w, \quad (14)$$

where ψ is a random phase, $\mathbf{c}(\phi_n)$ is the steering vector, and τ_n is the delay of the n -th wave. The receiver has an uniform linear array (ULA) with $N = 11$ antennas. For the generation of each parameter a distribution function was selected following the conclusions from measurement campaigns in the literature [6]. The random phase, ψ , is uniformly distributed in $[0, 2\pi]$. The angle of arrival ϕ_n follows a Von Mises distribution with parameters $\mu = 75^\circ$ and $\kappa = 26.9344$, corresponding to an angular spread of 8° . The delays were chosen from an exponential distribution with decay factor $\sigma_d = 2.34\mu s$, corresponding to a $2\mu s$ delay spread.

The sounding sequence is a complex length-127 M-sequence, with symbol duration $T_p = 1/4.096 \times 10^6$ and sampling period $T_s = T_p/2$. The random noise is generated from a complex white Gaussian process. The results are obtained during one sounding cycle, i.e., $I = 1$, and are averaged over 50 runs. In Table 1 the RMSE are shown for different values of signal-to-noise ratio (SNR), and compared to the results obtained by using the SAGE algorithm. The SNR is defined relative to the signal in each antenna at the receiver. A fixed number of 10 iterations were performed for each run of the SAGE algorithm, with the estimation of 30 waves. Once the estimated angles of arrival of the 30 waves are available, we estimate the Von Mises distribution that best fits the data, using a maximum-likelihood procedure. The obtained best fit is denoted by ϕ_{SAGE} and κ_{SAGE} . The results

Table 1
RMS error in estimating the parameters using SML and SAGE methods.

SNR (dB)	10	20
RMSE $\{\phi_{SML}\}$	0.89	0.77
RMSE $\{\kappa_{SML}\}$	2.86	0.36
RMSE $\{\sigma_{d,SML}\}$	1.77×10^{-7}	1.60×10^{-7}
RMSE $\{\phi_{SAGE}\}$	1.79	1.97
RMSE $\{\kappa_{SAGE}\}$	15.48	17.68
RMSE $\{\sigma_{d,SAGE}\}$	5.30×10^{-7}	5.65×10^{-7}

show that the stochastic maximum likelihood approach outperforms the SAGE algorithm in terms of RMSE, for both the mean angle of arrival and the angular dispersion. For the SAGE algorithm in particular it was observed that the estimation errors actually increased together with the SNR. This may be caused by bias in the estimation due to modeling errors.

Following the approach in [6], we will define the instantaneous power azimuth-delay spectrum as

$$P_I(\phi, \tau) = \sum_{n=1}^{N_w} \frac{1}{N_w} \delta(\phi - \phi_n, \tau - \tau_n), \quad (15)$$

and the power azimuth-delay spectrum (PAD) as $P(\phi, \tau) = E\{P_I(\phi, \tau)\}$. The power azimuth spectrum and the power delay spectrum are derived as

$$P_A(\phi) = \sum_{\tau} P(\phi, \tau), \quad P_D(\phi) = \sum_{\phi} P(\phi, \tau) \quad (16)$$

Figure 4 shows the true PAD and the estimates obtained using the method described in this work. The estimated spectra closely match the true ones. In Figure 5 we also show the estimated PAD obtained using the SAGE algorithm. From Figure 5 and Table 1, it is seen that the results of the SAGE algorithm were particularly poor for the κ parameter. This happens because in the process of identifying the discrete waves those with highest power are selected to be estimated, and since these correspond to waves close to the mean DOA, the concentration of waves around the mean appears to be higher than in reality.

VI. CONCLUSION

We derived a stochastic maximum likelihood approach for estimation of the spatio-temporal behavior of MIMO channels. We used the channel model derived in [4] 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. This stochastic model leads more compact representation of the propagation model. Moreover, the computational complexity of maximizing the likelihood function is reduced because fewer parameters are needed and the likelihood function is smoother. The obtained results were compared to the SAGE algorithm used in [1–3], and the proposed method outperformed the SAGE algorithm in the root

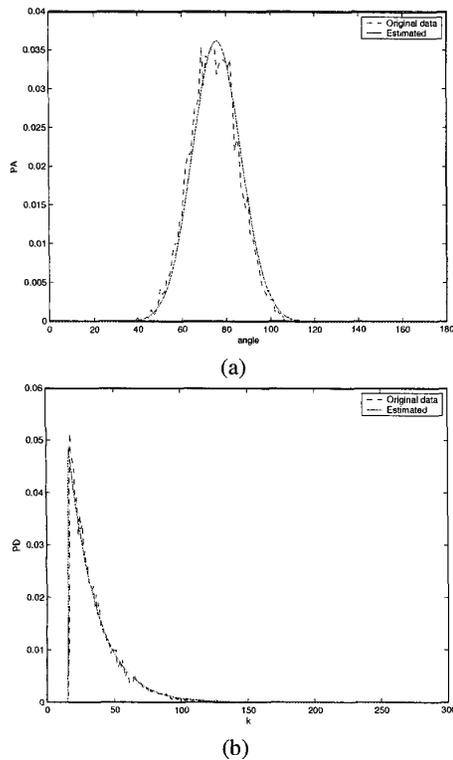


Fig. 4. Comparison between the true and estimated power azimuth and power delay spectra: (a) power azimuth spectrum; (b) power delay spectrum.

mean squared error of the estimates. The overall computational complexity is also reduced compared to SAGE algorithm, since the model dimension is lower in the proposed method.

ACKNOWLEDGEMENTS

This work was also supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, Brazilian governmental agency promoting scientific and technological development.

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] 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 bidirectional dispersion in the radio channel," in *IEEE 55th Vehicular Technology Conference*, May 2002, vol. 1, pp. 522–526.

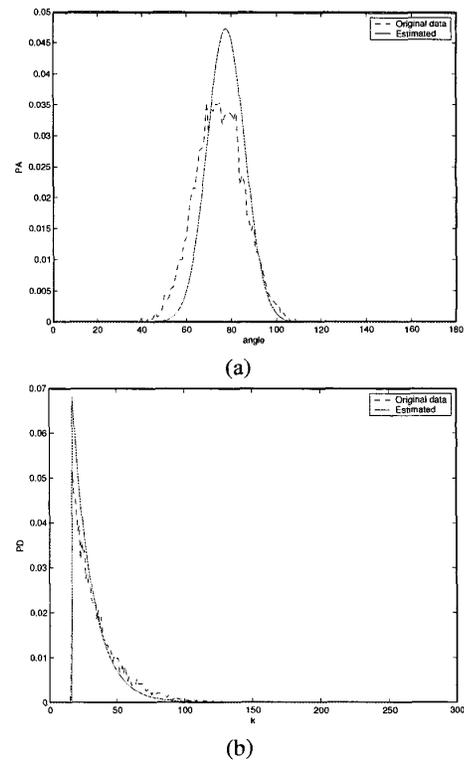


Fig. 5. Comparison between the true and estimated power azimuth-delay spectrum using the SAGE algorithm: (a) power azimuth spectrum; (b) power delay spectrum.

- [3] 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. GLOBECOM*, Dec. 2003.
- [4] 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.
- [5] K. V. Mardia, *Statistics of Directional Data*, Academic Press, London and New York, 1972.
- [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 Transactions on Vehicular Technology*, vol. 49, no. 2, pp. 437–447, March 2000.
- [7] K. Yu and B. Ottersten, "Models for MIMO Propagation Channels, A Review," *Special Issue on Adaptive Antennas and MIMO Systems, Wiley Journal on Wireless Communications and Mobile Computing*, vol. 2, no. 7, pp. 653–666, November 2002.
- [8] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the Nelder–Mead simplex method in low dimensions," *SIAM Journal on Optimization*, vol. 9, no. 1, pp. 112–147, 1998.