A GRADIENT BASED METHOD FOR MAXIMUM LIKELIHOOD CHANNEL PARAMETER ESTIMATION FROM MULTIDIMENSIONAL CHANNEL SOUNDING MEASUREMENTS

Andreas Richter(1), Markus Landmann(2), and Reiner S. Thomä(3)

(1) 98684 Ilmenau, Germany, POB 100565, Phone: +49 3677 691160, Fax: +49 3677 691113
Ilmenau University of Technology, Department of Electrical Engineering and Information Technology
E-mail: ric@e-technik.tu-ilmenau.de

(2) As (1) above, but E-mail: markus.landmann@tu-ilmenau.de

(3) As (1) above, but E-mail: reiner.thomae@tu-ilmenau.de

ABSTRACT

We describe a multidimensional maximum likelihood estimator for radio channel parameters. We also derive a data model to describe the complete data, that is virtually applicable to every antenna array geometry. The proposed iterative gradient based algorithm has been developed, since algorithms using component-wise updates of the likelihood function shows a slow convergence, if at least two propagation paths with nearly the same parameters exist in the measured radio propagation scenario, that means if super-resolution is necessary. The algorithm provides furthermore a variance estimate of the estimated parameters, since the Fisher-information matrix is calculated throughout the algorithm.

INTRODUCTION

The interest in the multidimensional structure of the mobile radio channel is growing rapidly. The initial motiva-tion was the investigation of the space-time structure at the base station (BS). In the recent time the double-directional modeling of the radio channel has attracted a lot of interest [1]. This is mainly due to two reasons. At one hand, double directional channel measurements gives a better physical insight into the wave propagation mechanism in real radio environments since it provides an enhanced multi-path resolution and it has the ability to remove the measurement antenna influence from the channel observation [2]. On the other hand, there is a growing interest in the exploitation of multiple antennas at both the BS and MS site. These MIMO (multiple-input-multiple-output) transmission systems promise a considerable increase in capacity [3]. Parametric MIMO channel models are required not only to estimate the achievable capacity from measurements [4] but also for realistic link-level simulations [5], [6] and to predict the long term channel parameters for controlling of the modem signal processing at the down-link.
Since resolution and accuracy of classical signal processing algorithms is limited by the available measurement aperture in the space-frequency-time domain, parametric super-resolution algorithms are applied to enhance the resolution by fitting an appropriate data model to the measured data. The achievable resolution is only limited by the signal to noise ratio (SNR), the remaining measurement device calibration error, and the limited validity of the data model. The algorithms applied to joint multidimensional radio channel parameter estimation from field experiments so far are the multidimensional ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) algorithm [7] and the SAGE (Space Alternating Generalized Expectation maximization) method [8] which, essentially, is an EM-based simplified ML parameter estimation procedure whereby the parameters are updated sequentially. Both algorithms have been applied to the problem [9],[10],[11],[12],[13],[14]. Important differences between the two algorithms to be considered is their applicability to certain antenna array architectures, calculation time in terms of convergence speed and statistical efficiency. It is well known that the ESPRIT-algorithm is an unbiased estimator for direction estimation only if the antenna arrays used for the measurements show a so called shift invariant structure (ULA, URA, CUBA) [7],[12],[14]. For other antenna array structures, i.e. UCA, UCPA, or spherical arrays, ESPRIT application is not possible or will at least result in biased estimates. Other drawbacks arise if we ask for a statistically efficient estimator and/or for the parameter estimation in a more complicated context such as colored measurement noise, non-ideal antenna-array-characteristics, etc. Therefore, we will focus on maximum likelihood parameter estimation in this paper.
As long as an sufficient accurate continuous parametric data model of the antenna array output signal is available, SAGE based channel parameter estimation can be applied for a large variety of antenna array architectures. The drawback is its slow convergence rate if two “closely spaced” propagation paths exist in the multi-path propagation scenario. Clearly, since we have only one transmitting source, all received paths have to be considered as potentially coherent. From our experience, this is a serious problem in typical micro- and pico-cell radio environments.

In the following Section we present our general data model for the measured radio channel, in the third Section an expression for the Fisher information matrix (FIM) of the parameters to be estimated is derived. In Section four we outline our gradient based multidimensional ML channel parameter estimator, and finally, in the last Section we present some concluding remarks.

GENERAL DATA MODEL

We will use the well known base-band representation of the double directional channel model [1],[12],[14] that approximates the narrow-band radio channel by the superposition of a finite number of propagation paths. Every propagation path is parameterized by 8 real values, the real- and imaginary part of the complex path weight γ, the transmit angles φ_T, ϑ_T (azimuth and elevation), the time-delay τ, the Doppler-shift α, and the receive angles φ_R, ϑ_R (azimuth and elevation).

$$\mathbf{h}(\alpha,\tau,\varphi_R,\vartheta_R,\varphi_T,\vartheta_T)=\sum_{k=1}^{K}\gamma_k\delta(\alpha-\alpha_k)\,\delta(\tau-\tau_k)\,\delta(\varphi_R-\varphi_{R_k})\,\delta(\vartheta_R-\vartheta_{R_k})\,\delta(\varphi_T-\varphi_{T_k})\,\delta(\vartheta_T-\vartheta_{T_k})$$

It is important to observe that every propagation path can be interpreted as a R-dimensional (6-D) shift operator on the transmit signal. It shifts the Tx-signal in the 4 independent angular domains, in the time-delay domain, and in the Doppler-frequency domain. A second observation important as well is, that the 6 related aperture domains frequency, time, and antenna array aperture are finite. The excitation signal is band-limited, the observation time interval is always finite, and the aperture of antenna arrays is limited too. Thirdly the parameters are or can be treated as bounded parameters. All angles are bounded by at least $(-\pi, +\pi)$, the time-delay is bounded by $(0, \tau_{max})$ where τ_{max} is a function of transmit power, free space loss, and receiver noise, and the Doppler-Shift is bounded by $(-\alpha_{max}, +\alpha_{max})$ whereby α_{max} is a function of the maximum velocity of the objects in the observed scenario and the carrier frequency.

Under this terms and considering that a shift in one domain can also be expressed by the multiplication with a complex exponential in the related aperture domain, the family of exponential functions is sufficient to construct a complete data model of the radio channel. For notational convenience we replace the shift-parameters of propagation path (component) k from the physical model using normalized shift parameters $\mu_k^{(i)}$, which are related to their physical counterparts by a unique projection. We collect all parameters $\mu_k^{(i)}$ belonging to one propagation path k in the vector $\boldsymbol{\mu}_k$, and construct the vector-valued basis function for one propagation path using the aperture sizes $N_1 ... N_R$ in the respective domains. Let

$$\mathbf{a}(\mu_k^{(i)})=N_i^{-\frac{1}{2}}\cdot\left[\exp\left(-\mathrm{j}\cdot\mu_k^{(i)}\cdot(N_i-1)/2\right)\ \ \dots\ \ 1\ \ \dots\ \ \exp\left(+\mathrm{j}\cdot\mu_k^{(i)}\cdot(N_i-1)/2\right)\right]^T$$

be the vector valued complex exponential related to the shift parameter $\mu_k^{(i)}$ in the dimension i of component k with length N_i, than $\mathbf{a}(\boldsymbol{\mu}_k)=\mathbf{a}(\mu_k^{(R)})\otimes\mathbf{a}(\mu_k^{(R-1)})\otimes\ldots\otimes\mathbf{a}(\mu_k^{(1)})$ is a vector valued function mapping the real shift parameters from $\mathbb{R}^R$ to a complex vector in $\mathbb{C}^N$ with unit length and size $N=N_1\cdot N_2\cdot\ldots\cdot N_R$. Introducing the linear projector **G,** describing the measurement system, we can express the observed MIMO impulse response of size M simply by

$$\mathbf{s}(\boldsymbol{\theta}_k)=(\gamma_{r,k}+\mathrm{j}\cdot\gamma_{i,k})\cdot\mathbf{G}\cdot\mathbf{a}(\boldsymbol{\mu}_k)=\gamma_k\cdot\mathbf{G}\cdot\mathbf{a}(\boldsymbol{\mu}_k)$$

with the complex weight γ_k and the parameter vector $\boldsymbol{\theta}_k=\begin{bmatrix}\gamma_{r,k} & \gamma_{i,k} & \boldsymbol{\mu}_k^T\end{bmatrix}$.

Let us clarify the meaning of **G** a little bit using an example. We assume the radio channel has been measured M_t times equally spaced over time using two-dimensional antenna arrays at both the Tx- and the Rx-site with M_T and M_R antenna elements respectively. The radio channel is measured in the frequency domain with a broadband-signal of M_f equally spaced lines around the carrier frequency f_c. Furthermore we assign the normalized shift parameters to the physical parameters in the following fashion ($\mu^{(1)}=f(\varphi_T),\mu^{(2)}=f(\vartheta_T),\mu^{(3)}=f(\tau),\mu^{(4)}=f(\alpha),\mu^{(5)}=f(\varphi_R),\mu^{(6)}=f(\vartheta_R)$). To keep things simple, we will assume narrow-band measurements in that sense, that it is sufficient in terms of the measurement accuracy to describe the directional characteristics of the antenna arrays at the carrier frequency, and that the Nyquist sampling theorem is strictly adhered to in all six dimensions. Now using the fact that the far-field beam-pattern of an antenna-array is the two-dimensional Fourier-transform of its aperture field, we can express the relation between the signals of the antenna array ports and the beam-patterns using their aperture fields. Collecting the aperture fields of the Tx- and Rx-array row-wise in the matrices $\mathbf{G}_T$ and $\mathbf{G}_R$ respectively, we can use $\mathbf{b}(\mu^{(1)},\mu^{(2)})=\mathbf{G}_T\cdot(\mathbf{a}(\mu^{(2)})\otimes\mathbf{a}(\mu^{(1)}))$ and $\mathbf{b}(\mu^{(5)},\mu^{(6)})=\mathbf{G}_R\cdot(\mathbf{a}(\mu^{(6)})\otimes\mathbf{a}(\mu^{(5)}))$ to express the relation between the signals on their ports and a far-field point source or drain in a fixed distance. Furthermore we use the diagonal matrix $\mathbf{G}_f$ to describe the frequency response of the measurement system and the identity matrix $\mathbf{G}_t=\mathbf{I}$ to describe the time sampling of the MIMO impulse responses. Altogether the matrix $\mathbf{G}=\mathbf{G}_R\otimes\mathbf{G}_t\otimes\mathbf{G}_f\otimes\mathbf{G}_T$ describes the measurement system of our example.

Taking into account that the radio channel is a linear system, we can say that the observation $\mathbf{s}(\boldsymbol{\theta})$ is up to a certain accuracy a superposition of a finite number K of propagation paths plus a complex vector **n** with i.i.d. Gaussian noise independent between real- and imaginary part. Introducing the complete parameter vector containing the parameters of all K components

$$\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\gamma}_r^T & \boldsymbol{\gamma}_i^T & \left(\boldsymbol{\mu}^{(1)}\right)^T & \ldots & \left(\boldsymbol{\mu}^{(R)}\right)^T \end{bmatrix}^T \in \mathbb{R}^{(2+R)\cdot K}$$

our observation can be expressed by

$$\mathbf{x} = \mathbf{n} + \sum_{k=1}^{K} \mathbf{s}(\boldsymbol{\theta}_k) = \mathbf{n} + \mathbf{s}(\boldsymbol{\theta}).$$

We will use this data model to derive the Fisher information matrix of the parameter vector $\boldsymbol{\theta}$ in the next Section.

STRUCTURE OF THE FISHER INFORMATION MATRIX

The conditional density of the observation $\mathbf{x}$ given $\boldsymbol{\theta}$ and σ is

$$pdf(\mathbf{x}|\boldsymbol{\theta},\sigma) = \frac{1}{(\pi\sigma^2)^M} e^{-\frac{1}{\sigma^2}(\mathbf{x}-\mathbf{s}(\boldsymbol{\theta}))^H \cdot (\mathbf{x}-\mathbf{s}(\boldsymbol{\theta}))}$$

taking the logarithm yields the log-likelihood function of the estimation problem as

$$L(\mathbf{x}|\boldsymbol{\theta},\sigma) = \ln\left(pdf(\mathbf{x}|\boldsymbol{\theta},\sigma)\right) = -M \cdot \ln(\pi\cdot\sigma^2) - \frac{1}{\sigma^2}(\mathbf{x}-\mathbf{s}(\boldsymbol{\theta}))^H \cdot (\mathbf{x}-\mathbf{s}(\boldsymbol{\theta})).$$

Defining the matrix with the first order partial derivatives of the complete data $\mathbf{s}(\boldsymbol{\theta})$

$$\mathbf{D}(\boldsymbol{\theta}) = \left[\left(\frac{\partial}{\partial\theta_1}\mathbf{s}(\boldsymbol{\theta})\right) \cdots \left(\frac{\partial}{\partial\theta_L}\mathbf{s}(\boldsymbol{\theta})\right)\right]$$

we get for the score function the expression

$$\mathbf{q}(\mathbf{x};\boldsymbol{\theta},\sigma^2) = 2\sigma^{-2}\cdot\Re\left\{\mathbf{D}(\boldsymbol{\theta})^H \cdot (\mathbf{x}-\mathbf{s}(\boldsymbol{\theta}))\right\} \tag{1}$$

and using some simple mathematics yields

$$\mathbf{J}(\boldsymbol{\theta},\sigma^2) = 2\sigma^{-2}\cdot\Re\left\{\mathbf{D}(\boldsymbol{\theta})^H \cdot \mathbf{D}(\boldsymbol{\theta})\right\} \tag{2}$$

for the Fisher information matrix of the parameter vector $\boldsymbol{\theta}$. Now we introduce the matrix containing the complex exponentials of all components in one dimension $\mathbf{A}_i$ and the related matrix with the first order derivatives $\boldsymbol{\Omega}_i$ as

$$\mathbf{A}_i = \left[\mathbf{a}(\mu_1^{(i)}) \;\; \mathbf{a}(\mu_2^{(i)}) \;\; \cdots \;\; \mathbf{a}(\mu_K^{(i)})\right], \text{ and } \boldsymbol{\Omega}_i = \left[\frac{\partial}{\partial\mu_1^{(i)}}\mathbf{a}(\mu_1^{(i)}) \;\; \frac{\partial}{\partial\mu_2^{(i)}}\mathbf{a}(\mu_2^{(i)}) \;\; \cdots \;\; \frac{\partial}{\partial\mu_K^{(i)}}\mathbf{a}(\mu_K^{(i)})\right].$$

For notational convenience we will use " $\Diamond$ " to express the column wise Kronecker-product of two matrices, leading to a compact matrix expression for our data model $\mathbf{s}(\boldsymbol{\theta}) = \mathbf{G}\cdot(\mathbf{A}_R\Diamond\ldots\Diamond\mathbf{A}_1)\cdot\boldsymbol{\gamma}$, where $\boldsymbol{\gamma} = [\gamma_1 \;\; \ldots \;\; \gamma_K]^T$. With $\mathbf{R}_\gamma = \boldsymbol{\gamma}\cdot\boldsymbol{\gamma}^H$, $\mathbf{B}_0 = \left[\mathbf{1}^{(1\times K)} \;\; \mathrm{j}\cdot\mathbf{1}^{(1\times K)} \;\; \mathbf{1}^{(1\times(R\cdot K))}\right]$, $\mathbf{B}_i = \left[\mathbf{1}^{(1\times(i+1))}\otimes\mathbf{A}_i \;\; \Omega_i \;\; \mathbf{1}^{(1\times(R+2-i))}\otimes\mathbf{A}_i\right]$, and $\mathbf{B} = \mathbf{B}_R\Diamond\ldots\Diamond\mathbf{B}_1\Diamond\mathbf{B}_0$ we can express (2) using the Schur-product " $\circ$ " by

$$\mathbf{J}(\boldsymbol{\theta},\sigma^2) = 2\cdot\sigma^{-2}\cdot\Re\left\{\left(\mathbf{1}^{((R+2)\times(R+2))}\otimes\mathbf{R}_\gamma\right)\circ\left(\mathbf{B}^H\cdot\mathbf{G}^H\cdot\mathbf{G}\cdot\mathbf{B}\right)\right\} \tag{3}$$

and if the matrix $\mathbf{G}$ is unitary we can simplify (3) to

$$\mathbf{J}(\boldsymbol{\theta},\sigma^2) = 2\cdot\sigma^{-2}\cdot\Re\left\{\left(\mathbf{1}^{((R+2)\times(R+2))}\otimes\mathbf{R}_\gamma\right)\circ\left(\mathbf{B}_0^H\cdot\mathbf{B}_0\right)\circ\left(\mathbf{B}_1^H\cdot\mathbf{B}_1\right)\circ\ldots\circ\left(\mathbf{B}_R^H\cdot\mathbf{B}_R\right)\right\}. \tag{4}$$

Recalling that the Cramer-Rao bound is the inverse of the Fisher information matrix, we can now easily calculate an lower bound for the error covariance matrix of any efficient unbiased estimator $\hat{\boldsymbol{\theta}}$. But the Fisher information provides furthermore some insight into the behavior of our problem in the vicinity of the solution as we shall see in the next Section.

MAXIMUM LIKELIHOOD ESTIMATION

Due to the large number of free parameters $(2+R)\cdot K$, the computational load involved in the direct maximization of the log-likelihood function by an $(R\cdot K)$-dimensional search algorithm is very heavy, and since the log-likelihood function is non-linear in the shift-parameters a general closed solution is, to our knowledge, not available. Therefore in [9],[10] the application of an EM-based iterative optimization the SAGE algorithm is proposed, unfortunately the computational attractive strategy to maximize the log-likelihood function coordinate-wise is also one of it's problems. To understand why, let us discuss another approach to the optimization task. Is it known that the log-likelihood function is quadratic near the maximum [15]. Therefore it is possible to calculate, in our case even analytically, a direction through the parameter space pointing in the direction of the local maximum, what leads us to an iterative optimization procedure using the Gauss-Newton algorithm. For this purpose we need the gradient and the Jacobian of the log-likelihood function at the actual point in the parameter space $\hat{\boldsymbol{\theta}}(n)$. Recalling that the gradient is the score function at $\hat{\boldsymbol{\theta}}(n)$ and the Jacobian is $\mathbf{D}(\hat{\boldsymbol{\theta}}(n))$, we obtain for the optimal step direction of the Gauss-Newton algorithm

$$\Delta\hat{\boldsymbol{\theta}}(n)=\left(\Re\left\{\mathbf{D}^{H}\left(\hat{\boldsymbol{\theta}}(n)\right)\cdot\mathbf{D}\left(\hat{\boldsymbol{\theta}}(n)\right)\right\}\right)^{-1}\cdot\Re\left\{\mathbf{D}^{H}\left(\hat{\boldsymbol{\theta}}(n)\right)\cdot\left(\mathbf{x}-\mathbf{s}\left(\hat{\boldsymbol{\theta}}(n)\right)\right)\right\}=\mathbf{J}\left(\mathbf{x};\hat{\boldsymbol{\theta}}(n)\right)^{-1}\cdot\mathbf{q}\left(\mathbf{x};\hat{\boldsymbol{\theta}}(n)\right)$$

and have now to maximize the log-likelihood function over λ to find the improved parameter set $\hat{\boldsymbol{\theta}}(n+1)$

$$\lambda_{opt}(n)=\arg\max_{\lambda}\left(L\left(\mathbf{x}\middle|\hat{\boldsymbol{\theta}}(n)-\lambda\cdot\Delta\hat{\boldsymbol{\theta}}(n)\right)\right)\Rightarrow\hat{\boldsymbol{\theta}}(n+1)=\hat{\boldsymbol{\theta}}(n)-\lambda_{opt}(n)\cdot\Delta\hat{\boldsymbol{\theta}}(n).$$

A good strategy is to start with a step size of $\lambda=1.0$ since it will send us right to the optimum for a function which is really quadratic. If we have improved the log-likelihood function we go to the next iteration or, if not, half the step size and try again. Considering that our log-likelihood function is quadratic at the optimum our algorithm will in general converge faster the closer we are to the optimum. Now lets discuss what happens if we use the SAGE strategy for the iterative maximization using two special cases. At first we assume a case where all components are separated in the parameter domain in such a way that all vectors $\mathbf{a}(\boldsymbol{\mu}_k)$ are orthogonal and furthermore that $\mathbf{G}$ is unitary, than the off diagonal elements of the FIM are very small and the optimal step direction for the Gauss-Newton algorithm is approx. a scaled version of the gradient. In other words we can also use a coordinate wise update strategy to find the optimum in an computational reasonable time. But if we change the scenario and assume we have at least two components which are close together in the parameter domain, meaning we have at least two vectors among all basis vectors $\mathbf{a}(\boldsymbol{\mu}_k)$ which are "highly correlated". Than the FIM has some strong off diagonal elements and the coordinate wise update strategy is not a good one anymore, especially if we consider that the closer we are to the optimum the higher the correlation between the vectors will be, leading to a decreasing convergence speed the closer we are to the maximum of the likelihood function.

Altogether, we propose the following general optimization strategy, i. calculate raw estimates of the parameters using some SAGE iteration, ii. maximize the likelihood function iteratively using the proposed Gauss-Newton algorithm until the optimum is found. After reaching the optimum one can calculate an estimate of the variance of $\hat{\boldsymbol{\theta}}_{ML}$ using

$$\hat{\boldsymbol{\sigma}}_{\theta}^{2}=\operatorname{diag}\left\{\mathbf{J}\left(\hat{\boldsymbol{\theta}}_{ML},\sigma^{2}\right)^{-1}\right\}.$$

It should be pointed out that although the algorithm is simple in concept, it still offers a lot of possibilities to reduce the computational costs further. One important point is for instance the calculation of the step direction $\Delta\boldsymbol{\theta}(n)$. One can collect components which are close together in the parameter domain in groups and break the problem into sub-problems. Another way to reduce the computational costs is e.g. to exploit the centro-symmetry of the vectors $\mathbf{a}\left(\mu_k^{(i)}\right)$.

Figure 1. shows exemplarily the convergence behavior of the proposed algorithm compared with the SAGE algorithm. Only two propagation paths has been simulated having the same parameters except the receive azimuth angle φ. An UCA with 16 elements and a diameter of two times the wavelength was used as the Rx antenna array. Our data model outlined above has been used for the SAGE and for the Gauss-Newton algorithm as well. The latter was initialized using one SAGE iteration.

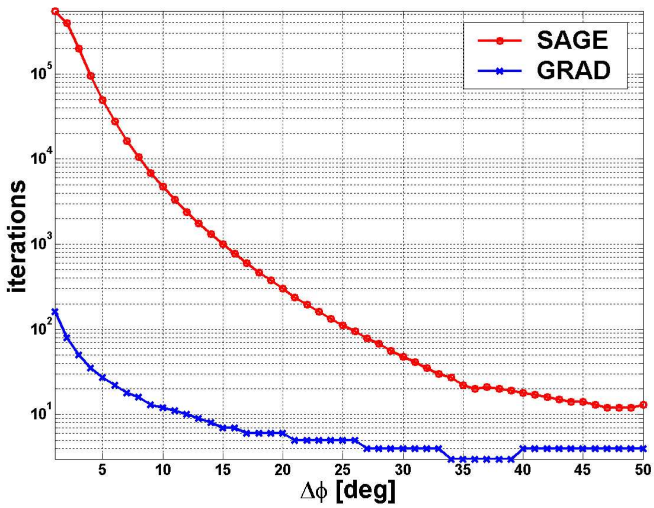

Figure 1.: Number of iterations vs. angular separation

CONCLUSION

We have presented an algorithm for multidimensional maximum likelihood radio channel parameter estimation from measurements that converges very fast to the optimum. The algorithm can be used with virtually every antenna array that is suitable for the parameter estimation problem itself. An issue not discussed in this paper is the initialization of the iterative ML algorithms. They will find a local maximum but it is not guaranteed that it will be the global maximum. This problem is still waiting for a solution.

REFERENCES

[1] M. Steinbauer, "A Comprehensive Transmission and Channel Model for Directional Radio Channels," EURO-COST, TD(98)027, Bern, 2nd - 4th Feb. 1998.

[2] R. S. Thomae, D. Hampicke, M. Landmann, G. Sommerkorn, A. Richter, „MIMO Measurement for Double-Directional Channel Modeling", Proc. IEE Technical Seminar on MIMO Communication Systems, December 12, 2001, London, UK

[3] G.J. Foschini and M.J. Gans, "On the limits of wireless communications in a fading environment when using multiple antennas," Wireless Personal Communications, Vol. 6, pp. 311-335, 1998.

[4] M. Steinbauer, A.F. Molisch, A. Burr, R. Thomae, "MIMO channel capacity based on measurement results," EWCS 2000, Oct. 2-6, Paris.

[5] U. Trautwein, D. Hampicke, G.Sommerkorn, R.S.Thomae, „Performance of Space-Time Processing for ISI- and CCI Suppression in Industrial Scenarios," Proc, IEEE VTC2000-Spring, Tokyo, May 2000.

[6] T. Yamada, S. Tomisato, T. Matsumoto, U. Trautwein, "Performance Evaluation of FTDL-Spatial/MLSE-Temporal Equalizers in the Presence of Co-Cannel Interference – Link-Level Simulation Results Using Field Measurement Data," IEICE Trans. Communications, vol. E 84-B, no. 7, July 2001, Tokyo

[7] M. Haardt, "Efficient One-, Two-, and Multidimensional High-Resolution Array Signal Processing," Ph.D. Thesis, Shaker Verlag, Aachen, Germany, 1996.

[8] J.A. Fessler, A.O. Hero, "Space Alternating Generalized Expectation Maximization Algorithm," IEEE Trans. on Signal Processing, Vol. 42, pp. 2664-2677, October 1994.

[9] B.H. Fleury, D. Dahlhaus, R. Heddergott, and M. Tschudin, "Wideband angle of arrival estimation using the SAGE algorithm," in Proc. IEEE Fourth Int. Symp. spread Spectrum techniques and Applications (ISSSTA'96), Mainz, Germany, pp. 79-85, Sept. 1996.

[10] Bernhard H. Fleury, Martin Tschudin, Ralf Heddergott, Dirk Dahlhaus, and Klaus Ingeman Pedersen, "Channel Parameter Estimation in Mobile Radio Environment Using the SAGE Algorithm," IEEE Journal on Selected Areas in Communications, Vol. 17, No. 3, pp. 434-450, March 1999.

[11] R.S. Thomä, D. Hampicke, A. Richter, G. Sommerkorn, U. Trautwein, "MIMO Vector Channel Sounder Measurement for Smart Antenna System Evaluation," European Transactions on Telecommunications, vol. 12, No. 5, September-October 2001

[12] Andreas Richter, Dirk Hampicke, Gerd Sommerkorn, and Reiner S. Thomae, "MIMO Measurement and Joint M-D Parameter Estimation of Mobile Radio Channels", VTC 2001-Spring, Rhodos, May 6-9, 2001

[13] M. Tschudin, R. Heddergott, and P. Truffer, "Validation of a high resolution measurement technique for estimating the parameters of impinging waves in indoor environments," in Proc. 9 th IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications (PIMRC'98) Boston, MA, pp. 1411-1416, Sept. 1998.

[14] A. Richter, D. Hampicke, G. Sommerkorn, and R.S. Thomae, "Joint estimation of DoD, time-delay, and DoA for high-resolution channel sounding," VTC 2000-Spring, Tokyo, May 15-18, 2000

[15] Louis L. Scharf, "Statistical Signal Processing: detection, estimation, and time series analysis", ISBN 0-201-19038-9