

EUROPEAN COOPERATION  
IN THE FIELD OF SCIENTIFIC  
AND TECHNICAL RESEARCH

---

COST 273 TD(03)198  
Prague, Czech Republic  
2003/Sept/24-26

EURO-COST

---

SOURCE: Institute of Communications and Measurement Engineering,  
Ilmenau University of Technology  
Ilmenau

## **Parametric Modelling and Estimation of Distributed Diffuse Scattering Components of Radio Channels**

Andreas Richter  
Technische Universität Ilmenau  
FG EMT  
POB 100565  
98684 Ilmenau  
GERMANY  
Phone: + 49-3677 69 1160  
Fax: + 49-3677 69 1113  
Email: [ric@e-technik.tu-ilmenau.de](mailto:ric@e-technik.tu-ilmenau.de)

# Parametric Modelling and Estimation of Distributed Diffuse Scattering Components of Radio Channels

Andreas Richter, Reiner S. Thomä

Ilmenau University of Technology, Ilmenau, Germany  
ric@e-technik.tu-ilmenau.de

*Abstract -- We introduce an extended data-model for high resolution channel parameter estimation and parametric channel modeling. Other than the well-known ray-optical based data models which contain only discrete (specular) propagation paths, we additionally introduce distributed diffuse scattering components. To this end a simple parametric data model of the diffuse scattering distribution in the delay domain is proposed. Furthermore, we develop an estimator for those model parameters and derive their Cramér-Rao lower bound. Finally we discuss implementation related issues, which arise if the extended channel model is integrated into existing high-resolution parameter estimation algorithms (such as ESPRIT, RIMAX, or SAGE) for the estimation of discrete propagation paths. It is demonstrated that also the reliability of high-resolution parameter estimation results in channel sounding measurements can be considerably enhanced.*

## I. Introduction

On every crossing between two propagation media, where the material properties the relative electric or magnetic permeability changes, a propagating wave is subdivided into parts. A part of the waves travels into the other medium, and another part is reflected or scattered back. The probability that a scattered wave reaches the receiver is generally higher than the probability that a reflected wave reaches him. This is due to the fact that a reflection requires a sufficiently large object with a reflecting surface. And if the reflection occurs it can only reach the receiver if the angles of incidence and the angle of reflection are appropriate to reach the next reflector or the receive antenna. Altogether the amount of specular (discrete/concentrated) propagation paths in a scenario is relatively small but their contribution to the total power transferred by them from the transmit antenna to the receive antenna is usually dominating the transmission. Although the contribution to the received power of a single scattered wave is small compared with the contribution of a reflected wave, the contribution of all scattered waves reaching the receiver is significant due to their large number. There exist scenarios where the distributed diffuse scattering is the dominating transmission mechanism, namely the industry scenario is a scenario where this happens often. Hence, we can not ignore the contribution of the distributed scattering to the radio transmission, if we want to model the radio channel and estimate the model parameters from channel observations.

This paper is devoted to the parametric modelling and estimation of the distributed diffuse scattering components of the observed radio channel in the time-delay domain.

## II. Continuous Data Model

Due to the finite measurement apertures available for the observation of the radio channel, it is not possible to resolve the large amount of scattered waves reaching the receiver. We can only resolve them within the Rayleigh-resolution determined by the apertures used during the measurements. The power of the diffuse components reaching the receiver within a time in-

terval  $t_0$  which is the reciprocal of the Measurement Bandwidth can't be resolved. The power of the components reaching the receiver within this time interval is largely determined by the free space attenuation which is approximately constant over this interval. So we can conclude that one complex delay-bin of the impulse response representing a observed time interval of  $t_0$  contains the superposition of some "propagation paths" caused by scattering with approximately the same power. But the phase of this "propagation paths" is due to the large difference in terms of the wavelength between their path lengths, approximately uniformly distributed within the interval  $\langle -\pi, +\pi \rangle$ . To clarify this a little bit, let us discuss an example. We assume the radio channel is observed with a bandwidth of 120MHz at a carrier frequency of 5.2GHz. Then a single delay-bin in the impulse-response represents the superposition of all propagation paths with an electrical length difference of approximately 2.5m. But since the wavelength at 5.2GHz is approximately 5.77cm this interval represents also propagation paths with an electrical length difference of up to 43 wave lengths.

Based on the central limit theorem we can assume that one complex delay-bin in the impulse response can be modeled as a complex circular normal distribution with zero mean, provided the relative bandwidth of the measurements is small. What we need furthermore, is a parametric model for the statistics of all delay-bins of the observed impulse response. To this end let us discuss a continuous model, which has been used to describe the power delay profile by various researchers [1], [2], [3]. The model is based on the observation that the power delay profile has an exponential decay over the delay-time and an base delay which is of course related to the distance between the transmit and the receive antenna. Except for the infinite bandwidth assumed, the proposed model (1) describes the power delay profile respectively the variance as a function of the time-delay of the distributed diffuse scattering components very good.

$$\Psi_x(\tau) = E\{|x(\tau)|^2\} = \begin{cases} 0 & \tau < \tau'_d \\ \alpha_1 \cdot \frac{1}{2} & \tau = \tau'_d \\ \alpha_1 \cdot e^{-B_d(\tau - \tau'_d)} & \tau > \tau'_d \end{cases} \quad (1)$$

The following table summarizes the parameters of the model (1).

$B_d$	- coherence bandwidth of the diffuse components
$\tau'_d$	- base delay of the diffuse components
$\beta_d = \frac{B_d}{B_m} = \frac{B_d}{M \cdot f_0}$	- coherence bandwidth of the diffuse components normalized to the measurement bandwidth
$M$	- number of frequency points measured within the measurement bandwidth
$\tau_d = \frac{\tau'_d}{t_m}$	- base delay of the diffuse components normalized to the total length of the observed impulse-response
$\alpha_1$	- power of the diffuse components at $\tau = \tau_d$

The related spectra, the Fourier-transform of (1) is

$$\Psi_x(f) = \frac{\alpha_1}{\beta + j2\pi f} \cdot e^{-j2\pi f \tau_d} \quad (2)$$

The model (1) for the distributed diffuse scattering components in the delay domain is incomplete insofar as it neglects the covariance between components at different time-delays. Let us assume the Fourier transform of the channel impulse response  $x(\tau)$ , the channel transfer function, exist and is  $X(f)$

$$x(\tau) \circ \bullet X(f) \quad (3)$$

Then assuming that  $E\{x(\tau_1)x(\tau_2)\}=0, \forall \tau_1 \neq \tau_2$  the covariance function of the channel impulse response  $\Psi_{xx}(\tau_1, \tau_2)$  and the covariance function of the channel transfer function  $\Psi_{xx}(f_1, f_2)$  are related in the following way

$$\begin{aligned} \Psi_{xx}(\tau_1, \tau_2) &= \Psi_x(\tau_2) \cdot \delta(\tau_1 - \tau_2) = E\{x(\tau_1) \cdot x(\tau_2)^*\} \\ &\begin{array}{c} \tau_1 \circ \\ | \\ \bullet f_1 \\ | \\ \bullet \tau_2 \\ | \\ \circ f_2 \end{array} = \Psi_x(\tau_2) \cdot e^{-j2\pi f_1 \tau_2} \\ \Psi_{xx}(f_1, f_2) &= \Psi_x(f_1 - f_2) = E\{X(f_1) \cdot X(f_2)^*\} \end{aligned} \quad (4)$$

So the function  $\Psi_x(f)$  (2) describes actually a spectral correlation between frequencies components with a certain distance  $f_1 - f_2$ . Since we observe the radio channel usually with a limited bandwidth  $B_m$  we introduce now the frequency response of the measurement system as  $G(f)$  and the related impulse response  $g(\tau)$ . Using the frequency response  $G(f)$  in (4) yields

$$\begin{aligned} \Psi_{xxgg}(f_1, f_2) &= \Psi_x(f_1 - f_2) \cdot G(f_1) \cdot G(f_2)^* = E\{X(f_1)G(f_1) \cdot X(f_2)^*G(f_2)^*\} \\ &\begin{array}{c} f_2 \circ \\ | \\ \bullet \tau_2 \\ | \\ \bullet f_1 \\ | \\ \circ \tau_1 \end{array} = (\Psi_x(\tau_2) \cdot e^{-j2\pi f_1 \tau_2} \cdot G(f_1))^*_{\tau_2} g(\tau_2)^* \\ \Psi_{xxgg}(\tau_1, \tau_2) &= \Psi_x(\tau_2) \cdot \delta(\tau_1 - \tau_2) *_{\tau_1} g(\tau_1)^* *_{\tau_2} g(\tau_2)^* = E\{x(\tau_1)x(\tau_1)^*\} *_{\tau_1} g(\tau_1)^* *_{\tau_2} g(\tau_2)^* \end{aligned} \quad (5)$$

where  $*_{\tau_1}$  and  $*_{\tau_2}$  denotes the convolution over  $\tau_1$  and  $\tau_2$  respectively. If we measure the channel using a measurement system with a bandwidth of  $B_m$  and a rectangular frequency response the covariance function of the observed channel impulse response becomes

$$\Psi_{xxgg}(\tau_1, \tau_2) = \Psi_{xx}(\tau_1, \tau_2) *_{\tau_1} \text{si}(\pi B_m \tau_1) *_{\tau_2} \text{si}(\pi B_m \tau_2). \quad (6)$$

Based on this continuous model we will now derive a statistic for a sampled version of the observed channel transfer function. But before we proceed one should observe that the convolution with the sinc-function in equation (6) assures that our assumption about the distribution of a single delay-bin is valid.

### III. Data Model for Parameter Estimation

Let us assume we measure the channel transfer function at  $M$  frequency points equidistantly over the measurement Bandwidth  $B_m$ . Furthermore we assume, that the observed channel  $\mathbf{x}$  contains distributed diffuse scattering components only, that means we will ignore the discrete components of the channel for starters. With this assumption and recalling the discussion about the nature of the band limited observed channel above, we model the distribution of the observed data as a multivariate circular normal distribution

$$p(\mathbf{x}|\boldsymbol{\theta}) = \frac{1}{\pi^M \det(\mathbf{R}(\boldsymbol{\theta}))} e^{-\mathbf{x}^H \mathbf{R}(\boldsymbol{\theta})^{-1} \mathbf{x}}. \quad (7)$$

The related log-likelihood function  $\mathcal{L}(\mathbf{x}|\boldsymbol{\theta}) = \ln(p(\mathbf{x}|\boldsymbol{\theta}))$  is

$$\mathcal{L}(\mathbf{x}|\boldsymbol{\theta}) = -M \cdot \ln(\pi) - \ln(\det(\mathbf{R}(\boldsymbol{\theta}))) - \mathbf{x}^H \cdot \mathbf{R}^{-1}(\boldsymbol{\theta}) \cdot \mathbf{x}. \quad (8)$$

Since an observation contains often multiple realizations of this process  $\mathbf{x}_1, \dots, \mathbf{x}_N$ , taken over time (sequence of observed transfer functions) or space (space sampling using an antenna ar-

ray) we collect multiple observations in the matrix  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N]$  and write for the distribution of the  $N$  independent realizations

$$p(\mathbf{X}|\boldsymbol{\theta}) = \frac{1}{\pi^{MN} \det(\mathbf{R}_{mm}(\boldsymbol{\theta}))^N} e^{-\text{tr}(\mathbf{X}^H \cdot \mathbf{R}_{mm}^{-1}(\boldsymbol{\theta}) \cdot \mathbf{X})}. \quad (9)$$

The related log-likelihood function is

$$\mathcal{L}(\mathbf{X}|\boldsymbol{\theta}) = -MN \cdot \ln(\pi) - N \ln(\det(\mathbf{R}(\boldsymbol{\theta}))) - \text{tr}(\mathbf{X}^H \cdot \mathbf{R}(\boldsymbol{\theta})^{-1} \cdot \mathbf{X}). \quad (10)$$

Equation (5) implies that the covariance matrix of the observations is a Toeplitz-matrix (in the frequency domain). Hence, assuming the measurement system has a flat transfer function over the bandwidth used to observe the channel, we can express the covariance matrix as

$$\mathbf{R}(\boldsymbol{\theta}) = \text{toep}(\boldsymbol{\kappa}(\boldsymbol{\theta}), \boldsymbol{\kappa}(\boldsymbol{\theta})^H). \quad (11)$$

Since the measurement noise can also be modeled as a multivariate circular normal distribution we include it in our model, assuming it has a covariance matrix of  $\alpha_0 \mathbf{I}$ . Consequently the sampled version of (2) is

$$\boldsymbol{\kappa}(\boldsymbol{\theta}) = \frac{\alpha_1}{M} \begin{bmatrix} 1 & \frac{e^{-j2\pi\tau_0}}{\beta_d + j2\pi\frac{1}{M}} & \dots & \frac{e^{-j2\pi(M-1)\tau_0}}{\beta_d + j2\pi\frac{M-1}{M}} \end{bmatrix}^T + \alpha_0 \mathbf{e}_0, \quad (12)$$

where  $\alpha_0$  is the variance of the circular independent identical normal distributed measurement noise, and  $\mathbf{e}_0 = [1 \ 0 \ \dots \ 0]^T$  is a unit vector.

Now we will derive the score function, the Fisher information matrix, and the Cramér-Rao bound for our model parameters.

#### IV. Cramér-Rao Bound

The derivation of the Cramér-Rao bound for the variance of any unbiased estimator  $\hat{\boldsymbol{\theta}}$  for the parameters  $\boldsymbol{\theta}$  of the covariance-matrix  $\mathbf{R}(\boldsymbol{\theta})$  is straight forward. We start with the score function, which is the gradient of the log-likelihood function [10] and yield

$$\frac{\partial}{\partial \theta_i} \mathcal{L}(\mathbf{X}|\boldsymbol{\theta}) = -N \text{tr} \left( \frac{1}{\det(\mathbf{R}(\boldsymbol{\theta}))} \text{adj}(\mathbf{R}(\boldsymbol{\theta})) \left( \frac{\partial}{\partial \theta_i} \mathbf{R}(\boldsymbol{\theta}) \right) \right) + \text{tr} \left( \mathbf{X}^H \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial}{\partial \theta_i} \mathbf{R}(\boldsymbol{\theta}) \right) \mathbf{R}(\boldsymbol{\theta})^{-1} \right). \quad (13)$$

Since  $\det(\mathbf{R}(\boldsymbol{\theta}))^{-1} \cdot \text{adj}(\mathbf{R}(\boldsymbol{\theta}))$  is just the inverse  $\mathbf{R}(\boldsymbol{\theta})^{-1}$  of the covariance matrix  $\mathbf{R}(\boldsymbol{\theta})$ , we can simplify expression (13) to

$$\frac{\partial}{\partial \theta_i} \mathcal{L}(\mathbf{X}|\boldsymbol{\theta}) = \text{tr} \left( \mathbf{X}^H \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial}{\partial \theta_i} \mathbf{R}(\boldsymbol{\theta}) \right) \mathbf{R}(\boldsymbol{\theta})^{-1} \mathbf{X} \right) - N \text{tr} \left( \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial}{\partial \theta_i} \mathbf{R}(\boldsymbol{\theta}) \right) \right). \quad (14)$$

Using the relation between the inner and outer product of two vectors  $\mathbf{x}^H \mathbf{y} = \text{tr}(\mathbf{y} \mathbf{x}^H)$  we write (14) as

$$\frac{\partial}{\partial \theta_i} \mathcal{L}(\mathbf{X}|\boldsymbol{\theta}) = N \text{tr} \left( \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial}{\partial \theta_i} \mathbf{R}(\boldsymbol{\theta}) \right) \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{1}{N} \mathbf{X} \mathbf{X}^H - \mathbf{R}(\boldsymbol{\theta}) \right) \right). \quad (15)$$

Now we note that the term  $1/N \cdot \mathbf{X} \cdot \mathbf{X}^H$  in (15) can be interpreted as an estimate of the covariance matrix  $\mathbf{R}(\boldsymbol{\theta})$  derived from the observation  $\mathbf{X}$ , and introduce the non-parametric estimate of the covariance matrix as

$$\hat{\mathbf{R}} = \frac{1}{N} \mathbf{X}\mathbf{X}^H. \quad (16)$$

Hence we get for the score function the expression

$$\frac{\partial}{\partial \theta_i} \mathcal{L}(\mathbf{X}|\boldsymbol{\theta}) = N \operatorname{tr} \left( \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial}{\partial \theta_i} \mathbf{R}(\boldsymbol{\theta}) \right) \mathbf{R}(\boldsymbol{\theta})^{-1} (\hat{\mathbf{R}} - \mathbf{R}(\boldsymbol{\theta})) \right) \quad (17)$$

or equivalently

$$\frac{\partial}{\partial \theta_i} \mathcal{L}(\mathbf{X}|\boldsymbol{\theta}) = N \operatorname{tr} \left( \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial}{\partial \theta_i} \mathbf{R}(\boldsymbol{\theta}) \right) (\mathbf{R}(\boldsymbol{\theta})^{-1} \hat{\mathbf{R}} - \mathbf{I}) \right). \quad (18)$$

The stochastic Fisher matrix is the second gradient of the log-likelihood function [10], hence we take the second partial derivative to the parameter  $\theta_k$  and obtain

$$\begin{aligned} \frac{\partial^2}{\partial \theta_i \partial \theta_k} \mathcal{L}(\mathbf{X}|\boldsymbol{\theta}) = N \operatorname{tr} & \left( -\mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial}{\partial \theta_k} \mathbf{R}(\boldsymbol{\theta}) \right) \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial}{\partial \theta_i} \mathbf{R}(\boldsymbol{\theta}) \right) (\mathbf{R}(\boldsymbol{\theta})^{-1} \hat{\mathbf{R}} - \mathbf{I}) + \right. \\ & \left. + \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial^2}{\partial \theta_i \partial \theta_k} \mathbf{R}(\boldsymbol{\theta}) \right) (\mathbf{R}(\boldsymbol{\theta})^{-1} \hat{\mathbf{R}} - \mathbf{I}) + \right. \\ & \left. - \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial}{\partial \theta_i} \mathbf{R}(\boldsymbol{\theta}) \right) \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial}{\partial \theta_k} \mathbf{R}(\boldsymbol{\theta}) \right) \mathbf{R}(\boldsymbol{\theta})^{-1} \hat{\mathbf{R}} \right) \end{aligned} \quad (19)$$

Now assuming an unbiased estimator for the covariance matrix

$$\mathbf{E}\{\hat{\mathbf{R}} - \mathbf{R}(\hat{\boldsymbol{\theta}})\} = \mathbf{0}, \text{ and } \mathbf{E}\{\mathbf{R}(\hat{\boldsymbol{\theta}})^{-1} \hat{\mathbf{R}}\} = \mathbf{I}$$

we yield for the elements of the *Fisher Information Matrix* the equation

$$J_{ik} = -\mathbf{E} \left\{ \frac{\partial^2}{\partial \theta_i \partial \theta_k} \mathcal{L}(\mathbf{X}|\boldsymbol{\theta}) \right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = N \operatorname{tr} \left\{ \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial}{\partial \theta_i} \mathbf{R}(\boldsymbol{\theta}) \right) \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial}{\partial \theta_k} \mathbf{R}(\boldsymbol{\theta}) \right) \right\}. \quad (20)$$

Recalling that the sum of the main diagonal elements (trace) of the product between an square hermitian matrix  $\mathbf{A} = \mathbf{A}^H$  and an arbitrary square matrix  $\mathbf{B}$  can be expressed using the vector operator  $\operatorname{vec}\{\bullet\}$  as

$$\operatorname{tr}(\mathbf{A} \cdot \mathbf{B}) = \operatorname{tr}(\mathbf{A}^H \cdot \mathbf{B}) = \sum_{n=1}^N \mathbf{a}_n^H \mathbf{b}_n = \operatorname{vec}\{\mathbf{A}\}^H \operatorname{vec}\{\mathbf{B}\},$$

we write for a single element of the *Fisher Information matrix* the expression

$$j_{ik} = -\mathbf{E} \left\{ \frac{\partial^2}{\partial \theta_i \partial \theta_k} \mathcal{L}(\mathbf{X}|\boldsymbol{\theta}) \right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = N \cdot \operatorname{vec} \left\{ \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial}{\partial \theta_i} \mathbf{R}(\boldsymbol{\theta}) \right) \right\}^H \cdot \operatorname{vec} \left\{ \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial}{\partial \theta_k} \mathbf{R}(\boldsymbol{\theta}) \right) \right\}.$$

So we use the vector operator  $\operatorname{vec}\{\bullet\}$  and define the matrix containing the partial derivatives to all  $L$  elements of the parameter vector  $\boldsymbol{\theta}$  as

$$\mathbf{D}(\boldsymbol{\theta}) = \left[ \operatorname{vec} \left\{ \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial}{\partial \theta_1} \mathbf{R}(\boldsymbol{\theta}) \right) \right\} \quad \dots \quad \operatorname{vec} \left\{ \mathbf{R}(\boldsymbol{\theta})^{-1} \left( \frac{\partial}{\partial \theta_L} \mathbf{R}(\boldsymbol{\theta}) \right) \right\} \right]. \quad (21)$$

Using the matrix  $\mathbf{D}(\boldsymbol{\theta})$  the *Fisher Information matrix*  $\mathcal{J}(\boldsymbol{\theta})$  of the parameters  $\boldsymbol{\theta}$  related to the covariance matrix  $\mathbf{R}(\boldsymbol{\theta})$  can be expressed in a compact form as

$$\mathbf{j}(\boldsymbol{\theta}) = -\mathbb{E}\left\{\frac{\partial^2}{\partial\theta_i\partial\theta_k}\mathcal{L}(\mathbf{X}|\boldsymbol{\theta})\right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = N \cdot \mathbf{D}(\boldsymbol{\theta})^H \cdot \mathbf{D}(\boldsymbol{\theta}). \quad (22)$$

Now we can easily derive the asymptotic Cramér-Rao bound [10], [11] on the variance of any unbiased estimator  $\hat{\boldsymbol{\theta}}$  for the parameters  $\boldsymbol{\theta}$  related to the covariance matrix  $\mathbf{R}(\boldsymbol{\theta})$  as

$$\text{CRB}(\boldsymbol{\theta}_0) = \frac{1}{N} \left(\mathbf{D}(\boldsymbol{\theta})^H \mathbf{D}(\boldsymbol{\theta})\right)^{-1} \leq \mathbb{E}\left\{\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right)^T\right\}. \quad (23)$$

Furthermore using the matrix  $\mathbf{D}(\boldsymbol{\theta})$  the score function (18) can be expressed as

$$\mathbf{q}(\mathbf{X}, \boldsymbol{\theta}) = \frac{\partial}{\partial\boldsymbol{\theta}} \mathcal{L}(\mathbf{X}, \boldsymbol{\theta}) = N \cdot \mathbf{D}(\boldsymbol{\theta})^H \cdot \text{vec}\left\{\mathbf{R}(\boldsymbol{\theta})^{-1} \hat{\mathbf{R}} - \mathbf{I}\right\} \quad (24)$$

So far, we have derived a general expressions for the score function, the Fisher information matrix, and the Cramér-Rao bound of the parameters  $\boldsymbol{\theta}$  related to the covariance matrix  $\mathbf{R}(\boldsymbol{\theta})$ . Now we will derive the expressions for the four parameters  $\boldsymbol{\theta} = [\alpha_0 \quad \alpha_1 \quad \beta \quad \tau_0]^T$ . Since  $\text{toep}(\bullet)$  is a linear operator the following equation holds

$$\frac{\partial}{\partial\theta_i} \mathbf{R}(\boldsymbol{\theta}) = \text{toep}\left(\frac{\partial}{\partial\theta_i} \boldsymbol{\kappa}(\boldsymbol{\theta}), \frac{\partial}{\partial\theta_i} \boldsymbol{\kappa}(\boldsymbol{\theta})^H\right).$$

Hence, we have only to calculate the first order partial derivatives of (12) to the four parameters in the parameter vector  $\boldsymbol{\theta} = [\alpha_0 \quad \alpha_1 \quad \beta \quad \tau_0]^T$  to derive the related Fisher Information matrix. The four partial derivatives are

$$\frac{\partial}{\partial\alpha_0} \boldsymbol{\kappa}(\boldsymbol{\theta}) = \mathbf{e}_0, \quad (25)$$

$$\frac{\partial}{\partial\alpha_1} \boldsymbol{\kappa}(\boldsymbol{\theta}) = \frac{1}{M} \begin{bmatrix} 1 & \frac{e^{-j2\pi\tau_0}}{\beta_d + j2\pi\frac{1}{M}} & \cdots & \frac{e^{-j2\pi(M-1)\tau_0}}{\beta_d + j2\pi\frac{M-1}{M}} \end{bmatrix}^T, \quad (26)$$

$$\frac{\partial}{\partial\beta} \boldsymbol{\kappa}(\boldsymbol{\theta}) = -\frac{\alpha_1}{M} \begin{bmatrix} 1 & \frac{e^{-j2\pi\tau_0}}{(\beta_d + j2\pi\frac{1}{M})^2} & \cdots & \frac{e^{-j2\pi(M-1)\tau_0}}{(\beta_d + j2\pi\frac{M-1}{M})^2} \end{bmatrix}^T, \quad (27)$$

and

$$\frac{\partial}{\partial\tau_0} \boldsymbol{\kappa}(\boldsymbol{\theta}) = \frac{\alpha_1}{M} \begin{bmatrix} 0 & -j2\pi \cdot \frac{e^{-j2\pi\tau_0}}{\beta_d + j2\pi\frac{1}{M}} & \cdots & -j2\pi(M-1) \cdot \frac{e^{-j2\pi(M-1)\tau_0}}{\beta_d + j2\pi\frac{M-1}{M}} \end{bmatrix}^T. \quad (28)$$

Using

$$\frac{\partial}{\partial\alpha_0} \mathbf{R}(\boldsymbol{\theta}) = \text{toep}\left(\frac{\partial}{\partial\alpha_0} \boldsymbol{\kappa}(\boldsymbol{\theta}), \frac{\partial}{\partial\alpha_0} \boldsymbol{\kappa}(\boldsymbol{\theta})^H\right) = \mathbf{I}, \quad (29)$$

$$\frac{\partial}{\partial\alpha_1} \mathbf{R}(\boldsymbol{\theta}) = \text{toep}\left(\frac{\partial}{\partial\alpha_1} \boldsymbol{\kappa}(\boldsymbol{\theta}), \frac{\partial}{\partial\alpha_1} \boldsymbol{\kappa}(\boldsymbol{\theta})^H\right), \quad (30)$$

$$\frac{\partial}{\partial\beta} \mathbf{R}(\boldsymbol{\theta}) = \text{toep}\left(\frac{\partial}{\partial\beta} \boldsymbol{\kappa}(\boldsymbol{\theta}), \frac{\partial}{\partial\beta} \boldsymbol{\kappa}(\boldsymbol{\theta})^H\right), \quad (31)$$

and

$$\frac{\partial}{\partial \tau_0} \mathbf{R}(\boldsymbol{\theta}) = \text{toep} \left( \frac{\partial}{\partial \tau_0} \boldsymbol{\kappa}(\boldsymbol{\theta}), \frac{\partial}{\partial \tau_0} \boldsymbol{\kappa}(\boldsymbol{\theta})^H \right) \quad (32)$$

in equation (21) the Cramér-Rao bound of any unbiased estimator for the parameters of the covariance matrix  $\mathbf{R}(\boldsymbol{\theta})$  can be calculated using equation (23). Similar expressions can be found in

## V. Parameter Estimation

Since the parametric covariance matrix (11), (12) is a nonlinear function in some of its parameters, it wasn't possible so far to derive a closed solution for the maximization problem

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}, \mathbf{X}). \quad (33)$$

Hence, we will discuss global and local search strategies to find the parameter vector  $\hat{\boldsymbol{\theta}}$  which maximizes (33).

### Local Search Strategies

Having derived the score function and the Fisher information matrix one can always use a gradient based local optimization strategy. Since the score function (24) is the gradient of the log-likelihood function and the Fisher information matrix is an approximation of the Hessian we can use the Gauss-Newton algorithm to iteratively maximize (33) using the iteration

$$\hat{\boldsymbol{\theta}}_{i+1} = \hat{\boldsymbol{\theta}}_i + \lambda \cdot \boldsymbol{\jmath}(\hat{\boldsymbol{\theta}}_i)^{-1} \cdot \mathbf{q}(\mathbf{X}, \hat{\boldsymbol{\theta}}_i). \quad (34)$$

To find a suitable step size  $\lambda$ , one can use a line search strategy or a trust region algorithm to ensure strict maximization of (33). The disadvantage of this local optimization strategy is the high memory and computation effort needed for a single iteration. For every iteration we have to invert the Toeplitz matrix  $\mathbf{R}(\hat{\boldsymbol{\theta}}_i)$ , if we use the Levinson algorithm this requires  $O(M^2)$  operations. To calculate the matrix (21) we need the product between this inverse and the Toeplitz matrices (29)-(32), an efficient algorithm for the product between the inverse of a Toeplitz matrix and an arbitrary vector requires  $O(M \log(M))$  operations using FFTs, hence we need for the calculation of (21)  $O(L \cdot M^2 \log(M))$  operations. The complexity needed to determine the score function and the Fisher Information matrix from (21) is  $O(M^2 \log(M)) + O(M^2 L)$  and  $O(M^2 L^2)$  respectively. Therefore the complexity of one iteration is of  $O(L \cdot M^2 \log(M))$ , under the assumption that  $L < \log(M)$ . Additionally we have to store the complete estimated non-parametric covariance matrix (16). A suitable method to reduce the memory requirements is to replace the non-parametric covariance matrix  $\hat{\mathbf{R}}$  with  $\text{toep}(\hat{\mathbf{r}}, \hat{\mathbf{r}}^H)$ , using the coefficients

$$\hat{\mathbf{r}}_l = \frac{1}{M-l+1} \sum_{i=1}^{M-l+1} \hat{r}_{i+l-1,i} \quad (l=1, \dots, M). \quad (35)$$

The elements of  $\hat{\mathbf{r}}$  are the mean values of the diagonals of the matrix  $\hat{\mathbf{R}}$ , and have an expected value of  $E\{\hat{\mathbf{r}}\} = \boldsymbol{\kappa}(\boldsymbol{\theta})$ . We can expect that an algorithm using  $\text{toep}(\hat{\mathbf{r}}, \hat{\mathbf{r}}^H)$  has the same performance statistically as an algorithm using  $\hat{\mathbf{R}}$ , since we average in equation (35) only values with the same distribution. Nevertheless, we have not reduced the computational complexity for the iterations so far.

Let us consider a circulant matrix  $\mathbf{C} = \text{circ}\{\mathbf{c}\}$ , it is completely determined by its first column  $\mathbf{c} \in \mathbb{C}^{M \times 1}$ , each column of  $\mathbf{C}$  is equal to the previous column rotated downwards by one element. Hence the circulant matrix  $\mathbf{C}$  contains every element of its first column  $M$  times.

Let us embed the smoothed covariance matrix  $\hat{\mathbf{R}}_s = \text{toep}(\hat{\mathbf{r}}, \hat{\mathbf{r}}^H)$  in a circulant matrix  $\hat{\mathbf{R}}_c$ , in such a way, that every information of  $\hat{\mathbf{r}}$  is equally often represented as in the smoothed covariance matrix itself

$$\hat{\mathbf{R}}_c = \text{circ} \left( \frac{1}{2M-1} \begin{bmatrix} \begin{bmatrix} M & 0 & \dots & \dots & 0 \\ 0 & M-1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \\ 0 & \dots & \dots & 0 & 1 \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & M-1 & 0 & \dots & 0 \end{bmatrix} \cdot \hat{\mathbf{r}} \\ \vdots \\ \begin{bmatrix} 0 & \dots & \dots & 0 & 1 \\ 0 & \dots & \dots & 0 & 1 \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & M-1 & 0 & \dots & 0 \end{bmatrix} \cdot \hat{\mathbf{r}}^* \end{bmatrix} \right). \quad (36)$$

Now we recall that the eigenvectors of an arbitrary circulant matrix are given by the DFT-matrix. In other words a circulant matrix is diagonalized by the DFT-matrix  $\mathbf{F}$

$$\text{diag}\{\hat{\mathbf{y}}_c\} = \mathbf{F}^H \cdot \hat{\mathbf{R}}_c \cdot \mathbf{F}.$$

The eigenvalues  $\hat{\mathbf{y}}_c$  are related to the first column of the circulant matrix  $\hat{\mathbf{R}}_c$  by

$$\hat{\mathbf{y}}_c = \frac{1}{\sqrt{2M-1}} \cdot \mathbf{F}^H \cdot \begin{bmatrix} \begin{bmatrix} M & 0 & \dots & \dots & 0 \\ 0 & M-1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \\ 0 & \dots & \dots & 0 & 1 \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & M-1 & 0 & \dots & 0 \end{bmatrix} \cdot \hat{\mathbf{r}} \\ \vdots \\ \begin{bmatrix} 0 & \dots & \dots & 0 & 1 \\ 0 & \dots & \dots & 0 & 1 \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & M-1 & 0 & \dots & 0 \end{bmatrix} \cdot \hat{\mathbf{r}}^* \end{bmatrix}.$$

So we define the transformation  $\mathbf{b} = \mathcal{T}_c\{\mathbf{a}\}$ , relating the eigenvalues  $\mathbf{b}$  to the vector  $\mathbf{a}$  as

$$\mathbf{b} = \frac{1}{\sqrt{2M-1}} \cdot \mathbf{F}^H \cdot \begin{bmatrix} \begin{bmatrix} M & 0 & \dots & \dots & 0 \\ 0 & M-1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \\ 0 & \dots & \dots & 0 & 1 \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & M-1 & 0 & \dots & 0 \end{bmatrix} \cdot \mathbf{a} \\ \vdots \\ \begin{bmatrix} 0 & \dots & \dots & 0 & 1 \\ 0 & \dots & \dots & 0 & 1 \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & M-1 & 0 & \dots & 0 \end{bmatrix} \cdot \mathbf{a}^* \end{bmatrix} = \mathcal{T}_c\{\mathbf{a}\}. \quad (37)$$

Now replacing the Toeplitz matrices  $\hat{\mathbf{R}}$  and  $\mathbf{R}(\boldsymbol{\theta})$  by their circulant counterparts in the log-likelihood function (10) yields the new cost function to maximize as

$$c_c(\hat{\mathbf{r}}, \boldsymbol{\theta}) = -N \cdot \sum_{m=1}^M \left( \ln(\mathcal{T}_c\{\boldsymbol{\kappa}(\boldsymbol{\theta})\}_m) + \frac{\mathcal{T}_c\{\hat{\mathbf{r}}\}_m}{\mathcal{T}_c\{\boldsymbol{\kappa}(\boldsymbol{\theta})\}_m} \right). \quad (38)$$

The gradient of the new cost function and a approximation of the Hessian can be derived in a similar fashion from (18) and (20) yielding

$$\mathbf{g}(\hat{\mathbf{r}}, \boldsymbol{\theta})_i = \sum_{m=1}^M \left( \left( \frac{\mathcal{T}_c \{\hat{\mathbf{r}}\}_m}{\mathcal{T}_c \{\boldsymbol{\kappa}(\boldsymbol{\theta})\}_m} - 1 \right) \frac{1}{\mathcal{T}_c \{\boldsymbol{\kappa}(\boldsymbol{\theta})\}_m} \mathcal{T}_c \left\{ \frac{\partial}{\partial \theta_i} \boldsymbol{\kappa}(\boldsymbol{\theta}) \right\}_m \right) \quad (39)$$

and

$$\tilde{\mathbf{H}}(\boldsymbol{\theta})_{ik} = \sum_{m=1}^M \left( \frac{\mathcal{T}_c \left\{ \frac{\partial}{\partial \theta_i} \boldsymbol{\kappa}(\boldsymbol{\theta}) \right\}_m \cdot \mathcal{T}_c \left\{ \frac{\partial}{\partial \theta_k} \boldsymbol{\kappa}(\boldsymbol{\theta}) \right\}_m}{\mathcal{T}_c \{\boldsymbol{\kappa}(\boldsymbol{\theta})\}_m^2} \right). \quad (40)$$

In summary the iteration (34) becomes

$$\hat{\boldsymbol{\theta}}_{i+1} = \hat{\boldsymbol{\theta}}_i + \lambda \cdot \tilde{\mathbf{H}}(\hat{\boldsymbol{\theta}}_i)^{-1} \cdot \mathbf{g}(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}_i), \quad (41)$$

and we have reduced the numerical complexity of one iteration by a factor of  $M$  since the complexity for the new iteration is  $O(L \cdot M \cdot \log(M))$  only. One should observe that  $\hat{\mathbf{y}}_c$  can be calculated directly from the observed channel transfer functions  $\mathbf{x}_i$  as follows

$$\hat{\mathbf{y}}_c = \frac{1}{N} \sum_{i=1}^N \left( \mathbf{F}^H \begin{bmatrix} \mathbf{x}_i \\ \mathbf{0}^{(M-1) \times 1} \end{bmatrix} \right) \circ \left( \mathbf{F}^H \begin{bmatrix} \mathbf{x}_i \\ \mathbf{0}^{(M-1) \times 1} \end{bmatrix} \right)^*.$$

What we still need is a starting point for the iteration the initial solution  $\hat{\boldsymbol{\theta}}_1$ . It should be close enough to the optimum solution to ensure convergences. If we process a sequence of channel observations taken over time, we can use the solution from the last preceding observation as a starting point for the actual iteration. But this leaves us still with the problem to determine an initial solution for the very first observation. Therefore we discuss a global search strategy in the next section.

### Global Search Strategies

To determine an initial solution  $\hat{\boldsymbol{\theta}}_1$  we compute an estimate of the power delay profile from  $\hat{\mathbf{R}}$  equation (16) by

$$\hat{\mathbf{y}} = \text{diag} \{ \mathbf{F}^H \cdot \hat{\mathbf{R}} \cdot \mathbf{F} \}. \quad (42)$$

Assuming that the impulse response is observed over a sufficiently long time, an estimate of the noise variance  $\hat{\alpha}_0$  is

$$\hat{\alpha}_0 = \min(\hat{\mathbf{y}}). \quad (43)$$

Furthermore an estimate of  $\hat{\alpha}_1$  can be derived from  $\hat{\mathbf{y}}$  using  $\hat{\alpha}_0$  by

$$\hat{\alpha}_1 = \max(\hat{\mathbf{y}}) - \hat{\alpha}_0. \quad (44)$$

Observing that the first element of  $\hat{\mathbf{r}}$  equation (35) has expected value

$$\mathbb{E}\{\hat{r}_1\} = \boldsymbol{\kappa}(\boldsymbol{\theta})_1 = \frac{\alpha_1}{M \cdot \beta_d} + \alpha_0,$$

we estimate  $\hat{\beta}_d$  using  $\hat{\alpha}_1$  and  $\hat{\alpha}_0$  from  $\hat{\mathbf{r}}$  using

$$\hat{\beta}_d = \frac{\hat{\alpha}_1}{M(\hat{r}_1 - \hat{\alpha}_0)}. \quad (45)$$

Finally we have to determine a initial solution for the base time delay of the diffuse components  $\hat{\tau}_d$ . At this point it is important to note that the determinant of  $\mathbf{R}(\boldsymbol{\theta})$  is independent from the parameter  $\tau_d$ . If we define the matrix valued function

$$\boldsymbol{\Omega}(\nu) = \text{diag}\left\{1 \quad e^{-j2\pi\nu} \quad \dots \quad e^{-j2\pi(M-1)\nu}\right\},$$

we can express  $\mathbf{R}(\boldsymbol{\theta})$  as

$$\begin{aligned} \mathbf{R} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \beta_d \\ \tau_d \end{pmatrix} &= \text{toep} \left( \boldsymbol{\kappa} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \beta_d \\ \tau_d \end{pmatrix}, \boldsymbol{\kappa} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \beta_d \\ \tau_d \end{pmatrix}^H \right) \\ &= \boldsymbol{\Omega}(\tau_d) \cdot \mathbf{R} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \beta_d \\ 0 \end{pmatrix} \cdot \boldsymbol{\Omega}(\tau_d)^H \end{aligned}$$

Since  $\boldsymbol{\Omega}(\tau_d)$  is unitary the equation

$$\det \left( \mathbf{R} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \beta_d \\ \tau_d \end{pmatrix} \right) = \det \left( \mathbf{R} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \beta_d \\ 0 \end{pmatrix} \right), \quad \forall \tau_d$$

holds. This is interesting insofar, that if the log-likelihood function is maximized over  $\tau_d$  only, the term  $\ln(\det(\mathbf{R}(\boldsymbol{\theta})))$  in equation (10) can be neglected and accordingly the term  $\ln(\mathcal{T}_c\{\boldsymbol{\kappa}(\boldsymbol{\theta})\})$  in the cost function (38), leading to the maximization problem

$$\hat{\tau}_d = \arg \max_{\tau_d} \sum_{m=1}^M \left( \frac{\mathcal{T}_c\{\hat{\mathbf{r}}\}_m}{\mathcal{T}_c\left\{\boldsymbol{\kappa}\left(\begin{bmatrix} \hat{\alpha}_0 & \hat{\alpha}_1 & \hat{\beta}_d & \tau_d \end{bmatrix}^T\right)\right\}_m} \right).$$

So we compute at first the elements of the vector  $\mathbf{z}$  from  $\boldsymbol{\kappa}\left(\begin{bmatrix} \hat{\alpha}_0 & \hat{\alpha}_1 & \hat{\beta}_d & 0 \end{bmatrix}^T\right)$

$$z_m = \frac{1}{\mathcal{T}_c\left\{\boldsymbol{\kappa}\left(\begin{bmatrix} \hat{\alpha}_0 & \hat{\alpha}_1 & \hat{\beta}_d & 0 \end{bmatrix}^T\right)\right\}_m}, \quad (46)$$

and after embedding this vector in the circulant matrix  $\mathbf{Z} = \text{circ}\{\mathbf{z}\}$ , the index  $l_{\max}$  of the largest element of  $\mathbf{c}_\tau = \mathbf{Z}^T \cdot \hat{\mathbf{y}}_c$  is an estimate of the base delay of the diffuse components

$$\hat{\tau}_d = \frac{l_{\max} - 1}{2M - 1}. \quad (47)$$

One should observe that the cyclic convolution between  $\mathbf{z}$  and  $\hat{\mathbf{y}}_c$  can be computed efficiently using the FFT by

$$\mathbf{c}_\tau = \mathbf{F}^H \cdot \left( (\mathbf{F} \cdot \hat{\mathbf{y}}_c) \circ (\mathbf{F} \cdot \mathbf{z})^* \right). \quad (48)$$

Equation (48) implies also that the cost function can be calculated with an arbitrary resolution using zero-padding. Altogether, to calculate an initial solution one has to calculate at first  $\hat{\mathbf{y}}$

using equation (42) and then using (43), (44), and (45) an estimate for  $\hat{\alpha}_0$ ,  $\hat{\alpha}_1$ , and  $\hat{\beta}_d$ . Using these estimates in equation (46) yields the vector  $\mathbf{z}$ , which can be used to determine an estimate for  $\hat{\tau}_d$  from equation (48) and (47). All four estimates together yield the initial solution  $\hat{\boldsymbol{\theta}}_1$  for the local search algorithms.

The global search algorithm gives bad results, if the number of independent observations is small and/or if the observations contain strong discrete (specular) components. A way to reduce the error influence, is to decompose the observation into overlapping observations with a smaller bandwidth. This increases the amount of observations, and reduces the influence of strong specular paths, that principle is similar to the subspace (subarray) smoothing algorithm used for the ESPRIT algorithm.

## VI. Joint Estimation of discrete and diffuse Components

In general the observed radio channel contains distributed diffuse scattering components as well as specular components (propagation paths). Since the specular components are mean values or shifts of the distribution  $\mathbf{s}(\boldsymbol{\theta})$  they can be easily incorporated into the data model (7). The only thing we have to consider is that we have only modeled the covariance matrix in the frequency respectively delay domain.

$$p(\mathbf{x}|\boldsymbol{\theta}) = \frac{1}{\pi^M \det(\mathbf{R}(\boldsymbol{\theta}))} e^{-\mathbf{(x-s(\boldsymbol{\theta}))}^H \mathbf{R}(\boldsymbol{\theta})^{-1} \mathbf{(x-s(\boldsymbol{\theta}))}}$$

Since the parameters of the mean  $\mathbf{s}(\boldsymbol{\theta})$  are not parameters of the covariance  $\mathbf{R}(\boldsymbol{\theta})$  the maximization of the log-likelihood function for the parameters of the specular components can still be treated as a least squares problem although it is now a weighted least squares problem. Therefore, the knowledge about the parameters of the distributed diffuse scattering components of the radio channel can be incorporated into existing high resolution channel parameter estimation algorithms, to improve their performance considerably. Appropriate high resolution channel parameter estimation algorithms for the parameters of the specular propagation paths are ESPRIT [4],[13], SAGE [5],[6],[7] and RIMAX [14],[15].

One should note that effective algorithms for the multiplication of the inverse of a Toeplitz matrix with an arbitrary vector have been developed using the FFT [17],[18].

It is clear that the estimation of the parameters of the covariance matrix  $\mathbf{R}(\boldsymbol{\theta})$  and of the mean  $\mathbf{s}(\boldsymbol{\theta})$  should be carried out jointly, since we have to remove the mean value (the specular paths) from the observation to estimate the parameters of the covariance matrix and we need the covariance matrix as a pre-whitening matrix for the estimation of the parameters of the mean. This can be easily implemented if someone uses the SAGE or RIMAX algorithm, simply by alternating between the optimization for the parameters of the covariance matrix and the parameters of the mean.

## VII. Example

The next two figures show the parameter estimation results estimated from a measurement in a street micro cell scenario. The channel has been measured with a RUSK™ channel sounder using an 8-element uniform linear array at the base station (receiver) and an omni-directional antenna at the transmitter. The parameter estimation has been carried out jointly for the parameter of the diffuse as well as the specular components of the observed channel. Figure 1 show the measured impulse response and the impulse response reconstructed from the estimated parameters of the specular paths. Figure 2 show the measured impulse response after removing the specular components estimated and the main diagonal elements of the covariance matrix in the time-domain reconstructed from the estimated parameters of the diffuse components of the observed channel.

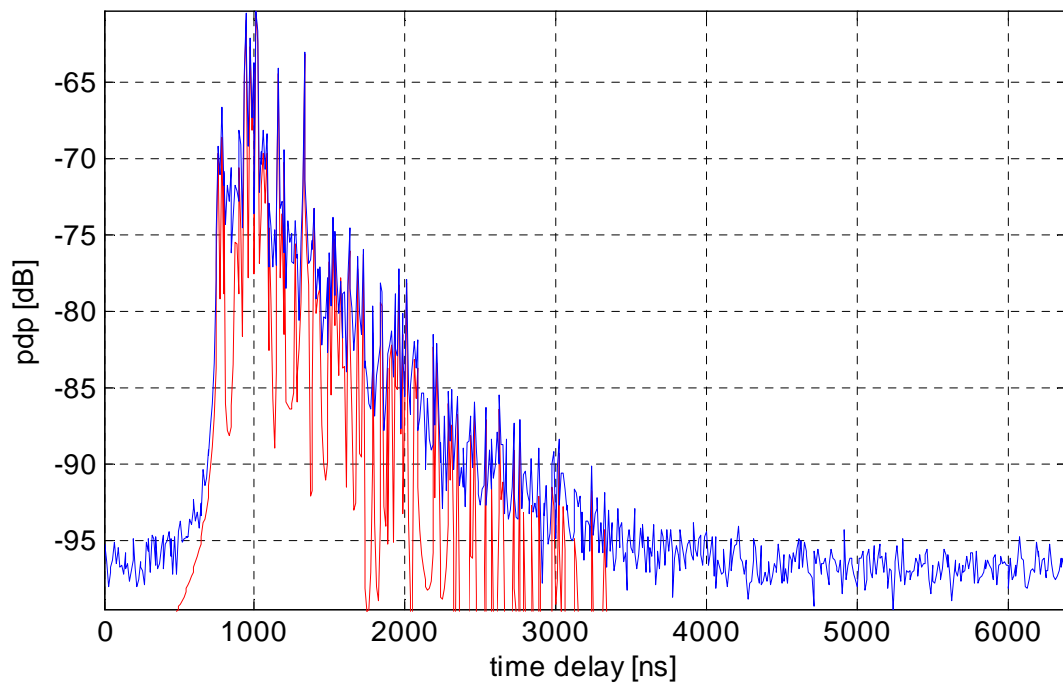


Figure 1: The measured impulse response (blue) and the related impulse response reconstructed from the parameters of the specular propagation paths estimated using SAGE (red).

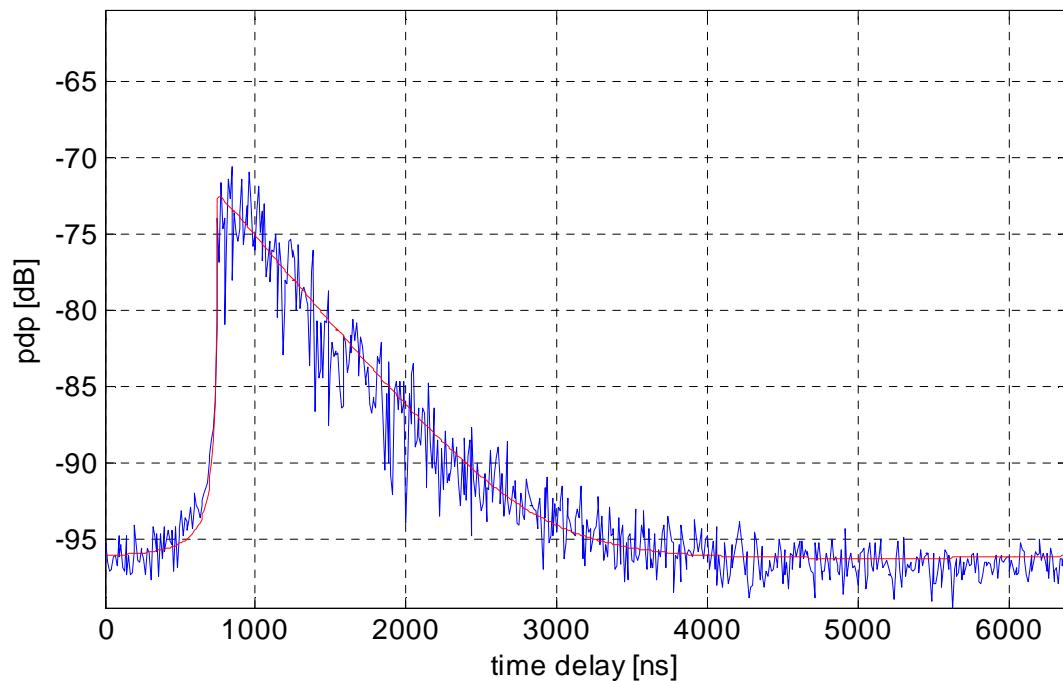


Figure 2: The measured impulse response after removing the specular components (blue) and the main diagonal elements of the estimated covariance matrix in the time domain (red).

## VIII. Further Research

Up to now we have concentrated our research to the time or frequency structure of the covariance matrix of the diffuse components of the observed channel. This is mainly motivated by the fact that the “noise coloring” due to the diffuse components in the frequency domain has a strong impact on the parameter estimation of the specular propagation paths. So the correlation of the diffuse scattering components in the space- respectively angular- as well as the Doppler-domain has to be investigated in further research work.

### References

- [1] K.I. Pedersen, P.E. Mogensen, B.H. Fleury, "A Stochastic Model of the Temporal and Azimuthal Dispersion Seen at the Base Station in Outdoor Environments", IEEE Trans. on Vehicular Technology, Vol. 49, No. 2, March 2000
- [2] Cassioli, D., Win, M.Z., Molisch, A.F., (2002), "The Ultra-Wide Bandwidth Indoor Channel: From Statistical Model to Simulations", IEEE Journal on Selected Areas in Communications, Vol. 20, NO. 6, August 2002, pp 1247-1257
- [3] V. Erceg et.al, "A model for the multipath delay profile of fixed wireless channels," IEEE J. Select Areas Commun., vol. 17, no.3, March 1999, pp. 399-410.
- [4] M. Haardt, "Efficient One-, Two-, and Multidimensional High-Resolution Array Signal Processing," Ph.D. Thesis, Shaker Verlag, Aachen, Germany, 1996.
- [5] J.A. Fessler, A.O. Hero, "Space Alternating Generalized Expectation Maximization Algorithm," IEEE Trans. on Signal Processing, Vol. 42, pp. 2664-2677, October 1994.
- [6] 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.
- [7] 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.
- [8] 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
- [9] 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.
- [10] Louis L. Scharf, "Statistical Signal Processing: detection, estimation, and time series analysis", ISBN 0-201-19038-9, Addison-Wesley
- [11] P. Stoica, A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation", IEEE Trans. on ASSP, Vol. 38, No. 10 pp 1783-1795
- [12] R.S. Thoma, D. Hampicke, A. Richter, G. Sommerkorn, U. Trautwein, "MIMO Vector Channel Sounder Measurement for Smart Antenna System Evaluation", European Transactions on Telecommunications ETT, Vol. 12, No. 5, Special Issue on Smart Antennas, pp. 427-438, Sept./Oct. 2001
- [13] M. Haardt, R.S. Thoma, A. Richter, "Multidimensional High-Resolution Parameter Estimation with Applications to Channel Sounding. In High-Resolution and Robust Signal Processing", Marcel Dekker, 2003
- [14] A. Richter, M. Landmann, R.S. Thoma, "A Gradient Based Method for Maximum Likelihood Channel Parameter Estimation from Multidimensional Channel Sounding Measurements," XXVIIth URSI General Assembly, Maastricht, NL, pp. CD, August 2002
- [15] A. Richter, M. Landmann, R. Thoma, "Maximum Likelihood Channel Parameter Estimation from Multidimensional Channel Sounding Measurements," IEEE VTC2003 Spring, Cheju, South Korea, April 2003
- [16] <http://www.channelsounder.de>

- [17] I. Gohberg, V. Olshevsky. “ Fast algorithms with preprocessing for matrix-vector multiplication problems.”, *Journal of Complexity*, No. 10, 1994, 411-427
- [18] I. Gohberg, I. Feldman, “Convolution equations and projection methods for their solutions”, *Translations of Mathematical Monographs*, 41, Amer. Math. Soc., 1974