Hindawi Publishing Corporation
EURASIP Journal on Advances in Signal Processing
Volume 2007, Article ID 49393, 13 pages
doi:10.1155/2007/49393

Research Article

New Approaches for Channel Prediction Based on Sinusoidal Modeling

Ming Chen,[1] Torbjörn Ekman,[2] and Mats Viberg[1]

[1] *Department of Signals and Systems, Chalmers University of Technology, SE 412 96 Göteborg, Sweden*
[2] *Department of Electronics and Telecommunications, Norwegian Institute of Science and Technology, NO-7491 Trondheim, Norway*

Received 4 December 2005; Revised 4 April 2006; Accepted 30 April 2006

Recommended by Kostas Berberidis

Long-range channel prediction is considered to be one of the most important enabling technologies to future wireless communication systems. The prediction of Rayleigh fading channels is studied in the frame of sinusoidal modeling in this paper. A stochastic sinusoidal model to represent a Rayleigh fading channel is proposed. Three different predictors based on the statistical sinusoidal model are proposed. These methods outperform the standard linear predictor (LP) in Monte Carlo simulations, but underperform with real measurement data, probably due to nonstationary model parameters. To mitigate these modeling errors, a joint moving average and sinusoidal (JMAS) prediction model and the associated joint least-squares (LS) predictor are proposed. It combines the sinusoidal model with an LP to handle unmodeled dynamics in the signal. The joint LS predictor outperforms all the other sinusoidal LMMSE predictors in suburban environments, but still performs slightly worse than the standard LP in urban environments.

1. INTRODUCTION

Link adaption techniques, such as multiuser diversity, adaptive modulation and coding, and fast scheduling hold great promise to improve spectrum efficiency. However, the improvement on the system capacity depends heavily on the predictability of the short-term channel fades [1, 2]. Extensive studies on this topic were made during the last several years by different researchers [1], [3–10]. It was found that a prediction horizon corresponding to a distance of half a wavelength traveled by the mobile is considered challenging [11]. In this paper, we assume the availability of a vector $\mathbf{y} = [y(t), y(t-1), \ldots, y(t-N+1)]^T$ containing the N channel observations (successive estimates of a particular channel coefficient)

$$\mathbf{y} = \mathbf{h} + \mathbf{e}, \tag{1}$$

where $\mathbf{h} = [h(t), h(t-1), \ldots, h(t-N+1)]^T$ is the time evolution of the *true* channel, and $\mathbf{e} = [e(t), e(t-1), \ldots, e(t-N+1)]^T$ is the additive estimation errors with Gaussian distribution $\mathcal{CN}(\mathbf{0}, \sigma_e^2 \mathbf{I}_N)$, where $\mathbf{I}_N$ is the $N \times N$ identity matrix. The value $h(t+L)$ is to be predicted from observations $\mathbf{y}$, where L is the prediction horizon. Such a scenario is presented in Figure 1, where the time index $t = 0$ and the length of the observation interval is $N = 100$. The published channel predictors are divided into two categories, which can be categorized as model-free predictors and model-based predictors, respectively.

1.1. Model-free channel predictors

The first category is essentially the class of linear predictors (LP), where the channel coefficient is predicted as a weighted sum of the previous channel observations. A dth order LP of $h(t+L)$, where $d < N$, is

$$\widehat{h}(t+L) = \sum_{k=0}^{d-1} \beta_k y(t-k) = \boldsymbol{\beta}_d^T \mathbf{y}_d = \mathbf{f}_l^H \mathbf{y}, \tag{2}$$

where

$$\mathbf{f}_l^H = [\boldsymbol{\beta}_d^T \mathbf{0}_{1xN-d}] \tag{3}$$

is the coefficient vector of the LP.

A large number of algorithms can be found in the literature to estimate $\boldsymbol{\beta}$ under various optimization criteria, for

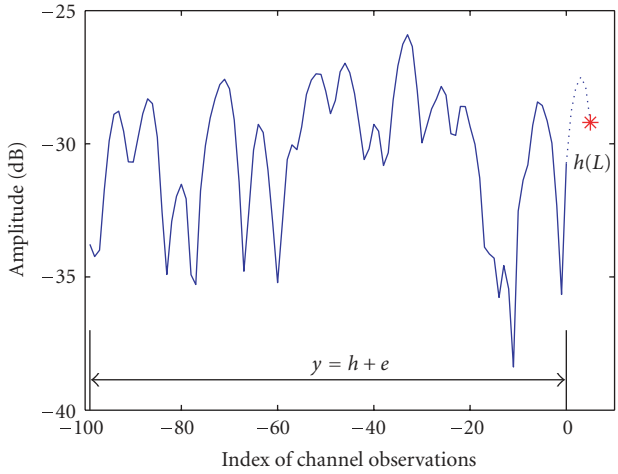

Figure 1: Prediction of Rayleigh fading channel.

instance, the least squares (LS) estimate, which is

$$\hat{\boldsymbol{\beta}}_d = \left(\mathbf{Y}_{ls}^H \mathbf{Y}_{ls}\right)^{-1} \mathbf{Y}_{ls}^H \mathbf{y}_{ls} = \mathbf{Y}_{ls}^{\dagger} \mathbf{y}_{ls}, \tag{4}$$

where

$$\mathbf{y}_{ls} = \left[y(t), y(t-1), \ldots, y(t-N+L+d)\right]^T \tag{5}$$

and the matrix $\mathbf{Y}_{ls}$ is a *Hankel* matrix, which is

$$\mathbf{Y}_{ls} = \begin{bmatrix} y(t-L) & y(t-L-1) & \cdots & y(t-L-d+1) \\ y(t-L-1) & y(t-L-2) & \cdots & y(t-L-d) \\ \vdots & \vdots & \ddots & \vdots \\ y(t-N+d) & y(t-N+d-1) & \cdots & y(t-N+1) \end{bmatrix}. \tag{6}$$

1.2. Model-based channel predictors

In the second category, the channel over a short observation interval is modeled as superimposed deterministic complex sinusoids, which correspond to the Doppler frequencies as in the Jakes model [12],

$$y(t) = h(t) + e(t) = \sum_{i=1}^{p} \rho_i e^{j\phi_i} e^{j\omega_i t} + e(t), \tag{7}$$

where p is the number of sinusoids, ρ_i, ϕ_i, and ω_i are the real amplitude, phase, and the Doppler frequency associated with the ith path, respectively, and $e(t)$ is the additive complex Gaussian estimation error, $e(t) \sim \mathcal{CN}(0, \sigma_e^2)$. Assuming there are no moving reflecting objects, the ith Doppler frequency is

$$\omega_i = \frac{2\pi v}{\lambda} \cos\theta_i, \tag{8}$$

where v is the velocity of the mobile terminal, and θ_i is the angle between the ith impinging ray and the mobile heading. In this model, there are $(3p+1)$ parameters to be estimated. They are

$$\boldsymbol{\psi} = \left[\boldsymbol{\rho}^T, \boldsymbol{\phi}^T, \boldsymbol{\omega}^T, \sigma_e^2\right]^T, \tag{9}$$

where

$$\begin{aligned} \boldsymbol{\rho} &= \left[\rho_1, \ldots, \rho_p\right]^T, \qquad \boldsymbol{\phi} = \left[\phi_1, \ldots, \phi_p\right]^T, \\ \boldsymbol{\omega} &= \left[\omega_1, \ldots, \omega_p\right]^T. \end{aligned} \tag{10}$$

These model parameters are assumed to be stationary over the observation interval. In these methods, the first step is to compute parameter estimates of $\boldsymbol{\rho}$, $\boldsymbol{\phi}$, and $\boldsymbol{\omega}$. Then, the straightforward prediction of $h(t+L)$ is

$$\hat{h}(t+L) = \sum_{i=1}^{p} \hat{\rho}_i e^{j\hat{\phi}_i} e^{j\hat{\omega}_i (t+L)}. \tag{11}$$

Let $s_i = \rho_i e^{j\phi_i}$ be the complex amplitude. Then, (7) can be written as

$$y(t) = \sum_{i=1}^{p} s_i e^{j\omega_i t} + e(t), \tag{12}$$

and the prediction of $h(t+L)$ is

$$\hat{h}(t+L) = \sum_{i=1}^{p} \hat{s}_i e^{j\hat{\omega}_i (t+L)}. \tag{13}$$

In vector form, the N observations can be represented as

$$\mathbf{y} = \mathbf{A}\mathbf{s} + \mathbf{e}, \tag{14}$$

where

$$\mathbf{A} = \left[\mathbf{a}_1, \ldots, \mathbf{a}_p\right], \qquad \mathbf{a}_k = \left[e^{j\omega_k t}, \ldots, e^{j\omega_k (t-N+1)}\right]^T, \tag{15}$$

$$\mathbf{s} = \left[s_1, \ldots, s_p\right]^T. \tag{16}$$

A thorough performance evaluation of linear predictions and channel predictions based on deterministic sinusoidal modeling were made by simulations and real world data in [6]. According to the reported results, the LP outperforms the deterministic sinusoidal modeling based methods. All these studies were performed in single-in single-out (SISO) systems. Later, similar results were reported in multiple-in multiple-out (MIMO) scenarios in [13]. It was also claimed that the calculation complexity of LP increases exponentially with the increase of the dimensions of the MIMO systems. This makes LP costly for high dimension MIMO channel prediction. This is one of the motivations to pursue channel prediction methods based on sinusoidal modeling, where a fewer number of model parameters is expected than the LP in MIMO scenarios.

Beside these prediction methods, a nonlinear prediction of mobile radio channel using the MARS modeling was studied in [14]. In this study, the channel prediction is studied in the frame of sinusoidal modeling of the fading channel. First, a statistical sinusoidal model of the Rayleigh fading channel is proposed. Based on this model, three sinusoidal LMMSE channel predictors are given. Later a joint moving average and sinusoidal (JMAS) channel predictor is proposed, which alleviates the influence of the nonsinusoidal model components and leads to the Joint LS predictors.

In this paper, the statistical sinusoidal model and the extensions into single-in multiple-out (SIMO) and MIMO systems are presented in Section 2. Three sinusoidal LMMSE predictors are given in Section 3. The JMAS prediction and Joint LS predictor are presented in Section 4. Section 5 presents an analysis of the measurement data and the stationarity properties of model parameters. The performance evaluations of these sinusoidal modeling-based methods by Monte Carlo simulations and real measurement data are presented in Section 6. Section 7 contains the conclusions.

2. STATISTICAL SINUSOIDAL MODELING

In SISO scenarios, the stochastic sinusoidal modeling of radio channel has the same form as the deterministic sinusoidal model in (14), but the complex amplitudes of the sinusoids $\mathbf{s}$ are modeled as zero-mean random variables with covariance matrix, $E[\mathbf{s}\mathbf{s}^H] = \sigma_s^2\mathbf{I}_p$. This extension leads to a number of LMMSE channel predictors, which are presented in the following section.

The statistical sinusoidal model for a SIMO system, with one transmit antenna and m receive antennas, is

$$\mathbf{Y} = \mathbf{A}\mathbf{S} + \mathbf{E}, \tag{17}$$

where

$$\begin{aligned} \mathbf{Y} &= \left[\mathbf{y}^T(t), \mathbf{y}^T(t-1), \ldots, \mathbf{y}^T(t-N+1)\right]^T, \\ \mathbf{S} &= \left[\mathbf{s}_1^T, \mathbf{s}_2^T, \ldots, \mathbf{s}_p^T\right]^T, \\ \mathbf{E} &= \left[\mathbf{e}^T(t), \mathbf{e}^T(t-1), \ldots, \mathbf{e}^T(t-N+1)\right]^T, \end{aligned} \tag{18}$$

where the m-vector $\mathbf{y}(t)$ is the channel observations at the sensor array at time t. Matrix $\mathbf{S}$ is $p \times m$, and the m-vector $\mathbf{s}_i$ is one realization of the complex random amplitude vector associated to the ith path. Assume these sinusoids have equal mean power σ_s^2, and $E[\mathbf{S}\mathbf{S}^H] = m\sigma_s^2\mathbf{I}_p$. The m-vector $\mathbf{e}(t)$ is the additive noise with covariance matrix $E[\mathbf{e}^H(k)\mathbf{e}(l)] = \sigma_e^2\mathbf{I}_m\delta_{k,l}$. Note that we have assumed that the angle of arrivals, $\boldsymbol{\theta}$, are identical at different antenna elements, but with independent amplitudes. In practice, inaccurate calibration of the array might give rise to such uncorrelated amplitudes.

To derive the signal model for a MIMO system with n transmit antennas and m receive antennas, first let the SIMO channel, associated to the kth transmit antenna, be

$$\mathbf{H}_k = \mathbf{A}\mathbf{S}_k, \tag{19}$$

where $\mathbf{A}$ and $\mathbf{S}_k$ are the same as in (17), except the subscripts. Vectorizing $\mathbf{H}_k$, we have

$$\begin{aligned} \mathbf{h}_k = \mathrm{vec}(\mathbf{H}_k) &= \mathrm{vec}(\mathbf{A}\mathbf{S}_k) = (\mathbf{I}_m \otimes \mathbf{A})\mathrm{vec}(\mathbf{S}_k), \\ &= (\mathbf{I}_m \otimes \mathbf{A})\mathfrak{s}_k, \end{aligned} \tag{20}$$

where $\otimes$ is the Kronecker product and $\mathrm{vec}(\cdot)$ is the vectorization operation [15]. Stacking $\mathbf{h}_k$, the sinusoidal signal model for the MIMO channel is

$$\mathcal{H} = \left[\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_n\right] = \left[\mathbf{I}_m \otimes \mathbf{A}\right]\left[\mathfrak{s}_1, \mathfrak{s}_2, \ldots, \mathfrak{s}_n\right] = \mathcal{A}\mathcal{S}, \tag{21}$$

where the matrices $\mathcal{H}$ are $(Nm \times n)$ and $\mathcal{S}$ is $(pm \times n)$, respectively. The power matrix is $E[\mathcal{S}\mathcal{S}^H] = n\sigma_s^2\mathbf{I}_{pm}$, where we have assumed, again, that the amplitudes are independent with equal mean power σ_s^2. Including channel estimation errors, the observed channel based on sinusoidal modeling for MIMO system becomes

$$\mathcal{Y} = \mathcal{A}\mathcal{S} + \mathcal{E}. \tag{22}$$

Note that we have assumed here that all mn subchannels share the same Doppler frequencies. The extension to the general case is obvious, but introduces considerably more unknown parameters to be estimated.

In the following sections, we restrict our study within SISO scenarios.

3. LMMSE PREDICTORS

In this paper we are not concerned with the frequency estimation problem, but focus on prediction only. A number of successful frequency estimation methods have been published in the literature [16–19]. The class of subspace methods provides high-performance estimators at low cost, and of these we have chosen the unitary ESPRIT method due to [20].

3.1. Conditional LMMSE channel predictor

Assume that the frequency estimates $\widehat{\boldsymbol{\omega}} = [\widehat{\omega_1}, \ldots, \widehat{\omega_p}]$ are given, the LMMSE estimate of $\mathbf{s}$ is then [4]

$$\widehat{\mathbf{s}} = \left(\widehat{\mathbf{A}}^H\widehat{\mathbf{A}} + \alpha\mathbf{I}\right)^{-1}\widehat{\mathbf{A}}^H\mathbf{y} = \mathbf{R}_{\mathrm{reg}}^{-1}\widehat{\mathbf{A}}^H\mathbf{y}, \tag{23}$$

where $\widehat{\mathbf{A}}$ is defined as $\mathbf{A}$, but using the estimated frequencies $\widehat{\omega}_k$ instead of the true frequencies, and $\alpha = \sigma_e^2/\sigma_s^2 = 1/SNR$. This can also be interpreted as a regularized LS estimator. Moreover, the LMMSE prediction of $h(t+L)$ (given that $\widehat{\boldsymbol{\omega}}$ is assumed to be correct) is given by

$$\begin{aligned} \widehat{h}(t+L) &= \left[e^{j\widehat{\omega}_1(t+L)}, \ldots e^{j\widehat{\omega_p}(t+L)}\right]\widehat{\mathbf{s}} \\ &= \widehat{\mathbf{a}}(L)^H\mathbf{R}_{\mathrm{reg}}^{-1}\widehat{\mathbf{A}}^H\mathbf{y} = \mathbf{f}_c^H\mathbf{y}, \end{aligned} \tag{24}$$

where $\widehat{\mathbf{a}}(L)^H = \left[e^{j\widehat{\omega}_1(t+L)}, \ldots, e^{j\widehat{\omega}_p(t+L)}\right]$,

$$\mathbf{f}_c^H = \widehat{\mathbf{a}}(L)^H\mathbf{R}_{\mathrm{reg}}^{-1}\widehat{\mathbf{A}}^H \tag{25}$$

is the conditional prediction filter. If the complex amplitudes, $\mathbf{s}$, are modeled as Gaussian random variables, $\mathcal{CN}(\mathbf{0}, \sigma_s^2\mathbf{I}_N)$, the statistical sinusoidal signal model (14) is termed *Bayesian general linear model* in [21], and the conditional LMMSE predictor is also the conditional MMSE predictor [10].

3.2. Adjusted conditional LMMSE channel predictor

Above, the frequency estimates $\widehat{\boldsymbol{\omega}}$ were regarded as exact. In practice, $\widehat{\boldsymbol{\omega}}$ is estimated by some method, such as unitary ESPRIT [20], and subject to errors. The error variance, in theory, is determined by the SNR and the number of observations [21]. Such a problem was investigated in [22, 23]. Define the residual signal $\boldsymbol{\epsilon}$=$[\epsilon(t), \ldots, \epsilon(t-N+1)]^T$ as the part

of the observations which cannot be accurately reconstructed by the estimated sinusoidal model,

$$\boldsymbol{\epsilon} = \mathbf{y} - \widehat{\mathbf{h}}, \tag{26}$$

where $\widehat{\mathbf{h}} = \widehat{\mathbf{A}}\widehat{\mathbf{s}}$. In the imperfect modeling case, a low-order LP is required to predict the colored residuals $\epsilon(t+L)$, which contain less signal dynamics than $\mathbf{y}$. So the adjusted conditional LMMSE predictor is

$$\widehat{h}_{\text{adj}}(t+L) = \widehat{h}(t+L) + \widehat{\epsilon}(t+L). \tag{27}$$

The predictor $\widehat{\epsilon}(t+L)$ is computed based on past residuals as

$$\widehat{\epsilon}(t+L) = \sum_{k=0}^{d-1} \beta_k \epsilon(t-k). \tag{28}$$

The coefficients $\boldsymbol{\beta}_d = [\beta_0, \ldots, \beta_{d-1}]^T$ are computed by solving an LS problem. To explicitly show the dependency of $\widehat{\epsilon}(t+L)$ on the channel observation $\mathbf{y}$, we rewrite (28) as

$$\widehat{\epsilon}(t+L) = [\boldsymbol{\beta}_d^T \; \mathbf{0}_{N-d}^T](\mathbf{I}_N - \widehat{\mathbf{A}}\mathbf{R}_{\text{reg}}^{-1}\widehat{\mathbf{A}}^H)\mathbf{y} = \mathbf{f}_\epsilon^H \mathbf{y}, \tag{29}$$

where $\mathbf{0}_{N-d}$ is an $(N-d)$ zero vector, and $\mathbf{f}_\epsilon^H$ is the residual predictor. Then, the adjusted conditional LMMSE predictor (27) can be expressed as

$$\widehat{h}_{\text{adj}}(t+L) = (\mathbf{f}_c^H + \mathbf{f}_\epsilon^H)\mathbf{y} = \mathbf{f}_a^H \mathbf{y}, \tag{30}$$

where $\mathbf{f}_a^H$ is the adjusted conditional LMMSE prediction filter.

3.3. Unconditional LMMSE predictor

Another approach to combat the frequency estimation error is motivated by introducing the frequency estimation uncertainty into the conditional LMMSE predictions. We can model ω_k as a random variable with mean $\widehat{\omega}_k$ and variance $E[\Delta\omega_k^2] = \sigma_{\Delta\omega_k}^2$,

$$\omega_k = \widehat{\omega}_k + \Delta\omega_k. \tag{31}$$

The classical form of the LMMSE predictor is given by the Wiener filter [21],

$$\widehat{h}(t+L) = \mathbf{R}_{hy}\mathbf{R}_{yy}^{-1}\mathbf{y}, \tag{32}$$

where

$$E\left\{\begin{bmatrix} h(t+L) \\ \mathbf{y} \end{bmatrix}\begin{bmatrix} h(t+L)^H & \mathbf{y}^H \end{bmatrix}\right\} = \begin{bmatrix} \sigma_h^2 & \mathbf{R}_{hy} \\ \mathbf{R}_{yh} & \mathbf{R}_{yy}l \end{bmatrix}. \tag{33}$$

The covariance matrix for N observations is

$$\mathbf{R}_{yy} = E\{\mathbf{y}\mathbf{y}^H\} = \mathbf{R}_{hh} + \mathbf{R}_{ee}. \tag{34}$$

Under the assumption that the amplitudes and frequency estimation errors are independent stochastic variables, we can write the (m,n)th element of $\mathbf{R}_{hh}$ as

$$\begin{aligned}[\mathbf{R}_{hh}]_{mn} &= E\left\{\left[\mathbf{A}\mathbf{s}\mathbf{s}^H \quad \mathbf{A}^H\right]_{mn}\right\} \\ &= \sum_{k=1}^{p} E\{|s_k|^2\} E\{e^{j(\widehat{\omega}_k+\Delta\omega_k)(n-m)}\}.\end{aligned} \tag{35}$$

The expectation over the frequency estimation error can be expressed as

$$\begin{aligned}E\{e^{j(\widehat{\omega}_k+\Delta\omega_k)(n-m)}\} &= e^{j\widehat{\omega}_k(n-m)} E\{e^{j\Delta\omega_k(n-m)}\} \\ &= e^{j\widehat{\omega}_k(n-m)}\Phi_{\Delta\omega_k}(n-m),\end{aligned} \tag{36}$$

where $\Phi_{\Delta\omega_k}(\tau)$ is the characteristic function of the frequency estimation error $\Delta\omega_k$. If we assume the frequency errors to be Gaussian distributed, then

$$\Phi_{\Delta\omega_k}(n-m) = e^{-(\sigma_{\Delta\omega_k}^2/2)(n-m)^2}. \tag{37}$$

The expectation over the ensemble of amplitudes is just the variance

$$E\{|s_k|^2\} = \sigma_s^2. \tag{38}$$

The (m,n)th element of the covariance matrix is thus obtained as

$$[\mathbf{R}_{hh}]_{mn} = \sum_{k=1}^{p} \sigma_s^2 e^{j\widehat{\omega}_k(n-m)} e^{-(\sigma_{\Delta\omega_k}^2/2)(n-m)^2}. \tag{39}$$

For simplicity, let all the frequency errors be IID with $\sigma_{\Delta\omega_k}^2 = \sigma_{\Delta\omega}^2$. We then have

$$[\mathbf{R}_{hh}]_{mn} = \sigma_s^2\left[\sum_{k=1}^{p} e^{j\widehat{\omega}_k(n-m)}\right] e^{-(\sigma_{\Delta\omega}^2/2)(n-m)^2}. \tag{40}$$

Define the $(N \times N)$ damping matrix $\boldsymbol{\Gamma}$ by

$$[\boldsymbol{\Gamma}]_{mn} = e^{-(\sigma_{\Delta\omega}^2/2)(n-m)^2}. \tag{41}$$

The covariance matrix can then be expressed as

$$\mathbf{R}_{hh} = \sigma_s^2 \widehat{\mathbf{A}}\widehat{\mathbf{A}}^H \odot \boldsymbol{\Gamma}, \tag{42}$$

where $\odot$ denotes the Hadamard product. Similarly, the cross covariance $\mathbf{R}_{hy}$ is given by

$$\begin{aligned}\mathbf{R}_{hy} = E\{\widehat{\mathbf{a}}(L)^H \mathbf{s}\mathbf{s}^H \mathbf{A}^H\} &= \sum_{k=1}^{p} E\{|s_k|^2\} \\ &\cdot E\{[e^{j(\widehat{\omega}_k+\Delta\omega_k)L}, \ldots, e^{j(\widehat{\omega}_k+\Delta\omega_k)(L+N-1)}]\}.\end{aligned} \tag{43}$$

Define the damping vector

$$\boldsymbol{\gamma} = [e^{(-\sigma_{\Delta\omega}^2/2)L^2}, \ldots, e^{(-\sigma_{\Delta\omega}^2/2)(L+N-1)^2}]. \tag{44}$$

The cross covariance is then obtained as

$$\begin{aligned}\mathbf{R}_{hy} &= \sigma_s^2 \sum_{k=1}^{p} \{[e^{j\widehat{\omega}_k L}, \ldots, e^{j\widehat{\omega}_k(L+N-1)}]\} \odot \boldsymbol{\gamma} \\ &= \sigma_s^2 \widehat{\mathbf{a}}(L)^H \widehat{\mathbf{A}}^H \odot \boldsymbol{\gamma}.\end{aligned} \tag{45}$$

The unconditional LMMSE predictor is finally given by (just some variation)

$$\begin{aligned}\widehat{h}(t+L) &= \mathbf{R}_{hy}\mathbf{R}_{yy}^{-1}\mathbf{y} \\ &= (\widehat{\mathbf{a}}(L)^H \widehat{\mathbf{A}}^H \odot \boldsymbol{\gamma})(\widehat{\mathbf{A}}\widehat{\mathbf{A}}^H \odot \boldsymbol{\Gamma} + \alpha\mathbf{I}_N)^{-1}\mathbf{y} \\ &= \mathbf{f}_u^H \mathbf{y},\end{aligned} \tag{46}$$

where

$$\mathbf{f}_u^H = \left(\hat{\mathbf{a}}(L)^H \hat{\mathbf{A}}^H \odot \boldsymbol{\gamma}\right)\left(\hat{\mathbf{A}}\hat{\mathbf{A}}^H \odot \boldsymbol{\Gamma} + \alpha \mathbf{I}_N\right)^{-1} \tag{47}$$

is the unconditional prediction filter. Clearly, the influence of old observations is reduced by the damping matrix $\boldsymbol{\Gamma}$ and the damping vector $\boldsymbol{\gamma}$, which are also dependent on the prediction range. The further ahead we are looking, the smaller the gain of the predictor is. The way the frequency error is taken into account can be interpreted as a convolution of the line spectrum of the signal with the error distribution. The filter design is thus done for distributed sources. The errors in the frequency estimates force the predictor to mistrust older data, as even a small frequency error can cause a large phase error if one waits long enough. As the special case, when $\sigma_{\Delta\omega}^2 = 0$, (46) degenerates to (24). When $\sigma_{\Delta\omega}^2 > 0$, (46), although being LMMSE (ignoring the influence of s_k on the estimate of ω_k), is not the MMSE due to the nonlinear dependency on the frequency estimates.

4. JMAS PREDICTION MODEL AND JOINT LS PREDICTOR

In order to represent signals with more general spectral characteristics, let the estimate of $h(t+L)$ be modeled by

$$\hat{h}(t+L) = \mathbf{y}_d^T \boldsymbol{\beta}_d + \hat{\mathbf{a}}(L)^H \mathbf{s} + e(t+L), \tag{48}$$

where $\hat{\boldsymbol{\omega}}$ in $\hat{\mathbf{a}}(L)$ contains only the frequencies of the stationary sinusoidal signals, which are determined by some model selection method. With this model, the channel is divided into two parts. One part contains the consistent sinusoidal signals, such as the direct rays in LOS scenarios. The second part captures all the remaining power in the observations. This prediction model is interpreted as the joint moving average and sinusoidal (JMAS) predictor. Note that the name refers to the predictor, but not to the signal model. Selecting which sinusoids to use in the predictor is not a trivial problem, and deserves further study. A possibility is to use frequency tracking and apply some stationarity measure for each frequency track.

Due to the linearity of the autoregressive and sinusoidal bases, the model parameters, $\boldsymbol{\theta} = [\boldsymbol{\beta}_d^T \;\; \mathbf{s}^T]^T$, can be estimated jointly. Given $\hat{\boldsymbol{\omega}}$, (48) can be rewritten as

$$\hat{h}(t+L) = \left[\mathbf{y}_d^T \;\; \hat{\mathbf{a}}(L)^H\right]\begin{bmatrix}\boldsymbol{\beta}_d \\ \mathbf{s}\end{bmatrix}. \tag{49}$$

In vector form, we have

$$\hat{\mathbf{h}}_{N-L-d+1} = \left[\mathbf{Y}_{ls} \; \hat{\mathbf{A}}_J\right]\boldsymbol{\beta} = \mathbf{H}\boldsymbol{\theta}, \tag{50}$$

where

$$\begin{aligned}\hat{\mathbf{h}}_{N-L-d+1} &= \left[\hat{h}(t), \hat{h}(t-1), \ldots, \hat{h}(t+L-N+d)\right]^T, \\ \hat{\mathbf{A}}_J &= \left[\hat{\mathbf{a}}_{J1}, \hat{\mathbf{a}}_{J2}, \ldots, \hat{\mathbf{a}}_{Jp}\right], \\ \hat{\mathbf{a}}_{Jk} &= \left[e^{j\hat{\omega}_k(t+L)}, e^{j\hat{\omega}_k(t+L-1)}, \ldots, e^{j\hat{\omega}_k(t+2L-N+d)}\right]^T.\end{aligned} \tag{51}$$

To take into account the prediction model error, (50) can be rewritten as

$$\begin{aligned}\mathbf{y}_{ls} &= \hat{\mathbf{h}}_{N-L-d+1} + \mathbf{u} \\ &= \mathbf{H}\boldsymbol{\theta} + \mathbf{u},\end{aligned} \tag{52}$$

where $\mathbf{u} = [u(t), u(t-1), \ldots, u(t-N+L+d)]^T$ with Gaussian distribution $\mathcal{CN}(\mathbf{0}, \sigma_u^2 \mathbf{I}_{N-L-d+1})$. Then, the LS estimate of $\boldsymbol{\theta}$ is

$$\hat{\boldsymbol{\theta}} = \left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{H}^H\mathbf{y}_{ls}. \tag{53}$$

The channel prediction is

$$\hat{h}(t+L) = \left[\mathbf{y}_d^T \;\; \hat{\mathbf{a}}(L)^H\right]\hat{\boldsymbol{\theta}}, \tag{54}$$

which is named Joint LS predictor. In the terminology of the previous sections, it is a conditional LS predictor, since $\hat{\boldsymbol{\omega}}$ is assumed to be given and $\mathbf{s}$ is modeled as deterministic.

Finally, substitute (53) into (54) to obtain

$$\begin{aligned}\hat{h}(t+L) &= \left[\mathbf{y}_d^T \;\; \hat{\mathbf{a}}(L)^H\right]\left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{H}^H\mathbf{y}_{N-L-d+1} \\ &= \left[\mathbf{y}_d^T \;\; \hat{\mathbf{a}}(L)^H\right]\left[\left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{H}^H \; \mathbf{0}_M\right]\mathbf{y} = \mathbf{f}_J^H\mathbf{y},\end{aligned} \tag{55}$$

where the $\mathbf{0}_M$ is a zero matrix of conformable dimensions, and

$$\mathbf{f}_J^H = \left[\mathbf{y}_d^T \;\; \hat{\mathbf{a}}(L)^H\right]\left[\left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{H}^H \; \mathbf{0}_M\right] \tag{56}$$

is the joint LS prediction filter. Note that this is a nonlinear predictor, since $\mathbf{H}$ depends on $\mathbf{y}$.

The difference between the joint LS method and the adjusted conditional LMMSE predictor in Section 3.2 can be seen by rewriting (26) as

$$\begin{aligned}\boldsymbol{\epsilon} &= \mathbf{y} - \hat{\mathbf{A}}\left(\hat{\mathbf{A}}^H\hat{\mathbf{A}}\right)^{-1}\hat{\mathbf{A}}^H\mathbf{y} \\ &= \left(\mathbf{I}_N - \hat{\mathbf{A}}\left(\hat{\mathbf{A}}^H\hat{\mathbf{A}}\right)^{-1}\hat{\mathbf{A}}^H\right)\mathbf{y} = \boldsymbol{\Pi}_{\hat{A}}^{\perp}\mathbf{y}.\end{aligned} \tag{57}$$

So that (27) becomes

$$\begin{aligned}\hat{h}_{\text{adj}}(t+L) &= \left[\boldsymbol{\epsilon}_d^T \;\; \hat{\mathbf{a}}(L)^H\right]\begin{bmatrix}\boldsymbol{\beta}_d \\ \mathbf{s}\end{bmatrix} \\ &= \left[\left(\left[\mathbf{1}_d^T \;\; \mathbf{0}_{N-d}^T\right]^T \odot \boldsymbol{\Pi}_{\hat{A}}^{\perp}\mathbf{y}\right)^T \;\; \hat{\mathbf{a}}(L)^H\right]\boldsymbol{\theta},\end{aligned} \tag{58}$$

where $\mathbf{1}_d^T$ is a d-vector with all ones. It can be seen that the autoregressive bases are orthogonal to the sinusoidal bases in (58), but they are not in (49).

In general, the total order of the joint LS prediction, $r = p + d$, can be determined by some classical criterion, that is, AIC [24] and MDL [25]. So a wide sense definition of *model selection* includes the selection of the stationary sinusoidal part of the signals, p, and the order of the LP, d, or both at once. In this paper, the *model selection* refers to the selection of the strong and stationary sinusoidal signals only.

In practice, it is probably a wise idea to overestimate the rank, and estimate unnecessarily many sinusoids. But only a subset of these is used in the predictor. The tricky part is to know which ones to keep. With only one data set, one would probably go for the ones with the largest estimated

amplitude. With several data sets, one could try to evaluate the stationarity of the various frequency estimates, but this requires a clever sorting (data association) to create frequency tracks.

5. ANALYSIS OF MEASUREMENT DATA

5.1. Measurement data

A measurement campaign was performed in the urban area of Stockholm, and the suburban area, Kista, a few kilometers north of central Stockholm. The measurements were performed at 1880 MHz, and the bandwidth is 5 MHz. In total, 45 and 31 effective measurements were performed at different urban and suburban locations, respectively. The velocities of the mobile station were between 30 to 90 km per hour during the measurements, except for a few stand still cases. Each measurement records channel sounding data over 156.4 ms and contains 1430 repetitions of channel impulse responses. This gives rise to the channel update rate of 9.1 KHz. Each channel impulse response (or power delay profile) is described by 120 taps with a sampling frequency of 6.4 MHz. The time interval between two neighboring taps is 0.156 μs.

In this study, the strongest channel tap in each measurement is used in the performance evaluations of the proposed channel prediction methods, which can be considered as a narrow band Rayleigh fading channel.

Obviously, the performance of the sinusoidal modeling-based channel prediction methods depends heavily on the stationarity properties of model parameters in the observation window and the prediction horizon. In this section, the stationarity of the model parameters is investigated by means of the short-term MUSIC pseudospectrum (MPS) and model parameter tracking.

5.2. MUSIC pseudospectrum

The MUSIC pseudospectrum (MPS) is defined as [16]

$$p_{mu}(\omega) = \frac{\mathbf{a}(\omega)^H \mathbf{a}(\omega)}{\mathbf{a}(\omega)^H \mathbf{\Pi}^{\perp} \mathbf{a}(\omega)}, \quad \omega \in [-f_{\max}, f_{\max}], \tag{59}$$

where $\mathbf{\Pi}^{\perp} = \mathbf{U}_n \mathbf{U}_n^H$, and the columns of $\mathbf{U}_n$ are the eigenvectors spanning the noise space of the covariance matrix of the data sequence $\mathbf{y}$, $\mathbf{R}_{yy} = E[\mathbf{y}\mathbf{y}^H]$. It can be obtained by taking eigenvalue decomposition of $\mathbf{R}_{yy}$ as

$$\begin{aligned} \mathbf{R}_{yy} &= \begin{bmatrix} \mathbf{U_s} & \mathbf{U_n} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_s & \\ & \mathbf{\Lambda}_n \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_n^H \end{bmatrix} \\ &= \mathbf{U}_s \mathbf{\Lambda}_s \mathbf{U}_s^H + \mathbf{U}_n \mathbf{\Lambda}_n \mathbf{U}_n^H, \end{aligned} \tag{60}$$

where a rank estimation algorithm is necessary, and the eigenvalues are ordered in nonincreasing order. In case of high SNR, MPS is not sensitive to under-estimated noise subspace dimension, while an overestimated noise space will attenuate certain weak frequency components. The $\mathbf{a}(\omega)$ is the DFT vector associated with frequency ω, and $f_{\max}$ is the maximum Doppler frequency.

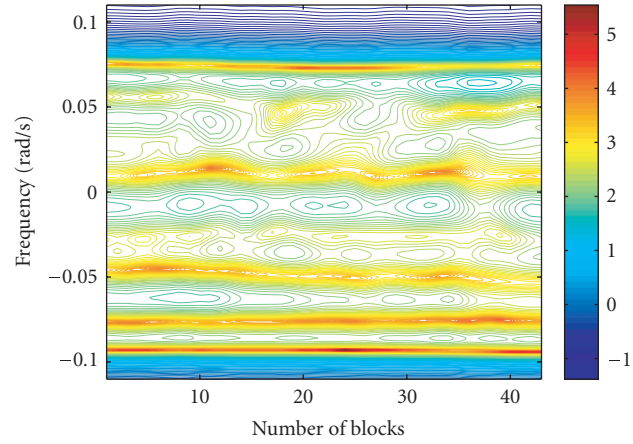

Figure 2: Example of a MUSIC pseudospectrum of a measurement in an urban area. A number of high-power frequency bins can be observed. They are distributed on both the positive and negative sides of the spectrum. But the relative power between these bins and the remaining frequency bins is low. Those bins close to the positive and negative maximum Doppler frequencies are more consistent over the measurements than the others. These bins have blurry edges.

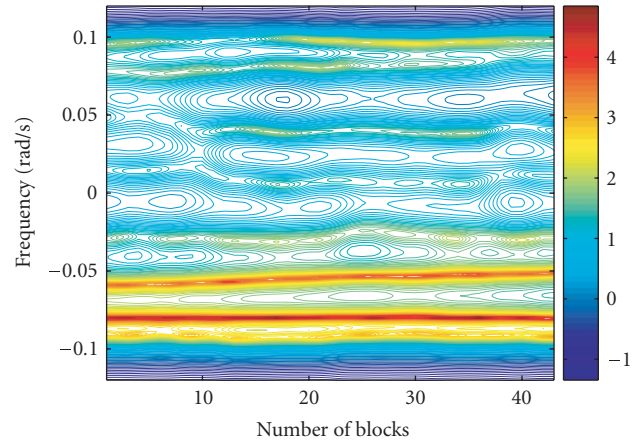

Figure 3: Example of a MUSIC pseudospectrum of a measurement in a suburban area. A fewer number of high-power frequency bins are observed compared to Figure 2. The locations of these frequency bins are biased to the negative side in the spectrum. The edges of these high-power frequency bins are much sharper than those in Figure 2. The power in the strongest frequency bin is much higher than the others and is consistent over the whole measurement interval.

The MPS is calculated over each segmented data block, where a sliding window is applied for data extraction. Examples of typical time varying MPSs in urban and suburban channels are presented in Figures 2 and 3, respectively. In these two measurements, the mobile speeds were around 30 km/h, the original data is downsampled by a factor of 10, and the window length (after downsampling) is 100,

which corresponds to a spatial observation interval of 6λ. The numbers of sinusoids are set to be 8 and 7, respectively, which is assumed to be slightly overestimated. In these figures, the grayscale of the contours indicates the powers in the frequency bins.

In Figure 2, the MPS from an urban measurement is given, where a number of frequency bins with deeper grayscale appear in the spectrum. A number of relatively high-power and stationary frequency bins, close to the positive and negative maximum Doppler frequencies, can be observed. Their edges are blurry, which implies that these bins are contributed by a cluster of (or more than one) close, but separated reflectors. This is coincident with our simulation results. The remaining frequency bins, close to zero frequency, are weaker and less stationary. This is due to that the low Doppler frequencies are associated with the impinging rays perpendicular to the mobile velocity. The observed frequencies and amplitudes belonging to these rays are more likely to experience large time variation over the observation interval compared to the maximum Doppler frequency components close to the direction or the inverse direction of the mobile velocity. Such a spectrum agrees with the typical Doppler spectrum in a rich scattering urban environment. These nonstationary frequency bins might render sinusoidal modeling-based channel prediction difficult.

The MPS in Figure 3 is obtained from a measurement in a suburban area. It has much fewer high-power frequency bins compared to what is observed in urban measurement as given in Figure 2. The distribution of these frequency bins in the spectrum is biased to the negative side in this example. This is probably due to that the mobile station was moving away from the base station or the reflection objects during the measurement. The edges of these high-power frequency bins are much sharper than those in Figure 2. This indicates that each bin might contain only one pure sinusoid. This is also coincident with a typical Doppler spectrum in a suburban or rural area in LOS. The power in the strongest frequency bin is much higher than the others, and is consistent over the whole measurement interval.

5.3. Model parameter tracking

Similar to the study of the MPSs, estimates of the model parameters (frequencies and amplitudes) are computed for each data block using the unitary ESPRIT and LS methods. The estimates of the frequencies and amplitudes using the same urban and suburban example measurement are plotted as a function of the number of data blocks in Figures 4 and 5, respectively. The numbers of sinusoids are assumed to be 8 and 7 for urban and suburban example measurements, respectively, as before. In these figures, the channel amplitudes and phases are given in subplots (a) and (b). The estimated frequencies and the associated amplitudes are given in subplots (c) and (d), where different markers are used for frequency and amplitude estimates associated with each sinusoid.

In Figure 4, which is the urban case, the dynamic range of the channel amplitude is larger than the one in the suburban measurement, as given in subplot (a) in Figure 5. The channel phase is linear in just a part of the measurement, which indicates that there might be just one dominant sinusoid in the linear part of the channel, but more comparable dominant sinusoids in the remaining part. In subplot (d), one amplitude is much higher than others, which is marked by an asterisk. Its frequency is close to -0.1 as seen in subplot (c). The amplitude of this ray is consistent over the observation interval. Besides this sinusoid, a number of sinusoids with well-separated frequencies are found in subplot (c), but their amplitudes are time varying and comparable. The variations of these amplitudes are much larger than that of the first frequency. The presence of these sinusoids might explain the large variation of the channel amplitudes and nonlinear phase in this measurement.

In Figure 5, the parameter tracking using the suburban example measurement is given. Both the average and the variation of the channel amplitude are smaller than in the urban measurement, as seen in subplot (a) in Figure 4. The phase of the channel is quite linear over the whole measurement interval. In subplot (d) there are two sinusoids marked by a triangle and a circle having the strongest amplitudes among all the estimated sinusoids. The frequencies associated with these two sinusoids are extremely close, as seen in subplot (c). This may be due to over estimated number of sinusoids and indicates that the channel could be well modeled by just one dominant sinusoidal component.

6. PERFORMANCE EVALUATIONS

6.1. Measures of prediction performance

The performance of the predictors is evaluated by Monte Carlo simulations and measurements. The normalized mean square error (NMSE) is adopted in this paper for performance evaluations. First, we define the normalized square error (NSE) in a single realization as

$$e^2_{\mathrm{NSE}}(t+L) = \frac{N\left|\left|h(t+L) - \hat{h}(t+L)\right|\right|^2}{\mathbf{h}^H\mathbf{h}}. \tag{61}$$

Then, the NMSE is the mean value of $e^2_{\mathrm{NSE}}(t+L)$ taken over the Monte Carlo simulations. However, an adjusted NMSE (ANMSE) is also introduced to get rid of the influence of a small amount of outliers which might ruin the whole averaged performance. So the outliers are dropped when the NSE of their power prediction is larger than 0.04, or in other words, when the prediction error of the complex amplitude is larger than 20% of the root mean square (RMS) channel amplitudes.

The concept of *level of confidence* is also used in the performance evaluation. For example, when we say the NMSE of a channel predictor is -10 dB with the level of confidence of 95%, we mean that the average of the NSEs of the best 95% of the channel predictions from simulations or measurements is less than -10 dB, while the worst 5% cases are excluded.

It is also worth noting that when the measurement data is used, the channel observation $y(t+L)$ is used as $h(t+L)$,

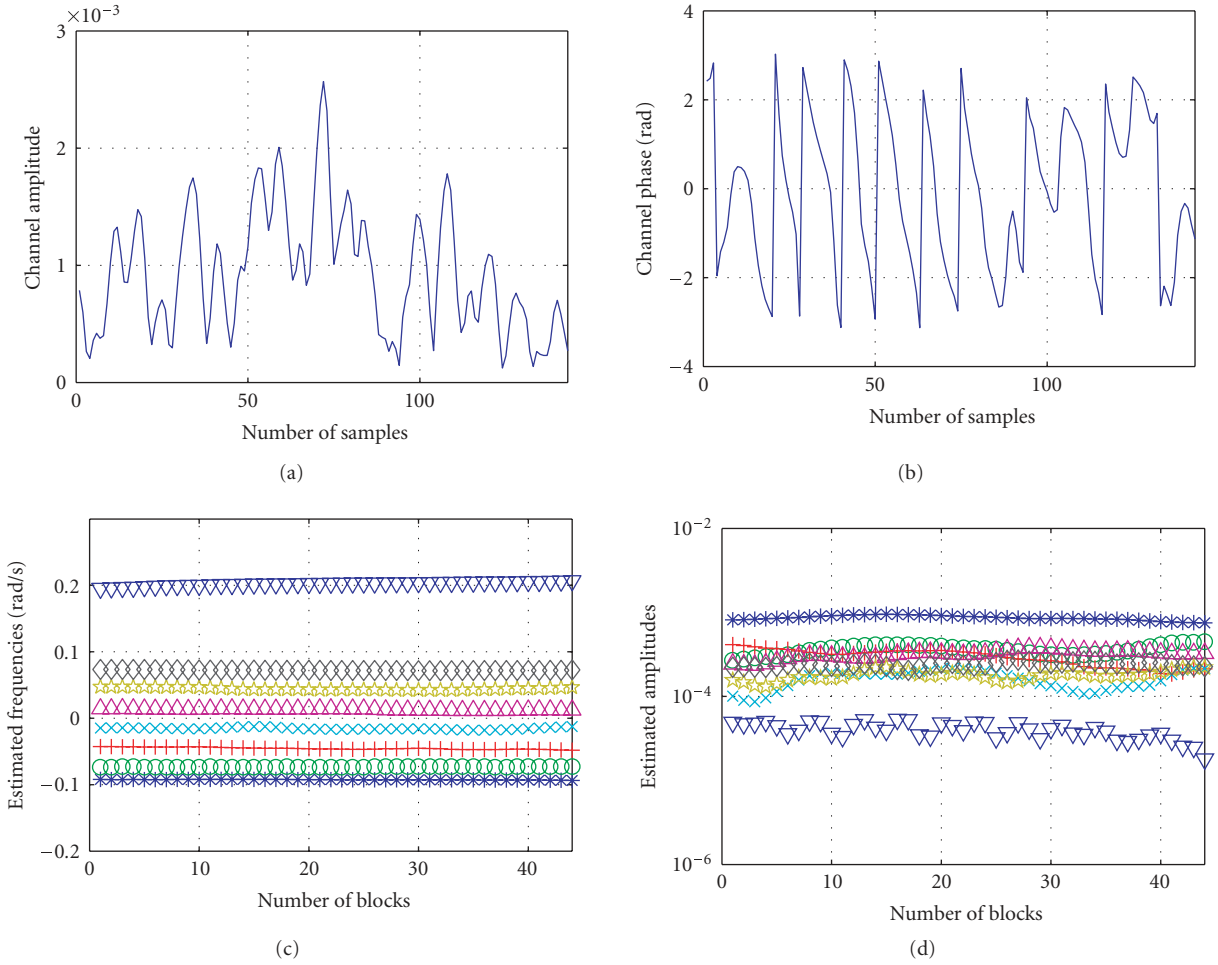

Figure 4: Example of the time variation of model parameters in an urban area: (a) the channel amplitude, whose dynamic range is large in this example; (b) the phase of the channel, which is linear in a part of this measurement; (c) the estimated frequencies of the sinusoids in each block, the number of sinusoids is assumed to be 8; (d) the amplitudes of the sinusoids associated to the frequencies in (c).

and the prediction error is normalized by the mean power of $\mathbf{y}$, instead of $\mathbf{h}$ as given in (61).

6.2. Model selection based on SVD

In this study, an ad hoc singular-value-decomposition-(SVD-) based model selection method is proposed. First, the number of the consistent sinusoidal signals is estimated using the ith largest gradient method. It chooses the index of the descending ordered singular values which gives the ith largest gradient in linear scale. For example, an m-vector

$$\boldsymbol{\sigma} = \left[\sigma_1, \ldots, \sigma_{r-1}, \sigma_r, \ldots, \sigma_m\right]^T \tag{62}$$

contains the singular values from the SVD of $\mathbf{Y}_{ls}$, where $L = 0$ and $d = m$. The value of m should be larger than the number of sinusoids. These singular values are arranged in descending order. If $(\sigma_{r-1} - \sigma_r)$ gives the ith largest gradient between adjacent singular values, σ_r is called the *break* point, and the index r is then selected as the number of consistent sinusoidal signals. In the second step, p frequencies are estimated from the data block, where $p > r$. The frequencies that have the r largest estimated amplitudes are, then, assumed to be consistent.

The selection of i in this simple method is environment dependent. In a suburban area, most rays can be expected to be consistent. We could therefore choose a larger i. But in an urban area, this method might not fit, since some sinusoids are strong, but not stationary. This means that a small i should be selected for urban measurements to be *conservative*. In the following performance evaluations, we choose $i = 1$ and $i = 3$ for urban measurements and suburban measurements, respectively, which provide the best performance for the methods in Section 2.

6.3. Simulation setup

The simulation parameter setting is as follows:

(i) SISO scenario;
(ii) spatial channel sampling interval = $0.1\,\lambda$ ($\Delta l = 0.1$);
(iii) prediction horizon = $0.5\,\lambda$ ($L = 5$);

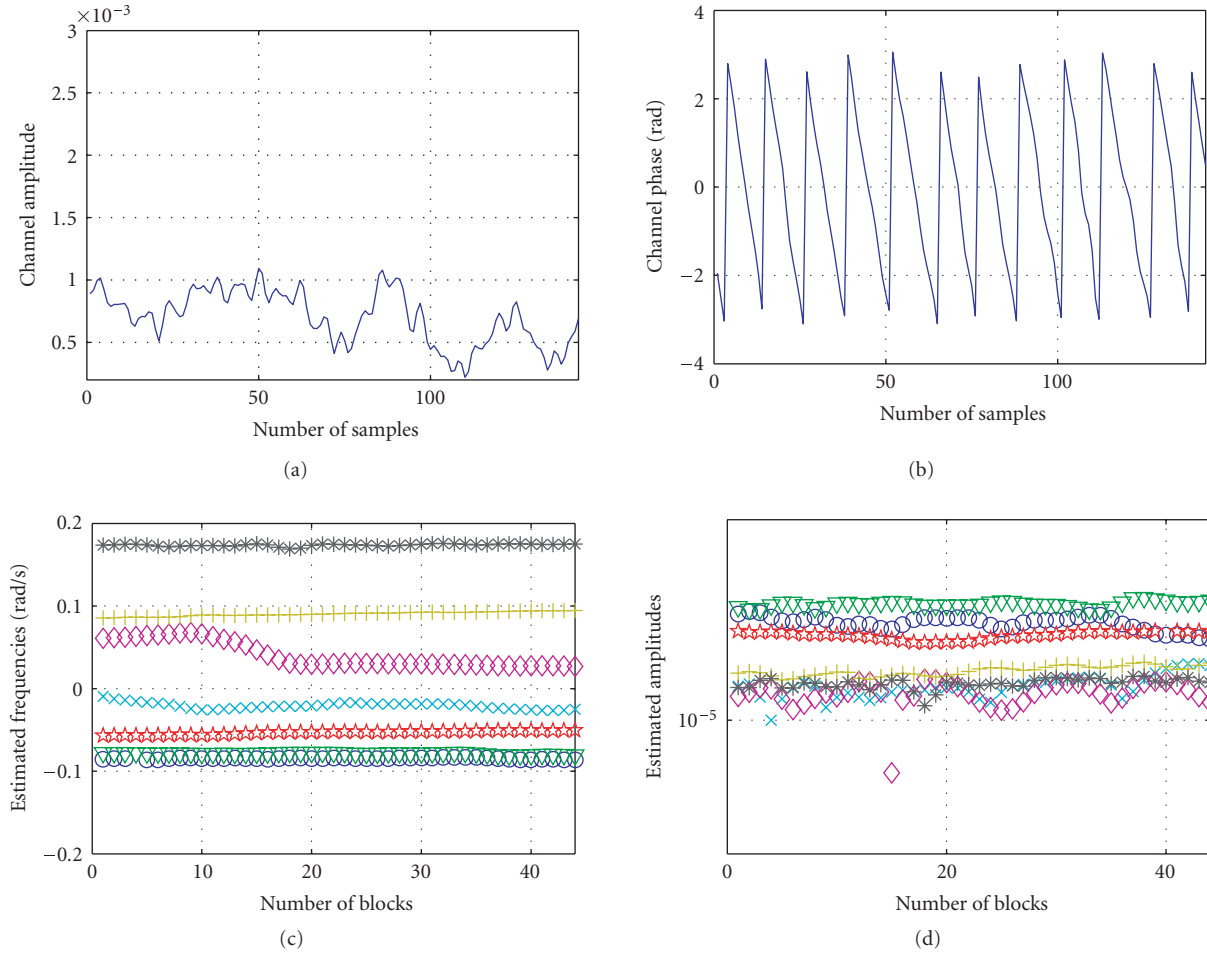

FIGURE 5: Example of the time variation of model parameters in a suburban area: (a) the channel amplitude, whose dynamic range is smaller than in Figure 4; (b) the phase of the channel, which is linear in the whole measurement interval; (c) the estimated frequencies of the sinusoids in each block, the number of sinusoids is assumed to be 7; (d) the amplitudes of the sinusoids associated to the frequencies in (c).

(iv) number of sinusoids = 8 ($p = 8$, which is assumed to be known);
(v) number of samples = 100 ($N = 100$);
(vi) number of Monte Carlo simulations = 1000;
(vii) order of LP is two times the number of sinusoids;
(viii) in the combined method, the model selection is based on the powers of sinusoids. Specifically the sinusoids with $|\hat{s}_k| > 2\sigma_e$ are predicted by the conditional LMMSE method, the residual is predicted using an LP with order 2 ($d = 2$);
(ix) the frequency error variance is used as a design parameter, and is taken to be $\sigma^2_{\Delta\omega} = 1/N^3$. This is somewhat arbitrary, but is motivated by the CRLB [21];
(x) the unitary ESPRIT method is used for frequency estimation.

In Figure 6, the cumulative density function (CDF) of the NMSE is presented. All sinusoidal model-based predictors have light tails, but the LP has a heavy tail.

The ANMSE of different predictors is given in Figure 7. It can be seen that the conditional LMMSE predictor with known frequency and the LP have the best and the worst performances, respectively. The conditional LMMSE predictor with estimated frequency has much better performance than the LP in the investigated scenarios. The unconditional LMMSE predictor performs slightly better than the conditional LMMSE predictor. The performance of the adjusted conditional LMMSE predictor approaches that of the conditional LMMSE predictor with known frequency with the increase of SNR.

Note that, in these simulations, for a given carrier frequency, f, and velocity, v, the spatial sampling interval can be easily converted into time sampling interval as $\Delta t = (\Delta l \cdot c)/(f \cdot v)$, where c is the speed of radio propagation.

6.4. *Performance evaluation of LMMSE predictors using measured data*

The original data is downsampled by a downsampling ratio (DSR) of 10 to reduce the calculation load and increase the prediction horizon, which gives rise to a sampling frequency

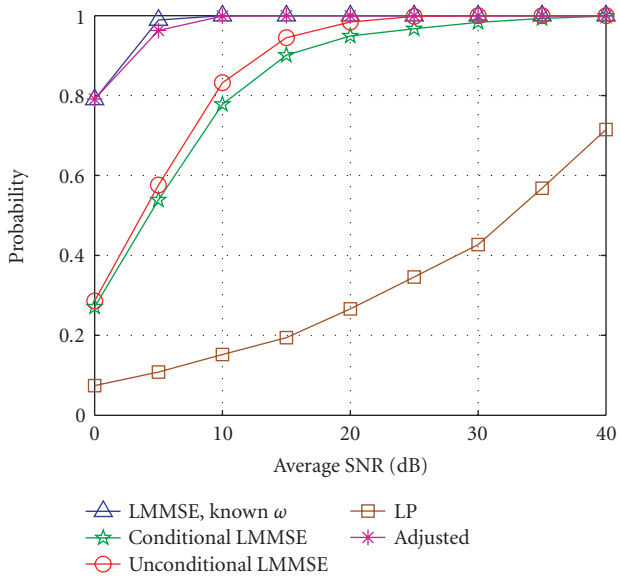

Figure 6: Probability of channel prediction error less than 20% ($L = 0.5\lambda$, $N = 100$, 8 rays).

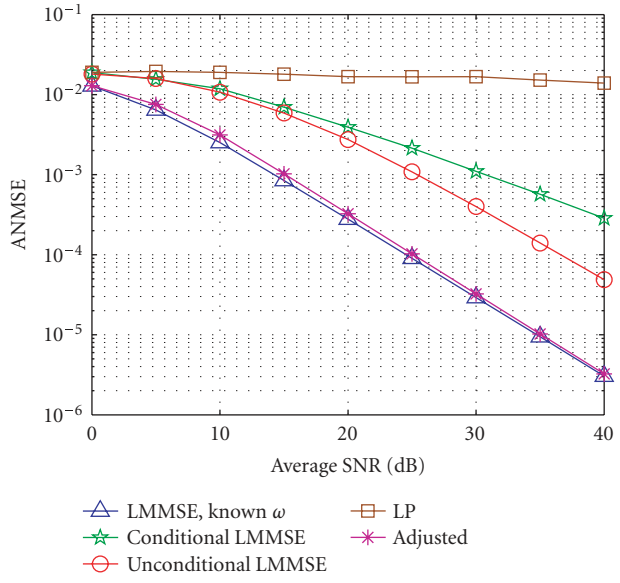

Figure 7: Adjusted NMSE of channel prediction ($L = 0.5\lambda$, $N = 100$, 8 rays).

of 910 Hz. After the downsampling, the evaluation parameter settings are

(i) number of samples = 100 ($N = 100$), which is corresponding to a measurement interval of 0.92 m (6λ) or 2.75 m (18λ) in distance at the speed of 30 km/h or 90 km/h, respectively;

(ii) the prediction horizon L is 5, which is $\lambda/3$ or 1λ at the above speeds and the frequency band of 1880 MHz; in time scale, it is a prediction of 5.5 ms into the future.

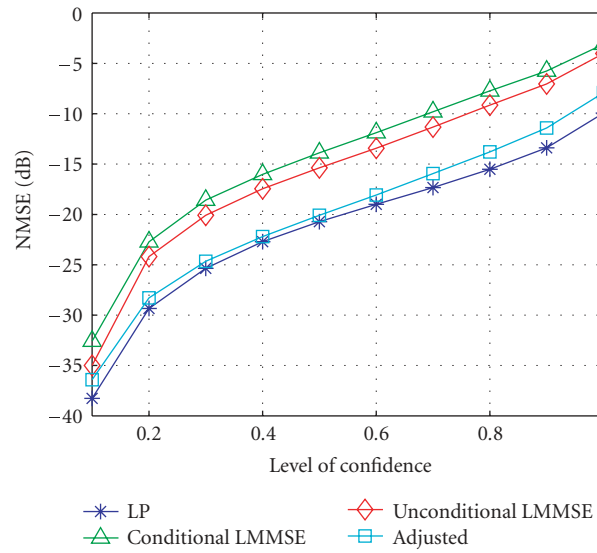

Figure 8: Performance evaluation of LMMSE prediction methods in an urban area.

(iii) the SNRs used in the conditional LMMSE prediction methods are estimated from the associated power delay profiles;

(iv) the order of the LP in the adjusted conditional LMMSE predictors is 2, which is fixed;

(v) the order of the standard LP is set to be the same as the signal order estimated for the data block using model selection method;

(vi) the 1st and 3rd largest gradients are used for the model selection based on SVD for urban and suburban environments, respectively;

(vii) other settings are the same as in Section 6.3.

With these settings there are 39 blocks of data in each measurement. In total, 1755 and 1209 blocks of data are obtained in urban and suburban areas, respectively.

In Figure 8 the overall NMSEs of different prediction methods using all measurement data in an urban area are presented versus the level of confidence. It is meaningless to discuss the achievable NMSEs in this case due to the limited number of measurements. We put our attention on the relative performance instead. In this figure, the LP has the best performance. The conditional LMMSE predictor has the worst performance. The unconditional LMMSE predictor outperforms the conditional LMMSE predictor. The adjusted conditional LMMSE predictor outperforms the conditional LMMSE predictor and the unconditional LMMSE predictor, and is slightly worse than LP.

In Figure 9, the results using suburban measurements are given. Similar results are observed. In these results, the performances of the conditional LMMSE and the unconditional LMMSE predictor are even closer compared to those in an urban area. This is mainly due to that the frequency separation is larger in suburban than in urban environments. The adjusted conditional LMMSE predictor outperforms the

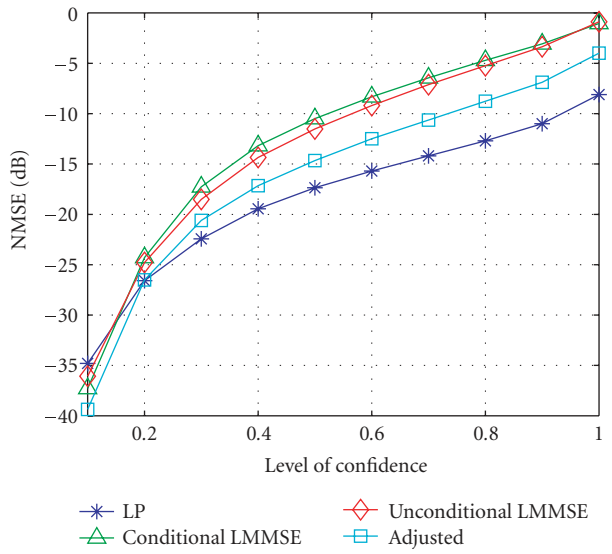

Figure 9: Performance evaluation of LMMSE prediction methods in a suburban area.

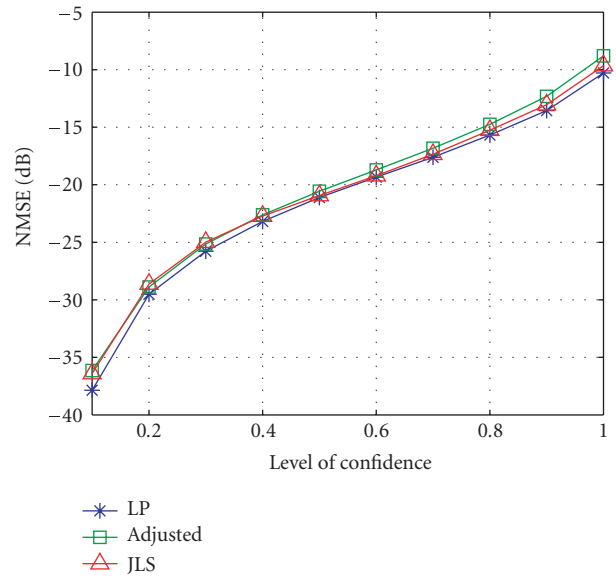

Figure 10: NMSE versus level of confidence (urban area).

conditional and unconditional LMMSE prediction methods and underperforms the LP.

6.5. *Performance evaluation of joint LS predictors using measured data*

In this study, most parameter settings are the same as in the previous section except the following.

(i) The order of the standard LP is 4 and fixed.
(ii) Both orders of the LPs in the joint LS predictor and the adjusted conditional LMMSE predictor are 3 and fixed.
(iii) The number of sinusoids is 1, where only the strongest sinusoidal component is selected.

In Figure 10, the overall NMSEs of different prediction methods using measurement data in an urban area is presented versus the level of confidence as before. In this case we see that the LP still performs (slightly) better than the others. This might be due to the nonstationarity of the model parameters. In addition, the chosen strategy for model selection might not be the best. The strong sinusoids might not have consistent parameters, and the consistent sinusoids are not necessarily the strongest, especially in urban environments. But the joint LS predictor is better than the adjusted conditional LMMSE predictors. This improvement is due to the joint estimate of the model parameters.

In Figure 11, the measurements from the suburban area are used. It can be seen that the joint LS predictor outperforms the other predictors. While the adjusted conditional LMMSE predictor outperforms the standard LP at a low level of confidence. This might be because the strongest path is better described by the sinusoidal model than the nonparametric method.

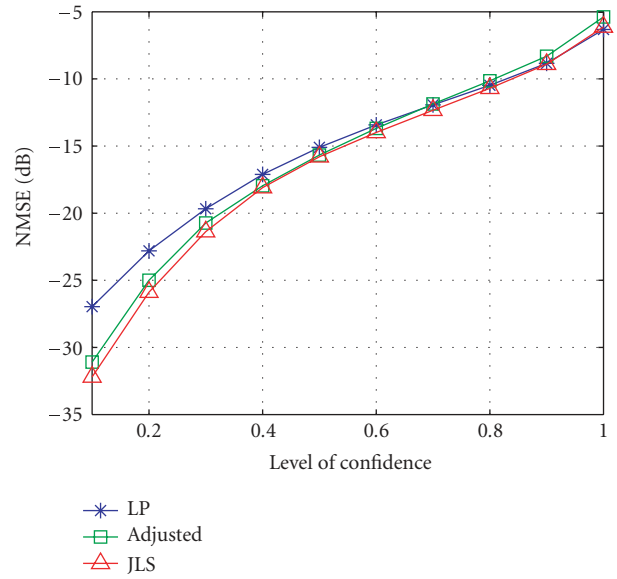

Figure 11: NMSE versus level of confidence (suburban area).

7. CONCLUSIONS

The problem of predicting flat Rayleigh fading channels in wireless communications is studied in the frame of sinusoidal modeling in this paper. A stochastic sinusoidal model for a Rayleigh fading channel is proposed, based on the nature of multipath propagation environment in wireless communications. Given frequency estimation, three sinusoidal LMMSE predictors are proposed. They are conditional LMMSE predictor, unconditional LMMSE predictor, and the adjusted conditional LMMSE predictor. They outperform

the standard LP in simulations, but underperform using real data. Thus the model-based estimators are significantly more tolerant to measurement noise (channel estimation errors), but suffer from sensitivity to modeling errors and nonstationarities. To take into account unmodeled, but stationary, signal components, a joint moving average and sinusoidal channel prediction model is proposed, which results in the joint LS predictor. Together with an SVD-based model selection method, the joint LS predictor outperforms the other predictors in a suburban environment. While, in the same environment, the adjusted conditional LMMSE predictor outperforms the standard LP at a low level of confidence. But both methods still perform slightly worse than LP in urban areas. Model selection based on frequency tracking might help to further improve the overall performance of the joint LS method and the adjusted conditional LMMSE predictor. Finally, we point out that the channel sounding measurements have a very high SNR, and in a practical use the channel estimates will be more noisy. The enhanced noise immunity of model-based channel predictors might then make them a better choice than standard LP.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Sören Andersson, Henrik Asplund, and Jan-Erik Berg at Ericsson Research for providing measurement data.

REFERENCES

[1] A. Duel-Hallen, S. Hu, and H. Hallen, "Long-range prediction of fading signals: enabling adaptive transmission for mobile radio channels," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 62–75, 2000.

[2] S. Falahati, A. Svensson, T. Ekman, and M. Sternad, "Adaptive modulation systems for predicted wireless channels," *IEEE Transactions on Communications*, vol. 52, no. 2, pp. 307–316, 2004.

[3] J. B. Andersen, J. Jensen, S. Holdt, and F. Frederiksen, "Prediction of future fading based on past measurements," in *Proceedings of the 50th Vehicular Technology Conference (VTC '99)*, vol. 1, pp. 151–155, Amsterdam, The Netherlands, September 1999.

[4] M. Chen and M. Viberg, "LMMSE channel predictor based on sinusoidal modeling," in *Proceedings of IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM '04)*, pp. 377–381, Barcelona, Spain, July 2004.

[5] J.-K. Hwang and J. H. Winters, "Sinusoidal modeling and prediction of fast fading processes," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '98)*, vol. 2, pp. 892–897, November 1998.

[6] S. Semmelrodt and R. Kattenbach, "Performance Analysis and Comparison of Different Fading Forecast Schemes for Flat Fading Radio Channels," *COST 273* TD(03)045, Barcelona, Spain, January 2003.

[7] T. Ekman, M. Sternad, and A. Ahlén, "Unbiased power prediction of Rayleigh fading channels," in *Proceedings of IEEE Vehicular Technology Conference (VTC '02)*, vol. 1, pp. 280–284, Vancouver, British Columbia, Canada, September 2002.

[8] M. Chen, T. Ekman, and M. Viberg, "Two new approaches for channel prediction based on sinusoidal modeling," *EURASIP Journal on Applied Signal Processing*, special issue on advances in subspace-based techniques for signal processing and communications, 2006.

[9] T. Ekman, "An FIR predictor interpretation of LS estimation of sinusoidal amplitudes followed by extrapolation," in *Proceedings of the 12th European Signal Processing Conference (EUSIPCO '04)*, Vienna, Austria, September 2004.

[10] M. Chen, "Channel prediction based on sinusoidal modeling," Licentiate thesis, Chalmers University of Technology, Göteborg, Sweden, August 2005.

[11] M. Sternad, T. Ekman, and A. Ahlén, "Power prediction on broadband channels," in *Proceedings of IEEE Vehicular Technology Conference (VTC '01)*, vol. 4, pp. 2328–2332, Rhodes, Greece, May 2001.

[12] W. C. Jakes Jr., *Microwave Mobile Communications*, IEEE Press, Piscataway, NJ, USA, 1974.

[13] T. Svantesson, J. W. Wallace, and S. Semmelrodt, "Performance evaluation of MIMO channel prediction algorithms using measurements," in *Proceedings of the 13th IFAC Symposium on System Identification*, Rotterdam, The Netherlands, August 2003.

[14] T. Ekman and G. Kubin, "Nonlinear prediction of mobile radio channels: measurements and MARS model designs," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '99)*, vol. 5, pp. 2667–2670, Phoenix, Ariz, USA, March 1999.

[15] H. Lütkepol, *Handbook of Matrices*, John Wiley & Sons, Chichester, UK, 1996.

[16] R. O. Schmidt, "A signal subspace approach to multiple emitter location and spectral estimation," Ph.D. dissertation, Stanford University, Stanford, Calif, USA, November 1981.

[17] A. J. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '83)*, vol. 1, pp. 336–339, Boston, Mass, USA, May 1983.

[18] R. Kumaresan, L. L. Scharf, and A. K. Shaw, "An algorithm for pole-zero modeling and spectral analysis," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, pp. 637–640, 1986.

[19] R. Roy, A. Paulraj, and T. Kailath, "ESPRIT—a subspace rotation approach to estimation of parameters of cisoids in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 5, pp. 1340–1342, 1986.

[20] M. Haardt and J. A. Nossek, "Unitary ESPRIT: how to obtain increased estimation accuracy with a reduced computational burden," *IEEE Transactions on Signal Processing*, vol. 43, no. 5, pp. 1232–1242, 1995.

[21] S. M. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory*, Prentice-Hall, Upper Saddle River, NJ, USA, 1993.

[22] P. Stoica and T. Söderström, "Statistical analysis of MUSIC and subspace rotation estimates of sinusoidal frequencies," *IEEE Transactions on Signal Processing*, vol. 39, no. 8, pp. 1836–1847, 1991.

[23] P. Tichavsky, "High-SNR asymptotics for signal-subspace methods in sinusoidal frequency estimation," *IEEE Transactions on Signal Processing*, vol. 41, no. 7, pp. 2448–2460, 1993.

[24] H. Akaike, "Fitting autoregressive models for prediction," *Annals of the Institute of Statistical Mathematics*, vol. 21, pp. 243–247, 1969.

[25] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.

Ming Chen received the B.Eng. degree in computer communications in 1995 from Xi'an University of Electronics Science and Technology, Xi'an, China. In 2001 and 2005, he received the M.S. degree and the Lic. degree in signal processing, respectively, both from Chalmers University of Technology, Göteborg, Sweden. Currently he is a Ph.D. candidate at Signal Processing Group, Chalmers University of Technology, Göteborg, Sweden. From 1996 to 1999, he was a technical representative at Allgon System AB, Beijing Office, Beijing, China. From 1999 to 2000, he was an STINT (Swedish Foundation for International Cooperation in Research and Higher Education) scholar for studying at Chalmers. From 2001 to 2004, he was a Research Engineer at Ericsson Research, Stockholm, Sweden. His current research interests include signal processing in wireless communications, channel estimation and channel prediction for MIMO systems, channel modeling, and mobile positioning.

Torbjörn Ekman was born in Västerås, Sweden, in 1969. He received the M.S. degree in engineering physics in 1994 and the Ph.D. degree in signal processing in 2002, both from Uppsala University, Uppsala, Sweden. Currently, he is an Associate Professor with the Department of Electronics and Telecommunications at the Norwegian University of Science and Technology (NTNU) in Trondheim, Norway. From 1997 to 1998 he was a visiting scientist at the Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology, Vienna, Austria, on a Marie Curie Grant. From 1999 to 2002, he was visiting the Digital Signal Processing Group, University of Oslo, Oslo, Norway. In 2002–2005, he made his postdoctoral studies at UniK, University Graduate Center, Kjeller, Norway. His current research interests include signal processing in wireless communications, scheduling of radio resources, and dynamic modeling, and prediction of radio channels.

Mats Viberg received the Ph.D. degree in automatic control from Linköping University, Sweden, in 1989. During 1989–1992, he was an Assistant Professor in Linköping University and in 1992–1993, he held visiting scholarships at the Information Systems Laboratory, Stanford University, and the Department of ECE, Brigham Young University, USA. He was appointed Full Professor at Chalmers University of Technology, Sweden, in 1993, and during 1999–2004, he served as the Head of the Department of Signals and Systems. His research interests are in statistical signal processing and its various applications, including sensor array signal processing, system identification, wireless communications, radar systems, automotive signal processing, and landmine detection. He has received 2 Paper Awards from the IEEE Signal Processing Society (1993 and 1999, resp.), and was the Chair of the IEEE SP Society's Technical Committee on Signal Processing Theory and Methods 2001–2003. He served as an Associate Editor of the IEEE Transactions on Signal Processing 2004-2005, and currently he is a Member of the IEEE SP Awards Board (2005 up todate), and the IEEE SP Technical Committee on Sensor Array and Multichannel from 2006 until now. He is a Fellow of the IEEE.