Minimum-Energy Band-Limited Predictor With Dynamic Subspace Selection for Time-Variant Flat-Fading Channels

Thomas Zemen, Christoph F. Mecklenbräuker, *Member, IEEE*, Florian Kaltenberger, and Bernard H. Fleury, *Senior Member, IEEE*

***Abstract*—In this paper, we develop and analyze the basic methodology for minimum-energy (ME) band-limited prediction of sampled time-variant flat-fading channels. This predictor is based on a subspace spanned by time-concentrated and band-limited sequences. The time-concentration of these sequences is matched to the length of the observation interval and the band-limitation is determined by the support of the Doppler power spectral density of the fading process. Slepian showed that discrete prolate spheroidal (DPS) sequences can be used to calculate the ME band-limited continuation of a finite sequence. We utilize this property to perform channel prediction. We generalize the concept of time-concentrated and band-limited sequences to a band-limiting region consisting of disjoint intervals. For a fading process with constant spectrum over its possibly discontiguous support we prove that the ME band-limited predictor is identical to a reduced-rank maximum-likelihood predictor which is a close approximation of a Wiener predictor. In current cellular communication systems the time-selective fading process is highly oversampled. The essential dimension of the subspace spanned by time-concentrated and band-limited sequences is in the order of two to five only. The prediction error mainly depends on the support of the Doppler spectrum. We exploit this fact to propose low-complexity time-variant flat-fading channel predictors using dynamically selected predefined subspaces. The subspace selection is based on a probabilistic bound on the reconstruction error. We compare the performance of the ME band-limited predictor with a predictor based on complex exponentials. For a prediction horizon of one eights of a wavelength the numerical simulation results show that the ME band-limited predictor with dynamic subspace selection performs better than, or similar to, a predictor based on complex exponentials with perfectly known frequencies. For a prediction horizons of three eights of a wavelength the performance of the ME band-limited predictor approaches that of a Wiener predictor with perfectly known Doppler bandwidth.**

***Index Terms*—Discrete prolate spheroidal sequences, minimum-energy band-limited predictor, time-variant channel prediction.**

Manuscript received May 25, 2006; revised December 6, 2006. The work of T. Zemen and F. Kaltenberger was funded by the Wiener Wissenschafts-, Forschungs-, und Technologiefonds (WWTF) in the ftw. project Future Mobile Communications Systems (Math+MIMO). The cooperation with B. Fleury was sponsored by the EU Network of Excellence in Wireless Communications (NEWCOM). Part of this work was presented at the Fifth Vienna Symposium in Mathematical Modeling (MATHMOD), Vienna, Austria, February 2006, at the IEEE International Conference on Communications (ICC), Istanbul, Turkey, June 2006, and at the Fourteenth European Signal Processing Conference (EUSIPCO), Florence, Italy, September 4–8, 2006 as an invited paper. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. A. P. Liavas.

T. Zemen is with ftw. Forschungszentrum Telekommunikation Wien (Telecommunications Research Center Vienna), 1220 Vienna, Austria (e-mail: thomas.zemen@ftw.at).

C. F. Mecklenbräuker was with ftw. Forschungszentrum Telekommunikation Wien (Telecommunications Research Center Vienna), 1220 Vienna, Austria. He is now with the Institute for Communications and Radio Frequency Engineering, Vienna University of Technology, Vienna, Austria (e-mail: cfm@nt.tuwien.ac.at).

F. Kaltenberger is with Austrian Research Centers GmbH, 1220 Vienna, Austria (e-mail: florian.kaltenberger@arcs.ac.at).

B. H. Fleury is with the Department of Communication Technology, Aalborg University, 9220 Aalborg, Denmark. He is also with the ftw. Forschungszentrum Telekommunikation Wien (Telecommunications Research Center Vienna), 1220 Vienna, Austria (e-mail: fleury@kom.aau.dk).

Digital Object Identifier 10.1109/TSP.2007.896063

I. Introduction

In mobile communication systems channel state information at the transmitter proves to be beneficial for increasing the system capacity. In a time-division duplex (TDD) system, channel state information can be obtained by exploiting channel reciprocity: While a data block is received, channel state information is obtained. This information can be utilized in the following transmission period. However, for moving users at vehicular speed the channel state information gets outdated rapidly. Thus, appropriate channel prediction is necessary.

Existing linear prediction algorithms for time-variant channels can be categorized into two groups. The first group of algorithms exploits the long-term correlation property of the fading process without considering a physical wave propagation model [1]. Thus, second-order statistics must be known in detail. In wireless communication systems, detailed second-order statistics are difficult to acquire due to the short time-interval over which the channel can be assumed to be stationary (in the wide sense) [2], [3]. Furthermore, bursty data transmission in multiuser systems pose another problem for obtaining channel observations over a sufficiently long time-interval.

The second group of algorithms take physical wave propagation models into account. The time-variant flat-fading channel is represented as the superposition of P propagation paths. Each path is characterized by its distinct complex weight and Doppler shift. A finite number of noisy channel observations is used to identify the parameters of all P paths.

1) The Doppler shift of each path is identified [4]–[8].
2) The complex weight of each path is estimated in the minimum mean square error (MMSE) sense.
3) Future channel values are predicted based on the above estimates.

The Doppler shift estimation for each individual path requires high computational complexity. Another drawback of the specular path models is rooted in the fact that the estimation error

for the complex weight increases drastically if the Doppler frequency difference between two paths becomes small [9]. In [6] a prediction algorithm assuming few discrete scatterers is analyzed. The results in [6] show good performance in simulations but reduced performance with measured channel data.

In this paper, we deal with the prediction of time-variant flat-fading channels in a single-input single-output wireless mobile communication system. We are concerned with low-complexity prediction of a fading process from noisy channel observations that are obtained while receiving a *single* data block. We consider a data block length such that the time-selective fading process is observed at most for one or two wavelength at the maximum user velocity.

The symbol rate, or equivalently the sampling rate of the fading process, in wireless communication systems is much higher than the Doppler bandwidth. Thus, time-limited snapshots of the sampled fading process span a subspace with small dimension [10]. The essential subspace dimension D' depends on the Doppler bandwidth B_{D} the sampling rate $1/T_{\mathrm{S}}$ and the number of observed samples M

$$D' = \lceil 2MB_{\mathrm{D}}T_{\mathrm{S}} \rceil + 1. \tag{1}$$

The number of maximal possible subspace dimensions M gets reduced by the factor $2B_{\mathrm{D}}T_{\mathrm{S}}$. In practical wireless communication systems D' is in the order of two to five only.

The same subspace is also spanned by index-limited discrete prolate spheroidal (DPS) sequences [11]. The band-limitation of the DPS sequences is chosen according to the support of the power spectral density of the time-selective fading process. The energy of the DPS sequences is most concentrated in an interval equal to the length of the observed data block. Thus, the DPS sequences allow to calculate the minimum-energy (ME) band-limited continuation of a finite sequence [10], hence, *predict* future samples.

Prediction of continuous-time channels is treated in [12]–[14]. In the present paper, we deal with discrete-time, i.e., sampled channels which are important for practical implementations using digital signal processing hardware. ME band-limited prediction for multidimensional energy-concentrated signals with *a priori known* band-limiting region is presented in [15].

Contributions of the Paper:

- We prove that ME band-limited prediction is equivalent to a reduced-rank approximation [16] of the Wiener predictor [17, Sec. 12.7] for a fading process with constant Doppler spectrum. A first result in the same direction but for the noiseless case is given in [18].
- Using analytic performance results we show for Clarke's fading model that the prediction error of a reduced-rank predictor is strongly dependent on the support of the Doppler power spectral density, while the actual shape of the Doppler spectrum is of minor importance [19].
- In mobile communication channels, fading processes frequently arise whose spectral support is the union of disjoint intervals. Examples for such short-time Doppler spectra are reported in [20], [21]. We generalize the concept of ME band-limited prediction to processes whose band-limit is defined by a region consisting of disjoint intervals.
- In practical systems, detailed information about the Doppler spectrum or its band-limiting region is not available. Instead, we exploit the small subspace dimension and define a finite set of precalculated subspaces. Each subspace represents a hypothesis about the disjoint band-limiting region of the fading process. We obtain probabilistic bounds on the reconstruction error of each predefined subspace using the method from [22] for the current data block. The subspace with the smallest reconstruction error is selected and utilized for channel prediction [23].

Notation: We denote a column vector by $\boldsymbol{a}$ and its ith element with $a[i]$. Similarly, we denote a matrix by $\boldsymbol{A}$ and its (i,ℓ)th element by $[\boldsymbol{A}]_{i,\ell}$. The transpose of $\boldsymbol{A}$ is given by $\boldsymbol{A}^{\mathrm{T}}$ and its conjugate transpose by $\boldsymbol{A}^{\mathrm{H}}$. A diagonal matrix with elements $a[i]$ is written as $\mathrm{diag}(\boldsymbol{a})$ and the $Q \times Q$ identity matrix as $\boldsymbol{I}_Q$. The absolute value of a is denoted by $|a|$ and its complex conjugate by a^*. The largest (smallest) integer that is lower (greater) or equal than $a \in \mathbb{R}$ is denoted by $\lfloor a \rfloor$ ($\lceil a \rceil$). We denote the set of all integers by $\mathbb{Z}$, the set of real numbers by $\mathbb{R}$ and the set of complex numbers by $\mathbb{C}$.

Organization of the Paper: The signal model for time-variant flat-fading channels is presented in Section II. In Section III, we explain the ME band-limited prediction method. The tight relation between the ME band-limited predictor and the Wiener predictor is highlighted in Section IV. In Section V, analytic expressions for the prediction error of the ME band-limited predictor are presented and an analytic performance comparison with the Wiener predictor is conducted in Section VI. In Section VII, a dynamic subspace selection scheme for the ME band-limited predictor is introduced and its computational complexity is assessed in Section VIII. An enhancement for discontiguous Doppler spectra is shown in Section IX. We provide Monte Carlo simulation results in Section X and draw conclusions in Section XI.

II. Signal Model for Time-Variant Flat-Fading Channels

We consider a TDD communication system transmitting data in blocks of length M over a time-variant channel. The symbol duration T_{S} is much longer than the delay spread T_{D} of the channel, i.e., $T_{\mathrm{S}} \gg T_{\mathrm{D}}$. Hence, we assume the channel as frequency-flat. Discrete time at rate $R_{\mathrm{S}} = 1/T_{\mathrm{S}}$ is denoted by m. The channel incorporates the transmit filter, the transmit antenna, the physical channel, the receive antenna, and the receive matched filter. The data symbols $b[m]$ are randomly and evenly drawn from a symbol alphabet with constant modulus. Without loss of generality $|b[m]| = 1$. The discrete-time signal at the matched filter output

$$h[m]b[m] + n'[m] \tag{2}$$

is the superposition of the data symbol multiplied by the sampled time-variant channel weight $h[m]$ and complex white Gaussian noise $n'[m]$ with variance σ_n^2. Without loss of generality $\{h[m]\}$ is a zero-mean, circularly symmetric, unit-variance (due to power control) process.

We assume an error-free decision feedback structure [24], [25]. Thus, we are able to obtain noisy channel observations [11] using the error-free data symbol estimates $\hat{b}[m] = b[m]$

$$\begin{aligned} y[m] &= (h[m]b[m] + n'[m])\,\hat{b}[m]^* \\ &= h[m] + n'[m]\hat{b}[m]^* = h[m] + n[m]. \end{aligned} \tag{3}$$

Note that $n[m]$ has the same statistical properties as $n'[m]$. The signal-to-noise ratio (SNR) is $\mathrm{SNR} = 1/\sigma_n^2$.

The transmission is block oriented. A data block spans the time interval $\mathcal{I}_M = \{0, \ldots, M-1\}$. The noisy channel observations $y[m]$, $m \in \mathcal{I}_M$ obtained during a single data block are used to predict the channel weight up to N symbols into the future.

The electromagnetic field at the receiver is the superposition of the contribution of the individual fields of P impinging wave fronts. Each wave front is conceived as originating from a specific scatterer. For a user moving with velocity v the time-variant fading process $\{h[m]\}$ is band-limited by the one-sided normalized Doppler bandwidth

$$\nu_{\mathrm{D}} = \frac{v f_{\mathrm{C}}}{c_0} T_{\mathrm{S}} \ll \frac{1}{2} \tag{4}$$

where f_{C} is the carrier frequency and c_0 stands for the speed of light. As indicated with the inequality in (4) the sampling rate $1/T_{\mathrm{S}}$ is much higher than the Nyquist sampling rate.

We assume a time-variant block-fading channel model. Hence, the fading process $\{h[m]\}$ is wide-sense stationary over the limited time interval $\mathcal{I}_{M+N}$ (cf. Section X-A) with covariance function

$$R_h[k] = \mathsf{E}\left\{h^*[m]h[m+k]\right\}. \tag{5}$$

III. Minimum-Energy Band-Limited Prediction

The samples of the channel weights in a single block $\mathcal{I}_M$ are collected in the vector

$$\boldsymbol{h} = [h[0], h[1], \ldots, h[M-1]]^{\mathrm{T}}. \tag{6}$$

The covariance matrix of $\boldsymbol{h}$ is defined as $\boldsymbol{\Sigma}_{\boldsymbol{h}} = \mathsf{E}\{\boldsymbol{h}\boldsymbol{h}^{\mathrm{H}}\}$ with elements $[\boldsymbol{\Sigma}_{\boldsymbol{h}}]_{\ell,m} = R_h[\ell - m]$ for $\ell, m \in \mathcal{I}_M$. The noisy observation vector $\boldsymbol{y} = [y[0], y[1], \ldots, y[M-1]]^{\mathrm{T}}$ is used for channel prediction. Its covariance matrix reads

$$\boldsymbol{\Sigma}_{\boldsymbol{y}} = \boldsymbol{\Sigma}_{\boldsymbol{h}} + \sigma_n^2 \boldsymbol{I}_M. \tag{7}$$

A. Reduced-Rank Channel Estimation

We consider a subspace-based approximation which expands the vector $\boldsymbol{h}$ in terms of D orthonormal basis vectors $\boldsymbol{u}_i = [u_i[0], u_i[1], \ldots, u_i[M-1]]^{\mathrm{T}}$, $i \in \{0, \ldots, D-1\}$

$$\boldsymbol{h} \approx \boldsymbol{U}\boldsymbol{\gamma} = \sum_{i=0}^{D-1} \gamma_i \boldsymbol{u}_i. \tag{8}$$

In this expression

$$\boldsymbol{U} = [\boldsymbol{u}_0, \ldots, \boldsymbol{u}_{D-1}] \tag{9}$$

contains the orthonormal basis vectors and $\boldsymbol{\gamma} = [\gamma_0, \ldots, \gamma_{D-1}]^{\mathrm{T}}$ collects the basis expansion coefficients. The least square estimate of $\boldsymbol{\gamma}$ simplifies to

$$\hat{\boldsymbol{\gamma}} = \boldsymbol{U}^{\mathrm{H}}\boldsymbol{y} \tag{10}$$

due to the orthogonality of the basis functions. The reconstruction error per data block is defined as

$$z = \frac{1}{M}\|\boldsymbol{h} - \hat{\boldsymbol{h}}\|^2 = \frac{1}{M}\|\boldsymbol{U}^{\mathrm{H}}\boldsymbol{n}\|^2 + \frac{1}{M}\|\boldsymbol{V}^{\mathrm{H}}\boldsymbol{h}\|^2 \tag{11}$$

where

$$\hat{\boldsymbol{h}} = \boldsymbol{U}\hat{\boldsymbol{\gamma}} \tag{12}$$

and $\boldsymbol{V} = [\boldsymbol{u}_D, \ldots, \boldsymbol{u}_{M-1}]$ contains the basis vectors spanning the subspace orthogonal to the signal subspace spanned by the columns of $\boldsymbol{U}$. The noise samples are collected in the vector $\boldsymbol{n} = [n[0], \ldots, n[M-1]]^{\mathrm{T}}$.

We define the mean square reconstruction error per sample

$$\mathrm{MSE}[m] = \mathsf{E}\left\{\left|h[m] - \hat{h}[m]\right|^2\right\}. \tag{13}$$

and the mean square reconstruction error per data block

$$\mathrm{MSE} = \mathsf{E}\{z\} = \frac{1}{M}\sum_{m=0}^{M-1} \mathrm{MSE}[m] \tag{14}$$

$$= \frac{D}{M}\sigma_n^2 + \frac{1}{M}\mathsf{E}\left\{\|\boldsymbol{V}^{\mathrm{H}}\boldsymbol{h}\|^2\right\}. \tag{15}$$

In the sequel, we seek basis vectors $\boldsymbol{u}_0, \ldots, \boldsymbol{u}_{D-1}$ and the subspace dimension D which minimize the reconstruction error per data block.

B. Time-Concentrated and Band-Limited Sequences

Slepian analyzed discrete prolate spheroidal (DPS) sequences $\{u_i[m]\}$ in [10]. DPS sequences time-limited to $\mathcal{I}_M$ form a set of orthogonal basis vectors $\boldsymbol{u}_i$. This set of basis vectors is used in [11] for the estimation of fading processes with symmetric spectral support $\mathcal{W} = (-\nu_{\mathrm{D}}, \nu_{\mathrm{D}})$ with $\nu_{\mathrm{D}} < 1/2$.

In mobile radio communication channels, the most significant part of the power in the estimated Doppler spectrum of the fading process is usually localized on the union of disjoint intervals in the frequency range $(-1/2, +1/2)$ as depicted in Fig. 1. Examples for such estimated (short-time) Doppler spectra are reported in [20] and [21]. In the sequel, we generalize the concept of time-concentrated and band-limited sequences from a symmetric band-limiting interval to the union of multiple disjoint intervals using results presented in [26].

Fig. 1 depicts a region $\mathcal{W} \subseteq (-1/2, 1/2)$ consisting of I disjoint intervals. Each interval is defined as $\mathcal{B}_i = (\nu_{i1}, \nu_{i2})$, $i \in \{1, \ldots, I\}$ and

$$\mathcal{W} = \bigcup_{i=1}^{I} \mathcal{B}_i = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \ldots \cup \mathcal{B}_I \tag{16}$$

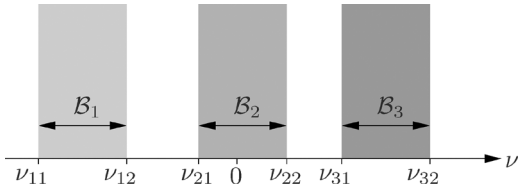

Fig. 1. Band-limiting region $\mathcal{W}$ (16) consisting of $I = 3$ disjoint intervals.

with $\nu_{11} \leq \nu_{12} \leq \ldots \leq \nu_{I1} \leq \nu_{I2}$. The Lebesgue measure of $\mathcal{W}$ reads

$$|\mathcal{W}| = \sum_{i=1}^{I} (\nu_{i2} - \nu_{i1}). \tag{17}$$

Definition 1: A sequence $\{u[m, \mathcal{W}]\}$ is band-limited to the region $\mathcal{W} \subseteq (-1/2, 1/2)$, if its spectrum

$$U(\nu, \mathcal{W}) = \sum_{m=-\infty}^{\infty} u[m, \mathcal{W}] \mathrm{e}^{-\mathrm{j}2\pi m\nu}, \quad |\nu| \leq \frac{1}{2} \tag{18}$$

vanishes outside $\mathcal{W}$. Thus,

$$u[m, \mathcal{W}] = \int_{\mathcal{W}} U(\nu, \mathcal{W}) \mathrm{e}^{\mathrm{j}2\pi m\nu} \mathrm{d}\nu, \quad m \in \mathbb{Z}. \tag{19}$$

Definition 2: The energy-concentration of a sequence $\{u[m, \mathcal{W}]\}$ in the interval $\mathcal{I}_M$ is defined as

$$\rho(\mathcal{W}) = \frac{\sum_{m=0}^{M-1} |u[m, \mathcal{W}]|^2}{\sum_{m=-\infty}^{\infty} |u[m, \mathcal{W}]|^2}. \tag{20}$$

Theorem 1: The sequences $\{u_i[m, \mathcal{W}]\}$, $i \in \{0, \ldots, M-1\}$ band-limited to the region $\mathcal{W}$ and with *most concentrated* energy in the interval $\mathcal{I}_M$ are the solutions to

$$\sum_{\ell=0}^{M-1} C[\ell - m, \mathcal{W}] u_i[\ell, \mathcal{W}] = \lambda_i(\mathcal{W}) u_i[m, \mathcal{W}], \quad m \in \mathbb{Z} \tag{21}$$

where

$$C[k, \mathcal{W}] = \int_{\mathcal{W}} \mathrm{e}^{\mathrm{j}2\pi k\nu} \mathrm{d}\nu. \tag{22}$$

Note that $C[k, \mathcal{W}]$ is proportional to the covariance function of a process exhibiting a constant spectrum with support $\mathcal{W}$. More details are given in Section III-C.

Equation (22) evaluates to

$$C[k, \mathcal{W}] = \frac{1}{\mathrm{j}2\pi k} \sum_{i=1}^{I} \left(\mathrm{e}^{\mathrm{j}2\pi k\nu_{i2}} - \mathrm{e}^{\mathrm{j}2\pi k\nu_{i1}} \right) \tag{23}$$

if the band-limiting region $\mathcal{W}$ consists of I disjoint intervals as defined in (16) and depicted in Fig. 1.

The actual energy-concentration of the sequence $\{u_i[m, \mathcal{W}]\}$ is given by

$$\rho_i(\mathcal{W}) = \frac{\sum_{m=0}^{M-1} |u_i[m, \mathcal{W}]|^2}{\sum_{m=-\infty}^{\infty} |u_i[m, \mathcal{W}]|^2} = \lambda_i(\mathcal{W}). \tag{24}$$

Proof: See Appendix II. □

Theorem 1 shows, that the eigenvalue $\lambda_i(\mathcal{W})$ is a direct measure of the energy-concentration of the sequence $\{u_i[m, \mathcal{W}]\}$ in the interval $\mathcal{I}_M$. The sequences $\{u_i[m, \mathcal{W}]\}$ and the eigenvalues $\lambda_i(\mathcal{W})$ depend on the region $\mathcal{W}$ and the interval length M. In the sequel, we omit the dependence on the block length M since this parameter is kept fixed throughout the paper.

Both, the sequences $\{u_i[m, \mathcal{W}]\}$ and their restrictions on $\mathcal{I}_M$ form orthogonal sets [27]. The eigenvalues $\lambda_i(\mathcal{W})$ decay exponentially for $i \geq D'(\mathcal{W})$ [26]. The essential subspace dimension is defined as

$$D'(\mathcal{W}) = \lceil |\mathcal{W}| M \rceil + 1. \tag{25}$$

Let us define the vectors $\boldsymbol{u_i}(\mathcal{W}) = [u_i[0, \mathcal{W}], \ldots, u_i[M-1, \mathcal{W}]]^{\mathrm{T}}$ and the matrix $[\boldsymbol{C}(\mathcal{W})]_{\ell,m} = C[\ell - m, \mathcal{W}]$ for $\ell, m \in \mathcal{I}_M$. It follows from (21) that

$$\boldsymbol{C}(\mathcal{W}) \boldsymbol{u_i}(\mathcal{W}) = \lambda_i(\mathcal{W}) \boldsymbol{u_i}(\mathcal{W}) \tag{26}$$

i.e., the vectors $\boldsymbol{u_i}(\mathcal{W})$ are the eigenvectors of $\boldsymbol{C}(\mathcal{W})$ and $\lambda_i(\mathcal{W})$ are their corresponding eigenvalues.

C. Relation to the Karhunen-Loève Identity

Identity (26) coincides with the Karhunen-Loève identity [28] in the case where the fading process $\{h[m]\}$ has a constant spectrum with support $\mathcal{W}$ given in (16)

$$S_h(\nu, \mathcal{W}) = \begin{cases} \frac{1}{|\mathcal{W}|} & \nu \in \mathcal{W}; \\ 0 & \text{otherwise.} \end{cases} \tag{27}$$

The covariance function of $\{h[m]\}$ reads in this case

$$R_h[k, \mathcal{W}] = \frac{1}{\mathrm{j}2\pi k |\mathcal{W}|} \sum_{i=1}^{I} \left(\mathrm{e}^{\mathrm{j}2\pi k\nu_{i2}} - \mathrm{e}^{\mathrm{j}2\pi k\nu_{i1}} \right). \tag{28}$$

Comparing (28) and (23) yields

$$R_h[k, \mathcal{W}] = \frac{1}{|\mathcal{W}|} C[k, \mathcal{W}]. \tag{29}$$

The restriction of the process $\{h[m]\}$ to the interval $\mathcal{I}_M$ has a covariance matrix proportional to $\boldsymbol{C}(\mathcal{W})$

$$\boldsymbol{\Sigma_h}(\mathcal{W}) = \frac{1}{|\mathcal{W}|} \boldsymbol{C}(\mathcal{W}) \tag{30}$$

The Karhunen-Loève identity writes in this case

$$\boldsymbol{\Sigma_h}(\mathcal{W}) \boldsymbol{u_i}(\mathcal{W}) = \frac{\lambda_i(\mathcal{W})}{|\mathcal{W}|} \boldsymbol{u_i}(\mathcal{W}). \tag{31}$$

From the Karhunen-Loève identity it follows that the basis vectors $\boldsymbol{u}_i(\mathcal{W})$ minimize the mean square reconstruction error (14) for a fading process with constant Doppler spectrum $S_h(\nu,\mathcal{W})$ with support $\mathcal{W}$. The subspace dimension minimizing the MSE per observation interval for a given SNR is found to be [29]

$$D(\mathcal{W}) = \underset{\mathcal{D}\in\{1,\ldots,M\}}{\arg\min} \left(\frac{1}{|\mathcal{W}|M} \sum_{i=\mathcal{D}}^{M-1} \lambda_i(\mathcal{W}) + \frac{\mathcal{D}}{M}\sigma_n^2 \right). \quad (32)$$

We implicitly assume here that the eigenvalues are ranked $\lambda_0(\mathcal{W}) \geq \lambda_1(\mathcal{W}) \geq \ldots \geq \lambda_{M-1}(\mathcal{W})$. Moreover, the term in the parentheses results immediately from (15) for the particular choice of $\boldsymbol{U}$. By optimizing $D(\mathcal{W})$ we perform a square bias-variance tradeoff, see also Section V.

D. Channel Prediction

So far we considered the channel estimation problem for a channel observed over a time interval $\mathcal{I}_M$. We used orthogonal basis vectors that result from time-limiting infinite sequences to the interval $\mathcal{I}_M$.

However, the main interest of this paper lies on channel prediction. Slepian points out [10, Sec. 3.1.4] that given the channel samples $h[m], m \in \mathcal{I}_M$ there are infinitely many ways to choose the channel samples $h[m], m \in \mathbb{Z}\setminus\mathcal{I}_M$ such that the infinite sequence $\{h[m]\}$ is band-limited. However, there exists only one way to extend a band-limited sequence in the sense of a ME continuation. This is achieved by using the time-concentrated and band-limited sequences $\{u_i[m,\mathcal{W}]\}$ because the sequences $\{u_i[m,\mathcal{W}]\}$ are most energy-concentrated in the interval $\mathcal{I}_M$.

Evaluating (26) we obtain $\boldsymbol{u}_i(\mathcal{W})$, i.e., we know $u_i[m,\mathcal{W}]$ for $m \in \mathcal{I}_M$. The sequences $\{u_i[m,\mathcal{W}]\}$ can be continued over $\mathbb{Z}$ in the ME band-limited sense by evaluating (21). Finally, we can express the ME band-limited prediction of a time-variant channel for any $m \in \mathbb{Z}$ as

$$\hat{h}[m] = \boldsymbol{f}[m,\mathcal{W}]^{\mathrm{T}}\hat{\boldsymbol{\gamma}} = \sum_{i=0}^{D(\mathcal{W})-1} \hat{\gamma}_i u_i[m,\mathcal{W}] \quad (33)$$

where

$$\boldsymbol{f}[m,\mathcal{W}] = [u_0[m,\mathcal{W}],\ldots,u_{D(\mathcal{W})-1}[m,\mathcal{W}]^{\mathrm{T}}. \quad (34)$$

In the next section we explain the relation between ME band-limited prediction and the Wiener predictor. The main benefit of ME band-limited prediction will become clear in Section VII. There we utilize the insights into the subspace structure of linear prediction, gained in this section, to design a new *low complexity* prediction scheme based on a *finite* set of hypothesis about the actual Doppler spectrum of the fading process. Each hypothesis is represented by a *precalculated* subspace spanned by the time-concentrated and band-limited sequences defined in this section.

IV. Relation to the Wiener Predictor

In [17, Sec. 12.7] a solution to the prediction problem is presented using a Wiener (linear minimum MSE) predictor. The Wiener predictor can be closely approximated by a reduced-rank predictor where the subspace dimensions with small eigenvalues are truncated [16]. Both, the Wiener predictor and its reduced rank approximation are defined for a process with general Doppler spectrum and the related covariance matrix $\boldsymbol{\Sigma}_h$.

A process $\{h[m]\}$ exhibiting a constant Doppler spectrum $S_h(\nu,\mathcal{W})$ with support $\mathcal{W}$ given in (27) (see Fig. 1) has the covariance function $R_h[k,\mathcal{W}]$ (28). The restriction of $\{h[m]\}$ on the interval $\mathcal{I}_M$ is described by the covariance matrix $\boldsymbol{\Sigma}_h(\mathcal{W})$ given in (30). For this class of processes we show that the reduced-rank maximum-likelihood (ML) predictor coincides with a ME band-limited predictor [19].

A. Wiener Predictor

With the definition

$$\boldsymbol{r}_h[m] = [R_h[m], R_h[m-1], \ldots, R_h[m-(M-1)]]^{\mathrm{T}} \quad (35)$$

for $m > M-1$ the ℓ-step Wiener predictor, $\ell = m-(M-1)$, is of the form

$$\hat{h}'[m] = \boldsymbol{r}_h[m]^{\mathrm{H}}\boldsymbol{\Sigma}_{\boldsymbol{y}}^{-1}\boldsymbol{y}. \quad (36)$$

B. Reduced-Rank Maximum-Likelihood Predictor

Similar to the reduced-rank ML *estimator* [16], [30] we are able to define a reduced-rank ML *predictor*. The reduced-rank ML channel estimation described in Section III-A uses a deterministic signal model: The channel weight vector $\boldsymbol{h}$ depends deterministically on $\boldsymbol{\gamma}$ given the matrix $\boldsymbol{U}$. However, as is pointed out in [16, Sec. 1] we implicitly exploit the long-term correlation properties: Matrix $\boldsymbol{U}$ (9) is comprised of the first D essential eigenvectors defined by the eigenvalue identity $\boldsymbol{\Sigma}_{\boldsymbol{h}}\boldsymbol{u}_i = \lambda_i\boldsymbol{u}_i$. The optimum dimension D is chosen according to

$$D = \underset{\mathcal{D}\in\{1,\ldots,M\}}{\arg\min} \left(\frac{1}{M} \sum_{i=\mathcal{D}}^{M-1} \lambda_i + \frac{\mathcal{D}}{M}\sigma_n^2 \right). \quad (37)$$

Note that D depends on the noise variance.

The reduced-rank ML *predictor* closely approximates (36) using a subspace of $\boldsymbol{\Sigma}_{\boldsymbol{h}}$ with the optimum subspace dimension D:

$$\hat{h}[m] = \underbrace{\boldsymbol{r}_h[m]^{\mathrm{H}}\boldsymbol{U}\boldsymbol{\Lambda}^{-1}}_{\boldsymbol{f}[m]} \underbrace{\boldsymbol{U}^{\mathrm{H}}\boldsymbol{y}}_{\hat{\boldsymbol{\gamma}}} = \sum_{i=0}^{D-1} u_i[m]\hat{\gamma}_i \quad (38)$$

where $\boldsymbol{\Lambda} = \mathrm{diag}([\lambda_0,\ldots,\lambda_{D-1}])$ and $\boldsymbol{f}[m] = [u_0[m],\ldots,u_{D-1}[m]]^{\mathrm{T}}$. The sequences $\{u_i[m]\}$ are defined as

$$\sum_{\ell=0}^{M-1} R_h[\ell-m]u_i[\ell] = \lambda_i u_i[m], \quad m \in \mathbb{Z} \quad (39)$$

generalizing (21). Again, the sequences $\{u_i[m]\}$ and the restrictions of these sequences on $\mathcal{I}_M$ form orthogonal sets. Note that $\boldsymbol{u}_i = [u_i[0],\ldots,u_i[M-1]]^{\mathrm{T}}$.

For reduced-rank ML prediction the situation is similar to reduced-rank ML channel estimation. The predicted channel

weights depend deterministically on $\boldsymbol{\gamma}$. However, again the deterministic sequences $\{u_i[m]\}$ are obtained by assuming a specific channel covariance matrix.

C. ME Band-Limited Prediction

For a fading process with constant Doppler spectrum $S_h(\nu, \mathcal{W})$ with support $\mathcal{W}$ we can show the similarity between the reduced-rank ML predictor (38) and the ME band-limited predictor (33). We recast (21) as

$$u_i[m, \mathcal{W}] = \frac{|\mathcal{W}|}{\lambda_i(\mathcal{W})} \sum_{\ell=0}^{M-1} R_h[m-\ell, \mathcal{W}]^* u_i[\ell, \mathcal{W}] \quad (40)$$

$$= \frac{|\mathcal{W}|}{\lambda_i(\mathcal{W})} \boldsymbol{r}_h[m, \mathcal{W}]^{\mathrm{H}} \boldsymbol{u}_i(\mathcal{W}) \quad (41)$$

where we use the fact that $R_h[k, \mathcal{W}] = R_h[-k, \mathcal{W}]^*$ and $C[k, \mathcal{W}] = |\mathcal{W}| R_h[k, \mathcal{W}]$. Moreover, inserting (40) in (34) yields

$$\boldsymbol{f}[m, \mathcal{W}]^{\mathrm{T}} = \boldsymbol{r}_h[m, \mathcal{W}]^{\mathrm{H}} \boldsymbol{U}(\mathcal{W}) \boldsymbol{\Lambda}^{-1}(\mathcal{W}) \quad (42)$$

where $\boldsymbol{r}_h[m, \mathcal{W}] = [R_h[m, \mathcal{W}], \ldots, R_h[m-(M-1), \mathcal{W}]]^{\mathrm{T}}$, $\boldsymbol{\Lambda}(\mathcal{W}) = 1/|\mathcal{W}| \mathrm{diag}([\lambda_0(\mathcal{W}), \ldots, \lambda_{D-1}(\mathcal{W})])$, and $\boldsymbol{U}(\mathcal{W}) = [\boldsymbol{u}_0(\mathcal{W}), \ldots, \boldsymbol{u}_{D-1}(\mathcal{W})]$.

Inserting (42) into the ME band-limited predictor (33) we obtain

$$\hat{h}[m] = \boldsymbol{f}[m, \mathcal{W}]^{\mathrm{T}} \hat{\boldsymbol{\gamma}} \quad (43)$$

$$= \boldsymbol{r}_h[m, \mathcal{W}]^{\mathrm{H}} \boldsymbol{U}(\mathcal{W}) \boldsymbol{\Lambda}^{-1}(\mathcal{W}) \boldsymbol{U}(\mathcal{W})^{\mathrm{H}} \boldsymbol{y} \quad (44)$$

which is identical to the reduced-rank ML predictor (38) for fading processes with constant Doppler spectrum $S_h(\nu, \mathcal{W})$. Hence, both predictors use a subspace spanned by time-concentrated and band-limited sequences.

V. Analytical Expressions for the Prediction Error

A. Wiener Predictor

The minimum MSE per sample, which is achieved with the Wiener predictor is given by [17, Sec. 12.7]

$$\mathrm{MSE}[m] = 1 - \boldsymbol{r}_h[m]^{\mathrm{H}} \boldsymbol{\Sigma}_{\boldsymbol{y}}^{-1} \boldsymbol{r}_h[m]. \quad (45)$$

Specializing to constant Doppler spectra $S_h(\nu, \mathcal{W})$ we can express the $\mathrm{MSE}[m]$ of the Wiener predictor in terms of time-concentrated and band-limited sequences $\{u_i[m, \mathcal{W}]\}$ and their eigenvalues $\lambda_i(\mathcal{W})$. Utilizing (42) we obtain

$$\mathrm{MSE}[m, \mathcal{W}] = 1 - \sum_{i=0}^{M-1} \frac{\lambda_i(\mathcal{W})^2/|\mathcal{W}|}{|\mathcal{W}|\sigma_n^2 + \lambda_i(\mathcal{W})} |u_i[m, \mathcal{W}]|^2. \quad (46)$$

For constant Doppler spectra $S_h(\nu, \mathcal{W})$ we can conclude that the prediction horizon of linear prediction methods is inherently limited because the energy of $\{u_i[m, \mathcal{W}]\}$ is most concentrated in the interval $\mathcal{I}_M$ and $\lim_{m\to\infty} u_i[m, \mathcal{W}] = 0$. This conclusion is a direct consequence of (46). Indeed, $\mathrm{MSE}[m, \mathcal{W}]$ directly depends on the absolute value of the time-concentrated and band-limited sequences $\{u_i[m, \mathcal{W}]\}$. Since the energy of these sequences is most concentrated in the interval $\mathcal{I}_M$ the absolute value of $u_i[m, \mathcal{W}]$ decays outside $\mathcal{I}_M$ and the MSE per sample $\mathrm{MSE}[m, \mathcal{W}]$ increases at the same time. A similar conclusion can be drawn from (45) and the decay of the covariance function $R_h[k]$.

B. Reduced-Rank ML Predictor

The MSE per sample of the reduced-rank ML predictor can be described as the sum of a square bias and a variance term

$$\mathrm{MSE}[m] = \mathrm{bias}^2[m] + \mathrm{var}[m]. \quad (47)$$

The expression for square bias and variance developed in [31, Sec. 6] and [11] can be extended for reduced-rank ML prediction [19]. Note that in all equations in this section $m \in \mathbb{Z}$.

The following (48)–(51) are valid for any Doppler power spectral density $S'_h(\nu)$ of the actual fading process $\{h[m]\}$. Thus, we can evaluate the prediction error for the mismatched case too. This mean, the predictor design is based on the assumed Doppler spectrum $S_h(\nu)$ but the actual fading process has Doppler spectrum $S'_h(\nu)$.

We define the instantaneous frequency response of the reduced-rank ML predictor according to

$$G[m, \nu] = \boldsymbol{f}^T[m] \sum_{\ell=0}^{M-1} \boldsymbol{f}^*[\ell] \mathrm{e}^{-\mathrm{j}2\pi\nu(m-\ell)} \quad (48)$$

where $|\nu| < 1/2$ and $\boldsymbol{f}[m]^{\mathrm{T}} = \boldsymbol{r}_h[m]^{\mathrm{H}} \boldsymbol{U} \boldsymbol{\Lambda}^{-1}$. In (48), the sum $\sum_{\ell=0}^{M-1} \boldsymbol{f}^*[\ell] \mathrm{e}^{-\mathrm{j}2\pi\nu(m-\ell)}$ projects the complex exponential onto the basis function, i.e., we calculate the inner product with every basis function. Then, the realization at time instant m is calculated by left multiplying with $\boldsymbol{f}^{\mathrm{T}}[m]$.

The complex exponential in (48) is shifted by m, thus, $|G[m, \nu]|$ is the instantaneous amplitude response of the reduced-rank ML predictor at time instant m. The design goal for the predictor is to have no amplitude error and no phase error. Therefore, the instantaneous error characteristic of the reduced rank Wiener predictor is defined as [31, Sec. 6.1.4]

$$E[m, \nu] = |1 - G[m, \nu]|^2. \quad (49)$$

The square bias per sample $\mathrm{bias}^2[m]$ of the reduced-rank ML predictor can be computed from the instantaneous error characteristic $E[m, \nu]$ and the power spectral density of $\{h[m]\}$ according to

$$\mathrm{bias}^2[m] = \int_{-\frac{1}{2}}^{\frac{1}{2}} E[m, \nu] S'_h(\nu) \mathrm{d}\nu. \quad (50)$$

The variance of reduced-rank ML prediction can be approximated according to

$$\mathrm{var}[m] \approx \sigma_n^2 \boldsymbol{f}^H[m] \boldsymbol{f}[m]. \quad (51)$$

C. ME Band-Limited Prediction

The MSE per sample

$$\mathrm{MSE}[m,\mathcal{W}] = \mathrm{bias}^2[m,\mathcal{W}] + \mathrm{var}[m,\mathcal{W}] \tag{52}$$

for ME band-limited prediction can be obtained by specializing the equations from the previous section by setting $\boldsymbol{f}[m] = \boldsymbol{f}[m,\mathcal{W}]$ (34).

VI. Analytic Performance Comparison

In this section, we show analytic performance results for the Wiener predictor (36) and the reduced-rank ML predictor (38) using full covariance information. These results are compared with the performance of ME band-limited prediction (33) where the knowledge of the support of the Doppler power spectral density is utilized only.

A. Channel Model and System Assumption

The time-variant flat-fading channel $\{h[m]\}$ is assumed to conform to Clarke's model [32]. The covariance function of $h[m]$ is $R_h[k] = \mathrm{J}_0(2\pi\nu_\mathrm{D}k)$ where $\mathrm{J}_0(\cdot)$ is the zeroth-order Bessel function of the first kind. The spectrum of $\{h[m]\}$ reads

$$S_h(\nu) = \frac{1}{\pi\nu_\mathrm{D}\sqrt{1-\left(\frac{\nu}{\nu_\mathrm{D}}\right)^2}} \tag{53}$$

for $|\nu| < \nu_\mathrm{D}$ and is zero elsewhere.

The symbol duration $T_\mathrm{S} = 20.57\ \mu$s is chosen according to the system parameters considered in [11]. The speed of the user varies in the range $0 \le v \le v_{\max} = 100\ \mathrm{km/h} = 27.8\ \mathrm{m/s}$. The carrier frequency is $f_\mathrm{C} = 2$ GHz. This results in a Doppler bandwidth range $0 \le B_\mathrm{D} \le 180$ Hz. Thus, the normalized Doppler bandwidth, $\nu_\mathrm{D} = B_\mathrm{D}T_\mathrm{S}$, ranges in $0 \le \nu_\mathrm{D} \le \nu_{\mathrm{D}\,\max} = 3.8\times 10^{-3}$. The channel is observed over $M = 256$ symbols. During the observation, interval the user travels a distance of at most one wavelength $0 < \nu_\mathrm{D}M < 1$. We are interested in the prediction error at a prediction horizon $\ell = m - M + 1 \in \{32, 64, 128\}$ symbols. At speed $v_{\max}$ the prediction horizon $\ell \in \{32, 64, 128\}$ corresponds to a distance of $\{\lambda/8, \lambda/4, \lambda/2\}$ where $\lambda = c_0/f_\mathrm{C}$ denotes the wavelength. For all simulations the SNR is 10 dB.

B. Analytic Results

In Fig. 2, we use the MSE of the Wiener predictor (denoted Wiener predictor) as lower bound and plot $\mathrm{MSE}[M-1+\ell]$ given by (45) for $\ell \in \{32, 64, 128\}$.

Second, we plot the MSE of the reduced-rank ML predictor (47) calculating the integral (50) numerically (denoted RR ML predictor). The reduced-rank ML predictor uses the exact Doppler spectrum of the fading process, $S'_h(\nu) = S_h(\nu)$. The steps in the curve corresponds to the transitions when the dimension of the (approximation) subspace is increased, see (37).

Third, we show the results for the ME band-limited predictor (denoted ME band-limited) given by $\mathrm{MSE}[M-1+\ell, \mathcal{W}(\nu_\mathrm{D})]$ in (52). For this predictor, we assume no knowledge about the detailed shape of the Doppler spectrum and assume exact knowledge of the Doppler bandwidth only represented by the band-limiting region $\mathcal{W}(\nu_\mathrm{D}) = (-\nu_\mathrm{D}, \nu_\mathrm{D})$. This assumption leads to a subspace spanned by DPS sequences. The dimension switching points obtained by (32) for this case are slightly suboptimal.

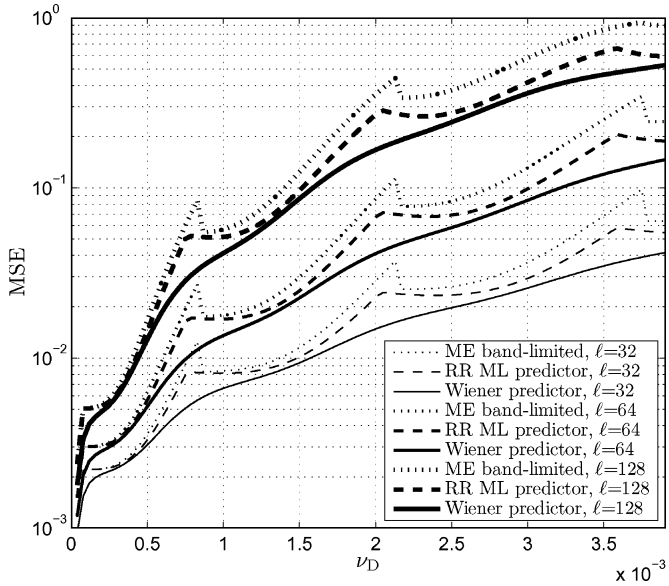

Fig. 2. Mean square prediction error $\mathrm{MSE}[M-1+\ell]$ at prediction horizon $\ell \in \{32, 64, 128\}$ versus the normalized Doppler bandwidth ν_D. The normalized Doppler bandwidth $\nu_\mathrm{D} = 0 \ldots 3.8 \times 10^{-3}$ is induced by a receiver moving with $v = 0 \ldots 27.8$ m/s. The block length $M = 256$. The measurement period relates to a distance traveled by the user in the range of $\nu_\mathrm{D}M = 0 \ldots 1$ wavelengths. The SNR equals 10 dB. We compare a Wiener predictor (Wiener predictor) with a reduced-rank ML predictor (RR ML predictor) and the ME band-limited predictor (ME band-limited).

We can see that the MSE for the reduced-rank ML predictor is larger than the MSE achieved with the Wiener predictor. If we assume knowledge of the Doppler bandwidth only the MSE is increased again. However, the MSE changes by several orders of magnitude with increasing Doppler bandwidth. Thus, for a practical implementation the ME band-limited predictor based on DPS sequences achieves close to optimum results while needing information about the Doppler-bandwidth only. The reason for the small loss incurred by utilizing the Doppler bandwidth information only is rooted in the extremely small dimension of the time-concentrated and band-limited channel subspace.

VII. Dynamic Subspace Selection

In practical systems, information about the Doppler bandwidth must be obtained from channel observations. In [33], a ME band-limited predictor is presented that utilizes the Doppler-bandwidth estimator from [34]. The estimator in [34] assumes a Doppler spectrum according to Clarke's model. Measured Doppler spectra deviate substantially from Clarke's model [20], [21]. Thus, specular Rice components lead to biased Doppler-bandwidth estimates as well as channels with a small number of specular propagation paths. The Doppler-bandwidth estimator proposed in [35] is less sensitive to deviations from Clarke's model, however it requires large observation intervals and an SNR larger than 30 dB.

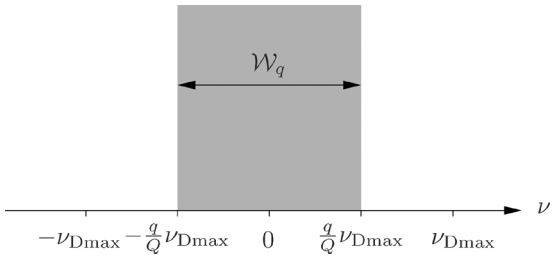

Fig. 3. Symmetric band-limiting region $\mathcal{W}_q=(-q/Q\nu_{\mathrm{D}\,\max},+q/Q\nu_{\mathrm{D}\,\max})$ for $q \in \{1,\ldots,Q\}$ used to define a set of Q subspaces.

We are interested in a low complexity implementation of the channel predictor. To this end we develop a dynamic subspace selection scheme for ME band-limited prediction that does not need an explicit Doppler-bandwidth estimate [23]:

- First, we define a finite number of hypotheses about the actual Doppler bandwidth in Section VII-A. Utilizing the theoretical results from Section III we represent each hypothesis by a subspace spanned by time-concentrated and band-limited sequences. The orthogonal basis vectors spanning each subspace are calculated once and then stored.
- Second, we propose a subspace selection method based on a probabilistic bound on the reconstruction error z (11) in Section VII-B. The subspace with the smallest reconstruction error is selected based on the observation of a single data block. This subspace is used for ME band-limited prediction.

A. Subspace Definition

We define the maximum Doppler bandwidth

$$\nu_{\mathrm{D}\,\max} = \frac{v_{\max} f_{\mathrm{C}} T_{\mathrm{S}}}{c_0} \tag{54}$$

as system parameter given by the maximum (supported) user velocity $v_{\max}$. Furthermore, we define a set of Q subspaces with spectral support

$$\mathcal{W}_q = \left(-\frac{q}{Q}\nu_{\mathrm{D}\,\max}, +\frac{q}{Q}\nu_{\mathrm{D}\,\max}\right) \tag{55}$$

for $q \in \{1,\ldots,Q\}$ as shown in Fig. 3. The selection of Q for a specific simulation scenario is treated in Section X-B. The time-concentrated and band-limited sequences $\{u_i[m,\mathcal{W}_q]\}$ corresponding to the band-limiting region $\mathcal{W}_q$ are calculated according to (21). We define the subspace $\boldsymbol{U}_q = [\boldsymbol{u}_0(\mathcal{W}_q),\ldots,\boldsymbol{u}_{D_q-1}(\mathcal{W}_q)]$ for $q \in \{1,\ldots,Q\}$. The subspace dimension $D_q = D(\mathcal{W}_q)$ is chosen according to (32). The dimension of the subspace spanned by $\boldsymbol{U}_q$ grows with increasing $q \in \{1,\ldots,Q\}$ due to the increasing spectral support $\mathcal{W}_q$ with Lebesgue measure $|\mathcal{W}_q| = 2q/Q\nu_{\mathrm{D}\,\max}$ [cf. (32)]. The subspace orthogonal to $\boldsymbol{U}_q$ is spanned by $\boldsymbol{V}_q = [\boldsymbol{u}_{D_q}(\mathcal{W}_q),\ldots,\boldsymbol{u}_{M-1}(\mathcal{W}_q)]$.

B. Subspace Selection

In [22], an information theoretic subspace selection scheme is proposed. This method uses the observable data error

$$x_q = \frac{1}{M}\|\boldsymbol{y} - \hat{\boldsymbol{h}}_q\|^2 \tag{56}$$

where

$$\hat{\boldsymbol{h}}_q = \boldsymbol{U}_q\boldsymbol{U}_q^{\mathrm{H}}\boldsymbol{y} \tag{57}$$

to obtain an estimate on the reconstruction error

$$z_q = \frac{1}{M}\|\boldsymbol{h} - \hat{\boldsymbol{h}}_q\|^2 \tag{58}$$

which cannot be observed directly. For the subspace selection, $\boldsymbol{h}$ is considered deterministic. The results in [22] are derived for real valued signals. They are adapted here for complex valued signals and noise.

a) Distribution of the Reconstruction Error: The reconstruction error z_q is a sample of a random variable Z_q which is distributed as [22, Lemma 1]

$$\frac{2M}{\sigma_n^2}\left(Z_q - \frac{1}{M}\left\|\boldsymbol{V}_q^{\mathrm{H}}\boldsymbol{h}\right\|^2\right) \sim \chi^2_{2D_q} \tag{59}$$

where $\chi^2_{2D_q}$ is a Chi-square random variable of order $2D_q$. Therefore, Z_q has expected value

$$\mathsf{E}\{Z_q\} = \frac{D_q}{M}\sigma_n^2 + \frac{1}{M}\|\boldsymbol{V}_q\boldsymbol{h}\|^2 \tag{60}$$

and variance

$$\mathrm{var}(Z_q) = \frac{D_q}{M^2}\left(\sigma_n^2\right)^2. \tag{61}$$

b) Distribution of the Data Error: The data error x_q is a sample of a random variable X_q which is distributed as [22, Lemma 2]

$$\frac{2M}{\sigma_n^2}X_q \sim \chi^2_{2(M-D_q)}. \tag{62}$$

Therefore, X_q has expected value

$$\mathsf{E}\{X_q\} = \left(\frac{1}{2} - \frac{D_q}{M}\right)\sigma_n^2 + \frac{1}{M}\|\boldsymbol{V}_q\boldsymbol{h}\|^2 \tag{63}$$

and variance

$$\mathrm{var}(X_q) = \frac{1}{M}\left(\frac{1}{2} - \frac{D_q}{M}\right)\left(\sigma_n^2\right)^2 + \frac{2\sigma_n^2}{M^2}\|\boldsymbol{V}_q\boldsymbol{h}\|^2. \tag{64}$$

c) Probabilistic Lower Bound On The Reconstruction Error: First, assuming $(1/M)\|\boldsymbol{V}_q^{\mathrm{H}}\boldsymbol{h}\|^2$ is known, the reconstruction error z_q is bounded with probability p_1 according to [22, Sec. III.C.]

$$\underline{z'_q}(p_1) \le z_q \le \overline{z'_q}(p_1), \tag{65}$$

where

$$\underline{z'_q}(p_1) = \frac{D_q}{M}\sigma_n^2 + \frac{1}{M}\|\boldsymbol{V}_q^{\mathrm{H}}\boldsymbol{h}\|^2 - G_q(p_1,\sigma_n,2D_q) \tag{66}$$

and

$$\overline{z'_q}(p_1) = \frac{D_q}{M}\sigma_n^2 + \frac{1}{M}\|\boldsymbol{V}_q^{\mathrm{H}}\boldsymbol{h}\|^2 + G_q(p_1,\sigma_n,2D_q). \tag{67}$$

The term $G_q(p_1,\sigma_n,2D_q)$ is calculated by solving

$$p_1 = F\left(2D_q + 2G_q\frac{M}{\sigma_n^2}, 2D_q\right) - F\left(2D_q - 2G_q\frac{M}{\sigma_n^2}, 2D_q\right) \tag{68}$$

numerically for G_q. In (68) $F(x,n)$ denotes the Chi-square cumulative distribution function with n degrees of freedom. Note that G_q is actually independent of the current channel realization [36] which will be important for a practical low complexity implementation as discussed in Section VIII.

Second, we utilize x_q to obtain a probabilistic bound on $(1/M)\|\boldsymbol{V}_q^{\mathrm{H}}\boldsymbol{h}\|^2$. Because $2(M-D_q)$ is large we can invoke the Central Limit Theorem to approximate X_q with a Gaussian random variable. The term $(1/M)\|\boldsymbol{V}_q^{\mathrm{H}}\boldsymbol{h}\|^2$ is bounded with probability p_2 according to [22, Theorem 1]

$$\underline{B_q}(x_q,p_2) \le \frac{1}{M}\|\boldsymbol{V}_q^{\mathrm{H}}\boldsymbol{h}\|^2 \le \overline{B_q}(x_q,p_2). \tag{69}$$

The lower bound $\underline{B_q}$ is zero if $(m_q - \alpha\sqrt{v_q}) \le x_q \le (m_q + \alpha\sqrt{v_q})$, where the probability p_2 is considered in the form $p_2 = \int_{-\alpha}^{\alpha}(1/\sqrt{2\pi})\mathrm{e}^{-x^2/2}\mathrm{d}x$, $m_q = (1/2 - D_q/M)\sigma_n^2$ and $v_q = 1/M(1/2 - D_q/M)\sigma_n^4$. Otherwise, the lower bound is

$$\underline{B_q}(x_q,p_2) = x_q - m_q + \frac{\alpha^2\sigma_n^2}{M} - K_q(\alpha) \tag{70}$$

where

$$K_q(\alpha) = 2\alpha\frac{\sigma_n}{\sqrt{2M}}\sqrt{\frac{\alpha^2\sigma_n^2}{2M} + x_q - \frac{1}{2}m_q}. \tag{71}$$

The upper bound is

$$\overline{B_q}(x_q,p_2) = x_q - m_q + \frac{\alpha^2\sigma_n^2}{M} + K_q(\alpha). \tag{72}$$

Finally, we use the upper bound $\overline{z_q}(x_q,p_1,p_2) \ge \overline{z_q'}(p_1)$,

$$\overline{z_q}(x_q,p_1,p_2) = \frac{D_q}{M}\sigma_n^2 + \overline{B_q}(x_q,p_2) + G_q(p_1,\sigma_n,2D_q) \tag{73}$$

on the reconstruction error to select the appropriate subspace q spanned by the columns of $\boldsymbol{U}_q$

$$\hat{q} = \arg\min_q \overline{z_q}(x_q,p_1,p_2). \tag{74}$$

The chosen subspace $\boldsymbol{U}_{\hat{q}}$ and the associated sequences $\{u_i[m,\mathcal{W}_q]\}$ are used for ME band-limited prediction. The probability p_1 is chosen as $p_1 = Q(\alpha)|_{\alpha=1}$ with $Q(\alpha) = \int_{-\alpha}^{\alpha}(1/\sqrt{2\pi})\mathrm{e}^{-x^2/2}\mathrm{d}x$. The probability $p_2 = Q(8) \approx 1 - 10^{-15}$.

VIII. Complexity

We assume the following known system parameters for the predictor: The block length M, the maximum velocity of the user $v_{\max}$, the maximum Doppler bandwidth $\nu_{\mathrm{D\,max}}$, the number of precalculated subspaces Q, and the operating range of the SNR $\in (\mathrm{SNR}_{\min}, \mathrm{SNR}_{\max})$. The noise variance σ_n^2 is represented by N discrete values over the operating range of the predictor.

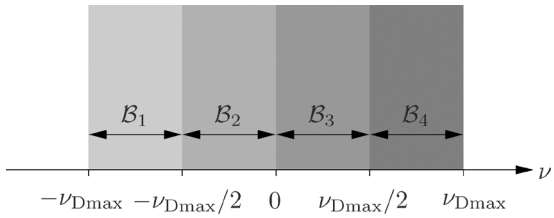

Fig. 4. Asymmetric band-limiting region. Example for $Q' = 4$.

The complexity of the proposed ME band-limited predictor with dynamic subspace selection is mainly determined by the complexity of projecting the observation vector $\boldsymbol{y}$ on all Q subspaces $\boldsymbol{U}_q$ in (57). This operation requires $2(\sum_{q=1}^{Q} D_q)M$ complex multiply accumulate instructions. The storage of all precalculated basis functions needs memory for $QD_q(M+N)$ values.

The computational effort needed to calculate the upper bound on the reconstruction error for each single subspace $\boldsymbol{U}_q$, $q \in \{1,\ldots,Q\}$ in (73) can be neglected since it involves the simple calculation of three terms (p_1, p_2 and α fixed): The first term in (73), $\sigma_n^2 D_q/M$, can be precalculated and needs storage of NQ values. The second term, $\overline{B_q}(x_q,p_2)$, depends directly on the data error x_q. For the calculation of the individual terms of $\overline{B_q}$ we need storage for $5NQ$ precalculated values and the calculation of Q square roots for $K_q(\alpha)$. The third term, $G_q(p_1,\sigma_n,2D_q)$, can be precalculated needing storage of NQ values.

A predictor based on complex exponentials needs Doppler shift estimates for all P paths. Most methods for Doppler shift estimation rely on an eigenvalue decomposition of the channel's sample covariance matrix [37]. The complexity of the eigenvalue decomposition grows with PM^2. Hence, the complexity the ME band-limited predictor with dynamic subspace selection is much smaller then the complexity of complex-exponential-based predictors if $PM > 2(\sum_{q=1}^{Q} D_q)$. For the simulation parameters used in Section X this relation is fulfilled.

IX. Adaption to Disjoint Doppler Spectra

In mobile communication channels, fading processes frequently arise whose spectral support is the union of disjoint intervals. Such short-time Doppler spectra are caused by a nonuniform scatterer distribution or by a small number of specular propagation paths, as reported in [20] and [21]. For such fading processes the set of subspaces defined based on symmetric constant Doppler spectra (55) are suboptimal.

If the number of specular propagation paths is small it is likely that all paths, for example, have either a positive or a negative Doppler shift only. The band-limiting region $\mathcal{W}_q$ defining the subspace $\boldsymbol{U}_q$ in Fig. 3 is symmetric to the origin, hence in this case the support of the band-limiting region is larger than necessary leading to a reduced prediction horizon, see the Monte Carlo simulation results in Section X.

Based on the above explanations we propose to partition the region $(-\nu_{\mathrm{D\,max}}, \nu_{\mathrm{D\,max}})$ into Q' spectral bins with equal length as depicted in Fig. 4. The spectral bin $i \in \{1,\ldots,Q'\}$ spans the interval

$$\mathcal{B}_i = \left[-\nu_{\mathrm{D\,max}} + (i-1)\frac{\nu_{\mathrm{D\,max}}}{Q'}, -\nu_{\mathrm{D\,max}} + i\frac{\nu_{\mathrm{D\,max}}}{Q'}\right]. \tag{75}$$

Using all possible binary combinations of $\mathcal{B}_i$ we can define $2^{Q'}-1$ band-limiting regions $\mathcal{W}'_{q'}$, $q' \in \{1,\ldots,2^{Q'}-1\}$

$$\mathcal{W}'_1 = \mathcal{B}_1 \tag{76}$$
$$\mathcal{W}'_2 = \mathcal{B}_2, \tag{77}$$
$$\mathcal{W}'_3 = \mathcal{B}_1 \cup \mathcal{B}_2 \tag{78}$$
$$\vdots$$
$$\mathcal{W}'_{(2^{Q'}-1)} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_{Q'} = (\nu_{\mathrm{D}\,\max}, \nu_{\mathrm{D}\,\max}). \tag{79}$$

Based on the band-limiting regions $\mathcal{W}'_{q'}$ we can define $\boldsymbol{U}'_{q'}$, $\boldsymbol{V}'_{q'}$ and $D'_{q'}$ similar to the definition in Section VII-A.

X. Monte Carlo Simulations

In this section, we present performance results for ME band-limited prediction using the dynamic subspace selection scheme derived in the previous section. We provide comparisons to a Wiener predictor that utilized the long-term Doppler spectrum and to a classic predictor based on complex exponentials. The predictor using complex exponentials is described in Appendix I.

A. Physical Wave Propagation Channel Model

We simulate the fading process $\{h[m]\}$ using physical wave propagation principles [32], [38]. The electromagnetic field at the receiving antenna is the superposition of the contribution of the individual fields of P impinging plane waves. Each plane wave is conceived as propagating along a specific path. Under these assumptions the channel weight is of the form

$$h[m] = \sum_{p=0}^{P-1} a_p \mathrm{e}^{\mathrm{j}2\pi f_p T_{\mathrm{S}} m} = \sum_{p=0}^{P-1} a_p \mathrm{e}^{\mathrm{j}2\pi \nu_p m}. \tag{80}$$

Here f_p is the Doppler shift of wave p. For easier notation we define the normalized Doppler frequency as $\nu_p = f_p T_{\mathrm{S}}$. Note that $|\nu_p| \leq \nu_{\mathrm{D}} < 1/2$. The gain and phase shift of path p are embodied in the complex weight $a_p \in \mathbb{C}$. We model the random parameter sets a_p and ν_p, $p \in \{0,\ldots,P-1\}$ as independent. The random variables in each set are independent and identically distributed. The path angles α_p are uniformly distributed over $[-\pi,\pi)$. The normalized Doppler shift per path is $\nu_p = \nu_{\mathrm{D}} \cos\alpha_p$. The path weights are defined as $a_p = (1/\sqrt{P})\mathrm{e}^{\mathrm{j}\psi_p}$ where ψ_p is uniformly distributed over $[-\pi,\pi)$. Under the above assumptions, the covariance function of $\{h[m]\}$ converges to $R_h[k] = \mathrm{J}_0(2\pi\nu_{\mathrm{D}}k)$, for $P \to \infty$, where J_0 is the zeroth-order Bessel function of the first kind [32].

We assume a time-variant block-fading channel model comprised of P paths. Hence, the random path parameters a_p and ν_p are assumed to be constant over a block of $M+N$ symbols. However, the path parameters a_p and ν_p change independently from block to block, therefore, the short-time spectrum changes as well [2].

We note that the overidealized simulation models from Jakes [39] or Zheng [40] are *not* suitable for the evaluation of channel prediction algorithms. This is because a symmetric distribution of the scatterers with equidistant spacing is assumed in [39], [40]. However real-world channels will not exhibit equidistantly spaced scatterers. Prediction algorithms assuming a finite number of specular paths [41] show optimistically biased performance due to this overidealized scatterer distribution.

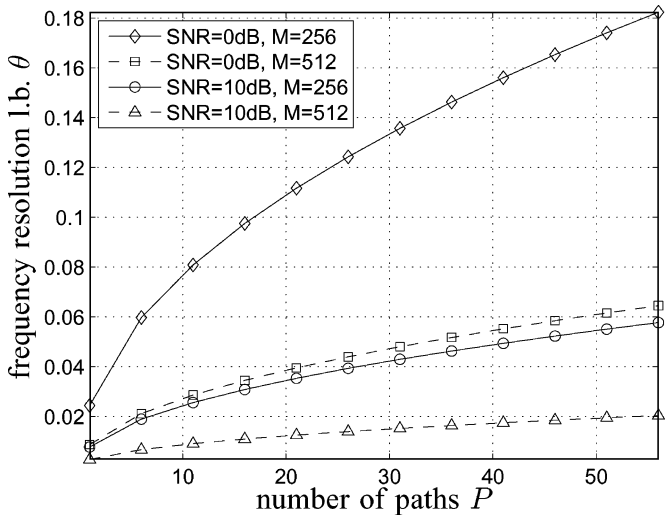

Fig. 5. Lower bound on the frequency resolution (82) versus number of propagation paths P. The $\mathrm{SNR} \in \{0, 10\}$ dB and the block length $M \in \{256, 512\}$.

B. Choice of the Number of Subspaces Q

The lower bound for the frequency estimation error of a single complex exponential in white Gaussian noise is given by the Cramér-Rao lower bound (CRLB) [17, Sec. 15.10],

$$\sqrt{\mathrm{var}(\hat{\nu})} \geq \sqrt{\frac{6\sigma_n^2}{(2\pi)^2 M(M^2-1)}}. \tag{81}$$

This bound applies to the problem of Doppler frequency estimation for $P = 1$ propagation path, too. With increasing number of paths P the estimation error for the Doppler frequency per path increases due to reduced energy per propagation path. The higher likelihood of closely spaced frequencies increases the CRLB additionally [9]. We conjecture that the CRLB for frequency estimation is a lower bound for the problem of Doppler bandwidth estimation for a fading process with $P > 1$ propagation paths.

We are interested to define a finite number of Q hypothesis about the Doppler bandwidth of the fading process (see Fig. 3). To this end we quantize the range $0 \leq \nu_{\mathrm{D}} \leq \nu_{\mathrm{D}\,\max}$ into Q sub-intervals. We define the relative frequency resolution as

$$\theta = \frac{1}{\nu_{\mathrm{D}\,\max}} \sqrt{\frac{6P\sigma_n^2}{(2\pi)^2 M(M^2-1)}} \tag{82}$$

and use its inverse to get an upper bound on the number of subspaces $Q \leq 1/\theta$. In Fig. 5 we plot θ versus the number of paths P for an $\mathrm{SNR} \in \{0, 10\}$ dB. Fig. 5 documents that the lower bound on the frequency resolution is in the order of $0.03\nu_{\mathrm{D}\,\max}$ to $0.11\nu_{\mathrm{D}\,\max}$ for $P = 20$ propagation paths and block length $M = 256$. Since frequencies can only be estimated with 10% accuracy, 10 hypothesis (subspaces) are sufficient to estimate the Doppler bandwidth. We partition the range $0 \leq \nu_{\mathrm{D}} \leq \nu_{\mathrm{D}\,\max}$ into $Q = 10$ intervals, see (55).

C. Simulation Setup and Results

We use the simulation parameters from Section VI-A. For all simulations the SNR is 10 dB. Monte Carlo simulations have been performed to contrast the performance of the three following predictors:

- The Wiener predictor (denoted Wiener predictor) is defined in (36) and utilizes knowledge of the long-term model covariance function $R_h[k] = \mathrm{J}_0(2\pi\nu_\mathrm{D}k)$. This predictor is a sensible choice for large number of propagation paths $P = 30$. However, for $P = 2$ paths this predictor will be suboptimal.
- The predictor based on complex exponential functions (denoted compl. exponential) which is derived based on the specular-path model (80) knows the number of paths P and the Doppler frequencies of each path perfectly. Only the complex path weights are estimated, see Appendix I. This allows us to obtain a lower bound on the performance of the predictor based on complex exponentials.
- The new ME band-limited predictor is presented in two configurations.
  1) We show simulation results (denoted ME band-limited sym.) using the set of symmetric subspaces $\bar{\boldsymbol{U}}_q$, $q \in \{1,\ldots,Q\}$ with $Q = 10$ defined according to Fig. 3 in Section VII-A.
  2) We combine the set of symmetric subspaces $\boldsymbol{U}_q$ with the set of asymmetric subspaces $\boldsymbol{U}'_{q'}$, $q' \in \{1,\ldots,2^{Q'}-1\}$ with $Q' = 4$ defined according to Fig. 4 in Section IX. Both sets contain two identical band-limiting regions, which are $\mathcal{W}_5 = \mathcal{W}'_6 = \mathcal{B}_2 \cup \mathcal{B}_3$ and $\mathcal{W}_{10} = (-\nu_{\mathrm{D}\,\max}, \nu_{\mathrm{D}\,\max}) = \mathcal{W}'_{15} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3 \cup \mathcal{B}_4$. Thus, $\boldsymbol{U}'_6$ and $\boldsymbol{U}'_{15}$ are not used. The subspace selection method from Section VII is applied on the combined set of subspaces. The simulation results are denoted ME band-limited sym.+asym.

The MSE of all three predictors are reported in Figs. 6–10 versus the normalized Doppler frequency ν_D. We present results for the prediction horizons $\ell = m - M + 1 \in \{32, 96\}$ and considering a channel model with $P \in \{2, 30\}$ propagation paths. At speed $v_{\max}$ the prediction horizon $\ell \in \{32, 96\}$ corresponds to a distance of $\{\lambda/8, 3\lambda/8, \}$ where $\lambda = c_0/f_\mathrm{C}$ denotes the wavelength.

All MSE results in the paper are given in terms of the sample mean $\widehat{\mathrm{MSE}}[m] = 1/B \sum_{b=0}^{B-1} |h_b[m] - \hat{h}_b[m]|^2$ where the index b denotes the block number. We average over $B = 1000$ independent channel realizations. We do *not* perform any further normalization or outlier removal as in [5], [6].

D. Discussion of Simulation Results

We can observe that the Wiener predictor performance is independent of the number of propagation paths P. In Fig. 6, we show the results for the ME band-limited predictor with dynamic subspace selection. This predictor shows slightly higher MSE than the Wiener predictor. Note that the Wiener predictor knows the model covariance function exactly while the ME band-limited predictor selects the best subspace based on the observation of a single data block. The results in Fig. 6 agree qualitatively with the analytic results presented in Fig. 2 which where obtained with perfect Doppler bandwidth knowledge.

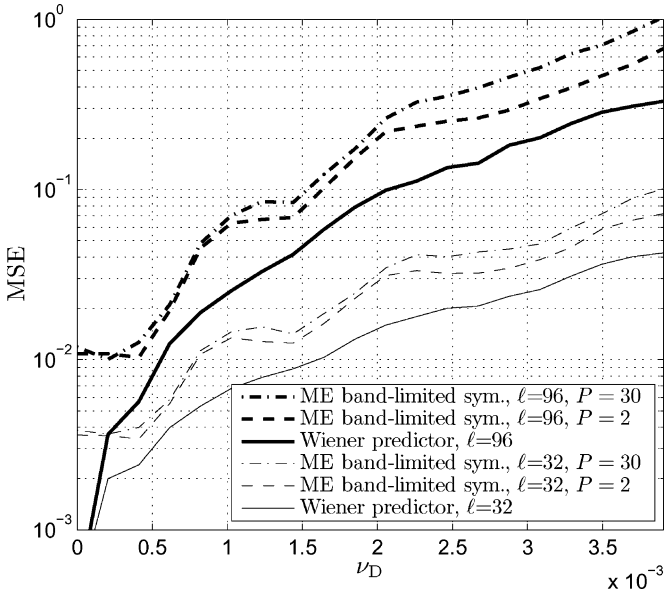

Fig. 6. Mean square prediction error $\mathrm{MSE}[M-1+\ell]$ at prediction horizon $\ell \in \{32, 96\}$ versus the normalized Doppler bandwidth ν_D. The normalized Doppler bandwidth $\nu_\mathrm{D} = 0 \ldots 3.8 \times 10^{-3}$ is induced by a receiver moving with $v = 0 \ldots 27.8$ m/s. The block length $M = 256$. The measurement period relates to a distance traveled by the user in the range of $\nu_\mathrm{D} M = 0 \ldots 1$ wavelengths. The SNR equals 10 dB. We compare a Wiener predictor (Wiener predictor) with a ME band-limited predictor with dynamic subspace selection (ME band-limited sym.).

By combining the set of symmetric and the set of asymmetric subspaces in Figs. 7 and 8 we are able to enhance the performance for small number of paths $P \leq 4$ and large Doppler bandwidth. The dynamic subspace selection has now added freedom to exclude empty regions from the signal subspace. Each spectral bin $\mathcal{B}_1 \ldots \mathcal{B}_4$ in Fig. 4 can be switched on or off individually leading to enhanced prediction performance.

Additionally, we show the performance of the predictor based on complex exponentials. This predictor performs poorly if the Doppler bandwidth ν_D is small, see Figs. 7 and 8. This is because the estimation error for the complex weight of each individual path increases if the Doppler shift difference between two paths becomes small. A detailed analysis of the Cramér-Rao lower bound (CRLB) for the situation of two paths is presented in [9]. For less than four paths and for large Doppler spread the predictor based on complex exponentials performs better than the Wiener predictor (see Figs. 7 and 8). We emphasize that we assumed perfect knowledge of the number of paths P and the Doppler shift of each path ν_p, $p \in \{1,\ldots,P\}$, which leads to highly optimistic performance for the predictor based on complex exponentials. We stick to this assumption to obtain a lower bound on the complex exponential predictor performance.

In Fig. 9, we plot the lower bound for a Wiener predictor by using the exact Doppler power spectral density of the current data block (denoted Wiener pred. inst. spectrum). In this case, the Doppler power spectral density is given as the sum of Dirac pulses at the Doppler shift of each path $S_h(\nu) = \sum_{p=0}^{P-1} E\{|a_p|^2\}\delta(\nu - \nu_p)$. Additionally, we plot the result for the most simple predictor that assumes a band-limiting

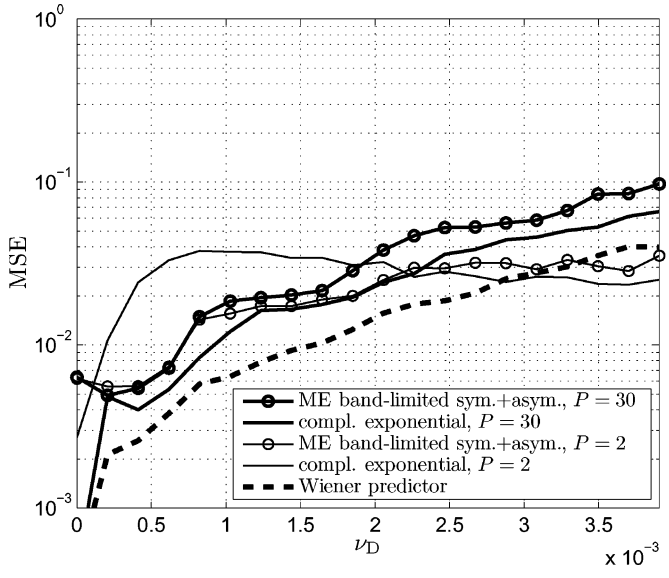

Fig. 7. Mean square prediction error $\mathrm{MSE}[M-1+\ell]$ versus Doppler bandwidth ν_{D} at prediction horizon $\ell = 32$ for a channel with $P \in \{2, 30\}$ propagation paths. We compare a predictor using complex exponential basis functions (compl. exponential) with perfectly known frequencies, the Wiener predictor (Wiener predictor) and the ME band-limited predictor with dynamically selected symmetric and asymmetric subspaces (ME band-limited sym.+asym.). The $\mathrm{SNR} = 10$ dB and the observation block length $M = 256$.

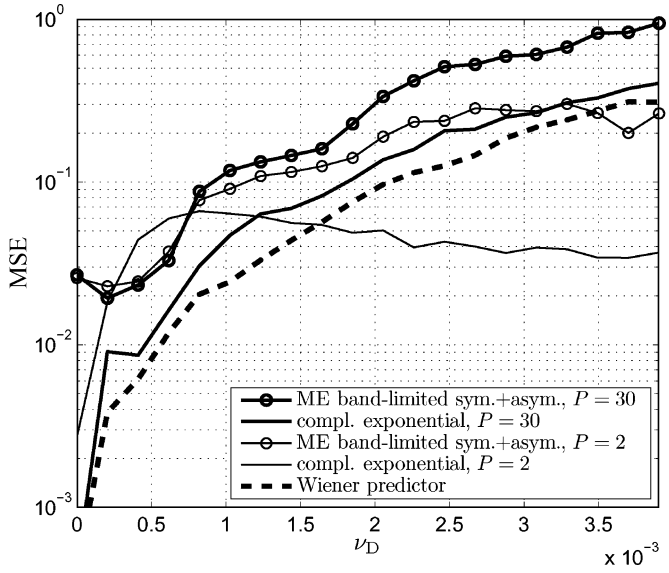

Fig. 8. Mean square prediction error $\mathrm{MSE}[M-1+\ell]$ versus Doppler bandwidth ν_{D} at prediction horizon $\ell = 96$ for a channel with $P \in \{2, 30\}$ propagation paths. The $\mathrm{SNR} = 10$ dB and the observation block length $M = 256$.

region $\mathcal{W} = (-\nu_{\mathrm{D\,max}}, +\nu_{\mathrm{D\,max}})$ according to the maximum velocity $v_{\max}$ (denoted ME band-limited $\nu_{\mathrm{D\,max}}$), see (54). For large Doppler bandwidth, the dynamic subspace selection scheme performs slightly worse than the predictor with the fixed bandlimiting region. This is due to the fact that the subspace selection scheme decides to use the wrong subspace sometimes.

In Fig. 10, we show results for a data block length of $M = 512$. By doubling the observation interval the prediction performance of the Wiener predictor and the predictor based on complex exponentials changes only slightly. However, the subspace

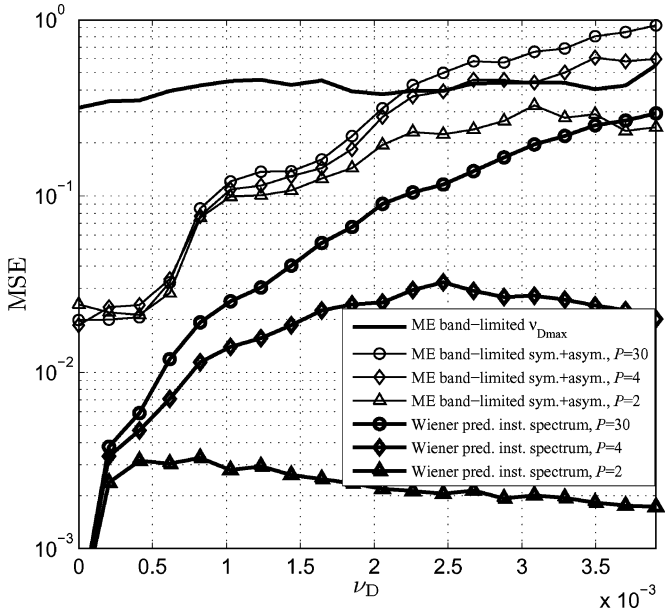

Fig. 9. Mean square prediction error $\mathrm{MSE}[M-1+\ell]$ versus Doppler bandwidth ν_{D} at prediction horizon $\ell = 96$ for a channel with $P \in \{2, 4, 30\}$ propagation paths. We compare the ME band-limited predictor with dynamically selected subspaces with a predictor that uses a fixed subspace according to the maximum Doppler spread (ME band-limited $\nu_{\mathrm{D\,max}}$). As lower bound we show results for a Wiener predictor that knows the instantaneous Doppler frequencies perfectly (Wiener pred. inst. spectrum). The $\mathrm{SNR} = 10$ dB and the observation block length $M = 256$.

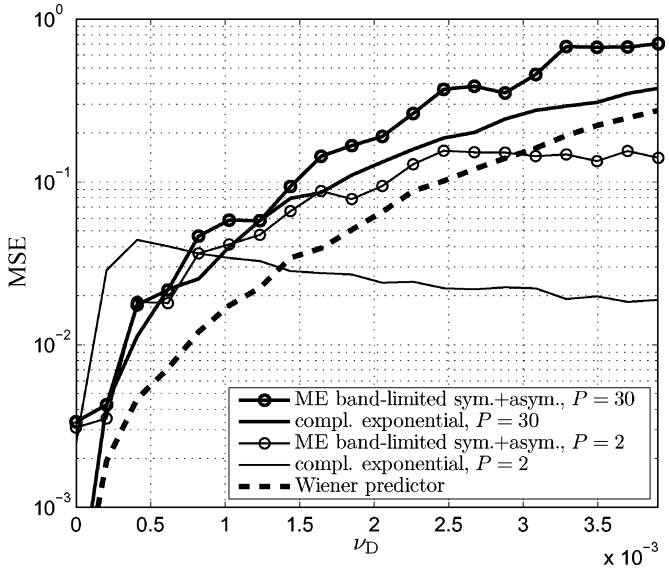

Fig. 10. Mean square prediction error $\mathrm{MSE}[M-1+\ell]$ versus Doppler bandwidth ν_{D} at prediction horizon $\ell = 96$ for a channel with $P \in \{2, 30\}$ propagation paths. The $\mathrm{SNR} = 10$ dB and the observation block length $M = 512$.

selection scheme benefits from the enlarged observation period because the probabilistic estimate of the upper bound on the reconstruction error (73) shows reduced variance.

XI. Conclusion

In this paper, we presented a new ME band-limited prediction method for a time-variant process with arbitrary power spectral density. The predictor is based on time-concentrated and band-

limited sequences. We obtain time-concentrated and band-limited sequences for a band-limiting region consisting of disjoint intervals by generalizing results from Slepian [10].

We showed that ME band-limited prediction is identical to reduced-rank ML prediction for fading processes with constant Doppler spectrum. For a fading process with constant Doppler spectrum, this equivalence allows the conclusion that the subspace underlying the linear prediction problem is energy-concentrated. This fact inherently limits the prediction horizon of linear prediction methods. A similar conclusion can be drawn from the decay of the covariance function $R_h(k)$.

We provided a performance analysis for reduced-rank ML prediction using full information about the channel covariance function and for ME band-limited prediction using information about the Doppler bandwidth only. Numeric evaluation of the prediction error shows that knowledge of the detailed power spectral density is not crucial. We conclude that the predictor performance primarily depends on the Doppler bandwidth, but is almost indifferent to other features of the Doppler spectrum.

We exploit these observations to design a set of subspaces spanned by sequences with fixed time-concentration but growing Doppler bandwidth. The sequences in each subspace exhibit a fixed time-concentration and a subspace-specific bandwidth. Each subspace is matched to the support of a certain Doppler power spectral density. The dimensions of the predefined subspaces are in the range from one to five for practical communication systems. The subspace applied for ME prediction is selected based on a probabilistic bound on the reconstruction error. For a prediction horizon of one eights of a wavelength the numerical simulation results show that the ME band-limited predictor with dynamic subspace selection performs better than or similar to a predictor based on complex exponentials with perfectly known frequencies. For a prediction horizons of three eights of a wavelength the performance of the ME band-limited predictor approaches that of a Wiener predictor with perfectly known Doppler bandwidth.

Appendix I
Channel Prediction Based on Complex Exponential Basis Functions

The ME band-limited predictor described in Section III uses time-concentrated and band-limited sequences to span the channel subspace. Classical channel prediction algorithms describe the channel subspace using complex exponential basis functions. In the method proposed in [4] and [5] the path parameters a_p and ν_p in (80) are estimated to enable channel prediction. We review here shortly the method, so that we are able to compare it with our ME band-limited predictor.

For a limited observation interval $\mathcal{I}_M$, we can rewrite (80) in vector matrix notation according to

$$\boldsymbol{h} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ w_0 & w_1 & \dots & w_{P-1} \\ \vdots & \vdots & & \vdots \\ w_0^{M-1} & w_1^{M-1} & \dots & w_{P-1}^{M-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{P-1} \end{bmatrix} = \boldsymbol{W}\boldsymbol{a} \tag{83}$$

where $w_p = \mathrm{e}^{\mathrm{j}2\pi\nu_p}$. In [4] ESPRIT [42] is used to estimate the Doppler shift ν_p of each single propagation path. ESPRIT requires $P \leq M$. The number of paths P is known as well.

The complex weight vector $\boldsymbol{a} = [a_1, \dots, a_P]^{\mathrm{T}}$ is estimated according to

$$\hat{\boldsymbol{a}} = \left(\hat{\boldsymbol{W}}^{\mathrm{H}}\hat{\boldsymbol{W}} + \sigma_n^2 \boldsymbol{I}_P\right)^{-1} \hat{\boldsymbol{W}}^{\mathrm{H}} \boldsymbol{y} \tag{84}$$

where $\hat{\boldsymbol{W}}$ results by inserting the Doppler estimates in $\boldsymbol{W}$. Finally, the time-variant channel is predicted via

$$\hat{h}[m] = \sum_{p=0}^{P-1} \hat{a}_p \mathrm{e}^{\mathrm{j}2\pi\nu_p m} \tag{85}$$

for $m \in \{M, \dots, M+N-1\}$. In this paper we assume that all ν_p, $p \in \{0, \dots, P-1\}$ are known exactly which allows us to obtain a lower bound on the performance of the predictor based on complex exponentials.

Appendix II
Proof of Theorem 1

Consider square-summable sequences $u[m] \in \ell^2(\mathbb{Z})$, which are band-limited to a region $\mathcal{W} \subseteq (-1/2, 1/2)$, i.e., whose Fourier transform $U(\nu)$ vanishes outside $\mathcal{W}$, see (18) and (19). Furthermore, the energy-concentration on an arbitrary but fixed index set $\mathcal{I} \subset \mathbb{Z}$ should be maximal, i.e.,

$$\rho = \frac{\sum_{m\in\mathcal{I}} |u[m]|^2}{\sum_{m=-\infty}^{\infty} |u[m]|^2} \to \max. \tag{86}$$

Using Parseval's theorem ρ can be recast as

$$\rho = \frac{\int_{\mathcal{W}} \int_{\mathcal{W}} \sum_{m\in\mathcal{I}} \mathrm{e}^{\mathrm{j}2\pi(\nu-\nu')m} U(\nu) U^*(\nu') \mathrm{d}\nu \mathrm{d}\nu'}{\int_{\mathcal{W}} |U(\nu)|^2 \mathrm{d}\nu}. \tag{87}$$

It can be seen, that ρ is maximum if, and only if, $U(\nu)$ satisfies the integral equation

$$\int_{\mathcal{W}} K(\nu, \nu') U(\nu) \mathrm{d}\nu = \gamma U(\nu') \tag{88}$$

with the Hermitian kernel $K(\nu, \nu')$ defined as

$$K(\nu, \nu') = \sum_{m\in\mathcal{I}} \mathrm{e}^{\mathrm{j}2\pi(\nu-\nu')m}. \tag{89}$$

Notice that inserting (88) into (87) yields

$$\gamma = \rho. \tag{90}$$

Since the kernel K is degenerate and has the specific form (89), the solutions of (88) are finite and can be found by writing the left-hand side (LHS) of (88) as [43]

$$\int_{\mathcal{W}} K(\nu, \nu') U_i(\nu) \mathrm{d}\nu = \sum_{\ell\in\mathcal{I}} u_i[\ell] e^{-\mathrm{j}2\pi\nu'\ell}. \tag{91}$$

In the above expression, i is an indexing of the solutions. Substituting in (88) and replacing ν' by ν results in

$$\sum_{\ell\in\mathcal{I}} u_i[\ell]e^{-\mathrm{j}2\pi\nu\ell} = \rho_i U_i(\nu). \tag{92}$$

Multiplying both sides with $e^{\mathrm{j}2\pi\nu m}$ and integrating with respect to ν over the region $\mathcal{W}$ yields

$$\sum_{\ell\in\mathcal{I}} u_i[\ell] \int_{\mathcal{W}} e^{\mathrm{j}2\pi\nu(m-\ell)}\mathrm{d}\nu = \rho_i u_i[m], \quad m\in\mathbb{Z}. \tag{93}$$

This identity is exactly the defining (21) of the band-limited and time-concentrated sequences in Theorem 1.

Note that in [27], identity (90) is derived assuming that the defining (21) is already known. The present proof yields all results by maximizing the energy-concentration of the sought sequence in the interval $\mathcal{I}$ only. In [10], identity (90) is proven for a symmetric band-limiting region $\mathcal{W} = (-\nu_\mathrm{D}, \nu_\mathrm{D})$.

The proof presented here is valid for any Lebesgue-measurable subset of $(-1/2, +1/2)$ and in particular when $\mathcal{W}$ is the union of disjoint intervals as considered in Section III-B. Moreover, throughout the paper we consider the special case $\mathcal{I} = \mathcal{I}_M$.

ACKNOWLEDGMENT

The authors thank J. Wehinger and E. Riegler for their helpful comments.

REFERENCES

[1] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Reading, MA: Addison-Wesley, 1991.

[2] I. Viering, *Analysis of Second Order Statistics for Improved Channel Estimation in Wireless Communications*, ser. Fortschritts-Berichte VDI Reihe. Düsseldorf, Germany: VDI Verlag GmbH, 2003, no. 733.

[3] G. Matz, "On non-WSSUS wireless fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2465–2478, Sep. 2005.

[4] J. B. Andersen, J. Jensen, S. H. Jensen, and F. Frederiksen, "Prediction of future fading based on past measurements," in *Proc. 50th IEEE Veh. Technol. Conf. (VTC)—Fall*, Sep. 19–22, 1999, vol. 1, pp. 151–155.

[5] M. Chen, M. Viberg, and T. Ekman, "Two new approaches to channel prediction based on sinusoidal modelling," presented at the 13th Workshop on Statistical Signal Processing (SSP), Bordeaux, France, Jul. 17–20, 2005.

[6] M. Chen, T. Ekman, and M. Viberg, "New approaches for channel prediction based on sinusoidal modelling," *EURASIP J. Appl. Signal Process.*, 2006, to be published.

[7] T. Ekman, "Prediction of mobile radio channels—modeling and design," Ph.D. dissertation, Uppsala Univ., Uppsala, Sweden, 2002.

[8] S. Semmelrodt, "Methoden zur prädiktiven Kanalschätzung für adaptive Übertragunstechniken im mobilfunk," Ph.D. dissertation, Kassel Univ., Kassel, Germany, 2003.

[9] B. H. Fleury, M. Tschdin, R. Heddergott, D. Dalhaus, and K. I. Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 434–450, Mar. 1999.

[10] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty—V: The discrete case," *The Bell Syst. Tech. J.*, vol. 57, no. 5, pp. 1371–1430, May–Jun. 1978.

[11] T. Zemen and C. F. Mecklenbräuker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Trans. Signal Process.*, vol. 53, no. 9, pp. 3597–3607, Sep. 2005.

[12] R. J. Lyman, W. W. Edmonson, S. McCullough, and M. Rao, "The predictability of continuous-time, bandlimited processes," *IEEE Trans. Signal Process.*, vol. 48, no. 2, pp. 311–316, Feb. 2000.

[13] R. J. Lyman and W. W. Edmonson, "The prediction of bandlimited processes with flat spectral densities," *IEEE Trans. Signal Process.*, vol. 49, no. 7, pp. 1564–1569, Jul. 2001.

[14] R. J. Lyman, "Optimal mean-square prediction of the mobile-radio fading envelope," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 819–824, Mar. 2003.

[15] S. Dharanipragada and K. S. Arun, "Bandlimited extrapolation using time-bandwidth dimension," *IEEE Trans. Signal Process.*, vol. 45, no. 12, pp. 2951–2966, Dec. 1997.

[16] F. A. Dietrich and W. Utschik, "Pilot-assisted channel estimation based on second-order statistics," *IEEE Trans. Signal Process.*, vol. 53, no. 3, pp. 1178–1193, Mar. 2005.

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

[18] A. K. Jain and S. Ranganath, "Extrapolation algorithms for discrete signals with application in spectral estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-29, no. 4, pp. 830–845, Aug. 1981.

[19] T. Zemen, C. F. Mecklenbräuker, and B. H. Fleury, "Time-variant channel prediction using time-concentrated and band-limited sequences—Analytic results," in *5th Vienna Symp. Math. Model. (MATHMOD)*, Vienna, Austria, Feb. 8–10, 2006, (Invited Paper).

[20] G. Acosta and M. A. Ingram, "Model development for the wideband vehicle-to-vehicle 2.4 GHz channel," presented at the IEEE Wireless Commun. Netw. Conf. (WCNC), Las Vegas, NV, Apr. 3–6, 2006.

[21] X. Zhao, J. Kivinen, P. Vainikainen, and K. Skog, "Characterization of Doppler spectra for mobile communications at 5.3 GHz," *IEEE Trans. Veh. Technol.*, vol. 52, no. 1, pp. 14–23, Jan. 2003.

[22] S. Beheshti and M. A. Dahleh, "A new information-theoretic approach to signal denoising and best basis selection," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3613–3624, Oct. 2005.

[23] T. Zemen, C. F. Mecklenbräuker, and B. H. Fleury, "Minimum-energy bandlimited time-variant channel prediction with dynamic subspace selection," presented at the Proc. 14th Eur. Signal Process. Conf. (EUSIPCO), Florence, Italy, Sep. 4–8, 2006, (Invited Paper).

[24] T. Zemen, C. F. Mecklenbräuker, J. Wehinger, and R. R. Müller, "Iterative joint time-variant channel estimation and multi-user detection for MC-CDMA," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1469–1478, Jun. 2006.

[25] T. Zemen, C. F. Mecklenbräuker, J. Wehinger, and R. R. Müller, "Iterative multi-user decoding with time-variant channel estimation for MC-CDMA," in *Proc. 5th Int. Conf. 3G Mobile Commun. Technol.*, London, U.K., Oct. 18–20, 2004, pp. 88–92, (Invited Paper).

[26] H. J. Landau and H. Widom, "Eigenvalue distribution of time and frequency limiting," *J. Math. Anal. Applicat.*, vol. 77, pp. 469–481, 1980.

[27] S. Dharanipragada, "Time-bandwidth dimension and its application to signal reconstruction," Master's thesis, Univ. Illinois at Urbana-Champaign, Urbana-Champaign, 1991.

[28] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, Singapore: McGraw-Hill, 1991.

[29] L. L. Scharf and D. W. Tufts, "Rank reduction for modeling stationary signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 3, pp. 350–355, Mar. 1987.

[30] M. Nicoli, O. Simenone, and U. Spagnolini, "Multislot estimation of fast-varying space-time communication channels," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1184–1195, May 2003.

[31] M. Niedzwiecki, *Identification of Time-Varying Processes*. New York: Wiley, 2000.

[32] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell Syst. Tech. J.*, p. 957, Jul.–Aug. 1968.

[33] T. Zemen, C. F. Mecklenbräuker, and B. H. Fleury, "Time-variant channel prediction using time-concentrated and band-limited sequences," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Istanbul, Turkey, May 2006, pp. 5660–5665.

[34] J. Holtzmann and A. Sampath, "Adaptive averaging methodology for handoffs in cellular systems," *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 59–66, Feb. 1995.

[35] K. E. Baddour and N. C. Beaulieu, "Robust Doppler spread estimation in nonisotropic fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2677–2682, Nov. 2005.

[36] S. Beheshti, 2006, private communication.

[37] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Upper Saddle River, NJ: Prentice-Hall, 2000.

[38] H. Hofstetter and G. Steinböck, "A geometry based stochastic channel model for MIMO systems," presented at the ITG Workshop Smart Antennas, Munich, Germany, Jan. 2004.

[39] W. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[40] Y. R. Zheng and C. Xiao, "Simulation models with correct statistical properties for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 51, no. 6, pp. 920–928, Jun. 2003.
[41] I. C. Wong and B. L. Evans, "Joint channel estimation and prediction for OFDM systems," in *Proc. IEEE Global Telecommun. Conf. (IEEE GLOBECOM)*, St. Louis, MO, Nov. 2005, vol. 4, pp. 2255–2259.
[42] R. Roy and T. Kailath, "ESPRIT—Estimation of signal parameters by rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.
[43] D. Porter and D. S. G. Stirling, *Integral Equations*. Cambridge, U.K.: Cambridge Univ. Press, 1990.

Thomas Zemen was born in Mödling, Austria, in 1970. He received the Dipl.-Ing. degree (with distinction) in electrical engineering and the doctoral degree (with distinction) both from Vienna University of Technology, Vienna, Austria, in 1998 and 2004, respectively.

He joined Siemens, Austria, in 1998 where he worked as hardware engineer and project manager for the Radio Communication Devices Department. From October 2001 to September 2003, he was delegated by Siemens Austria as a researcher to the mobile communications group at the Telecommunications Research Center Vienna (ftw.). Since October 2003, he has been with the Telecommunications Research Center Vienna, working as researcher in the strategic I0 project. His research interests include orthogonal frequency division multiplexing (OFDM), multiuser detection, time-variant channel estimation, iterative MIMO receiver structures, and distributed signal processing. Since May 2005, he has led the project "Future Mobile Communications Systems—Mathematical Modeling, Analysis, and Algorithms for Multi Antenna Systems," which is funded by the Vienna Science and Technology Fund (Wiener Wissenschafts-, Forschungs- und Technologiefonds, WWTF). He also teaches MIMO Communications as external lecturer at the Vienna University of Technology.

Christoph F. Mecklenbräuker (S'88–M'97) was born in Darmstadt, Germany, in 1967. He received the Dipl.-Ing. degree in electrical engineering from Technische Universität Wien, Austria, in 1992 and the Dr.-Ing. degree from Ruhr-Universität Bochum, Germany, in 1998, both with distinction.

He was with Siemens, Vienna, from 1997 to 2000, where he participated in the European framework of ACTS 90 Future Radio Wideband Multiple Access System (FRAMES). He was a delegate to the Third Generation Partnership Project (3GPP) and engaged in the standardization of the radio access network for the Universal Mobile Telecommunications System (UMTS). From 2000 to 2006, he has held a senior research position with the Forschungszentrum Telekommunikation Wien (ftw.), Wien, Austria, in the field of mobile communications. In 2006, he joined the Faculty of Electrical Engineering and Information Technology as a Full Professor with the Technische Universität Wien, Austria. He has authored approximately 80 papers in international journals and conferences, for which he has also served as a reviewer, and holds eight patents in the field of mobile cellular networks. His current research interests include antenna array and MIMO-signal processing for wireless systems, and ultrawideband radio.

Dr. Mecklenbräuker is a member of the IEEE Signal Processing, Antennas and Propagation, and Vehicular Technology Societies, and EURASIP. His doctoral dissertation on matched field processing received the Gert-Massenberg Prize in 1998.

Florian Kaltenberger was born in Vienna, Austria, in 1978. He received the Diploma degree (Dipl.-Ing.) and the Ph.D. degree, both in technical mathematics, from the Vienna University of Technology, Vienna, Austria, in 2002 and 2007, respectively.

During summer 2001, he held an internship position with British Telecom, BT Exact Technologies, Ipswich, U.K., where he was working on mobile video conferencing applications. After his studies, he started as a Research Assistant with the Vienna University of Technology, Institute for Advanced Scientific Computing, working on distributed signal processing algorithms. In 2003, he joined the Wireless Communications Group, Austrian Research Centers GmbH, where he is currently working on the development of low-complexity smart antenna and MIMO algorithms, as well as on the ARC SmartSim real-time hardware channel emulator.

Bernard H. Fleury (SM'99) received the Diploma degree in electrical engineering and mathematics in 1978 and 1990, respectively, and the doctoral degree in electrical engineering in 1990, all from the Swiss Federal Institute of Technology Zurich (ETHZ), Switzerland.

From 1978 to 1985 and 1988 to 1992, he was a Teaching Assistant and Research Assistant, respectively, at the Communication Technology Laboratory and at the Statistical Seminar at ETHZ. In 1992, he joined again the former laboratory as Senior Research Associate. Since 1997, he has been with the Department of Communication Technology, Aalborg University, Denmark, where he is currently Professor in digital communications. He has also been affiliated with the Telecommunication Research Center, Vienna (ftw.) since April 2006. He is presently Chairman of Department 2 Radio Channel Modelling for Design Optimisation and Performance Assessment of Next Generation Communication Systems of the ongoing FP6 network of excellence NEWCOM (Network of Excellence in Communications). His general fields of interest cover numerous aspects within communication theory and signal processing mainly for wireless communications. His current areas of research include stochastic modelling and estimation of the radio channel, characterization of multiple-input multiple-output (MIMO) channels, and iterative (turbo) techniques for joint channel estimation and data detection/decoding in multiuser communication systems.