Adaptive Reduced-Rank Estimation of Nonstationary Time-Variant Channels Using Subspace Selection

Thomas Zemen, *Senior Member, IEEE*, and Andreas F. Molisch, *Fellow, IEEE*

***Abstract*—In this paper, we focus on adaptive time-variant channel estimation for vehicle-to-vehicle (V2V) communications in intelligent transportation systems (ITS) using the IEEE 802.11p physical layer. The IEEE 802.11p pilot pattern is identical to that in the well-known IEEE 802.11a/g (WiFi) standard, which was initially designed for indoor environments with little or no mobility. However, in a V2V drive-by situation, the channel impulse response rapidly changes due to the high relative velocity between transmitter and receiver, as well as the changes in the scattering environment. Hence, for such V2V channels, advanced decision directed channel estimation methods are needed to reach a frame error rate (FER) smaller than 10^{-1}. Even more importantly, the channels are nonstationary, which implies that the Doppler power spectral density (DSD) and the power delay profile (PDP) change on a timescale comparable with the frame length, which complicates the estimator design. In this paper, we develop an adaptation method for the channel estimation filter that is suitable for the following: 1) the short frame length in IEEE 802.11p; 2) the given pilot structure; and 3) the requirement of only a single received short frame for filter adaptation. We define a set of hypotheses on the support of the DSD and a second set of hypotheses on the support of the PDP. Each hypothesis is represented by a specific subspace spanned by orthogonal basis vectors. For basis vector calculation, we develop a numerically stable algorithm utilizing generalized discrete prolate spheroidal sequences. The adaptation algorithm chooses a hypothesis from both sets such that a probabilistic bound on the channel estimation error is minimized. We implement the hypothesis test by means of a novel subspace selection algorithm that allows utilizing correlated observations of a time- and frequency-selective (2-D) fading process. We validate the adaptive channel estimation scheme in an IEEE 802.11p compliant link level simulation for a relative velocity range from 0 to 111 m/s $\approx$ 400 km/h $\approx$ 249 mi/h. Adaptive filtering enables an up to fourfold reduction in the number of required iterations to reach an FER below 10^{-1} for an $E_b/N_0 = 12$ dB.**

***Index Terms*—Adaptive filter, discrete prolate spheroidal sequences, hypothesis test, IEEE 802.11p, intelligent transportation system (ITS), non-stationary fading process, vehicular communications.**

Manuscript received April 2, 2012; revised June 18, 2012; accepted July 25, 2012. Date of publication August 1, 2012; date of current version November 6, 2012. The work of T. Zemen was supported by the FTW project NFN SISE S106 funded by the Austrian Science Fund, as well as NOWIRE, funded by the Vienna Science and Technology Fund. FTW Forschungszentrum Telekommunikation Wien was supported by the Austrian Government and the City of Vienna within the competence center program COMET. The review of this paper was coordinated by Dr. T. Taniguchi.

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

A. F. Molisch is with the University of Southern California, Los Angeles, CA 90089 USA (e-mail: molisch@usc.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2012.2211053

I. Introduction

WIRELESS vehicle-to-vehicle (V2V) communications is the foundation for intelligent transportation systems (ITS), which allow, *inter alia*, the reduction of traffic accidents and the mitigation of congestion [1]. The main candidate for the physical layer of ITS is the IEEE 802.11p standard [2], which is based on orthogonal frequency-division multiplexing (OFDM).

A main challenge for V2V communications are the rapidly changing radio propagation conditions that strongly differ from cellular wireless networks. In V2V communications, both the transmitter and the receiver are mobile, and the scattering environment can rapidly change. Many measurement campaigns have demonstrated high Doppler and delay spreads and even nonstationarities of the channel statistics [3]–[7]. Consequently, adaptive channel estimation is particularly important, but particularly difficult as well, in such an environment.

The situation for adaptive channel estimation is further exacerbated by the following three aspects.

1) The training signal in the IEEE 802.11p standard is identical to that in the well-known IEEE 802.11a/g (WiFi) standard, which was initially designed for indoor environments with little or no mobility. An IEEE 802.11p frame starts with two pilot OFDM symbols containing all subcarriers, whereas throughout the frame only four subcarriers are allocated to carry pilot information. Hence, for V2V propagation channels that have both a large delay spread and Doppler spread, the sampling theorem is violated in the time and frequency domain [5], and advanced decision directed channel estimation techniques are required [8]–[12].
2) The relative velocity on (European) motorways for V2V communications can reach up to 400 km/h $\approx$ 111 m/s $\approx$ 249 mi/h, leading to high channel dynamics in drive-by situations.
3) In safety-related ITS applications, data are typically transmitted by short frames with a low duty cycle. Hence, the data stream in ITS is bursty. Consequently, channel estimation has to happen on a frame-by-frame basis and cannot rely on interpolation between frames.

Related Literature: Adaptive filters for time-variant channels [13]–[17] estimate the covariance function of the channel fading process iteratively. They offer good performance at moderate complexity; however, an initial training phase on the order of more than 150 symbols is required. This prohibits their application to frame-based communication systems like

IEEE 802.11p, where the frame length is only in the order of $M = 21, \ldots, 137$ OFDM symbols.

In [18]–[20], compressed sensing methods are investigated for time-variant frequency-selective channel estimation. These algorithms perform well for channels with a small number of scatterers or clusters of scatterers. For V2V channels, which are the focus of this paper, a channel estimator needs to deal with a mixture of specular reflections and substantial diffuse components [21]. These substantial diffuse contributions stem from reflections along the roadside. Compressed sensing methods also have specific requirements like a randomized pilot pattern, which is not possible in IEEE 802.11p without substantial nonbackward-compatible standard modifications.

Another approach is the application of a robust Wiener filter for the estimation of the time- and frequency-selective channel [22], [23]. The filter is designed according to two *a priory* known system parameters, the maximum support of the Doppler power spectral density (DSD), and the maximum support of the power delay profile (PDP). A delay-Doppler scattering function [24] prototype with flat spectrum is used to calculate the filter coefficients.

Such a robust Wiener filter design is clearly mismatched to the actual shape of the scattering function. However, its performance in terms of mean square error (MSE) shows only a minor degradation for a wide parameter range compared with the matched Wiener filter [22]. A low-rank approximation of the robust Wiener filter [25], [26] allows for a substantial computational complexity reduction and is studied in a number of papers [27]–[30]. In vehicle-to-infrastructure propagation channels, such as that treated by Zemen *et al.* [30], the difference of the actual support and the assumed fixed maximum support of DSD and PDP causes only a minor degradation of the iterative channel estimator.

The iterative channel estimation presented in [30] allows achieving a frame error rate (FER) that is essentially the same as that obtained with perfect channel state information (CSI). However, due to the specific pilot pattern of IEEE 802.11p, the number of iterations needed in this scheme increases with increasing velocity and increasing frame length.

In this paper, we are concerned with channels where the support of the scattering function changes quickly over a wide range. This situation occurs typically in V2V channels due to the high relative velocity between two vehicles in drive-by scenarios as well as the fast propagation environment changes. The approach of [30] to assume a fixed maximum support of the DSD and PDP would lead to a substantial increase in the channel estimation error after the first iteration as well as an increased number of iterations for V2V channel estimation to reach convergence. To improve the channel estimation, an estimate for the support of the DSD and PDP is required for each single short frame. This fact introduces a new set of problems whose solution is at the center of this paper.

In [31], a first step toward a frame-based DSD support estimate for the prediction of a frequency-flat and time-selective channel is presented. The algorithm from [31] is successfully validated in [32] with narrow band vehicular channel measurement data. However, it does not handle frequency-selective channels, which usually occur in V2V communications.

Contributions of This Paper:

1) We provide a new formulation of the subspace-based channel estimator from [30] in the framework of robust reduced-rank Wiener filters. We analyze the subspace structure of the related covariance matrix prototype and provide a factorization of its eigenvalue decomposition into time- and frequency-dependent terms. This factorization enables a numerically stable calculation of the eigenvectors and eigenvalues of the covariance matrix utilizing generalized discrete prolate spheroidal (DPS) sequences. The robust reduced-rank Wiener filter furthermore allows complexity reduction of more than three orders of magnitude in an IEEE 802.11p system compared with the full rank filter.
2) We derive a subspace selection algorithm for correlated time- and frequency-selective (2-D) channel observations. This algorithm is based on the numerically stable eigenvector factorization of the covariance matrix and allows a frame-by-frame adaptation of the robust Wiener filter to varying propagation conditions. We also design a specific partitioning of the IEEE 802.11p pilot pattern for the subspace selection. To the best of our knowledge, this is the first time that a rapid adaptation method is demonstrated for time- and frequency-selective nonstationary channel estimation in V2V communication channels that is only based on a single short frame.
3) We provide a performance analysis of subspace selection for time-varying and frequency-selective fading processes with a frame length constrained and study its numerical complexity for the IEEE 802.11p standard in terms of floating point operations (FLOPS).
4) We show analytically as well as numerically that the MSE of the robust, but mismatched, Wiener filter approaches the MSE of the exact Wiener filter for small support of the DSD and small block length.
5) We validate our adaptive reduced-rank channel estimator in a practical vehicular communication scenario using an IEEE 802.11p compliant link level simulation [5], [33]. We use the FER results from [30] as base line and demonstrate the improvements obtained with the new 2-D subspace selection algorithm. These improvements are either in terms of reduced FER when keeping the number of iterations fixed or in a reduced number of iterations for a constant FER target.

Notation: We denote a scalar by a, 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^*. For the discrete set I, $|I|$ denotes the number of elements of I, and for the continuous region W, $|W|$ denotes the Lebesgue measure of W. The Frobenius (2-norm) of a matrix or vector is denoted by $\|\boldsymbol{A}\|$. The largest (smallest) integer that is lower (greater) than or equal to $a \in \mathbb{R}$ is denoted by $\lfloor a \rfloor$ ($\lceil a \rceil$). We denote the set of all integers by $\mathbb{Z}$, the set of

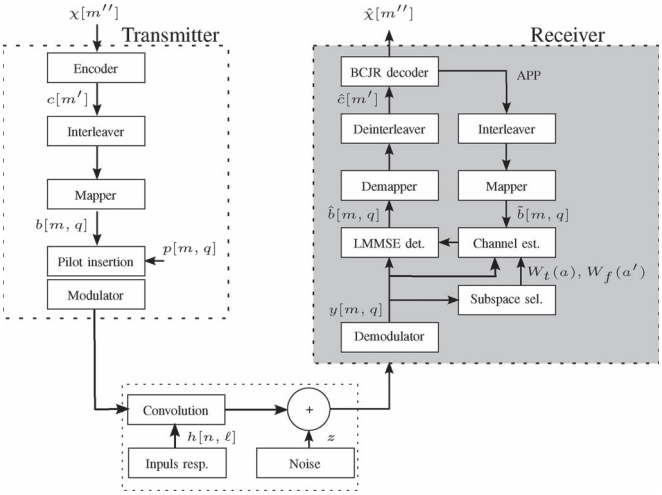

Fig. 1. Schematic structure of the 802.11p link-level simulator with adaptive iterative channel estimation.

real numbers by $\mathbb{R}$, and the set of complex numbers by $\mathbb{C}$. The Kronecker product is denoted by $\otimes$.

Organization of This Paper: The OFDM signal model for time-variant frequency-selective channels is presented in Section II. In Section III, we introduce the robust reduced-rank Wiener filter for iterative channel estimation and analyze the eigenvalue structure of its covariance matrix. Based on this analysis, we develop a hypothesis design and a hypothesis test for subspace selection for time- and frequency-selective (2-D) channel observations in Section IV. The MSE of the robust reduced-rank Wiener filter is analyzed in Section V. In Section VI, detailed numerical performance results of the hypothesis test are presented as well as FER results of the IEEE 802.11p system with the adaptive channel estimation algorithm. The numerical complexity of the hypothesis test and the robust reduced-rank Wiener filter is analyzed in Section VII. We draw conclusions in Section VIII.

II. Signal Model

We consider an OFDM-based communication system utilizing an even number of subcarriers N and a cyclic prefix with length G. Frames (data packets) are transmitted at random times, where we assume here (in line with the usual V2V scenarios) that interframe transmission times are much larger than the duration of a frame and much larger than the channel coherence time. We furthermore assume that the interframe transmission times are larger than the time over which the channel statistics remain stationary. Each frame has a duration of M samples and utilizes a bandwidth B, and the sampling rate at the receiver side is chosen as $1/T_C = B$. The OFDM symbol duration is given by $T_S = (N+G)T_C$.

Each frame contains $S = |\mathcal{S}|$ coded data symbols $b[m,q]$ $\forall\ [m,q] \in \mathcal{S}$, where $\mathcal{S}$ denotes the 2-D data symbol position index set in the time–frequency plane, $m \in \{0,\dots,M-1\}$ denotes discrete time, and $q \in \{0,\dots,N-1\}$ denotes the subcarrier index, respectively. For $[m,q] \notin \mathcal{S}$, we define $b[m,q] = 0$.

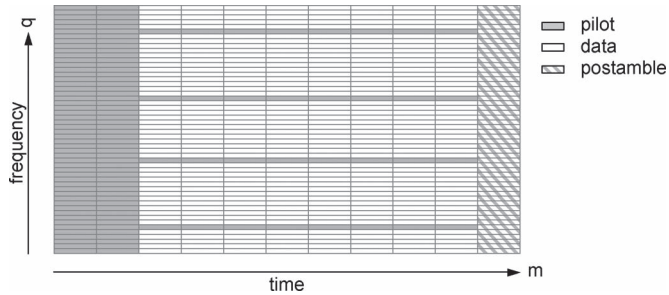

Fig. 2. Pilot pattern in an 802.11 OFDM frame, extended by an OFDM pilot symbol at the end of the frame (postamble) [30].

The binary information sequence $\chi[m'']$ of length $2SR_C$ is convolutionally encoded with code rate R_C, resulting in a sequence of code bits $c[m']$ (see Fig. 1). After interleaving and quadrature phase shift keying (QPSK) modulation with Gray labeling, the data symbols are mapped onto the OFDM time–frequency grid

$$b[\mathcal{S}_s] = \frac{1}{\sqrt{2}}\left(c[2s] + \mathrm{j}c[2s+1]\right) \quad \forall s \in \{0,\dots,S-1\} \quad (1)$$

where $\mathcal{S}_s$ denotes the sth element $[m_s, q_s]$ of the set $\mathcal{S}$ in the OFDM time–frequency grid with $s \in \{0,\dots,S-1\}$.

In each frame, $P = |\mathcal{P}|$ pilot symbols $p[m,q]$ $\forall\ [m,q] \in \mathcal{P}$ are transmitted, where $\mathcal{P}$ denotes the pilot symbol position index set. For $[m,q] \notin \mathcal{P}$, we define $p[m,q] = 0$. The two sets $\mathcal{P}$ and $\mathcal{S}$ are nonoverlapping. Hence, by adding the pilot symbols $p[m,q]$, we can perform a multiplexing operation, giving

$$d[m,q] = b[m,q] + p[m,q]. \quad (2)$$

The pilot pattern for IEEE 802.11p with an additional postamble OFDM pilot symbol[1] (as suggested in [30]) is depicted in Fig. 2.

The time- and frequency-selective V2V communication channel, including the transmit and receive filter, is described by the impulse response $h(t,\tau)$. The sampled impulse response is defined as

$$h[n,\ell] := h(nT_C, \ell T_C) \quad (3)$$

where n denotes discrete time at rate $1/T_C$, and ℓ denotes discrete delay at rate $1/T_C$. We assume $h[n,\ell]$ has its essential support within $0 \le \ell \le L-1 = \lceil \tau_{\mathrm{Pmax}}/T_C \rceil$, where τ_{Pmax} denotes the maximum path delay and L the corresponding number of samples. To avoid intersymbol interference, we also assume that the maximum number of samples of the impulse response is shorter than or equal to the cyclic prefix length $L \le G$. The time-variant frequency response sampled at subcarriers with spacing $1/T_S$ is defined as

$$g[m,q] = \sum_{\ell=0}^{G-1} e^{\frac{-\mathrm{j}2\pi q\ell}{N}} \quad h[m(N+G),\ell], \quad q \in \{0,\dots,N-1\}. \quad (4)$$

[1]To reduce the required number of iterations and, at the same time, the chip set complexity, [30] proposes a backward compatible pilot pattern improvement for IEEE 802.11p in the form of a postamble. With this modified pilot pattern, the channel estimation complexity can be reduced by a factor of 2–3.

The received signal after cyclic prefix removal and discrete Fourier transform can be expressed as

$$y[m,q] = g[m,q]d[m,q] + z[m,q] \tag{5}$$

and $z[m,q] \sim \mathcal{CN}(0,\sigma_z^2)$ denotes symmetric complex additive white Gaussian noise with zero mean and covariance σ_z^2. See [30] for a discussion of the intercarrier interference robustness of the IEEE 802.11p standard.[2]

The output of a linear minimum MSE (LMMSE) detector

$$\hat{b}[m,q] = \frac{y[m,q]\hat{g}[m,q]^*}{\sigma_z^2 + |\hat{g}[m,q]|^2} \quad \forall \quad (m,q) \in \mathcal{S} \tag{6}$$

provides symbol estimates $\hat{b}[m,q]\,\forall\,(m,q) \in \mathcal{S}$. In (6), we denote by $\hat{g}[m,q]$ the channel estimate at time index m and subcarrier index q. After demapping and deinterleaving, we calculate log-likelihood ratios as input for the max-log maximum *a posteriori* (MAP) decoder[3] [34], [35] as $2\hat{c}[m']/\hat{\sigma}_z^2$, assuming equally likely code bits $\{-1,+1\}$. The noise variance estimate $\hat{\sigma}_z^2$ is calculated as

$$\hat{\sigma}_z^2 = \frac{1}{2S}\sum_{m'=0}^{2S-1}\left(|\hat{c}[m']| - \hat{\mu}_c\right)^2 \tag{7}$$

where

$$\hat{\mu}_c = \frac{1}{2S}\sum_{m'=0}^{2S-1}\hat{c}[m']. \tag{8}$$

These estimates are required to compensate for the channel estimation error in the first iterations, resulting in a biased LMMSE detection in (6) [36].

III. Iterative Reduced-Rank Channel Estimation

We start by defining the linear Wiener filter for the iterative channel estimator utilizing soft symbol feedback [36] from the max-log MAP decoder. We define the receive vector

$$\boldsymbol{y} = [y[0,0],\ldots,y[0,N-1],\ldots, y[M-1,0],\ldots,y[M-1,N-1]]^{\mathrm{T}}. \tag{9}$$

Similarly, we define the symbol $\boldsymbol{d}$, channel $\boldsymbol{g}$, and noise vector $\boldsymbol{z}$, respectively. With these prerequisites, we can define the signal model for channel estimation as

$$\boldsymbol{y} = \boldsymbol{D}\boldsymbol{g} + \boldsymbol{z} \tag{10}$$

where $\boldsymbol{D} = \mathrm{diag}(\boldsymbol{d})$.

[2]Note that for the numeric link level simulations (see Section VI-D), we implement a time-variant convolution with $h[n,\ell]$ (3); hence, all intercarrier interference effects are still fully present in the received signal.

[3]The max-log MAP decoder is advantageous because it allows a robust operation in the log domain from a numerical point of view. The max-log MAP decoder could be also replaced by a soft-output Viterbi algorithm. This would lead to some performance degradation but allows a complexity reduction of about a factor of 2. See [34] for a detailed comparison and analysis.

We use the soft symbol feedback from the max-log MAP decoder to enhance the channel estimates iteratively. The soft symbols $\tilde{b}[\mathcal{S}_s]$ are defined according to

$$\tilde{b}[\mathcal{S}_s] = \mathop{\mathbb{E}}_{b}^{(\mathrm{APP})}\{b[\mathcal{S}_s]\} \tag{11}$$

$$= \frac{1}{\sqrt{2}}\left(\mathop{\mathbb{E}}_{c}^{(\mathrm{APP})}\{c[2s]\} + \mathrm{j}\mathop{\mathbb{E}}_{c}^{(\mathrm{APP})}\{c[2s+1]\}\right) \tag{12}$$

where $\mathcal{S}_s$ denotes the sth element $[m_s,q_s]$ of the data symbol index set $\mathcal{S}$, with $s \in \{0,\ldots,S-1\}$ [see the definition in (1)]. The two expectation terms in (12)

$$\mathop{\mathbb{E}}_{c}^{(\mathrm{APP})}\{c[m']\} = 2\Pr^{(\mathrm{APP})}\{c[m'] = +1 \,|\, \hat{c}[m']\} - 1 \tag{13}$$

calculate the expectation over the alphabet of the code bits c, which is $\{-1,+1\}$. By $\Pr^{(\mathrm{APP})} \in [0,1]$, we denote the *a posteriori* probability (APP) for the code symbol being $+1$ if $\hat{c}[m']$ is observed, where $\hat{c}[m']$ is obtained from the equalizer output $\hat{b}[m,q]$ after demapping and deinterleaving [36] (see Fig. 1).

The absolute value of the soft symbol is bounded as $0 \leq |\tilde{b}[m,q]| \leq 1$. Hence, soft symbols that have a high degree of certainty have an absolute value close to 1. Soft symbols with low confidence have a small absolute value close to zero.

In the first iteration, only pilot symbols are used for channel estimation. From the second iteration, pilot symbols together with soft symbols are used. We define $\tilde{\boldsymbol{\mathcal{D}}} = \mathrm{diag}(\tilde{\boldsymbol{d}})$, where $\tilde{\boldsymbol{d}}$ has elements

$$\tilde{d}[m,q] = \tilde{b}[m,q] + p[m,q] \tag{14}$$

containing the soft symbol feedback $\tilde{b}[m,q]$ [cf. with (2)].

A. Wiener Filter

Following the derivation in [36, (30)–(39)], we obtain the Wiener filter for $\boldsymbol{g}$ in (10) as

$$\hat{\boldsymbol{g}} = \boldsymbol{R}_g\tilde{\boldsymbol{D}}^{\mathrm{H}}\left(\tilde{\boldsymbol{D}}\boldsymbol{R}_g\tilde{\boldsymbol{D}}^{\mathrm{H}} + \boldsymbol{\Lambda} + \sigma_z^2\boldsymbol{I}_{MN}\right)^{-1}\boldsymbol{y} \tag{15}$$

where $\boldsymbol{R}_g = \mathbb{E}\{\boldsymbol{g}\boldsymbol{g}^{\mathrm{H}}\}$ denotes the covariance matrix of $\boldsymbol{g}$, and

$$[\boldsymbol{\Lambda}]_{m+Mq,m+Mq} = [\boldsymbol{R}_g]_{m+Mq,m+Mq}\left(1 - \left|\tilde{d}[m,q]\right|^2\right) \tag{16}$$

for $m \in \{0,\ldots,M-1\}$ and $q \in \{0,\ldots,N-1\}$. The entries on the diagonal of $\boldsymbol{\Lambda}$ become 0 for time–frequency grid positions $[m,q]$, where either a pilot symbol is transmitted or for which we have high confidence in the decision of the max-log MAP decoder on the transmitted data symbol.

The complexity of the Wiener filter (15) is dominated by the matrix inversion and scales with $\mathcal{O}((MN)^3)$ [37]. For the parameter set of IEEE 802.11p, this becomes prohibitive for a chip set implementation. Therefore, we are interested in analyzing the subspace structure of the covariance matrix $\boldsymbol{R}_g$ and use only the dominant subspace for a reduced-rank low-complexity implementation.

B. Robust Reduced-Rank Wiener Filter

As pointed out in Section I, $\boldsymbol{R}_g$ is not known at the receiver side. To tackle this problem, we define a robust (mismatched) Wiener filter [22], [23]. We approximate $\boldsymbol{R}_{\boldsymbol{g}}$ with a covariance matrix $\tilde{\boldsymbol{R}}_g \approx \boldsymbol{R}_{\boldsymbol{g}}$ assuming a delay-Doppler scattering function prototype [24] with flat spectrum in a 2-D region defined by the Cartesian product

$$\mathcal{W} = W_t \times W_f = [-\nu_D, \nu_D] \times [0, \theta_P] \tag{17}$$

where W_t defines the support region of the DSD with $0 \leq \nu_D \leq \nu_{\mathrm{Dmax}}$, and W_f defines the support region of the PDP with $0 \leq \theta_P \leq \theta_{\mathrm{Pmax}}$. The maximum normalized (one sided) Doppler bandwidth $\nu_{\mathrm{Dmax}} = f_{\mathrm{Dmax}} T_S$, and the maximum normalized path delay $\theta_{\mathrm{Pmax}} = \tau_{\mathrm{Pmax}}/(NT_C)$.

The fading process is observed on a finite index set

$$\mathcal{I} = I_t \times I_f = [0, \ldots, M-1] \times [-N'/2, \ldots, N'/2-1] \tag{18}$$

in the time and frequency domain $[m, q] \in \mathcal{I}$, with $N' = N - N_g$, where N_g denotes the guard band width, I_t the observation interval in the time domain, and I_f the observation interval in the frequency domain, respectively.

For a flat delay-Doppler scattering function prototype with support $\mathcal{W}$, the covariance matrix $\tilde{\boldsymbol{R}}_g$ can be factorized as [38]

$$\tilde{\boldsymbol{R}}_g = \boldsymbol{R}(W_t, I_t) \otimes \boldsymbol{R}(W_f, I_f) \tag{19}$$

where the elements of $\boldsymbol{R}(W, I)$ are defined as

$$[\boldsymbol{R}(W, I)]_{k,\ell} = \frac{1}{|W|} C[k - \ell, W] \tag{20}$$

for $k, \ell \in I$ with

$$C[k, W] = \int_W e^{\mathrm{j}2\pi k\nu} d\nu = \frac{1}{\mathrm{j}2\pi k} (e^{\mathrm{j}2\pi k\nu_2} - e^{\mathrm{j}2\pi k\nu_1}) \tag{21}$$

and $W = [\nu_1, \nu_2]$.

The matrix $\tilde{\boldsymbol{R}}_g$ is a full rank matrix, but it is not directly accessible by conventional eigensolvers due to its specific eigenvalue spectrum (high condition number). However, as we will see, its eigenvector decomposition is a crucial aspect for both the low complexity formulation of the reduced-rank Wiener filter and the filter adaptation using the subspace selection algorithm.

1) Numerically Stable Eigenvector Factorization: To obtain a numerically stable eigenvector factorization of $\tilde{\boldsymbol{R}}_g$, we exploit its factorization (19) into a time- and a frequency-dependent term. The eigenvectors of $\boldsymbol{R}(W, I)$ are the generalized DPS sequences $u_i[m, W, I]$, $i \in I$ for the band limit W and time-limited to the observation interval $m \in I$ [31], [38], [39]. They fulfill the identity

$$\boldsymbol{R}(W, I)\boldsymbol{u}_i(W, I) = \lambda_i(W, I)\boldsymbol{u}_i(W, I) \tag{22}$$

where

$$\boldsymbol{u}_i(W, I) = [u_i(0, W, I), \ldots, u_i(|I| - 1, W, I)]^{\mathrm{T}}. \tag{23}$$

The eigenvalues $\lambda_i(W, I)$ are close to $1/|W|$ for $i \leq D'$, where D' denotes the time–bandwidth product

$$D'(W, I) = \lceil |W||I| \rceil + 1. \tag{24}$$

For index values $i > D'$, the eigenvalues rapidly decay. In compact notation, we can write

$$\boldsymbol{R}(W, I) = \boldsymbol{U}(W, I)\boldsymbol{\sigma}(W, I)\boldsymbol{U}(W, I)^{\mathrm{H}} \tag{25}$$

where the columns of $\boldsymbol{U}(W, I)$ are given by the generalized DPS sequences $\boldsymbol{u}_i(W, I)$, and the diagonal matrix $\boldsymbol{\sigma}(W, I)$ has elements $\lambda_i(W, I)$.

To solve the eigenvalue problem for $\boldsymbol{R}(W, I)$ (and $\tilde{\boldsymbol{R}}_g$), we can make use of the classical result of Slepian *et al.* [39] that identified a specific tridiagonal matrix that commutes with $\boldsymbol{R}(W, I)$. This tridiagonal matrix enables the stable eigenvector calculation since it has the following: 1) a better conditioned eigenvalue spectrum and 2) the same eigenvectors as $\boldsymbol{R}(W, I)$. For more details, see [39] and [40].

The eigenvector decomposition of $\tilde{\boldsymbol{R}}_g$ is formally defined as

$$\tilde{\boldsymbol{R}}_g = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{U}^{\mathrm{H}}. \tag{26}$$

It follows from (19) that the eigenvector matrix $\boldsymbol{U}$ can be factorized as

$$\boldsymbol{U} = \boldsymbol{U}(\mathcal{W}, \mathcal{I}) = \Pi\left(\boldsymbol{U}(W_t, I_t) \diamond \boldsymbol{U}(W_f, I_f)\right) \tag{27}$$

where the operator $\diamond$ denotes the Tracy–Singh product of column-wise partitioned matrices [41], [42]. The diagonal eigenvalue matrix $\boldsymbol{\Sigma}$ is given by

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\mathcal{W}, \mathcal{I}) = \Pi\left(\mathrm{diag}\left(\boldsymbol{\sigma}(W_t, I_t) \otimes \boldsymbol{\sigma}(W_f, I_f)\right)\right). \tag{28}$$

The permutation operator $\Pi(\cdot)$ is chosen such that the columns of $\boldsymbol{\Sigma}$ (and $\boldsymbol{U}$) are sorted according to $\lambda_0(\mathcal{W}, \mathcal{I}) \geq \lambda_1(\mathcal{W}, \mathcal{I}) \geq \cdots \geq \lambda_{(|\mathcal{I}|-1)}(\mathcal{W}, \mathcal{I})$.

2) Low-Complexity Reduced-Rank Filter: We use D dominant eigenvectors to approximate the covariance matrix $\tilde{\boldsymbol{R}}_g$ as

$$\tilde{\boldsymbol{R}}_g \approx \breve{\boldsymbol{R}}_g = \boldsymbol{U}_D\boldsymbol{\Sigma}_D\boldsymbol{U}_D^{\mathrm{H}} \tag{29}$$

where $\boldsymbol{U}_D$ and $\boldsymbol{\Sigma}_D$ contain the first D columns of $\boldsymbol{U}$ and $\boldsymbol{\Sigma}$, respectively. The dimension $D = D(\mathcal{W}, \mathcal{I})$ minimizing the MSE for a given noise variance σ_z^2 is found to be [43]

$$D = \underset{\mathcal{D} \in \{1, \ldots, |\mathcal{I}|\}}{\arg\min} \left(\frac{1}{|\mathcal{W}||\mathcal{I}|} \sum_{i=\mathcal{D}}^{|\mathcal{I}|-1} \lambda_i(\mathcal{W}, \mathcal{I}) + \frac{\mathcal{D}}{|\mathcal{I}|}\sigma_z^2 \right) \tag{30}$$

where the two summands on the right-hand side represent the square bias and the variance term.

Inserting (29) into (15), we obtain the robust reduced-rank Wiener filter

$$\hat{\boldsymbol{g}} = \boldsymbol{U}_D\boldsymbol{\Sigma}_D\boldsymbol{U}_D^{\mathrm{H}}\tilde{\boldsymbol{D}}^{\mathrm{H}} \left(\tilde{\boldsymbol{D}}\boldsymbol{U}_D\boldsymbol{\Sigma}_D\boldsymbol{U}_D^{\mathrm{H}}\tilde{\boldsymbol{D}}^{\mathrm{H}} + \underbrace{\boldsymbol{\Lambda} + \sigma_z^2\boldsymbol{I}_{MN}}_{\triangleq \boldsymbol{\Psi}} \right)^{-1} \boldsymbol{y}. \tag{31}$$

In this expression, it is necessary to invert an $MN \times MN$-dimensional matrix. Therefore, we apply the matrix inversion lemma, and the final expression becomes

$$\hat{\boldsymbol{g}} = \boldsymbol{U}_D \left(\boldsymbol{U}_D^{\mathrm{H}} \tilde{\boldsymbol{D}}^{\mathrm{H}} \boldsymbol{\Psi}^{-1} \tilde{\boldsymbol{D}} \boldsymbol{U}_D + \boldsymbol{\Sigma}_D^{-1} \right)^{-1} \boldsymbol{U}_D^{\mathrm{H}} \tilde{\boldsymbol{D}}^{\mathrm{H}} \boldsymbol{\Psi}^{-1} \boldsymbol{y}. \tag{32}$$

In (32), the first matrix inversion must be calculated for a $D \times D$ dimensional matrix, where $D \ll MN$, allowing for a substantial complexity reduction in terms of FLOPS. The second inversion of the diagonal matrix $\boldsymbol{\Psi} = \boldsymbol{\Lambda} + \sigma_z^2 \boldsymbol{I}_{MN}$ in (32) does not contribute significantly to the computational effort. See Section VII for a detailed complexity analysis.

The foregoing derivation leading to (32) provides a sound motivation for the factorization of $\tilde{\boldsymbol{R}}_g$ in a frequency- and time-domain subspace. In [30], we conjectured that this factorization is possible and used ad hoc subspace arguments to obtain equivalent results [30, eq. (18) and (31)].

After having established (32) and the factorization of $\boldsymbol{U}$ (27) and $\boldsymbol{\Sigma}$ (28), we are interested in obtaining an estimate of the actual support of the DSD $W_t = [-\nu_D, \nu_D] \subset [-\nu_{\mathrm{Dmax}}, \nu_{\mathrm{Dmax}}]$ and of the support of the PDP $W_f = [0, \theta_P] \subset [0, \theta_{\mathrm{Pmax}}]$. How this can be achieved using a single observed short frame and a given nonregular pilot pattern, like that in IEEE 802.11p, will be discussed in the next section.

IV. Hypothesis Test for Correlated Two-Dimensional Channel Observations

We are interested in a low-complexity channel estimator that is well adapted to the situation in ITS applications. The data are typically transmitted by short frames with a low duty cycle.

The direct estimation of the support of the DSD and PDP would lead to high estimation errors [44]–[46] due to the short frame length. Hence, we define a finite set of hypotheses on the support of the DSD $\{W_t(1), \ldots, W_t(A)\}$ and on the support of the PDP $\{W_f(1), \ldots, W_f(A')\}$. These hypotheses will be tested for each single received frame. Based on the result of the hypothesis test, we can choose the best parameterization of the robust Wiener filter and perform channel estimation.

A. Hypothesis Design

For the symmetric support of the DSD, we define A hypotheses

$$W_t(a) = \left(-\frac{a}{A}\nu_{\mathrm{Dmax}}, \frac{a}{A}\nu_{\mathrm{Dmax}} \right) \tag{33}$$

for $a \in \{1, \ldots, A\}$ (see Fig. 3).

For the asymmetric support of the PDP, we define A' hypotheses

$$W_f(a') = \left(0, \frac{a'}{A'}\theta_{\mathrm{Pmax}} \right) \tag{34}$$

for $a' \in \{1, \ldots, A'\}$ (see Fig. 4).

Each hypothesis is represented by a subspace spanned by the columns of $\boldsymbol{U}(W_t(a), I_t)$ for $a \in \{1, \ldots, A\}$ and $\boldsymbol{U}(W_f(a'), I_f)$ for $a' \in \{1, \ldots, A'\}$. These $A + A'$ matrices are precalculated and stored.

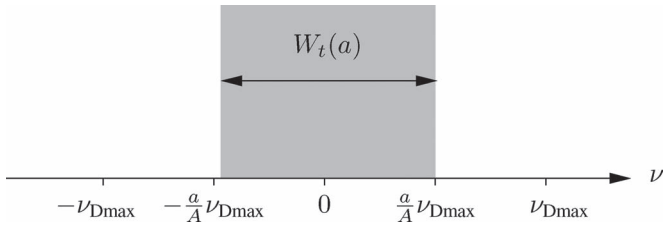

Fig. 3. Symmetric support hypothesis for the DSD $W_t(a) = (-a/A\nu_{\mathrm{Dmax}}, a/A\nu_{\mathrm{Dmax}})$, $a \in \{1, \ldots, A\}$.

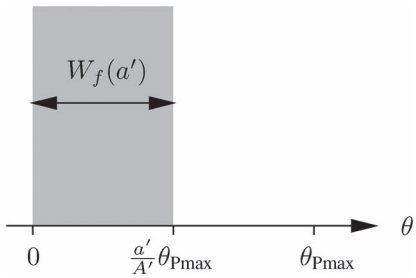

Fig. 4. Asymmetric support hypothesis for the PDP $W_f(a') = (0, a'/A'\theta_{\mathrm{Pmax}})$, $a' \in \{1, \ldots, A'\}$.

B. Hypothesis Test by Subspace Selection

We base our hypothesis test on the subspace selection algorithm introduced in [47] and used in [31] for a frequency-flat channel prediction algorithm. Here, we extend the subspace selection algorithm to correlated 2-D observations of a time- and frequency-selective channel with a nonregular pilot pattern.

For the hypothesis test, we define the signal model for channel estimation only at the pilot positions

$$\boldsymbol{y}^{(\mathcal{P})} = \boldsymbol{D}^{(\mathcal{P})} \boldsymbol{g}^{(\mathcal{P})} + \boldsymbol{z}^{(\mathcal{P})} \tag{35}$$

with $\boldsymbol{D}^{(\mathcal{P})} = \mathrm{diag}(\boldsymbol{d}^{(\mathcal{P})})$, where $\boldsymbol{y}^{(\mathcal{P})}$, $\boldsymbol{d}^{(\mathcal{P})}$, $\boldsymbol{g}^{(\mathcal{P})}$, and $\boldsymbol{z}^{(\mathcal{P})}$ contain the respective elements for $[m, q] \in \mathcal{P}$ in the same order as in $\boldsymbol{y}$ (9). The channel observations at the pilot positions $\mathcal{P}$ are given by

$$\boldsymbol{w}^{(\mathcal{P})} = \boldsymbol{D}^{(\mathcal{P})^{\mathrm{H}}} \boldsymbol{y}^{(\mathcal{P})} = \boldsymbol{g}^{(\mathcal{P})} + \boldsymbol{D}^{(\mathcal{P})^{\mathrm{H}}} \boldsymbol{z}^{(\mathcal{P})} = \boldsymbol{g}^{(\mathcal{P})} + \boldsymbol{z}'^{(\mathcal{P})} \tag{36}$$

where $\boldsymbol{z}'^{(\mathcal{P})} \sim \mathcal{CN}(0, \sigma_z^2 \boldsymbol{I}_{|\mathcal{P}|})$ has the same statistics as $\boldsymbol{z}^{(\mathcal{P})}$.

The 2-D observations of the fading process at the pilot positions $\mathcal{P}$ collected in the vector $\boldsymbol{g}^{(\mathcal{P})}$ span a certain subspace $\mathcal{G}$. According to the Karhunen–Loéve expansion [48], the same subspace $\mathcal{G}$ is also spanned by the eigenvectors of the covariance matrix

$$\tilde{\boldsymbol{R}}_{\boldsymbol{g}^{(\mathcal{P})}} = \mathbb{E}\left\{ \boldsymbol{g}^{(\mathcal{P})} \boldsymbol{g}^{(\mathcal{P})^{\mathrm{H}}} \right\}. \tag{37}$$

The pilot pattern $\mathcal{P}$ of IEEE 802.11p (see Fig. 2) cannot directly be factorized in a Cartesian product of a time and a frequency component. Hence, the numerically stable computation of the eigenvalues and eigenvectors of $\tilde{\boldsymbol{R}}_{\boldsymbol{g}^{(\mathcal{P})}}$, as discussed in Section III-B1 with (27) and (28), is not directly applicable. However, by inspecting Fig. 2, it becomes clear that we can split the pilot pattern $\mathcal{P}$ into two parts that can indeed be factorized.

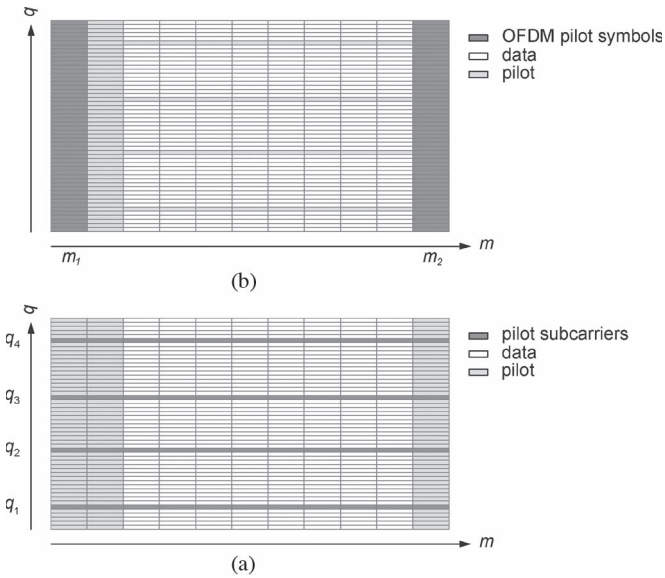

Fig. 5. (a) Pilot subcarriers $\{q_1, q_2, q_3, q_4\}$ defining the index set $\mathcal{T}$ used to obtain an estimate of the DSD support. (b) OFDM pilot symbols $\{m_1, m_2\}$ defining the index set $\mathcal{F}$ used to obtain an estimate of the PDP support.

1) Pilot Pattern Partitioning: The first part consists out of the four pilot subcarriers at frequency index $I_{\mathcal{P}_f} = \{q_1, q_2, q_3, q_4\}$. The subcarrier pilot index set is defined as

$$\mathcal{T} = I_t \times I_{\mathcal{P}_f}. \tag{38}$$

We have depicted the pilot subcarrier index set in Fig. 5(a) in dark gray. We can utilize (35) and (37) by replacing $^{(\mathcal{P})}$ with $^{(\mathcal{T})}$ to define the signal model

$$\boldsymbol{y}^{(\mathcal{T})} = \boldsymbol{D}^{(\mathcal{T})}\boldsymbol{g}^{(\mathcal{T})} + \boldsymbol{z}^{(\mathcal{T})} \tag{39}$$

for the pilot subcarrier index set $\mathcal{T}$, as well as the correlation matrix

$$\tilde{\boldsymbol{R}}_{\boldsymbol{g}^{(\mathcal{T})}} = \mathbb{E}\left\{\boldsymbol{g}^{(\mathcal{T})}\boldsymbol{g}^{(\mathcal{T})^{\mathrm{H}}}\right\}. \tag{40}$$

The subspace spanned by the fading process $\boldsymbol{g}^{(\mathcal{T})}$ is also spanned by the columns of the matrix $\boldsymbol{U}(\mathcal{W}, \mathcal{T})$.

The second part consists out of two OFDM pilot symbols at time index $I_{\mathcal{P}_t} = \{m_1, m_2\}$ [see Fig. 5(b)]. The OFDM pilot symbol index set is defined as

$$\mathcal{F} = I_{\mathcal{P}_t} \times I_f. \tag{41}$$

Similarly as for the first part the subspace spanned by the fading process, $\boldsymbol{g}^{(\mathcal{F})}$ is also spanned by the columns of the matrix $\boldsymbol{U}(\mathcal{W}, \mathcal{F})$. The full pilot set $\mathcal{P} = \mathcal{T} \cap \mathcal{F}$.

2) Reduced Complexity Subspace Selection: With the definition of $\boldsymbol{U}(\mathcal{W}, \mathcal{T})$ and $\boldsymbol{U}(\mathcal{W}, \mathcal{F})$ spanning the subspace of the fading process observed via pilots on the index set $\mathcal{T}$ and $\mathcal{F}$, respectively, we have everything in place to develop the hypothesis test for subspace selection.

We are interested to find the best hypothesis for the DSD support from the hypotheses $\{W_t(a), a \in \{1, \ldots, A\}\}$. Similarly, we are interested to find the best PDP support hypothesis from the finite set $\{W_f(a'), a' \in \{1, \ldots, A'\}\}$. This requires the test of AA' different hypotheses. To avoid this high number of tests, we propose the following reduced complexity algorithm that needs to test only $A + A'$ different hypotheses.

1) We test the support of the DSD with the subcarrier pilot index set $\mathcal{T}$. For now, we assume that the PDP has maximum support $W_f(A')$. The subspace spanned by the fading process $\boldsymbol{g}^{(\mathcal{T})}$ can be approximated by the columns of the matrix $\boldsymbol{U}(W_t(a) \times W_f(A'), I_t \times I_{\mathcal{P}_f})$. We collect its first $D_a = D(W_t(a) \times W_f(A'), I_t \times I_{\mathcal{P}_f})$ columns in matrix $\boldsymbol{U}_a$ and the remaining columns in matrix $\boldsymbol{V}_a$.With $\boldsymbol{U}_a$, $a \in \{1, \ldots, A\}$, we can obtain A reduced-rank maximum likelihood channel estimates [26]

$$\hat{\boldsymbol{g}}_a = \boldsymbol{U}_a\boldsymbol{U}_a^{\mathrm{H}}\boldsymbol{w}^{(\mathcal{T})} \tag{42}$$

testing all A hypotheses for the support of the DSD. We can express the data error as

$$x_a = \frac{1}{|\mathcal{T}|}\|\boldsymbol{w}^{(\mathcal{T})} - \hat{\boldsymbol{g}}_a\|^2. \tag{43}$$

However, the metric that we want to minimize is not the data error x_a but the reconstruction error

$$z_a = \frac{1}{|\mathcal{T}|}\|\boldsymbol{g}^{(\mathcal{T})} - \hat{\boldsymbol{g}}_a\|^2 \tag{44}$$

$$= \frac{1}{|\mathcal{T}|}\left(\left\|\boldsymbol{U}_a^{\mathrm{H}}\boldsymbol{z}^{(\mathcal{T})}\right\|^2 + \left\|\boldsymbol{V}_a^{\mathrm{H}}\boldsymbol{g}^{(\mathcal{T})}\right\|^2\right) \tag{45}$$

which cannot directly be observed at the receiver side. Knowing x_a, we are interested to obtain a probabilistic upper bound on z_a as follows:

$$z_a < \overline{z_a}(x_a, p_1, p_2) \tag{46}$$

that only depends on x_a and some constants p_1 and p_2. This bound enables us to select the best hypothesis $W_t(\hat{a})$, i.e.,

$$\hat{a} = \arg\min_a \overline{z_a}(x_a, p_1, p_2) \tag{47}$$

which minimizes the reconstruction error. The detailed algorithm to calculate the bound (47) is shown in the Appendix.

2) After we have obtained $\hat{a}$, we proceed in the same way to test the support of the PDP using the OFDM pilot index set $\mathcal{F}$. The subspace spanned by the fading process $\boldsymbol{g}^{(\mathcal{F})}$ can be approximated by the columns of the matrix $\boldsymbol{U}(W_t(\hat{a}) \times W_f(a'), I_{\mathcal{P}_t} \times I_f)$. We collect its first $D_{a'} = D(W_t(\hat{a}) \times W_f(a'), I_{\mathcal{P}_t} \times I_f)$ columns in matrix $\boldsymbol{U}_{a'}$ and the remaining columns in matrix $\boldsymbol{V}_{a'}$. The algorithms proceed now as in step 1), replacing $\mathcal{T}$ with $\mathcal{F}$ and a with a' such that we finally obtain

$$\hat{a}' = \arg\min_{a'} \overline{z_{a'}}(x_{a'}, p_1, p_2). \tag{48}$$

The described algorithm is motivated by the following: 1) the specific pilot pattern structure of IEEE 802.11p; 2) a numerical stable implementation; and 3) a good tradeoff between performance and complexity. Clearly, the algorithm provides an approximation, but we will demonstrate its good performance by numeric simulation results in Sections VI and

VI-D. Performing the hypothesis test iteratively is possible; however, we did not observe a substantial performance improvement by doing so.

C. Adaptive Channel Estimation

With the results of the hypothesis test $\hat{a}$ and $\hat{a}'$, we select the corresponding precalculcated matrices $\boldsymbol{U}(W_t(\hat{a}), I_t)$ and $\boldsymbol{U}(W_f(\hat{a}'), I_f)$. These matrices are inserted into (27) to obtain $\boldsymbol{U}$. After choosing the appropriate dimension D with (30), we are able to obtain the robust reduced-rank channel estimates using (32).

V. On the Mismatch of the Robust Wiener Filter

The robust Wiener filter (15) with the covariance matrix factorization (19) is by design mismatched to the actual delay-Doppler scattering function of the fading process. In this section, we want to establish bounds on this mismatch such that the designer of a channel estimation algorithm is able to assess if the robust Wiener filter is the right choice for a given operating range.

The MSE matrix of the matched Wiener filter [49, Sec. 15.8] is given by

$$\boldsymbol{R}_e = \boldsymbol{R}_g - \boldsymbol{R}_g \left(\boldsymbol{R}_g + \sigma_z^2 \boldsymbol{I}_M\right)^{-1} \boldsymbol{R}_g^{\mathrm{H}}. \tag{49}$$

The MSE matrix of the mismatched robust reduced-rank Wiener filter [22, Sec. 5.3] [50] can be expressed as

$$\begin{aligned}\tilde{\boldsymbol{R}}_e = \boldsymbol{R}_g &+ \tilde{\boldsymbol{R}}_g \left(\tilde{\boldsymbol{R}}_g + \sigma_z^2 \boldsymbol{I}_M\right)^{-1} \left(\boldsymbol{R}_g + \sigma_z^2 \boldsymbol{I}_M\right) \\ &\cdot \left(\tilde{\boldsymbol{R}}_g + \sigma_z^2 \boldsymbol{I}_M\right)^{-1} \tilde{\boldsymbol{R}}_g^{\mathrm{H}} - \tilde{\boldsymbol{R}}_g \left(\tilde{\boldsymbol{R}}_g + \sigma_z^2 \boldsymbol{I}_M\right)^{-1} \tilde{\boldsymbol{R}}_g^{\mathrm{H}} \\ &- \boldsymbol{R}_g \left(\tilde{\boldsymbol{R}}_g + \sigma_z^2 \boldsymbol{I}_M\right)^{-1} \tilde{\boldsymbol{R}}_g^{\mathrm{H}}. \end{aligned} \tag{50}$$

Finally, we define the relative error as

$$\Delta_e = \frac{1}{M} \mathrm{tr}(\tilde{\boldsymbol{R}}_e - \boldsymbol{R}_e). \tag{51}$$

A. Numerical Time-Domain Evaluation

In Fig. 6, we plot the MSE of the exact Wiener filter, the robust Wiener filter, and the reduced-rank Wiener filter versus ν_D, $0 < \nu_D \leq \nu_{\mathrm{Dmax}} = 0.0172$ for $M = 73$. These values correspond to a maximum relative velocity of 111.1 m/s $\approx$ 400 km/h and a typical IEEE 802.11p frame length (see Section VI).

We set $\tilde{\boldsymbol{R}}_g = \boldsymbol{R}(W_t, I_t)$. The fading process in this example has a Clarke's spectrum [51], with covariance function $[\boldsymbol{R}_g]_{k,\ell} = R_g[k-\ell] = J_0(2\pi\nu_D(k-\ell))$, while the robust Wiener filter assumes a flat spectrum with covariance function $\tilde{R}_g[k] = 1/(2\pi(\nu_{\mathrm{Dmax}}\overline{a}/A)k)\sin(2\pi(\nu_{\mathrm{Dmax}}\overline{a}/A)k)$, where

$$\overline{a} = \lfloor A\nu_D/\nu_{\mathrm{Dmax}} + 0.5 \rfloor. \tag{52}$$

The results in Fig. 6 indicate that the gradual abstraction from Clarke's spectrum using 1) a flat spectrum with quantized support and 2) a reduced-rank filter causes only a minor increase in the MSE. Hence, the shape of the DSD does not have a big impact on the MSE, but the support does (and a similar argument holds for the PDP in the frequency domain as well).

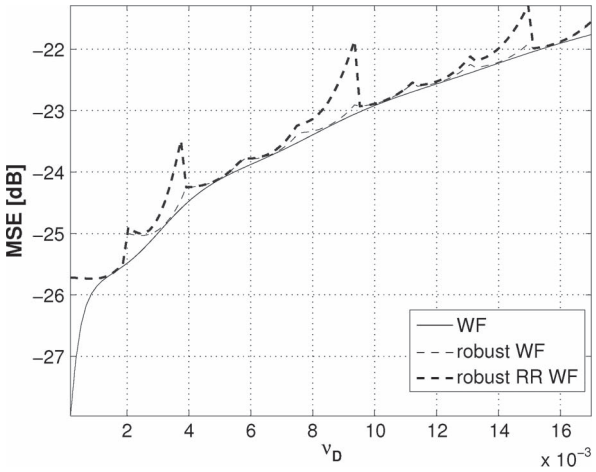

Fig. 6. MSE versus ν_D for $M = 73$, SNR $= 10$ dB. We plot the MSE for the exact Wiener filter (WF) matched to $R_g[k, \nu_D]$, the robust WF matched to $\tilde{R}_g[k, \nu_{\mathrm{Dmax}} a/A]$, and the robust reduced-rank (RR) WF.

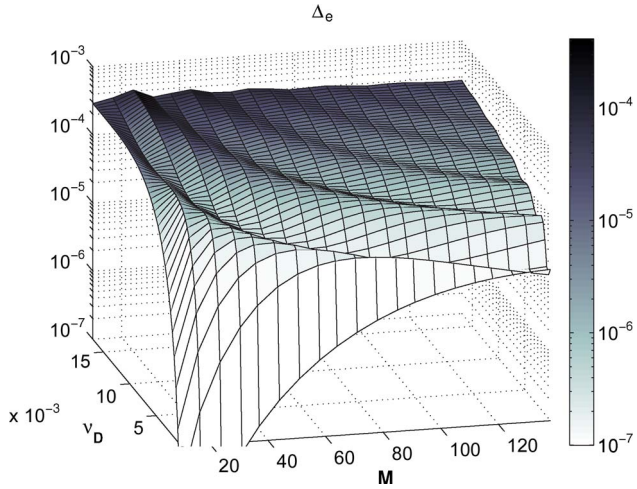

Fig. 7. Relative error Δ_e versus M and ν_D for SNR $= 10$ dB. Clarke's DSD.

In Fig. 7, we plot Δ_e versus ν_D and M. Comparing Fig. 7 with Fig. 6, one can see that for the whole operation range of IEEE 802.11p, the error increase Δ_e is more than one order of magnitude smaller than the actual filter error $(1/M)\mathrm{tr}(\boldsymbol{R}_e)$.

The frequency-domain results show a similar behavior and are omitted due to space limitations.

B. Rank One Region

For small frame length M and small Doppler bandwidth ν_D, the relative error Δ_e converges to zero exponentially, as shown in Fig. 7. Analyzing this behavior in more detail, we find that in this regime the essential subspace dimension $D_t = D(W_t, I_t) = 1$. From (30), we find that $D_t = 1$ if

$$\lambda_1(W, I) < |W|\sigma_z^2. \tag{53}$$

Using the eigenvalue asymptotic from [39, Sec. 2.5] for small $|W|$, we can express $\lambda_1(W, I)$ as

$$\lambda_1(W, I) = \frac{1}{\pi}(2\pi|W|)^3 \frac{1}{36}(|I|-1)(|I|)(|I|+1). \tag{54}$$

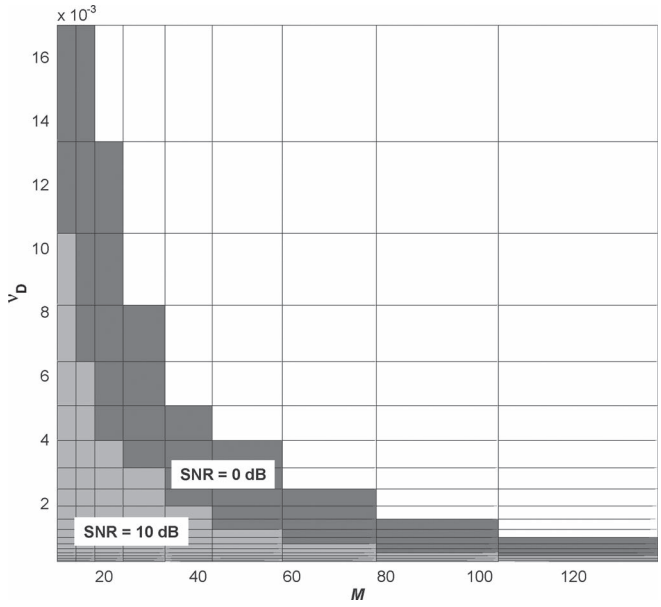

Fig. 8. Parameter range in terms of Doppler bandwidth ν_D and block length M, where the effective subspace dimension $D_t = 1$.

Inserting this expression into (53), we can fully characterize the operating region, defined by W and I, where the essential subspace dimension $D_t = 1$

$$\frac{8}{36}\left(\pi|W|\right)^2\left(|I|-1\right)\left(|I|\right)\left(|I|+1\right) < \sigma_z^2. \tag{55}$$

It is also known that the DPS sequences converge to the Legendre polynomial for small ν_D [52]. Hence, for the regime described by (55), the Wiener filter essentially assumes a constant channel. This explains also why the relative error vanishes. Please note that the rank one region does also depend on the noise variance σ_z^2. For the IEEE 802.11p system, we display the rank one region as the gray areas in Fig. 8 for SNR $\in \{0, 10\}$ dB.

VI. Numerical Performance Results for Subspace Selection

In this section, we present numerical performance results for subspace selection in the time and frequency domain. We use the IEEE 802.11p standard [2] for concrete parameterization of the OFDM communication system with bandwidth $B =$ 10 MHz, $N = 64$ subcarriers, $G = 16$ cyclic prefix length, and a carrier frequency of $f_C = 5.8$ GHz.

To simplify the notation in our derivation, we have assumed some minor deviations from the actual standard, which does not significantly affect the performance. Specifically, we assume that data are also transmitted on the dc subcarrier and that the guard bands are symmetrical with $N_g = 6$.

We furthermore note that the pilot symbol at the end of the packet is not (yet) part of the official standard. We believe that without such a backward-compatible pilot pattern modification, reliable communication in V2V scenarios in nonline-of-sight (NLOS) scenarios can be only obtained with large complexity at the receiver side. For the current IEEE 802.11p standard without postamble, the results presented in this paper can be seen as a lower bound.

We assume a maximum relative velocity between transmitter and receiver of $v_{\max} = 111.1$ m/s ≈ 400 km/h ≈ 249 mi/h. The data frame has length $M = 73$ OFDM symbols containing 400 B payload for coding and modulation scheme 3 at 6 Mb/s. This scheme uses QPSK (with symbol mapper rate $R_S = 2$) and a convolutional code with constraint length 7 and code rate $R_C = 1/2$. With an OFDM symbol duration of $T_S = (N + G)/B = 8$ μs, the maximum normalized Doppler frequency $\nu_{\mathrm{Dmax}} = v_{\max} f_C T_S / c_0 = 0.0172$, where c_0 is the speed of light. We assume a maximum path delay of $\tau_{\mathrm{Pmax}} = 1.6$ μs so that no intersymbol interference is caused, resulting in a maximum normalized excess delay of $\theta_{\mathrm{Pmax}} = 0.25$. This maximum path delay results in a strongly frequency-selective channel since $\tau_{\mathrm{Pmax}} \gg T_C = 1/B$.

A. Channel Model

We employ a four times oversampled Rayleigh fading channel model with an exponentially decaying PDP with normalized support $[0, \theta_P]$ and normalized root mean square (RMS) delay spread θ_P/β with $\beta = 5$ [53]. This choice is well in line with the results obtained from the analysis of extensive V2V channel measurements [54], [55].

At the receiver side, the angles of incidence of the impinging waves relative to the velocity vector are drawn from the interval $(0, 2\pi)$ according to Clarke's model [51], giving a normalized Doppler bandwidth ν_D. Note that ν_D is the instantaneously maximum Doppler frequency, i.e., the maximum Doppler frequency occurring in a particular stationarity region [56], associated with a particular maximum relative velocity. Due to the nonstationarity of the V2V channel, this ν_D can take on different values up to what we define as the "maximum" Doppler frequency ν_{Dmax}.

On the transmitter and receiver sides, a root-raised cosine pulse shaping filter with roll off factor of 0.2 is used. This model is similar to that described in [57], although using more realistic real valued path delays. The generated channel impulse responses are normalized to have average energy 1.

B. Time Domain

To validate the subspace selection algorithm in the time domain, we generate a time- and frequency-selective fading process. We set the PDP support to $\theta_P = \theta_{\mathrm{Pmax}}$ and vary the Doppler bandwidth in the range $0 \le \nu_D \le \nu_{\mathrm{Dmax}} = 0.017$. This relates to a velocity range $v \in (0, 111.1)$ m/s $\approx (0, 400)$ km/h. The used subcarrier pilot index set is defined as $\mathcal{T} = I_t \times I_{\mathcal{P}_f} = \{0, \ldots, 72\} \times \{8, 22, 44, 58\}$, representing four pilot subcarrier in an IEEE 802.11p frame. For each evaluated Doppler bandwidth ν_D, we transmitted 500 frames and use the subspace selection algorithm to perform the hypothesis test.

In Fig. 9, we plot the MSE at the pilot pattern $\mathcal{T}$ versus Doppler bandwidth ν_D at an SNR $\in \{0, 10\}$ dB. We depict the MSE $z_{\hat{a}}$ (44) for the selected hypothesis $\hat{a}$ obtained from (47) as solid line. Additionally, we show the MSE for the largest subspace z_A and the MSE $z_{\overline{a}}$ that is obtained for perfect knowledge of the Doppler bandwidth ν_D. Clearly, the subspace

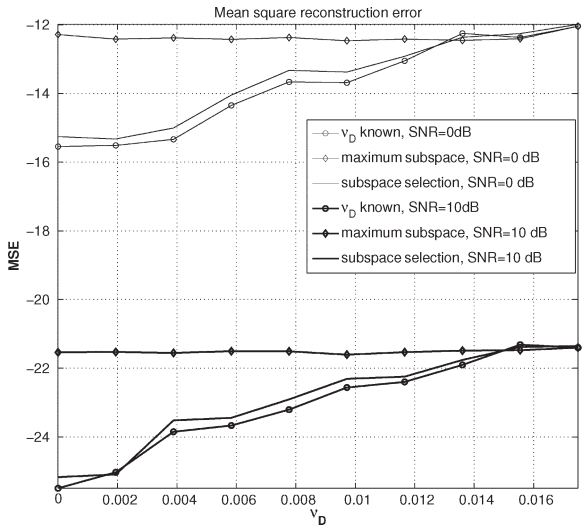

Fig. 9. MSE versus Doppler bandwidth ν_D at an SNR $\in \{0, 10\}$ dB using subspace selection in the time domain. The velocity ranges from $v \in (0, 400)$ km/h with $\nu_D \in (0, 0.017)$. Additionally, we plot the MSE for the largest subspace $W_t(A) = (-\nu_{\mathrm{Dmax}}, \nu_{\mathrm{Dmax}})$ and the MSE obtained with perfect knowledge of ν_D.

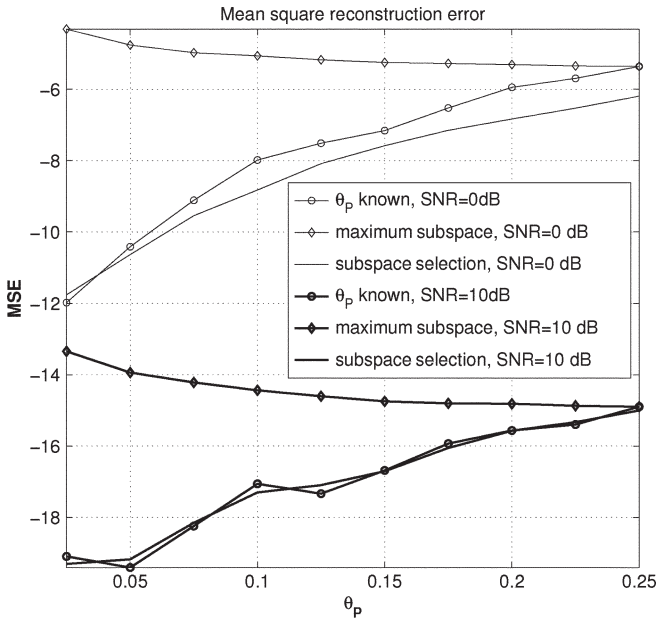

Fig. 10. MSE versus normalized PDP support θ_P for an SNR $\in \{0, 10\}$ dB. The maximum path delay ranges from $\theta_P \in (0.075, 0.25)$. Additionally, we plot the MSE for the largest subspace $W_f(A') = (-\theta_{\mathrm{Pmax}}, \theta_{\mathrm{Pmax}})$ and the MSE obtained with perfect knowledge of θ_P.

selection algorithm gains up to 4 dB at low velocities. The MSE with subspace selection closely follows the curve for perfectly known ν_D.

For comparison purposes, we implemented the Doppler bandwidth estimators from [44] and [45]. Both are not useable for short frame length and an SNR value of 10 dB. Their results coincide with the maximum subspace in Fig. 9 due to biased estimates.

C. Frequency Domain

To test the subspace selection algorithm in the frequency domain, we generate the same time- and frequency-selective channel as in Section VI-B. We set the DSD support to $\nu_D = \nu_{\mathrm{Dmax}}$ and vary the PDP support in the range $0 \leq \theta_P \leq \theta_{\mathrm{Pmax}} = 0.25$. This relates to a maximum excess delay range $\tau_P \in (0, 1.6)\,\mu$s. The used OFDM pilot index set is defined as $\mathcal{F} = I_{\mathcal{P}_t} \times I_f = \{0, 72\} \times \{6, \ldots, 58\}$, representing two pilot OFDM symbols at the beginning and end of the IEEE 802.11p frame. For each evaluated PDP support θ_P, we transmitted 500 frames and use the subspace selection algorithm to perform the hypothesis test.

In Fig. 10, we plot the MSE at the OFDM pilot symbol pattern $\mathcal{F}$ versus the support of the PDP θ_P for an SNR $\in \{0, 10\}$ dB. We depict the MSE with the subspace selection algorithm $z_{\hat{a}'}$ as a solid line. Additionally, we show the MSE for the largest subspace $z_{A'}$ and the MSE that is obtained for perfect knowledge of the support of the PDP $z_{\overline{a}'}$, where $\overline{a}' = \lceil A'\theta_P/\theta_{\mathrm{Pmax}} + 0.5 \rceil$. Clearly, the subspace selection algorithm gains up to 7–8 dB for small PDP support values. The MSE with subspace selection closely follows the curve for perfectly known θ_P for an SNR $= 10$ dB. For an SNR $= 0$ dB, the subspace selection results are even slightly better. This is possible because the PDP is exponentially decaying, and therefore, the filter is mismatched. The subspace selection can now improve the MSE by selecting a smaller subspace to reduce the mismatch.

The MSE for the largest subspace $z_{A'}$ slightly increases as θ_P becomes smaller. This is because the filter is designed for $[-\nu_{\mathrm{Dmax}}, \nu_{\mathrm{Dmax}}] \times [0, \theta_{\mathrm{Pmax}}]$, but the actual support of the channel scattering function in the delay domain shrinks to $[0, \theta_P]$ with $\theta_P < \theta_{\mathrm{Pmax}}$. The MSE increases since the mismatch between the Wiener filter and the actual fading process grows. Hence, a larger amount of noise affects the channel estimates and degrades the MSE.

D. Two-Dimensional Subspace Selection for Time-Variant Frequency-Selective Channels

Finally, we apply the subspace selection algorithm jointly to the time and the frequency domain such that we can adapt the robust Wiener filter for iterative channel estimation during an IEEE 802.11p transmission.

We show results for the following two channel models:

1) A Rayleigh/Rician fading channel model with exponentially decaying PDP, as described in Section VI-A for urban scenarios [57]. To speed up the calculation, we use a four times oversampled tap delay line implementation with Clarke's spectrum for each tap. The first tap can have a specified Rician K-factor to model a line-of-sight (LOS) scenario.
2) A geometry-based stochastic channel model (GSCM) [21] implementing a V2V link on a rural highway with a dominant LOS propagation path. The scenario depicted in [5, Fig. 18] is adapted to the V2V setting by placing the transmitter in a vehicle driving in opposite direction to the receiver. We use the GSCM parameterization from [30]. For every frame, a new random placement of all scattering objects is calculated. The distance of the two vehicles at the start of every frame transmission is fixed to x.

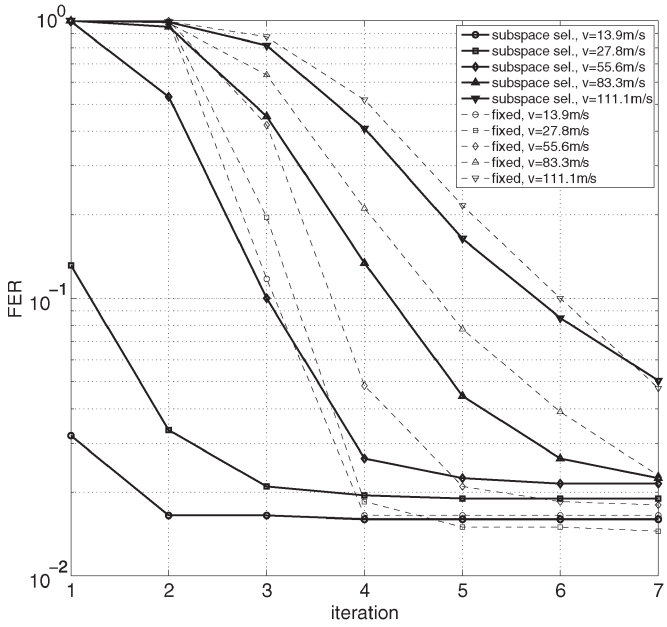

Fig. 11. FER versus number of iterations for a fixed $E_b/N_0 = 12$ dB for velocities $v \in \{13.9, 27.8, 55.6, 83.3, 111.1\}$ m/s $\approx \{50, 100, 200, 300, 400\}$ km/h and an RMS delay spread of 400 ns in an NLOS Rayleigh fading scenario. The frame length $M = 73$, containing 400 B of payload. The thick solid lines depict the results using the subspace selection algorithm. The dashed lines are the results using a fixed subspace with $W_t(A) = [-\nu_{\mathrm{Dmax}}, \nu_{\mathrm{Dmax}}]$ and $W_f(A') = [0, \theta_{\mathrm{Pmax}}]$.

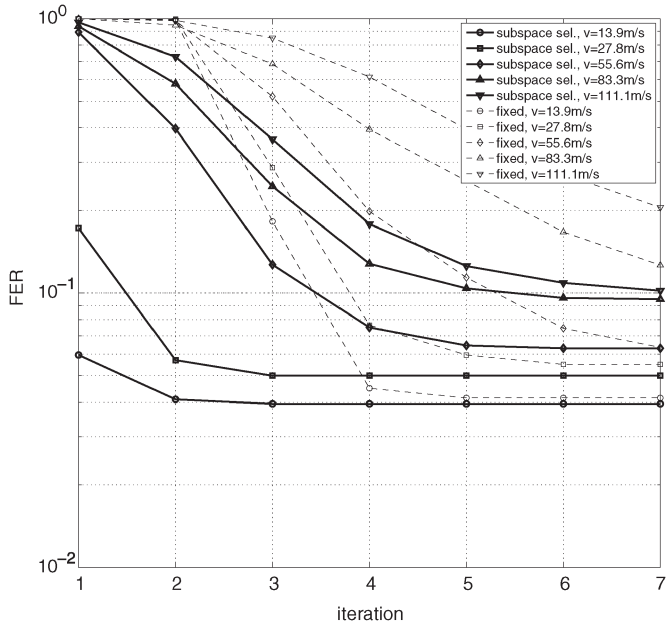

Fig. 12. FER versus number of iterations for a fixed $E_b/N_0 = 14$ dB for velocities $v \in \{13.9, 27.8, 55.6, 83.3, 111.1\}$ m/s $\approx \{50, 100, 200, 300, 400\}$ km/h and an RMS delay spread of 100 ns in an LOS scenario with a Rician K-factor of $K = 30$. The frame length $M = 73$, containing 400 B of payload. The thick solid lines depict the results using the subspace selection algorithm. The dashed lines are the results using a fixed subspace with $W_t(A) = [-\nu_{\mathrm{Dmax}}, \nu_{\mathrm{Dmax}}]$ and $W_f(A') = [0, \theta_{\mathrm{Pmax}}]$.

The FER results will be shown versus E_b/N_0, where E_b denotes the energy per bit, and N_0 denotes the noise power spectral density. Hence, we calculate the variance of the additive symmetric complex white Gaussian noise according to

$$\frac{1}{\sigma_z^2} = \frac{E_b}{N_0} R_{\mathrm{S}} R_C \frac{N}{N+G} \frac{MN-P}{MN}. \tag{56}$$

This corrects for the additional transmit energy used for the cyclic prefix and pilots.

We compare the results of the new adaptive reduced-rank Wiener filter to the results obtained with a fixed maximum subspace with $W_t(A) = [-\nu_{\mathrm{Dmax}}, \nu_{\mathrm{Dmax}}]$ and $W_f(A') = [0, \theta_{\mathrm{Pmax}}]$, which is the performance that can be obtained with the channel estimation algorithm from [30]. All numeric FER simulation results will be averaged over 2000 frames. We selected an E_b/N_0 value for each simulation such that a FER $= 10^{-1}$ could be reached at the highest relative velocity of $v = 111.1$ m/s ≈ 400 km/h.

In Fig. 11, we show the FER versus the number of iterations for a fixed $E_b/N_0 = 12$ dB for relative velocities $v \in \{13.9, 27.8, 55.6, 83.3, 111.1\}$ m/s $\approx \{50, 100, 200, 300, 400\}$ km/h and an RMS delay spread of 400 ns in an NLOS Rayleigh fading scenario. The frame length $M = 73$, containing 400 B of payload. The thick solid lines depict the results using the subspace selection algorithm. The dashed lines are the results using the fixed maximum subspace with $W_t(A) = [-\nu_{\mathrm{Dmax}}, \nu_{\mathrm{Dmax}}]$ and $W_f(A') = [0, \theta_{\mathrm{Pmax}}]$. Clearly, the subspace selection algorithm allows for faster convergence of the iterative receiver, requiring fewer iterations. For a velocity of {50, 100, 200, 300, 400} km/h, we need {1, 2, 3, 5, 6} iterations to reach an FER smaller than 10^{-1}. Without subspace selection, {4, 4, 4, 5, 7} iterations are needed. As expected, the number of iterations can be reduced most for slower velocities; in that case, the reduction is up to a factor of 4.

Due to the iterative channel estimation, the performance achieved with the largest subspace and the perfectly matched subspace is identical after more than seven iterations. In both cases, the performance converges to that with perfectly known CSI. In the regime with more than six iterations, we can also observe a slightly increased FER of the subspace selection algorithm if compared with the maximum subspace results. This is due to rare but possible wrong decisions in the hypothesis test toward a too small subspace.

In Fig. 12, we show results for a LOS scenario with a reduced RMS delay spread of 100 ns and a Rician K-factor of $K = 30$ for the first channel tap. Due to the reduced frequency diversity, an $E_b/N_0 = 14$ dB is required to reach an FER smaller than 10^{-1}. For a relative velocity of {50, 100, 200, 300, 400} km/h, we need {1, 2, 4, 5, 7} iterations to reach an FER smaller than 10^{-1}. Without subspace selection, {4, 4, 6, −, −} iterations are needed. For $v \geq 300$ km/h, substantially more than seven iterations are needed. Hence, the advantage of the subspace selection algorithm requiring less iterations is even more pronounced than in Fig. 11 due to the smaller support of the PDP that the subspace selection algorithm is able to exploit.

In Fig. 13, we show results for an urban LOS highway scenario [21], [58] modeled with a GSCM. The distance between both vehicles is $x = 50$ m at the beginning of each frame transmission. Due to the reduced multipath propagation in this scenario, we chose an $E_b/N_0 = 14$ dB. For a relative velocity of {50, 100, 200, 300, 400} km/h, we need {2, 2, 3, 3, 4} iterations to reach a FER smaller than 10^{-1}, whereas without subspace selection, {4, 4, 4, 4, 5} iterations are needed.

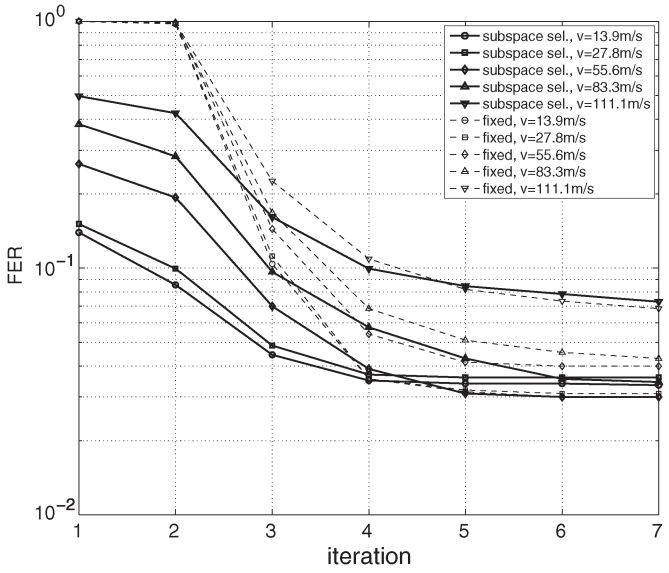

Fig. 13. FER versus number of iterations for a fixed $E_b/N_0 = 14$ dB for relative velocities $v \in \{13.9, 27.8, 55.6, 83.3, 111.1\}$ m/s $\approx \{50, 100, 200, 300, 400\}$ km/h using the GSCM from [21] with a distance between the two vehicles of $x = 50$ m. The frame length $M = 73$, containing 400 B of payload. The thick solid lines depict the results using the subspace selection algorithm. The dashed lines are the results using a fixed subspace with $W_t(A) = [-\nu_{\mathrm{Dmax}}, \nu_{\mathrm{Dmax}}]$ and $W_f(A') = [0, \theta_{\mathrm{Pmax}}]$.

VII. Numerical Complexity

The complexity of the iterative channel estimation algorithm presented in this paper is determined by four factors: 1) the reduced-rank Wiener filter; 2) the subspace selection algorithm; 3) the number of iterations I, and 4) the soft-input soft-output max-log MAP decoder.[4]

A. Reduced-Rank Wiener Filter

The robust reduced-rank Wiener filter (32) allows reducing the complexity of a normal Wiener filter exploiting only the dominant subspace. Hence, the complexity of (32) is mainly determined by the dimension of matrix $\boldsymbol{U}_D$ given by $D(\mathcal{W}, \mathcal{I})$ (30). We express the numerical complexity of the reduced-rank filter in terms of FLOPS [37], [60] as follows:

$$C_{\mathrm{RR}} \approx 8MND(\mathcal{W}, \mathcal{I})^2 + \frac{8}{3} D(\mathcal{W}, \mathcal{I})^3. \tag{57}$$

Clearly, the number of FLOPS now depends on the DSD support $|W_t| = 2\nu_D$ and the PDP support $|W_f| = \theta_P$ since $|\mathcal{W}| = |W_t||W_f|$.

For the maximum DSD and PDP support $\nu_D = \nu_{\mathrm{Dmax}} = 0.016$ and $\theta_P = \theta_{\mathrm{Pmax}} = 0.25$, we obtain the smallest complexity reduction. For a frame length of $M = 73$ containing 400-B payload and an $E_b/N_0 = 14$ dB, we obtain a subspace dimension of $D(\mathcal{W}, \mathcal{I}) = 75$, which gives a complexity reduction relative to the normal Wiener filter of $C_{\mathrm{W}}/C_{\mathrm{RR}} = 5.1 \cdot 10^3$ (see also [30]).

[4]For the max-log MAP decoder efficient, very large scale integration implementations are available [59] and are currently state of the art for turbo decoding in current long-term evolution receivers.

For small DSD and PDP support, the subspace dimension $D(\mathcal{W}, \mathcal{I}) = 1$ (see Section V-B). Hence, for the full DSD and PDP support range, the complexity reduction is bounded as

$$5.1 \cdot 10^3 \leq \frac{C_{\mathrm{W}}}{C_{\mathrm{RR}}} \leq 2.9 \cdot 10^7. \tag{58}$$

This complexity reduction of three to seven orders of magnitude is very desirable for a low-complexity and power-efficient chip set implementation, allowing the power consumption to be adapted to the requirements of the actual channel condition.

The robust reduced-rank Wiener filter must be calculated for all I iterations.

B. Subspace Selection Algorithm

The complexity of the subspace selection algorithm is primarily determined by (42), which must be evaluated for all $A + A'$ hypotheses per received frame, giving

$$C_{\mathrm{HT}} = \sum_{a=1}^{A} |\mathcal{T}| D_a^2 + \sum_{a'=1}^{A'} |\mathcal{F}| D_{a'}^2. \tag{59}$$

For the foregoing parameters with $A = A' = 10$, $D_a = 5$ and $D_{a'} = 15$, we obtain the complexity bound

$$4.2 \cdot 10^3 \leq C_{\mathrm{HT}} \leq 3.6 \cdot 10^5 \tag{60}$$

in terms of FLOPS, where the maximum value is required for the largest DSD and PDP support $\nu_D = \nu_{\mathrm{Dmax}} = 0.016$ and $\theta_P = \theta_{\mathrm{Pmax}} = 0.25$.

The term G_a in (69) and (70) can be precomputed by using an appropriate quantization of σ_z^2, and the bound (79) involves only elementary operations for the computation of the Gaussian approximation. These computations have negligible complexity compared with (59).

The subspace selection algorithm is only performed once per received frame; hence, its complexity is independent of the number of iterations I.

VIII. Conclusion

In this paper, we have presented a method to adapt a robust reduced-rank Wiener channel estimation filter based on the observation of a single IEEE 802.11p frame with short length for V2V communication scenarios. In a V2V drive-by situation, the DSD and PDP rapidly change due to the high relative velocity between transmitter and receiver, as well as changes in the scattering environment.

We utilize soft symbol feedback from a max-log MAP decoder to perform iterative channel estimation. The covariance matrix of the robust Wiener filter is designed according to a delay-Doppler scattering function prototype with flat spectrum and support in $[-\nu_D, \nu_D] \times [0, \theta_P]$. Due to the Cartesian product structure of the delay-Doppler scattering function, we are able to factorize the covariance matrix into two terms, depending on the support of the DSD and PDP. This factorization enables a numerically stable eigenvector decomposition of the covariance matrix utilizing generalized DPS sequences.

We define a set of hypotheses on the support of the DSD and a second set of hypotheses on the support of the PDP. Each hypothesis is represented by a specific subspace spanned by orthogonal basis vectors constructed from generalized DPS sequences. The adaptation algorithm chooses a hypothesis from both sets such that a probabilistic bound on the channel estimation error is minimized. We implement the hypothesis test by means of a novel subspace selection algorithm that allows utilizing correlated observations of a time- and frequency-selective (2-D) fading process.

We provide a performance analysis of the hypothesis test for a time-variant and frequency-selective fading process with a constrained frame length. We demonstrated a gain in MSE of 4 dB in the time domain and up to 8 dB in the frequency domain for small velocities and small delay spreads, respectively.

We validate the subspace selection scheme in an IEEE 802.11p compliant link level simulation for a relative velocity range from 0 to 111 m/s $\approx$ 400 km/h $\approx$ 248 mi/h. Up to fourfold reduction in number of iterations was demonstrated by numeric Monte-Carlo simulations to reach the same performance as with perfect CSI. An FER below 10^{-1} was reached for an $E_b/N_0 =$ 12 dB in an NLOS scenario and for an $E_b/N_0 = 14$ dB in an LOS scenario over the full relative velocity range.

APPENDIX

Here, we provide the derivation of the probabilistic upper bound $\overline{z_a}(x_a, p_1, p_2)$ used in (47). The derivation uses the results of [47] and [31] and adapts them to the notation of this paper.

1) We obtain expressions for the distribution of the reconstruction error z_a, i.e.,

$$\frac{2|\mathcal{T}|}{\sigma_z^2}\left(z_a - \frac{1}{P}\left\|\boldsymbol{V}_a^{\mathrm{H}}\boldsymbol{g}^{\mathcal{T}}\right\|^2\right) \sim \chi^2_{2D_a} \tag{61}$$

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

$$\mathbb{E}\{z_a\} = \frac{D_a}{P}\sigma_z^2 + \frac{1}{P}\left\|\boldsymbol{V}_a^{\mathrm{H}}\boldsymbol{g}^{\mathcal{T}}\right\|^2 \tag{62}$$

and variance

$$\mathrm{var}(z_a) = \frac{D_a}{|\mathcal{T}|^2}\left(\sigma_z^2\right)^2. \tag{63}$$

2) The data error x_a is a sample of a random variable that is distributed as

$$\frac{2P}{\sigma_z^2}x_a \sim \chi^2_{2(|\mathcal{T}|-D_a)}. \tag{64}$$

Therefore, x_a has expected value

$$\mathbb{E}\{x_a\} = \left(1 - \frac{D_a}{|\mathcal{T}|}\right)\sigma_z^2 + \frac{1}{|\mathcal{T}|}\left\|\boldsymbol{V}_a^{\mathrm{H}}\boldsymbol{g}^{\mathcal{T}}\right\|^2 \tag{65}$$

and variance

$$\mathrm{var}(x_a) = \frac{1}{|\mathcal{T}|}\left(1 - \frac{D_a}{|\mathcal{T}|}\right)\left(\sigma_z^2\right)^2 + \frac{2\sigma_z^2}{|\mathcal{T}|^2}\left\|\boldsymbol{V}_a^{\mathrm{H}}\boldsymbol{g}^{\mathcal{T}}\right\|^2. \tag{66}$$

3) According to [47, Sec. III.C], z_a is near its mean with probability p_1, i.e.,

$$\Pr\{|z_a - \mathbb{E}\{z_a\}| < G_a\} = p_1. \tag{67}$$

To proceed, we assume for now that $(1/|\mathcal{T}|)\|\boldsymbol{V}_a^{\mathrm{H}}\boldsymbol{g}^{\mathcal{T}}\|^2$ is known. In step 4, we will obtain a bound on this expression. The reconstruction error z_a is bounded with probability p_1 according to

$$\underline{z'_a}(p_1) \le z_a \le \overline{z'_a}(p_1) \tag{68}$$

where

$$\underline{z'_a}(p_1) = \frac{D_a}{|\mathcal{T}|}\sigma_z^2 + \frac{1}{|\mathcal{T}|}\left\|\boldsymbol{V}_a^{\mathrm{H}}\boldsymbol{g}^{\mathcal{T}}\right\|^2 - G_a(p_1, \sigma_z, 2D_a) \tag{69}$$

$$\overline{z'_a}(p_1) = \frac{D_a}{|\mathcal{T}|}\sigma_z^2 + \frac{1}{|\mathcal{T}|}\left\|\boldsymbol{V}_a^{\mathrm{H}}\boldsymbol{g}^{\mathcal{T}}\right\|^2 + G_a(p_1, \sigma_z, 2D_a). \tag{70}$$

The term $G_a(p_1, \sigma_z, 2D_a)$ is calculated by solving

$$p_1 = F\left(2D_a + 2G_a\frac{|\mathcal{T}|}{\sigma_z^2}, 2D_a\right) - F\left(2D_a - 2G_a\frac{|\mathcal{T}|}{\sigma_z^2}, 2D_a\right) \tag{71}$$

numerically for G_a. In (71), $F(x, i)$ denotes the Chi-square cumulative distribution function with i degrees of freedom. Note that G_a is independent of the actual channel realization $\boldsymbol{g}^{\mathcal{T}}$ and can be precomputed using an appropriate quantization of σ_z^2.

4) We utilize x_a and its known distribution to obtain a probabilistic bound on the square bias term $(1/|\mathcal{T}|)\|\boldsymbol{V}_a^{\mathrm{H}}\boldsymbol{g}^{\mathcal{T}}\|^2$ [47, Sec. VI.C]. For large $2(|\mathcal{T}| - D_a)$, we can invoke the central limit theorem to approximate x_a with a Gaussian random variable. The term

$$\frac{1}{|\mathcal{T}|}\left\|\boldsymbol{V}_a^{\mathrm{H}}\boldsymbol{g}^{\mathcal{T}}\right\|^2$$

is bounded with probability p_2 considered in the form

$$p_2 = \int_{-\alpha}^{\alpha}(1/\sqrt{2\pi})e^{-x^2/2}dx. \tag{72}$$

According to [47, Th. 1]

$$\underline{B_a}(x_a, p_2) \le \frac{1}{|\mathcal{T}|}\left\|\boldsymbol{V}_a^{\mathrm{H}}\boldsymbol{g}^{\mathcal{T}}\right\|^2 \le \overline{B_a}(x_a, p_2). \tag{73}$$

We, respectively, define the shorthand notation for the first term in (65) as

$$m_a = \left(1 - \frac{D_a}{|\mathcal{T}|}\right)\sigma_z^2 \tag{74}$$

and in (66) as

$$v_a = \frac{1}{|\mathcal{T}|}\left(1 - \frac{D_a}{|\mathcal{T}|}\right)\sigma_z^4. \tag{75}$$

The lower bound $\underline{B_a}$ is zero if

$$(m_a - \alpha\sqrt{v_a}) \leq x_a \leq (m_a + \alpha\sqrt{v_a}). \tag{76}$$

Otherwise, the lower bound is

$$\underline{B_a}(x_a, p_2) = x_a - m_a + \frac{\alpha^2\sigma_z^2}{|\mathcal{T}|} - K_a(\alpha) \tag{77}$$

where

$$K_a(\alpha) = 2\alpha\frac{\sigma_z}{\sqrt{2|\mathcal{T}|}}\sqrt{\frac{\alpha^2\sigma_z^2}{2|\mathcal{T}|} + x_a - \frac{1}{2}m_a}. \tag{78}$$

The upper bound is

$$\overline{B_a}(x_a, p_2) = x_a - m_a + \frac{\alpha^2\sigma_z^2}{|\mathcal{T}|} + K_a(\alpha). \tag{79}$$

5) Inserting (79) into (70) for the square bias term $(1/|\mathcal{T}|)\|\boldsymbol{V}_a^{\mathrm{H}}\boldsymbol{g}^{\mathcal{T}}\|^2$, we finally obtain the upper bound $\overline{z_a}(x_a, p_1, p_2) \geq \overline{z_a'}(p_1)$ as

$$\overline{z_a}(x_a, p_1, p_2) = \frac{D_a}{|\mathcal{T}|}\sigma_z^2 + \overline{B_a}(x_a, p_2) + G_a(p_1, \sigma_z, 2D_a). \tag{80}$$

ACKNOWLEDGMENT

The authors would like to thank F. Kaltenberger, L. Bernadó, C. Mecklenbräuker, and N. Czink for inspiring and helpful discussions and comments.

REFERENCES

[1] "Intelligent transport systems (ITS); Vehicular communications; basic set of applications; definitions," European Telecommunications Standards Institute, France, ETSI TR 102 638 V1.1.1, Jun. 2009.

[2] Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Amendment 6: Wireless Access in Vehicular Environments, IEEE P802.11p, Jul. 2010.

[3] A. F. Molisch, F. Tufvesson, J. Karedal, and C. F. Mecklenbräuker, "A survey on vehicle-to-vehicle propagation channels," *IEEE Wireless Commun. Mag.*, vol. 16, no. 6, pp. 12–22, Dec. 2009.

[4] O. Renaudin, V.-M. Kolmonen, P. Vainikainen, and C. Oestges, "Non-stationary narrowband MIMO inter-vehicle channel characterization in the 5-GHz band," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 2007–2015, May 2010.

[5] C. F. Mecklenbräuker, A. F. Molisch, J. Karedal, F. Tufvesson, A. Paier, L. Bernadó, T. Zemen, O. Klemp, and N. Czink, "Vehicular channel characterization and its implications for wireless system design and performance," *Proc. IEEE*, vol. 99, no. 7, pp. 1189–1212, Jul. 2011.

[6] G. Acosta, K. Tokuda, and M. A. Ingram, "Measured joint Doppler-delay power profiles for vehicle-to-vehicle communications at 2.4 GHz," in *Proc. IEEE GLOBECOM*, Dallas, TX, Dec. 2004, vol. 6, pp. 3813–3817.

[7] J. Maurer, T. Fügen, and W. Wiesbeck, "Narrow-band measurement and analysis of the inter-vehicle transmission channel at 5.2 GHz," in *Proc. 55th VTC*, Brimingham, AL, May 2002, pp. 1274–1278.

[8] L. Cheng, B. Henty, R. Cooper, D. Stancil, and F. Bai, "A measurement study of time-scaled 802.11 a waveforms over the mobile-to-mobile vehicular channel at 5.9 GHz," *IEEE Commun. Mag.*, vol. 46, no. 5, pp. 84–91, May 2008.

[9] A. Paier, R. Tresch, A. Alonso, D. Smely, P. Meckel, Y. Zhou, and N. Czink, "Average downstream performance of measured IEEE 802.11p infrastructure-to-vehicle links," in *Proc. IEEE ICC*, Cape Town, South Africa, May 23–27, 2010, pp. 1–5.

[10] J. Fernandez, D. Stancil, and F. Bai, "Dynamic channel equalization for IEEE 802.11p waveforms in the vehicle-to-vehicle channel," in *Proc. 48th Annu. Allerton Conf. Commun., Control, Comput.*, Oct. 1–29, 2010, pp. 542–551.

[11] J. Fernandez, K. Borries, L. Cheng, V. Bhagavatula, D. Stancil, and F. Bai, "Performance of the 802.11p physical layer in vehicle-to-vehicle environments," *IEEE Trans. Veh. Technol.*, vol. 61, no. 1, pp. 3–14, Jan. 2012.

[12] A. Bourdoux, H. Cappelle, and A. Dejonghe, "Channel tracking for fast time-varying channels in IEEE802.11p systems," in *Proc. GLOBECOM*, Houston, TX, Dec. 2011, pp. 1–6.

[13] M. Honig and J. Goldstein, "Adaptive reduced-rank interference suppression based on the multistage Wiener filter," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 986–994, Jun. 2002.

[14] H. Qian and S. Batalama, "Data record-based criteria for the selection of an auxiliary vector estimator of the MMSE/MVDR filter," *IEEE Trans. Commun.*, vol. 51, no. 10, pp. 1700–1708, Oct. 2003.

[15] D. Schafhuber, G. Matz, and F. Hlawatsch, "Adaptive Wiener filters for time-varying channel estimation in wireless OFDM systems," in *Proc. IEEE ICASSP*, Hong Kong, Apr. 2003, vol. 4, pp. IV-688–IV-691.

[16] R. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank processing based on joint and iterative interpolation, decimation, and filtering," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2503–2514, Jul. 2009.

[17] R. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank equalization algorithms based on alternating optimization design techniques for MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2482–2494, Jul. 2011.

[18] G. Tauböck, F. Hlawatsch, D. Eiwen, and H. Rauhut, "Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 255–271, Apr. 2010.

[19] C. Berger, S. Zhou, J. Preisig, and P. Willett, "Sparse channel estimation for multicarrier underwater acoustic communication: From subspace methods to compressed sensing," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1708–1721, Mar. 2010.

[20] D. Eiwen, G. Tauböck, F. Hlawatsch, and H. Feichtinger, "Group sparsity methods for compressive channel estimation in doubly dispersive multicarrier systems," in *Proc. IEEE 11th Int. Workshop SPAWC*, 2010, pp. 1–5.

[21] J. Karedal, F. Tufvesson, N. Czink, A. Paier, C. Dumard, T. Zemen, C. F. Mecklenbräuker, and A. F. Molisch, "A geometry-based stochastic MIMO model for vehicle-to-vehicle communications," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3646–3657, Jul. 2009.

[22] S. Kaiser, *Multi-Carrier CDMA Mobile Radio Systems—Analysis and Optimization of Detection, Decoding, and Channel Estimation*. Düsseldorf, Germany: VDI Verlag GmbH, 1998, ser. Fortschritts-Berichte VDI Reihe.

[23] S. Kaiser and P. Hoeher, "Performance of multi-carrier CDMA systems with channel estimation in two dimensions," in *Proc. 8th IEEE Int. Symp. PIMRC*, Helsinki, Finnland, Sep. 1997, vol. 1, pp. 115–119.

[24] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun. Syst.*, vol. CS-11, no. 4, pp. 360–393, Dec. 1963.

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

[26] 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.

[27] S. Y. Park, Y. G. Kim, and C. G. Kang, "Iterative receiver for joint detection and channel estimation in OFDM systems under mobile radio channels," *IEEE Trans. Veh. Technol.*, vol. 53, no. 2, pp. 450–460, Mar. 2004.

[28] 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.

[29] P. S. Rossi and R. R. Müller, "Slepian-based two-dimensional estimation of time–frequency variant MIMO-OFDM channels," *IEEE Signal Process. Lett.*, vol. 15, pp. 21–24, Jan. 2008.

[30] T. Zemen, L. Bernado, N. Czink, and A. F. Molisch, "Iterative time-variant channel estimation for 802.11p using generalized discrete prolate spheroidal sequences," *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 1222–1233, Mar. 2012.

[31] T. Zemen, C. F. Mecklenbräuker, B. H. Fleury, and F. Kaltenberger, "Minimum-energy band-limited predictor with dynamic subspace selection for time-variant flat-fading channels," *IEEE Trans. Signal Process.*, vol. 55, no. 9, pp. 4534–4548, Sep. 2007.

[32] T. Zemen, S. Caban, N. Czink, and M. Rupp, "Validation of minimum-energy band-limited prediction using vehicular channel measurements," in *Proc. 17th EUSIPCO*, Glasgow, U.K., Aug. 2009.

[33] L. Bernadó, N. Czink, T. Zemen, and P. Belanovic, "Physical layer simulation results for IEEE 802.11p using vehicular non-stationary channel model," in *Proc. IEEE ICC*, Cape Town, South Africa, May 2010, pp. 1–5.
[34] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal map decoding algorithms operating in the log domain," in *Proc. IEEE ICC Seattle—Gateway to Globalization*, Jun. 1995, vol. 2, pp. 1009–1013.
[35] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 2, pp. 284–287, Mar. 1974.
[36] 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.
[37] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.
[38] F. Kaltenberger, T. Zemen, and C. Ueberhuber, "Low-complexity geometry-based MIMO channel simulation," *EURASIP J. Adv. Signal Process.*, vol. 2007, pp. 95281-1–95281-17, 2007.
[39] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty—V: The discrete case," *Bell Syst. Tech. J.*, vol. 57, no. 5, pp. 1371–1430, May/Jun. 1978.
[40] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications*. Cambridge, U.K.: Cambridge Univ. Press, 1963.
[41] D. S. Tracy and R. P. Singh, "A new matrix product and its applications in matrix differentiation," *Statistica Neerlandica*, vol. 26, no. 4, pp. 143–157, Dec. 1972.
[42] Shuangzhe and Liu, "Matrix results on the Khatri–Rao and Tracy–Singh products," *Linear Algebra Appl.*, vol. 289, no. 1–3, pp. 267–277, Mar. 1999.
[43] 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.
[44] 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.
[45] 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.
[46] H. Zhang and A. Abdi, "Nonparametric mobile speed estimation in fading channels: Performance analysis and experimental results," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1683–1692, Apr. 2009.
[47] 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.
[48] A. Papoulis, *Probability, Random Variables and Stochastic Processes*. Singapore: McGraw-Hill, 1991.
[49] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice-Hall, 1993.
[50] A. Ispas, L. Bernadó, M. Dörpinghaus, G. Ascheid, and T. Zemen, "On the use of mismatched Wiener filtering for the characterization of non-stationary channels," in *Proc. 44th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Nov. 7–10, 2010, pp. 1971–1975.
[51] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell Syst. Tech. J.*, vol. 47, no. 6, pp. 957–1000, Jul./Aug. 1968.
[52] F. Hlawatsch and G. Matz, *Wireless Communications over Rapidly Time-Varying Channels*. New York: Academic, 2011.
[53] A. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Sel. Areas Commun.*, vol. SAC-5, no. 2, pp. 128–137, Feb. 1987.
[54] L. Bernadó, A. Roma, T. Zemen, N. Czink, J. Karedal, A. Paier, A. Thiel, O. Klemp, F. Tufvesson, A. F. Molisch, and C. F. Mecklenbräuker, "In-tunnel vehicular radio channel characterization," in *Proc. IEEE 73rd Veh. Technol. Conf.*, Budapest, Hungary, May 2011, pp. 1–5.
[55] L. Bernadó, "Non-stationarity in vehicular wireless channels," Ph.D. dissertation, Vienna Univ. Technol., Vienna, Austria, Apr. 2012.
[56] A. Paier, T. Zemen, L. Bernadó, G. Matz, J. Karedal, N. Czink, C. Dumard, F. Tufvesson, A. F. Molisch, and C. F. Mecklenbräuker, "Non-WSSUS vehicular channel characterization in highway and urban scenarios at 5.2 GHz using the local scattering function," in *Proc. WSA*, Darmstadt, Germany, Feb. 2008, pp. 9–15.
[57] G. Acosta-Marum and M. A. Ingram, "Six time- and frequency-selective empirical channel models for vehicular wireless LANs," *IEEE Veh. Technol. Mag.*, vol. 2, no. 4, pp. 4–11, Dec. 2007.
[58] J. Karedal, "Measurement-based modeling of wireless propagation channels—MIMO and UWB," Ph.D. dissertation, Lund Univ., Lund, Sweden, 2009.
[59] L. Sabeti, M. Ahmadi, and K. Tepe, "Low-complexity BCJR decoder for turbo decoders and its VLSI mplementation in 0.18-μm CMOS," in *Proc. IEEE 62nd VTC*, Sep. 2005, vol. 2, pp. 912–916.
[60] C. Dumard and T. Zemen, "Low-complexity MIMO multiuser receiver: A joint antenna detection scheme for time-varying channels," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2931–2940, Jul. 2008.

Thomas Zemen (S'03–M'05–SM'10) received the Dipl.-Ing. and doctoral degrees (both with distinction) in electrical engineering from Vienna University of Technology, Vienna, Austria, in 1998 and 2004, respectively.

From 1998 to 2003, he was a Hardware Engineer and Project Manager with the Radio Communication Devices Department, Siemens Austria. Since October 2003, he has been with FTW Forschungszentrum Telekommunikation Wien, Vienna, where he has been leading the "Signal and Information Processing" Department since 2008. He is the speaker of the national research network for "Signal and Information Processing in Science and Engineering" funded by the Austrian Science Fund. He is the author or coauthor of four book chapters, 17 journal papers, and more than 65 conference communications. His research interests include vehicular channel measurements and modeling, time-variant channel estimation, orthogonal frequency division multiplexing, iterative multiple-input–multiple-output receiver structures, cooperative communication systems, and interference management.

Dr. Zemen is an External Lecturer with the Vienna University of Technology and serves as Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.

Andreas F. Molisch (S'89–M'95–SM'00–F'05) received the Dipl. Ing., Ph.D., and habilitation degrees from the Technical University of Vienna, Vienna, Austria, in 1990, 1994, and 1999, respectively.

He is a Professor of electrical engineering with the University of Southern California, Los Angeles. was previously with the AT&T (Bell) Laboratories Research, Lund University, Lund, Sweden, Mitsubishi Electric Research Labs, and Vienna University of Technology, Vienna, Austria. His current research interests are measurement and modeling of mobile radio channels, ultra-wideband communications and localization, cooperative communications, multiple-input–multiple-output systems, and wireless systems for healthcare. He has authored, coauthored, or edited four books (among them the textbook *Wireless Communications*: New York: Wiley-IEEE Press), 14 book chapters, more than 140 journal papers, and numerous conference contributions, as well as more than 70 patents and 60 standards contributions.

Dr. Molisch has been an editor of a number of journals and special issues, General Chair, Technical Program Committee Chair, or Symposium Chair of multiple international conferences, as well as Chairman of various international standardization groups. He is a Fellow of the Institution of Engineering and Technology, an IEEE Distinguished Lecturer, and a member of the Austrian Academy of Sciences. He has received numerous awards, most recently the Donald Fink Prize and the Eric Sumner Award, both from the IEEE.