Characterization of Non-Stationary Channels Using Mismatched Wiener Filtering

Adrian Ispas, *Student Member, IEEE*, Meik Dörpinghaus, *Member, IEEE*, Gerd Ascheid, *Senior Member, IEEE*, and Thomas Zemen, *Senior Member, IEEE*

***Abstract*—A common simplification in the statistical treatment of linear time-varying (LTV) wireless channels is the approximation of the channel as a stationary random process inside certain time-frequency regions. We develop a methodology for the determination of local quasi-stationarity (LQS) regions, i.e., local regions in which a channel can be treated as stationary. Contrary to previous results relying on, to some extent, heuristic measures and thresholds, we consider a finite-length Wiener filter as realistic channel estimator and relate the size of LQS regions in time to the degradation of the mean square error (MSE) of the estimate due to outdated and thus mismatched channel statistics. We show that for certain power spectral densities (PSDs) of the channel a simplified but approximate evaluation of the matched MSE based on the assumption of an infinite filtering length yields a lower bound on the actual matched MSE. Moreover, for such PSDs, the actual MSE degradation is upper-bounded and the size of the actual LQS regions is lower-bounded by the approximate evaluation. Using channel measurements, we compare the evolution of the LQS regions based on the actual and the approximate MSE; they show strong similarities.**

I. Introduction

An important simplification in the statistical modeling of linear time-varying (LTV) wireless channels is the assumption of stationarity of the fading process in time and frequency. A restriction to first- and second-order stationarity results in wide-sense stationary (WSS) and uncorrelated scattering (US) channels, i.e., WSSUS channels [1]. On the one hand, this assumption is used due to the resulting mathematical simplifications in the treatment of such channels, and, on the other hand, it has some reasonable physical justification. The physical justification for the WSSUS assumption is that large-scale effects such as shadow fading change the statistics of the channel only slowly over time and frequency in comparison to the coherence time and the coherence frequency, respectively. In [1], this leads to the quasi-WSSUS model where the channel is divided into WSSUS regions. In [2], a framework for the treatment of non-stationary channels, which satisfy the doubly underspread condition, is presented. Doubly underspread channels are channels that are dispersion- and correlation-underspread. The framework in [2] can be considered as a natural extension of the quasi-WSSUS model and the work in [1].

We define local quasi-stationarity (LQS) regions as regions inside which the channel can be treated as a stationary process. From the analysis of wireless channels measurements, it is known that the assumption of LQS regions is a reasonable approach to perform, e.g., channel estimation [3]–[5]. An important open problem is the determination of the size of these LQS regions. Note that elaborate tests for stationarity of a random process exist, see [6]–[10] and references therein. However, we seek for a method characterizing the degree of non-stationarity [11] of doubly underspread random processes since the wireless channel is inherently non-stationary. Several such methods using various measures are available, see, e.g., [3]–[5], [12], [13] as well as [14] for an overview. The usual approach of selecting, to some extent, heuristic measures and to compare them to suitably chosen thresholds is far from being satisfactory [10]. We overcome this problem by relating the non-stationarity characterization of the channel to an algorithmic view. Our approach is connected to [15] and [16] in the sense that they also consider the effects of mismatched statistics in estimation theory. However, they focus on minimum mean square error (MMSE) estimation, whereas we restrict to a linear MMSE (LMMSE) estimator which corresponds to a typical channel estimator.

Contributions of the Paper: We describe a methodology for the determination of LQS regions of the channel fading process, a non-stationary random process. The method is based on the performance degradation of a realistic channel estimation algorithm due to mismatched channel statistics. Our concept is similar in nature to the comparison of power spectral densities (PSDs) of stationary random processes presented for the case of noiseless observations in [17]. We base our approach on a Wiener, i.e., an LMMSE, filter using noisy observations to estimate a non-stationary LTV frequency-flat fading channel. Our contributions are as follows.

- We apply mismatched Wiener filtering to non-stationary doubly underspread channels with an effectively finite

This work was supported by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMmunications NEWCOM++. Additionally, the work of Adrian Ispas, Meik Dörpinghaus, and Gerd Ascheid was supported by the Ultra high-speed Mobile Information and Communication (UMIC) research centre. The work of M. Dörpinghaus was partially supported by the DFG under Grant DO1568/1-1. The work of Thomas Zemen was supported by the project COCOMINT funded by the Vienna Science and Technology Fund (WWTF) as well as NFN SISE (S10607) funded by the Austrian Science Fund (FWF). Moreover, the FTW is supported by the Austrian Government and the City of Vienna within the competence center program COMET. Parts of this work were presented at the 44th Annual Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, November 2010.

Adrian Ispas and Gerd Ascheid are with the Chair for Integrated Signal Processing Systems, RWTH Aachen University, Germany; email: {ispas, ascheid}@iss.rwth-aachen.de.

Meik Dörpinghaus is with the Institute for Theoretical Information Technology, RWTH Aachen University, Germany; email: doerpinghaus@ti.rwth-aachen.de.

Thomas Zemen is with FTW Telecommunications Research Center Vienna, Austria; email: thomas.zemen@ftw.at.

correlation, i.e., an autocorrelation function that can be approximated by zero for time differences outside a finite interval. We obtain the mismatched mean square error (MSE) of the channel estimate based on finite-length (FL) filtering. Then we define the MSE degradation as a measure to assess the performance degradation in comparison to the matched case. This, in turn, yields LQS regions that are related to the update interval of the channel statistics. The update is required to operate the channel estimator below a certain MSE degradation.
- As the MSE evaluation requires a large computational effort, we also give a simplified but approximate expression based on infinite-length (IL) Wiener filtering. We show that for certain rectangular PSDs of the channel fading process the approximate matched MSE based on infinite-length filtering is a lower bound to the actual matched MSE based on finite-length filtering. We then show that for this type of PSDs an upper bound on the MSE degradation and a lower bound on the size of the LQS regions, respectively, is obtained by the approximate evaluation.
- We provide an exemplary analysis of the introduced concepts using a rectangular PSD of the channel fading process, on the one hand, and for actual channel measurements, on the other hand. We observe that the actual and the approximate evaluation of the MSE degradation show a similar behavior regarding the size of the LQS regions. Moreover, we compare the measurement-based LQS regions to those using the widely used correlation matrix distance (CMD) [14].

Organization: After introducing the system model in Section II, we recall and investigate mismatched finite- and infinite-length Wiener filtering for WSS channels in Section III. In Section IV, the use of mismatched Wiener filtering is extended to non-stationary wireless channels that are doubly underspread and that have an effectively finite correlation. These results yield and relate the actual and the approximate MSE degradation which are used to evaluate the corresponding LQS regions of the channel. Finally, in Section V, we provide an exemplary analysis of the introduced concepts using a rectangular PSD of the channel fading process as well as for actual channel measurements.

Notation: We use lowercase and uppercase boldface letters to designate vectors and matrices, respectively. $\mathbf{A}^*$, $\mathbf{A}^T$, and $\mathbf{A}^H$ are the (element-wise) complex conjugate, the transpose, and the conjugate transpose of the matrix $\mathbf{A}$, respectively. $\mathrm{tr}\{\mathbf{A}\}$ denotes the trace of the matrix $\mathbf{A}$, and $\|\mathbf{A}\|_F$ is the Frobenius norm of $\mathbf{A}$. For two matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A} \odot \mathbf{B}$ is the Hadamard (element-wise) product and $\mathbf{A} \otimes \mathbf{B}$ denotes the Kronecker product. $\mathbf{I}_N$ is the $N \times N$ identity matrix. $[\mathbf{A}]_{k,l}$ is the element of the matrix $\mathbf{A}$ in the kth row and lth column, and $[\mathbf{a}]_k$ is the kth element of the vector $\mathbf{a}$. $|\mathcal{A}|$ denotes the cardinality of the set $\mathcal{A}$. The circular convolution $x(\nu) \circledast y(\nu), -B/2 \leq \nu < B/2$ is defined as $\int_{-B/2}^{B/2} x(\check{\nu}) \sum_{k=-\infty}^{\infty} \check{y}(\nu - \check{\nu} - kB) d\check{\nu}$ with $\check{y}(\nu) = y(\nu)$ for $-B/2 \leq \nu < B/2$ and $\check{y}(\nu) = 0$ else. The Kronecker delta is represented by $\delta[m]$ for an integer m. $\hat{x}$ denotes the LMMSE estimate of x, and $\tilde{x}$ indicates a mismatched variant of x or $\hat{x}$. $\lfloor x \rfloor$ and $\lceil x \rceil$ denote rounding x down and up to the next integer, respectively. Sums and integrals are taken from $-\infty$ to ∞, unless otherwise specified. The imaginary unit is denoted by j.

II. System Model

We assume transmission over a non-stationary LTV frequency-flat fading channel. In the complex baseband, the matched-filtered, symbol-sampled received signal is modeled by

$$y[m] = h[m]x[m] + n[m], \ m \in \mathbb{Z} \tag{1}$$

where the additive noise process $\{n[m]\}$ is a zero-mean white proper [18] complex Gaussian random process with known variance $\sigma_n^2 > 0$ and the time-varying channel process $\{h[m]\}$ is a proper complex zero-mean random process. The transmitted sequence $\{x[m]\}$ consists of random data symbols and periodically inserted deterministic pilot, i.e., training, symbols with period L at positions $m = kL$ for $k \in \mathbb{Z}$. The received signal at the pilot positions will be used to estimate the channel process $\{h[m]\}$. The processes $\{h[m]\}$, $\{n[m]\}$, and $\{x[m]\}$ are assumed to be mutually independent. We assume real- and positive-valued pilot symbols with magnitude $\sigma_p > 0$. This causes no loss of generality, as the MSE does not depend on the phase of the pilot symbols. Furthermore, we define the ratio $\gamma = \sigma_p^2/\sigma_n^2$.[1]

We introduce the (normalized) Doppler shift $\nu = \nu' T$ with the unnormalized Doppler shift ν' and the symbol duration T in order to simplify notation. In the remainder of this work, we assume that channel sampling by pilot symbols satisfies the Nyquist criterion

$$L < \frac{1}{2\nu_{\max}} \tag{2}$$

where L is the pilot spacing, i.e., every Lth symbol is a pilot symbol, and $\nu_{\max} = \nu'_{\max} T$ is the maximal Doppler shift. Here, $\nu_{\max}$ is defined by the support of the autocorrelation function $R_S(\nu_1, \nu_2) = \mathrm{E}\{S_h(\nu_1) S_h^*(\nu_2)\}$ in ν_1 and ν_2 for $-1/2 \leq \nu_1, \nu_2 < 1/2$ with $S_h(\nu) = \sum_m h[m] e^{-j2\pi m\nu}$.[2] The assumption of a maximal Doppler shift is reasonable as any movement, be it of the transmitter, the receiver, or the scatterers in the environment, occurs with a finite velocity. Therefore, the channel process $\{h[m]\}$ is a bandlimited random process[3] and a sufficient statistic of $\{h[m]\}$ is obtained by regularly sampling $\{h[m]\}$ with period L.

A. Doubly Underspread Channels

In [2], the class of doubly underspread channels is introduced for time- and frequency-varying channels. The doubly

[1] Strictly speaking, (1) is not a sufficient statistic of the received continuous-time signal, as the channel widens the bandwidth of the (noise-free) received signal. However, (1) holds approximately for *Doppler* dispersion-underspread channels that are described in Section II-A. For a thorough discussion on the discretization of *delay-Doppler* dispersion-underspread WSSUS time- and frequency-selective LTV channels, see [19].

[2] In [20], a simple derivation of the sampling theorem for non-stationary random processes is given.

[3] In this context, the term *bandlimited* refers to the support of $R_S(\nu_1, \nu_2)$ in ν_1 and ν_2 for $-1/2 \leq \nu_1, \nu_2 < 1/2$.

underspread condition is usually satisfied for wireless channels [21], see [3] for an example in an urban macrocell scenario. We adapt the definition of doubly underspread channels in [2] to the special case of only time-varying (frequency-flat fading) channels as considered here. Such doubly underspread channels are, on the one hand, *dispersion*-underspread with a maximal Doppler $\nu_{\max} \ll 1$, and, on the other hand, they are *correlation*-underspread with a maximal (effective) correlation in Doppler $\Delta_{\nu,\max} \ll \nu_{\max}$. Here, $\Delta_{\nu,\max}$ is defined by the effective support[4] of the autocorrelation function $R_S(\nu+\Delta_\nu,\nu)$ in the Doppler difference Δ_ν for $-1/2 \leq \nu+\Delta_\nu, \nu < 1/2$. The doubly underspread condition essentially means that the stationarity time of the channel defined as $N_\mathrm{s} = 1/\Delta_{\nu,\max}$ is much larger than the coherence time of the channel $N_\mathrm{c} = 1/\nu_{\max}$, which itself is much larger than one, i.e., $1 \ll N_\mathrm{c} \ll N_\mathrm{s}$. This definition of the doubly underspread condition is thus connected to the symbol duration, which is a system parameter. Note that LQS regions can be of significantly larger size than the above mentioned (local) stationarity regions.

III. Mismatched Wiener Filtering

We know from Section II-A that wireless channels can be assumed as WSS inside certain regions. Therefore, we first recall the concept of *mismatched* Wiener filtering for WSS channels in the finite- and infinite-length setting. In this work, mismatched Wiener filtering refers to the use of wrong statistical knowledge of the channel, but not of the noise. The mismatched statistics of the channel are also assumed to correspond to a bandlimited WSS random process which is sampled by the pilot symbols such that (2) is satisfied. The Wiener filter considered here uses the noisy observations at the pilot positions for the estimation of the channel process. In the matched case, the Wiener filter is an LMMSE estimator and thus minimizes the MSE among all linear estimators. In this section, we restrict to WSS channels to allow for a comprehensive introduction to mismatched Wiener filtering. After recalling results on mismatched Wiener filtering, we give an alternative derivation of the MSE for infinite-length filtering. Then, we study under which circumstances an increasing noise power can lead to a decreasing MSE. Finally, we show how to quantify the loss in MSE due to the linearity of the estimator.

A. Finite-Length Filtering

We first consider the finite-length filtering case with N_p pilot symbols, i.e., training symbols used to estimate the channel, and the pilot spacing L. Denoting the block length as N, we have $N_d = N - N_p$ data symbols. Without loss of generality, we assume that the block starts with the pilot symbol at $m = 0$. In this section, we assume that the channel weights $h[m]$ form a WSS random process. The length-N column vector $\mathbf{h} \in \mathbb{C}^{N\times 1}$ contains the channel weights $h[m]$ with $[\mathbf{h}]_{m+1} = h[m], m = 0,\ldots,N-1$. Similarly, $\mathbf{y} \in \mathbb{C}^{N\times 1}$ is the vector of noisy observations given by $[\mathbf{y}]_{m+1} = y[m], m = 0,\ldots,N-1$. The vectors $\mathbf{y}_p \in \mathbb{C}^{N_p\times 1}$, $\mathbf{h}_p \in \mathbb{C}^{N_p\times 1}$, and $\mathbf{n}_p \in \mathbb{C}^{N_p\times 1}$ contain the noisy observations, the channel, and the noise, respectively, at pilot positions only. They are given by $[\mathbf{y}_p]_{k+1} = y[kL]$, $[\mathbf{h}_p]_{k+1} = h[kL]$, and $[\mathbf{n}_p]_{k+1} = n[kL]$ for $k = 0,\ldots,N_p-1$, respectively. We thus obtain

$$\mathbf{y}_p = \mathbf{h}_p\sigma_p + \mathbf{n}_p \tag{3}$$

where σ_p is the absolute value of the pilot symbols. Based on the orthogonality principle for LMMSE estimation [22][5] $\mathrm{E}\{(\hat{\mathbf{h}}-\mathbf{h})\mathbf{y}_p^H\} = \mathbf{0}$ with the linear estimate of the zero-mean channel $\mathbf{h}$ denoted by $\hat{\mathbf{h}} = \mathbf{W}\mathbf{y}_p$, we obtain the Wiener filter coefficients $\mathbf{W} = \mathbf{R}_{h;y_p}\mathbf{R}_{y_p}^{-1}$ where $\mathbf{R}_{h;y_p} = \mathrm{E}\left\{\mathbf{h}\mathbf{y}_p^H\right\}$ is the cross-correlation matrix between the channel at all positions and the observations at pilot positions, and $\mathbf{R}_{y_p} = \mathrm{E}\left\{\mathbf{y}_p\mathbf{y}_p^H\right\}$ is the Hermitian autocorrelation matrix of the observations at pilot positions. Denoting the corresponding mismatched correlation matrices used for the mismatched estimate $\tilde{\mathbf{h}} = \tilde{\mathbf{W}}\mathbf{y}_p$ as $\tilde{\mathbf{R}}_{h;y_p}$ and $\tilde{\mathbf{R}}_{y_p}$, respectively, the mismatched Wiener filter coefficients are obtained as

$$\tilde{\mathbf{W}} = \tilde{\mathbf{R}}_{h;y_p}\tilde{\mathbf{R}}_{y_p}^{-1}. \tag{4}$$

We estimate the channel at all positions inside the considered block, i.e., at pilot and data positions. With the autocorrelation matrix of the channel $\mathbf{R}_h = \mathrm{E}\left\{\mathbf{h}\mathbf{h}^H\right\}$ and using the fact that the noise is white and independent of the channel, we obtain the mismatched MSE matrix [24, Section 5.3]

$$\begin{aligned}\tilde{\mathbf{R}}_e &= \mathrm{E}\left\{\left(\tilde{\mathbf{h}}-\mathbf{h}\right)\left(\tilde{\mathbf{h}}-\mathbf{h}\right)^H\right\}\\ &= \mathbf{R}_h + \tilde{\mathbf{R}}_{h;h_p}\left(\tilde{\mathbf{R}}_{h_p}+\mathbf{I}_{N_p}\gamma^{-1}\right)^{-1}\left(\mathbf{R}_{h_p}+\mathbf{I}_{N_p}\gamma^{-1}\right)\\ &\times\left(\tilde{\mathbf{R}}_{h_p}+\mathbf{I}_{N_p}\gamma^{-1}\right)^{-1}\tilde{\mathbf{R}}_{h;h_p}^H - \tilde{\mathbf{R}}_{h;h_p}\left(\tilde{\mathbf{R}}_{h_p}+\mathbf{I}_{N_p}\gamma^{-1}\right)^{-1}\mathbf{R}_{h;h_p}^H\\ &- \mathbf{R}_{h;h_p}\left(\tilde{\mathbf{R}}_{h_p}+\mathbf{I}_{N_p}\gamma^{-1}\right)^{-1}\tilde{\mathbf{R}}_{h;h_p}^H\end{aligned} \tag{5}$$

where $\mathbf{R}_{h_p} = \mathrm{E}\left\{\mathbf{h}_p\mathbf{h}_p^H\right\}$ is the autocorrelation matrix of the channel at pilot positions, and $\mathbf{R}_{h;h_p} = \mathrm{E}\left\{\mathbf{h}\mathbf{h}_p^H\right\}$ is the cross-correlation matrix between the channel at all positions and the channel at pilot positions only. $\tilde{\mathbf{R}}_h$ and $\tilde{\mathbf{R}}_{h;h_p}$ are the corresponding mismatched correlation matrices. For matched a-priori information about the channel, (5) simplifies to the well-known result [25, Section 15.8]

$$\mathbf{R}_e = \mathbf{R}_h - \mathbf{R}_{h;h_p}\left(\mathbf{R}_{h_p}+\mathbf{I}_{N_p}\gamma^{-1}\right)^{-1}\mathbf{R}_{h;h_p}^H. \tag{6}$$

In the finite-length filtering case, the MSE on the diagonal of $\mathbf{R}_e$ depends on the position of the estimate. We use the average mismatched and matched MSE over all positions:

$$\tilde{\sigma}_{e,N,L}^2 = \frac{1}{N}\operatorname{tr}\left\{\tilde{\mathbf{R}}_e\right\}\ ; \qquad \sigma_{e,N,L}^2 = \frac{1}{N}\operatorname{tr}\left\{\mathbf{R}_e\right\}. \tag{7}$$

Using the real-valued and non-negative PSD of the channel process $C_h\left(e^{j2\pi\nu}\right)$, we obtain $\mathbf{R}_h$ from

$$[\mathbf{R}_h]_{k,l} = \int_{-\frac{1}{2}}^{\frac{1}{2}} C_h\left(e^{j2\pi\nu}\right)e^{j2\pi(k-l)\nu}d\nu,\ k,l = 1,\ldots,N. \tag{8}$$

[4]The *effective support* of a function $f(x)$ in x is the domain $[x_1,x_2]$ of smallest size with $|f(x)| \leq \epsilon, \forall x \notin [x_1,x_2]$ for small ϵ.

[5]Note that [22] only introduces the orthogonality principle for real random vectors; we use the extension to proper complex random vectors [23].

The correlation matrices $\mathbf{R}_{h_p}$ and $\mathbf{R}_{h;h_p}$ are obtained as submatrices of $\mathbf{R}_h$ by extracting L-spaced rows and/or columns. More specifically, we have $[\mathbf{R}_{h_p}]_{k,l} = [\mathbf{R}_h]_{(k-1)L+1,(l-1)L+1}$ for $k,l = 1,\ldots,N_p$ and $[\mathbf{R}_{h;h_p}]_{k,l} = [\mathbf{R}_h]_{k,(l-1)L+1}$ for $k = 1,\ldots,N$ and $l = 1,\ldots,N_p$.

B. Infinite-Length Filtering

In the following, we derive the MSE of the infinite-length Wiener filter using observations at pilot positions only. We start by recalling the derivation of the Wiener filter transfer function based on the orthogonality principle. The noisy observations at pilot positions are

$$y[kL] = h[kL]\sigma_p + n[kL],\ k \in \mathbb{Z} \tag{9}$$

and estimation is performed on the lth position relative to the pilot symbol at position kL, with $l = 0,\ldots,L-1$. From the orthogonality principle [22] $\mathrm{E}\left\{\left(\hat{h}[kL+l] - h[kL+l]\right) y^*[kL-sL]\right\} = 0,\ \forall s \in \mathbb{Z}$ and the linear estimate of the zero-mean channel $h[kL+l]$ denoted by $\hat{h}[kL+l] = \sum_i w_l[iL]\ y[kL-iL]$, we can obtain the transfer function of the Wiener filter coefficients $w_l[kL]$. The transfer function $W_l\left(e^{j2\pi L\nu}\right) = \sum_k w_l[kL]e^{-j2\pi kL\nu}$ is a function of the PSD of the channel based on the L-spaced pilot symbols

$$C_h\left(e^{j2\pi L\nu}\right) = \sum_{\Delta_k} R_h[\Delta_k L]e^{-j2\pi\Delta_k L\nu} \tag{10}$$

where $R_h[\Delta_m] = \mathrm{E}\left\{h[m]h^*[m-\Delta_m]\right\}$ is the autocorrelation function of the channel process.[6] The mismatched Wiener filter, resulting in the mismatched estimate $\tilde{h}[m]$, uses possibly erroneous statistical a-priori information on the channel process, i.e., $\tilde{C}_h\left(e^{j2\pi L\nu}\right)$ [27]:

$$\tilde{W}_l\left(e^{j2\pi L\nu}\right) = \frac{1}{\sigma_p}\ \frac{\tilde{C}_h\left(e^{j2\pi L\nu}\right)}{\tilde{C}_h\left(e^{j2\pi L\nu}\right)+\gamma^{-1}}\ e^{j2\pi l\nu}. \tag{11}$$

The PSD of the mismatched error process at a shifted position l relative to the pilot grid is given by

$$\tilde{C}_{e,l}\left(e^{j2\pi L\nu}\right) = \lim_{N_p\to\infty}\frac{1}{N_p}\mathrm{E}\left\{\left|\tilde{W}_l\left(e^{j2\pi L\nu}\right)\left(\sigma_p H_{0,N_p}\left(e^{j2\pi L\nu}\right)\right.\right.\right.$$
$$\left.\left.\left. + N_{N_p}\left(e^{j2\pi L\nu}\right)\right) - H_{l,N_p}\left(e^{j2\pi L\nu}\right)\right|^2\right\} \tag{12}$$

for an absolutely summable channel autocorrelation function and a stable filter. Here, we have the filter length N_p and

$$H_{l,N_p}\left(e^{j2\pi L\nu}\right) = \sum_{k=-\lfloor N_p/2\rfloor}^{\lceil N_p/2\rceil-1} h[kL+l]e^{-j2\pi kL\nu} \tag{13}$$

$$N_{N_p}\left(e^{j2\pi L\nu}\right) = \sum_{k=-\lfloor N_p/2\rfloor}^{\lceil N_p/2\rceil-1} n[kL]e^{-j2\pi kL\nu}. \tag{14}$$

[6]Note that (10) and its inverse hold under certain technical conditions, see, e.g., [26].

Using $\tilde{W}_l\left(e^{j2\pi L\nu}\right) = \left|\tilde{W}_l\left(e^{j2\pi L\nu}\right)\right| e^{j2\pi l\nu}$ from (11), we obtain

$$\tilde{C}_{e,l}\left(e^{j2\pi L\nu}\right) = \sigma_p^2\left|\tilde{W}_l\left(e^{j2\pi L\nu}\right)\right|^2\left(C_h\left(e^{j2\pi L\nu}\right)+\gamma^{-1}\right)$$
$$+ C_h\left(e^{j2\pi L\nu}\right) - 2\sigma_p\left|\tilde{W}_l\left(e^{j2\pi L\nu}\right)\right| C_h\left(e^{j2\pi L\nu}\right). \tag{15}$$

By inserting (11) into (15), the PSD of the mismatched error process follows as

$$\tilde{C}_{e,l}\left(e^{j2\pi L\nu}\right) = C_h\left(e^{j2\pi L\nu}\right) + \left(C_h\left(e^{j2\pi L\nu}\right)+\gamma^{-1}\right)$$
$$\times\ \frac{\tilde{C}_h^2\left(e^{j2\pi L\nu}\right)}{\left(\tilde{C}_h\left(e^{j2\pi L\nu}\right)+\gamma^{-1}\right)^2} - \frac{2C_h\left(e^{j2\pi L\nu}\right)\tilde{C}_h\left(e^{j2\pi L\nu}\right)}{\tilde{C}_h\left(e^{j2\pi L\nu}\right)+\gamma^{-1}}. \tag{16}$$

The resulting mismatched error PSD is independent of the position l of the estimate due to infinite-length filtering and the Nyquist criterion. We thus drop the index l from now on.

1) Symbol-Based Reformulation: By definition $C_h\left(e^{j2\pi L\nu}\right)$ and $\tilde{C}_h\left(e^{j2\pi L\nu}\right)$ are zero for $|\nu| \geq 1/(2L)$; therefore, $\tilde{C}_e\left(e^{j2\pi L\nu}\right)$ is also zero for $|\nu| \geq 1/(2L)$, see (16). We thus have

$$\tilde{C}_h\left(e^{j2\pi L\nu}\right) = \frac{1}{L}\tilde{C}_h\left(e^{j2\pi\nu}\right),\ -\frac{1}{2L} \leq \nu < \frac{1}{2L} \tag{17}$$

$$\tilde{C}_e\left(e^{j2\pi L\nu}\right) = \frac{1}{L}\tilde{C}_e\left(e^{j2\pi\nu}\right),\ -\frac{1}{2L} \leq \nu < \frac{1}{2L} \tag{18}$$

and accordingly for the matched case. Here the factor $1/L$ follows from the L-spaced discrete-time Fourier transform (DTFT). The (position-independent) mismatched and matched MSEs are obtained by an inverse DTFT, i.e., an integration over the PSD of the corresponding error process, as

$$\tilde{\sigma}^2_{e,\infty,L} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \tilde{C}_e\left(e^{j2\pi\nu}\right) d\nu\ ;\ \sigma^2_{e,\infty,L} = \int_{-\frac{1}{2}}^{\frac{1}{2}} C_e\left(e^{j2\pi\nu}\right) d\nu. \tag{19}$$

The symbol-based mismatched error PSD $\tilde{C}_e\left(e^{j2\pi\nu}\right)$ is obtained from (18) and (16) as

$$\tilde{C}_e\left(e^{j2\pi\nu}\right) = C_h\left(e^{j2\pi\nu}\right) + \left(C_h\left(e^{j2\pi\nu}\right)+L\gamma^{-1}\right)$$
$$\times\ \frac{\tilde{C}_h^2\left(e^{j2\pi\nu}\right)}{\left(\tilde{C}_h\left(e^{j2\pi\nu}\right)+L\gamma^{-1}\right)^2} - \frac{2C_h\left(e^{j2\pi\nu}\right)\tilde{C}_h\left(e^{j2\pi\nu}\right)}{\tilde{C}_h\left(e^{j2\pi\nu}\right)+L\gamma^{-1}} \tag{20}$$

$$= \frac{\left(L\gamma^{-1}\right)^2 C_h\left(e^{j2\pi\nu}\right) + L\gamma^{-1}\tilde{C}_h^2\left(e^{j2\pi\nu}\right)}{\left(\tilde{C}_h\left(e^{j2\pi\nu}\right)+L\gamma^{-1}\right)^2} \tag{21}$$

and the symbol-based matched error PSD is [27]

$$C_e\left(e^{j2\pi L\nu}\right) = \frac{C_h\left(e^{j2\pi\nu}\right)}{\frac{\gamma}{L}\ C_h\left(e^{j2\pi\nu}\right)+1}. \tag{22}$$

Note that the pilot spacing L can be interpreted as a noise power increase by a factor of L with respect to the $L = 1$ case.

2) Alternative Derivation of the MSE: Alternatively, we can derive the MSE in the infinite-length case by studying the asymptotic finite-length filtering MSE, i.e., by considering $N \to \infty$ in (7). Assuming that we have a block length $N = N_p L$, we give this derivation in Appendix A, i.e., we show that

$$\lim_{N\to\infty}\tilde{\sigma}^2_{e,N,L} = \tilde{\sigma}^2_{e,\infty,L}\ ;\quad \lim_{N\to\infty}\sigma^2_{e,N,L} = \sigma^2_{e,\infty,L}. \tag{23}$$

3) Properties of the Mismatched MSE: By differentiating (21) with respect to $L\gamma^{-1}$, we obtain a decreasing error PSD for an increasing noise power at a certain ν whenever

$$\frac{\gamma}{L} < \frac{\tilde{C}_h\left(e^{j2\pi\nu}\right) - 2C_h\left(e^{j2\pi\nu}\right)}{\tilde{C}_h^2\left(e^{j2\pi\nu}\right)} \tag{24}$$

is satisfied. Depending on the set of ν satisfying (24), a decreasing MSE for an increasing noise power might occur. This cannot be the case in the matched setting; however, it can occur for a mismatched estimation. A similar observation is made in [16] for mismatched continuous-time causal MMSE estimation.

We note that in the noise-dominated case, i.e., $L\gamma^{-1} \gg C_h\left(e^{j2\pi\nu}\right)$ and $L\gamma^{-1} \gg \tilde{C}_h\left(e^{j2\pi\nu}\right)$, with channel power gain $\int_{-\frac{1}{2}}^{\frac{1}{2}} C_h\left(e^{j2\pi\nu}\right) d\nu = \int_{-\frac{1}{2}}^{\frac{1}{2}} \tilde{C}_h\left(e^{j2\pi\nu}\right) d\nu$ a mismatch has only a minor influence on the resulting MSE, see (21) and (22).

C. Implications of the Linearity of the Estimator

Within this work we restrict to linear estimators. It is thus of interest to assess the limitations of this approach by quantifying the performance loss due to LMMSE estimation instead of MMSE estimation. More specifically, we want to characterize the additional MSE incurred by using the LMMSE instead of the MMSE estimator. Adapting results from [15] to our setting with complex random variables, we obtain for a pilot spacing $L = 1$:

$$D(P\|Q) = \int_0^\infty \mathrm{mse}_{Q|P}(\gamma) - \mathrm{mse}_{P|P}(\gamma) d\gamma \tag{25}$$

where $D(P\|Q)$ is the Kullback-Leibler divergence between the distributions P and Q. Here, $\mathrm{mse}_{Q|P}(\gamma)$, a function of $\gamma = \sigma_p^2/\sigma_n^2$, is the total MSE of an unrestricted, i.e., not necessarily linear, estimator treating the random vector to be estimated as distributed according to Q instead of its actual distribution P.[7] Note that in this case the MSE is not normalized, i.e., the sum MSE at all estimated time slots is used. We denote by P_h the actual distribution of the relevant part of the channel process $\{h[m]\}$ and by P_h^G the proper Gaussian distribution with the same first- and second-order moments as P_h. We define the matched LMMSE $\mathrm{lmmse}_{P_h}(\gamma) = \mathrm{mse}_{P_h^G|P_h}(\gamma)$, the matched MMSE $\mathrm{mmse}_{P_h}(\gamma) = \mathrm{mse}_{P_h|P_h}(\gamma)$, the mismatched LMMSE $\widetilde{\mathrm{lmmse}}_{P_h}(\gamma) = \mathrm{mse}_{\tilde{P}_h^G|P_h}(\gamma)$, and the mismatched MMSE $\widetilde{\mathrm{mmse}}_{P_h}(\gamma) = \mathrm{mse}_{\tilde{P}_h|P_h}(\gamma)$. With (25), it follows for $L = 1$ that

$$\int_0^\infty \mathrm{lmmse}_{P_h}(\gamma) - \mathrm{mmse}_{P_h}(\gamma) d\gamma = D(P_h\|P_h^G) \tag{26}$$

$$\int_0^\infty \widetilde{\mathrm{lmmse}}_{P_h}(\gamma) - \mathrm{mmse}_{P_h}(\gamma) d\gamma = D(P_h\|\tilde{P}_h^G) \tag{27}$$

$$\int_0^\infty \widetilde{\mathrm{lmmse}}_{P_h}(\gamma) - \widetilde{\mathrm{mmse}}_{P_h}(\gamma) d\gamma = D(P_h\|\tilde{P}_h^G) - D(P_h\|\tilde{P}_h). \tag{28}$$

We can thus characterize the degradation of the MSE due to, possibly mismatched, linear estimation compared to an unrestricted estimation. In [16], the implications of mismatched causal and non-causal linear estimation have been treated for the continuous-time setting in a similar manner. Finally, it should be mentioned that the (matched) LMMSE estimator is optimal in the sense that it minimizes the maximal MSE over all joint channel and observation distributions $P_{h,y}$ with fixed first- and second-order moments [28]. In other words, the LMMSE estimator is the solution of [28, Theorem 1]

$$\min_{\hat{\mathbf{h}} \in \mathcal{S}_y} \max_{P_{h,y} \in \mathcal{S}_P} \mathrm{E}\left\{ \left\| \hat{\mathbf{h}} - \mathbf{h} \right\|^2 \right\} \tag{29}$$

where $\mathcal{S}_P$ is the set of all joint channel and observation distributions having the fixed first- and second-order moments and $\mathcal{S}_y$ is the set of random vectors that are arbitrary functions of the observation.

IV. Application to Non-Stationary Channels

In Section III, we assumed the channel process to be WSS. In a real scenario, the wireless channel is non-stationary. However, the channel process can be assumed to be stationary inside small regions, see Section II-A. Wiener filtering in these assumed stationarity regions, i.e., $N \leq N_\mathrm{s}$, is appropriate. Furthermore, due to the change of the channel statistics over time, mismatched Wiener filtering with, e.g., outdated channel statistics, might occur. Based on the mismatched MSE, we have the possibility to evaluate the performance degradation of a realistic channel estimator due to wrong statistical knowledge of the channel. However, the resulting expressions for the MSE, i.e., (7) with (5)/(6), require a large computational effort, e.g., matrix inversions are required.

In contrast, the infinite-length filtering MSE, i.e., (19) with (21)/(22), allows for a simplified evaluation. The infinite-length Wiener filtering approach is strictly speaking not appropriate because the statistics of the channel process can change significantly over an infinite-length block. However, as we will see in Section IV-A, common wireless channels have an effectively finite correlation; thus, infinite-length filtering can be considered as meaningful. In the subsequent analysis, we will give an approximate evaluation of the MSE based on infinite-length filtering. Furthermore, we will show how the approximate evaluation of the MSE is related to the actual MSE based on finite-length filtering.

A. Channels with an Effectively Finite Correlation

Due to the correlation-underspread property of the channel, i.e., $N_\mathrm{c} \ll N_\mathrm{s}$, the coherence time N_c is much smaller than the stationarity time N_s. With a small coherence time, we can motivate a time-varying autocorrelation function of the channel that is approximately zero outside a finite interval $-N'', \ldots, N''$ of length $N' = 2N'' + 1$, i.e., that is *effectively* supported on $-N'', \ldots, N''$. Thus, we can assume the channel to be effectively correlated over a finite interval with $N' \gg N_\mathrm{c}$.[8] Furthermore, we will see in Section V that the estimation of the statistics of non-stationary channels from a single measurement run is based on a windowing over the channel process and, thus, on the assumption of a

[7] A proof of (25) for the multivariate case can be found in [16].

[8] Note that a *strictly* bandlimited signal cannot be *strictly* timelimited, but only *effectively* timelimited. See [29] for a detailed discussion.

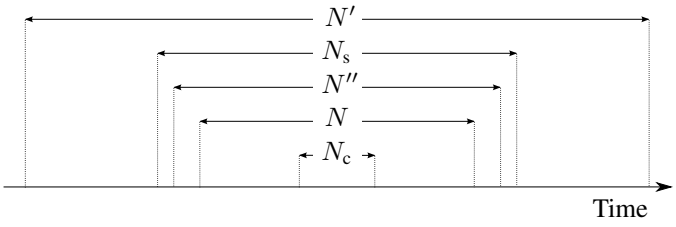

Fig. 1. Visualization of the relevant time quantities: N_c is the coherence time of the channel, N_s is the stationarity time of the channel, $N' = 2N'' + 1$ is the effective support of the autocorrelation function of the channel, and N is the block length. Exemplarily, the case $N_\mathrm{s} > N''$ is shown.

finite correlation of the channel. Therefore, we perform an evaluation for an effectively finite correlation of the channel. We relate the block length N to the assumed finite correlation length of the channel $N' = 2N'' + 1$ such that $N'' \geq N - 1$ is satisfied. This implies that every pair of symbols in the block is correlated.[9] Moreover, we assume N' to be a multiple of L. Recall that we perform estimation inside a stationarity region; therefore, we choose $N \leq N_\mathrm{s}$. See Fig. 1 for a visualization of the relevant time quantities.

In order to simplify the exposition, we restrict to a WSS channel for the remainder of Section IV-A. Defining

$$C_h^{(\mathrm{f})}\left(e^{j2\pi\nu}\right) = \sum_{\Delta_m=-N''}^{N''} R_h[\Delta_m]\; e^{-j2\pi\Delta_m\nu} \tag{30}$$

with $R_h[\Delta_m] = \mathrm{E}\left\{h[m]h^*[m-\Delta_m]\right\}$, we have the following approximate relation due to the assumption of an effectively finite correlation of the channel, cf. (10) for $L = 1$:

$$C_h\left(e^{j2\pi\nu}\right) \approx C_h^{(\mathrm{f})}\left(e^{j2\pi\nu}\right). \tag{31}$$

We insert (30) into (31) and note that (30) is a discrete Fourier transform (DFT); therefore, we can use samples of $C_h^{(\mathrm{f})}\left(e^{j2\pi\nu}\right)$ to calculate $R_h[\Delta_m]$ based on an inverse DFT. By substituting $R_h[\Delta_m]$ with the inverse DFT of $C_h^{(\mathrm{f})}\left(e^{j2\pi\frac{k}{N'}}\right)$ for $k = -N'', \ldots, N''$, we obtain the Doppler-continuous PSD of the channel as

$$\begin{aligned} &C_h\left(e^{j2\pi\nu}\right) \\ &\approx \sum_{\Delta_m=-N''}^{N''} \frac{1}{N'} \sum_{i=-N''}^{N''} C_h^{(\mathrm{f})}\left(e^{j2\pi\frac{i}{N'}}\right) e^{j2\pi\Delta_m\frac{i}{N'}} e^{-j2\pi\Delta_m\nu} \\ &= \frac{1}{N'} \sum_{i=-N''}^{N''} C_h^{(\mathrm{f})}\left(e^{j2\pi\frac{i}{N'}}\right) \sum_{\Delta_m=-N''}^{N''} e^{j2\pi\Delta_m\left(\frac{i}{N'}-\nu\right)}. \end{aligned} \tag{32}$$

The mismatched case follows analogously.

1) Finite-Length Filtering: In the finite-length filtering case of length N, we can insert (32) into (8). We then obtain the correlation matrices using samples of the matched PSD of the channel process for $k, l = 1, \ldots, N$ as

$$[\mathbf{R}_h]_{k,l} \approx \frac{1}{N'} \sum_{i=-N''}^{N''} C_h^{(\mathrm{f})}\left(e^{j2\pi\frac{i}{N'}}\right) e^{j2\pi(k-l)\frac{i}{N'}}. \tag{33}$$

[9]The finite-length filtering approach with a block length N performs a block-wise estimation. It thus only makes use of the correlation properties of the channel for time differences of at most $-(N-1), \ldots, N-1$, i.e., a maximum of $N-1$ time instants in each time direction, see Section III-A.

Analogously, we have $[\mathbf{R}_{h_p}]_{k,l} = [\mathbf{R}_h]_{(k-1)L+1,(l-1)L+1}$ for $k, l = 1, \ldots, N_p$ and $[\mathbf{R}_{h;h_p}]_{k,l} = [\mathbf{R}_h]_{k,(l-1)L+1}$ for $k = 1, \ldots, N$ and $l = 1, \ldots, N_p$. The mismatched case follows accordingly. In matrix notation, we obtain

$$\begin{aligned} \mathbf{R}_h &\approx \check{\mathbf{F}}_h^H \mathbf{C}_h^{(\mathrm{f})} \check{\mathbf{F}}_h \\ \mathbf{R}_{h_p} &\approx \check{\mathbf{F}}_{h,L}^H \mathbf{C}_h^{(\mathrm{f})} \check{\mathbf{F}}_{h,L}\;; \qquad \tilde{\mathbf{R}}_{h_p} \approx \check{\mathbf{F}}_{h,L}^H \tilde{\mathbf{C}}_h^{(\mathrm{f})} \check{\mathbf{F}}_{h,L} \\ \mathbf{R}_{h;h_p} &\approx \check{\mathbf{F}}_h^H \mathbf{C}_h^{(\mathrm{f})} \check{\mathbf{F}}_{h,L}\;; \qquad \tilde{\mathbf{R}}_{h;h_p} \approx \check{\mathbf{F}}_h^H \tilde{\mathbf{C}}_h^{(\mathrm{f})} \check{\mathbf{F}}_{h,L}. \end{aligned} \tag{34}$$

The matrices $\check{\mathbf{F}}_h$ and $\check{\mathbf{F}}_{h,L}$ are sub-matrices of the $N' \times N'$ DFT matrix $\check{\mathbf{F}}$. The $N' \times N$ matrix $\check{\mathbf{F}}_h$ is defined, for $k = 1, \ldots, N'$ and $l = 1, \ldots, N$, by

$$\left[\check{\mathbf{F}}_h\right]_{k,l} = \frac{1}{\sqrt{N'}} \exp\left(-j2\pi\frac{(k-1)(l-1)}{N'}\right) \tag{35}$$

such that it contains the first N columns of $\check{\mathbf{F}}$. The $N' \times N_p$ matrix $\check{\mathbf{F}}_{h,L}$ is defined, for $k = 1, \ldots, N'$ and $l = 1, \ldots, N_p$, by

$$\left[\check{\mathbf{F}}_{h,L}\right]_{k,l} = \frac{1}{\sqrt{N'}} \exp\left(-j2\pi\frac{(k-1)(l-1)L}{N'}\right) \tag{36}$$

i.e., it contains only the columns of $\check{\mathbf{F}}_h$ coinciding with pilot positions. The following properties hold for $k = 1, \ldots, N'$:

$$\check{\mathbf{F}}_h^H \check{\mathbf{F}}_h = \mathbf{I}_N\;; \qquad \left[\check{\mathbf{F}}_h \check{\mathbf{F}}_h^H\right]_{k,k} = \frac{N}{N'} \tag{37}$$

$$\check{\mathbf{F}}_{h,L}^H \check{\mathbf{F}}_{h,L} = \mathbf{I}_{N_p}\;; \qquad \left[\check{\mathbf{F}}_{h,L} \check{\mathbf{F}}_{h,L}^H\right]_{k,k} = \frac{N_p}{N'}. \tag{38}$$

The $N' \times N'$ matrices $\mathbf{C}_h^{(\mathrm{f})}$ and $\tilde{\mathbf{C}}_h^{(\mathrm{f})}$ are diagonal matrices containing regular samples of the matched and mismatched PSD of the channel process, respectively:

$$\left[\mathbf{C}_h^{(\mathrm{f})}\right]_{k,k} = \begin{cases} C_h^{(\mathrm{f})}\left(e^{j2\pi\frac{k-1}{N'}}\right), & \text{for } k = 1, \ldots, \frac{N'+1}{2} \\ C_h^{(\mathrm{f})}\left(e^{j2\pi\frac{k-1-N'}{N'}}\right), & \text{for } k = \frac{N'+3}{2}, \ldots, N' \end{cases} \tag{39}$$

and accordingly for the mismatched case.

2) Infinite-Length Filtering: In the infinite-length filtering case, we can insert (32) and its mismatched equivalent into (21) for the mismatched case and into (22) for the matched case. We then obtain the PSD of the corresponding error process. As this would still necessitate the integration over the PSDs of the error processes to obtain both MSEs in (19), we choose a different approach. Rewriting (30), we obtain

$$\begin{aligned} C_h^{(\mathrm{f})}\left(e^{j2\pi\nu}\right) &= \sum_{\Delta_m} R_h[\Delta_m] \operatorname{rect}_{N''}[\Delta_m]\, e^{-j2\pi\Delta_m\nu} \\ &= C_h\left(e^{j2\pi\nu}\right) \circledast \frac{\sin(\pi\nu N')}{\sin(\pi\nu)}, \quad -\frac{1}{2} \leq \nu < \frac{1}{2} \end{aligned} \tag{40}$$

with $\operatorname{rect}_{N''}[m] = 1, |m| \leq N''$ and $\operatorname{rect}_{N''}[m] = 0$ else. From (21), we can deduce that the maximal Doppler of $\tilde{C}_e\left(e^{j2\pi\nu}\right)$ is equal to the maximum of the maximal Dopplers of $C_h\left(e^{j2\pi\nu}\right)$ and $\tilde{C}_h\left(e^{j2\pi\nu}\right)$. This implies that the corresponding coherence times, given by the inverses of the maximal Dopplers, are the same as well.

However, having equal coherence times of two random processes does not mean that the random processes have an

effectively finite correlation over the same interval. Thus, an effectively finite correlation of the channel process over a certain interval does not imply that the error process can be accurately described by a finite correlation over the same interval. We want to describe the implications of approximating the error process as finitely correlated over the same interval. To this end, we will relate the resulting MSE to the one of the finite-length case in Section IV-D. Approximating the error process as finitely correlated over an interval of length N', i.e.,

$$\tilde{C}_e\left(e^{j2\pi\nu}\right) \approx \tilde{C}_e^{(\mathrm{f})}\left(e^{j2\pi\nu}\right) \tag{41}$$

with

$$\tilde{C}_e^{(\mathrm{f})}\left(e^{j2\pi\nu}\right) = \tilde{C}_e\left(e^{j2\pi\nu}\right) \circledast \frac{\sin(\pi\nu N')}{\sin(\pi\nu)}, \quad -\frac{1}{2} \leq \nu < \frac{1}{2} \tag{42}$$

we obtain analogously to (32)

$$\begin{aligned}\tilde{C}_e\left(e^{j2\pi\nu}\right) &\approx \frac{1}{N'}\sum_{k=-N''}^{N''} \tilde{C}_e^{(\mathrm{f})}\left(e^{j2\pi\frac{k}{N'}}\right) \sum_{\Delta_m=-N''}^{N''} e^{j2\pi\Delta_m\left(\frac{k}{N'}-\nu\right)} \\ &= \frac{1}{N'}\sum_{k=-N''}^{N''} \tilde{C}_e^{(\mathrm{f})}\left(e^{j2\pi\frac{k}{N'}}\right) \frac{\sin\left(\pi\left(\nu N'-k\right)\right)}{\sin\left(\pi\frac{\nu N'-k}{N'}\right)}.\end{aligned} \tag{43}$$

From (43), we see that only samples of $\tilde{C}_e^{(\mathrm{f})}\left(e^{j2\pi\nu}\right)$ are required for the approximate reconstruction of $\tilde{C}_e\left(e^{j2\pi\nu}\right)$. The required samples of $\tilde{C}_e^{(\mathrm{f})}\left(e^{j2\pi\nu}\right)$ are given with (41), (21), (31), and the mismatched version of (31), for $k = -N'', \ldots, N''$, as

$$\begin{aligned}\tilde{C}_e^{(\mathrm{f})}\left(e^{j2\pi\frac{k}{N'}}\right) &\approx \tilde{C}_e\left(e^{j2\pi\frac{k}{N'}}\right) \\ &\approx \frac{\left(L\gamma^{-1}\right)^2 C_h^{(\mathrm{f})}\left(e^{j2\pi\frac{k}{N'}}\right) + L\gamma^{-1}\left(\tilde{C}_h^{(\mathrm{f})}\left(e^{j2\pi\frac{k}{N'}}\right)\right)^2}{\left(\tilde{C}_h^{(\mathrm{f})}\left(e^{j2\pi\frac{k}{N'}}\right) + L\gamma^{-1}\right)^2}\end{aligned} \tag{44}$$

and accordingly for the matched case.

B. Time-Dependent Power Spectral Density

The local scattering function (LSF) is an extension of the scattering function in the context of WSSUS channels to the non-stationary case [2]. In the special case of a frequency-flat fading channel, it is thus a time-dependent PSD in the Doppler domain. We adapt the continuous-time approach in [2] to discrete-time channels. This is reasonable as we assume to have a maximum Doppler shift that allows to sufficiently sample the underlying continuous-time channel, i.e., to yield a sufficient statistic by sampling. Furthermore, channel measurements rely on the same argument and are only available at discrete time instants. We thus obtain the LSF

$$C_h\left(m; e^{j2\pi\nu}\right) = \sum_{\Delta_m} R_h[m;\Delta_m] e^{-j2\pi\Delta_m\nu} \tag{45}$$

where $R_h[m;\Delta_m] = \mathrm{E}\left\{h[m]h^*[m-\Delta_m]\right\}$ is the time-dependent autocorrelation function of the channel. Note that a WSS channel has a constant LSF over time, i.e., it is independent of m, and thus (45) reduces to $C_h\left(m; e^{j2\pi\nu}\right) = C_h\left(e^{j2\pi\nu}\right)$. Furthermore, a (zero-mean) WSS channel is uncorrelated in the Doppler domain [2].

In [19], the PSD of a discrete-time channel is related to the PSD of the underlying continuous-time channel. The straightforward generalization to the non-stationary setting is

$$C_h\left(m; e^{j2\pi\nu}\right) = \frac{1}{T}\sum_{\Delta_m} \Gamma_h\left(m; \frac{\nu-\Delta_m}{T}\right) \tag{46}$$

where $\Gamma_h(m;\nu') = \int \mathrm{E}\left\{h(t)h^*(t-\Delta_t)\right\} e^{-j2\pi\Delta_t\nu'} d\Delta_t$ denotes the time-dependent PSD of the continuous-time channel $h(t)$ and T is the symbol period, i.e., $h[m] = h(mT)$.

The LSF has some deficiencies with respect to the PSD of a stationary process, e.g., it is not guaranteed to be real-valued and non-negative. For doubly underspread channels, it is possible to define generalized LSFs (GLSFs) which are smoothed versions of the LSF and do not have the above deficiencies [2]. We adapt the approach in [2], [13] to the discrete-time and frequency-flat fading case; we define the GLSF by a 2-dimensional convolution:

$$\begin{aligned}C_h^{(\Phi)}\left(m; e^{j2\pi\nu}\right) = \sum_{\check{m}} \int_{-\frac{1}{2}}^{\frac{1}{2}} & C_h\left(\check{m}; e^{j2\pi\check{\nu}}\right) \\ & \times \Phi\left(m-\check{m}; e^{j2\pi(\nu-\check{\nu})}\right) d\check{\nu}\end{aligned} \tag{47}$$

with

$$\Phi\left(m; e^{j2\pi\nu}\right) = \sum_{s=1}^{S} \gamma_s \sum_{\Delta_m} g_s^*[-m] g_s[-m-\Delta_m] e^{-j2\pi\Delta_m\nu} \tag{48}$$

the windowing functions $g_s[m]$ normalized to unit-energy, the number of windows S, and $\gamma_s \geq 0$ with $\sum_{s=1}^{S}\gamma_s = 1$ for $s = 1, \ldots, S$. The parametrization of the windowing functions is particularly important for the estimation of the GLSF: they should be localized in time and Doppler, whereas their length and number are subject to bias-variance trade-offs [2], [30]. In [2], it is shown that GLSFs of doubly underspread channels are real-valued, non-negative, and approximately equivalent to the LSF. For further details on the meaning of approximate equivalence, we refer to [2]. In the following, we use the term (time-dependent) PSD instead of the term GLSF.

C. Mean Square Error

The MSE evaluation for non-stationary channels is based on the time-dependent PSD of the channel process, i.e., $C_h^{(\Phi)}\left(m; e^{j2\pi\nu}\right)$. Due to the effectively finite correlation of the channel process, we only need samples of the PSD in the Doppler domain, i.e., $C_h^{(\Phi)}[m;k] = C_h^{(\Phi)}\left(m; e^{j2\pi\frac{k}{N'}}\right)$ for $k = -N'', \ldots, N''$, see (32).

1) Finite-Length Filtering: For the finite-length filtering case, we use (5) with (34). We obtain the *actual FL-based mismatched MSE* at time instant m using statistical knowledge

at time instant m' as

$$\begin{aligned}\tilde{\sigma}^2_{e,N,L}[m,m'] &= \frac{1}{N}\operatorname{tr}\left\{\tilde{\mathbf{R}}_e[m,m']\right\}\\ &\approx \frac{1}{N}\operatorname{tr}\Big\{\check{\mathbf{F}}_h^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_h+\check{\mathbf{F}}_h^H\tilde{\mathbf{C}}_h^{(\Phi)}\check{\mathbf{F}}_{h,L}\\ &\times\left(\check{\mathbf{F}}_{h,L}^H\tilde{\mathbf{C}}_h^{(\Phi)}\check{\mathbf{F}}_{h,L}+\mathbf{I}_{N_p}\gamma^{-1}\right)^{-1}\left(\check{\mathbf{F}}_{h,L}^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_{h,L}+\mathbf{I}_{N_p}\gamma^{-1}\right)\\ &\times\left(\check{\mathbf{F}}_{h,L}^H\tilde{\mathbf{C}}_h^{(\Phi)}\check{\mathbf{F}}_{h,L}+\mathbf{I}_{N_p}\gamma^{-1}\right)^{-1}\check{\mathbf{F}}_{h,L}^H\tilde{\mathbf{C}}_h^{(\Phi)}\check{\mathbf{F}}_h\\ &-\check{\mathbf{F}}_h^H\tilde{\mathbf{C}}_h^{(\Phi)}\check{\mathbf{F}}_{h,L}\left(\check{\mathbf{F}}_{h,L}^H\tilde{\mathbf{C}}_h^{(\Phi)}\check{\mathbf{F}}_{h,L}+\mathbf{I}_{N_p}\gamma^{-1}\right)^{-1}\check{\mathbf{F}}_{h,L}^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_h\\ &-\check{\mathbf{F}}_h^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_{h,L}\left(\check{\mathbf{F}}_{h,L}^H\tilde{\mathbf{C}}_h^{(\Phi)}\check{\mathbf{F}}_{h,L}+\mathbf{I}_{N_p}\gamma^{-1}\right)^{-1}\check{\mathbf{F}}_{h,L}^H\tilde{\mathbf{C}}_h^{(\Phi)}\check{\mathbf{F}}_h\Big\}\end{aligned} \tag{49}$$

where $\tilde{\mathbf{R}}_e[m,m']$ is the mismatched MSE matrix. Here, we use that the time-dependent PSD is real-valued. The difference to Section IV-A1 lies in the substitution of (39) by the (diagonal) time-dependent PSD matrices with

$$\left[\mathbf{C}_h^{(\Phi)}\right]_{k,k} = \begin{cases} C_h^{(\Phi)}[m;k-1], & \text{for } k=1,\ldots,\frac{N'+1}{2}\\ C_h^{(\Phi)}[m;k-1-N'], & \text{for } k=\frac{N'+3}{2},\ldots,N'\end{cases} \tag{50}$$

$$\left[\tilde{\mathbf{C}}_h^{(\Phi)}\right]_{k,k} = \begin{cases} C_h^{(\Phi)}[m';k-1], & \text{for } k=1,\ldots,\frac{N'+1}{2}\\ C_h^{(\Phi)}[m';k-1-N'], & \text{for } k=\frac{N'+3}{2},\ldots,N'\end{cases} \tag{51}$$

where we drop the time argument for the matrices $\mathbf{C}_h^{(\Phi)}$ and $\tilde{\mathbf{C}}_h^{(\Phi)}$ to simplify notation. Based on (6) and (34), the *actual FL-based matched MSE* at time instant m follows as

$$\begin{aligned}\sigma^2_{e,N,L}[m] &= \frac{1}{N}\operatorname{tr}\{\mathbf{R}_e[m]\}\\ &\approx \frac{1}{N}\operatorname{tr}\Big\{\check{\mathbf{F}}_h^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_h-\check{\mathbf{F}}_h^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_{h,L}\\ &\quad\times\left(\check{\mathbf{F}}_{h,L}^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_{h,L}+\mathbf{I}_{N_p}\gamma^{-1}\right)^{-1}\check{\mathbf{F}}_{h,L}^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_h\Big\}\end{aligned} \tag{52}$$

where $\mathbf{R}_e[m]$ is the matched MSE matrix which follows from the mismatched one for $\tilde{\mathbf{C}}_h^{(\Phi)}=\mathbf{C}_h^{(\Phi)}$.

2) Infinite-Length Filtering: For the infinite-length filtering case, we define the time-dependent error PSD at time instant m using statistical knowledge of the channel from time instant m' as $\tilde{C}_e\left(m,m';e^{j2\pi\nu}\right)$. Using (43), we obtain the mismatched MSE based on (19) as

$$\begin{aligned}\tilde{\sigma}^2_{e,\infty,L}[m,m'] &= \int_{-\frac{1}{2}}^{\frac{1}{2}}\tilde{C}_e\left(m,m';e^{j2\pi\nu}\right)d\nu\\ &\approx \int_{-\frac{1}{2}}^{\frac{1}{2}}\frac{1}{N'}\sum_{k=-N''}^{N''}\tilde{C}_e^{(\mathrm{f})}\left(m,m';e^{j2\pi\frac{k}{N'}}\right)\frac{\sin\left(\pi\left(\nu N'-k\right)\right)}{\sin\left(\pi\frac{\nu N'-k}{N'}\right)}d\nu\\ &= \frac{1}{N'}\sum_{k=-N''}^{N''}\tilde{C}_e^{(\mathrm{f})}\left(m,m';e^{j2\pi\frac{k}{N'}}\right)\end{aligned} \tag{53}$$

i.e., the integration in (19) is replaced by a summation. We denote this approximate evaluation, i.e., the use of (41), by "ap". Thus, with (44), the *approximate IL-based mismatched MSE* at time instant m using channel statistics of time instant m' is

$$\begin{aligned}&\tilde{\sigma}^2_{e,\mathrm{ap},L}[m,m']\\ &= \frac{1}{N'}\sum_{k=-N''}^{N''}\frac{\left(L\gamma^{-1}\right)^2C_h^{(\Phi)}[m;k]+L\gamma^{-1}C_h^{(\Phi)^2}[m';k]}{\left(C_h^{(\Phi)}[m';k]+L\gamma^{-1}\right)^2}.\end{aligned} \tag{54}$$

The *approximate IL-based matched MSE* at time instant m

$$\sigma^2_{e,\mathrm{ap},L}[m] = \frac{1}{N'}\sum_{k=-N''}^{N''}\frac{C_h^{(\Phi)}[m;k]}{\frac{\gamma}{L}C_h^{(\Phi)}[m;k]+1} \tag{55}$$

is obtained analogously. With $\tilde{\sigma}^2_{e,\mathrm{ap},L}[m,m']$ and $\sigma^2_{e,\mathrm{ap},L}[m]$, we have approximate, but simplified, expressions of the MSEs.

D. Relations Between the Actual FL-Based and the Approximate IL-Based MSE

In the following, we justify the use of the approximate IL-based MSE by relating it to the actual FL-based MSE. More specifically, we give a condition on the sampled PSD of the channel $C_h^{(\Phi)}[m;k]$ for which $\sigma^2_{e,N,L}[m]\geq\sigma^2_{e,\mathrm{ap},L}[m]$ is satisfied.

We consider a rectangular sampled PSD of the channel $C_h^{(\Phi)}[m;k]$ with a maximal Doppler satisfying[10]

$$\frac{1}{2L}-\frac{1}{2N'}<\nu_{\max}[m]<\frac{1}{2L}. \tag{56}$$

Note that a rectangular PSD should be seen as an approximative PSD since its corresponding correlation function (sinc) is only slowly diminishing. The limitation to maximal Doppler frequencies close to the Nyquist limit in (56) is not a severe restriction. It is relevant in the case of joint processing of pilot and data symbols [31], e.g., using iterative receivers. The reason is that then channel sampling is performed close to the limit imposed by the Nyquist criterion (2). Rectangular PSDs of the channel restricted to (56) can be used to model a change in the channel power gain $P_h[m]=\sum_{k=-N''}^{N''}C_h^{(\Phi)}[m;k]/N'$ as well as a small variation in the maximal Doppler $\nu_{\max}[m]$. In Appendix B, we show that the approximate IL-based matched MSE $\sigma^2_{e,\mathrm{ap},L}[m]$ is a lower bound to the actual FL-based matched MSE $\sigma^2_{e,N,L}[m]$ for a rectangular PSD of the channel satisfying (56).

Consider the case of a rectangular mismatched PSD of the channel with (56), an arbitrary matched PSD of the channel, and a pilot spacing $L=1$. Under these conditions, it can be shown that the actual and the approximate MSE are equivalent, i.e., $\tilde{\sigma}^2_{e,N,L=1}[m]=\tilde{\sigma}^2_{e,\mathrm{ap},L=1}[m]$. This result is not surprising as the correlation matrices $\tilde{\mathbf{R}}_{h_p}$ and $\tilde{\mathbf{R}}_{h;h_p}$ in (34) are, in this case, scaled identity matrices; therefore, each channel entry is estimated based on its own observation only, see (4).

A sampled PSD of the channel with a rectangular shape satisfying (56) is of particular interest. Subject to a fixed channel power gain $P_h[m]$, it results in a maximal approximate

[10]We use $\nu_{\max}[m]$ with a time argument to indicate that the maximal Doppler can change over time.

IL-based matched MSE. In contrast, a sampled PSD of the channel with a single peak minimizes the approximate IL-based matched MSE subject to a fixed channel power gain. The proofs are provided in Appendix C. Note that the closer $N-1$ is to N'', the more similar the approximate MSE is to the actual MSE. The reason is that the additional observations used by infinite-length filtering are then reduced.

E. Local Quasi-Stationarity

We aim to relate the size of LQS regions to the performance degradation of a channel estimator. Therefore, we introduce the degradation of the MSE due to the use of mismatched channel statistics, i.e., channel statistics from another time instant. The MSE degradation at time instant m using, possibly mismatched, statistical knowledge of the channel corresponding to time instant m' is defined as

$$\eta_{N,L}[m,m'] = \frac{\tilde{\sigma}^2_{e,N,L}[m,m']}{\sigma^2_{e,N,L}[m]} - 1. \tag{57}$$

The approximate IL-based MSE degradation $\eta_{\text{ap},L}[m,m']$ follows accordingly. We define the difference between the approximate IL-based and the actual FL-based MSE degradation as

$$\begin{aligned}\Delta_{\eta,N,L}[m,m'] &= \eta_{\text{ap},L}[m,m'] - \eta_{N,L}[m,m'] \\ &= \frac{\tilde{\sigma}^2_{e,\text{ap},L}[m,m']\sigma^2_{e,N,L}[m] - \tilde{\sigma}^2_{e,N,L}[m,m']\sigma^2_{e,\text{ap},L}[m]}{\sigma^2_{e,N,L}[m]\sigma^2_{e,\text{ap},L}[m]}.\end{aligned} \tag{58}$$

With (58), it follows that the approximate IL-based MSE degradation upper-bounds the actual FL-based MSE degradation if and only if

$$\frac{\tilde{\sigma}^2_{e,\text{ap},L}[m,m']}{\sigma^2_{e,\text{ap},L}[m]} \geq \frac{\tilde{\sigma}^2_{e,N,L}[m,m']}{\sigma^2_{e,N,L}[m]}. \tag{59}$$

Consider now the case that the matched and the mismatched PSD are both rectangular with a maximal Doppler $\nu_{\max}[m]$ satisfying (56). This is representative for a setting with a mismatch in the channel power gain $P_h[m]$ or a small mismatch in $\nu_{\max}[m]$. In this case, it can be shown that the MSEs are equal to

$$\sigma^2_{e,N,L}[m] = P_h[m] - P_{\times,N,L}[m] \tag{60}$$

$$\tilde{\sigma}^2_{e,N,L}[m,m'] = P_h[m] - \tilde{P}_{\times,N,L}[m,m'] \tag{61}$$

$$\sigma^2_{e,\text{ap},L}[m] = P_h[m] - d_{N,L}P_{\times,N,L}[m] \tag{62}$$

$$\tilde{\sigma}^2_{e,\text{ap},L}[m,m'] = P_h[m] - d_{N,L}\tilde{P}_{\times,N,L}[m,m'] \tag{63}$$

where $P_{\times,N,L}[m]$ and $\tilde{P}_{\times,N,L}[m,m']$ are defined in accordance with (52) and (49), respectively. With (91) from Appendix B, we have

$$d_{N,L} = \frac{N}{L^2 \operatorname{tr}\left\{\bar{\mathbf{F}}_h^H \bar{\mathbf{F}}_{h,L} \bar{\mathbf{F}}_{h,L}^H \bar{\mathbf{F}}_h\right\}} \geq 1 \tag{64}$$

which is a factor arising from the bounding of the last term(s) of the actual FL-based MSEs. See also Appendix B. With (58), it then follows that

$$\begin{aligned}&\Delta_{\eta,N,L}[m,m'] \\ &= (d_{N,L}-1)P_h[m]\frac{\tilde{\sigma}^2_{e,N,L}[m,m'] - \sigma^2_{e,N,L}[m]}{\sigma^2_{e,N,L}[m]\sigma^2_{e,\text{ap},L}[m]} \geq 0\end{aligned} \tag{65}$$

holds, i.e., the actual FL-based MSE degradation is always upper-bounded by the approximate IL-based MSE degradation if the matched and the mismatched PSD are both rectangular and satisfy (56).

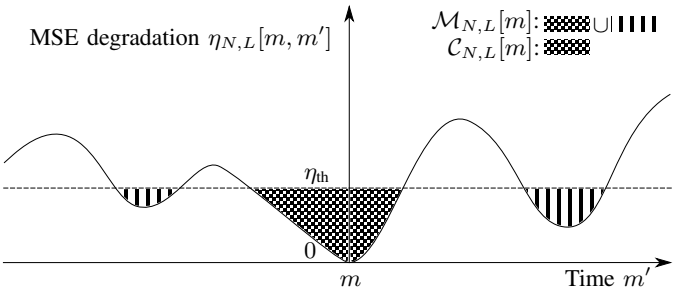

Fig. 2. Visualization of the definition of LQS regions.

Having defined the MSE degradation as a measure to characterize non-stationarity, we proceed to the definition of LQS regions. In order to reflect the performance degradation the system engineer is willing to accept, we define a threshold on the MSE degradation. This threshold is chosen as the maximal MSE degradation, due to mismatched statistics of the channel, that is deemed acceptable. The resulting LQS regions in time are then related to the required update rate of the channel statistics for the channel estimator. Defining the set

$$\mathcal{M}_{N,L}[m] = \{m' \mid \eta_{N,L}[m,m'] < \eta_{\text{th}}\} \tag{66}$$

we obtain the (time-dependent) actual FL-based LQS time

$$T_{\text{LQS},N,L}[m] = |\mathcal{C}_{N,L}[m]|\, T \tag{67}$$

where T is the symbol duration and $\mathcal{C}_{N,L}[m]$ is the connected subset of $\mathcal{M}_{N,L}[m]$ containing m and having maximum cardinality.[11] The approximate IL-based LQS time $T_{\text{LQS},\text{ap},L}[m]$ follows accordingly. Fig. 2 depicts the definition of LQS regions. We mentioned that for certain PSDs of the channel the actual FL-based MSE degradation is upper-bounded by the approximate IL-based MSE degradation. Therefore, in such cases, the size of the actual FL-based LQS regions is lower-bounded by the size of the approximate IL-based LQS regions.

V. Analysis

In this section, we will exemplarily apply the developed concepts for a non-stationarity analysis to a rectangular PSD with varying $\nu_{\max}[m]$, on the one hand, and to a realistic setting based on a measured channel, on the other hand.

[11]Note that our definition of locally quasi-stationary random processes is different to the definition of locally stationary random processes introduced in [32]. In contrast to this work, [32] imposes a structure on the covariance of the random process.

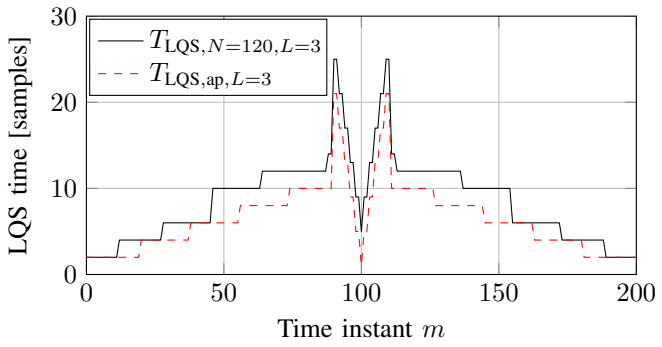

Fig. 3. LQS times for the SNR $\rho = 10$ dB and $\eta_{\mathrm{th}} = 0.1$.

A. Exemplary Analysis for a Rectangular PSD of the Channel

We consider a setting with a rectangular PSD and varying maximal Doppler shift $\nu_{\max}[m]$. More specifically, we assume a fixed channel power gain $P_h = P_h[m] = \sum_{k=-N''}^{N''} C_h^{(\Phi)}[m;k]/N'$ and the following evolution of $\nu_{\max}[m]$ for $m = 0, \ldots, N_{\mathrm{track}} - 1$:

$$\nu_{\max}[m] = \begin{cases} \nu_{\mathrm{inc}} m + \nu_0, & \text{for } 0 \le m \le \left\lceil \frac{N_{\mathrm{track}}}{2} \right\rceil - 1 \\ \nu_{\mathrm{inc}} \left(N_{\mathrm{track}} - 1 - m\right) + \nu_0, & \text{for } \left\lceil \frac{N_{\mathrm{track}}}{2} \right\rceil \le m < N_{\mathrm{track}} \end{cases} \tag{68}$$

where N_{track} is the total track time, ν_0 is the initial Doppler, and ν_{inc} describes the increase in the Doppler. This evolution is representative for, e.g., a fixed transmitter and fixed scatterers in the environment with only a mobile receiver. The receiver starts with an initial velocity and then has a constant acceleration until a peak velocity is reached; afterwards, the velocity linearly reduces back to the initial velocity.

Consider the following parameters: the block length $N = 120$, the pilot spacing $L = 3$, the signal-to-noise ratio (SNR) $\rho = P_h \gamma = 10$ dB, the fixed channel power gain $P_h = 1$, the total track time $N_{\mathrm{track}} = 201$, the effective correlation length of the channel $N' = 303$, and $\nu_0 = \frac{1}{2N'}$. Finally, ν_{inc} is chosen such that, exemplarily, $\nu_{\max}[m]$ reaches a peak value of $1/(2L) - 1/(4N')$, see (56).[12] With these parameters, we obtain the size of the LQS regions in time as shown in Fig. 3. It can be observed that the size of the actual FL-based LQS regions is always lower-bounded by the size of the approximate IL-based LQS regions. Furthermore, for a fixed channel power gain, the LQS times generally decrease with a decreasing Doppler. The reason is that for a small maximal Doppler the channel power gain is concentrated inside a small fraction of the samples of the PSD of the channel. Thus, the MSE is more sensitive to a mismatch in the maximal Doppler when the maximal Doppler is small. Finally, we observe strong variations in the LQS times around the middle of the track, where the peak value of $\nu_{\max}[m]$ is reached.

B. Exemplary Analysis of a Measured Channel

We now use the developed concepts for the non-stationarity analysis of a measured channel. First, we have to find an estimate of the time-dependent PSD of the channel sampled in the Doppler domain, i.e., $C_h^{(\Phi)}[m;k]$, from a single measurement run of the channel. We use the doubly underspread condition and the effectively finite correlation of the channel to perform a weighted averaging inside local regions, i.e., finite intervals. The application of the multitaper estimator [13], [30], [33] to our setting with frequency-flat fading yields the time-dependent PSD estimate

$$\hat{C}_h^{(\Phi)}[m;k] = \frac{1}{S} \sum_{s=0}^{S-1} \left| H^{(g_s)}[m;k] \right|^2 \tag{69}$$

for $k = -\lfloor N'/(2L) \rfloor, \ldots, \lceil N'/(2L) \rceil - 1$, and $\hat{C}_h^{(\Phi)}[m;k] = 0$ else. Furthermore, we have

$$H^{(g_s)}[m;k] = \sqrt{\frac{T_{\mathrm{m}}}{T}} \sum_{\check{m}=-\lfloor N_{\mathrm{w}}/2 \rfloor}^{\lceil N_{\mathrm{w}}/2 \rceil - 1} g_s^*[\check{m}] h((m+\check{m})T_{\mathrm{m}}) e^{-j2\pi L \frac{k\check{m}}{N'}} \tag{70}$$

where $h(mT_{\mathrm{m}})$ denotes a measured sample of the continuous-time channel. We choose the measurement samples to be the pilot symbols, i.e., $T_{\mathrm{m}} = LT$. This choice leads to the normalization factor $\sqrt{T_{\mathrm{m}}/T} = \sqrt{L}$, cf. (46). Denoting the minimum value of N'' by $N''_{\min}$, the window length in time N_{w} has to satisfy $N''_{\min} \le (N_{\mathrm{w}} - 1)L < N_{\mathrm{s}}$, i.e., the estimator has to capture the relevant correlation coefficients of the channel while being restricted to an interval of constant statistics.

The channel measurements used in the subsequent analysis were performed in an urban macrocell scenario in Ilmenau, Germany; they are also available for download [34]. For the sake of completeness, we shortly describe the considered scenario; a more elaborate presentation is given in [3]. We consider vertically polarized propagation at a carrier frequency of 2.505 GHz. We choose a uniform linear array at base station 1 at a height of 25 m as the transmitter and a uniform circular array (the lower one) of the mobile terminal on the mobile terminal reference track 9a-9b [3] as the receiver. The specific link consists of one antenna element of the uniform linear array at the base station and the antenna element at the mobile terminal oriented towards the right with respect to the direction of motion.

We first verify the doubly underspread condition $\Delta_{\nu,\max} \ll \nu_{\max} \ll 1$ to ensure a meaningful local averaging over time [2]. In [3], we have shown that in the considered scenario the doubly underspread condition is satisfied with $\Delta_{\nu,\max} \approx 0.03/L$ and $\nu_{\max} \approx 0.31/L$, or equivalently $N_{\mathrm{s}} \approx 32.89L$ and $N_{\mathrm{c}} \approx 3.26L$. For the windows $g_k[m]$ used in (69), we choose discrete prolate spheroidal sequences [35] with length $N_{\mathrm{w}} = 31$ and a time-halfbandwidth product of 2. We obtain a coherence time $T_{\mathrm{c}} = N_{\mathrm{c}} T = T_{\mathrm{m}}/(\nu_{\max} L) \approx 0.04$ s, a windowing time $N_{\mathrm{w}} T_{\mathrm{m}} = N_{\mathrm{w}} LT = 0.41$ s, and a stationarity time $T_{\mathrm{s}} = N_{\mathrm{s}} T = T/\Delta_{\nu,\max} \approx 0.43$ s. Therefore, the measurement data satisfies $N_{\mathrm{c}} \ll N_{\mathrm{s}}$, and we have $(N_{\mathrm{w}} - 1)L < N_{\mathrm{s}}$. The number of windows is $S = 2$. We preprocess the data by estimating a noise level in the time-delay domain of the channel; all channel weights below this threshold are discarded. Furthermore, we remove the very small time-dependent mean of the channel. It is estimated using 31 time instants and a

[12] Note that, strictly speaking, the choice of N' is only appropriate when $\nu_{\max}[m]$ is large.

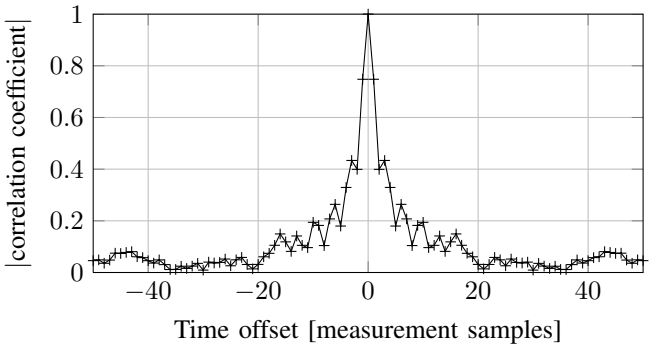

Fig. 4. Magnitude of the average correlation coefficients of the channel.

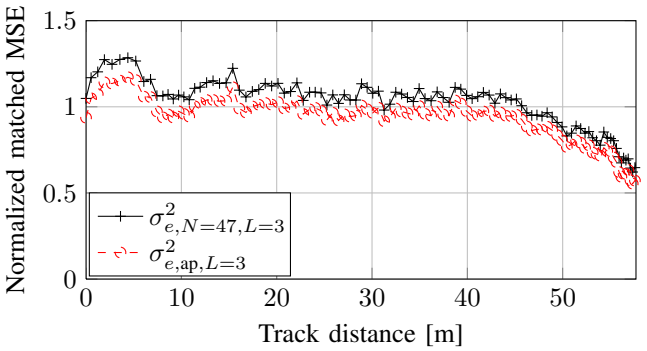

Fig. 5. Normalized matched MSEs for the average SNR $\rho_{\mathrm{avg}} = 10$ dB.

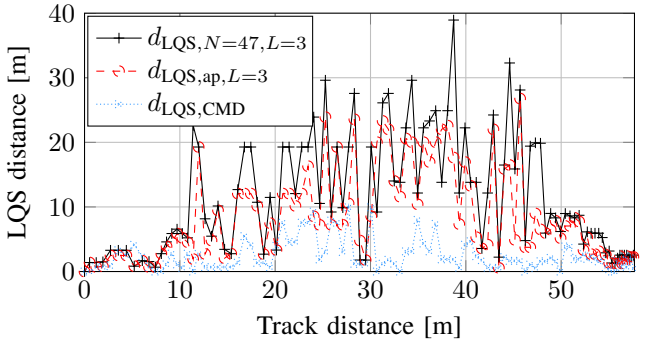

Fig. 6. LQS distances for the average SNR $\rho_{\mathrm{avg}} = 10$ dB and $\eta_{\mathrm{th}} = 0.1$.

frequency bandwidth of 5 MHz, i.e., 32 samples in frequency. Inside this bandwidth, the channel statistics can be assumed to be constant [3]. Then we use (69) to estimate the PSD and average the estimate over a bandwidth of 5 MHz to improve the estimation. No claims of optimality are made for these parameters. The magnitude of the correlation coefficients of the channel process averaged over the whole track are shown in Fig. 4. A strong decrease in the correlation coefficients to values below 0.1 can already be observed for a time offset of ± 20 measurement samples, i.e., pilot symbols. Therefore, the assumption on an effectively finite correlation of the channel process with $N'' \geq N''_{\mathrm{min}} = 20L$ is justified. Furthermore, $(N_{\mathrm{w}} - 1)L \geq N''_{\mathrm{min}} \gg N_{\mathrm{c}}$ holds.

In the following analysis, we choose the pilot spacing $L = 3$, the block length $N = 47$, the threshold $\eta_{\mathrm{th}} = 0.1$, and the average SNR $\rho_{\mathrm{avg}} = \gamma \sum_{m=1}^{M} |h(mT_{\mathrm{m}})|^2 / M = 10$ dB, where M is the total number of samples. The correlation length of the channel is chosen as $N' = 2N'' + 1 = 183$, i.e., $N'' = 91 \geq N''_{\mathrm{min}}$. Note that the estimated correlation coefficients for symbol offsets close to $(N_{\mathrm{w}} - 1)L = 90$ suffer from a reduced accuracy; we thus choose the block length $N = 47$. In Fig. 5, the actual FL-based matched MSE is compared to the approximate IL-based matched MSE; both are normalized to the mean value of the actual FL-based matched MSE over the track. One can see that here the approximate matched MSE lower-bounds the actual one, see Section IV-D. In Fig. 6, we show an exemplary evolution of the size of the LQS regions d_{LQS}, i.e., T_{LQS} mapped to the driven distance. From Fig. 6, we can see that both LQS regions, i.e., the actual one $d_{\mathrm{LQS},N=47,L=3}$ and the approximate one $d_{\mathrm{LQS,ap},L=3}$ show strong similarities in their evolution over the distance. As discussed in Section IV-E, the size of the actual FL-based LQS regions is lower-bounded by the size of the approximate IL-based LQS regions up to a few exceptions. Furthermore, we evaluate the LQS regions $d_{\mathrm{LQS,CMD}}$ based on the CMD, which is commonly used to analyze the spatial properties of the channel [14], [36]. To this end, we use the measure

$$\eta_{\mathrm{CMD}}[m, m'] = 1 - \frac{\mathrm{tr}\left\{ \mathbf{C}_h^{(\Phi)} \tilde{\mathbf{C}}_h^{(\Phi)} \right\}}{\left\| \mathbf{C}_h^{(\Phi)} \right\|_F \left\| \tilde{\mathbf{C}}_h^{(\Phi)} \right\|_F} \tag{71}$$

and the same threshold $\eta_{\mathrm{th}} = 0.1$. In this example, we observe significant differences between the LQS distances based on the CMD and those based on the MSE degradation.

VI. Conclusion

We have developed a methodology for the determination of LQS regions, i.e., local regions, for example in time, in which a channel can be treated as stationary. In contrast to previous results in the literature, we consider a realistic channel estimation algorithm. We relate the size of LQS regions to the degradation of the MSE of the channel estimate due to mismatched channel statistics, i.e., channel statistics from another time instant. The channel estimator is a finite-length Wiener filter estimating a time-varying frequency-flat fading channel. As the evaluation of the MSE turns out to be computationally complex, we also give an approximate expression based on infinite-length filtering. We have shown that for certain PSDs of the channel the approximate evaluation of the matched MSE based on infinite-length filtering is a lower bound to the actual matched MSE based on finite-length filtering. Moreover, for such PSDs of the channel, the actual MSE degradation is upper-bounded and the size of the actual LQS regions is lower-bounded by the approximate evaluation. Exemplarily, we have evaluated the actual and the approximate MSE degradation using channel measurements in an urban macrocell scenario. We have observed that the size of the resulting LQS regions based on the actual and the approximate MSE degradation show strong similarities in their evolution.

APPENDIX A
ALTERNATIVE DERIVATION OF THE INFINITE-LENGTH FILTERING MSE FOR WSS CHANNELS

In order to show (23) for a block length $N = N_p L$, we start by decomposing the correlation matrices as

$$\begin{aligned} \mathbf{R}_h &= \mathbf{F}_h^H \mathbf{\Lambda}_h \mathbf{F}_h \\ \mathbf{R}_{h_p} &= \mathbf{F}_{h_p}^H \mathbf{\Lambda}_{h_p} \mathbf{F}_{h_p} \; ; \qquad \tilde{\mathbf{R}}_{h_p} = \mathbf{F}_{h_p}^H \tilde{\mathbf{\Lambda}}_{h_p} \mathbf{F}_{h_p} \\ \mathbf{R}_{h;h_p} &= \mathbf{F}_h^H \mathbf{\Lambda}_h \mathbf{F}_{h,L} \; ; \qquad \tilde{\mathbf{R}}_{h;h_p} = \mathbf{F}_h^H \tilde{\mathbf{\Lambda}}_h \mathbf{F}_{h,L}. \end{aligned} \tag{72}$$

We use the $N \times N$ discrete Fourier transform (DFT) matrix $\mathbf{F}_h$ with $[\mathbf{F}_h]_{k,l} = \sqrt{1/N} \exp(-j2\pi(k-1)(l-1)/N)$ for $k, l = 1, \ldots, N$ and the $N_p \times N_p$ DFT matrix $\mathbf{F}_{h_p}$ with $[\mathbf{F}_{h_p}]_{k,l} = \sqrt{1/N_p} \exp(-j2\pi(k-1)(l-1)/N_p)$ for $k, l = 1, \ldots, N_p$. The $N \times N_p$ matrix $\mathbf{F}_{h,L}$ contains L-spaced columns of $\mathbf{F}_h$, i.e., $[\mathbf{F}_{h,L}]_{k,l} = \sqrt{1/N} \exp(-j2\pi(k-1)(l-1)L/N)$ for $k = 1, \ldots, N$ and $l = 1, \ldots, N_p$. The following property holds:

$$\mathbf{F}_{h,L} \mathbf{F}_{h_p}^H = \mathbf{1}_L \otimes \mathbf{I}_{N_p} \frac{1}{\sqrt{L}} \tag{73}$$

with the length-L column-vector $\mathbf{1}_L$ containing 1 in every entry. Note that $\mathbf{\Lambda}_h$ and $\tilde{\mathbf{\Lambda}}_h$ have dimension $N \times N$, and that $\mathbf{\Lambda}_{h_p}$ and $\tilde{\mathbf{\Lambda}}_{h_p}$ have dimension $N_p \times N_p$. With the decompositions in (72) and $\mathbf{F}_{h_p}^H \tilde{\mathbf{\Lambda}}_{h_p} \mathbf{F}_{h_p} + \mathbf{I}_{N_p} \gamma^{-1} = \mathbf{F}_{h_p}^H \left(\tilde{\mathbf{\Lambda}}_{h_p} + \mathbf{I}_{N_p} \gamma^{-1} \right) \mathbf{F}_{h_p}$, the mismatched MSE in (5) can be expressed as

$$\begin{aligned} \tilde{\mathbf{R}}_e &= \mathbf{F}_h^H \mathbf{\Lambda}_h \mathbf{F}_h + \mathbf{F}_h^H \tilde{\mathbf{\Lambda}}_h \mathbf{F}_{h,L} \mathbf{F}_{h_p}^H \left(\tilde{\mathbf{\Lambda}}_{h_p} + \mathbf{I}_{N_p} \gamma^{-1} \right)^{-1} \\ &\times \left(\mathbf{\Lambda}_{h_p} + \mathbf{I}_{N_p} \gamma^{-1} \right) \left(\tilde{\mathbf{\Lambda}}_{h_p} + \mathbf{I}_{N_p} \gamma^{-1} \right)^{-1} \mathbf{F}_{h_p} \mathbf{F}_{h,L}^H \tilde{\mathbf{\Lambda}}_h^H \mathbf{F}_h \\ &- \mathbf{F}_h^H \tilde{\mathbf{\Lambda}}_h \mathbf{F}_{h,L} \mathbf{F}_{h_p}^H \left(\tilde{\mathbf{\Lambda}}_{h_p} + \mathbf{I}_{N_p} \gamma^{-1} \right)^{-1} \mathbf{F}_{h_p} \mathbf{F}_{h,L}^H \mathbf{\Lambda}_h^H \mathbf{F}_h \\ &- \mathbf{F}_h^H \mathbf{\Lambda}_h \mathbf{F}_{h,L} \mathbf{F}_{h_p}^H \left(\tilde{\mathbf{\Lambda}}_{h_p} + \mathbf{I}_{N_p} \gamma^{-1} \right)^{-1} \mathbf{F}_{h_p} \mathbf{F}_{h,L}^H \tilde{\mathbf{\Lambda}}_h^H \mathbf{F}_h. \end{aligned} \tag{74}$$

Defining $\tilde{\mathbf{\Lambda}}_Z = \tilde{\mathbf{\Lambda}}_{h_p} + \mathbf{I}_{N_p} \gamma^{-1}$, $\mathbf{\Lambda}_Z = \mathbf{\Lambda}_{h_p} + \mathbf{I}_{N_p} \gamma^{-1}$, and

$$\begin{aligned} \tilde{\mathbf{Z}}_\Lambda &= \mathbf{\Lambda}_h + \frac{1}{L} \tilde{\mathbf{\Lambda}}_h \left(\left(\mathbf{1}_L \mathbf{1}_L^H \right) \otimes \left(\tilde{\mathbf{\Lambda}}_Z^{-1} \mathbf{\Lambda}_Z \tilde{\mathbf{\Lambda}}_Z^{-1} \right) \right) \tilde{\mathbf{\Lambda}}_h^H - \frac{1}{L} \tilde{\mathbf{\Lambda}}_h \\ &\times \left(\left(\mathbf{1}_L \mathbf{1}_L^H \right) \otimes \tilde{\mathbf{\Lambda}}_Z^{-1} \right) \mathbf{\Lambda}_h^H - \frac{1}{L} \mathbf{\Lambda}_h \left(\left(\mathbf{1}_L \mathbf{1}_L^H \right) \otimes \tilde{\mathbf{\Lambda}}_Z^{-1} \right) \tilde{\mathbf{\Lambda}}_h^H \end{aligned} \tag{75}$$

and using (73), the mismatched MSE matrix can be written as $\tilde{\mathbf{R}}_e = \mathbf{F}_h^H \tilde{\mathbf{Z}}_\Lambda \mathbf{F}_h$. The average mismatched MSE over all positions in (7) can thus be expressed as

$$\tilde{\sigma}_{e,N,L}^2 = \frac{1}{N} \operatorname{tr} \left\{ \mathbf{F}_h^H \tilde{\mathbf{Z}}_\Lambda \mathbf{F}_h \right\} = \frac{1}{N} \operatorname{tr} \left\{ \tilde{\mathbf{Z}}_\Lambda \right\}. \tag{76}$$

Autocorrelation matrices, which are Hermitian and Toeplitz for WSS processes, have an asymptotically equivalent[13] circulant matrix for absolutely summable autocorrelation functions [37, Lemma 4.1 & 4.6]. With the asymptotic equivalence of, e.g., $\tilde{\mathbf{R}}_h$ and the circulant matrix $\mathbf{F}_h^H \tilde{\mathbf{C}}_h \mathbf{F}_h$, we have

$$\lim_{N_p \to \infty} \frac{\| \tilde{\mathbf{R}}_{h_p} - \mathbf{F}_{h_p}^H \tilde{\mathbf{C}}_{h_p} \mathbf{F}_{h_p} \|_F}{\sqrt{N_p}} = \lim_{N_p \to \infty} \frac{\| \tilde{\mathbf{\Lambda}}_{h_p} - \tilde{\mathbf{C}}_{h_p} \|_F}{\sqrt{N_p}} = 0 \tag{77}$$

$$\lim_{N \to \infty} \frac{\| \tilde{\mathbf{R}}_h - \mathbf{F}_h^H \tilde{\mathbf{C}}_h \mathbf{F}_h \|_F}{\sqrt{N}} = \lim_{N \to \infty} \frac{\| \tilde{\mathbf{\Lambda}}_h - \tilde{\mathbf{C}}_h \|_F}{\sqrt{N}} = 0 \tag{78}$$

where the diagonal matrices $\tilde{\mathbf{C}}_{h_p}$ and $\tilde{\mathbf{C}}_h$ of size $N_p \times N_p$ and $N \times N$, respectively, are defined as

$$\left[\tilde{\mathbf{C}}_{h_p} \right]_{k,k} = \begin{cases} \tilde{C}_h \left(e^{j2\pi L \frac{k-1}{N}} \right), & \text{for } k = 1, \ldots, \left\lceil \frac{N_p}{2} \right\rceil \\ \tilde{C}_h \left(e^{j2\pi L \frac{k-1-N_p}{N}} \right), & \text{for } k = \left\lceil \frac{N_p}{2} \right\rceil + 1, \ldots, N_p \end{cases} \tag{79}$$

$$\left[\tilde{\mathbf{C}}_h \right]_{k,k} = \begin{cases} L\tilde{C}_h \left(e^{j2\pi L \frac{k-1}{N}} \right), & \text{for } k = 1, \ldots, \left\lceil \frac{N_p}{2} \right\rceil \\ L\tilde{C}_h \left(e^{j2\pi L \frac{k-1-N}{N}} \right), & \text{for } k = N - \left\lfloor \frac{N_p}{2} \right\rfloor + 1, \ldots, N \\ 0, & \text{else} \end{cases} \tag{80}$$

with $\tilde{C}_h \left(e^{j2\pi L\nu} \right) < \infty$. The asymptotic equivalence for the matched case is obtained accordingly.[14] We now define

$$\begin{aligned} \tilde{\mathbf{Z}}_C &= \mathbf{C}_h + \frac{1}{L} \tilde{\mathbf{C}}_h \left(\left(\mathbf{1}_L \mathbf{1}_L^H \right) \otimes \left(\tilde{\mathbf{C}}_Z^{-1} \mathbf{C}_Z \tilde{\mathbf{C}}_Z^{-1} \right) \right) \tilde{\mathbf{C}}_h^H - \frac{1}{L} \tilde{\mathbf{C}}_h \\ &\times \left(\left(\mathbf{1}_L \mathbf{1}_L^H \right) \otimes \tilde{\mathbf{C}}_Z^{-1} \right) \mathbf{C}_h^H - \frac{1}{L} \mathbf{C}_h \left(\left(\mathbf{1}_L \mathbf{1}_L^H \right) \otimes \tilde{\mathbf{C}}_Z^{-1} \right) \tilde{\mathbf{C}}_h^H \end{aligned} \tag{81}$$

with $\tilde{\mathbf{C}}_Z = \tilde{\mathbf{C}}_{h_p} + \mathbf{I}_{N_p} \gamma^{-1}$ and $\mathbf{C}_Z = \mathbf{C}_{h_p} + \mathbf{I}_{N_p} \gamma^{-1}$. Using [37, Theorem 2.1] with $\sigma_n^2 > 0$, it follows that $\tilde{\mathbf{\Lambda}}_Z^{-1}$ and $\tilde{\mathbf{C}}_Z^{-1}$ are asymptotically equivalent. With $\left\| \left(\mathbf{1}_L \mathbf{1}_L^H \right) \otimes \mathbf{A} \right\|_F = L \left\| \mathbf{A} \right\|_F$ for an arbitrary matrix $\mathbf{A}$, it follows that the matrices resulting from the Kronecker products in (75) are asymptotically equivalent to the corresponding ones in (81). Now, invoking again [37, Theorem 2.1], it follows that the matrices resulting from the matrix multiplications in (75) are asymptotically equivalent to the corresponding ones in (81). Thus, $\tilde{\mathbf{Z}}_C$ is asymptotically equivalent to $\tilde{\mathbf{Z}}_\Lambda$, and we have

$$\lim_{N \to \infty} \frac{\| \tilde{\mathbf{Z}}_\Lambda - \tilde{\mathbf{Z}}_C \|_F}{\sqrt{N}} = 0. \tag{82}$$

With [37, Theorem 2.2] and (82), the average mismatched MSE over all positions in (76) for $N \to \infty$ follows as

$$\lim_{N \to \infty} \tilde{\sigma}_{e,N,L}^2 = \lim_{N \to \infty} \frac{1}{N} \operatorname{tr} \left\{ \tilde{\mathbf{Z}}_\Lambda \right\} = \lim_{N \to \infty} \frac{1}{N} \operatorname{tr} \left\{ \tilde{\mathbf{Z}}_C \right\}. \tag{83}$$

We now rewrite the trace of (81) without Kronecker products. Instead, we will obtain a sum of products of diagonal matrices. Due to the zero entries and the factor L in (80), we can use

[13] A formal definition of asymptotically equivalent sequences of matrices is given in [37, Section 2.3].

[14] It is shown in [38] that (77) and (78) also hold for square summable autocorrelation functions and a different construction of the circulant matrices.

(81) with (79) and (80) to state

$$\begin{aligned}\operatorname{tr}\left\{\tilde{\mathbf{Z}}_C\right\} &= \operatorname{tr}\Big\{L\mathbf{C}_{h_p} + L\tilde{\mathbf{C}}_{h_p}\tilde{\mathbf{C}}_Z^{-1}\mathbf{C}_Z\tilde{\mathbf{C}}_Z^{-1}\tilde{\mathbf{C}}_{h_p}^H \\ &\quad -L\tilde{\mathbf{C}}_{h_p}\tilde{\mathbf{C}}_Z^{-1}\mathbf{C}_{h_p}^H - L\mathbf{C}_{h_p}\tilde{\mathbf{C}}_Z^{-1}\tilde{\mathbf{C}}_{h_p}^H\Big\} \\ &= \operatorname{tr}\Big\{L\mathbf{C}_{h_p} + L\tilde{\mathbf{C}}_{h_p}\left(\tilde{\mathbf{C}}_{h_p} + \mathbf{I}_{N_p}\gamma^{-1}\right)^{-1}\left(\mathbf{C}_{h_p} + \mathbf{I}_{N_p}\gamma^{-1}\right) \\ &\quad \times\left(\tilde{\mathbf{C}}_{h_p} + \mathbf{I}_{N_p}\gamma^{-1}\right)^{-1}\tilde{\mathbf{C}}_{h_p}^H - L\tilde{\mathbf{C}}_{h_p}\left(\tilde{\mathbf{C}}_{h_p} + \mathbf{I}_{N_p}\gamma^{-1}\right)^{-1}\mathbf{C}_{h_p}^H \\ &\quad -L\mathbf{C}_{h_p}\left(\tilde{\mathbf{C}}_{h_p} + \mathbf{I}_{N_p}\gamma^{-1}\right)^{-1}\tilde{\mathbf{C}}_{h_p}^H\Big\}. \end{aligned} \tag{84}$$

Finally, we evaluate (83) with (84), and we obtain

$$\begin{aligned}\lim_{N\to\infty}\tilde{\sigma}_{e,N,L}^2 = \lim_{N\to\infty}\frac{L}{N}\sum_{k=-\lfloor\frac{N}{2L}\rfloor}^{\lceil\frac{N}{2L}\rceil-1}\Bigg(&C_h\left(e^{j2\pi L\frac{k}{N}}\right) \\ &+\frac{\left(C_h\left(e^{j2\pi L\frac{k}{N}}\right)+\gamma^{-1}\right)\tilde{C}_h^2\left(e^{j2\pi L\frac{k}{N}}\right)}{\left(\tilde{C}_h\left(e^{j2\pi L\frac{k}{N}}\right)+\gamma^{-1}\right)^2} \\ &-\frac{2C_h\left(e^{j2\pi L\frac{k}{N}}\right)\tilde{C}_h\left(e^{j2\pi L\frac{k}{N}}\right)}{\tilde{C}_h\left(e^{j2\pi L\frac{k}{N}}\right)+\gamma^{-1}}\Bigg). \end{aligned} \tag{85}$$

Identifying (16) in (85), we can write

$$\begin{aligned}\lim_{N\to\infty}\tilde{\sigma}_{e,N,L}^2 &= \lim_{N\to\infty}\frac{L}{N}\sum_{k=-\lfloor\frac{N}{2L}\rfloor}^{\lceil\frac{N}{2L}\rceil-1}\tilde{C}_e\left(e^{j2\pi L\frac{k}{N}}\right) \\ &= L\int_{-\frac{1}{2L}}^{\frac{1}{2L}}\tilde{C}_e\left(e^{j2\pi L\nu}\right)d\nu = \tilde{\sigma}_{e,\infty,L}^2 \end{aligned} \tag{86}$$

for Riemann integrable error processes. The proof for matched Wiener filtering follows as a special case. In conclusion, we have proven (23).

Appendix B
Lower Bound on the Actual FL-Based Matched MSE for a Rectangular PSD of the Channel

We derive a lower bound on the actual FL-based *matched* MSE (52) for the special case of a *rectangular* PSD satisfying (56). We will show that this lower bound is equal to the approximate IL-based matched MSE (55). Remember that we have assumed a finite correlation of the error process in the derivation of the approximate MSE, see (53). Starting from (52), we obtain

$$\begin{aligned}\sigma_{e,N,L}^2[m] &= \frac{1}{N}\operatorname{tr}\Big\{\check{\mathbf{F}}_h^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_h - \check{\mathbf{F}}_h^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_{h,L} \\ &\quad\times\left(\check{\mathbf{F}}_{h,L}^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_{h,L} + \mathbf{I}_{N_p}\gamma^{-1}\right)^{-1}\check{\mathbf{F}}_{h,L}^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_h\Big\} \\ &\overset{(a)}{=} \frac{1}{N}\operatorname{tr}\Big\{\check{\mathbf{F}}_h^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_h - \left[\mathbf{C}_h^{(\Phi)}\right]_{1,1}^2\bar{\mathbf{F}}_h^H\bar{\mathbf{F}}_{h,L} \\ &\quad\times\left(\bar{\mathbf{F}}_{h,L}^H\bar{\mathbf{F}}_{h,L}\left[\mathbf{C}_h^{(\Phi)}\right]_{1,1} + \mathbf{I}_{N_p}\gamma^{-1}\right)^{-1}\bar{\mathbf{F}}_{h,L}^H\bar{\mathbf{F}}_h\Big\} \\ &\overset{(b)}{=} \frac{1}{N}\operatorname{tr}\Big\{\check{\mathbf{F}}_h^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_h - L\left[\mathbf{C}_h^{(\Phi)}\right]_{1,1}^2\bar{\mathbf{F}}_h^H\bar{\mathbf{F}}_{h,L} \\ &\quad\times\left(\mathbf{I}_{N_p}\left[\mathbf{C}_h^{(\Phi)}\right]_{1,1} + \mathbf{I}_{N_p}L\gamma^{-1}\right)^{-1}\bar{\mathbf{F}}_{h,L}^H\bar{\mathbf{F}}_h\Big\} \\ &\overset{(c)}{\geq} \frac{1}{N}\operatorname{tr}\Big\{\check{\mathbf{F}}_h^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_h - \frac{NL}{N_p}\left[\mathbf{C}_h^{(\Phi)}\right]_{1,1}^2\bar{\mathbf{F}}_{h,L}^H\bar{\mathbf{F}}_{h,L} \\ &\quad\times\left(\mathbf{I}_{N_p}\left[\mathbf{C}_h^{(\Phi)}\right]_{1,1} + \mathbf{I}_{N_p}L\gamma^{-1}\right)^{-1}\bar{\mathbf{F}}_{h,L}^H\bar{\mathbf{F}}_{h,L}\Big\} \\ &\overset{(d)}{=} \frac{1}{N}\operatorname{tr}\Big\{\check{\mathbf{F}}_h\check{\mathbf{F}}_h^H\mathbf{C}_h^{(\Phi)} \\ &\quad - \frac{N}{N_pL}\left[\mathbf{C}_h^{(\Phi)}\right]_{1,1}^2\left(\left[\mathbf{C}_h^{(\Phi)}\right]_{1,1} + L\gamma^{-1}\right)^{-1}\mathbf{I}_{N_p}\Big\} \\ &\overset{(e)}{=} \frac{1}{N'}\operatorname{tr}\left\{\mathbf{C}_h^{(\Phi)}\right\} - \frac{1}{L}\left[\mathbf{C}_h^{(\Phi)}\right]_{1,1}^2\left(\left[\mathbf{C}_h^{(\Phi)}\right]_{1,1} + L\gamma^{-1}\right)^{-1} \\ &\overset{(f)}{=} \frac{1}{N'}\sum_{k=-N''}^{N''}\frac{C_h^{(\Phi)}[m;k]}{\frac{\gamma}{L}C_h^{(\Phi)}[m;k]+1} = \sigma_{e,\mathrm{ap},L}^2[m]. \end{aligned} \tag{87}$$

We first state

$$\check{\mathbf{F}}_{h,L} = \mathbf{1}_L\otimes\bar{\mathbf{F}}_{h,L}\ ; \qquad \bar{\mathbf{F}}_{h,L}^H\bar{\mathbf{F}}_{h,L} = \mathbf{I}_{N_p}\frac{1}{L} \tag{88}$$

where the $\frac{N'}{L}\times N_p$ matrix $\bar{\mathbf{F}}_{h,L}$ is specified by the entries $\left[\bar{\mathbf{F}}_{h,L}\right]_{k,l} = \sqrt{1/N'}\exp(-j2\pi(k-1)(l-1)L/N')$ for $k = 1,\ldots,N'/L$ and $l = 1,\ldots,N_p$. In (a), we use (88) to obtain

$$\begin{aligned}\check{\mathbf{F}}_{h,L}^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_{h,L} &= \bar{\mathbf{F}}_{h,L}^H\bar{\mathbf{C}}_h\bar{\mathbf{F}}_{h,L} \\ \check{\mathbf{F}}_h^H\mathbf{C}_h^{(\Phi)}\check{\mathbf{F}}_{h,L} &= \bar{\mathbf{F}}_h^H\bar{\mathbf{C}}_h\bar{\mathbf{F}}_{h,L} \end{aligned} \tag{89}$$

where the $(N'/L)\times N$ matrix $\bar{\mathbf{F}}_h$ contains only the rows of $\check{\mathbf{F}}_h$ corresponding to the non-zero columns of $\mathbf{C}_h^{(\Phi)}$, and the $(N'/L)\times(N'/L)$ diagonal matrix $\bar{\mathbf{C}}_h$ contains only the non-zero rows and columns of $\mathbf{C}_h^{(\Phi)}$. Furthermore, in (a), we use $\bar{\mathbf{C}}_h = \mathbf{I}_{\frac{N'}{L}}\left[\mathbf{C}_h^{(\Phi)}\right]_{1,1}$ due to a rectangular PSD of the channel with (56). In (b), we use (88). In order to show (c), we define $\bar{\mathbf{f}}_{h,k}$ and $\bar{\mathbf{f}}_{h,L,k}$ to be the kth column of $\bar{\mathbf{F}}_h$ and $\bar{\mathbf{F}}_{h,L}$, respectively. Note that $\left\|\bar{\mathbf{f}}_{h,k}\right\|_F^2 = \left\|\bar{\mathbf{f}}_{h,L,k}\right\|_F^2 = 1/L$ holds and that all non-zero eigenvalues of $\bar{\mathbf{F}}_{h,L}\bar{\mathbf{F}}_{h,L}^H = \sum_{k=1}^{N_p}\bar{\mathbf{f}}_{h,L,k}\bar{\mathbf{f}}_{h,L,k}^H$ are equal to $1/L$. With the Rayleigh-Ritz theorem [39, Theorem 4.2.2], we have

$$\begin{aligned}\left\|\bar{\mathbf{F}}_{h,L}^H\mathbf{x}\right\|_F^2 = \mathbf{x}^H\bar{\mathbf{F}}_{h,L}\bar{\mathbf{F}}_{h,L}^H\mathbf{x} &\leq \frac{1}{L^2} = \bar{\mathbf{f}}_{h,L,k}^H\bar{\mathbf{F}}_{h,L}\bar{\mathbf{F}}_{h,L}^H\bar{\mathbf{f}}_{h,L,k} \\ &= \left\|\bar{\mathbf{F}}_{h,L}^H\bar{\mathbf{f}}_{h,L,k}\right\|_F^2,\ \forall k = 1,\ldots,N_p \end{aligned} \tag{90}$$

for an arbitrary vector $\mathbf{x}$ with $\|\mathbf{x}\|_F^2 = 1/L$. It follows that

$$\begin{aligned}\frac{1}{N}\operatorname{tr}\left\{\bar{\mathbf{F}}_h^H\bar{\mathbf{F}}_{h,L}\bar{\mathbf{F}}_{h,L}^H\bar{\mathbf{F}}_h\right\} &= \frac{1}{N}\sum_{k=1}^{N}\bar{\mathbf{f}}_{h,k}^H\bar{\mathbf{F}}_{h,L}\bar{\mathbf{F}}_{h,L}^H\bar{\mathbf{f}}_{h,k} \\ &\overset{(90)}{\leq} \bar{\mathbf{f}}_{h,L,1}^H\bar{\mathbf{F}}_{h,L}\bar{\mathbf{F}}_{h,L}^H\bar{\mathbf{f}}_{h,L,1} = \frac{1}{N_p}\operatorname{tr}\left\{\bar{\mathbf{F}}_{h,L}^H\bar{\mathbf{F}}_{h,L}\bar{\mathbf{F}}_{h,L}^H\bar{\mathbf{F}}_{h,L}\right\}. \end{aligned} \tag{91}$$

In (d) of (87), we use (88). In (e), we use (37) and the following equality, which holds for an arbitrary square matrix $\mathbf{A}$ and a diagonal matrix $\mathbf{D}$ of appropriate size:

$$\operatorname{tr}\{\mathbf{AD}\} = \operatorname{tr}\{\mathbf{A}\odot\mathbf{D}\}. \tag{92}$$

Finally, in (f), we reformulate the second term in (e) by making use of the fact that only N'/L values of $C_h^{(\Phi)}[m;k]$ are non-zero.

APPENDIX C
CHANNEL PSDS MINIMIZING AND MAXIMIZING THE APPROXIMATE IL-BASED MATCHED MSE

We show that a PSD with one peak minimizes and a rectangular PSD satisfying (56) maximizes the approximate IL-based matched MSE (55) subject to a fixed channel power gain $P_h = P_h[m] = \sum_{k=-N''}^{N''} C_h^{(\Phi)}[m;k]/N'$ and (2). We can rewrite the approximate IL-based matched MSE (55) by defining the concave and monotonically increasing function $f(x) = xL/(x\gamma + L), x \geq 0$:

$$\sigma^2_{e,\mathrm{ap},L}[m] = g(\mathbf{c}) = \frac{1}{N'} \sum_{k=-N''}^{N''} f\left(c[k]\right) \tag{93}$$

where we define the vector $\mathbf{c}$ with $[\mathbf{c}]_{k+N''+1} = c[k] = C_h^{(\Phi)}[m;k]$ for $k = -N'', \ldots, N''$ and we drop the time argument to simplify notation. As $f(x)$ is concave in x, $g(\mathbf{x})$ is also concave in $\mathbf{x}$ [40, Section 3.2.1].

To show which PSD minimizes $\sigma^2_{e,\mathrm{ap},L}[m]$, we define the vector $\mathbf{b}$ with $[\mathbf{b}]_{k+N''+1} = b[k] = c[k]/(P_h N')$ for $k = -N'', \ldots, N''$ and $\sum_{k=-N''}^{N''} b[k] = 1$. We can thus write $\mathbf{c} = P_h N'\ \mathbf{b}$. Therefore, we have

$$\begin{aligned} g(\mathbf{c}) = g\left(P_h N'\ \mathbf{b}\right) &\overset{(a)}{\geq} \sum_{k=-N''}^{N''} b[k]\ g\left(P_h N'\ \mathbf{1}_k\right) \\ &\overset{(b)}{=} \sum_{k=-N''}^{N''} b[k]\ \frac{1}{N'}\ f\left(P_h N'\right) = \frac{1}{N'}\ f\left(P_h N'\right) \end{aligned} \tag{94}$$

where $\mathbf{1}_k$ is a zero-vector except for a 1 in the kth entry, (a) follows from Jensen's inequality, and (b) from (93). Thus, a single peak with amplitude $P_h N'$ minimizes (93).

We now show which PSD maximizes $\sigma^2_{e,\mathrm{ap},L}[m]$. With the concavity of $f(x)$ and Jensen's inequality, it follows that

$$\begin{aligned} g(\mathbf{c}) &= \frac{1}{N'} \sum_{k=-N''}^{N''} f\left(c[k]\right) = \frac{1}{L}\frac{L}{N'} \sum_{k=-\left\lfloor \frac{N'}{2L} \right\rfloor}^{\left\lfloor \frac{N'}{2L} \right\rfloor} f\left(c[k]\right) \\ &\leq \frac{1}{L} f\left(\frac{L}{N'} \sum_{k=-\left\lfloor \frac{N'}{2L} \right\rfloor}^{\left\lfloor \frac{N'}{2L} \right\rfloor} c[k] \right) = \frac{1}{N'} \sum_{k=-\left\lfloor \frac{N'}{2L} \right\rfloor}^{\left\lfloor \frac{N'}{2L} \right\rfloor} f\left(LP_h\right) \end{aligned} \tag{95}$$

holds since only N'/L elements of $\mathbf{c}$ are allowed to be non-zero. Therefore, we can state that (93) is maximized by choosing $c[k] = LP_h, k = -\lfloor N'/(2L) \rfloor, \ldots, \lfloor N'/(2L) \rfloor$ and $c[k] = 0$ else.

REFERENCES

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

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

[3] A. Ispas, C. Schneider, G. Ascheid, and R. Thomä, "Analysis of local quasi-stationarity regions in an urban macrocell scenario," in *Proc. 71st IEEE Veh. Technol. Conf. (VTC)*, Taipei, Taiwan, May 2010.

[4] L. Bernadó, T. Zemen, A. Paier, G. Matz, J. Karedal, N. Czink, C. Dumard, F. Tufvesson, M. Hagenauer, A. F. Molisch, and C. F. Mecklenbräuker, "Non-WSSUS vehicular channel characterization at 5.2 GHz - Spectral divergence and time-variant coherence parameters," in *Proc. XXIXth General Assembly Scientific Symp. Int. Union Radio Science (URSI)*, Chicago, IL, USA, Aug. 2008.

[5] L. Bernadó, T. Zemen, A. Paier, and J. Karedal, "Complexity reduction for vehicular channel estimation using the filter divergence measure," in *Proc. 44th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 2010.

[6] M. B. Priestley and T. S. Rao, "A test for non-stationarity of time-series," *J. Roy. Stat. Soc. B (Methodologic.)*, vol. 31, no. 1, pp. 140–149, 1969.

[7] S. Kay, "A new nonstationarity detector," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1440–1451, Apr. 2008.

[8] P. Borgnat, P. Flandrin, P. Honeine, C. Richard, and J. Xiao, "Testing stationarity with surrogates: A time-frequency approach," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3459–3470, Jul. 2010.

[9] T. J. Willink, "Wide-sense stationarity of mobile MIMO radio channels," *IEEE Trans. Veh. Technol.*, vol. 57, no. 2, pp. 704–714, Mar. 2008.

[10] J. W. Wallace and M. A. Jensen, "Time-varying MIMO channels: Measurement, analysis, and modeling," *IEEE Trans. Antennas Propag.*, vol. 54, no. 11, pp. 3265–3273, Nov. 2006.

[11] W. Martin, "Measuring the degree of non-stationarity by using the Wigner-Ville spectrum," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, San Diego, CA, USA, Mar. 1984, pp. 262–265.

[12] L. Dossi, G. Tartara, and F. Tallone, "Statistical analysis of measured impulse response functions of 2.0 GHz indoor radio channels," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 3, pp. 405–410, Apr. 1996.

[13] 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. Int. ITG Workshop Smart Antennas (WSA)*, Darmstadt, Germany, Feb. 2008.

[14] M. Herdin, "Non-stationary indoor MIMO radio channels," Ph.D. dissertation, Vienna Univ. of Technology, Vienna, Austria, 2004.

[15] S. Verdú, "Mismatched estimation and relative entropy," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3712–3720, Aug. 2010.

[16] T. Weissman, "The relationship between causal and noncausal mismatched estimation in continuous-time AWGN channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4256–4273, Sep. 2010.

[17] T. T. Georgiou, "Distances and Riemannian metrics for spectral density functions," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 3995–4003, Aug. 2007.

[18] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.

[19] G. Durisi, U. G. Schuster, H. Bölcskei, and S. Shamai (Shitz), "Noncoherent capacity of underspread fading channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 367–395, Jan. 2010.

[20] W. Gardner, "A sampling theorem for nonstationary random processes," *IEEE Trans. Inf. Theory*, vol. 18, no. 6, pp. 808–809, Nov. 1972.

[21] G. Matz and F. Hlawatsch, "Time-varying communication channels: Fundamentals, recent developments, and open problems," in *Proc. 14th European Signal Process. Conf. (EUSIPCO)*, Florence, Italy, Sep. 2006.

[22] R. M. Gray and L. D. Davisson, *An Introduction to Statistical Signal Processing*. Cambridge, UK: Cambridge Univ. Press, 2004.

[23] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.

[24] S. Kaiser, "Multi-carrier CDMA mobile radio systems - Analysis and optimization of detection, decoding, and channel estimation," Ph.D. dissertation, TU Kaiserslautern, Kaiserslautern, Germany, 1998.

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

[26] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*. Upper Saddle River, NJ, USA: Prentice Hall, 1999.

[27] J. Baltersee, G. Fock, and H. Meyr, "An information theoretic foundation of synchronized detection," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2115–2123, Dec. 2001.

[28] T. Michaeli and Y. C. Eldar, "Hidden relationships: Bayesian estimation with partial knowledge," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 1933–1948, May 2011.

[29] D. Slepian, "On bandwidth," *Proc. IEEE*, vol. 64, no. 3, pp. 292–300, Mar. 1976.

[30] G. Matz, "Doubly underspread non-WSSUS channels: Analysis and estimation of channel statistics," in *Proc. 4th IEEE Workshop Signal*

Process. Advances Wireless Commun. (SPAWC), Rome, Italy, Jun. 2003, pp. 190–194.
[31] M. Dörpinghaus, A. Ispas, and H. Meyr, "On the gain of joint processing of pilot and data symbols in stationary Rayleigh fading channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2963–2982, May 2012.
[32] R. Silverman, "Locally stationary random processes," *IRE Trans. Inf. Theory*, vol. 3, no. 3, pp. 182–187, Sep. 1957.
[33] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proc. IEEE*, vol. 70, no. 9, pp. 1055–1096, Sep. 1982.
[34] Institute of Information Technology, Ilmenau Univ. of Technology. (2009, Apr.) Ilmenau reference scenario. [Online]. Available: http://www-emt.tu-ilmenau.de/ReferenceScenario/
[35] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty - V: The discrete case," *Bell System Technical Journal*, vol. 57, no. 5, pp. 1371–1430, Jun. 1978.
[36] 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.
[37] R. M. Gray, "Toeplitz and circulant matrices: A review," in *Foundations and Trends in Communications and Information Theory*. Delft, The Netherlands: now Publishers, 2006, vol. 2, no. 3, pp. 155–239.
[38] J. Pearl, "On coding and filtering stationary signals by discrete Fourier transforms," *IEEE Trans. Inf. Theory*, vol. 19, no. 2, pp. 229–232, Mar. 1973.
[39] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge Univ. Press, 1990.
[40] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge Univ. Press, 2004.

Adrian Ispas (S'11) obtained the Dipl.-Ing. degree in electrical engineering and information technology (with distinction) from RWTH Aachen University, Germany, in 2007. He is currently working as a research assistant and pursuing a doctorate degree at the Chair for Integrated Signal Processing Systems, RWTH Aachen University. His main research interests lie in the areas of channel modeling and characterization and statistical signal processing.

Adrian Ispas received the Henry Ford II Award 2005 for academic achievements during his studies and the SEW-EURODRIVE-Stiftung Award 2008 in recognition of his diploma thesis.

Meik Dörpinghaus (S'04-M'10) received the Dipl.-Ing. degree in electrical engineering and information technology (with distinction) and the Dr.-Ing. degree in electrical engineering and information technology (summa cum laude) both from RWTH Aachen University, Aachen, Germany, in 2003 and 2010, respectively. Currently, he is a postdoctoral researcher at RWTH Aachen University. In 2007, he was a visiting researcher at ETH Zurich, Switzerland. His research interests are in the areas of communication and information theory.

Dr. Dörpinghaus received the Friedrich-Wilhelm Preis of RWTH Aachen in 2011 for an outstanding Ph.D. thesis, and the Friedrich-Wilhelm Preis of RWTH Aachen in 2004 and the Siemens Preis in 2004 for an excellent diploma thesis.

Gerd Ascheid (SM'08) received his Diploma and PhD degrees in Electrical Engineering (Communications Eng.) from RWTH Aachen University. In 1988 he started as a co-founder CADIS GmbH which successfully brought the system simulation tool COSSAP to the market. From 1994-2003 Gerd Ascheid was Director / Senior Director with Synopsys, a California-based EDA market leader. In 2002 he was co-founder of LisaTek whose processor design tools are now part of the Synopsys product portfolio. Since April 2003 Gerd Ascheid heads the Institute for Integrated Signal Processing Systems of RWTH Aachen University. He is also coordinator of the UMIC (Ultra-high speed Mobile Information and Communication) Research Centre at RWTH Aachen University. His research interest is in wireless communication algorithms and application specific integrated platforms, in particular, for mobile terminals.

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

From 1998 to 2003 he worked as Hardware Engineer and Project Manager for the Radio Communication Devices Department, Siemens Austria. Since October 2003 Thomas Zemen has been with FTW Telecommunications Research Center Vienna, he leads the department "Signal and Information Processing" 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 (FWF). Thomas Zemen is the author or coauthor of four book chapters, 18 journal papers and more than 65 conference communications. His research interests include vehicular channel measurements and modelling, time-variant channel estimation, iterative multiple-input multiple-output (MIMO) receiver structures, 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.