Maximally Robust 2-D Channel Estimation for OFDM Systems

Muhammad Danish Nisar, *Student Member, IEEE*, Wolfgang Utschick, *Senior Member, IEEE*, and Thomas Hindelang, *Member, IEEE*

***Abstract*—Two-dimensional minimum mean square error (MMSE) channel estimation for orthogonal frequency division multiplexing (OFDM) systems is known to perform better than the least squares, as well as the conventional 1-D MMSE estimation, owing to its ability of exploiting both, the time and the frequency correlations among the channel frequency response (CFR) coefficients. Its superior performance comes however at the price of increased requirements—the knowledge of observation noise power and that of the channel frequency, as well as time correlation functions. In practical transmission scenarios, where channel correlation functions are not known or cannot be easily estimated, it is desirable to have an estimator that is robust to mismatches between the assumed and the actual channel correlation function. While such a robust estimator, for the case of an infinite number of observations, is well known for various uncertainty classes, not much attention has been paid to the practical case of a finite number of observations. We derive in this paper, the maximally robust (MR) 2-D channel estimator for the case of a finite number of pilot observations under some realistic constraints on the uncertainty class to which the 2-D channel correlation sequence belongs. We demonstrate that the correlation sequence associated with the MR estimator can be obtained by a simple semi-definite optimization procedure and is interestingly different from the well-known heuristic proposals. Simulation results establish the superiority of the proposed MR estimator over commonly employed heuristic robust estimator by as much as 3 dB in terms of the worst-case estimation MSE and around 1 dB in terms of the average BER performance under different practical transmission scenarios of interest.**

***Index Terms*—Channel estimation, minimax optimization, orthogonal frequency division multiplexing (OFDM) systems, multicarrier systems, robust signal processing, worst-case estimation.**

I. Introduction

PILOT-AIDED channel estimation for wireless orthogonal frequency division multiplexing (OFDM) systems dates back to some of the early contributions from Edfors *et al.* [1], [2], where the authors proposed the use of minimum mean square error (MMSE) filter and its low rank approximation as an efficient low complexity channel estimation technique. This approach of exploiting the frequency correlations via MMSE filtering was later extended to the 2-D estimation techniques, whereby correlations across both frequency and time were exploited to yield superior performances [3].

In their seminal paper [4], Li *et al.* made an extensive performance analysis of the 2-D MMSE channel estimation for OFDM systems and especially studied the effect of model mismatches, i.e., imperfect knowledge of channel frequency and time correlations—a scenario not uncommon in mobile communication environments. Their analysis for mean square error (MSE) degradation in such cases led to the proposal of a robust 2-D MMSE channel estimator, highly insensitive to such model mismatches.

Precisely speaking, [4] proposed the use of a time correlation function corresponding to the ideal band limited (rectangular Doppler) spectrum. However, an inherent assumption in their analysis was the availability of an infinite number of time observations. Obviously, the same could not hold true for the number of observations along the frequency direction due to the limited number of subcarriers, so only an approximate analysis was pursued for the frequency correlation function, suggesting finally the use of a uniform power delay profile (PDP) based correlation function as a robust choice. The authors eventually relaxed the assumption of an infinite number of time observations as well, and heuristically proposed to employ the truncated sinc-based time correlation function. In a nutshell, the recommendation in [4] and [5] to use the rectangular channel correlation spectrum, along both the time and the frequency direction for robust channel estimation performance resulted, at least for the case of a finite number of observations, from an approximate and heuristic analysis.

In this paper, we try to arrive at the maximally robust (MR) MSE-based channel estimator with *a finite number of pilot observations* in a typical OFDM system as described in Section II. A description of the problem under consideration follows in Section III, where besides formulating the associated mathematical preliminaries, we intuitively present the problem as an optimization over a set of possible channel correlation sequences. In Section IV, we characterize this set of channel correlation sequences (the uncertainty class) in terms of practical transmission parameters and prove a property that is vital for subsequent analysis. Finally, in Section V, unlike the approach in [4] and [5], we pursue the minimax optimization approach and demonstrate that being constrained with a *finite* number of observations along time and frequency, the sinc-based heuristic correlation function is no longer the MR choice. This is in contrast to the minimax optimization results for the case of an *infinite* number

Manuscript received June 03, 2009; accepted January 13, 2010. Date of publication February 17, 2010; date of current version May 14, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Milica Stojanovic. The work of M. Danish Nisar was supported by Nokia Siemens Networks GmbH & Co. KG, Munich, Germany.

M. D. Nisar and W. Utschick are with the Associate Institute for Signal Processing (MSV), Technische Universität München (TUM), 80290, Munich, Germany (e-mail: mdanishnisar@ieee.org; utschick@tum.de).

T. Hindelang was with Nokia Siemens Networks GmbH & Co. KG, Munich, Germany. He is now with the European Patent Office (EPO), Munich, Germany. (e-mail: hindelang@ieee.org).

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

Digital Object Identifier 10.1109/TSP.2010.2043126

of observations as presented, for instance, in the comprehensive survey papers [6], [7].

It turns out, that the minimax problem of finding the MR channel estimator for pilot-aided OFDM systems can be reduced to a semidefinite optimization problem in terms of the pilot grid parameters, namely the pilot spacings in time and frequency directions, and the estimation parameters, namely the number of observations along time and frequency to be used for estimation. We conclude the paper in Section VI with a comprehensive analysis and discussion on the performance of the proposed MR estimator against that of the heuristic estimator [4] and the reference 2-D MMSE estimator with perfect knowledge of the channel correlation spectrum. Note that, other than the brief numerical example in Section IV-A and the simulation example in Section VI-B, we keep the discussion general enough to be applicable to any orthogonal multicarrier system.

II. System Model

We consider a typical OFDM system (see, for instance, [8, Ch. 12]) where the transmission takes place over a fading wireless multipath channel. Under the typical assumptions of cyclic prefix (CP) being longer than the channel impulse response (CIR) length and the channel coherence time being sufficiently larger than the OFDM symbol duration, the effective channel matrix after CP removal at the receiver, happens to be circulant in the time domain or equivalently diagonal in the frequency domain. As such, there exists a scalar relationship between the transmitted and the received frequency domain samples—a scalar multiplication with the channel frequency response (CFR) coefficient at each subcarrier and corruption by the additive Gaussian noise. The system model can therefore be simply written in the scalar form as under

$$Y_{f,t} = H_{f,t}X_{f,t} + \zeta_{f,t} \tag{1}$$

where $X_{f,t}$ and $Y_{f,t}$ denote, respectively, the transmitted and received frequency domain samples, while $H_{f,t}$ denotes the CFR coefficient and $\zeta_{f,t}$ the Gaussian noise. Note that the subscript f denotes the frequency index that runs over all the N subcarriers of the system and the subscript t denotes the time or OFDM symbol index. This simple frequency domain system model is often visualized as a 2-D grid of symbols with time along horizontal and frequency along vertical direction. In order to allow for pilot-aided channel estimation at receiver, almost all contemporary OFDM-based systems, such as WiMax [9] and 3G LTE [10], recommend the transmission of intermittent pilot symbols along both the time and the frequency directions. Optimal design and placement of these pilot symbols has been an active area of research during the past decade [11], [12]. In this paper, we consider the transmission of arbitrary pilot symbols placed at regular intervals forming a pilot grid with a constant frequency and time spacing of Δ_{F} and Δ_{T}, respectively. Furthermore, for notational convenience, we focus on aligned (nonstaggered) pilot grid that matches for instance to the LTE uplink specifications [10] by setting $\Delta_{\mathrm{F}} = 1$ (no spacing) and $\Delta_{\mathrm{T}} = 7$. It is worth mentioning here, that the assumption of nonstaggered grids is made just to simplify the presentation of the paper. The underlying principle of transforming from the minimax to maximin problem and then arriving at a convex optimization problem for finding the least-favorable (LF) correlation spectra stays essentially the same even in the case of staggered pilot grids [13].

Two parameters of interest for subsequent discussion are the channel's delay and Doppler spread. The delay spread measures the duration of the channel impulse response so that if τ denotes the CIR length in seconds, the CIR length in taps can be obtained as $L = \lceil \tau/T_{\mathrm{s}} \rceil$, where $\lceil \bullet \rceil$ is the operator for rounding to the nearest larger integer, and $T_{\mathrm{s}} = T_{\mathrm{b}}/(N + \nu)$ denotes the sampling interval, with T_{b} being the duration of one OFDM symbol and ν being the number of CP samples.

The Doppler spread, on the other hand, measures the channel variations along time, often characterized in terms of the Doppler power spectrum [8, Ch. 3]. An overview and some recent results on the statistical modeling of time varying channels can be found in [14] and [15]. The highest frequency for which the channel's Doppler power spectrum is nonzero is labeled here as $f_{\mathrm{d,max}}$. In the simple scenario, where the transmitter as well as the scattering environment is stationary, while the receiver is moving with a velocity v, the maximum Doppler spread can be analytically expressed as $f_{\mathrm{d,max}} = (v/c)f_{\mathrm{c}}$, where f_{c} denotes the carrier frequency and $c = 3 \times 10^8$ m/s is the velocity of wave propagation.

Notation: The operators $\mathrm{E}[\bullet], |\bullet|^2, (\bullet)^{\mathrm{T}}, (\bullet)^*, (\bullet)^{\mathrm{H}}, \mathrm{vec}(\bullet)$ stand for expectation, absolute value square, transpose, complex conjugate, Hermitian, and vectorization (vertically stacking the columns of a matrix below each other), respectively. The operator $|\bullet|$ with a set as its argument denotes the cardinality of set. $\mathbb{C}$ denotes the complex plane. The set of $Q \times Q$ positive semidefinite matrices is denoted by $\mathbb{P}^Q$, while $\mathbb{T}^Q$ denotes the set of Toeplitz positive semidefinite matrices of dimension Q.

III. Problem Formulation

The problem that we address in this paper is to find a finite pilot observation-based MR 2-D MMSE channel estimator, where robustness is desired against mismatches between the assumed and the actual channel correlation function. Since the problem of 2-D MMSE channel estimation at data positions from the available pilot observations decouples, in the context of a system model like (1), to the straightforward least squares (LS) channel estimation at pilot positions followed by a 2-D MMSE channel interpolation to data positions, we focus only on the second step (the MMSE interpolation) that actually requires the knowledge of the channel correlation function. Thus, intuitively speaking, after pursuing LS estimation of CFR coefficients at pilot positions, we look for a 2-D linear filter that performs an optimum weighted combination of the available pilot channel estimates to form an estimate of the channel at the data position of interest in presence of an uncertainty about the channel correlation function.[1]

As aforementioned, we restrict ourselves to the case of a finite number of observations (pilot channel estimates) to be acted upon by the interpolation filter. The motivation comes not only from the availability of limited pilot observations in practice,

[1]Although interpolation is only a special case of estimation, as we will see in Section V, the underlying idea of this paper—finding the LF 2-D channel correlation sequence—can also be obviously applied to the problems of finite dimensional robust MMSE channel smoothing and channel prediction as well.

but also from the limited computational budget of a practical receiver. Let K_{T} and K_{F} denote the number of pilot channel estimates, on each side of the data position of interest, to be employed for interpolation, respectively, along the time and the frequency directions. Thus, the relevant LS pilot channel estimates can be expressed as

$$\tilde{H}_{f,t} = H_{f,t} + \eta_{f,t} \quad \text{for } (f,t) \in \mathcal{P} \tag{2}$$

with $\mathcal{P}$ being the set of frequency and time indexes of pilot positions to be employed in the interpolation at the data position of interest, so that $|\mathcal{P}| = N_{\mathrm{T}}N_{\mathrm{F}}$ where N_{T} and N_{F} denote the total number of pilot CFRs, respectively, along the time and frequency directions falling inside the interpolation window[2]. To simplify the notations we stack the relevant pilot channel estimates (observations), cf. (2), into a $N_{\mathrm{T}}N_{\mathrm{F}}$ dimensional vector $\tilde{\boldsymbol{h}}_{\mathrm{p}}$, i.e., we have

$$\tilde{\boldsymbol{h}}_{\mathrm{p}} = \boldsymbol{h}_{\mathrm{p}} + \eta \tag{3}$$

where $\eta \in \mathbb{C}^{N_{\mathrm{T}}N_{\mathrm{F}}}$ denotes the pilot channel estimation error (observation noise) vector. Now let $\boldsymbol{w}^{\mathrm{H}} \in \mathbb{C}^{1\times N_{\mathrm{T}}N_{\mathrm{F}}}$ denote the vector containing the 2-D filter coefficients, so that the CFR estimate at the desired data positions $\mathcal{D}$ can be obtained as

$$\hat{H}_{f,t} = \boldsymbol{w}^{\mathrm{H}}\tilde{\boldsymbol{h}}_{\mathrm{p}} \quad \text{for } (f,t) \in \mathcal{D}. \tag{4}$$

The corresponding estimation error can be expressed as a function of the estimator $\boldsymbol{w}$ and the 2-D channel correlation sequence $\{r_H(i,j)\}$, where the scalar $r_H(i,j) = \mathrm{E}\left[H_{f,t}H^*_{f+i,t+j}\right]$ represents the CFR correlation value at frequency offset i and time offset j. Note that we assume the wide sense stationarity of the random process $H_{f,t}$ throughout this paper so that the correlation function is independent of the indexes f and t. The estimation error $\mathrm{E}\left[|H_{f,t} - \hat{H}_{f,t}|^2\right]$ reads as

$$\varepsilon(\boldsymbol{w}, \{r_H(i,j)\}) = r_H(0,0) + \boldsymbol{w}^{\mathrm{H}}\left(\boldsymbol{R}_{h_{\mathrm{p}}} + \boldsymbol{R}_{\eta}\right)\boldsymbol{w} - \boldsymbol{w}^{\mathrm{H}}\boldsymbol{r}_{h_{\mathrm{p}}} - \boldsymbol{r}^{\mathrm{H}}_{h_{\mathrm{p}}}\boldsymbol{w} \tag{5}$$

where $\boldsymbol{R}_{\eta}$ denotes the observation noise covariance matrix, while $\boldsymbol{R}_{h_{\mathrm{p}}} = \mathrm{E}[\boldsymbol{h}_{\mathrm{p}}\boldsymbol{h}^{\mathrm{H}}_{\mathrm{p}}] \in \mathbb{P}^{N_{\mathrm{T}}N_{\mathrm{F}}}$ denotes the pilot CFR correlation matrix and the vector $\boldsymbol{r}_{h_{\mathrm{p}}} = \mathrm{E}[\boldsymbol{h}_{\mathrm{p}}H^*_{f,t}] \in \mathbb{C}^{N_{\mathrm{T}}N_{\mathrm{F}}}$ denotes the correlation vector between pilot CFRs and the CFR at the current data position of interest. Both $\boldsymbol{R}_{h_{\mathrm{p}}}$ and $\boldsymbol{r}_{h_{\mathrm{p}}}$ are completely specified by the 2-D channel correlation sequence $\{r_H(i,j)\}$ for frequency offsets $i = -i_0,\ldots,i_0$ with $i_0 = 2K_{\mathrm{F}}\Delta_{\mathrm{F}}$ and time offsets $j = -j_0,\ldots,j_0$ with $j_0 = 2K_{\mathrm{T}}\Delta_{\mathrm{T}}$. Finally, minimization of the MSE in (5) w.r.t. the filter coefficients yields the well-known MMSE solution

$$\boldsymbol{w}_{\mathrm{MMSE}} = \left(\boldsymbol{R}_{h_{\mathrm{p}}} + \boldsymbol{R}_{\eta}\right)^{-1}\boldsymbol{r}_{h_{\mathrm{p}}} \tag{6}$$

[2]Given the definition of K_{T} above, N_{T} is $2K_{\mathrm{T}}+1$ or $2K_{\mathrm{T}}$ depending upon whether the current CFR to be estimated lies in the OFDM symbol containing pilots or not. N_{F} follows similarly.

leading to the minimum MSE that can be expressed (by substituting $\boldsymbol{w} = \boldsymbol{w}_{\mathrm{MMSE}}$ in (5)) as a function of the 2-D channel correlation sequence

$$\varepsilon_{\mathrm{MMSE}}(\{r_H(i,j)\}) = r_H(0,0) - \boldsymbol{r}^{\mathrm{H}}_{h_{\mathrm{p}}}\left(\boldsymbol{R}_{h_{\mathrm{p}}} + \boldsymbol{R}_{\eta}\right)^{-1}\boldsymbol{r}_{h_{\mathrm{p}}}. \tag{7}$$

Now a conceptual version of the problem at hand is as follows. Lacking perfect knowledge of the 2-D channel correlation sequence $\{r_H(i,j)\}$, we would like to design the MR filter that leads to the minimum MSE in the worst case scenario (worst-possible mismatch between the assumed and actual correlation sequence). To this end, we aim to minimize (via MMSE filter) the maximum MSE (because of the worst-possible mismatch). We show in Section V that this minimax optimization problem can be recasted into an equivalent maximin problem, whereby we try to maximize the minimum MSE (attained by the MMSE filter) through the selection of the worst-possible mismatch. This in fact corresponds to the problem of finding the LF 2-D correlation sequence and then determining the corresponding 2-D MMSE filter coefficients.

Note that, given no other constraints, finding the LF 2-D correlation sequence is rather trivial, since zero correlation among the neighboring CFR coefficients, i.e., $r_H(i,j) = 0$ for $(i,j) \neq (0,0)$, is indeed the LF scenario and leads to the maximum MSE. But as discussed in the next section, such a correlation sequence may not be a valid correlation sequence for the transmission scenarios of interest.

IV. Characterizing the Uncertainty Class of Correlation Spectrum

In this section, we show that the underlying phenomena for CFR variations along frequency and time directions limit the set of correlation sequences in such a way that they have a bandlimited spectrum for both, the correlation along time as well as the correlation along frequency direction. This provides us with a natural way of characterizing the uncertainty class—the set to which the unknown correlation sequence belongs.

A. Bandlimitedness of Correlation Spectrum

1) Correlation Along Time: Since Doppler spread underlies the evolution of the CFR coefficients along time from one block to another, the spectrum of channel temporal correlation function is bandlimited by the transmission's highest Doppler spread $f_{\mathrm{d,max}}$. If T_{b} denotes the time separation between consecutive transmission blocks and $f_{\mathrm{b}} = 1/T_{\mathrm{b}}$, then the normalized maximum angular frequency component in the spectrum of the time correlation function can be given as

$$\omega_{\mathrm{t,max}} = 2\pi\frac{f_{\mathrm{d,max}}}{f_{\mathrm{b}}} = 2\pi\frac{v}{c}\frac{f_{\mathrm{c}}}{f_{\mathrm{b}}} \tag{8}$$

so that as long as $f_{\mathrm{d,max}}$ is less than $f_{\mathrm{b}}/2$—its highest value under the Nyquist theorem—the spectrum of the time correlation function is bandlimited to $\omega_{\mathrm{t,max}} < \pi$.

2) Correlation Along Frequency: Under the Wiener Khinchin theorem [16, pp. 406–409], the channel frequency correlation spectrum can be obtained by averaging the magnitude squared Fourier transform of different CFR realizations. To this end, we note that the channel frequency correlation spectrum is one sided, has a shape that corresponds to the channel power delay profile, and its bandwidth therefore has a direct relationship with the CIR length. Thus, the spectrum of the frequency correlation function is bandlimited by the normalized angular frequency[3]

$$\omega_{\mathrm{f,max}} = 2\pi \frac{L}{N} = 2\pi \tau_{\max} \frac{f_{\mathrm{s}}}{N}. \tag{9}$$

Practical Example. In order to have a numerical idea, we see that for LTE Uplink specifications [10] ($f_{\mathrm{c}} = 2.6$ GHz, $f_{\mathrm{b}} \simeq 14$ Hz), we have a bandlimited time correlation spectrum (assuming a stationary environment) as long as the terminal velocity $v < (c/2)(f_{\mathrm{b}}/f_{\mathrm{c}}) \simeq 808$ /s. Similarly the frequency correlation spectrum (for $f_{\mathrm{s}} \simeq 30.72$ Hz and $N = 2048$ in the LTE Uplink 20 MHz scenario) is bandlimited as long as the channel's maximum delay spread $\tau_{\max} < (N/(2f_{\mathrm{s}}) = 33\ \mu$s.

Thus, for practical scenarios with velocities less than 100 m/s and delay spread less than 10 μs, the channel correlation function can safely be assumed to have a bandlimited spectrum in both directions. In the sequel, we assume a knowledge of $\omega_{\mathrm{t,max}}$ and $\omega_{\mathrm{f,max}}$, or at least their upper bounds for the transmission scenarios of interest, thereby limiting our uncertainty about the unknown channel correlation spectrum.

B. Convexity and Compactness of Uncertainty Class

Given the bandlimitedness of a realistic correlation spectrum (as described earlier), we characterize the uncertainty class $\mathcal{U}_{r_H}$ as the set of all valid 2-D channel correlation sequences with known bandwidth restrictions of $\omega_{\mathrm{t,max}}$ and $\omega_{\mathrm{f,max}}$ on their spectrum in the time and frequency directions, respectively, and furthermore having a bounded L_1 norm.[4] Put mathematically

$$\mathcal{U}_{r_H} = \left\{ \begin{array}{l} \{r_H(i,j)\}_{-i_0,-j_0}^{i_0,j_0} : r_H(0,0) \leq \beta, \quad S_H(\omega_{\mathrm{f}},\omega_{\mathrm{t}}) \geq 0, \\ S_H(\omega_{\mathrm{f}},\omega_{\mathrm{t}}) = 0, \quad \text{for } \omega_{\mathrm{f,max}} < |\omega_{\mathrm{f}}| \leq \pi \\ \qquad\qquad \text{and } \omega_{\mathrm{t,max}} < |\omega_{\mathrm{t}}| \leq \pi, \\ S_H(\omega_{\mathrm{f}},\omega_{\mathrm{t}}) = \mathcal{F}_{\mathrm{2-D}}\left(\mathcal{E}_{\mathrm{2-D}}^{\mathrm{PSD-BL}}\left(\{r_H(i,j)\}_{-i_0,-j_0}^{i_0,j_0}\right)\right) \end{array} \right\} \tag{10}$$

where $\mathcal{F}_{\mathrm{2-D}}$ denotes 2-D Fourier transformation, while $\mathcal{E}_{\mathrm{2-D}}^{\mathrm{PSD-BL}}$ represents a bandlimited positive semidefinite (PSD)[5] extrapolation. Such an extrapolation is necessary in the definition of the uncertainty class, because otherwise the Fourier transform of a finite dimensional sequence cannot lead to a bandlimited spectrum. We mention that, although a method proposed in [17] may be used for computing such a bandlimited extrapolation, in practice it is not needed anywhere to solve the optimization problem presented in Section V. We now state a proposition about the uncertainty class defined above.

Proposition 1: *The uncertainty class $\mathcal{U}_{r_H}$ of all valid 2-D channel correlation sequences with a bounded L_1 norm and bandwidth restrictions of $\omega_{\mathrm{t,max}}$ and $\omega_{\mathrm{f,max}}$ on their spectrum in the time and frequency directions, respectively, as stipulated in (10) forms a convex and compact set.*

Proof: To prove convexity, let us consider two member sequences $\{r_{H1}(i,j)\}_{-i_0,-j_0}^{i_0,j_0}$ and $\{r_{H2}(i,j)\}_{-i_0,-j_0}^{i_0,j_0}$ of the uncertainty class $\mathcal{U}_{r_H}$. With some abuse of notation, their convex combination for any $\alpha \in [0,1]$ can be expressed as $\{r_{H3}(i,j)\}_{-i_0,-j_0}^{i_0,j_0} = \alpha\{r_{H1}(i,j)\}_{-i_0,-j_0}^{i_0,j_0} + (1-\alpha)\{r_{H2}(i,j)\}_{-i_0,-j_0}^{i_0,j_0}$ and can easily be seen to satisfy the L_1 norm constraint, i.e., $r_{H3}(0,0) \leq \beta$. Moreover the convex combination of the PSD extensions of the class members $\{r_{H1}(i,j)\}_{-i_0,-j_0}^{i_0,j_0}$ and $\{r_{H2}(i,j)\}_{-i_0,-j_0}^{i_0,j_0}$ also happens to be a valid PSD extension of the sequence $\{r_{H3}(i,j)\}_{-i_0,-j_0}^{i_0,j_0}$ and as such the bandwidth restrictions are satisfied as well. Thus, the uncertainty class $\mathcal{U}_{r_H}$ is proven to be convex.

To prove compactness, we need to prove that the set is closed and bounded [18, Sec. 2.2]. A set is bounded if some distance metric between its members is bounded i.e., it can be contained in a ball of sufficiently large radius. This can be seen to hold from the definition of the uncertainty class in (10) because of the constraints $r_H(0,0) \leq \beta$ and $|r_H(i,j)| \leq r_H(0,0) \forall i,j$ (which follows from the PSD property). The closedness of the set can be observed from the fact that the set includes all its boundary points. Thus, we see that the uncertainty class $\mathcal{U}_{r_H}$ in (10) constitutes a convex and compact set. ∎

V. Minimax Optimization Formulation

Following the line of discussion at the end of Section III, we note that the problem of finding the MR 2-D estimator can be expressed as a maximization followed by a minimization operation. We first maximize the MSE in (5) over the set of all valid 2-D channel correlation sequences to arrive at the worst case scenario and then minimize the resultant MSE via MMSE optimization for the filter, i.e.

$$\min_{\boldsymbol{w} \in C^{N_{\mathrm{T}} N_{\mathrm{F}}}} \quad \max_{\{r_H(i,j)\} \in \mathcal{U}_{r_H}} \quad \varepsilon(\boldsymbol{w}, \{r_H(i,j)\}) \tag{11}$$

where $\mathcal{U}_{r_H}$ refers to the uncertainty class defined in (10) for the 2-D channel correlation sequences. We now use the following theorem from [19] on the equivalence of finite dimensional minimax and maximin problems under consideration. For more insight into this theorem we refer the reader to [20, Sec. 2.4.1.2], while for a historical perspective n minimax theorems [21, Ch. 1 by S. Simons] is highly recommended.

Theorem 2: *Given the estimation problem of a zero mean parameter θ from observation γ, with unknown covariance matrix $\boldsymbol{K} = \begin{bmatrix} \boldsymbol{R}_{\gamma\gamma} & \boldsymbol{R}_{\gamma\theta} \\ \boldsymbol{R}_{\theta\gamma} & \boldsymbol{R}_{\theta\theta} \end{bmatrix}$ belonging to a convex and compact uncertainty class $\mathcal{U}_K$, and given that the estimator is linear, viz.*

[3]Alternatively, the relationship follows from (8) by exploiting the channel duality concepts of Doppler spread and coherence time versus delay spread and coherence bandwidth (see, for instance, [8]).

[4]L_1 norm of a positive valued continuous function is defined to be the area under its curve, which is equal to $r_H(0,0)$ in the context of correlation spectrum.

[5]A complex valued conjugate symmetric sequence $\{x(n)\}$ is said to be PSD, if $\sum_{m=1}^{N}\sum_{n=1}^{N} a(m)x(m-n)a^*(n) \geq 0$ for all complex valued sequences $a(n)$ or equivalently if the sequence $x(n)$ has a real-valued nonnegative Fourier spectrum.

the estimate can be expressed as $\hat{\theta} = \boldsymbol{w}^{\mathrm{H}}\gamma$*, the solution of the following two problems coincide*

$$\min_{\boldsymbol{w}} \max_{\boldsymbol{K}\in\mathcal{U}_K} \mathrm{E}\left[\|\theta - \hat{\theta}\|_2^2\right]$$

$$\max_{\boldsymbol{K}\in\mathcal{U}_K} \min_{\boldsymbol{w}} \mathrm{E}\left[\|\theta - \hat{\theta}\|_2^2\right]$$

i.e., the solutions, the LF covariance matrix $\boldsymbol{K}_* \in \mathcal{U}_K$ *and the corresponding MMSE estimator* $\boldsymbol{w}_*$ *form a saddle point.*

Adapted to our scenario, the parameter θ to be estimated is the data CFR $H_{f,t}$, while $\gamma = \tilde{\boldsymbol{h}}_{\mathrm{p}}$ forms the observation vector. The covariance matrix $\boldsymbol{K}$ is completely characterized by the correlation sequence $\{r_H(i,j)\}_{-i_0,-j_0}^{i_0,j_0}$. Thus, the condition of convexity and compactness of $\mathcal{U}_K$ in our scenario translates into the convexity and compactness of the uncertainty class $\mathcal{U}_{r_H}$ which holds already from Proposition 1. Consequently, the original minimax problem in (11) can be reformulated into the equivalent maximin problem

$$\max_{\{r_H(i,j)\}\in\mathcal{U}_{r_H}} \min_{\boldsymbol{w}} \varepsilon(\boldsymbol{w}, \{r_H(i,j)\}). \tag{12}$$

In essence, the problem of finding the MR estimator is now casted into the one of finding the LF 2-D correlation sequence. Note that the inner minimization problem in (12) is nothing else than the conventional MMSE optimization problem discussed already in Section III, leading to the following residual problem [cf. (6) and (7)]

$$\max_{\{r_H(i,j)\}\in\mathcal{U}_{r_H}} r_H(0,0) - \boldsymbol{r}_{h_{\mathrm{p}}}^{\mathrm{H}}\left(\boldsymbol{R}_{h_{\mathrm{p}}} + \boldsymbol{R}_\eta\right)^{-1}\boldsymbol{r}_{h_{\mathrm{p}}}. \tag{13}$$

Thus, we arrived from a minimax optimization problem down to a pure (still nonconvex) maximization problem. We note that, given the objective function in (13) is monotonically increasing in $r_H(0,0)$, the maximization is reached once the L_1 norm bound in (10) is fulfilled with equality, i.e., $r_H(0,0) = \beta$. Thus, instead of maximizing the complete objective function, we may just minimize the subtracter under the additional equality constraint $r_H(0,0) = \beta$, so that the problem in (13) reduces to

$$\min_{\{r_H(i,j)\}\in\mathcal{U}_{r_H}} \boldsymbol{r}_{h_{\mathrm{p}}}^{\mathrm{H}}\left(\boldsymbol{R}_{h_{\mathrm{p}}} + \boldsymbol{R}_\eta\right)^{-1}\boldsymbol{r}_{h_{\mathrm{p}}} \quad \text{s.t.} \quad r_H(0,0) = \beta \tag{14}$$

which can be transformed via the epigraph notation [22, Sec. 3.1.7] into

$$\min_{t,\{r_H(i,j)\}\in\mathcal{U}_{r_H}} t \quad \text{s.t.} \quad t - \boldsymbol{r}_{h_{\mathrm{p}}}^{\mathrm{H}}\left(\boldsymbol{R}_{h_{\mathrm{p}}} + \boldsymbol{R}_\eta\right)^{-1}\boldsymbol{r}_{h_{\mathrm{p}}} \geq 0, \quad r_H(0,0) = \beta \tag{15}$$

and now by employing the Schur complement positive semidefiniteness theorem,[6] the optimization problem reads as

$$\min_{t,\{r_H(i,j)\}\in\mathcal{U}_{r_H}} t \quad \text{s.t.} \begin{bmatrix} t & \boldsymbol{r}_{h_{\mathrm{p}}}^{\mathrm{H}} \\ \boldsymbol{r}_{h_{\mathrm{p}}} & \boldsymbol{R}_{h_{\mathrm{p}}} + \boldsymbol{R}_\eta \end{bmatrix} \succeq 0, \quad r_H(0,0) = \beta. \tag{16}$$

[6]Given a matrix $\boldsymbol{M} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix}$, and the Schur complement of $\boldsymbol{A}_{22}$ as $\boldsymbol{S}_{22} = \boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}$, we have $\boldsymbol{M} \succeq 0$ if and only if $\boldsymbol{A}_{22} \succ 0$ and $\boldsymbol{S}_{22} \succeq 0$. See, for instance, [20, Thm A.2].

Thus we get rid of the matrix inverse in the original problem formulation. Next comes the crucial step of decomposing the $\{r_H(i,j)\} \in \mathcal{U}_{r_H}$ constraint into individual analytical constraints. To this end,

- we note that the sequence $\{r_H(i,j)\}$ being a 2-D autocorrelation sequence must satisfy the positive semidefiniteness properties. Furthermore,
- the sequence $\{r_H(0,j)\}$, i.e., the time correlation function, besides being positive semidefinite must be bandlimited to the interval $[-\omega_{\mathrm{t,max}}, \omega_{\mathrm{t,max}}]$ for reasons outlined in Section IV-A–1.
- Similarly, the frequency correlation function $\{r_H(i,0)\}$ should have a real valued positive spectrum that is one sided and bandlimited to the interval $[0, \omega_{\mathrm{f,max}}]$. See Section IV-A–2 for a discussion.

The positive semidefiniteness of the finite dimensional correlation sequence can be ensured by a simple positive semidefinite constraint on the appropriately defined correlation matrix $\boldsymbol{R}_H$ containing all frequency and time offsets up to i_0 and j_0, respectively.

Constraining the finite support time and frequency autocorrelation functions to be bandlimited besides being positive semidefinite is, however, nontrivial and we use the following theorem from [23] on the existence and uniqueness of bandlimited positive semidefinite extensions of partially known positive semidefinite sequences.

Theorem 3: *A positive semidefinite sequence* $\{x(m)\}$ *for* $m = -M, \ldots, M$ *of length* $2M+1$ *has a positive semidefinite extension bandlimited to* $[\omega_l, \omega_h]$*, if and only if the sequence*

$$\breve{x}(m) = e^{\iota(\omega_h+\omega_l)/2}x(m-1) + e^{-\iota(\omega_h+\omega_l)/2}x(m+1) - 2\cos((\omega_h - \omega_l)/2)x(m)$$

evaluated at $m = 0, 1, \ldots, M-1$ *forms a* $M \times M$ *positive semidefinite Toeplitz matrix.*

To impose the respective bandwidth constraints, this theorem can be applied separately to the time and frequency autocorrelation sequences. For the time correlation sequence, we need to ensure the existence of a positive semidefinite extrapolation bandlimited to $[-\omega_{\mathrm{t,max}}, \omega_{\mathrm{t,max}}]$, which means that the sequence

$$r_T(j) = r_H(0,j-1) + r_H(0,j+1) - 2\cos(\omega_{\mathrm{t,max}})r_H(0,j) \tag{17}$$

for $j = 0, 1, \ldots, 2K_{\mathrm{T}}\Delta_{\mathrm{T}} - 1$ leads to a positive semidefinite Toeplitz matrix $\boldsymbol{R}_T \in bbT^{2K_{\mathrm{T}}\Delta_{\mathrm{T}}\times 2K_{\mathrm{T}}\Delta_{\mathrm{T}}}$. Similarly, for the frequency correlation sequence to be bandlimited to $[0, \omega_{\mathrm{f,max}}]$, the sequence

$$r_F(i) = e^{\iota\omega_{\mathrm{f,max}}/2}r_H(i-1,0) + e^{-\iota\omega_{\mathrm{f,max}}/2}r_H(i+1,0) - 2\cos(\omega_{\mathrm{f,max}}/2)r_H(i,0) \tag{18}$$

for $i = 0, 1, \ldots, 2K_{\mathrm{F}}\Delta_{\mathrm{F}} - 1$ must result in a positive semidefinite Toeplitz matrix $\boldsymbol{R}_F \in \mathbb{T}^{2K_{\mathrm{F}}\Delta_{\mathrm{F}}\times 2K_{\mathrm{F}}\Delta_{\mathrm{F}}}$.

Finally, given the three positive semidefiniteness constraints on $\boldsymbol{R}_H, \boldsymbol{R}_T$and $\boldsymbol{R}_F$ to characterize the uncertainty class, the desired overall optimization problem from (16) can now be posed as

$$\left\{\bar{t},\left\{r_H^{\mathrm{LF}}(i,j)\right\}\right\}=\underset{t,\{r_H(i,j)\}}{\operatorname{argmin}} \quad t \quad \text{s.t.} \begin{bmatrix} t & \boldsymbol{r}_{h_\mathrm{p}}^{\mathrm{H}} \\ \boldsymbol{r}_{h_\mathrm{p}} & \boldsymbol{R}_{h_\mathrm{p}}+\boldsymbol{R}_\eta \end{bmatrix} \succeq 0$$
$$r_H(0,0)=\beta, \boldsymbol{R}_H \succeq 0,$$
$$\boldsymbol{R}_T \succeq 0, \boldsymbol{R}_F \succeq 0. \quad (19)$$

Thus, we arrive at a convex minimization problem with a linear cost function, an equality constraint, and some positive semidefiniteness constraints. As such, the problem can be solved via any semidefinite problem solver like SeDuMi [24]. The solution of this problem yields the LF 2-D CFR correlation sequence $\{r_H^{\mathrm{LF}}(i,j)\}$ with bandwidth constraints of $\omega_{\mathrm{f,max}}$ and $\omega_{\mathrm{t,max}}$ on its (positive semidefinite extrapolation's) spectrum in the frequency and time direction, respectively. This LF correlation sequence can now be used for the computation of the MR 2-D MMSE estimation filter coefficients, i.e.

$$\boldsymbol{w}_{\mathrm{MMSE}}^{\mathrm{MR}}=\left(\boldsymbol{R}_{h_\mathrm{p}}^{\mathrm{LF}}+\boldsymbol{R}_\eta\right)^{-1}\boldsymbol{r}_{h_\mathrm{p}}^{\mathrm{LF}} \quad (20)$$

with $\boldsymbol{r}_{h_\mathrm{p}}^{\mathrm{LF}}$ and $\boldsymbol{R}_{h_\mathrm{p}}^{\mathrm{LF}}$ as defined in Section III. The superscript $(\bullet)^{\mathrm{LF}}$ indicates that they are constructed from the optimized LF correlation sequence $\{r_H^{\mathrm{LF}}(i,j)\}$. Some remarks about the result follow.

Remark 1: An analysis of the MSE degradation owing to the correlation model mismatches pursued in [4] and [5] led to the proposal of a sinc correlation-based robust estimator. But the entire analysis was conditioned on the availability of an infinite number of observations. In this work, the analysis for the considered scenario of a limited number of pilot observations along time and frequency directions and bounded L_1 norm demonstrates that the sinc-based heuristic estimator is no longer the MR estimator.

Remark 2: Unlike the heuristic estimator for a finite number of observations [4], the MR estimator proposed in this work is based on the LF 2-D correlation sequence, that unfortunately can not be expressed explicitly rather it results from the solution of the semidefinite optimization problem in (19).

Remark 3: For the case of white Gaussian noise with $\boldsymbol{R}_\eta = \sigma_\eta^2\boldsymbol{I}$, the solution of the optimization problem in (19) can easily be shown to be independent of the individual values β or σ_η^2, but depend only on their ratio β/σ_η^2 which equals $\mathrm{E}[|H_{f,t}|^2]/\sigma_\eta^2$ and can be seen (cf. (2)) as the operational signal-to-noise ratio (SNR).

Remark 4: The computationally expensive optimization procedure in (19) needs to be carried out offline only once for a given set of pilot grid and estimation parameters, so that the associated computational effort is not considered to be a problem for the real time *estimation procedure* itself. Precisely speaking, the solution—the LF correlation sequence—depends only on the pilot spacings Δ_T and Δ_F, the number of pilot observations N_T and N_F, and the operational SNR. Thus, the 2-D LF correlation sequence or even the corresponding robust estimator coefficients can easily be precomputed offline for a particular system at quantized values of SNR and at upper bounds on the channel's delay and Doppler spread. As such, the online estimation complexity remains identical to that of any other 2-D MMSE estimator.

Remark 5: It can be demonstrated, that the 2-D correlation sequence corresponding to a rectangular spectrum results as the solution of a minimax optimization problem over finite observations, but with bounded L_∞ norm instead of L_1 norm. However, in practice it might be hard to estimate a bound on spectrum's L_∞ norm. Moreover, given such a bound, the hypothesis of a rectangular spectrum would obviously lead to a rather loose upper bound of the corresponding L_1 norm ($=\mathrm{bandwidth}\times \mathrm{L}_\infty$ norm), thus overestimating the actually encountered L_1 norm of most channels and resulting in a significantly more conservative MSE performance.

Remark 6: The derivations in current work, are based on the assumption of a rectangular aligned pilot grid with fixed pilot spacings along frequency and time direction. The results therefore can not be directly applied to the more general pilot grid structures. In the conference version of the work, we address some aspects of practical importance: handling the pilot grid edge effects in [25] and the case of staggered (skewed) pilot grids in [13].

Remark 7: The proposed MR estimator promises the best *worst case* MSE performance which in practical systems often dominates the *average* bit error rate (BER) performance. Thus, we conjecture that the proposed MR estimator, besides its best worst case MSE performance, offers a superior average BER performance as compared to the heuristic estimators. This is indeed confirmed by the BER simulation results in Section VI-B.

VI. Simulation Results and Further Insights

In this section, we present a comparison of the proposed estimator against one of the most commonly employed heuristic robust estimator based on the rectangular correlation spectrum. The performance comparison is presented in terms of the worst case estimation MSE in Section VI-A—which has been our optimization criteria in this paper—and also in terms of the average BER performance in Section VI-B—which is of prime interest in many practical communication systems. Finally in Section VI-C, an interesting link between the two estimators is established.

We consider an OFDM system with the total number of subcarriers being $N=2048$ and a regular rectangular pilot grid, as described in Section II, with pilot spacing along time $\Delta_\mathrm{T}=3$ and along frequency $\Delta_\mathrm{F}=2$. Among the estimation parameters we set $K_\mathrm{F}=2$ and $K_\mathrm{T}=1$ as the number of pilot observations available on either side along frequency and time axis, respectively. The channel's L_1 norm is bounded by $\beta=1$, while the observation noise power is fixed at $\sigma_\eta^2=10^{-3}$ or $\sigma_\eta^2=10^{-1}$. These β/σ_η^2 ratios correspond to operating SNR points of 30 dB and 10 dB, respectively.

For the sake of clarity, we reiterate that in all subsequent discussions, the Heuristic Robust (HR) estimator refers to a 2-D MMSE estimator based on a rectangular Doppler spectrum and uniform power delay profile as proposed in [4], while the proposed Maximally Robust (MR) estimator is based on the LF 2-D correlation sequence obtained as a result of the finite observation based minimax optimization in Section V. The online estimation process itself is completely identical in both cases and has the same computational effort as that of any other 2-D MMSE estimator. As a reference, we also consider the optimal 2-D MMSE filter based on the perfect knowledge of the channel correlation spectrum. It determines basically the lower bound on the performance achieved by any robust filter.

A. *Worst Case Estimation MSE Performance Analysis*

A comparison of the two robust estimators in terms of their worst case estimation MSE is pursued here via Monte Carlo simulations, i.e., we investigate the guaranteed maximum MSE that they offer in the case of worst possible mismatch between the assumed and the actual channel correlation spectrum. To this end, we maximize (via Monte Carlo simulations) the MSE of the estimators over different randomly generated correlation spectra with bounded energy and being bandlimited in accordance with the uncertainty class (10). Thus, we obtain

$$\max_{\{r_H(i,j)\}\in\mathcal{U}_{r_H}} \mathrm{E}[|H_{f,t}-\hat{H}_{f,t}|^2] \quad \text{s.t.} \quad \boldsymbol{w}=\{\boldsymbol{w}_{\mathrm{MR}} \text{or } \boldsymbol{w}_{\mathrm{HR}}\}$$

with $\boldsymbol{w}_{\mathrm{MR}}$ and $\boldsymbol{w}_{\mathrm{HR}}$ representing the fixed 2-D MMSE estimator based on the proposed LF and the heuristic rectangular correlation spectrum, respectively.

Fig. 1 shows the worst case estimation MSE, as described above, for both heuristic and proposed MR estimator as a function of the increasing bandwidths of the time and frequency correlation spectra. Note that increasing the bandwidths of the time and frequency correlation spectrum corresponds to increasing the relative terminal velocity and channel impulse response length as described by (8) and (9), respectively. Also shown is the performance of the reference MMSE estimator with perfect knowledge of the channel correlation function.

It can readily be seen that the proposed MR estimator, as expected, leads to a smaller worst case MSE as compared to the heuristic estimator in the various scenarios considered. Thus, in terms of their worst case estimation MSE, the MR estimator achieves gains of as much as 1–2 dB over the performance of heuristic estimator. Moreover, as the bandwidths of time and frequency correlation spectra increase, the gains easily approach 3–4 dB.

Fig. 2 shows similar worst case MSE performance of the two robust estimators at a higher noise level of $\sigma_\eta^2 = 10^{-1}$. Again, we note a similar trend in all the scenarios; the performance of the MR estimator is superior by at least 1 dB as compared to that of the heuristic estimator.

B. *BER Performance Analysis*

The result in Section VI.A that the proposed MR estimator offers a superior worst case performance as compared to the heuristic estimator is somewhat expected. The MR estimator in fact promises the best worst case MSE performance than any other heuristic choice for the given uncertainty class. More interestingly, in this subsection we explore the performance of the two estimators in terms of the coded BER over certain practical transmission channels. To this end, we consider the LTE uplink system and choose the 20 MHz band for various relevant parameters such as symbol duration and cyclic prefix length [10]. The LTE uplink pilot grid can be described by $\Delta_{\mathrm{T}} = 7$ and $\Delta_{\mathrm{F}} = 1$. We choose $K_{\mathrm{T}} = 1$ and $K_{\mathrm{F}} = 2$ to determine the

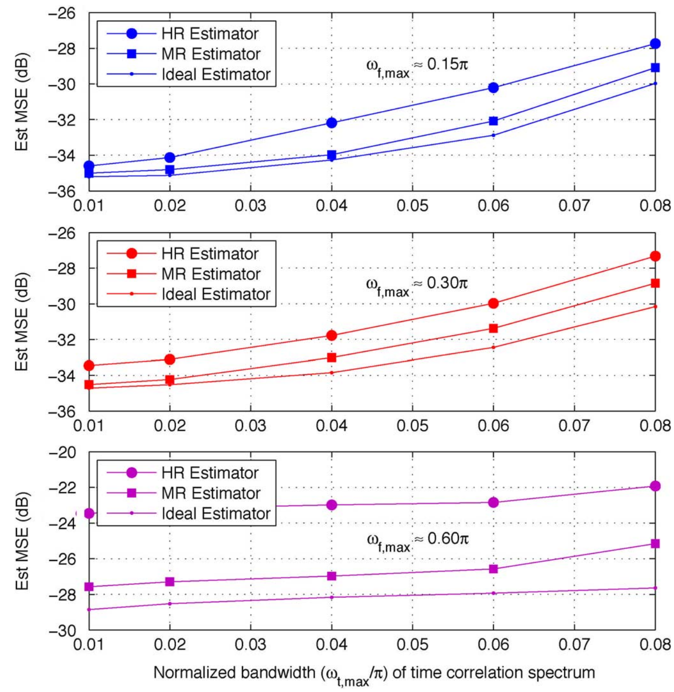


Fig. 1. Worst case channel estimation MSE (at $\sigma_\eta^2 = 10^{-3}$) of the heuristic and the proposed maximally robust estimator along with that of the reference estimator having perfect knowledge of channel correlation function. The three subplots show the performance at different frequency correlation spectra bandwidths $\omega_{\mathrm{f,max}}$, whereas the time correlation spectra bandwidths $\omega_{\mathrm{t,max}}$ run along the horizontal axis. Under, for instance LTE uplink parameters, the values of $\omega_{\mathrm{f,max}}$ and $\omega_{\mathrm{t,max}}$ correspond, respectively, to CIR lengths of 150, 300, and 600 taps, and terminal velocities of 30 to 240 kmph. Gains of 1–4 dB are achieved by the MR estimator over the HR estimator.

interpolation window[7] and employ the standard frequency domain MMSE equalizer and a rate 1/3 turbo code from the LTE specifications to show the overall system performance for both the estimators. Thus, the difference only comes from the choice of the fixed channel correlation sequence employed by the two MMSE channel estimators.

We note that the degree of mismatch between the actual simulated correlation spectrum and the one assumed by the estimator has direct impact on the final BER performance. Thus, we study the BER performance of the estimators under different correlation spectra simulated by having different channel Power Delay Profiles (PDP) and Doppler Spectra (DS). Specifically, for the channel frequency correlation spectra we use the standard Vehicular-A (Veh-A) and Typical Urban (TU) power delay profiles [26], while for the time correlation spectra we use the well known Jakes DS [27] and a Bandpass DS with evenly distributed power at the extreme 10% Doppler frequencies.

Fig. 3 shows the comparison under two different transmission scenarios. For QPSK modulation at a speed of 120 kmph with Bandpass DS and Veh-A PDP, we observe a gain of 1.05 dB at the coded BER of 10^{-3} by employing the MR channel estimator instead of the heuristic estimator. Similarly for 16-QAM modulation scheme at 60 kmph with TU PDP and the standard Jakes DS, a gain of 0.9 dB is achieved at the coded BER of 10^{-3} by

[7]The choice $K_{\mathrm{T}} = 1$ gets its motivation from the proposed slot-based frequency hopping in LTE uplink specifications [10], which means that a choice of $K_{\mathrm{T}} > 1$ is not feasible.

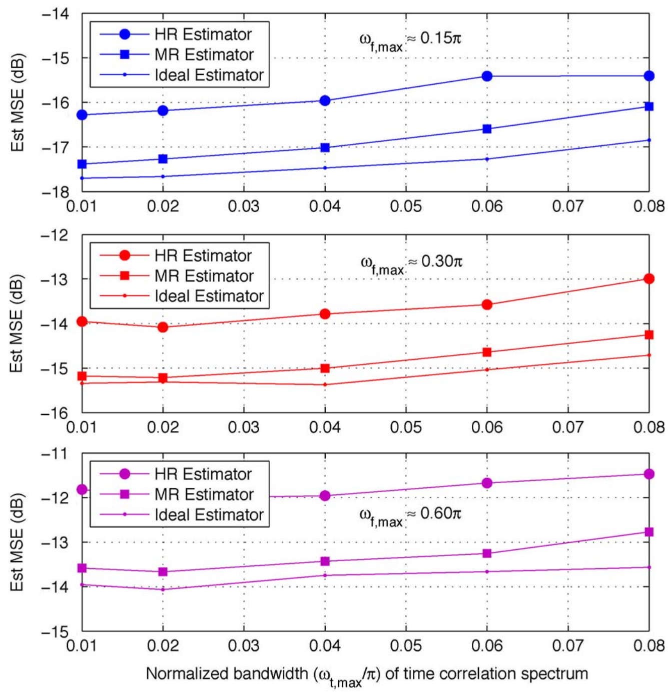
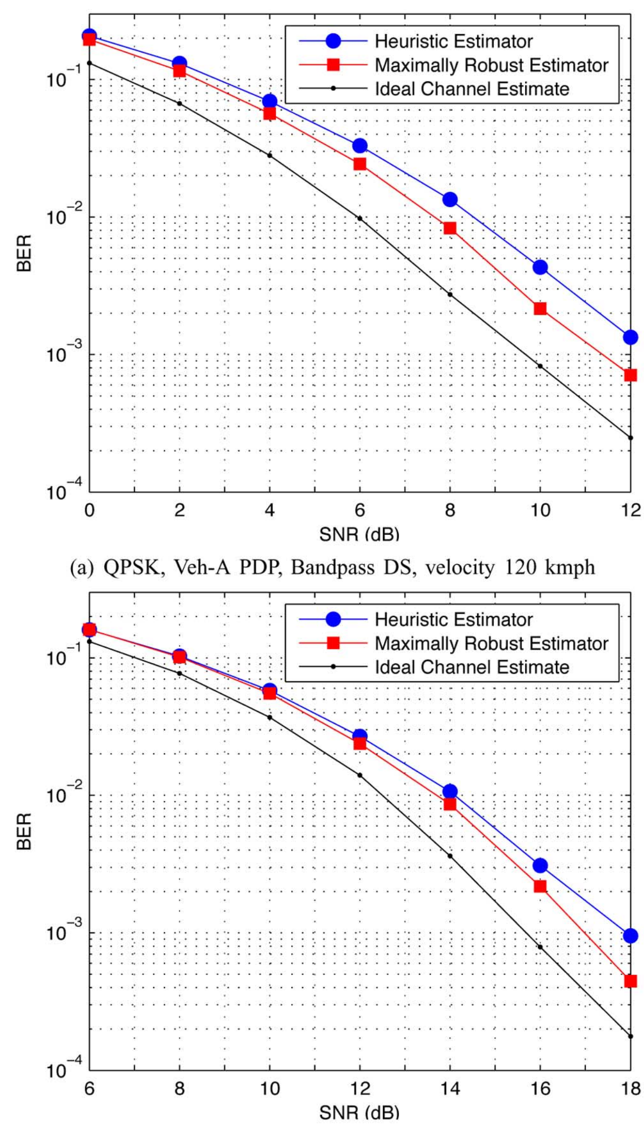


Fig. 2. Worst case channel estimation MSE (at $\sigma_\eta^2 = 10^{-1}$) of the heuristic and proposed estimator. Please refer to the caption of Fig. 1 for further details. Gains of around 1 dB are achieved by the MR estimator.



(a) QPSK, Veh-A PDP, Bandpass DS, velocity 120 kmph

(b) 16-QAM, TU PDP, Jakes DS, velocity 60 kmph

Fig. 3. Performance comparison of heuristic and MR estimators against the performance with ideal channel estimation in terms of coded BER for different transmission scenarios. A gain of around 1 dB is achieved by the MR estimator over the heuristic one in both scenarios, which speaks of the advantage of the worst case design approach (cf. Remark 7).

the MR estimator. Thus, we note that, as remarked at the end of Section V, guarding against the worst case MSE performance turns out to be a good choice even with regard to the averaged BER performance.

It may be pointed out here that the BER performance comparison, as above, naturally depends on the underlying simulation channel model. For instance, a channel model with uniform power delay profile and a rectangular (or a Jakes) Doppler spectrum is favorable for the rectangular spectrum-based heuristic estimator, since no (or only a slight) model mismatch occurs. On the other hand, the use of a negative exponential or a some other nonuniform power delay profile combined with a nonuniform Doppler spectrum is not as much favorable for the heuristic estimator. Nevertheless, in practical transmission scenarios with rapid unpredictable channel variations, the proposed robust estimator is poised to outperform the heuristic one, because of its better guaranteed worst case performance.

C. Convergence of Estimators at an Infinite Number of Observations

In this section, we present an additional result that gives further insight into the characteristics of the MR 2-D MMSE estimator presented in this paper and relates it to the previously proposed heuristic robust estimator [4]. To this end, we show in Fig. 4 the effect of increasing the number of observations along time direction to be used for CFR estimation at the data position of interest while keeping other parameters fixed at $\Delta_\mathrm{T} = 2$, $\Delta_\mathrm{F} = 2$ and $K_\mathrm{F} = 1$. Thus, we try to approach the assumption of infinite number of time observations ($K_\mathrm{T} \to \infty$) and observe the impact on the resulting LF time correlation spectrum.

Note that given a finite number of time observations, the optimization problem (19) only identifies some of the initial values of the time autocorrelation function, *viz.* offsets $0, 1, \ldots, (2K_\mathrm{T} - 1)\Delta_\mathrm{T}$. Nevertheless, the optimization setup (Theorem 3) ensures that this sequence has a positive semidefinite extension bandlimited to $[-\omega_{\mathrm{t,max}}, \omega_{\mathrm{t,max}}]$. In order to arrive at this extrapolation, we use the Levinson's Algorithm [28] based method outlined in [17].[8] Once an extrapolation up to a sufficiently large time offset is obtained, we take the Fourier transform to view the spectrum of the LF time correlation function.

[8]Although the original random process is obviously regular, for purposes of analysis the underlying random process of the outcoming autocorrelation function w.r.t. the optimization problem (19) can be interpreted as a deterministic process which is a key assumption for applying the proposed extrapolation technique.

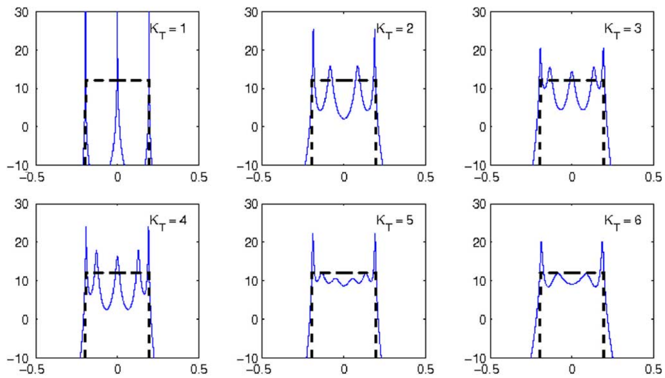


Fig. 4. Effect of increasing $K_{\rm T}$, the number of observations along time direction on the LF time correlation spectrum. The solid lines plot the spectra of LF time correlation sequence whereas the dashed lines plot the rectangular (heuristic) time correlation spectrum for reference. The vertical axis displays the spectrum magnitude in dB while the horizontal axis shows the angular frequency ($-\pi$ to π). Note that we show only the central portion of the spectra.

This spectrum is plotted in Fig. 4 for increasing values of $K_{\rm T}$, i.e., $K_{\rm T} = 1, 2, \ldots, 6$ along with the heuristic rectangular spectrum. Ignoring the Gibb's phenomena [29] at the spectrum edges, one can easily note the interesting effect of the convergence of the LF time correlation (Doppler) spectrum to the rectangular spectrum as $K_{\rm T}$ increases. Thus, we conjecture that the proposed optimization setup leads us in the limiting case of $K_{\rm T} \rightarrow \infty$ to the rectangular Doppler spectrum as was proposed in [4] and [5].

A similar effect, not shown here, is observed once we increase the number of observations along the frequency direction, while keeping the number of time observations fixed. The proposed optimization setup again leads us in the limiting case of $K_{\rm F} \rightarrow \infty$ to a rectangular spectrum (corresponding to the uniform power delay profile) in line with the heuristic proposal of Li *et al.* in [4].

VII. CONCLUSION

The paper discussed the idea of robust 2-D MMSE channel estimation with a finite number of available observations. It has been shown that the heuristic robust estimator based on the rectangular (uniform) correlation spectrum along both frequency and time direction, proposed in [4], is not MR if the number of observations is finite. The MR estimator has been shown to be obtained as a result of a semidefinite optimization procedure for finding the LF 2-D correlation sequence under certain natural constraints on the uncertainty class. The proposed MR estimator has been shown to exhibit a 1–4 dB superior worst case estimation MSE performance, as well as around 1 dB gain in terms of average coded BER for a practical wireless system under realistic transmission scenarios. The paper concluded with the demonstration of an interesting phenomena about the convergence of the proposed MR estimator to the previously proposed heuristic estimator as the number of observations increase.

REFERENCES

[1] J. J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesson, "On channel estimation in OFDM systems," in *Proc. IEEE Veh. Technol. Conf.*, Jul. 1995, vol. 2, pp. 815–819.
[2] O. Edfors, M. Sandell, J. J. van de Beek, S. K. Wilson, and P. O. Borjesson, "OFDM channel estimation by singular value decomposition," in *Proc. IEEE Veh. Technol. Conf.*, May 1996, vol. 2, pp. 923–927.
[3] P. Hoeher, S. Kaiser, and P. Robertson, "Pilot-symbol-aided channel estimation in time and frequency," in *Proc. Int. Conf. Acoust., Speech Signal Process. (ICASSP'97)*, Munich, Germany, Apr. 1997, vol. 3, pp. 1845–1848.
[4] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 902–915, Jul. 1998.
[5] Y. Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 1207–1215, Jul. 2000.
[6] S. Verdu and H. V. Poor, "On minimax robustness: A general approach and applications," *IEEE Trans. Inf. Theory*, vol. 30, pp. 328–340, Mar. 1984.
[7] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proc. IEEE*, vol. 73, pp. 433–481, Mar. 1985.
[8] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[9] The IEEE 802.16 Working Group on Broadband Wireless Access Standards. [Online]. Available: http://www.ieee802.org/16 2010
[10] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation*, TS 36.211, Standardization Committee 3GPP, 2010 [Online]. Available: http://www.3gpp.org
[11] R. Nilsson, O. Edfors, M. Sandell, and P. O. Borjesson, "An analysis of two-dimensional pilot-symbol assisted modulation for OFDM," in *Proc. IEEE Int. Conf. Pers. Wireless Commun.*, Dec. 1997, pp. 71–74.
[12] T. Strohmer and S. Beaver, "Optimal OFDM design for time-frequency dispersive channels," *IEEE Trans. Commun.*, vol. 51, pp. 1111–1122, Jul. 2003.
[13] M. D. Nisar, W. Utschick, and J. Forster, "Robust 2-D Channel Estimation for Staggered Pilot Grids in Multi-Carrier Systems: LTE Downlink as an example," *Lecture Notes in Elect. Eng. (Vol 41), Multi-Carrier Syst. Solutions (MCSS), Herrsching, Germany*, pp. 113–121, May 2009.
[14] G. Matz and F. Hlawatsch, "Time-varying communication channels: Fundamentals, recent developments, and open problems," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Florence, Italy, Sep. 2006.
[15] W. Kozek and A. F. Molisch, "On the eigenstructure of underspread WSSUS channels," in *Proc. IEEE Workshop on Signal Process. Adv. Wireless Commun. (SPAWC)*, Paris, France, Apr. 1997, pp. 325–328.
[16] L. W. Couch. II, *Digital and Analog Communications Systems*, 6th ed. Englewood Cliffs, NJ: Prentice-Hall, 2001.
[17] A. Papoulis, "Levinson's algorithm, wold's decomposition and spectral estimation," *SIAM Rev.*, vol. 27, pp. 405–441, Sep. 1985.
[18] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, 3rd ed. New York: Wiley , 2006.
[19] V. N. Soloviov, "Towards the theory of minimax-Bayesian estimation," *Theory Probab. Appl.*, vol. 44, pp. 739–754, 2000.
[20] F. A. Dietrich, *Robust Signal Processing for Wireless Communications*, 1st ed. New York: Springer, 2008.
[21] *Minimax and Applications*, D. Du and P. M. Pardalos, Eds., 1st ed. New York: Springer, 1995.
[22] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[23] K. S. Arun and L. C. Potter, "Existence and uniqueness of band-limited, positive semidefinite extrapolations," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 48, pp. 547–549, Mar. 1990.
[24] J. F. Sturm, "Using SeDuMi 1.02, A MATLAB toolbox for optimization over symmetric cones," *Optim. Methods and Softwares*, vol. 11–12, pp. 625–653, 1999.
[25] M. D. Nisar, W. Utschick, and T. Hindelang, "Robust 2-D channel estimation for multi-carrier systems with finite dimensional pilot grid," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2009, pp. 2685–2688.
[26] *Guidelines for Evaluation of Radio Transmission Technologies for IMT-2000*, ITU-R Rec. M. 1225, Standardization Committee ITU.
[27] W. C. Jakes, *Mobile Microwave Communication*. New York: Wiley , 1974.
[28] N. Levinson, "The Wiener RMS error criterion in filter design and prediction," *J. Math. Phys.*, vol. 25, pp. 261–278, 1947.
[29] H. Carslaw, "A historical note on Gibb's phenomenon in Fourier series and integrals," *Bull. Amer. Math. Soc.*, vol. 31, pp. 420–424, 1925.

Muhammad Danish Nisar (S'05) was born in Karachi, Pakistan, in 1982. He received the Bachelor's degree in electrical engineering from the National University of Science and Technology (NUST), Pakistan, in 2004, and the Master's degree in communication engineering from the Technical University Munich (TUM), Germany, in 2006, both with honors.

Since 2007, he has been pursuing research toward the doctoral degree at the Associate Institute for Signal Processing, Technical University Munich (TUM). His research is being funded by Nokia Siemens Networks GmbH & Co. KG, Munich. His research interests focus on robust signal processing for communications, primarily from the perspective of physical layer techniques in mobile communication transceivers.

Wolfgang Utschick (SM'06) was born on May 6, 1964. He completed several industrial education programs before he received the diploma and doctoral degrees, both with honors, in electrical engineering from the Technische Universität München, Germany (TUM), in 1993 and 1998, respectively. In this period he held a scholarship of the Bavarian Ministry of Education for exceptional students.

From 1998 to 2002, he codirected the Signal Processing Group of the Institute of Circuit Theory and Signal Processing, TUM. Since 2000, he has been consulting in 3GPP standardization in the field of multielement antenna systems. In 2002, he was appointed Professor at the TUM where he is Head of the Fachgebiet Methoden der Signalverarbeitung. He teaches courses on signal processing, stochastic processes, and optimization theory in the field of digital communications.

Dr. Utschick was awarded in 2006 for his excellent teaching records at TUM, and in 2007, he received the ITG Award of the German Society for Information Technology (ITG). He is a senior member of the German VDE/ITG where he has been appointed in the Expert Committee for Information and System Theory in 2009. He is currently also serving as a Chairman of the national DFG Focus Program "Communications in Interference Limited Networks" (COIN). He is an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and an Editor and Founder of the Springer Book Series "Foundations in Signal Processing, Communications and Networking."

Thomas Hindelang (S'97–M'02) received the Dipl.-Ing. and the Ph.D. degree in electrical engineering from Technische Universität München (TUM), Germany, in 1994 and 2001, respectively.

He worked as a Research Assistant with the Institute of Communications Engineering from 1994 to 2000, focusing primarily on the field of combined source and channel coding for mobile communications. While on leave from TUM, he served as a consultant with the Communications Research Department, AT&T Labs, Florham Park, NJ, from March to August 1999. From 2000 to 2008, he was with Siemens AG and Nokia Siemens Networks GmbH & Co. KG, first in the area of mobile devices and later in the area of mobile networks. From 2002 to 2006, he actively participated and contributed to the 3GPP standardization body on physical layer issues for the Multimedia Broadcast and Multicast Service (MBMS) and 3G Long Term Evolution (LTE) standards. From 2006 to 2008, he led the group "Baseband Algorithms and Simulations," where he was responsible for the physical layer algorithms and simulations for GSM, UMTS, LTE, and WiMAX hardware development. Since October 2007, he has been a Senior Lecturer at TUM, and since November 2008, he has been an examiner in the area Audio Video Media at the European Patent Office. He holds approximately 40 patents and is the author of more than 30 published conference and journal papers. He contributed to the book *Advances in Digital Speech Transmission*, R. Martin, U. Heute, and C. Antweiler, Eds. His research interest includes the entire physical layer from the source until the channel, i.e., speech and video coding, error correction coding, equalization, and channel estimation techniques.