Doubly-Selective MMSE Channel Estimation and ICI Mitigation for OFDM Systems

Ronald Nissel and Markus Rupp
Institute of Telecommunications, Vienna University of Technology, Vienna, Austria
Email: {rnissel, mrupp}@nt.tuwien.ac.at

***Abstract*—In high mobility orthogonal frequency division multiplexing systems, subcarriers are no longer orthogonal, causing Inter-Carrier Interference (ICI). Equalization then becomes more challenging and requires an accurate estimate of the time-variant channel. In this paper, we propose a novel Minimum Mean Squared Error (MMSE) estimation of the sampled time-variant transfer function. Based on such channel estimation, we propose an iterative three step ICI mitigation technique whereby each step increases the channel estimation accuracy and, consequently, the performance of our MMSE equalization. In Step 1, we consider ICI as an additional noise term. In Step 2, we reduce the ICI at pilot positions and finally, in Step 3, we treat all estimated data symbols, obtained from Step 2, as if they were pilot symbols. We evaluate the Bit Error Ratio (BER) of our ICI mitigation technique by means of simulation and testbed measurements (up to 400 km/h). In both cases, we achieve a BER close to perfect channel knowledge and zero ICI.**

I. Introduction

Orthogonal Frequency Division Multiplexing (OFDM) is currently employed by many wireless communication standards, such as LTE, DAB and WIFI 802.11, because it can efficiently deal with frequency-selective channels, caused by multi-path delays. However, in high mobility scenarios, as they appear for example in car/train to infrastructure communication by LTE, the channel is additionally time-selective, i.e., it changes significantly within one OFDM symbol, so that subcarriers are no longer orthogonal, leading to Inter-Carrier Interference (ICI). The signal detection then becomes more challenging and requires an accurate estimation of the time-variant channel, typically based on pilot symbols. We can, in principle, distinguish between channel estimation methods utilizing either one OFDM symbol [1] or multiple OFDM symbols [2]–[4]. The first method usually employs a basis expansion model to reduce the number of unknown variables and requires a clustered pilot symbol structure, which is not compliant with most standards. The second method, on the other hand, initially ignores ICI and interpolates the estimated time-averaged channel impulse responds of several OFDM symbols, e.g., linearly [2], by Least Squares (LS) polynomial fitting [3] or by LS discrete prolate spheroidal fitting [4]. Once the channel is estimated, the ICI at pilot positions can be reduced in order to increase the accuracy of an iterative channel estimation. We follow such second approach, based on multiple OFDM symbols. By assuming a continuous transmit signal, a discrete receiver structure and a sufficient long Cyclic Prefix (CP), the sampled time-variant transfer function straightforwardly describes the OFDM transmission. Compared with the usually considered impulse response, which includes frequency components not relevant for data transmission, the transfer function allows to ignore these irrelevant frequency components. In contrast to [2]–[4], we interpolate the estimated channel of several OFDM symbols according to the Minimum Mean Squared Error (MMSE) criteria, thus obtaining the MMSE channel estimation of the sampled time-variant transfer function. MMSE channel estimation was investigated in [5] and [6]. However, they consider only one OFDM symbol and their estimation process includes ICI, making it rather complex. We, on the other hand, treat ICI as an additional noise term which is then stepwise reduced by our iterative ICI mitigation technique that combines channel estimation, equalization and ICI cancellation. Iterative channel estimation was for example discussed in [7] where they model the channel variations by a polynomial basis expansion and use the estimated data symbols to estimate the ICI. We, however, use the estimated data symbols only to cancel ICI and, in the final iteration step, as pilot symbols. Many authors [8]–[10] proposed low-complexity equalizers which exploit the underlying structure of the ICI. However, the main focus of this paper is channel estimation so that we employ an ordinary full block MMSE equalizer [11]. For our measurements, we utilize the Vienna Wireless Testbed in combination with a rotation wheel unit [12], i.e, the receive antenna rotates around a central pivot, allowing high velocity measurements of up to 400 km/h.

II. System Model

In our OFDM transmission model, $x_{l,k} \in \mathbb{C}$ denotes the transmitted, unit-power data-symbol at subcarrier-position l ($l = 1, 2 \ldots, L$) and time-position k ($k = 1, 2 \ldots, K$), chosen from a Gray coded Quadrature Amplitude Modulation (QAM) signal constellation. We denote T_s as the length of one OFDM symbol in the time domain. For the interval $(k-1)T_\mathrm{s} \leq t < kT_\mathrm{s}$, the continuous baseband signal in the time domain can be written as:

$$s(t) = \frac{1}{\sqrt{T}} \sum_{l=1}^{L} x_{l,k} \mathrm{e}^{\mathrm{j}2\pi l \Delta f (t - kT_\mathrm{cp})}, \tag{1}$$

where Δf denotes the subcarrier spacing, T_cp the length of the CP and T the length of one OFDM symbol without CP, i.e., $T = T_\mathrm{s} - T_\mathrm{cp} = 1/\Delta f$. Convolution of the transmitted signal $s(t)$ with the time-variant channel impulse response $\tilde{h}(\tau, t)$ delivers the received signal $r(t)$ in the time domain, given for the interval $(k-1)T_\mathrm{s} + T_\mathrm{cp} \leq t < kT_\mathrm{s}$ by:

$$r(t) = \frac{1}{\sqrt{T}} \sum_{l=1}^{L} x_{l,k} \mathrm{e}^{\mathrm{j}2\pi l \Delta f (t - kT_\mathrm{cp})} \underbrace{\int_0^{\tau_\mathrm{max}} \tilde{h}(\tau, t) \mathrm{e}^{-\mathrm{j}2\pi l \Delta f \tau} \, \mathrm{d}\tau}_{H(l\Delta f, t)}. \tag{2}$$

We assumed that the CP is sufficient long, so that OFDM symbol $k-1$ does not influence Equation (2) and also means that the time-variant transfer function $H(f,t)$ straightforwardly describes the channel behavior. For the demodulation step, we sample the received signal $r(t)$ with $\Delta t = \frac{T}{N}$ and apply a discrete Fourier transformation, so that the received data symbol $y_{l,k}$ can be obtained as:

$$y_{l,k} = \frac{\Delta t}{\sqrt{T}} \sum_{n=0}^{N-1} r(n\Delta t + (k-1)T_\mathrm{s} + T_\mathrm{cp})\, \mathrm{e}^{-\mathrm{j}2\pi \frac{ln}{N}}, \quad (3)$$

Placing the transmitted data symbols $x_{l,k}$ and the received data symbols $y_{l,k}$ in a vector $\mathbf{y}_k = [\ y_{1,k} \quad \dots \quad y_{L,k}\ ]^T \in \mathbb{C}^{L\times 1}$, respectively $\mathbf{x}_k = [\ x_{1,k} \quad \dots \quad x_{L,k}\ ]^T \in \mathbb{C}^{L\times 1}$ and including a noise vector $\mathbf{z}_k \sim \mathcal{CN}(0, P^\mathrm{Z}\mathbf{I}_L)$, we obtain the transmission model in matrix notation as:

$$\mathbf{y}_k = \mathbf{D}_k \mathbf{x}_k + \mathbf{z}_k, \quad (4)$$

where the l-th row and d-th column of the OFDM matrix $\mathbf{D}_k \in \mathbb{C}^{L\times L}$ is given by:

$$[\mathbf{D}_k]_{l,d} = \frac{1}{N} \sum_{n=0}^{N-1} H[d, n + kN_\mathrm{s} - N + 1]\, \mathrm{e}^{-\mathrm{j}2\pi \frac{l-d}{N} n}. \quad (5)$$

The discrete function $H[l,m]$ is the sampled (in both frequency and time) time-variant transfer function $H(f,t)$ and defined for $m = 1, 2, \dots N_\mathrm{s}K$ (see Figure 1). Furthermore, we assume unit power Rayleigh fading, i.e., $H[l,m]$ is jointly complex Gaussian with zero mean and unit variance. If the channel changes within one OFDM symbol, matrix $\mathbf{D}_k$ is no longer a diagonal matrix so that different subcarrier interfere each other, which is called ICI. In contrast to the time-invariant case (for $N \geq L$), the performance also depends on the number of samples N due to aliasing in combination with non-orthogonality. The diagonal elements of the OFDM matrix $\mathbf{D}_k$ correspond to the piecewise time average of the channel $H[l,m]$, are denoted by $\bar{\mathbf{h}}_k \in \mathbb{C}^{L\times 1}$ and characterize the desired signal components, while the non-diagonal elements describe the ICI. We thus decompose Equation (4) in a signal, an ICI and a noise part:

$$\mathbf{y}_k = \mathrm{diag}\{\bar{\mathbf{h}}_k\}\mathbf{x}_k + \mathbf{y}_k^\mathrm{ICI} + \mathbf{z}_k \quad (6)$$

$$y_{l,k} = \bar{h}_{l,k} x_{l,k} + y_{l,k}^\mathrm{ICI} + z_{l,k}, \quad (7)$$

whereas the signal power is given by $P_{l,k}^\mathrm{S} = \mathbb{E}\{\bar{h}_{l,k}\bar{h}_{l,k}^*\}$, the ICI power by $P_{l,k}^\mathrm{ICI} = \mathbb{E}\{y_{l,k}^\mathrm{ICI}(y_{l,k}^\mathrm{ICI})^*\}$ and the noise power by $P^\mathrm{Z} = \mathbb{E}\{z_{l,k}z_{l,k}^*\}$. For simplicity, we assume that the ICI terms $y_{l,k}^\mathrm{ICI}$ are uncorrelated and Gaussian distributed, even if the real distribution is given by a weighted Gaussian mixture due to random data symbols [13].

III. ICI Power

In order to calculate the ICI power, we define a new vector $\mathbf{h}_k \in \mathbb{C}^{LN\times 1}$ in which we stack all elements of the sampled time-variant transfer function, relevant to OFDM symbol k,

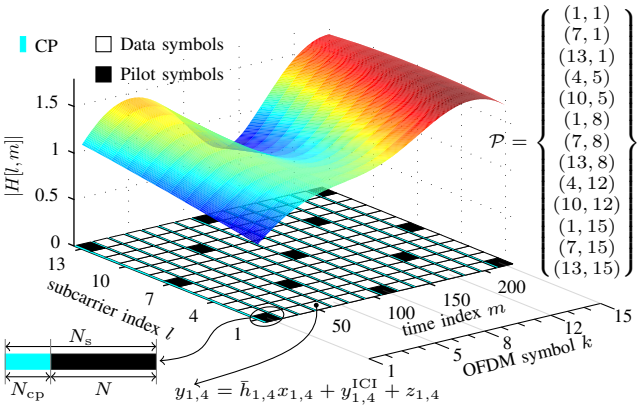

Fig. 1. A possible realization of the sampled time-variant transfer function $H[l,m]$ for $L = 13$, $K = 15$, $N = 13$ and $N_\mathrm{cp} = 2$. The data symbols $x_{l,k}$ are transmitted over a $L \times K$ frequency-time grid, whereas $\bar{h}_{l,k}$ denotes the piecewise time average of the channel $H[l,m]$. The pilot symbol positions are chosen according to the LTE standard.

according to:

$$\mathbf{h}_k = \begin{bmatrix} H[1, kN_\mathrm{s} + N_\mathrm{cp} + 1] \\ \vdots \\ H[L, kN_\mathrm{s} + N_\mathrm{cp} + 1] \\ H[1, kN_\mathrm{s} + N_\mathrm{cp} + 2] \\ \vdots \\ H[L, kN_\mathrm{s} + N_\mathrm{cp} + N] \end{bmatrix}. \quad (8)$$

We can then find a sparse matrix $\mathbf{W}_l \in \mathbb{C}^{L\times LN}$ ($100\frac{L-1}{L}$ percent of the elements are zero), so that Equation (5) can be rewritten in matrix notation as:

$$\mathbf{D}_k = \begin{bmatrix} (\mathbf{W}_1\mathbf{h}_k)^\mathrm{T} \\ \vdots \\ (\mathbf{W}_L\mathbf{h}_k)^\mathrm{T} \end{bmatrix}, \quad (9)$$

with $\mathbf{W}_l$ $(l = 1, 2 \dots, L)$ given by:

$$\mathbf{W}_l = \frac{1}{N} \exp\left\{ -\frac{\mathrm{j}2\pi}{N} \left(\mathbf{1}_{1\times N} \otimes \begin{bmatrix} l-1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & l-L \end{bmatrix} \right) \circ \right. \cdot \left. \circ ([0 \quad \cdots \quad N-1] \otimes \mathbf{I}_L) \right\} \circ (\mathbf{1}_{1\times N} \otimes \mathrm{diag}\{\mathbf{1}_{1\times L}\}) \quad (10)$$

Here, the $\exp\{\cdot\}$ operator is applied at each element of the matrix, the operations $\otimes$ and $\circ$ are the Kronecker product, respectively the Hadamard (point-wise) product, $\mathbf{I}_L$ denotes the identity matrix of size L and $\mathbf{1}_{1\times N}$ an all-one vector of size $1 \times N$.

Let us further split the matrix $\mathbf{W}_l$ into a signal part $\mathbf{W}_l^\mathrm{S}$ and an ICI part $\mathbf{W}_l^\mathrm{ICI}$:

$$\mathbf{W}_l = \mathbf{W}_l^\mathrm{S} + \mathbf{W}_l^\mathrm{ICI}. \quad (11)$$

Matrix $\mathbf{W}_l^\mathrm{ICI}$ consists of nearly the same elements as $\mathbf{W}_l$, with the difference that the l-th row is set to zero.

The signal power $P_{l,k}^\mathrm{S}$ and the ICI power $P_{l,k}^\mathrm{ICI}$ are then given by:

$$P_{l,k}^\mathrm{S} = \mathrm{tr}\left\{ \mathbf{W}_l^\mathrm{S} \mathbf{R}_{\mathbf{h}_k} \left(\mathbf{W}_l^\mathrm{S}\right)^\mathrm{H} \right\} \quad (12)$$

$$P_{l,k}^\mathrm{ICI} = \mathrm{tr}\left\{ \mathbf{W}_l^\mathrm{ICI} \mathbf{R}_{\mathbf{h}_k} \left(\mathbf{W}_l^\mathrm{ICI}\right)^\mathrm{H} \right\}, \quad (13)$$

whereas $\mathbf{R}_{\mathbf{h}_k} = \mathbb{E}\{\mathbf{h}_k \mathbf{h}_k^{\mathrm{H}}\} \in \mathbb{C}^{LN \times LN}$ denotes the correlation matrix (see Section VII for an example). For a Wide-Sense Stationary Uncorrelated Scattering (WSSUS) channel, the signal power is independent of the subcarrier position l and the time position k, i.e., $P^{\mathrm{S}}_{l,k} = P^{\mathrm{S}}$. On the other hand, the ICI power depends on l and k if the number of samples N is larger than L, due to aliasing. For infinitely many subcarriers, we can find analytical expressions for the signal power and the ICI power. For example, assuming a Jakes Doppler spectral density, Equation (12) approaches a generalized hypergeometric function [14].

IV. MMSE Channel Estimation

Our goal is to estimate the channel $H[l,m]$ out of known pilot symbols $\mathbf{x}_{\mathcal{P}} \in \mathbb{C}^{|\mathcal{P}| \times 1}$. The pilot symbol pattern can be arbitrary, e.g., diamond shaped, as defined in the LTE standard (see Figure 1). Similar to Equation (8), we stack all elements of the channel $H[l,m]$ in a large vector $\mathbf{h} \in \mathbb{C}^{LN_{\mathrm{s}}K \times 1}$:

$$\mathbf{h} = [H[1,1] \quad \cdots \quad H[L,1] \quad H[1,2] \quad \cdots \quad H[L,N_{\mathrm{s}}K]]^{\mathrm{T}} \tag{14}$$

The assumption of Rayleigh fading implies that $\mathbf{h}$ is jointly complex Gaussian, i.e., $\mathbf{h} \sim \mathcal{CN}(0, \mathbf{R}_{\mathbf{h}})$. The channel is normalized so that the diagonal elements of $\mathbf{R}_{\mathbf{h}}$ become one.

According to Equation (6) and (7), an LS estimation of the channel (Equation (17)) delivers only the piecewise mean $\bar{h}_{l,k}$, corrupted by ICI and noise. The relationship between the piecewise mean channel $\bar{\mathbf{h}} \in \mathbb{C}^{LK \times 1}$ ($\bar{\mathbf{h}} = \begin{bmatrix} \bar{\mathbf{h}}_1^{\mathrm{T}} & \cdots & \bar{\mathbf{h}}_K^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$) and the channel vector $\mathbf{h} \in \mathbb{C}^{LN_{\mathrm{s}}K \times 1}$ can be described by a sparse matrix $\mathbf{M} \in \mathbb{R}^{LK \times LN_{\mathrm{s}}K}$ ($100[1 - \frac{N}{LN_{\mathrm{s}}K}]$ percent of the elements are zero):

$$\bar{\mathbf{h}} = \mathbf{M}\mathbf{h}. \tag{15}$$

with $\mathbf{M}$ given by:

$$\mathbf{M} = \left(\mathbf{I}_K \otimes \begin{bmatrix} \mathbf{0}_{1 \times N_{\mathrm{cp}}} & \mathbf{1}_{1 \times N} \end{bmatrix} \frac{1}{N} \right) \otimes \mathbf{I}_L. \tag{16}$$

The received data symbols at pilot positions $\mathcal{P}$ are stacked in a vector $\mathbf{y}_{\mathcal{P}} \in \mathbb{C}^{|\mathcal{P}| \times 1}$ and element wise divided by the pilot symbols $\mathbf{x}_{\mathcal{P}}$, so that the LS estimate of the piecewise mean channel $\hat{\bar{\mathbf{h}}}^{\mathrm{LS}}_{\mathcal{P}} \in \mathbb{C}^{|\mathcal{P}| \times 1}$ at pilot positions becomes:

$$\hat{\bar{\mathbf{h}}}^{\mathrm{LS}}_{\mathcal{P}} = \mathrm{diag}\{\mathbf{x}_{\mathcal{P}}\}^{-1} \mathbf{y}_{\mathcal{P}}. \tag{17}$$

An estimation of the channel vector $\mathbf{h}$ can then be found by:

$$\hat{\mathbf{h}} = \mathbf{A}\hat{\bar{\mathbf{h}}}^{\mathrm{LS}}_{\mathcal{P}}. \tag{18}$$

The matrix $\mathbf{A} \in \mathbb{C}^{LN_{\mathrm{s}}K \times |\mathcal{P}|}$ represents a general description of a linear estimation and includes, for example, basis expension models [4] or the MMSE solution (Equation (24)). The correlation matrix $\mathbf{R}_{\hat{\mathbf{h}}} = \mathbb{E}\{\hat{\mathbf{h}}\hat{\mathbf{h}}^{\mathrm{H}}\} \in \mathbb{C}^{LN_{\mathrm{s}}K \times LN_{\mathrm{s}}K}$ of such channel estimation can be written as:

$$\mathbf{R}_{\hat{\mathbf{h}}} = \mathbf{A} \left(\mathbf{M}_{\mathcal{P}} \mathbf{R}_{\mathbf{h}} \mathbf{M}_{\mathcal{P}}^{\mathrm{H}} + \mathrm{diag}\{\mathbf{p}^{\mathrm{ICI}}_{\mathcal{P}}\} + P^{\mathrm{Z}} \mathbf{I}_{|\mathcal{P}|} \right) \mathbf{A}^{\mathrm{H}}, \tag{19}$$

whereas the vector $\mathbf{p}^{\mathrm{ICI}}_{\mathcal{P}} \in \mathbb{R}^{|\mathcal{P}| \times 1}$ consists of all ICI power elements (Equation (13)), corresponding to the correct pilot positions. In order to determine the matrix $\mathbf{M}_{\mathcal{P}} \in \mathbb{R}^{|\mathcal{P}| \times LN_{\mathrm{s}}K}$, we take $|\mathcal{P}|$ rows, at correct pilot positions, out of the matrix $\mathbf{M} \in \mathbb{R}^{LK \times LN_{\mathrm{s}}K}$. The Mean Squared Error (MSE) of our channel estimation is then given as the diagonal elements of the following correlation matrix:

$$\mathbf{R}_{\mathbf{h}-\hat{\mathbf{h}}} = \mathbf{R}_{\mathbf{h}} - \mathbf{A}\mathbf{M}_{\mathcal{P}}\mathbf{R}_{\mathbf{h}} - \mathbf{R}_{\mathbf{h}}\mathbf{M}_{\mathcal{P}}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}} + \mathbf{R}_{\hat{\mathbf{h}}}. \tag{20}$$

For the derivation of the MMSE channel estimation, we consider the i-th element of the channel estimation $\hat{\mathbf{h}}$ individually ($i = l + (k-1)L$). Let us denote the i-th row of matrix $\mathbf{A}$ as $\tilde{\mathbf{a}}_i^{\mathrm{T}} \in \mathbb{R}^{1 \times |\mathcal{P}|}$. Equation (18) then transforms to:

$$[\hat{\mathbf{h}}]_i = \tilde{\mathbf{a}}_i^{\mathrm{T}} \hat{\bar{\mathbf{h}}}^{\mathrm{LS}}_{\mathcal{P}} \tag{21}$$

$$= \tilde{\mathbf{a}}_i^{\mathrm{T}} \left(\mathbf{M}_{\mathcal{P}} \mathbf{h} + \mathrm{diag}\{\mathbf{x}_{\mathcal{P}}\}^{-1} (\mathbf{y}^{\mathrm{ICI}}_{\mathcal{P}} + \mathbf{z}_{\mathcal{P}}) \right). \tag{22}$$

The orthogonal projection theorem requires:

$$\mathbb{E}\left\{ ([\hat{\mathbf{h}}]_i - [\mathbf{h}]_i)[\hat{\mathbf{h}}]_i^* \right\} = 0, \tag{23}$$

so that the MMSE solution $\mathbf{A}^{\mathrm{MMSE}} \in \mathbb{C}^{LN_{\mathrm{s}}K \times |\mathcal{P}|}$ becomes:

$$\mathbf{A}^{\mathrm{MMSE}} = \mathbf{R}_{\mathbf{h}} \mathbf{M}_{\mathcal{P}}^{\mathrm{H}} \left(\mathbf{M}_{\mathcal{P}} \mathbf{R}_{\mathbf{h}} \mathbf{M}_{\mathcal{P}}^{\mathrm{H}} + \mathrm{diag}\{\mathbf{p}^{\mathrm{ICI}}_{\mathcal{P}}\} + P^{\mathrm{Z}} \mathbf{I}_{|\mathcal{P}|} \right)^{-1} \tag{24}$$

Equation (24) together with (18) delivers all $LN_{\mathrm{s}}K$ MMSE estimates of the channel $H[l,m]$ which can be inserted in Equation (5) or (9) to estimate the behavior of our OFDM system. Multiplying matrix $\mathbf{M}$ with Equation (24) results in the MMSE filter matrix $\bar{\mathbf{A}}^{\mathrm{MMSE}} \in \mathbb{C}^{LK \times |\mathcal{P}|}$, which delivers the MMSE estimate of the piecewise mean channel $\hat{\bar{\mathbf{h}}}$:

$$\bar{\mathbf{A}}^{\mathrm{MMSE}} = \mathbf{M}\mathbf{A}^{\mathrm{MMSE}}. \tag{25}$$

This is the conventional two-dimensional MMSE solution [15] where usually the matrices $\mathbf{R}_{\bar{\mathbf{h}}_{\mathcal{P}}} = \mathbf{M}_{\mathcal{P}} \mathbf{R}_{\mathbf{h}} \mathbf{M}_{\mathcal{P}}^{\mathrm{H}} \in \mathbb{C}^{|\mathcal{P}| \times |\mathcal{P}|}$ and $\mathbf{R}_{\bar{\mathbf{h}},\bar{\mathbf{h}}_{\mathcal{P}}} = \mathbf{M}\mathbf{R}_{\mathbf{h}}\mathbf{M}_{\mathcal{P}}^{\mathrm{H}} \in \mathbb{C}^{LK \times |\mathcal{P}|}$ are assumed to be known explicitly. We, on the other hand, showed in this paper how these matrices ($\mathbf{R}_{\bar{\mathbf{h}}_{\mathcal{P}}}$ and $\mathbf{R}_{\bar{\mathbf{h}},\bar{\mathbf{h}}_{\mathcal{P}}}$) can be decomposed to describe OFDM systems over doubly-selective channels.

V. ICI Mitigation

Our ICI mitigation method combines the OFDM matrix estimation in Equation (28), the MMSE equalization in Equation (29) and the ICI cancellation in Equation (30).

Let us denote $\mathbf{A}_k \in \mathbb{C}^{LN \times |\mathcal{P}|}$ as the filter matrix, consisting of those row elements of $\mathbf{A} \in \mathbb{C}^{LN_{\mathrm{s}}K \times |\mathcal{P}|}$ which deliver the estimated channel $\hat{\mathbf{h}}_k$ (see Equation (8)):

$$\hat{\mathbf{h}}_{\mathbf{k}} = \mathbf{A}_k \hat{\bar{\mathbf{h}}}^{\mathrm{LS}}_{\mathcal{P}}. \tag{26}$$

We now denote the column i of $\mathbf{A}_k$ as $\mathbf{a}_k^{(i)} \in \mathbb{C}^{LN \times 1}$:

$$\mathbf{A}_k = \begin{bmatrix} \mathbf{a}_k^{(1)} & \cdots & \mathbf{a}_k^{(|\mathcal{P}|)} \end{bmatrix}. \tag{27}$$

Combining Equation (9), (26) and (27), the estimated OFDM matrix $\hat{\mathbf{D}}_k$ becomes a linear combination of the basis matrices $\mathbf{C}_k^{(i)} \in \mathbb{C}^{L \times L}$:

$$\hat{\mathbf{D}}_k = \sum_{i=1}^{|\mathcal{P}|} \frac{y_{\mathcal{P}_i}}{x_{\mathcal{P}_i}} \underbrace{\begin{bmatrix} \left(\mathbf{W}_1 \mathbf{a}_k^{(i)}\right)^{\mathrm{T}} \\ \vdots \\ \left(\mathbf{W}_L \mathbf{a}_k^{(i)}\right)^{\mathrm{T}} \end{bmatrix}}_{\mathbf{C}_k^{(i)}}. \tag{28}$$

Note that the matrices $\mathbf{C}_k^{(i)}$ can be precomputed. Once we have the estimated OFDM matrix $\hat{\mathbf{D}}_k$, the equalization is performed as follows:

$$\hat{\mathbf{x}}_k = \hat{\mathbf{D}}_k^{\mathrm{H}} \left(\hat{\mathbf{D}}_k \hat{\mathbf{D}}_k^{\mathrm{H}} + P^{\mathrm{Z}} \mathbf{I}_L \right)^{-1} \mathbf{y}_k. \tag{29}$$

For perfect channel knowledge, i.e., $\hat{\mathbf{D}}_k = \mathbf{D}_k$, Equation (29) delivers the MMSE estimation of the transmitted data symbols $\hat{\mathbf{x}}_k$ [11]. However, we will call it MMSE equalization even if the channel estimation is not perfect. Note that the complexity of Equation (29) can be reduced by exploiting the banded structure of the OFDM matrix $\mathbf{D}_k$ [9]. It is also possible to reduce the amount of ICI by means of ICI cancellation [16]:

$$\mathbf{y}_k^{\mathrm{ICICancel}} = \mathbf{D}_k \mathbf{x}_k + \mathbf{z}_k - \hat{\mathbf{D}}_k^{\mathrm{ICI}} Q\left(\hat{\mathbf{x}}_k\right), \tag{30}$$

where matrix $\hat{\mathbf{D}}_k^{\mathrm{ICI}}$ consists of nearly the same elements as $\hat{\mathbf{D}}_k$, with the difference that the diagonal elements are all zero, thus reflecting only the ICI relevant terms. The function $Q(\cdot)$ indicates quantization of the estimated data symbols $\hat{\mathbf{x}}_k$, i.e., hard decision according to the minimum distance criteria. The quantization process is nonlinear, so that the ICI power after cancellation $P_{l,k}^{\mathrm{ICICancel}}$ cannot be straightforwardly calculated. We approximate this power by assuming that the estimated data symbols in Equation (30) are perfectly known $\hat{\mathbf{x}}_k = \mathbf{x}_k$. The ICI power after cancellation can then be found as:

$$P_{l,k}^{\mathrm{ICICancel}} = \mathrm{tr}\left\{ \mathbf{W}_l^{\mathrm{ICI}} \mathbf{R}_{\mathbf{h}_k - \hat{\mathbf{h}}_k} \left(\mathbf{W}_l^{\mathrm{ICI}} \right)^{\mathrm{H}} \right\}, \tag{31}$$

whereas the correlation matrix $\mathbf{R}_{\mathbf{h}_k - \hat{\mathbf{h}}_k}$ is given by Equation (20). The three basic steps of our ICI mitigation method are:

Step 1: We estimate the OFDM matrix $\hat{\mathbf{D}}_k$ according to Equation (28) in combination with (24) whereby the ICI power $\mathbf{p}_{\mathcal{P}}^{\mathrm{ICI}}$ is given by Equation (13). The MMSE equalization, Equation (29), then delivers the estimated data symbols $\hat{\mathbf{x}}_k$ of the first step.

Step 2: The estimated data symbols $\hat{\mathbf{x}}_k$ together with the estimated matrix $\hat{\mathbf{D}}_k$, both obtained from the first step, reduce the ICI at pilot position, see Equation (30), so that the piecewise mean LS channel estimates at pilot positions $\hat{\bar{\mathbf{h}}}_{\mathcal{P}}^{\mathrm{LS}}$ become more accurate: $\hat{\bar{\mathbf{h}}}_{\mathcal{P}}^{\mathrm{LS}} = \mathrm{diag}\{\mathbf{x}_{\mathcal{P}}\}^{-1} \mathbf{y}_{\mathcal{P}}^{\mathrm{ICICancel}}$. This improved channel estimation is then used in Equation (28) together with (24) to estimate the OFDM matrix $\hat{\mathbf{D}}_k$. Compared with the first step, the ICI power $\mathbf{p}_{\mathcal{P}}^{\mathrm{ICI}}$ in Equation (24) is now lower and given by Equation (31). After MMSE equalization, Equation (29), we obtain the estimated data symbols $\hat{\mathbf{x}}_k$ of the second step.

Step 3: Here, we treat all estimated data symbols $\hat{\mathbf{x}}_k$ as if they were pilot symbols. After ICI cancellation, using $\hat{\mathbf{x}}_k$ and $\hat{\mathbf{D}}_k$ from the second step, see Equation (30), we obtain the piecewise mean LS channel estimates at all LK positions by: $\hat{\bar{\mathbf{h}}}^{\mathrm{LS}} = \mathrm{diag}\{\hat{\mathbf{x}}\}^{-1} \mathbf{y}^{\mathrm{ICICancel}}$. In this step, we assume LK "pilot symbols", so that matrix $\mathbf{M}_{\mathcal{P}}$ in Equation (24) becomes $\mathbf{M}$ and $|\mathcal{P}|$ in Equation (28) becomes LK. Additionally, the ICI power $\mathbf{p}_{\mathcal{P}}^{\mathrm{ICI}}$ in Equation (24) is further decreased compared to Step 2, because the ICI cancellation uses more accurate estimations of the OFDM matrix. Note that, if there exist data symbols of nonunit magnitude, e.g., 16-QAM, the ICI power and noise power are enhanced due to the LS estimation process. Again, MMSE estimation of the OFDM matrix, Equation (24) and (28), in combination with MMSE equalization, Equation (28), delivers the estimated data symbols $\hat{\mathbf{x}}_k$ of the third step.

VI. Bit Error Probability

We evaluate the performance of our ICI mitigation technique by means of uncoded Bit Error Ratio (BER) and compare it to conventional methods, i.e., zero-forcing one-tap equalizers, given by:

$$\hat{\mathbf{x}} = \mathrm{diag}\{\hat{\bar{\mathbf{h}}}\}^{-1} \mathbf{y}, \tag{32}$$

In Rayleigh fading channels, the one-tap equalizer results in a complex Gaussian ratio so that there exists a closed form solution for the cumulative distribution function. Using this fact, we derived in [17] a closed-form solution for the Bit Error Probability (BEP) under arbitrary linear channel estimation $\bar{\mathbf{A}}$. For example, the BEP for the MMSE channel estimation $\bar{\mathbf{A}}^{\mathrm{MMSE}}$, Equation (25), becomes ($i = l + (k-1)L$):

$$\mathrm{BEP}_{l,k}^{4\mathrm{QAM,MMSE}} = \frac{1}{2} - \frac{1}{2\sqrt{2 \frac{P_{l,k}^{\mathrm{S}} + P_{l,k}^{\mathrm{ICI}} + P^{\mathrm{Z}}}{\left[\bar{\mathbf{A}}^{\mathrm{MMSE}} \mathbf{M}_{\mathcal{P}} \mathbf{R}_{\mathbf{h}} \mathbf{M}^{\mathrm{H}}\right]_{i,i}} - 1}}. \tag{33}$$

For perfect channel knowledge and no ICI, the zero-forcing one-tap equalizer corresponds also to the Maximum Likelihood (ML) detection and outperforms MMSE equaliziation in terms of BEP (but not in terms of MSE). For 4-QAM, however, these two equalizers show the same performance. Its BEP is given by:

$$\mathrm{BEP}_{l,k}^{4\mathrm{QAM,perfect,NoICI}} = \frac{1}{2} - \frac{1}{2\sqrt{1 + 2\frac{P^{\mathrm{Z}}}{P_{l,k}^{\mathrm{S}}}}}, \tag{34}$$

which corresponds to the lowest possible BEP that can be achieved (ignoring the small loss of signal power to adjacent subcarriers).

VII. Numerical Results

As illustrated in Figure 1, we use 13 subcarriers, 15 OFDM symbols and an LTE pilot symbol pattern. Furthermore, we assume 4-QAM, a subcarrier spacing of 15 kHz and a carrier frequency of 2.5 GHz.

A. Simulations

In order to obtain a sampled time-variant transfer function $H[l,m]$, as in Equation (5), we utilize the following WSSUS channel model:

$$H[l,m] = \frac{1}{\sqrt{I}} \sum_{i=1}^{I} \mathrm{e}^{\mathrm{j}\left(\theta_i + 2\pi \frac{\nu_i}{\Delta f} \frac{m}{N} - 2\pi \frac{\tau_i}{T} l\right)}, \tag{35}$$

whereas the random Doppler shifts ν_i are distributed according to a Jakes Doppler spectral density, the random normalized delays $\frac{\tau_i}{T}$ are uniformly distributed between 0 and $\frac{\tau_{\max}}{T} = 0.01$ and the random phases θ_i are uniformly distributed between 0 and 2π, ensuring circularly symmetry. We assume $N = 13$ samples, $I = 200$ channel paths and consider 50 000 realizations over which we average. Additionally, all random variables are assumed to be statistically independent, so that the frequency-time correlation function becomes:

$$\mathbb{E}\left\{H[l_1,m_1] H^*[l_2,m_2]\right\} = r_{H_f}[l_1 - l_2] r_{H_t}[m_1 - m_2]. \tag{36}$$

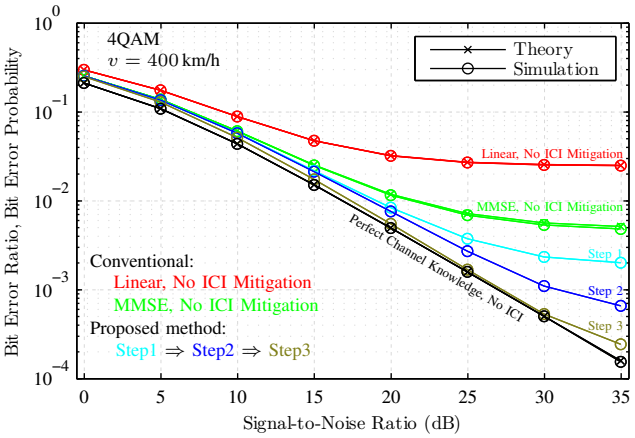

Fig. 2. Simulated BER and theoretical BEP as a function of SNR. Simulation confirms our analytical prediction. At high SNR, our ICI mitigation method avoids an early saturation of the BER. Each step improves the BER and Step 3 achieves a performance close to perfect channel knowledge and no ICI.

with (J_0 is the zero-th order Bessel function):

$$r_{H_t}[m_1 - m_2] = \mathrm{J}_0\left(2\pi \frac{\nu_{\max}}{\Delta f}\frac{m_1 - m_2}{N}\right) \tag{37}$$

$$r_{H_f}[l_1 - l_2] = \mathrm{sinc}\left(\frac{\tau_{\max}}{T}(l_1 - l_2)\right) \mathrm{e}^{\mathrm{j}\pi \frac{\tau_{\max}}{T}(l_1 - l_2)}. \tag{38}$$

If the channel elements are stacked according to Equation (14), correlation matrix $\mathbf{R_h} \in \mathbb{C}^{LN_\mathrm{s}K \times LN_\mathrm{s}K}$ becomes:

$$\mathbf{R_h} = \begin{bmatrix} r_{H_t}[0] & \cdots & r_{H_t}[(KN_\mathrm{s}-1)] \\ \vdots & \ddots & \vdots \\ r_{H_t}[-(KN_\mathrm{s}-1)] & \cdots & r_{H_t}[0] \end{bmatrix} \otimes \begin{bmatrix} r_{H_f}[0] & \cdots & r_{H_f}[(L-1)] \\ \vdots & \ddots & \vdots \\ r_{H_f}[-(L-1)] & \cdots & r_{H_f}[0] \end{bmatrix}. \tag{39}$$

Figure 2 shows that our theoretical BEP expressions for one-tap equalizers (Section VI) coincide with the simulated BER. The illustrated conventional methods ignore ICI and estimate the channel either by MMSE (ICI is considered as additional noise) or by linear interpolation of the three closest LS estimates at pilot position (no statistical knowledge necessary). For a high Signal-to-Noise Ratio (SNR), the ICI becomes the limiting factor so that the BEP of conventional methods soon saturates. This effect can be mitigated by our proposed ICI mitigation technique. The BER after the third step comes very close to perfect channel knowledge and no ICI which corresponds to the lowest possible BER.

Figure 3 shows the BER relative to the BER of Step 3, thereby delivering the improvement factor of our proposed method. Clearly, for higher velocities, the ICI becomes more and more an impediment so that the performance of our ICI mitigation technique increases relatively to conventional methods. For time-invariant channels, i.e., a velocity of zero, Step 1 and Step 2 exhibit the same BER as the one-tap equalizer using MMSE channel estimation (no ICI mitigation).

B. Measurements

Our measurements were conducted on the Vienna Wireless Testbed that has been augmented by a rotating wheel unit to

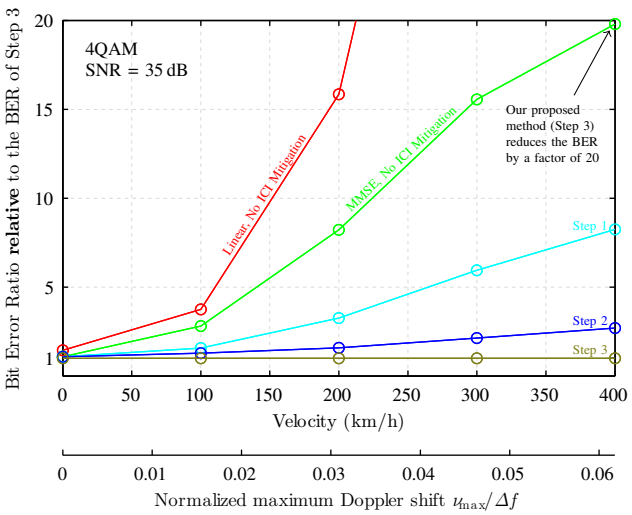

Fig. 3. Simulated BER **relative** to the BER of Step 3 as a function of velocity, respectively normalized maximum Doppler shift. The higher the velocity, the better our methods performs compared to conventional methods. For a velocity of zero, the BER of MMSE (no ICI mitigation), Step 1 and Step 2 coincide.

generate high speed movements [12]. Transmitter and receiver are both placed indoors, in different rooms, and the number of samples N after demodulation is set to 26. For a given transmit power, corresponding to a specific average SNR, we obtain different channel realizations by relocating the transmit antenna to 144 positions, equidistantly spaced over a 3×3 wavelength grid. The receive antenna moves at up to 400 km/h and we measure always over the same 3 wavelength movement interval. To estimate the noise power at the receiver, we include two all-zero OFDM symbols, $\mathbf{x}_0 = \mathbf{0}_{L\times 1}$ and $\mathbf{x}_{K+1} = \mathbf{0}_{L\times 1}$. Similar, the ICI-plus-noise power is estimated by using all-zero subcarriers, $x_{0,k} = 0$ and $x_{L+1,k} = 0$. Because the second order statistics, required for the MMSE estimation, are not perfectly known, we obtain only a mismatched solution. We assume a WSSUS channel and a separable correlation function, as in Equation (36). For the MMSE estimation, we cannot directly insert the estimated correlation functions because they are corrupted by noise, leading to large errors due to the matrix inversion. We thus model the delays as well as the Doppler shifts by a uniform distribution, so that the frequency correlation is given by Equation (38) and the time correlation by:

$$r_{H_t}[m_1 - m_2] = \mathrm{sinc}\left(2\frac{\nu_{\max}}{\Delta f}\frac{m_1 - m_2}{N}\right). \tag{40}$$

From the estimated frequency correlation function $\hat{r}_{H_f}$ we conclude that $\tau_{\max} = 0$. The maximum Doppler shifts $\nu_{\max}$, on the other hand, are chosen so that the theoretical ICI power and the measured ICI power coincide. At a velocity of 100 km/h, we measure the same ICI power we would expect at 98 km/h for a uniform distributed Doppler spectral density, so that the normalized maximum Doppler shift in Equation (40) is set to $\frac{\nu_{\max}}{\Delta f} = 0.015$. Similar, the measurement velocity of 200 km/h corresponds to an equivalent (uniform Doppler spectral density) velocity of 187 km/h, 300 km/h to 280 km/h and 400 km/h to 372 km/h.

Figure 4 shows that our theoretical expressions for one-tap equalizers (Section VI) accurately model the true physical

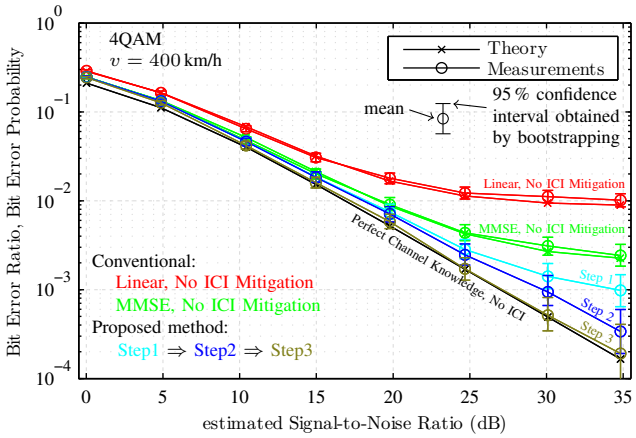

Fig. 4. Measured BER and theoretical BEP as a function of estimated SNR, similar to Figure 2. Also in testbed measurements, our proposed ICI mitigation methods comes close to perfect channel knowledge and no ICI.

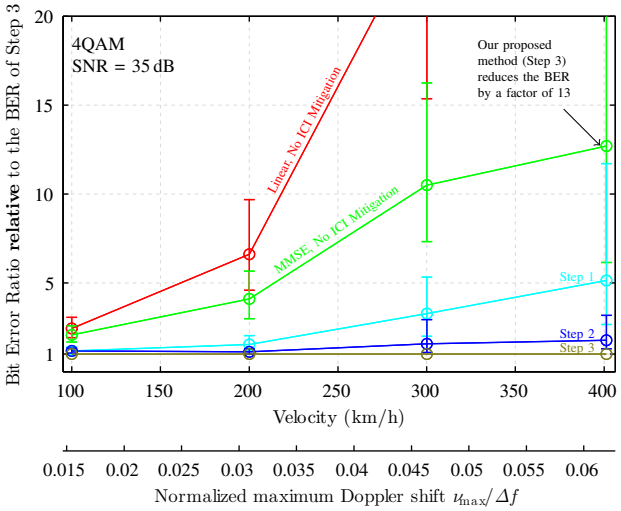

Fig. 5. Measured BER **relative** to the BER of Step 3 as a function of velocity. Compared to Figure 3, the potential improvement of our ICI mitigation method is lower because the underlying channel parameters differ between simulation and measurements (lower delay spread due to indoor measurements, lower Doppler spread and $N > L$).

behavior. Each step of our proposed mitigation technique reduces the BER whereas the third step achieves a performance close to perfect channel knowledge and no ICI.

Figure 5 illustrates the measured BER relative to the BER of Step 3. Compared to Figure 3, the one-tap equalizers perform better due to different channel parameters (lower delay spread, lower Doppler spread and $N > L$), lowering the potential improvement of our ICI mitigation method.

VIII. Conclusion

We derived the MMSE channel estimation of the sampled time-variant transfer function by exploiting the channel correlation of several OFDM symbols. We then proposed an iterative three step ICI mitigation technique based on our MMSE channel estimation, MMSE equalization and ICI cancellation. Simulation as well as testbed measurements show that our proposed method achieves a BER close to perfect channel knowledge cancelling perfectly the ICI.

Acknowledgment

This work has been funded by the Christian Doppler Laboratory for Wireless Technologies for Sustainable Mobility, the A1 Telekom Austria AG, and the KATHREIN-Werke KG. The financial support by the Federal Ministry of Economy, Family and Youth and the National Foundation for Research, Technology and Development is gratefully acknowledged.

References

[1] Z. Tang, R. C. Cannizzaro, G. Leus, and P. Banelli, "Pilot-assisted time-varying channel estimation for OFDM systems," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 2226–2238, 2007.

[2] Y. Mostofi and D. C. Cox, "ICI mitigation for pilot-aided OFDM mobile systems," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 2, pp. 765–774, 2005.

[3] H. Hijazi and L. Ros, "Polynomial estimation of time-varying multipath gains with intercarrier interference mitigation in OFDM systems," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 1, pp. 140–151, 2009.

[4] M. Simko, C. Mehlführer, T. Zemen, and M. Rupp, "Inter-carrier interference estimation in MIMO OFDM systems with arbitrary pilot structure," in *IEEE Vehicular Technology Conference (VTC Spring), 2011*.

[5] C. Sgraja and J. Lindner, "Estimation of rapid time-variant channels for OFDM using Wiener filtering," in *IEEE International Conference on Communications (ICC)*, vol. 4, May 2003, pp. 2390–2395 vol.4.

[6] C. Shin, J. Andrews, and E. Powers, "An efficient design of doubly selective channel estimation for OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 6, no. 10, pp. 3790–3802, October 2007.

[7] S. Tomasin, A. Gorokhov, H. Yang, and J.-P. Linnartz, "Iterative interference cancellation and channel estimation for mobile OFDM," *IEEE Transactions on Wireless Communications*, vol. 4, no. 1, pp. 238–245, 2005.

[8] P. Schniter, "Low-complexity equalization of OFDM in doubly selective channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 4, pp. 1002–1011, 2004.

[9] L. Rugini, P. Banelli, and G. Leus, "Simple equalization of time-varying channels for OFDM," *IEEE Communications Letters*, vol. 9, no. 7, pp. 619–621, July 2005.

[10] K. Fang, L. Rugini, and G. Leus, "Low-complexity block turbo equalization for OFDM systems in time-varying channels," *IEEE Transactions on Signal Processing*, vol. 56, no. 11, pp. 5555–5566, 2008.

[11] Y.-S. Choi, P. Voltz, and F. Cassara, "On channel estimation and detection for multicarrier signals in fast and selective Rayleigh fading channels," *IEEE Transactions on Communications*, vol. 49, no. 8, pp. 1375–1387, Aug 2001.

[12] R. Nissel, M. Lerch, and M. Rupp, "Experimental validation of the OFDM bit error probability for a moving receive antenna," in *IEEE Vehicular Technology Conference (VTC Fall)*, Sept 2014.

[13] T. Wang, J. Proakis, E. Masry, and J. Zeidler, "Performance degradation of OFDM systems due to Doppler spreading," *IEEE Transactions on Wireless Communications*, vol. 5, no. 6, pp. 1422–1432, June 2006.

[14] P. Robertson and S. Kaiser, "The effects of Doppler spreads in OFDM(A) mobile radio systems," in *IEEE 50th Vehicular Technology Conference (VTC Fall)*, 1999, pp. 329–333, vol.1.

[15] P. Hoeher, S. Kaiser, and P. Robertson, "Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 1845–1848 vol.3, 1997.

[16] A. Molisch, M. Toeltsch, and S. Vermani, "Iterative methods for cancellation of intercarrier interference in OFDM systems," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 4, pp. 2158–2167, July 2007.

[17] R. Nissel, M. Lerch, M. Šimko, and M. Rupp, "Bit error probability for pilot-symbol-aided OFDM channel estimation in doubly-selective channels," in *18th International ITG Workshop on Smart Antennas (WSA 2014)*, Erlangen, Germany, Mar. 2014.