

PILOT-ASSISTED TIME-VARYING OFDM CHANNEL ESTIMATION BASED ON MULTIPLE OFDM SYMBOLS

Zijian Tang, Geert Leus*

Delft University of Technology - Fac. EEMCS
Mekelweg 4, 2628 CD Delft, The Netherlands
{tang, leus}@cas.et.tudelft.nl

Paolo Banelli

University of Perugia - DIEI
Via G. Duranti 93, 06125 Perugia, Italy
banelli@diei.unipg.it

ABSTRACT

In this paper, we deal with channel estimation for Orthogonal Frequency-Division Multiplexing (OFDM) systems. The channels are assumed to be Time-Varying (TV) and approximated by a Basis Expansion Model (BEM). Due to the time-variation, the resulting channel matrix in the frequency domain is no longer diagonal, but can be approximated as banded. Based on this band approximation, we propose a channel estimator that can combat both the additive noise and the out-of-band interference. Compared to our previous work, the proposed channel estimator can span multiple OFDM symbol intervals such that more time-correlation information can be explored to improve the estimation accuracy.

keywords: OFDM, BEM, time-varying channels, pilot-assisted modulation.

1. INTRODUCTION

In mobile communications, a high mobile speed causes the carrier frequency to spread out and leads to Time-Varying (TV) channels. Basis Expansion Models (BEM) can be used to approximate the time-variation within a certain observation window. Examples of such BEMs are the Complex Exponential BEM (CE-BEM) in [1], the Generalized CE-BEM (GCE-BEM) in [2], the Discrete Prolate Spheroidal BEM (DPS-BEM) in [3], the Polynomial BEM (P-BEM) in [4], and the Karhunen-Loeve BEM (KL-BEM) in [5].

Focusing on the estimation of channels that are modeled by a BEM, we basically only need to estimate the BEM coefficients. [6] proposes pilot-assisted estimators based on a CE-BEM assumption, where pilots are clustered in the time-domain to combat the Inter-Symbol Interference (ISI). Likewise, for OFDM systems, it is also useful to cluster the pilots in the frequency-domain to combat the Inter-Carrier Interference (ICI) induced by the Doppler spread [7]. Indeed, as pointed out in [8], most ICI is concentrated in adjacent subcarriers, which implies that the channel matrix in the frequency-domain is roughly banded, a situation comparable to the channel in the time-domain. Based on this (banded channel) observation, [7, 9] propose channel estimators that also employ clustered pilots in the frequency-domain, with the difference that [7] relies exclusively on the CE-BEM and assumes a fixed bandwidth of the channel matrix, while [9] can be applied to any other BEM, and assumes a flexible bandwidth, which can be optimized for any situation.

In this paper, we will extend the results from [9] to the case where multiple OFDM symbols are utilized. This means that not only the target OFDM symbol but also its neighboring OFDM symbols will be invoked for channel estimation. This is beneficial since the Doppler-induced time-variation of the channel is not random but

dictated by a certain statistical model [10]. As a consequence, the channel gains in the neighboring OFDM intervals can still bear a strong correlation with those in the target OFDM interval, which could be helpful to reduce the channel estimator's variance [11]. To make use of this correlation information in the context of a BEM, we need to employ a bigger BEM window to account for several adjacent OFDM intervals. By doing so, we can take more pilots into account without reducing the bandwidth efficiency. On the other hand, a larger BEM window should entail more basis functions to maintain the same modeling capability, which implies that more BEM coefficients need to be estimated. By simulations, we show that for a given bandwidth efficiency, increasing the BEM window has an overall positive effect on estimation performance.

Notation: We use upper (lower) bold face letters to denote matrices (column vectors). $(\cdot)^T$ and $(\cdot)^H$ represent transpose and complex conjugate transpose (Hermitian), respectively. $\mathcal{E}_{\mathbf{x}}\{\cdot\}$ stands for the expected value with respect to \mathbf{x} . \otimes represents the Kronecker product. \dagger represents the pseudo inverse. \mathbf{I}_N stands for the $N \times N$ identity matrix. $\mathbf{1}_{M \times N}$ and $\mathbf{0}_{M \times N}$ stand for the $M \times N$ all-one and all-zero matrix, respectively. Further, we use $[\mathbf{x}]_p$ to indicate the $(p+1)$ st element of the vector \mathbf{x} , and $[\mathbf{X}]_{p,q}$ to indicate the $(p+1, q+1)$ st entry of the matrix \mathbf{X} .

2. OFDM SYSTEM AND BEM

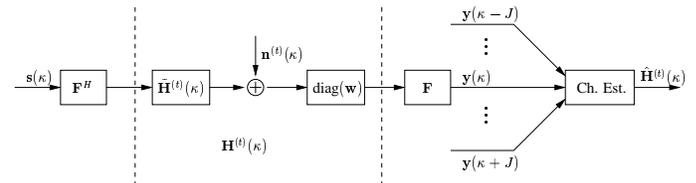


Fig. 1. The transceiver block diagram for channel estimation

Let us consider an OFDM system with N subcarriers with its transceiver block diagram sketched in Fig. 1. For the n th OFDM symbol, the information symbols $\mathbf{s}(n)$ are first modulated on N subcarriers as $\mathbf{s}^{(t)}(n) = \mathbf{F}^H \mathbf{s}(n)$, where \mathbf{F} stands for the N -point unitary Discrete Fourier Transform (DFT) matrix with $[\mathbf{F}]_{p,q} = 1/\sqrt{N} \exp(-j2\pi pq/N)$. Making abstraction of the digital-to-analog and analog-to-digital conversions, $\mathbf{s}^{(t)}(n)$ is next concatenated by a cyclic prefix (CP) of length L_{cp} , sent over the channel, stripped from the CP, and finally demodulated. The received data stream resulting from the n th OFDM symbol can be expressed as

$$\mathbf{y}(n) = \mathbf{F} \mathbf{H}^{(t)}(n) \mathbf{F}^H \mathbf{s}(n) + \mathbf{F} \mathbf{n}^{(t)}(n) = \mathbf{H}(n) \mathbf{s}(n) + \mathbf{n}(n), \quad (1)$$

*This research was supported in part by NWO-STW under the VICI program (DTC.5893) and the VIDI program (DTC.6577).

where $\mathbf{H}^{(t)}(n)$ and $\mathbf{H}(n) := \mathbf{F}\mathbf{H}^{(t)}(n)\mathbf{F}^H$ denote the channel matrix for the n th OFDM symbol in the time-domain and frequency-domain, respectively; $\mathbf{n}^{(t)}(n)$ and $\mathbf{n}(n) := \mathbf{F}\mathbf{n}^{(t)}(n)$ represent the noise in the time-domain and frequency-domain, respectively. Defining $h_{p,l}^{(t)}$ as the time-domain channel gain for the p th time-interval at the l th lag, we assume the channel is of Finite Impulse Response (FIR) with order L , i.e., $h_{p,l}^{(t)} = 0$ for $l < 0$ or $l > L$. Due to the CP, $\mathbf{H}^{(t)}(n)$ has entries $[\mathbf{H}^{(t)}(n)]_{p,q} = h_{n(N+L_{cp})+p+L_{cp}, \text{mod}(p-q, N)}$.

In this paper, we will focus on estimating the channel corresponding to the specific 0th OFDM symbol. A BEM will be adopted to model the channel's variation not only within the 0th OFDM interval, but also its J adjacent OFDM intervals on both sides. Note that the target OFDM symbol is situated in the middle of the BEM window, thereby avoiding possible edge effects due to BEM modeling. More specifically, by collecting all the channel gains within this observation window in a $(2J+1)(L+1)(N+L_{cp}) \times 1$ vector: $\mathbf{h}^{(t)} := [\mathbf{h}^{(t)T}(-J), \dots, \mathbf{h}^{(t)T}(J)]^T$, with

$$\mathbf{h}^{(t)}(n) := [h_{n(N+L_{cp}),0}^{(t)}, \dots, h_{n(N+L_{cp}),L}^{(t)}, \dots, h_{n(N+L_{cp})+N+L_{cp}-1,0}^{(t)}, \dots, h_{n(N+L_{cp})+N+L_{cp}-1,L}^{(t)}]^T, \quad (2)$$

we can adopt the BEM approximation $\mathbf{h}^{(t)} \approx (\mathbf{B} \otimes \mathbf{I}_{L+1})\mathbf{h}$, where $\mathbf{B} := [\mathbf{b}_0, \dots, \mathbf{b}_Q]$ is an $(2J+1)(N+L_{cp}) \times (Q+1)$ matrix that collects $Q+1$ orthonormal basis functions \mathbf{b}_q as columns; the BEM coefficients are collected in the $(L+1)(Q+1) \times 1$ vector $\mathbf{h} := [h_{0,0}, \dots, h_{0,L}, \dots, h_{Q,0}, \dots, h_{Q,L}]^T$, with $h_{q,l}$ standing for the q th BEM coefficient for the l th channel lag, which is obtained in a Least Squares (LS) sense and shall remain invariant within these $2J+1$ OFDM symbol intervals. With aid of the BEM and assuming that the CP length satisfies $L_{cp} \geq L$, we can express the time-domain channel matrix $\mathbf{H}^{(t)}(n)$ as¹ $\mathbf{H}^{(t)}(n) = \sum_{q=0}^Q \text{diag}\{\mathbf{P}_n \mathbf{b}_q\} \mathbf{H}_q^c$, where \mathbf{H}_q^c is an $N \times N$ circulant matrix composed with the q th BEM coefficients $[\mathbf{H}_q^c]_{m,n} = h_{q, \text{mod}(m-n, N)}$; \mathbf{P}_n is the matrix that selects within the BEM window the N elements corresponding to the n th OFDM symbol $\mathbf{P}_n := [\mathbf{0}_{N \times (J+n)(N+L_{cp})+L_{cp}}, \mathbf{I}_N, \mathbf{0}_{N \times (J-n)(N+L_{cp})}]$. By introducing the BEM, (1) becomes

$$\mathbf{y}(n) = \sum_{q=0}^Q \mathbf{D}_q(n) \Delta_q \mathbf{s}(n) + \mathbf{n}(n). \quad (3)$$

In the above equality,

$$\mathbf{D}_q(n) := \mathbf{F} \text{diag}\{\mathbf{P}_n \mathbf{b}_q\} \mathbf{F}^H, \\ \Delta_q := \mathbf{F} \mathbf{H}_q^c \mathbf{F}^H = \text{diag}\{\mathbf{F}_L [h_{q,0}, \dots, h_{q,L}]^T\},$$

where \mathbf{F}_L stands for the first $L+1$ columns of the matrix $\sqrt{N}\mathbf{F}$.

It is noteworthy that (1) can implicitly include the effect of a receiver window: $\mathbf{H}^{(t)}(n) = \text{diag}\{\mathbf{w}\} \tilde{\mathbf{H}}^{(t)}(n)$, where \mathbf{w} is the adopted window and $\tilde{\mathbf{H}}^{(t)}(n)$ represents the original channel (see Fig. 1). Such a receiver window has been recently reported in [12, 13] to improve the performance of low-complexity equalizers that exploit the banded approximation of the frequency-domain channel matrix $\mathbf{H}(n)$. To approximate such a windowed channel, we differentiate between two options in the BEM design. First, if a CE-BEM is considered, we can just stick to the original design for the unwindowed case as in [1]. For the other BEMs, it turns out to be beneficial if we adapt the BEM to the window: more specifically, we

¹From now on, we assume the BEM modeling error is negligible.

design \mathbf{B} as

$$\mathbf{B} := \{\tilde{\mathbf{w}}\} \tilde{\mathbf{B}} \mathbf{Q}, \quad (4)$$

$$\tilde{\mathbf{w}} := \mathbf{1}_{(2J+1) \times 1} \otimes [\mathbf{0}_{L_{cp} \times 1}^T, \mathbf{w}^T]^T, \quad (5)$$

where $\tilde{\mathbf{B}}$ yields one of the traditional BEM designs presented in [2–4] and \mathbf{Q} is a square matrix to make the columns of \mathbf{B} orthonormal. Note that due to $\mathbf{0}_{L_{cp} \times 1}$ in (5), the channel taps corresponding to the CP are in effect discarded.

3. DATA MODEL FOR CHANNEL ESTIMATION

Our task is thus to obtain the BEM coefficient estimates collected in $\hat{\mathbf{h}}$ to recover the time-domain channel related to the 0th OFDM symbol $(\mathbf{P}_0 \mathbf{B} \otimes \mathbf{I}_{L+1}) \hat{\mathbf{h}}$.

Pilots are inserted in dedicated subcarriers in each OFDM symbol. They are grouped in M clusters of length P with each cluster denoted as $\mathbf{s}_m^{(p)}(n)$. Collecting all the pilot clusters in the vector $\mathbf{s}^{(p)}(n) := [\mathbf{s}_0^{(p)T}(n), \dots, \mathbf{s}_{M-1}^{(p)T}(n)]^T$, we differentiate them from the data symbols collected in the vector $\mathbf{s}^{(d)}(n)$.

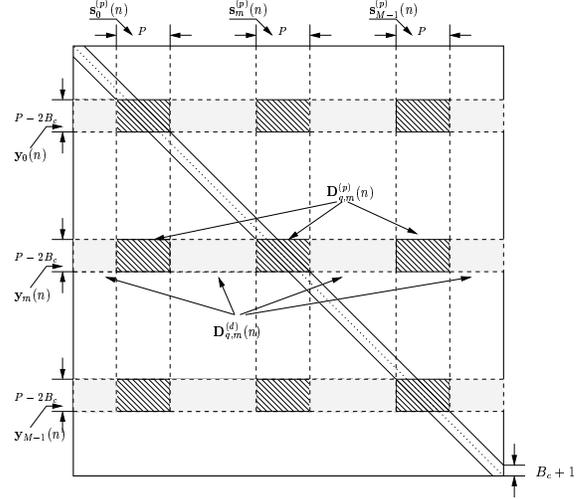


Fig. 2. Structure of $\mathbf{D}_q(n)$.

For such clustered pilots, it is up to the receiver to decide which of the received samples must be used for channel estimation. To clarify the notations that will come forth, we plot the structure of $\mathbf{D}_q(n)$ in Fig. 2. Obviously, the columns of $\mathbf{D}_q(n)$ are related to the positions of the pilots and information symbols, which operate on $\mathbf{D}_q(n)$ through the diagonal matrix Δ_q . The row positions are related to the observation samples. For the m th pilot cluster $\mathbf{s}_m^{(p)}(n) = [[\mathbf{s}(n)]_{P_m}, \dots, [\mathbf{s}(n)]_{P_m+P-1}]^T$, where P_m stands for its begin position, let us consider the following observation samples:

$$\mathbf{y}_m(n) := [[\mathbf{y}(n)]_{P_m+B_c}, \dots, [\mathbf{y}(n)]_{P_m+P-B_c-1}]^T. \quad (6)$$

Clearly, if $\mathbf{D}_q(n)$ were 'strictly' banded with $2B_c+1$ non-zero diagonals, $\mathbf{y}_m(n)$ would be the vector of maximal length that exclusively depends on the pilot cluster $\mathbf{s}_m^{(p)}(n)$. In this sense, B_c can be interpreted as the assumed bandwidth of $\mathbf{D}_q(n)$ as suggested in Fig. 2. However, we must be cautious with this interpretation, because $\mathbf{D}_q(n)$ is not strictly banded in general. Later on, it will become more clear that B_c actually provides a handle on the amount

of out-of-band interference that we want to take into account. Note that B_c can even be negative, in which case the bandwidth physical interpretation cannot be directly accounted for.

To formulate the above discussion in mathematical expressions with notations indicated in Fig. 2, we obtain

$$\mathbf{y}_m(n) = \sum_{q=0}^Q \mathbf{D}_{q,m}^{(p)}(n) \mathbf{\Delta}_q^{(p)} \mathbf{s}^{(p)}(n) + \underbrace{\sum_{q=0}^Q \mathbf{D}_{q,m}^{(d)}(n) \mathbf{\Delta}_q^{(d)} \mathbf{s}^{(d)}(n)}_{\mathbf{d}_m(n)} + \mathbf{n}_m(n),$$

where $\mathbf{D}_{q,m}^{(p)}(n)$ is an $(P - 2B_c) \times MP$ matrix, representing the hatched parts of $\mathbf{D}_q(n)$ in Fig. 2; $\mathbf{\Delta}_q^{(p)}$ is an $MP \times MP$ diagonal matrix, which is carved out of $\mathbf{\Delta}_q$ corresponding to the pilot-carrying subcarriers; $\mathbf{D}_{q,m}^{(d)}(n)$ is an $(P - 2B_c) \times (N - MP)$ matrix, representing the shaded parts of $\mathbf{D}_q(n)$ in Fig. 2; $\mathbf{\Delta}_q^{(d)}$ is an $(N - MP) \times (N - MP)$ diagonal matrix, which is carved out of $\mathbf{\Delta}_q$ corresponding to the information-carrying subcarriers; finally, $\mathbf{n}_m(n)$ stands for the noise related to $\mathbf{y}_m(n)$. In the above equation, we have thus uncoupled the effect of the information symbols from the pilots, and put it in a separate term $\mathbf{d}_m(n)$. Collecting all the observation samples in one vector $\mathbf{y}^{(p)}(n) := [\mathbf{y}_0^T(n), \dots, \mathbf{y}_{M-1}^T(n)]^T$, we can easily derive that

$$\mathbf{y}^{(p)}(n) = \mathcal{P}(n)\mathbf{h} + \mathbf{d}(n) + \mathbf{n}^{(p)}(n), \quad (7)$$

with

$$\mathcal{P}(n) := \begin{bmatrix} \mathbf{D}_{0,0}^{(p)}(n) & \cdots & \mathbf{D}_{Q,0}^{(p)}(n) \\ \vdots & \ddots & \vdots \\ \mathbf{D}_{0,M-1}^{(p)}(n) & \cdots & \mathbf{D}_{Q,M-1}^{(p)}(n) \end{bmatrix} (\mathbf{I}_{Q+1} \otimes \text{diag}\{\mathbf{s}^{(p)}(n)\}) \mathbf{F}_L^{(p)}.$$

Here, $\mathbf{F}_L^{(p)}$ standing for the rows of \mathbf{F}_L corresponding to the positions of the pilots, and $\mathbf{d}(n) := [\mathbf{d}_0^T(n), \dots, \mathbf{d}_{M-1}^T(n)]^T$ and $\mathbf{n}^{(p)}(n) := [\mathbf{n}_0^T(n), \dots, \mathbf{n}_{M-1}^T(n)]^T$. From (7), we observe that the observation samples are not only contaminated by the additive noise $\mathbf{n}^{(p)}(n)$, but also the out-of-band interference term $\mathbf{d}(n)$. The latter contains also the information \mathbf{h} as can be seen

$$\begin{aligned} \mathbf{d}(n) &= \mathbf{D}^{(d)}(n) \mathcal{S}^{(d)}(n) \mathbf{h}, \\ \mathbf{D}^{(d)}(n) &:= \begin{bmatrix} \mathbf{D}_{0,0}^{(d)}(n) & \cdots & \mathbf{D}_{Q,0}^{(d)}(n) \\ \vdots & \ddots & \vdots \\ \mathbf{D}_{0,M-1}^{(d)}(n) & \cdots & \mathbf{D}_{Q,M-1}^{(d)}(n) \end{bmatrix}, \\ \mathcal{S}^{(d)}(n) &:= \mathbf{I}_{Q+1} \otimes (\text{diag}\{\mathbf{s}^{(d)}(n)\}) \mathbf{F}_L^{(d)}. \end{aligned} \quad (8)$$

Here, $\mathbf{F}_L^{(d)}$ stands for the rows of \mathbf{F}_L corresponding to the positions of the information symbols. From its definition, we understand that $\mathbf{d}(n)$ can be made smaller by increasing the value of B_c , which corresponds to a more accurate band assumption of the channel matrix. However, this leads at the same time to less observation samples that can be fed to the channel estimator, thus a ‘fatter’ $\mathcal{P}(n)$ in (7). Optimizing B_c will be discussed in the next section.

Finally, since the BEM holds for a total of $2J + 1$ OFDM symbols we have

$$\mathbf{y}^{(p)} = \mathcal{P}\mathbf{h} + \mathbf{d} + \mathbf{n}^{(p)}, \quad (9)$$

with $\mathbf{y}^{(p)} := [\mathbf{y}^{(p)T}(-J), \dots, \mathbf{y}^{(p)T}(J)]^T$, $\mathcal{P} := [\mathcal{P}^T(-J), \dots, \mathcal{P}^T(J)]^T$, and \mathbf{d} and $\mathbf{n}^{(p)}$ similarly defined as $\mathbf{y}^{(p)}$.

4. CHANNEL ESTIMATION AND B_C OPTIMIZATION

Below we propose a Least Squares (LS) and a Linear Minimum Mean Square Error (LMMSE) estimator. For both of them, we assume the channel, data and noise are uncorrelated with each other. Based on the resulting estimation error, we put forward criteria to find the optimal B_c .

4.1. The LS Estimator

The LS estimator treats \mathbf{h} as a deterministic variable and assumes further no statistical knowledge about the data or noise. It can be obtained by seeking

$$\hat{\mathbf{h}}_{\text{LS}} = \arg \min_{\{\mathbf{h}\}} \|\mathbf{y}^{(p)} - \mathcal{P}\mathbf{h}\|^2, \quad (10)$$

which leads to the LS estimator $\mathbf{W}_{\text{LS}} := \mathcal{P}^\dagger$ such that

$$\hat{\mathbf{h}}_{\text{LS}} = \mathbf{W}_{\text{LS}} \mathbf{y}^{(p)} = \mathbf{h} + \mathcal{P}^\dagger (\mathbf{d} + \mathbf{n}^{(p)}).$$

The LS estimator is robust in cases where knowledge about the channel and noise statistics is unreliable or even absent. However, it suffers from an inferior performance especially when the interference is prominent as evident from the resulting MSE:

$$\begin{aligned} \text{MSE}_{\text{LS}} &:= \mathcal{E}_{\mathbf{h}, \mathbf{s}^{(d)}, \mathbf{n}^{(p)}} \{ \text{trace}\{(\hat{\mathbf{h}}_{\text{LS}} - \mathbf{h})(\hat{\mathbf{h}}_{\text{LS}} - \mathbf{h})^H\} \\ &= \mathcal{E}_{\mathbf{h}, \mathbf{s}^{(d)}, \mathbf{n}^{(p)}} \{ \text{trace}\{\mathcal{P}^\dagger (\mathbf{d} + \mathbf{n}^{(p)}) (\mathbf{d} + \mathbf{n}^{(p)})^H \mathcal{P}^{\dagger H}\} \} \\ &= \text{trace}\{\mathcal{P}^\dagger (\mathbf{R}_d + \mathbf{R}_n^{(p)}) \mathcal{P}^{\dagger H}\}, \end{aligned} \quad (11)$$

where $\mathbf{R}_d := \mathcal{E}_{\mathbf{h}, \mathbf{s}^{(d)}} \{\mathbf{d}\mathbf{d}^H\}$ and $\mathbf{R}_n^{(p)} := \mathcal{E}_{\mathbf{n}} \{\mathbf{n}^{(p)} \mathbf{n}^{(p)H}\}$, whose computation can be similarly found in [9]. It is not difficult to understand that MSE_{LS} relies heavily on the condition number of \mathcal{P} , which is in its turn at the choice of B_c .

4.2. The LMMSE Estimator

The LMMSE estimator treats the channel as stochastic. Besides, we assume that the data symbols are zero-mean white with unit variance. The LMMSE channel estimates $\hat{\mathbf{h}} = \mathbf{W}\mathbf{y}^{(p)}$ are found by minimizing the following cost function

$$\begin{aligned} \text{MSE}_{\text{LMMSE}} &:= \text{trace}\{\mathcal{E}_{\mathbf{h}, \mathbf{s}^{(d)}, \mathbf{n}} \{(\mathbf{W}\mathbf{y}^{(p)} - \mathbf{h})(\mathbf{W}\mathbf{y}^{(p)} - \mathbf{h})^H\}\} \\ &= \mathbf{W} (\mathcal{P}\mathbf{R}_h\mathcal{P}^H + \mathbf{R}_d + \mathbf{R}_n^{(p)} + 2\Re\{\mathcal{E}_{\mathbf{h}, \mathbf{s}^{(d)}} \{\mathbf{d}\mathbf{h}^H\} \mathcal{P}^H\}) \mathbf{W}^H \\ &\quad - 2\Re\{\mathbf{R}_h\mathcal{P}^H\mathbf{W}^H + \mathcal{E}_{\mathbf{h}, \mathbf{s}^{(d)}} \{\mathbf{h}\mathbf{d}^H\} \mathbf{W}^H\} + \mathbf{R}_h, \end{aligned} \quad (12)$$

with $\mathbf{R}_h := \mathcal{E}_{\mathbf{h}} \{\mathbf{h}\mathbf{h}^H\}$ whose computation is given in [9]. Due to the assumption that the data symbols are zero-mean and uncorrelated with the channel, $\mathcal{E}_{\mathbf{h}, \mathbf{s}^{(d)}} \{\mathbf{d}\mathbf{h}^H\} \mathcal{P}^H = \mathbf{0}$ and $\mathcal{E}_{\mathbf{h}, \mathbf{s}^{(d)}} \{\mathbf{h}\mathbf{d}^H\} = \mathbf{0}$. Hence, (12) becomes

$$\text{MSE}_{\text{LMMSE}} = \mathbf{W} (\mathcal{P}\mathbf{R}_h\mathcal{P}^H + \mathbf{R}_d + \mathbf{R}_n^{(p)}) \mathbf{W}^H - 2\Re\{\mathbf{R}_h\mathcal{P}^H\mathbf{W}^H\} + \mathbf{R}_h,$$

which, by minimizing with respect to \mathbf{W} , leads to

$$\mathbf{W}_{\text{LMMSE}} = \mathbf{R}_h\mathcal{P}^H (\mathcal{P}\mathbf{R}_h\mathcal{P}^H + \mathbf{R}_d + \mathbf{R}_n^{(p)})^{-1}, \quad (13)$$

From the above steps, we remark that although in our data model the out-of-band interference contains the information \mathbf{h} itself, its cross-correlation with other terms can be averaged out and thus the resulting channel estimator still bears the same form as the classical Wiener filter [14]. By substituting (13) back to (12), we can obtain the corresponding MSE as

$$\text{MSE}_{\text{LMMSE}} = \text{trace}\{(\mathcal{P}^H (\mathbf{R}_d + \mathbf{R}_n^{(p)})^{-1} \mathcal{P} + \mathbf{R}_h^{-1})^{-1}\}. \quad (14)$$

which is again a function of B_c .

4.3. Optimization of B_c

First of all, we frame the possible values of B_c to a certain range by claiming the following lemma (a proof can be found as in [9]):

Lemma 1 *Practical values of B_c must satisfy:*

$$\frac{P}{2} - \frac{N}{2M} \leq B_c \leq \frac{P}{2} - \frac{(L+1)(Q+1)}{2M(2J+1)}, \quad (15)$$

as we recall that M is the number of pilot clusters and P is the size of each pilot cluster.

We optimize B_c within the above range in terms of the MSE expressions given in (11) and (14), which are defined for the LS and LMMSE estimators, respectively. However, a closed-form solution is difficult to find. Fortunately, Lemma 1 implies only a limited exhaustive computer search, and the resulting MSE-versus- B_c curves by simulation exhibit a monotonous trail (see e.g., [9]), although in an opposite direction for the LS and LMMSE estimator. In particular for the LS estimator, the maximum $B_c = \frac{P}{2} - \frac{(L+1)(Q+1)}{2M(2J+1)}$ is optimal. Due to its lack of statistical knowledge, the LS estimator requires the out-of-band interference to be minimized, or in other words, the channel must be viewed as banded as possible. On the other hand, the minimum $B_c = \frac{P}{2} - \frac{N}{2M}$ is optimal for the LMMSE estimator. For practical setups (e.g., in Sec.5), this often leads to a negative B_c , which implies that in some of the observation samples, the power of the out-of-band interference can even overrule the power of the pilots. This forms, however, no serious problem to the LMMSE estimator, which is weaponed with the statistical knowledge to suppress the interference efficiently.

5. NUMERICAL RESULTS

In this section, we test the proposed algorithms for true Jakes' channels, using the simulator that is given in [15]. We assume the channel is an FIR channel with $L+1=4$ channel taps, each being an independent Gaussian random variable with an exponential power decay $\sigma_l^2 = e^{-l/P_l}$ for $l \in \{0, \dots, 3\}$ and $P_l = 1$.

We consider an OFDM system with $N = 128$ subcarriers, where the pilot subcarriers are grouped in M equidistant clusters, each containing P pilot tones. Inside each cluster, we adopt the scheme referred to as "Frequency-Domain Kronecker Delta" (FDKD) in [7], where a non-zero pilot is located in the middle of the cluster with zero guard bands on both sides.

We characterize the TV channel by the Doppler frequency normalized to the subcarrier spacing $f_D := \frac{vf_c}{c} T_c N = 0.14$, where v denotes the mobile velocity, f_c the carrier frequency, $T_c N$ the OFDM symbol duration, and c the speed of light. We differentiate between an unwindowed channel and a windowed channel: for the latter, we adopt the MBAE-SOE window proposed in [12], which will be a sum of five exponentials. The performance will be compared in terms of the following criterion

$$\text{MSE-CH} := \mathcal{E}\{\|(\mathbf{P}_0 \otimes \mathbf{I}_{L+1})(\mathbf{h}^{(t)} - (\mathbf{B} \otimes \mathbf{I}_{L+1})\hat{\mathbf{h}})\|^2\}. \quad (16)$$

Note that since we compare to a realistic TV channel, not only the channel estimator's MSE but also the BEM's modeling error are taken into account in the above criterion.

Throughout the test, we will adopt the GCE-BEM (other BEMs are also applicable, but will not be examined here due to space restrictions), for which the BEM matrix $\hat{\mathbf{B}}$ in (4) has entries $[\hat{\mathbf{B}}]_{p,q} = \exp\{j \frac{2\pi}{(2J+1)NK} p(q - \frac{Q}{2})\}$. Here $(2J+1)NK$ determines the period of the BEM. First of all, by extending the BEM window to multiple OFDM symbols $J > 0$ the BEM has a larger period, thereby

avoiding possible edge effects [2]. In addition, by increasing J more pilots can be taken into account. Increasing the oversampling factor K has a similar effect as increasing J , but has no influence on the amount of pilot symbols. Further, by enlarging the BEM period (increasing J or K), the BEM probably needs a larger Q to maintain the same modeling capability. On the other hand, increasing Q will increase the number of unknowns, which could have a detrimental effect on pilot-assisted channel estimation. For the LMMSE estimator, we see that the performance saturates at a certain Q , whereas for the LS estimator, the performance first improves and then deteriorates, leading to an optimal Q . We will not study these effects in depth here, but always take a close to optimal combination of $[K, Q]$ for every J in the simulations.

Test case 1. The LS estimator. In Fig. 3 and Fig. 4, we plot the performance of the LS estimators for the unwindowed and windowed channel, respectively. We observe that (i) the estimators based on three consecutive OFDM symbols ($J = 1$) have a superior performance to the estimators based on a single OFDM symbol ($J = 0$); (ii) despite the roughly same amount of overhead, the pilot structure (M and P) plays a distinctive role in channel estimation; (iii) the LS could suffer from an error floor at high SNR due to the out-of-band interference, which can be better combatted by basing the LS estimator on multiple OFDM symbols and/or impose the MBAE-SOE window.

Test case 2. The LMMSE estimator. In Fig. 5 and Fig. 6, we plot the performance of the LMMSE estimators for the unwindowed and windowed channel, respectively. Similar observations as regards with the LS estimator can be made here as well. Besides, it is obvious to see that the LMMSE estimator has a better performance than the LS estimator due to the exploration of the statistical knowledge.

Test case 3 The effect of mismatched statistical knowledge. The statistical knowledge necessary for the LMMSE estimator can be inexact in practice (see e.g., [16] though in a different context). In Fig. 7, we construct an LMMSE estimator based on a fixed assumption $f_D = 0.14$, and compare its performance for a range of true Doppler frequencies. It loses its precision, especially when the assumed f_D deviates too much from the true value. In Fig. 8, we construct LMMSE estimators that are based on multipath power profiles assuming $P_l = 1$ and $P_l = \infty$, and compare their performance for a range of multipath power profiles. It can be observed that the LMMSE estimator has a robust performance by assuming a uniform power profile, i.e. let $P_l = \infty$.

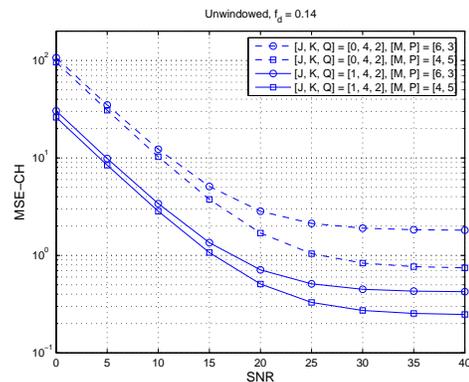


Fig. 3. The LS estimator for unwindowed TV channels.

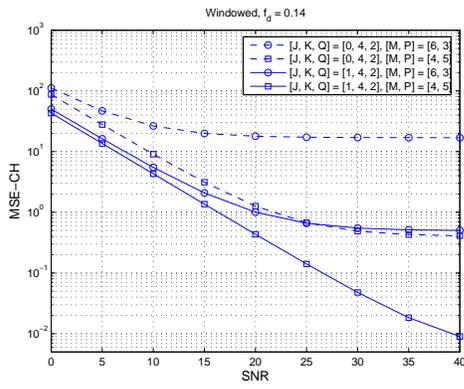


Fig. 4. The LS estimator for windowed TV channels.

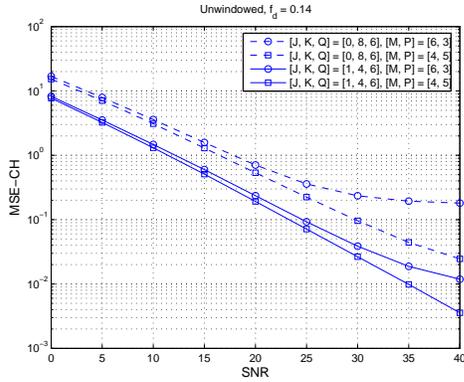


Fig. 5. The LMMSE estimator for unwinded TV channels.

6. REFERENCES

- [1] M. K. Tsatsanis and G. B. Giannakis, "Modeling and equalization of rapidly fading channels," *International Journal of Adaptive Control and Signal Processing*, vol. 10, no. 2/3, pp. 159–176, Mar. 1996.
- [2] G. Leus, "On the estimation of rapidly time-varying channels," *European Signal Processing Conference, EUSIPCO*, Sept. 2004.
- [3] T. Zemen and C. F. Mecklenbräuker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Transactions on Signal Processing*, vol. 53, no. 9, pp. 3597–3607, Sept. 2005.
- [4] D. K. Borah and B. D. Hart, "Frequency-selective fading channel estimation with a polynomial time-varying channel model," *IEEE Transactions on Communications*, vol. 47, no. 6, pp. 862–873, June 1999.

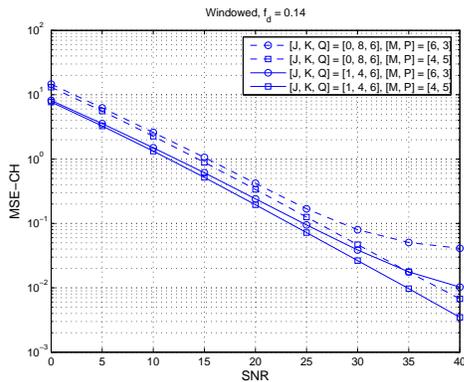


Fig. 6. The LMMSE estimator for windowed TV channels.

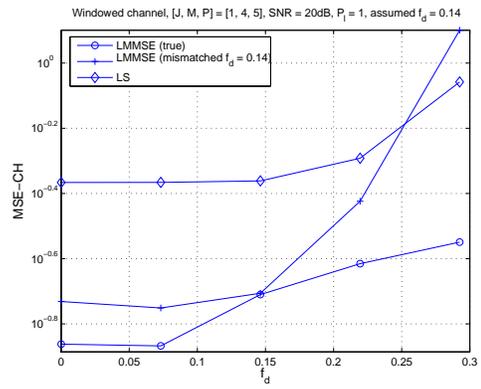


Fig. 7. The impact of incorrect Doppler frequency knowledge.

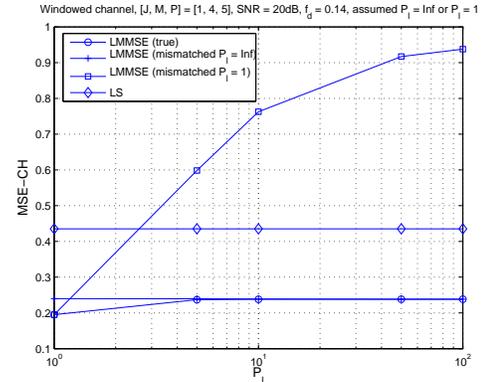


Fig. 8. The impact of incorrect multipath power profile.

- [5] K. D. Teo and S. Ohno, "Optimal MMSE finite parameter model for doubly-selective channels," *IEEE Global Telecommunications Conference, GLOBECOM*, 2005.
- [6] X. Ma, G. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly-selective fading channels," *IEEE Transactions on Signal Processing*, vol. 51, no. 5, pp. 1351–1366, May 2003.
- [7] A. P. Kannu and P. Schniter, "MSE-optimal training for linear time-varying channels," *International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, Mar. 2005.
- [8] A. Stamoulis, S. N. Diggavi, and N. Al-Dhahir, "Intercarrier interference in MIMO OFDM," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2451–2464, Oct. 2002.
- [9] Z. Tang, R. C. Cannizzaro, G. Leus, and P. Banelli, "Pilot-assisted time-varying channel estimation for OFDM systems," *submitted to IEEE Transactions on Signal Processing*.
- [10] W. C. Jakes, *Microwave Mobile Channels*, New York: Wiley, 1974.
- [11] Y.-S. Choi, P. J. Voltz, and F. A. 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] L. Rugini, P. Banelli, and G. Leus, "Block DFE and windowing for Doppler-affected OFDM systems," *IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications, SPAWC*, pp. 470–474, June 2005.
- [13] P. Schniter, "Low-complexity equalization of OFDM in doubly-selective channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 4, pp. 1002–1011, Apr 2004.
- [14] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, New Jersey, USA, 1993.
- [15] Y. R. Zheng and C. Xiao, "Simulation models with correct statistical properties for Rayleigh fading channels," *IEEE Transactions on Communications*, vol. 51, no. 6, pp. 920–928, June 2003.
- [16] P. Schniter, "On doubly dispersive channel estimation for pilot-aided pulse-shaped multi-carrier modulation," *submitted to Conference on Information Sciences and Systems, CISS*, 2006.