

PILOT-ASSISTED TIME-VARYING OFDM CHANNEL ESTIMATION

Zijian Tang, Geert Leus

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

Rocco Claudio Cannizzaro, 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 novel channel estimators to combat both the noise and the out-of-band interference. Our claims are supported by simulation results, which are obtained based on realistic TV channels with a fairly high Doppler spread.

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

1. INTRODUCTION

In mobile communications, a high vehicle speed causes the carrier frequency to spread out. This so-called Doppler spread yields a Time-Varying (TV) channel, whose channel taps vary with time. Basis Expansion Models (BEMs) 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, 2] amongst others, the Generalized CE-BEM (GCE-BEM) in [3], the Discrete Prolate Spheroidal BEM (DPS-BEM) in [4], and the Polynomial BEM (P-BEM) in [5]. Other modeling approaches are also reported, e.g., a first-order Gauss-Markov process is utilized in [6]. Such a model is interesting for sequential time-domain processing. When we deal with block transmission/precoding schemes, such as an Orthogonal Frequency-Division Multiplexing (OFDM) system, it is often more convenient to use a block-based channel model like a BEM.

Focusing on the estimation of channels that are modeled by a BEM, we basically only need to estimate the BEM coefficients. [2, 7] propose pilot-assisted estimators based on a CE-BEM assumption. Usually, pilots can be clustered in the time-domain [7] to combat the Inter-Symbol Interference (ISI). Likewise, for OFDM systems, it is also useful to cluster the pilots in the frequency-domain [2] to combat the Inter-Carrier Interference (ICI) induced by the Doppler spread. 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.

In this paper, we will adopt a similar pilot structure as in [2], but consider channel estimation for a general BEM assumption. However, as opposed to the CE-BEM channel that is strictly banded in the frequency-domain [2], the ICI of other BEM channels de-

grades only gradually, just like a realistic TV channel with a finite number of subcarriers. In that case, if we artificially select a clear-cut bandwidth for the channel matrix, the out-of-band entries will give rise to interference. We propose three channel estimators to combat this interference as well as the additive noise. We will show that for each channel estimator the influence of the out-of-band interference is distinctive, and as a result, the bandwidth selection must be optimized for each channel estimator, individually.

Notation: We use upper (lower) bold face letters to denote matrices (column vectors). $(\cdot)^H$ represents complex conjugate transpose (Hermitian). $\mathcal{E}_x\{\cdot\}$ 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. 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 MODEL

Let us consider an OFDM system with N subcarriers. The OFDM symbol \mathbf{s} is first modulated to N subcarriers as $\mathbf{s}^{(t)} = \mathbf{F}^H \mathbf{s}$, 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, $\mathbf{s}^{(t)}$ is next concatenated by a cyclic prefix (CP), sent over the channel, stripped from the CP, and finally demodulated. The resulting data stream can be expressed as

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

where $\mathbf{H}^{(t)}$ and $\mathbf{H} := \mathbf{F} \mathbf{H}^{(t)} \mathbf{F}^H$ denote the channel matrix in the time-domain and frequency-domain, respectively, and \mathbf{n} represents the noise in the frequency-domain. Defining $h_{n,l}^{(t)}$ as the l th time-domain channel tap at the n th sample, we assume the channel is Finite Impulse Response (FIR) with order L , i.e., $h_{n,l}^{(t)} = 0$ for $l < 0$ or $l > L$. If the CP length L_{cp} then satisfies $L_{cp} \geq L$, $\mathbf{H}^{(t)}$ is 'pseudo-circulant' and free from inter-block interference with $[\mathbf{H}^{(t)}]_{p,q} = h_{p+L_{cp}, \text{mod}(p-q, N)}^{(t)}$.

The channel taps are approximated by a BEM. Collecting the time-variation of the l th channel tap within the considered OFDM symbol in an $N \times 1$ vector $\mathbf{h}_l^{(t)} := [h_{L_{cp}, l}^{(t)}, \dots, h_{N+L_{cp}-1, l}^{(t)}]^T$, we can express $\mathbf{h}_l^{(t)}$ as $\mathbf{h}_l^{(t)} \approx \mathbf{B} \mathbf{h}_l$, where $\mathbf{B} := [\mathbf{b}_0, \dots, \mathbf{b}_Q]$ is an $N \times (Q+1)$ matrix that collects $Q+1$ orthonormal basis functions \mathbf{b}_q as columns, and $\mathbf{h}_l := [h_{0,l}, \dots, h_{Q,l}]^T$ with $h_{q,l}$ representing the q th BEM coefficient for the l th channel tap, which is obtained in a Mean Square Error (MSE) sense and shall remain invariant within the OFDM symbol. Adopting the same BEM approximation for all the channel taps $l = 0, \dots, L$, we can

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

rewrite (1) after some algebra as¹

$$\mathbf{y} = \sum_{q=0}^Q \mathbf{D}_q \mathbf{\Delta}_q \mathbf{s} + \mathbf{n}, \quad (2)$$

where $\mathbf{D}_q := \mathbf{F} \text{diag}(\mathbf{b}_q) \mathbf{F}^H$ and $\mathbf{\Delta}_q$ is a diagonal matrix $\mathbf{\Delta}_q := \text{diag}\{\mathbf{F}_L[h_{q,0}, \dots, h_{q,L}]\}^T$ with \mathbf{F}_L standing 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 possible receiver window. In that case, the time-domain channel matrix $\mathbf{H}^{(t)}$ can be written as $\mathbf{H}^{(t)} = \text{diag}\{\mathbf{w}\} \tilde{\mathbf{H}}^{(t)}$, where \mathbf{w} is the adopted window and $\tilde{\mathbf{H}}^{(t)}$ represents the unwinded channel. Such a receiver window has been recently reported in [9, 10] to improve the performance of low-complexity equalizers that exploit the banded approximation of the frequency-domain channel matrix \mathbf{H} . To approximate such a windowed channel, we differentiate two options in the BEM design. First, if a CE-BEM is to be used, we can just stick to the original design for the unwinded case as in [1]. For other BEMs, it turns out to be beneficial to adapt the BEM to the window, i.e., we design \mathbf{B} as

$$\mathbf{B} := \text{diag}\{\mathbf{w}\} \tilde{\mathbf{B}} \mathbf{Q} \quad (3)$$

where $\tilde{\mathbf{B}}$ yields one of the traditional BEM designs presented in [3–5] and \mathbf{Q} is a square matrix to ensure that the columns of \mathbf{B} are orthonormal. For more details on the BEM design see [11].

3. DATA MODEL FOR CHANNEL ESTIMATION

Instead of estimating the true bulky channel matrix \mathbf{H} , we will estimate the in total $(L+1)(Q+1)$ BEM coefficients collected in $\mathbf{h} := [h_{0,0}, \dots, h_{0,L}, \dots, h_{Q,0}, \dots, h_{Q,L}]^T$. We assume there are M pilot clusters of length L_p denoted as $\mathbf{s}_m^{(p)}$, $m = 0, 1, \dots, M-1$. They are collected in the vector $\mathbf{s}^{(p)} := [\mathbf{s}_0^{(p)T}, \dots, \mathbf{s}_{M-1}^{(p)T}]^T$, and interleaved with information symbols collected in the information symbol vector $\mathbf{s}^{(d)}$.

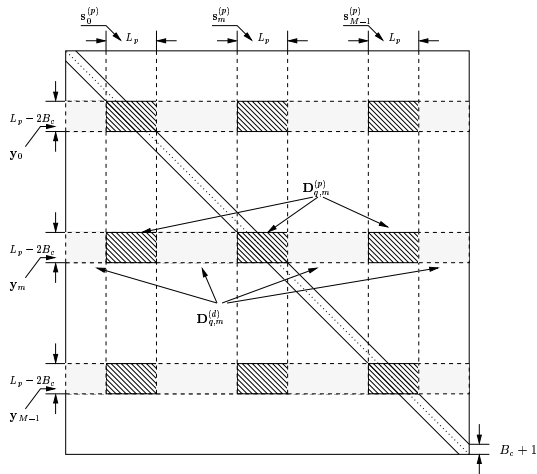


Fig. 1. The structure of \mathbf{D}_q in relation with the pilots and the received samples.

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

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 in Fig. 1. Obviously, the columns of \mathbf{D}_q are related to the positions of the pilots and information symbols, which operate on \mathbf{D}_q through the diagonal matrix $\mathbf{\Delta}_q$. The row positions are related to the observation samples. For the m th pilot cluster $\mathbf{s}_m^{(p)} = [[\mathbf{s}]_{P_m}, \dots, [\mathbf{s}]_{P_m+L_p-1}]^T$, where P_m stands for its begin position, let us consider the following observation samples:

$$\mathbf{y}_m := [[\mathbf{y}]_{P_m+B_c}, \dots, [\mathbf{y}]_{P_m+L_p-B_c-1}]^T. \quad (4)$$

It is not hard to imagine that if \mathbf{D}_q were ‘strictly’ banded with $2B_c+1$ non-zero diagonals (this occurs if we assume a CE-BEM), then \mathbf{y}_m would be the vector of maximal length that exclusively depends on the pilot cluster $\mathbf{s}_m^{(p)}$. In this sense, B_c can be interpreted as the assumed bandwidth of \mathbf{D}_q as suggested in Fig. 1. However, we must be cautious with this interpretation, because \mathbf{D}_q is not strictly banded for most BEMs. 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. 1, we obtain

$$\mathbf{y}_m = \sum_{q=0}^Q \mathbf{D}_{q,m}^{(p)} \mathbf{\Delta}_q^{(p)} \mathbf{s}_m^{(p)} + \underbrace{\sum_{q=0}^Q \mathbf{D}_{q,m}^{(d)} \mathbf{\Delta}_q^{(d)} \mathbf{s}^{(d)}}_{\mathbf{d}_m} + \mathbf{n}_m, \quad (5)$$

where $\mathbf{D}_{q,m}^{(p)}$ is an $(L_p - 2B_c) \times ML_p$ matrix, representing the hatched parts of \mathbf{D}_q in Fig. 1; $\mathbf{\Delta}_q^{(p)}$ is an $ML_p \times ML_p$ diagonal matrix, which is carved out of $\mathbf{\Delta}_q$ corresponding to the pilot-carrying subcarriers; $\mathbf{D}_{q,m}^{(d)}$ is an $(L_p - 2B_c) \times (N - ML_p)$ matrix, representing the shaded parts of \mathbf{D}_q in Fig. 1; $\mathbf{\Delta}_q^{(d)}$ is an $(N - ML_p) \times (N - ML_p)$ diagonal matrix, which is carved out of $\mathbf{\Delta}_q$ corresponding to the information-carrying subcarriers; finally, \mathbf{n}_m stands for the noise related to \mathbf{y}_m . 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 . Collecting all the observation samples in one vector $\mathbf{y}^{(p)} := [\mathbf{y}_0^T, \dots, \mathbf{y}_{M-1}^T]^T$ and denoting

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

with $\mathbf{F}_L^{(p)}$ standing for the rows of \mathbf{F}_L corresponding to the positions of the pilots, we can easily derive

$$\mathbf{y}^{(p)} = \mathcal{P} \mathbf{h} + \mathbf{d} + \mathbf{n}^{(p)}, \quad (6)$$

where $\mathbf{d} := [\mathbf{d}_0^T, \dots, \mathbf{d}_{M-1}^T]^T$ and $\mathbf{n}^{(p)} := [\mathbf{n}_0^T, \dots, \mathbf{n}_{M-1}^T]^T$. From the last equality, we observe that the observation samples are not only contaminated by the additive noise $\mathbf{n}^{(p)}$, but also by the interference term \mathbf{d} . The latter contains also the information \mathbf{h} as $\mathbf{d} = \mathcal{G} \mathbf{h}$, where

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

with $\mathbf{F}_L^{(d)}$ standing for the rows of \mathbf{F}_L corresponding to the positions of the information symbols.

Clearly, the pilot-related \mathcal{P} and the interference-related \mathcal{G} both have the dimension $M(L_p - 2B_c) \times (Q + 1)(L + 1)$. Intuitively, one would like to reduce the power of the interference \mathbf{d} by setting B_c as large as possible. The same idea is adopted in [8] though the authors address the problem from a different point of view. To explain this using the physical interpretation of B_c : a larger B_c corresponds to a more accurate band approximation, and thus to a smaller out-of-band interference. However, at the same time when B_c is enlarged, less observation samples are taken into account, leading to a ‘fatter’ \mathcal{P} . This is usually detrimental if we are to deploy a linear channel estimator.

4. CHANNEL ESTIMATION

In the following, we assume the information symbols $\mathbf{s}^{(d)}$ and the noise $\mathbf{n}^{(p)}$ are both zero-mean and uncorrelated with each other. The channel vector \mathbf{h} will be treated either as a random variable or a deterministic variable. This leads to three different channel estimators. The performance of each channel estimator is sensitive to the choice of B_c , which will be examined individually.

4.1. The Linear Minimum Mean Square Error (LMMSE) Estimator

Assuming \mathbf{h} to be stochastic and uncorrelated with $\mathbf{s}^{(d)}$ and $\mathbf{n}^{(p)}$, we can find a linear filter \mathbf{W} to minimize the MSE between the estimated and true BEM coefficients. [11] gives its expression as

$$\hat{\mathbf{h}}_{\text{LMMSE}} = \mathbf{R}_h \mathcal{P}^H (\mathcal{P} \mathbf{R}_h \mathcal{P}^H + \mathbf{R}_{\mathcal{I}})^{-1} \mathbf{y}^{(p)}, \quad (7)$$

where $\mathbf{R}_{\mathcal{I}} := \mathbf{R}_d + \mathbf{R}_n^{(p)}$ with $\mathbf{R}_d := \mathcal{E}_{\mathbf{h}, \mathbf{s}^{(d)}} \{\mathbf{d} \mathbf{d}^H\}$ and $\mathbf{R}_n^{(p)} := \mathcal{E}_{\mathbf{n}^{(p)}} \{\mathbf{n}^{(p)} \mathbf{n}^{(p)H}\}$ whose derivations are given in [11]. Note that although (7) is similar to the canonical LMMSE estimator [12], we underline the difference that in the considered case, the interference term contains the information \mathbf{h} itself and is mitigated by the LMMSE estimator resorting to the Second-Order Statistics (SOS) of the channel. From (7), we can show that the MSE of this LMMSE estimator is:

$$\text{MSE}_{\text{LMMSE}} = \text{trace}\{(\mathcal{P}^H \mathbf{R}_{\mathcal{I}}^{-1} \mathcal{P} + \mathbf{R}_h^{-1})^{-1}\}. \quad (8)$$

Since the choice of B_c determines the content of both \mathcal{P} and $\mathbf{R}_{\mathcal{I}}$, (8) implies that the LMMSE’s performance is subject to B_c .

4.2. The Least Squares Estimator

The Least Squares (LS) estimator treats \mathbf{h} as a deterministic variable and assumes further no knowledge about the channel and noise statistics:

$$\hat{\mathbf{h}}_{\text{LS}} = \mathcal{P}^\dagger \mathbf{y}^{(p)} = \mathbf{h} + \mathcal{P}^\dagger (\mathbf{d} + \mathbf{n}^{(p)}). \quad (9)$$

For the LS estimator, we can easily derive its MSE as

$$\text{MSE}_{\text{LS}} = \text{trace}\{\mathcal{P}^\dagger \mathbf{R}_{\mathcal{I}} \mathcal{P}^\dagger H\}, \quad (10)$$

Compared to (8), one can observe that the LS estimator’s performance is also influenced by B_c , but in a different manner.

4.3. An Iterative Best Linear Unbiased Estimator (BLUE)

From (6), a BLUE can be derived [12, Appendix 6B] by treating the interference \mathbf{d} and noise $\mathbf{n}^{(p)}$ as a whole such that

$$\hat{\mathbf{h}}_{\text{BLUE}} = (\mathcal{P}^H \tilde{\mathbf{R}}_{\mathcal{I}}^{-1}(\mathbf{h}) \mathcal{P})^{-1} \mathcal{P}^H \tilde{\mathbf{R}}_{\mathcal{I}}^{-1}(\mathbf{h}) \mathbf{y}^{(p)}, \quad (11)$$

with $\tilde{\mathbf{R}}_{\mathcal{I}}(\mathbf{h}) := \mathcal{E}_{\mathbf{s}^{(d)}, \mathbf{n}^{(p)}} \{(\mathbf{d} + \mathbf{n}^{(p)})(\mathbf{d} + \mathbf{n}^{(p)})^H\} = \tilde{\mathbf{R}}_d(\mathbf{h}) + \mathbf{R}_n^{(p)}$. Here $\tilde{\mathbf{R}}_d(\mathbf{h})$ is the covariance matrix of \mathbf{d} by treating \mathbf{h} as deterministic $\tilde{\mathbf{R}}_d(\mathbf{h}) := \mathcal{E}_{\mathbf{s}^{(d)}} \{\mathbf{d} \mathbf{d}^H\}$. For more details, see [11].

However, (11) is not implementable since its computation entails the information of \mathbf{h} itself. A recursive approach is therefore proposed: suppose at the k th iteration, an estimate for \mathbf{h} is available and denoted as $\hat{\mathbf{h}}_{\text{BLUE}}^{(k)}$. Next, we use this estimate to update the covariance matrix $\tilde{\mathbf{R}}_{\mathcal{I}}(\mathbf{h})$, which is in turn used to produce the BLUE for the next iteration and so on:

$$\hat{\mathbf{h}}_{\text{BLUE}}^{(k+1)} = (\mathcal{P}^H \tilde{\mathbf{R}}_{\mathcal{I}}^{-1}(\hat{\mathbf{h}}_{\text{BLUE}}^{(k)}) \mathcal{P})^{-1} \mathcal{P}^H \tilde{\mathbf{R}}_{\mathcal{I}}^{-1}(\hat{\mathbf{h}}_{\text{BLUE}}^{(k)}) \mathbf{y}^{(p)}, \quad (12)$$

The convergence can be easily ensured by initializing the iteration with $\hat{\mathbf{h}}_{\text{BLUE}}^{(0)} = \mathbf{0}$. In that case, we already obtain a weighted LS estimate $\hat{\mathbf{h}}_{\text{BLUE}}^{(1)} = (\mathcal{P}^H (\mathbf{R}_n^{(p)})^{-1} \mathcal{P})^{-1} \mathcal{P}^H (\mathbf{R}_n^{(p)})^{-1} \mathbf{y}^{(p)}$, which could be close to the global minimum if the interference term is not too eminent. Note that a similar idea is also echoed in [13] amongst others, though in the context of superimposed training for time-invariant channel estimation.

Assuming that $\hat{\mathbf{h}}_{\text{BLUE}}^{(k)} \rightarrow \hat{\mathbf{h}}_{\text{BLUE}}$, we use (11) to find a lower bound on the estimator’s MSE

$$\text{MSE}_{\text{BLUE}} = \mathcal{E}_{\mathbf{h}} \{\text{trace}((\mathcal{P}^H \tilde{\mathbf{R}}_{\mathcal{I}}^{-1}(\mathbf{h}) \mathcal{P})^{-1})\}. \quad (13)$$

Because the expression above is difficult to evaluate in closed form, we need to resort to the Monte Carlo method. As will be evident later on, the MSE of the BLUE also depends on the choice of B_c .

4.4. Optimization of B_c

First of all, we restrict the possible values of B_c to a certain range using the following lemma (for a proof see [11]):

Lemma 1 *Practical values of B_c must satisfy:*

$$\frac{L_p}{2} - \frac{N}{2M} \leq B_c \leq \frac{L_p}{2} - \frac{(L+1)(Q+1)}{2M}, \quad (14)$$

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

We optimize B_c within the above range in terms of the MSE expressions given in (8), (10) and (13), which are defined for the LMMSE, LS, and the BLUE, respectively. In other words, we solve $B_c = \arg \min_{\{B_c\}} \text{MSE}$. It is difficult to find a closed-form solution for this problem. However, [11] shows by simulation that the MSE versus B_c curve for the LS estimator is monotonous descending, while the curves for the LMMSE estimator and the BLUE are both monotonous ascending, which implies that we can simply choose the largest possible B_c for the LS estimator and the smallest possible B_c for the LMMSE estimator and the BLUE. To explain this, we recall that a larger B_c is equivalent to a smaller out-of-band interference power. This is in favor of the LS estimator, which is not good at suppressing the interference due to the lack of statistical knowledge. For the LMMSE estimator and BLUE on the other hand, the out-of-band interference is not a big concern,

and we could therefore minimize B_c in order to process more observation samples. However, since the computational complexity of the BLUE increases significantly for a decreasing B_c , we generally take the optimal B_c for the BLUE somewhere in the middle of the range in (14). More details about this issue can be found in [11].

5. NUMERICAL RESULTS

We test the proposed algorithms for realistic Jakes' channels. We assume an FIR channel with $L + 1 = 6$ taps. The variation of the channel taps is determined by the Doppler frequency, which is normalized to the subcarrier spacing $f_D := \frac{v}{c} T_s N$, where v denotes the vehicle's velocity, f_c the carrier frequency, $T_s N$ the OFDM symbol duration, and c the speed of light. Further, we window the received signal by the MBAE-SOE window proposed in [9], which is a sum of three exponentials. To approximate such a windowed channel with a GCE-BEM, we set $Q = 4$ and design the GCE-BEM following (3).

We assume the OFDM system has $N = 256$ subcarriers, where roughly 80% of the subcarriers are used for transmitting information symbols that are QPSK modulated. The remaining subcarriers are reserved for pilots, which are grouped in $M = 6$ equidistant clusters, each containing $L_p = 9$ pilot tones. Inside each cluster, we adopt the pilot scheme used in [2], where a non-zero pilot is located in the center of the cluster with zero guard bands on both sides.

Test Case 1. The estimator performance. In line with the analysis in Sect. 4.4, we select $B_c = -16$ for the LMMSE estimator, $B_c = 2$ for the LS estimator, and $B_c = -3$ for the BLUE, and inspect their channel estimation performance for two types of TV channels (i) $f_D = 1$, and (ii) $f_D = 0.2$. The proposed estimators are compared in Fig. 2 with the channel estimator presented in [2], which resembles the LMMSE estimator but is based on a CE-BEM assumption. Despite its MSE optimality proven in [2], the CE-BEM based estimator is inferior to the proposed LMMSE estimator and BLUE due to a larger BEM modeling error with respect to the considered GCE-BEM. The performance is also compared with the CRB (see [11] for a derivation). The CRB is obtained based on many Monte Carlo runs, thereby exploring the channel statistics. It is clear that the performance of the BLUE is very close to the CRB.

Test Case 2. Equalization performance based on the estimated channel. The influence of the channel estimation error on the BER performance of the considered OFDM system is examined next. To this end, we apply the banded-MMSE (B-MMSE) equalizer proposed in [14], which is characterized by a good compromise between performance and complexity. For the sake of comparison, we also list the equalization performances, which are based on the estimated CE-BEM channel and the perfect CSI. From Fig. 3, we can observe that the equalization performance to a great extent follows the channel estimation performance depicted in Fig. 2.

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, pp. 159–176, Mar. 1996.
- [2] 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.

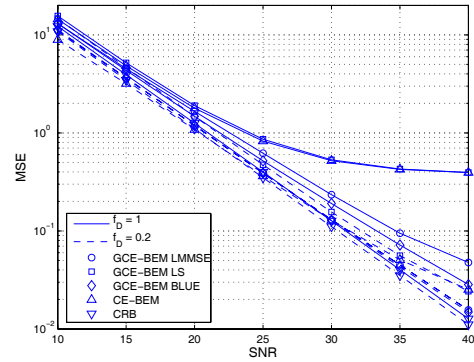


Fig. 2. Channel estimator performance.

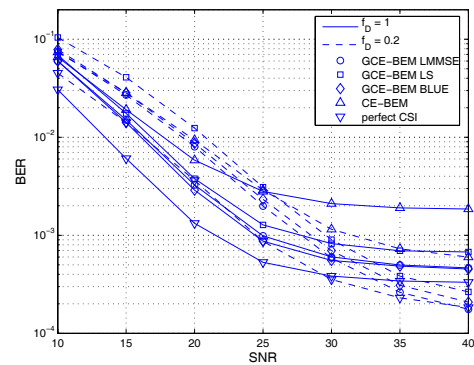


Fig. 3. Equalization performance.

- [3] G. Leus, "On the estimation of rapidly time-varying channels," *European Signal Processing Conference, EUSIPCO*, Sept. 2004.
- [4] T. Zemen and C. F. Mecklenbräuker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Transactions on Signal Processing*, vol. 53, pp. 3597–3607, Sept. 2005.
- [5] 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, pp. 862–873, June 1999.
- [6] M. Dong, L. Tong, and B. Sadler, "Optimal insertion of pilot symbols for transmissions over time-varying flat fading channels," *IEEE Transactions on Signal Processing*, vol. 52, pp. 1403–1418, May 2004.
- [7] X. Ma, G. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly-selective fading channels," *IEEE Transactions on Signal Processing*, vol. 51, pp. 1351–1366, May 2003.
- [8] A. Stamoulis, S. N. Diggavi, and N. Al-Dahir, "Inter-carrier interference in MIMO OFDM," *IEEE Transactions on Signal Processing*, vol. 50, pp. 2451–2464, Oct. 2002.
- [9] 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.
- [10] P. Schniter, "Low-complexity equalization of OFDM in doubly-selective channels," *IEEE Transactions on Signal Processing*, vol. 52, pp. 1002–1011, Apr 2004.
- [11] 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*.
- [12] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. New Jersey, USA, 1993.
- [13] M. Ghogho and A. Swami, "Improved channel estimation using superimposed training," *IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications, SPAWC*, pp. 110–114, July 2004.
- [14] L. Rugini, P. Banelli, and G. Leus, "Simple equalization of time-varying channels for OFDM," *IEEE Communications Letters*, vol. 9, pp. 619–621, July 2005.