A Novel Receiver Architecture for Single-Carrier Transmission over Time-Varying Channels

Zijian Tang and Geert Leus

***Abstract*—In this paper, we present a single-carrier transceiver for rapidly time-varying channels, where the equalization step is implemented in the frequency domain. When the channel abides with both fast fading and severe inter-block interference, our equalizer relies on a band approximation of the frequency-domain channel matrix to maintain low complexity. We will show that the band approximation error can be associated in the time domain to a critically-sampled complex exponential basis expansion modeling error. Based on this property, we propose a novel receiver architecture that extends the original data model by inserting zeros at the receiver. The resulting effective channel can be characterized by an oversampled complex exponential basis expansion model, which has a considerably reduced modeling error compared to the critically-sampled one. In other words, the band assumption that is essential to the equalizer will be made more accurate and thus the equalization performance can be improved.**

***Index Terms*—single-carrier, basis expansion model, time-varying channels, inter-block interference**

I. Introduction

In a single-carrier transmission system over a lengthy channel, it is more efficient to equalize the channel in the frequency domain utilizing a simple one-tap equalizer [1]. The underlying consideration is that the frequency-domain (FD) channel is a diagonal matrix, i.e., the subcarriers remain orthogonal to each other. However, this is only true if there is no inter-block interference (IBI) present and the channel stays invariant during at least one block. In the presence of IBI, a sufficient number of redundant symbols (guard interval) needs to be inserted between the blocks, which might not be affordable in practice due to a stringent bandwidth constraint. In a scenario of high-mobility, on the other hand, the channel time variation within a block cannot be neglected, and induces Doppler spread in the frequency domain. In these cases, the orthogonality among the subcarriers is corrupted and the FD channel becomes actually a full matrix. A reliable FD equalizer for such a channel will be much more expensive, which is the key issue considered in this paper.

To restore the orthogonality among the subcarriers, pre-processing at the receiver is indispensable. For instance, a channel shortening technique, in the form of a finite impulse response (FIR) filter, is proposed in [2] for time-invariant channels, with the aim of shortening the effective channel within the given guard interval. In [3], an FIR filter is adopted to "flatten" the channel's fluctuation, which can be considered as the dual of channel shortening. In [4], both schemes are combined. Such an approach works well for channels that are moderately spread in delay and Doppler dimensions.

Often, a perfectly diagonal FD channel matrix is too difficult to achieve. In a realistic transmission system, the Doppler-induced channel has most of its power concentrated in the vicinity of the diagonal in a circular sense, with those entries that are far away from the diagonal decreasing fast [5], [6]. This implies that it is more practical to assume a banded FD channel matrix[1]. Many equalizers exploit this banded (rather than diagonal) structure to lower the complexity, e.g., the block linear zero-forcing (ZF) equalizer in [7], the block linear minimum mean square error (MMSE) equalizer in [8], [9], the iterative serial MMSE equalizer in [10]–[12], the maximum likelihood (ML) equalizer in [13]–[15], etc. It can be imagined that to enhance the equalization performance, especially at a moderate to high signal-to-noise ratio (SNR), the band approximation error must be reduced as much as possible. One solution can be the FIR filter of [4], but it generally requires a multiple antenna assumption and can still be too complicated. Since we need not to enforce a diagonal FD channel matrix but a banded one, a reduced-order FIR filter with just a single tap could be adequate. Such a filter is referred to as a receiver window in [9]–[12], [15].

In this paper, we will present two receiver architectures in combination with windowing to counteract the channel time variation as well as the IBI for a single-carrier transmission system. The first receiver will be based on the original data model (ODM), which describes the actual channel input/output (I/O) relationship. The second receiver will be based on the so-called extended data model (EDM), which extends the ODM by inserting zeros at the receiver. The advantage of the EDM is that by inserting zeros at the receiver, we are endowed with some extra design freedom to shape the Doppler effect better. More specifically, we can choose freely the part of the channel that corresponds to the inserted zeros. In both data models, the full FD channel matrix will be approximated by a banded matrix for the sake of complexity. Note that unlike [9]–[12], [15] where the banded matrix is obtained by extracting the significant diagonals from the original FD channel matrix, the band approximation in this paper is achieved in a different manner: we will device a banded matrix that is close to the full FD channel matrix only in terms of the Frobenius norm. For the ODM, where the considered full FD channel matrix is the same as the original one, these two approaches are the same. However, for the EDM, where part of the considered full FD channel matrix

Manuscript received March 31, 2007; revised September 10, 2007. This research was supported in part by NWO-STW under the VICI program (DTC.5893) and the VIDI program (DTC.6577).

The authors are with Delft University of Technology - Fac. EEMCS, Mekelweg 4, 2628 CD Delft, The Netherlands (e-mail: {z.tang, g.leus}@tudelft.nl).

Digital Object Identifier 10.1109/JSAC.2008.080213.

[1]Strictly speaking, we actually mean "circularly-banded" here. However, we will use the term "banded" in the sequel for the sake of brevity.

can be chosen freely and thus not completely corresponds to the original one, the proposed approach has an advantage. A more profound rationale is that we have translated the band approximation error in the frequency domain into a basis expansion modeling (BEM) error in the time domain. Recall that the idea of the BEM is initially documented in [16] to reduce the number of parameters of a time-varying channel at the cost of a small modeling error. In this paper, we can show that the band approximation error in the ODM corresponds to a modeling error resulting from the critically-sampled complex exponential BEM ((C)CE-BEM) [17], [18][2], while the band approximation error in the EDM corresponds to a modeling error resulting from the oversampled complex exponential BEM ((O)CE-BEM) [19]–[21]. This idea will be reflected in our window design. Since the (O)CE-BEM in general can yield a much tighter fit to a realistic time-varying channel than the (C)CE-BEM, as reported in [22], it is not hard to understand that the equalizer for the EDM will be subject to a much smaller band approximation error than for the ODM, and could thus be able to render a better performance.

We assume in this paper that the channel state information (CSI) is known. Channel estimation for time-varying channels has for instance been discussed in [18]. In the simulations section, we will include some results using this channel estimator.

Notation: We use upper (lower) bold face letters to denote matrices (column vectors). $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ represent conjugate, transpose and complex conjugate transpose (Hermitian), respectively. $\mathcal{E}_x\{\cdot\}$ stands for the expectation with respect to x. $\odot$ represents the Schur-Hadamard (element-wise) product. $\mathrm{mod}(a,b)$ gives the remainder of a divided by b. $\mathrm{tr}(\mathbf{X})$ and $\|\mathbf{X}\|$ denote the trace and Frobenius norm of $\mathbf{X}$, respectively. $\mathbf{X}^\dagger$ denotes the pseudo inverse of $\mathbf{X}$. $\boldsymbol{\mathcal{D}}\{\mathbf{x}\}$ stands for a diagonal matrix with $\mathbf{x}$ as the diagonal. We use $[\mathbf{x}]_p$ to indicate the $(p+1)$st element of $\mathbf{x}$, and $[\mathbf{X}]_{p,q}$ to indicate the $(p+1,q+1)$st entry of $\mathbf{X}$. Further, we let $\mathbf{I}_N$ denote an $N\times N$ identity matrix, $\mathbf{0}_{M\times N}$ an $M\times N$ all-zero matrix, and $\mathbf{1}_{M\times N}$ an $M\times N$ all-one matrix. $\mathbf{e}_k$ stands for a unit vector with a one at the $(k+1)$st position. $\mathbf{F}_N$ denotes the unitary N-point DFT matrix with $[\mathbf{F}_N]_{p,q}=\frac{1}{\sqrt{N}}e^{-j\frac{2\pi}{N}pq}$.

II. System Model

Let us consider the discrete-time baseband model of a communication system, where the channel is assumed to be an FIR filter with order L, i.e., if we use $h_{p,l}$ to denote the lth channel tap at the pth time index then $h_{p,l}=0$ if $l<0$ or $l>L$. Conform the FIR assumption, we can express the I/O relationship as

$$y_p = w_p\sum_{l=0}^{L} h_{p,l}s_{p-l} + v_p, \tag{1}$$

where w_p stands for the pth element of the window that is deployed at the receiver; y_p and v_p denote the (windowed) observation sample and noise at the pth time index, respectively; and s_p denotes the pth data symbol.

[2]As a matter of fact, such a link also underlies the equalizer design in [10], [12], but is not straightforward to observe.

For this data model, we adopt the following assumptions.

Assumption 1: We deal in this paper with time-varying channels, which implies that $h_{p,l}\neq h_{q,l}$ if $p\neq q$. We assume that the channel can be statistically characterized by a wide-sense stationary uncorrelated scattering (WSSUS) model. To be specific, we assume that

$$\mathcal{E}_h\{h_{p,l}h_{p-m,l-n}\} = \sigma_l^2\gamma_m\delta_n, \tag{2}$$

where δ_n denotes the Kronecker delta, σ_l^2 the variance of the lth channel tap, and γ_m the normalized time correlation, i.e., $\gamma_0=1$.

Assumption 2: We assume that the data symbols are zero-mean white with unit variance, i.e., $\mathcal{E}_s\{s_p s_{p-m}^*\}=\delta_m$, and the noise prior to windowing is zero-mean white with variance σ^2. With the window taken into account, this means that $\mathcal{E}_v\{v_p v_{p-m}^*\}=\sigma^2\delta_m w_p w_{p-m}^*$.

III. FD Equalization Based on the ODM

A. Equalization Scheme

Suppose that the received samples are parsed into (possibly overlapping) blocks of size N. For instance, let us define $\mathbf{y}_{t,N}$ as a vector collecting the observation samples from time index 0 to $N-1$, $\mathbf{y}_{t,N}:=[y_0,\cdots,y_{N-1}]^T$. Conform (1), the I/O relationship for $\mathbf{y}_{t,N}$ can be expressed in matrix/vector form as

$$\mathbf{y}_{t,N} = \boldsymbol{\mathcal{D}}\{\mathbf{w}\}\mathbf{H}[\mathbf{s}_{\text{pre}}^T,\mathbf{s}_{N-L}^T,\mathbf{s}_{\text{post}}^T]^T + \mathbf{v}_{t,N}, \tag{3}$$

where $\mathbf{v}_{t,N}$ is similarly defined as $\mathbf{y}_{t,N}$; $\mathbf{w}$ is the $N\times 1$ window vector, $\mathbf{w}:=[w_0,\cdots,w_{N-1}]^T$; $\mathbf{s}_{N-L}$ is an $(N-L)\times 1$ vector collecting data symbols, $\mathbf{s}_{N-L}:=[s_0,\cdots,s_{N-L-1}]^T$; and $\mathbf{s}_{\text{pre}}$ and $\mathbf{s}_{\text{post}}$ represent the L data symbols that are contiguous to $\mathbf{s}_{N-L}$, $\mathbf{s}_{\text{pre}}:=[s_{-L},\cdots,s_{-1}]^T$ and $\mathbf{s}_{\text{post}}:=[s_{N-L},\cdots,s_{N-1}]^T$. The $N\times(N+L)$ matrix $\mathbf{H}$ stands for the convolutive channel matrix with entries $[\mathbf{H}]_{p,n}:=h_{p,p-n+L}$.

In Fig. 1, it is shown how the relationship (3) can be reformulated as a quasi-circulant relationship[3]. In mathematics, this leads to

$$\mathbf{y}_{t,N} = \boldsymbol{\mathcal{D}}\{\mathbf{w}\}\mathbf{H}_{t,N}\mathbf{s}_N + \boldsymbol{\epsilon}_{t,N} + \mathbf{v}_{t,N}, \tag{4}$$

where

$$\mathbf{s}_N := [\mathbf{s}_{N-L}^T,\mathbf{s}_{\text{post}}^T]^T, \tag{5}$$

and the $N\times N$ matrix $\mathbf{H}_{t,N}$ has entries $[\mathbf{H}_{t,N}]_{p,n}:=h_{p,\mathrm{mod}(p-n,N)}$. The term $\boldsymbol{\epsilon}_{t,N}$ represents the IBI, which can be expressed as

$$\boldsymbol{\epsilon}_{t,N} := \boldsymbol{\mathcal{D}}\{\mathbf{w}\}\mathbf{H}_{i,N}(\mathbf{s}_{\text{pre}}-\mathbf{s}_{\text{post}}), \tag{6}$$

where $\mathbf{H}_{i,N}$ is an $N\times L$ matrix with entries $[\mathbf{H}_{i,N}]_{p,n}:=h_{p,p-n+L}$. Transformed into the frequency domain, (4) becomes

$$\mathbf{y}_{f,N} := \mathbf{F}_N\mathbf{y}_{t,N} = \mathbf{H}_{f,N}\mathbf{F}_N\mathbf{s}_N + \boldsymbol{\epsilon}_{f,N} + \mathbf{v}_{f,N}, \tag{7}$$

where $\boldsymbol{\epsilon}_{f,N}$ and $\mathbf{v}_{f,N}$ are similarly defined as $\mathbf{y}_{f,N}$, and $\mathbf{H}_{f,N}:=\mathbf{F}_N\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\mathbf{H}_{t,N}\mathbf{F}_N^H$ stands for the FD channel matrix. Since $\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\mathbf{H}_{t,N}$ is quasi-circulant, and not circulant due to the channel time variation, $\mathbf{H}_{f,N}$ is not diagonal.

[3]We use the term "quasi-circulant" when dealing with a time-varying circular convolution.

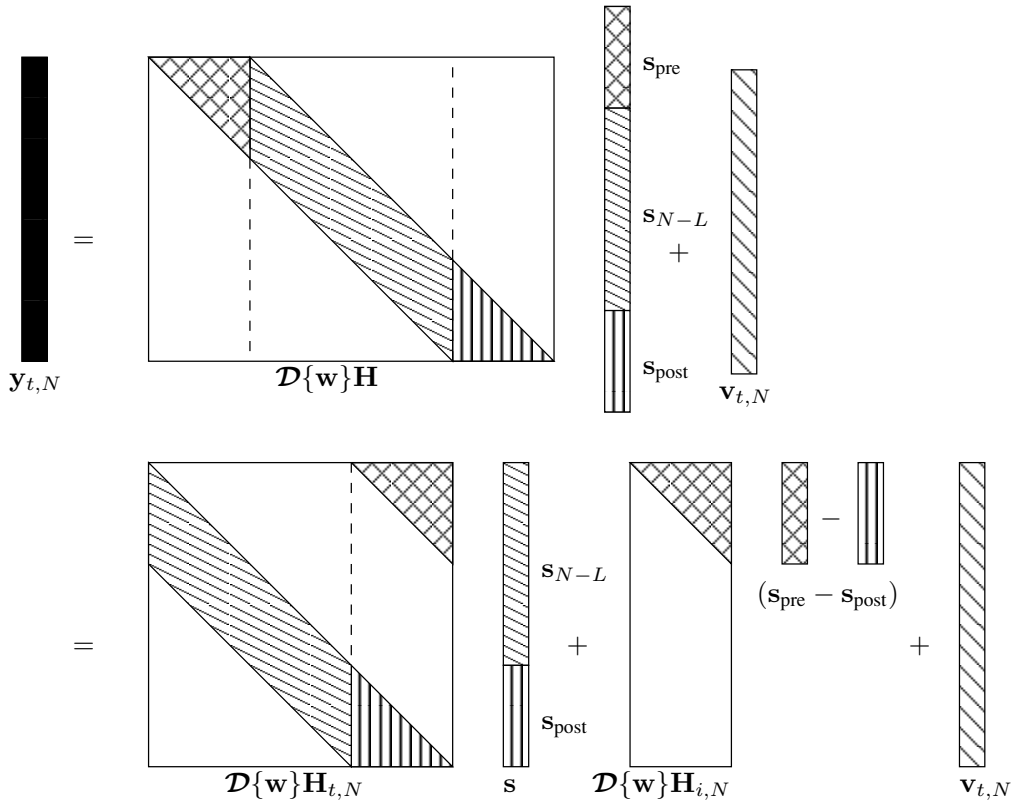

Fig. 1. The original data model.

In (7), except for the windowing and the DFT, we did not apply any other processing on the received samples, and the data model is identical to what has actually happened in reality. We therefore call this data model the original data model (ODM). It is typical to the ODM that the DFT size equals the number of observation samples. This will be in contrast with the EDM scheme discussed in the next section, where the DFT size is larger than the number of observation samples.

The IBI in the frequency domain $\boldsymbol{\epsilon}_{f,N}$ can be mitigated by the utility of a guard interval of length L_z, e.g., a cyclic-prefix (CP), a zero-postfix (ZP) [23] or a non-zero postfix (NZP) [24]. In the CP case, we let $[s_{-L_z}, \cdots, s_{-1}] = [s_{N-L}, \cdots, s_{N-L+L_z-1}]$, while in the ZP and NZP case, we let $[s_{-L_z}, \cdots, s_{-1}]^T = [s_{N-L}, \cdots, s_{N-L+L_z-1}]^T = \mathbf{p}$, with $\mathbf{p}$ being a zero or non-zero pilot vector, respectively. When $L_z \geq L$, IBI is completely removed. When there is no guard, i.e., $L_z = 0$, IBI is present and has to be dealt with. However, it is in that case still possible to reduce the amount of IBI. We can for instance apply a sliding window approach, which only estimates N_s data symbols out of $\mathbf{s}_{N-L}$ at the time. After that, the equalizer shifts N_s observation samples forward to estimate the next N_s data symbols. This implies that the consecutive received sample blocks $\mathbf{y}_{t,N}$ will overlap with each other over $N - N_s$ samples. This is beneficial because the IBI usually undermines the reliability of the data symbols at the edges of the block. Note that a similar approach has been proposed in [12].

The non-zero off-diagonal elements of $\mathbf{H}_{f,N}$ prevent the viability of a simple one-tap equalizer. To facilitate a low-complexity equalizer, we will approximate $\mathbf{H}_{f,N}$ with a banded matrix $\hat{\mathbf{H}}_{f,N}$, which has only non-zero entries on the main diagonal, the first $Q/2$ upper and the first $Q/2$ lower diagonals in a circular sense. Here, Q is a design parameter that can be chosen to make a trade-off between complexity and performance. The smaller Q, the cheaper the equalizer, but the worse the performance. A good value for Q is related to the maximum Doppler spread [11]. It is noteworthy that the banded matrix $\hat{\mathbf{H}}_{f,N}$ is not simply obtained by taking the $Q+1$ most significant diagonals of $\mathbf{H}_{f,N}$ in a circular sense. We desire that $\hat{\mathbf{H}}_{f,N}$ should be close to $\mathbf{H}_{f,N}$ in terms of the Frobenius norm. Therefore, $\hat{\mathbf{H}}_{f,N}$ is designed as

$$\min_{\hat{\mathbf{H}}_{f,N}} \|\mathbf{H}_{f,N} - \hat{\mathbf{H}}_{f,N}\|^2, \quad \text{s.t. } \hat{\mathbf{H}}_{f,N} = \hat{\mathbf{H}}_{f,N} \odot \mathbf{T}_Q. \tag{8}$$

where $\mathbf{T}_Q$ is a matrix of proper dimensions, which has ones on the main diagonal, the first $Q/2$ upper and the first $Q/2$ lower diagonals in a circular sense, and zeros on the remaining entries.

Let us now make things more concrete by applying the above ideas to a block linear MMSE equalizer. This is just one possible option, and similar studies can for instance be made for a serial linear MMSE equalizer or for the iterative versions of these two MMSE equalizers. Note that a similar study for the iterative serial MMSE equalizer has been presented in [10], [12]. Suppose we want to estimate the middle N_s data symbols of $\mathbf{s}_{N-L}$ by applying a block linear MMSE equalizer on (7), where we neglect the IBI and replace $\mathbf{H}_{f,N}$ by $\hat{\mathbf{H}}_{f,N}$:

$$\hat{\mathbf{s}}_{N_s} = \boldsymbol{\Omega}_N \mathbf{F}_N^H \hat{\mathbf{H}}_{f,N}^H (\hat{\mathbf{H}}_{f,N} \hat{\mathbf{H}}_{f,N}^H + \mathbf{R}_{v,N})^{-1} \mathbf{y}_{f,N}, \tag{9}$$

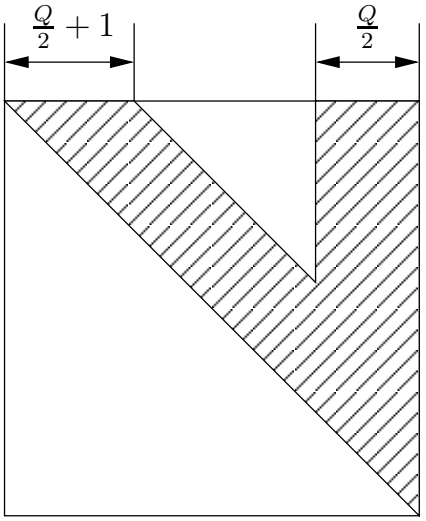

Fig. 2. The V-shape of $\mathbf{G}_N$.

with

$$\mathbf{\Omega}_N := \left[\mathbf{0}_{N_s\times\frac{N-L-N_s}{2}}\mathbf{I}_{N_s}\mathbf{0}_{N_s\times\frac{N+L-N_s}{2}}\right], \quad (10)$$

$$\mathbf{R}_{v,N} := \mathcal{E}_v\{\mathbf{v}_{f,N}\mathbf{v}_{f,N}^H\} = \sigma^2\mathbf{F}_N\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\boldsymbol{\mathcal{D}}\{\mathbf{w}^*\}\mathbf{F}_N^H. \quad (11)$$

Since $\hat{\mathbf{H}}_{f,N}\hat{\mathbf{H}}_{f,N}^H$ is banded with bandwidth $2Q+1$ and assuming that $\mathbf{R}_{v,N}$ is also banded with bandwidth $2Q+1$ (we come back to this issue later on), we can apply a Cholesky factorization [25] on the covariance matrix in (9) such that

$$\hat{\mathbf{H}}_{f,N}\hat{\mathbf{H}}_{f,N}^H + \mathbf{R}_{v,N} = \mathbf{G}_N\mathbf{G}_N^H, \quad (12)$$

where the upper-triangular matrix $\mathbf{G}_N$ will assume a sparse V-shape structure as illustrated in Fig. 2 (note that a similar structure was observed in [15]). Applying the inverse of $\hat{\mathbf{H}}_{f,N}\hat{\mathbf{H}}_{f,N}^H + \mathbf{R}_{v,N}$ can then be implemented by applying the inverses of $\mathbf{G}_N$ and $\mathbf{G}_N^H$ separately using, e.g., Gaussian elimination. It can be shown that this approach inflicts a complexity of $\mathcal{O}(NQ^2)$, i.e., the complexity is linear in N and square in Q.

It is worth mentioning that in a single-carrier system, the channel can also be equalized in the time domain. For instance, we can apply a block linear MMSE equalizer directly on (3), for which the complexity can be shown to be $\mathcal{O}(N_sL^2)$. However, since N/N_s is generally much smaller than L/Q, it is more appealing to equalize the channel in the frequency domain than in the time domain.

From the above derivations, it can be understood that to enhance the equalization performance of the ODM, we need to design the window $\mathbf{w}$ and the banded matrix $\hat{\mathbf{H}}_{f,N}$ such that the IBI $\|\boldsymbol{\epsilon}_{f,N}\|^2$ as well as the band approximation error $\|\mathbf{H}_{f,N}-\hat{\mathbf{H}}_{f,N}\|^2$ will be minimized in some average sense. In addition, the window should also be able to make the noise covariance matrix $\mathbf{R}_{v,N}$ banded. These issues will be discussed next.

B. Window Design for the ODM

We begin the window design with its noise shaping behavior. Considering Assumption 2, we adopt the approach of [9], which is summarized in the following proposition (see [9] for a proof).

Proposition 1: The noise covariance matrix $\mathbf{R}_{v,N}$ will be banded with bandwidth $2Q+1$ if we let the window $\mathbf{w}$ be a weighted sum of $Q+1$ complex exponentials:

$$\mathbf{w} = \mathbf{B}_N\mathbf{d}, \quad (13)$$

where $\mathbf{B}_N$ is comprised of the first $Q/2+1$ and the last $Q/2$ columns of $\mathbf{F}_N$; and $\mathbf{d}$ is a $(Q+1)\times 1$ vector containing all the weighting coefficients.

It is worth mentioning that the matrix $\mathbf{B}_N$ tallies with the classical definition of the (C)CE-BEM, whose period equals the BEM window size N [17], [18].

Next, to minimize the band approximation error $\|\mathbf{H}_{f,N}-\hat{\mathbf{H}}_{f,N}\|^2$, we need to design the window $\mathbf{w}$ and the banded matrix $\hat{\mathbf{H}}_{f,N}$ jointly. The following theorem proves to be important (see Appendix A for a proof).

Theorem 1: The minimization of the band approximation error $\|\mathbf{H}_{f,N}-\hat{\mathbf{H}}_{f,N}\|^2$ in the frequency domain can be transformed in the time domain as the minimization of the (C)CE-BEM modeling error. In mathematics, this can be expressed as

$$\min_{\hat{\mathbf{H}}_{f,N}}\|\mathbf{H}_{f,N}-\hat{\mathbf{H}}_{f,N}\|^2 = \min_{\boldsymbol{\mathcal{C}}}\|\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\boldsymbol{\mathcal{H}}-\mathbf{B}_N\boldsymbol{\mathcal{C}}\|^2. \quad (14)$$

In the above, $\boldsymbol{\mathcal{H}}$ stands for the $N\times(L+1)$ matrix collecting all the channel taps, $[\boldsymbol{\mathcal{H}}]_{n,l}=h_{n,l}$, and $\boldsymbol{\mathcal{C}}$ for the $(Q+1)\times(L+1)$ matrix collecting all the BEM coefficients, $[\boldsymbol{\mathcal{C}}]_{q,l}=c_{q,l}$.

Theorem 1 establishes the equivalence between the band approximation error in the ODM and the (C)CE-BEM modeling error.

Regarding the IBI, the following theorem is needed (see Appendix B for a proof).

Theorem 2: The average power of the IBI $\boldsymbol{\epsilon}_{f,N}$ in the ODM is a function of the window as

$$\mathcal{E}_{h,s}\{\|\boldsymbol{\epsilon}_{f,N}\|^2\} = 2\mathbf{w}^T\mathbf{R}_{\epsilon,N}\mathbf{w}^*, \quad (15)$$

where $\mathbf{R}_{\epsilon,N}$ denotes a diagonal matrix with diagonal entries given by

$$[\mathbf{R}_{\epsilon,N}]_{n,n} = \begin{cases}\sum_{l=n+L_z+1}^{L}\sigma_l^2 & \text{if } n\le L-L_z-1,\\ 0 & \text{otherwise.}\end{cases} \quad (16)$$

Using Proposition 1 as well as Theorems 1 and 2, the design problem of finding a fixed window that minimizes both the average minimal band approximation error and the average IBI can now be formulated as

$$\min_{\mathbf{w}}\ \mathcal{E}_h\{\min_{\boldsymbol{\mathcal{C}}}\|\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\boldsymbol{\mathcal{H}}-\mathbf{B}_N\boldsymbol{\mathcal{C}}\|^2\} + 2\mathbf{w}^T\mathbf{R}_{\epsilon,N}\mathbf{w}^*, \quad \text{s.t. } \mathbf{w}=\mathbf{B}_N\mathbf{d} \text{ and } \|\mathbf{w}\|^2=N. \quad (17)$$

Note that the constraint $\|\mathbf{w}\|^2=N$ is imposed to avoid the trivial all-zero window.

We first solve (17) for $\boldsymbol{\mathcal{C}}$ leading to $\boldsymbol{\mathcal{C}}=\mathbf{B}_N^{\dagger}\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\boldsymbol{\mathcal{H}}$. Plugging this result into the first term of (17), we obtain

$$\begin{aligned}&\mathcal{E}_h\{\min_{\boldsymbol{\mathcal{C}}}\|\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\boldsymbol{\mathcal{H}}-\mathbf{B}_N\boldsymbol{\mathcal{C}}\|^2\}\\ &= \mathrm{tr}\big(\boldsymbol{\mathcal{P}}_{\mathbf{B}_N}\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\mathbf{R}_{\boldsymbol{\mathcal{H}},N}\boldsymbol{\mathcal{D}}\{\mathbf{w}^H\}\boldsymbol{\mathcal{P}}_{\mathbf{B}_N}^H\big)\\ &= \mathbf{w}^T\Big(\sum_{n=0}^{N-1}\boldsymbol{\mathcal{D}}\{\boldsymbol{\mathcal{P}}_{\mathbf{B}_N}^T\mathbf{e}_n\}\mathbf{R}_{\boldsymbol{\mathcal{H}},N}\boldsymbol{\mathcal{D}}\{\boldsymbol{\mathcal{P}}_{\mathbf{B}_N}^H\mathbf{e}_n\}\Big)\mathbf{w}^*, \end{aligned} \quad (18)$$

where

$$\boldsymbol{\mathcal{P}}_{\mathbf{B}_N} := \mathbf{I}_N - \mathbf{B}_N \mathbf{B}_N^{\dagger}, \tag{19}$$

$$\mathbf{R}_{\boldsymbol{\mathcal{H}},N} := \mathcal{E}_h\{\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^H\}. \tag{20}$$

Note that using Assumption 1, the entries of $\mathbf{R}_{\boldsymbol{\mathcal{H}},N}$ can be expressed as $[\mathbf{R}_{\boldsymbol{\mathcal{H}},N}]_{m,n} = \sum_{l=0}^{L} \sigma_l^2 \gamma_{m-n}$.

Substituting (13) and (18) in (17) finally leads to

$$\min_{\mathbf{d}} \mathbf{d}^T \boldsymbol{\mathcal{X}}_N \mathbf{d}^*, \quad \text{s.t. } \|\mathbf{d}\|^2 = N, \tag{21}$$

with

$$\boldsymbol{\mathcal{X}}_N := \mathbf{B}_N^T \Big(\sum_{n=0}^{N-1} \boldsymbol{\mathcal{D}}\{\boldsymbol{\mathcal{P}}_{\mathbf{B}_N}^T \mathbf{e}_n\} \mathbf{R}_{\boldsymbol{\mathcal{H}},N} \times \boldsymbol{\mathcal{D}}\{\boldsymbol{\mathcal{P}}_{\mathbf{B}_N}^H \mathbf{e}_n\} + 2\mathbf{R}_{\epsilon,N} \Big) \mathbf{B}_N^*. \tag{22}$$

As a result, $\mathbf{d}$ can be computed as the least significant eigenvector of $\boldsymbol{\mathcal{X}}_N^*$.

Remark 1: We can show that the banded matrix $\hat{\mathbf{H}}_{f,N}$ that is obtained after minimizing the band approximation error actually corresponds to the $Q+1$ most significant diagonals of $\mathbf{H}_{f,N}$ in a circular sense, i.e., $\hat{\mathbf{H}}_{f,N} = \mathbf{H}_{f,N} \odot \mathbf{T}_Q$ (see [26] for a proof). It is noteworthy that although this result coincides with the canonical band approximation approach in [9]–[12], [15], the underlying consideration [c.f. (8)] is obviously different.

Remark 2: The windows designed in [10], [12] maximize the signal to interference (band approximation error and IBI) and noise ratio directly in the frequency domain. As a matter of fact, the band approximation error considered in [10], [12] can also be translated as the (C)CE-BEM modeling error just like in this paper. Indeed, if there is no noise and IBI, and the window length is restricted to be equal to the observation block length, we can show that the window of [10], [12] will admit the same expression as the window of the ODM (the proof can be found in [26])[4]. As will become evident from the simulations, the performance of the windowing strategies of [10], [12] is very close to that of the ODM. It is thus not difficult to understand that a possible drawback of the ODM window, and that of [10], [12] as well, is associated with a relatively large modeling error inherent to the (C)CE-BEM as reported in [22]. For the case no guard band is present, i.e., $L_z = 0$, this can partially be solved by taking N_s much smaller than $N - L$, so that the edge effects of the (C)CE-BEM are avoided. But this is more difficult to carry out when a guard band is present. A more general approach to avoid this problem consists of extending the data model, as will be discussed in the next section.

IV. FD Equalization Based on the EDM

In the previous section, we have shown that the band approximation error in the ODM can be translated into the (C)CE-BEM modeling error. While the (C)CE-BEM suffers from a relatively large modeling error, it is proposed in [19], [20] that a more generalized form, the (O)CE-BEM, can yield a much better modeling performance [22]. The (O)CE-BEM is achieved by simply enlarging the period of the (C)CE-BEM from N to K with $K > N$, maintaining the BEM window length at N. However, the (O)CE-BEM can not be straightforwardly applied to the ODM, because the (O)CE-BEM channel matrix, if transformed into the frequency domain by $\mathbf{F}_N$ to the left and $\mathbf{F}_N^H$ to the right, will not be banded. We can solve this problem by enlarging the data model from size N to size K, meanwhile keeping the observation block length equal to N. This is explained in more detail next.

A. Equalization Scheme

To derive a larger data model, let us first rewrite the ODM given in (4) as

$$\mathbf{y}_{t,N} = \boldsymbol{\mathcal{D}}\{\mathbf{w}\}\bar{\mathbf{H}}_N \mathbf{s}_{N-L} + \boldsymbol{\mathcal{D}}\{\mathbf{w}\}\bar{\mathbf{H}}_{i,N} \begin{bmatrix} \mathbf{s}_{\text{pre}} \\ \mathbf{s}_{\text{post}} \end{bmatrix} + \mathbf{v}_{t,N}, \tag{23}$$

where $\bar{\mathbf{H}}_N$ is an $N \times (N-L)$ matrix with entries $[\bar{\mathbf{H}}_N]_{p,n} := h_{p,p-n}$, and $\bar{\mathbf{H}}_{i,N}$ stands for an $N \times 2L$ matrix constructed as

$$\bar{\mathbf{H}}_{i,N} := \begin{bmatrix} \mathbf{A} & \mathbf{0}_{L\times L} \\ \mathbf{0}_{(N-2L)\times 2L} & \\ \mathbf{0}_{L\times L} & \mathbf{B} \end{bmatrix}, \tag{24}$$

where the $L \times L$ matrix $\mathbf{A}$ has entries $[\mathbf{A}]_{m,n} = h_{m,L-n+m}$, and the $L \times L$ matrix $\mathbf{B}$ has entries $[\mathbf{B}]_{m,n} = h_{N-L+m,m-n}$. The relationship (23) is illustrated in the upper part of Fig. 3, which should be compared with the upper part of Fig. 1.

We now want to extend the data model in (23) from size N to size K, with $K \geq N$. This can be done as shown in the bottom part of Fig. 3, where the $K \times (K-N+L)$ matrix $\mathbf{U}$, the $(K-N+L) \times 1$ vector $\mathbf{s}_{\text{vir}}$, and the $(K-N) \times 1$ vector $\mathbf{v}_{\text{vir}}$ are subject to design, and where we have further included the $K \times 1$ vector $\mathbf{y}_{\text{vir}}$, which equals

$$\mathbf{y}_{\text{vir}} = \mathbf{U}\mathbf{s}_{\text{vir}} + [\mathbf{0}_{1\times N}, \mathbf{v}_{\text{vir}}^T]^T. \tag{25}$$

In compact form, we get the following expression:

$$\underbrace{\begin{bmatrix} \mathbf{y}_{t,N} \\ \mathbf{0}_{(K-N)\times 1} \end{bmatrix} + \mathbf{y}_{\text{vir}}}_{\mathbf{y}_{t,K}} = \underbrace{\begin{bmatrix} \begin{pmatrix} \boldsymbol{\mathcal{D}}\{\mathbf{w}\}\bar{\mathbf{H}}_N \\ \mathbf{0}_{(K-N)\times(N-L)} \end{pmatrix} & \mathbf{U} \end{bmatrix}}_{\mathbf{H}_{t,K}} \underbrace{\begin{bmatrix} \mathbf{s}_{N-L} \\ \mathbf{s}_{\text{vir}} \end{bmatrix}}_{\mathbf{s}_K} + \underbrace{\begin{bmatrix} \boldsymbol{\mathcal{D}}\{\mathbf{w}\}\bar{\mathbf{H}}_{i,N} \\ \mathbf{0}_{(K-N)\times 2L} \end{bmatrix} \begin{bmatrix} \mathbf{s}_{\text{pre}} \\ \mathbf{s}_{\text{post}} \end{bmatrix}}_{\boldsymbol{\epsilon}_{t,K}} + \underbrace{\begin{bmatrix} \mathbf{v}_{t,N} \\ \mathbf{v}_{\text{vir}} \end{bmatrix}}_{\mathbf{v}_{t,K}}. \tag{26}$$

Note that in contrast to what we were saying before, we do not only add zeros at the receiver, but we possibly also include an additional known vector $\mathbf{y}_{\text{vir}}$. In any case, we have introduced some redundancy only at the receiver, which is completely transparent to the transmitter. Hence, the data rate is not compromised. A direct implication is that the ODM in (23) remains valid: it simply becomes a part of the larger data model in (26). For this reason, we will refer to the resulting relationship as the extended data model (EDM).

The second term on the RHS of (26), $\boldsymbol{\epsilon}_{t,K}$, is due to the IBI. Removing its last $K-N$ zero elements, we obtain

$$\bar{\boldsymbol{\epsilon}}_{t,N} = \boldsymbol{\mathcal{D}}\{\mathbf{w}\}\bar{\mathbf{H}}_{i,N} \begin{bmatrix} \mathbf{s}_{\text{pre}} \\ \mathbf{s}_{\text{post}} \end{bmatrix}. \tag{27}$$

[4]Note that [10] considers a CP with $L_z \geq L$ to remove the IBI and restricts the length of the window to the observation block length as in the ODM. On the other hand, [12] considers no guard, i.e., $L_z = 0$, but does not necessarily restrict the length of the window to the observation block length.

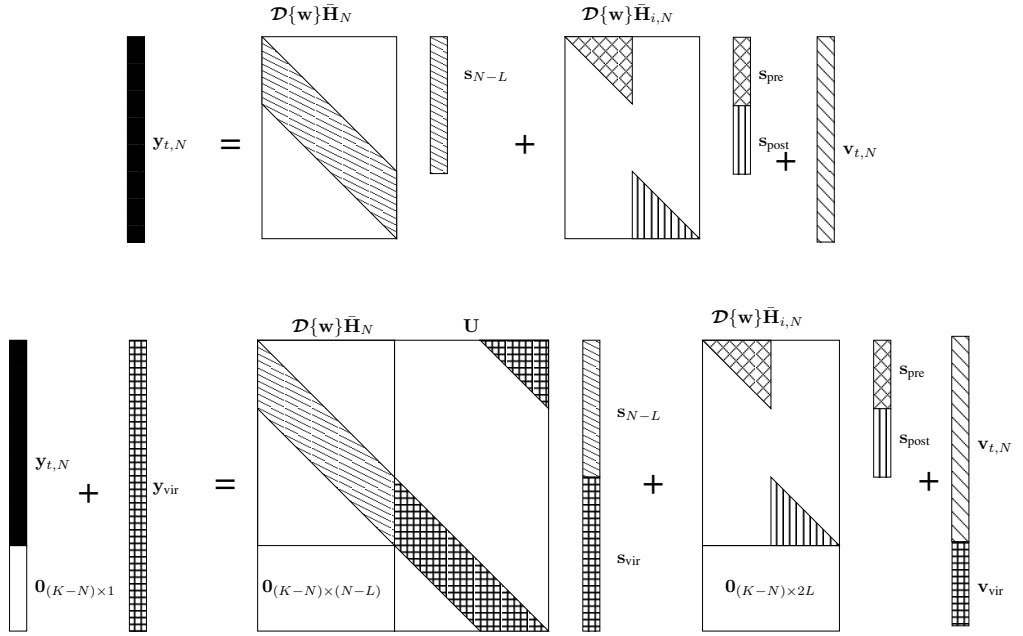

Fig. 3. Derivation of the extended data model.

Note that it can only be eliminated by the ZP or the NZP, which is in contrast to the ODM, where we could also eliminate the IBI using a CP.

Transformed into the frequency domain, the EDM in (26) becomes

$$\mathbf{y}_{f,K} := \mathbf{F}_K \mathbf{y}_K = \mathbf{H}_{f,K}\mathbf{F}_K \mathbf{s}_K + \boldsymbol{\epsilon}_{f,K} + \mathbf{v}_{f,K}, \qquad (28)$$

where $\boldsymbol{\epsilon}_{f,K}$ and $\mathbf{v}_{f,K}$ are similarly defined as $\mathbf{y}_{f,K}$, and $\mathbf{H}_{f,K} := \mathbf{F}_K \mathbf{H}_{t,K} \mathbf{F}_K^H$ stands for the FD channel matrix, which is again a full matrix. Like in the previous section, we use a banded matrix $\hat{\mathbf{H}}_{f,K}$ to replace $\mathbf{H}_{f,K}$, with $\hat{\mathbf{H}}_{f,K}$ having non-zero entries only on the main diagonal, the first $Q/2$ upper and the first $Q/2$ lower diagonals in a circular sense. Let us focus again on the block linear MMSE equalizer, and let us choose $\mathbf{s}_{\text{vir}}$ to contain some random symbols that have the same distribution as $\mathbf{s}_{N-L}$ such that $\mathcal{E}_s\{\mathbf{s}_K \mathbf{s}_K^H\} = \mathbf{I}_K$. Estimating the middle N_s data symbols of $\mathbf{s}_{N-L}$, where we neglect the IBI and replace $\mathbf{H}_{f,N}$ by $\hat{\mathbf{H}}_{f,N}$, we then obtain

$$\hat{\mathbf{s}}_{N_s} = \boldsymbol{\Omega}_K \mathbf{F}_K^H \hat{\mathbf{H}}_{f,K}^H (\hat{\mathbf{H}}_{f,K}\hat{\mathbf{H}}_{f,K}^H + \mathbf{R}_{v,K})^{-1} \mathbf{y}_{f,K}, \qquad (29)$$

where

$$\boldsymbol{\Omega}_K := \left[\mathbf{0}_{N_s \times \frac{N-L-N_s}{2}} \mathbf{I}_{N_s} \mathbf{0}_{N_s \times \frac{2K-N+L-N_s}{2}}\right], \qquad (30)$$

$$\begin{aligned}\mathbf{R}_{v,K} &:= \mathcal{E}_v\{\mathbf{v}_{f,K}\mathbf{v}_{f,K}^H\} \\ &= \mathbf{F}_K \begin{bmatrix} \sigma^2 \mathcal{D}\{\mathbf{w}\}\mathcal{D}\{\mathbf{w}^*\} & \\ & \mathcal{E}\{\mathbf{v}_{\text{vir}}\mathbf{v}_{\text{vir}}^H\} \end{bmatrix} \mathbf{F}_K^H. \end{aligned} \qquad (31)$$

As explained in the previous section, since $\hat{\mathbf{H}}_{f,K}\hat{\mathbf{H}}_{f,K}^H$ is banded with bandwidth $2Q+1$ and assuming that $\mathbf{R}_{v,K}$ is also banded with bandwidth $2Q+1$ (we come back to this issue later on), the required complexity to carry out (29) is $\mathcal{O}(KQ^2)$.

In summary, to enhance the equalization performance, the window of the EDM should take a three-fold task: 1) to make the noise covariance matrix $\mathbf{R}_{v,K}$ banded; 2) to minimize the IBI $\|\boldsymbol{\epsilon}_{f,K}\|^2$ in some average sense; and 3) to minimize the band approximation error $\|\mathbf{H}_{f,K} - \hat{\mathbf{H}}_{f,K}\|^2$ in some average sense.

B. Window Design for the EDM

Regarding the noise-shaping behavior of the window, we have the following proposition (the proof is similar to the proof of Proposition 1).

Proposition 2: The noise covariance matrix $\mathbf{R}_{v,K}$ will be banded with bandwidth $2Q+1$ if the window can be constructed as a weighted sum of $Q+1$ complex exponentials:

$$\mathbf{w} = \bar{\mathbf{B}}_N^{(0)} \mathbf{d}, \qquad (32)$$

where $\bar{\mathbf{B}}_N^{(0)}$ is an $N \times (Q+1)$ matrix with entries $[\bar{\mathbf{B}}_N^{(0)}]_{p,q} = \frac{1}{\sqrt{K}} e^{j\frac{2\pi}{K}p(q-\frac{Q}{2})}$; and $\mathbf{d}$ is a $(Q+1)\times 1$ vector containing all the weighting coefficients. In addition, the virtual noise $\mathbf{v}_{\text{vir}}$ must be designed such that

$$\mathcal{E}\{\mathbf{v}_{\text{vir}}\mathbf{v}_{\text{vir}}^H\} = \sigma^2 \mathcal{D}\{\bar{\mathbf{B}}_{K-N}^{(N)}\mathbf{d}\}\mathcal{D}\{\bar{\mathbf{B}}_{K-N}^{(N)}\mathbf{d}\}^H, \qquad (33)$$

where $\bar{\mathbf{B}}_{K-N}^{(N)}$ is a $(K-N)\times(Q+1)$ matrix with entries $[\bar{\mathbf{B}}_{K-N}^{(N)}]_{p,q} = \frac{1}{\sqrt{K}} e^{j\frac{2\pi}{K}(p+N)(q-\frac{Q}{2})}$.

The notations $\bar{\mathbf{B}}_N^{(0)}$ and $\bar{\mathbf{B}}_{K-N}^{(N)}$ stem from a more general notation $\bar{\mathbf{B}}_M^{(l)}$, which is defined next. Let us first define $\mathbf{B}_K$ as a $K \times (Q+1)$ matrix, which is comprised of the first $Q/2+1$ and the last $Q/2$ columns of the K-point DFT matrix $\mathbf{F}_K$. Then $\bar{\mathbf{B}}_M^{(l)}$ will denote an $M \times (Q+1)$ matrix consisting of the lth until $(l+M-1)$st row of $\mathbf{B}_K$. In mathematics, this means that $\bar{\mathbf{B}}_M^{(l)}$ has entries

$$[\bar{\mathbf{B}}_M^{(l)}]_{p,q} = \frac{1}{\sqrt{K}} e^{j\frac{2\pi}{K}(p+l)(q-\frac{Q}{2})}. \qquad (34)$$

Note that if we stack the two matrices $\bar{\mathbf{B}}_N^{(0)}$ and $\bar{\mathbf{B}}_{K-N}^{(N)}$ from Proposition 2 on top of each other, then we obtain $\mathbf{B}_K$, i.e., $[\bar{\mathbf{B}}_N^{(0)T}, \bar{\mathbf{B}}_{K-N}^{(N)T}]^T = \mathbf{B}_K$. That is why the proof of Proposition 2 can easily be derived from the proof of Proposition 1. The matrix $\bar{\mathbf{B}}_M^{(l)}$ tallies with the definition of the (O)CE-BEM, which uses an exponential period K and a BEM window size M with $K > M$ [19], [20]. The (O)CE-BEM plays an important role in minimizing the band approximation

error $\|\mathbf{H}_{f,K} - \hat{\mathbf{H}}_{f,K}\|^2$ as is evident from the following theorem (see Appendix C for a proof).

Theorem 3: The minimization of the band approximation error $\|\mathbf{H}_{f,K} - \hat{\mathbf{H}}_{f,K}\|^2$ in the frequency domain can be transformed in the time domain as the minimization of the (O)CE-BEM modeling error. More explicitly, we can write

$$\min_{\mathbf{U},\hat{\mathbf{H}}_{f,K}} \|\mathbf{H}_{f,K}-\hat{\mathbf{H}}_{f,K}\|^2 = \sum_{l=0}^{L} \min_{\mathbf{c}_l} \|\boldsymbol{\mathcal{D}}\{\boldsymbol{\Upsilon}_l\mathbf{w}\}\mathbf{h}_l-\bar{\mathbf{B}}_{N-L}^{(l)}\mathbf{c}_l\|^2, \tag{35}$$

where $\boldsymbol{\Upsilon}_l$ stands for an $(N-L)\times N$ selection matrix, $\boldsymbol{\Upsilon}_l := [\mathbf{0}_{(N-L)\times l}, \mathbf{I}_{N-L}, \mathbf{0}_{(N-L)\times(L-l)}]$; $\mathbf{h}_l$ for the $(N-L)\times 1$ vector collecting the lth channel tap from time index l to $N-L+l-1$, $\mathbf{h}_l := [h_{l,l}, \cdots, h_{N-L+l-1,l}]^T$; and $\mathbf{c}_l$ for the $(Q+1)\times 1$ vector collecting the BEM coefficients for the lth channel tap, $\mathbf{c}_l := [c_{0,l}, \cdots, c_{Q,l}]^T$.

In Theorem 3, we use for each channel tap $h_{n,l}$ a slightly different (O)CE-BEM matrix $\bar{\mathbf{B}}_{N-L}^{(l)}$ to approximate the time variation of that channel tap in the time interval from time index l to $N-L+l-1$. Adding the resulting BEM modeling errors for all channel taps accounts for the band approximation error.

Next, to minimize the IBI, the following theorem is useful (see Appendix D for a proof).

Theorem 4: The average power of the IBI $\boldsymbol{\epsilon}_{f,K}$ for the EDM is related to the window as

$$\mathcal{E}_{h,s}\{\|\boldsymbol{\epsilon}_{f,K}\|^2\} = \mathbf{w}^T\bar{\mathbf{R}}_{\epsilon,N}\mathbf{w}^*, \tag{36}$$

with $\bar{\mathbf{R}}_{\epsilon,N}$ denoting an $N\times N$ diagonal matrix with diagonal entries given by

$$[\bar{\mathbf{R}}_{\epsilon,N}]_{n,n} = \begin{cases} \sum_{l=n+L_z+1}^{L}\sigma_l^2 & \text{if } n \le L-L_z-1, \\ \sum_{l=0}^{n-N+L-L_z}\sigma_l^2 & \text{if } N-L+L_z \le n \le N-1, \\ 0 & \text{otherwise.} \end{cases} \tag{37}$$

The window design problem that jointly minimizes the average minimal band approximation error and the average IBI can then be formulated as

$$\begin{aligned} \min_{\mathbf{w}} \sum_{l=0}^{L} \mathcal{E}_h\big\{\min_{\mathbf{c}_l} \{\|\boldsymbol{\mathcal{D}}\{\boldsymbol{\Upsilon}_l\mathbf{w}\}\mathbf{h}_l - \bar{\mathbf{B}}_{N-L}^{(l)}\mathbf{c}_l\|^2\}\big\} \\ +\mathbf{w}^T\bar{\mathbf{R}}_{\epsilon,N}\mathbf{w}^*, \\ \text{s.t. } \mathbf{w} = \bar{\mathbf{B}}_N^{(0)}\mathbf{d} \text{ and } \|\mathbf{w}\|^2 = N. \end{aligned} \tag{38}$$

We solve the above first for $\mathbf{c}_l$ resulting into $\mathbf{c}_l = \bar{\mathbf{B}}_{N-L}^{(l)\dagger}\boldsymbol{\mathcal{D}}\{\boldsymbol{\Upsilon}_l\mathbf{w}\}\mathbf{h}_l$. Plugging this result into the lth term of (38), and using the property $\boldsymbol{\mathcal{D}}\{\boldsymbol{\Upsilon}_l\mathbf{w}\} = \boldsymbol{\Upsilon}_l\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\boldsymbol{\Upsilon}_l^H$, we obtain

$$\begin{aligned} &\mathcal{E}_h\big\{\min_{\mathbf{c}_l} \{\|\boldsymbol{\mathcal{D}}\{\boldsymbol{\Upsilon}_l\mathbf{w}\}\mathbf{h}_l - \bar{\mathbf{B}}_{N-L}^{(l)}\mathbf{c}_l\|^2\}\big\} \\ &= \mathrm{tr}\big(\boldsymbol{\mathcal{P}}_{\bar{\mathbf{B}}_{N-L}^{(l)}}\boldsymbol{\Upsilon}_l\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\boldsymbol{\Upsilon}_l^H\mathbf{R}_{\mathbf{h}_l,N-L}\boldsymbol{\Upsilon}_l\boldsymbol{\mathcal{D}}\{\mathbf{w}^*\}\boldsymbol{\Upsilon}_l^H\boldsymbol{\mathcal{P}}_{\bar{\mathbf{B}}_{N-L}^{(l)}}^H\big) \\ &= \mathbf{w}^T\Big(\sum_{n=0}^{N-L-1}\boldsymbol{\Upsilon}_l^T\boldsymbol{\mathcal{D}}\{\boldsymbol{\mathcal{P}}_{\bar{\mathbf{B}}_{N-L}^{(l)}}^T\mathbf{e}_n\}\mathbf{R}_{\mathbf{h}_l,N-L} \\ &\qquad \times \boldsymbol{\mathcal{D}}\{\boldsymbol{\mathcal{P}}_{\bar{\mathbf{B}}_{N-L}^{(l)}}^H\mathbf{e}_n\}\boldsymbol{\Upsilon}_l^*\Big)\mathbf{w}^*, \end{aligned} \tag{39}$$

where

$$\boldsymbol{\mathcal{P}}_{\bar{\mathbf{B}}_{N-L}^{(l)}} := \mathbf{I}_{N-L} - \bar{\mathbf{B}}_{N-L}^{(l)}\bar{\mathbf{B}}_{N-L}^{(l)\dagger} \tag{40}$$

$$\mathbf{R}_{\mathbf{h}_l,N-L} := \mathcal{E}_h\{\mathbf{h}_l\mathbf{h}_l^H\}. \tag{41}$$

Note that using Assumption 1, the entries of $\mathbf{R}_{\mathbf{h}_l,N-L}$ can be expressed as $[\mathbf{R}_{\mathbf{h}_l,N-L}]_{m,n} = \sigma_l^2\gamma_{m-n}$. Substituting (32) and (39) in (38), finally results into

$$\min_{\mathbf{d}} \mathbf{d}^T\bar{\boldsymbol{\mathcal{X}}}_N\mathbf{d}^*, \quad \text{s.t. } \|\bar{\mathbf{B}}_N^{(0)}\mathbf{d}\|^2 = N, \tag{42}$$

with

$$\begin{aligned} \bar{\boldsymbol{\mathcal{X}}}_N := \bar{\mathbf{B}}_N^{(0)T}\Big(\sum_{l=0}^{L}\sum_{n=0}^{N-L-1}\boldsymbol{\Upsilon}_l^T\boldsymbol{\mathcal{D}}\{\boldsymbol{\mathcal{P}}_{\bar{\mathbf{B}}_{N-L}^{(l)}}^T\mathbf{e}_n\}\mathbf{R}_{\mathbf{h}_l,N-L} \\ \times \boldsymbol{\mathcal{D}}\{\boldsymbol{\mathcal{P}}_{\bar{\mathbf{B}}_{N-L}^{(l)}}^H\mathbf{e}_n\}\boldsymbol{\Upsilon}_l^* + \bar{\mathbf{R}}_{\epsilon,N}\Big)\bar{\mathbf{B}}_N^{(0)*}. \end{aligned} \tag{43}$$

To resolve (42), we note that the columns of $\bar{\mathbf{B}}_N^{(0)}$ are not orthonormal to each other. Therefore, we have to compute $\mathbf{d}$ as the least significant generalized eigenvector of the matrix pair $(\bar{\boldsymbol{\mathcal{X}}}_N^*, \bar{\mathbf{B}}_N^{(0)H}\bar{\mathbf{B}}_N^{(0)})$ [25].

Remark 3: For the EDM, the band approximation error is minimized by tuning not only the banded matrix $\hat{\mathbf{H}}_{f,K}$ itself but also the matrix $\mathbf{U}$ as part of $\mathbf{H}_{f,K}$ [c.f. (35)], instead of tuning only the banded matrix $\hat{\mathbf{H}}_{f,N}$ as in the ODM case. This already shows that the EDM has a better band approximation error than the ODM. Related to this, for the EDM, we are able to transform the band approximation error to the (O)CE-BEM modeling error [c.f. (35)], instead of to the (C)CE-BEM modeling error as in the ODM case. Since the (O)CE-BEM is much tighter than the (C)CE-BEM [22], this also explains why the EDM has an improved band approximation error over the ODM.

V. Numerical results

We test the proposed algorithms over a time-varying channel following Jakes' Doppler profile [27] using the time-varying channel generator given in [28]. The channel is assumed to have $L+1 = 31$ channel taps with the lth tap having variance $\sigma_l^2 = e^{-\frac{l}{10}}$. The Jakes' Doppler profile is characterized by the normalized time correlation $\gamma_m = J_0(2\pi\nu m)$, where $J_0(\cdot)$ denotes the zeroth-order Bessel function of the first kind, and ν stands for the normalized Doppler spread, which is obtained as $\nu = \frac{vf}{c}T$, where v denotes the vehicle velocity, f the carrier frequency, T the data symbol duration, and c the speed of light. We will test two types of time-varying channels, $\nu = 0.002$ and $\nu = 0.004$, through which QPSK data symbols are transmitted.

Test case 1. Without IBI. In the first test case, we let the observation block length in the ODM and EDM be $N = 128$, and a sufficiently long ZP of length $L_z = L = 30$ is used such that the IBI is completely removed. As a result, we can estimate the first $N_s = 98$ data symbols in one shot. Further, we set $Q = 4$ for the ODM, while $Q = 2$ and $K = 256$ for the EDM. With those parameters chosen, the complexity of the ODM $\mathcal{O}(NQ^2)$ is higher than that of the EDM $\mathcal{O}(KQ^2)$. For the sake of simplicity, we will set $\mathbf{s}_{\text{vir}} = \mathbf{0}$ and $\mathbf{v}_{\text{vir}} = \mathbf{0}$ for the EDM. Although this breaks some conditions that we assumed to hold, we have seen that this simplification only brings small performance differences (not shown here).

The windows of the ODM and EDM are depicted in Fig. 4, where we can see that the windows of the ODM take on a bell-shape with the edges tending to zero, which is beneficial

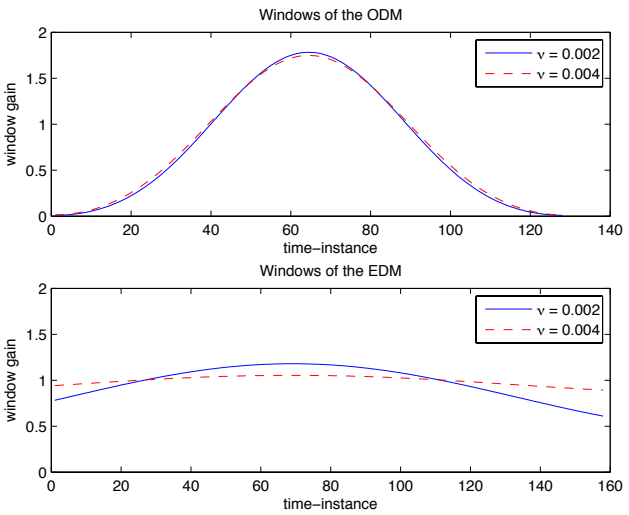

Fig. 4. The windows in the absence of IBI.

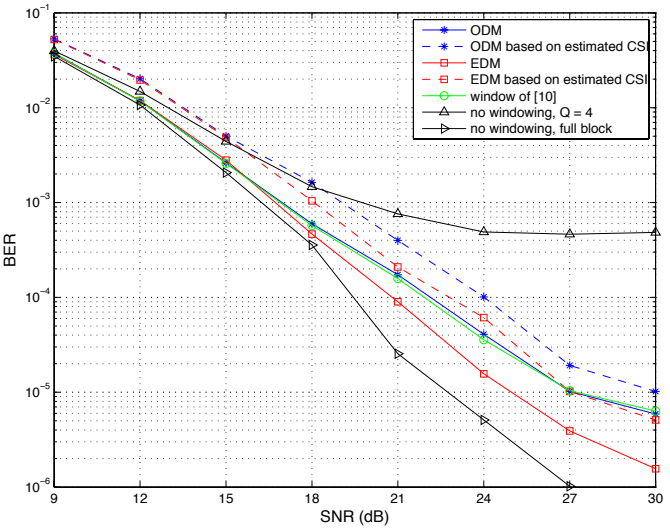

Fig. 5. BER in the absence of IBI, $\nu = 0.002$.

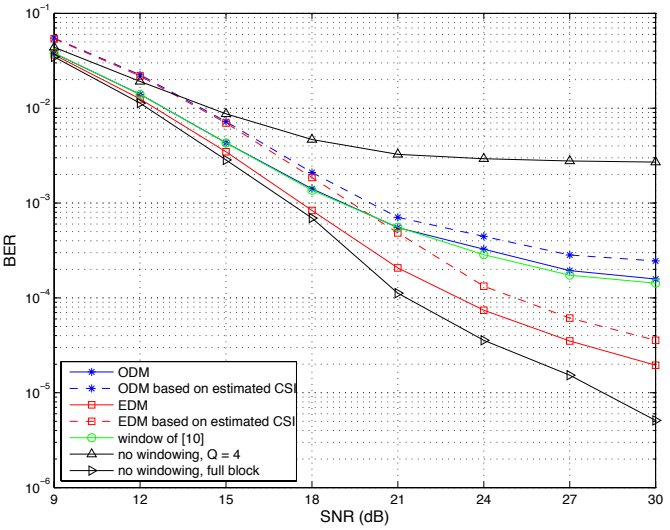

Fig. 6. BER in the absence of IBI, $\nu = 0.004$.

to the modeling performance of the (C)CE-BEM. In contrast, the windows of the EDM are almost flat, suggesting that the (O)CE-BEM itself is quite accurate in modeling the channel time variation.

The bit error rate (BER) is compared in Fig. 5 and Fig. 6 for $\nu = 0.002$ and $\nu = 0.004$, respectively. For comparison, we have also plotted the performance of the same block linear MMSE equalizer but using the window of [10], which has the same length as the window of the ODM. Besides, we also show the performance of the ODM without windowing and the performance of the block linear MMSE equalizer applied in the time domain, as discussed at the end of Section III-A. Note that the latter does not suffer from any band approximation errors. From Fig. 5 and Fig. 6, we can see that the ODM without windowing performs the worst while the time-domain approach performs the best. It is obvious that the band approximation error plays a significant role. For the same reason, the EDM renders a better performance than the ODM even with a lower complexity. This performance lead is larger at a higher Doppler spread, where the band approximation error is more pronounced. As we have predicted, the window of [10] yields a performance that is very close to that of the ODM. Note though that [10] focuses on an iterative serial MMSE equalizer, whereas we only implement block linear MMSE equalizers.

The above equalizers are constructed based on perfect channel state information (CSI). The performance of the equalizers of the ODM and EDM based on estimated CSI is also exhibited in Fig. 5 and Fig. 6. The channel is estimated in the time domain with the aid of pilots, which are interleaved with data symbols as described in Fig. 7. We refer the interested reader to [18] for more details about the considered channel estimation scheme. Note that the proposed training scheme is not very spectrally efficient, but this is the price we have to pay for estimating such highly time-varying channels. The spectral efficiency could be somewhat improved by using iterative channel estimation schemes.

Test case 2. With IBI. Here, we examine the performance when no guard interval is embedded. To combat the impact of the IBI, we adopt the sliding window approach, where in each time only the middle $N_s = 64$ data symbols of a block of $N - L$ data symbols are estimated, and afterwards, the equalizer moves forward to estimate the next $N_s = 64$ data symbols. The final BER is an average of all these data estimates. In an effort to present a complete picture, we compare four different schemes for the ODM and EDM, respectively. The parameters for these schemes are summarized in Table I and Table II. Note that the third column represents the order of magnitude of the complexity per estimated data symbol. For comparison, we also include the performance of the window proposed in [12], which only has a slightly longer window length as the ODM ($N + L$ instead of N). The windows of the ODM and EDM for this test case are depicted in Fig. 8, where we can see that the windows of the EDM are not flat anymore but also take on a bell-shape to account for the IBI. The windows of the ODM still have their typical bell-shape.

The performance for channels with $\nu = 0.002$ is plotted in Fig. 9, where we find that the EDM actually performs worse than the other methods. Compared with Test Case 1, the advantage of the EDM due to a better band approximation

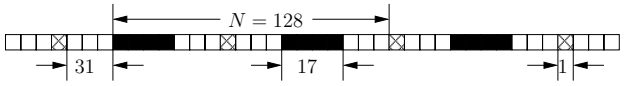

Fig. 7. An illustration of the pilot structure: the black boxes represent data; the blank boxes zero pilots; and the hatched boxes non-zero pilots.

TABLE I
ODM PARAMETERS

	N	Q	$NQ^2/64$
ODM-I	128	2	8
ODM-II	128	4	32
ODM-III	256	2	16
ODM-IV	256	4	64

is now nullified by the presence of IBI. Indeed, for the same DFT size, the observation block length of the EDM is chosen to be much shorter than for the other methods, and thus the EDM is more susceptible to the impact of IBI.

A different story is depicted in Fig. 10 where the performance is shown for channels with $\nu = 0.004$. For such a high Doppler spread, the band approximation error is more prominent while the impact of IBI remains unchanged. In this case, the ODM degrades, and the EDM regains its lead in some cases.

From Fig. 9 and Fig. 10, it can also be remarked that the window of [12] renders a performance that is similar to that of the ODM.

Test case 3. With Partial IBI. In a practical system, it is reasonable to assume that the IBI from the previous block can be completely removed by perfectly estimating $\mathbf{s}_{\text{pre}}$. Then only the IBI resulting from $\mathbf{s}_{\text{post}}$ needs to be combatted and the window design can be adapted accordingly. In this case, we can see from Fig. 11 that the windows of the EDM are only bended to zero at the right edge where the IBI is still present, while the windows of the ODM keep their typical bell-shape. With less influence from the IBI, the EDM is able to produce a better performance than the ODM in all situations, as we can observe from Fig. 12 and Fig. 13.

VI. CONCLUSIONS

In this paper, we have discussed how a single-carrier system that is plagued by fast fading and IBI can be effectively equalized in the frequency domain. For the sake of complexity, the FD channel matrix is approximated to be banded. Two data models have been discussed that can both reduce the band approximation error: 1) the original data model (ODM); 2) the extended data model (EDM). We have established a link between the band approximation error in the ODM and the modeling error of the (C)CE-BEM, and a link between the band approximation error in the EDM and the modeling error of the (O)CE-BEM. The (O)CE-BEM is known to yield a much tighter fit than the (C)CE-BEM. It has been shown in the simulations that although the EDM is not really effective in combating the IBI, its superior band approximation performance makes it still an appealing alternative for equalizing a fast varying channel in the frequency domain.

APPENDIX A
PROOF OF THEOREM 1

It can be shown that for the banded $\hat{\mathbf{H}}_{f,N}$, its time-domain counterpart $\hat{\mathbf{H}}_{t,N} := \mathbf{F}_N^H \hat{\mathbf{H}}_{f,N} \mathbf{F}_N$ can be uniquely expressed as

$$\hat{\mathbf{H}}_{t,N} = \sum_{q=0}^{Q} \boldsymbol{\mathcal{D}}\{\mathbf{B}_N \mathbf{e}_q\} \mathbf{C}_q, \tag{44}$$

where $\mathbf{C}_q$ stands for a circulant matrix with $[c_{q,0}, \cdots, c_{q,N-1}]^T$ as its first yet-to-be-designed column. Hence, the band approximation error in the frequency domain can be expressed in the time domain as

$$\|\mathbf{H}_{f,N} - \hat{\mathbf{H}}_{f,N}\|^2 = \|\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\mathbf{H}_{t,N} - \hat{\mathbf{H}}_{t,N}\|^2. \tag{45}$$

If we want to minimize (45) it is clear that since $\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\mathbf{H}_{t,N}$ has zeros outside its first $L+1$ lower diagonals in a circular sense, we should also design $\hat{\mathbf{H}}_{t,N}$ to have zeros outside its first $L+1$ lower diagonals in a circular sense (note that the main diagonal is assumed to be included here), which can be realized by taking $c_{q,L+1}, \cdots, c_{q,N-1} = 0$. So, $\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\mathbf{H}_{t,N}$ and $\hat{\mathbf{H}}_{t,N}$ are different only in these $L+1$ diagonals, which for $\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\mathbf{H}_{t,N}$ can be written as $\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\boldsymbol{\mathcal{H}}$ and for $\hat{\mathbf{H}}_{t,N}$ can be written as $\mathbf{B}_N\boldsymbol{\mathcal{C}}$. This concludes the proof.

TABLE II
EDM PARAMETERS

	N	K	Q	$KQ^2/64$
EDM-I	94	128	2	8
EDM-II	94	128	4	32
EDM-III	158	256	2	16
EDM-IV	158	256	4	64

APPENDIX B
PROOF OF THEOREM 2

By the definition of $\boldsymbol{\epsilon}_{t,N}$ in (6) and using Assumption 2, we understand that

$$\begin{aligned}\mathcal{E}_{h,s}\{\|\boldsymbol{\epsilon}_{f,N}\|^2\} &= \mathcal{E}_{h,s}\{\|\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\mathbf{H}_{i,N}(\mathbf{s}_{\text{pre}} - \mathbf{s}_{\text{post}})\|^2\} \\ &= 2\text{tr}\big(\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\mathcal{E}_h\{\mathbf{H}_{i,N}\boldsymbol{\Phi}_L\mathbf{H}_{i,N}\}\boldsymbol{\mathcal{D}}\{\mathbf{w}^H\}\big),\end{aligned} \tag{46}$$

where an $L \times L$ diagonal matrix $\boldsymbol{\Phi}_L$ is introduced to account for a possible guard interval in $\mathbf{s}_{\text{pre}}$ and $\mathbf{s}_{\text{post}}$:

$$\boldsymbol{\Phi}_L := \boldsymbol{\mathcal{D}}\{[\mathbf{1}_{1\times(L-L_z)}, \mathbf{0}_{1\times L_z}]^T\}. \tag{47}$$

It can be shown that under Assumption 1, we obtain

$$\mathcal{E}_h\{\mathbf{H}_{i,N}\boldsymbol{\Phi}_L\mathbf{H}_{i,N}^H\} = \mathbf{R}_{\epsilon,N}, \tag{48}$$

with $\mathbf{R}_{\epsilon,N}$ defined as in (16). Substituting the above in (46) concludes the proof.

APPENDIX C
PROOF OF THEOREM 3

Similar to Appendix A, we can show that for the banded $\hat{\mathbf{H}}_{f,K}$, its time-domain counterpart $\hat{\mathbf{H}}_{t,K} := \mathbf{F}_K^H \hat{\mathbf{H}}_{f,K} \mathbf{F}_K$ can be expressed as

$$\hat{\mathbf{H}}_{t,K} = \sum_{q=0}^{Q} \boldsymbol{\mathcal{D}}\{\mathbf{B}_K \mathbf{e}_q\} \mathbf{C}_q, \tag{49}$$

where $\mathbf{C}_q$ stands for a circulant matrix with $[c_{q,0}, \cdots, c_{q,K-1}]^T$ as its first yet-to-be-designed column. As a result, the band approximation error in the frequency domain can be transformed in the time domain as

$$\|\mathbf{H}_{f,K} - \hat{\mathbf{H}}_{f,K}\|^2 = \|\mathbf{H}_{t,K} - \hat{\mathbf{H}}_{t,K}\|^2. \tag{50}$$

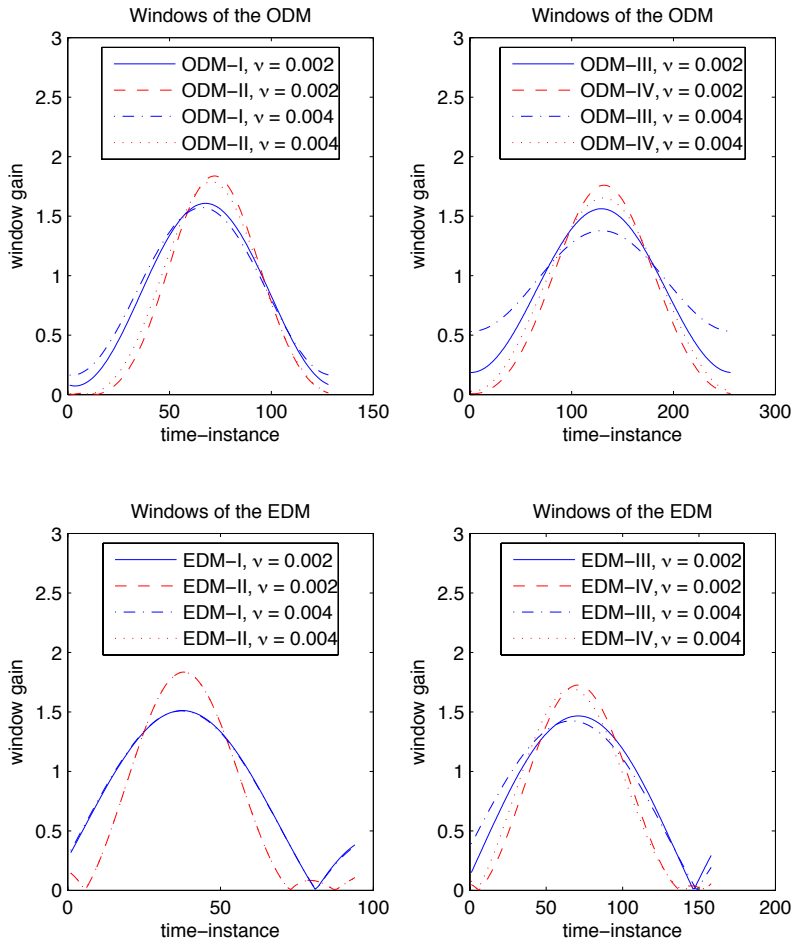

Fig. 8. The windows in the presence of IBI.

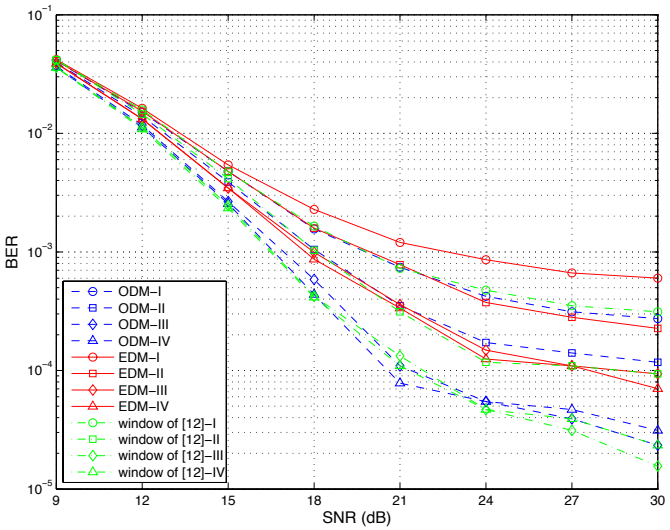

Fig. 9. BER in the presence of IBI, $\nu = 0.002$.

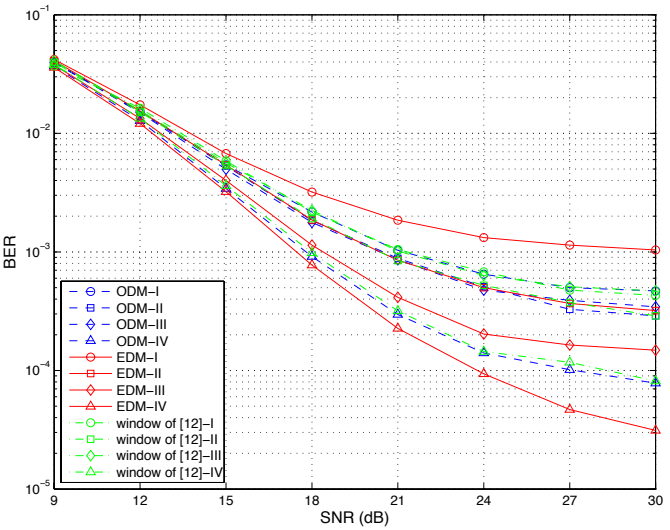

Fig. 10. BER in the presence of IBI, $\nu = 0.004$.

Recall from (26) that $\mathbf{H}_{t,K}$ contains the $K \times (K-N+L)$ matrix $\mathbf{U}$, which is subject to design. Hence, if we want to minimize (50), we should take $\mathbf{U}$ equal to the last $K-N+L$ columns of $\hat{\mathbf{H}}_{t,K}$. In this way, $\mathbf{H}_{t,K}$ and $\hat{\mathbf{H}}_{t,K}$ will be different only in their first $N-L$ columns. Moreover, since $\mathbf{H}_{t,K}$ has zeros outside its $L+1$ most significant diagonals of its first $N-L$ columns, we should also design $\hat{\mathbf{H}}_{t,K}$ to have zeros outside its $L+1$ most significant diagonals of its first $N-L$ columns, which can be realized by taking $c_{q,L+1}, \cdots, c_{q,K-1} = 0$. This way, we come to a point where $\mathbf{H}_{t,K}$ and $\hat{\mathbf{H}}_{t,K}$ are different only in their $L+1$ most significant diagonals of their first $N-L$ columns.

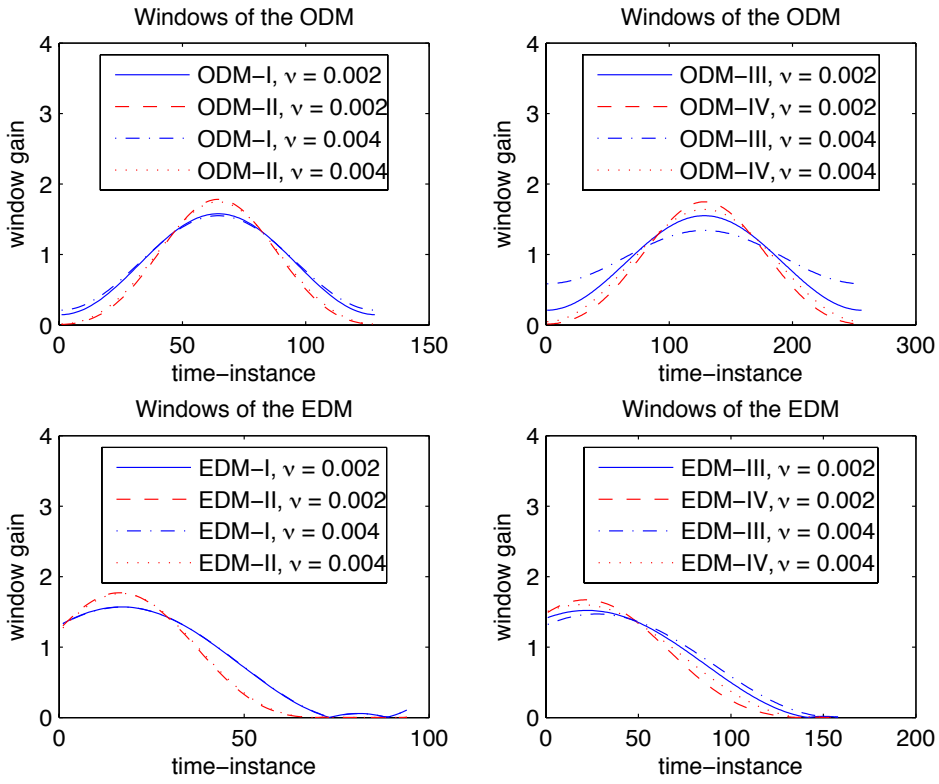

Fig. 11. The windows in the presence of partial IBI.

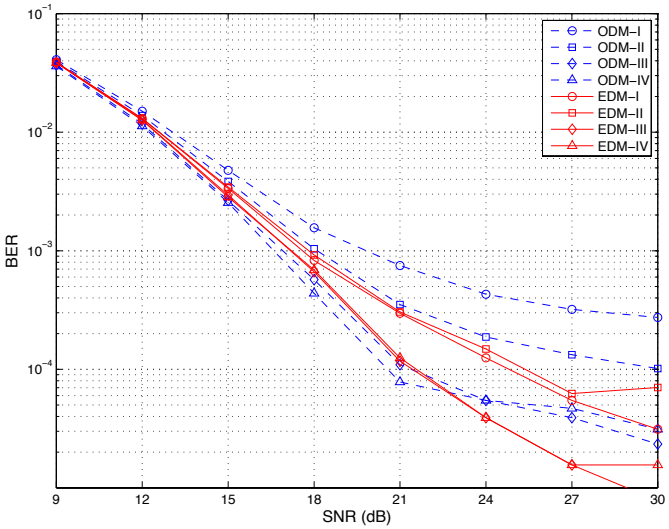

Fig. 12. BER in the presence of partial IBI, $\nu = 0.002$.

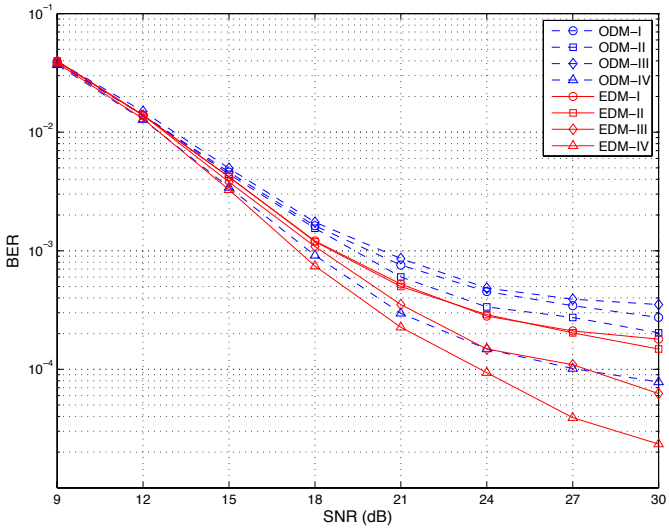

Fig. 13. BER in the presence of partial IBI, $\nu = 0.004$.

The lth diagonal of the first $N-L$ columns of $\mathbf{H}_{t,K}$ corresponds to the lth channel tap from time index l to $(N-L+l)$, which can thus be expressed as

$$[w_l h_{l,l}, \cdots, w_{N-L+l} h_{N-L+l,L}]^T = \boldsymbol{\mathcal{D}}\{\boldsymbol{\Upsilon}_l \mathbf{w}\}\mathbf{h}_l. \quad (51)$$

From (49), it can further be derived that the lth diagonal of the first $N-L$ columns of $\hat{\mathbf{H}}_{t,K}$ can be expressed as

$$\left[[\hat{\mathbf{H}}_{t,K}]_{l,0}, \cdots, [\hat{\mathbf{H}}_{t,K}]_{N-L-1,N-L-l-1}\right]^T = \bar{\mathbf{B}}^{(l)}_{N-L}\mathbf{c}_l. \quad (52)$$

Hence, if we want to minimize (50) over both $\mathbf{U}$ and $\hat{\mathbf{H}}_{t,K}$, we have to minimize the norm squared of the difference between (51) and (52) over $\mathbf{c}_l$ for every $l = 0, 1, \cdots, L$. This is quivalent to (35), which concludes the proof.

Appendix D
Proof of Theorem 4

By the definition of $\bar{\boldsymbol{\epsilon}}_{t,N}$ in (27) and using Assumption 2, we understand that

$$\begin{aligned}\mathcal{E}_{h,s}\{\|\boldsymbol{\epsilon}_{f,K}\|^2\} &= \mathcal{E}_{h,s}\{\|\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\bar{\mathbf{H}}_{i,N}\begin{bmatrix}\mathbf{s}_{\text{pre}}\\ \mathbf{s}_{\text{post}}\end{bmatrix}\|^2\}\\ &= \text{tr}\big(\boldsymbol{\mathcal{D}}\{\mathbf{w}\}\mathcal{E}_h\{\bar{\mathbf{H}}_{i,N}\bar{\boldsymbol{\Phi}}_N\bar{\mathbf{H}}^H_{i,N}\}\boldsymbol{\mathcal{D}}\{\mathbf{w}^H\}\big),\end{aligned} \quad (53)$$

with

$$\bar{\mathbf{\Phi}}_N := \boldsymbol{\mathcal{D}}\{[\mathbf{1}_{1\times(L-L_z)}, \mathbf{0}_{1\times 2L_z}, \mathbf{1}_{1\times(L-L_z)}]^T\}, \qquad (54)$$

which is introduced to account for the presence of a guard interval (ZP or NZP). Under Assumption 1, it can be shown that

$$\mathcal{E}_h\{\bar{\mathbf{H}}_{i,N}\bar{\mathbf{\Phi}}_N\bar{\mathbf{H}}_{i,N}^H\} = \bar{\mathbf{R}}_{\epsilon,N}, \qquad (55)$$

with $\bar{\mathbf{R}}_{\epsilon,N}$ defined as in (37). Substituting the above in (53) concludes the proof.

REFERENCES

[1] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, pp. 58–66, Apr. 2002.

[2] K. Van Acker, G. Leus, M. Moonen, and T. Pollet, "Improved initialization for time domain equalization in adsl," *Elsevier Signal Processing*, vol. 84, pp. 1895–1908, Oct 2004.

[3] I. Barhumi, G. Leus, and M. Moonen, "Time-domain and frequency-domain per-tone equalization for OFDM in doubly-selective channels," *Elsevier Signal Processing*, vol. 84, pp. 2055–2066, Nov. 2004.

[4] I. Barhumi, G. Leus, and M. Moonen, "Equalization for OFDM over doubly-selective channels," *IEEE Trans. Signal Processing*, vol. 54, pp. 1445–1458, Apr. 2006.

[5] A. Stamoulis, S. N. Diggavi, and N. Al-Dhahir, "Intercarrier interference in MIMO OFDM," *IEEE Trans. Signal Processing*, vol. 50, pp. 2451–2464, Oct. 2002.

[6] X. Cai and G. B. Giannakis, "Bounding performance and suppressing intercarrier interference in wireless mobile OFDM," *IEEE Trans. Commun.*, vol. 51, pp. 2047–2056, Dec. 2003.

[7] W. G. Jeon, K. H. Chang, and Y. S. Cho, "An equalization technique for orthogonal frequency-division multiplexing systems in time-variant multipath channels," *IEEE Trans. Commun.*, vol. 47, pp. 27–32, Jan. 1999.

[8] L. Rugini, P. Banelli, and G. Leus, "Simple equalization of time-varying channels for OFDM," *IEEE Commun. Lett.*, vol. 9, pp. 619–621, July 2005.

[9] L. Rugini, P. Banelli, and G. Leus, "Low-complexity banded equalizers for OFDM systems in Doppler spread channels," *EURASIP Journal on Applied Signal Processing*, pp. Article ID 67404, 13 pages, 2006.

[10] P. Schniter and H. Liu, "Iterative equalization for single-carrier cyclic-prefix in doubly dispersive channels," *Asilomar Conference on Signals, Systems, and Computers*, pp. 502–506, Nov 2003.

[11] P. Schniter, "Low-complexity equalization of OFDM in doubly-selective channels," *IEEE Trans. Signal Processing*, vol. 52, pp. 1002–1011, Apr 2004.

[12] P. Schniter and H. Liu, "Iterative frequency-domain equalization for single-carrier systems in doubly-dispersive channels," *Asilomar Conference on Signals, Systems, and Computers*, pp. 667–671, Nov 2004.

[13] S. Ohno, "Maximum likelihood inter-carrier interference suppression for wireless OFDM with null subcarriers," *International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, vol. 3, pp. 849 – 852, Mar. 2005.

[14] J. Kim, J. R. W. Heath, and E. Powers, "Receiver designs for Alamouti coded OFDM systems in fast fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 550–559, Mar. 2005.

[15] S.-J. Hwang and P. Schniter, "Efficient sequence detection of multicarrier transmissions over doubly dispersive channels," *EURASIP J. Applied Signal Processing*, vol. 2006, pp. Article ID 93638, 17 pages, 2006.

[16] M. K. Tsatsanis and G. B. Giannakis, "Modeling and equalization of rapidly fading channels," *International J. Adaptive Control and Signal Processing*, vol. 10, pp. 159–176, Mar. 1996.

[17] G. B. Giannakis and C. Tepedelenlioğlu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proc. IEEE*, vol. 86, pp. 1969–1986, Oct. 1998.

[18] X. Ma, G. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly-selective fading channels," *IEEE Trans. Signal Processing*, vol. 51, pp. 1351–1366, May 2003.

[19] T. A. Thomas and F. W. Vook, "Multi-user frequency-domain channel identification, interference suppression, and equalization for time-varying broadband wireless communications," *Proceedings of the 2000 IEEE Sensor Array and Multichannel Signal Processing Workshop*, pp. 444–448, Mar. 2000.

[20] G. Leus, "On the estimation of rapidly time-varying channels," *European Signal Processing Conference, EUSIPCO*, pp. 2227–2230, Sept. 2004.

[21] T. Cui, C. Tellambura, and Y. Wu, "Low-complexity pilot-aided channel estimation for OFDM systems over doubly-selective channels," *IEEE International Conference on Communications, ICC*, vol. 3, pp. 1980–1984, May 2005.

[22] 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.

[23] Z. Wang, X. Ma, and G. B. Giannakis, "OFDM or single-carrier block transmissions?," *IEEE Trans. Commun.*, vol. 52, pp. 380–394, Mar. 2004.

[24] L. Deneire, B. Gyselinckx, and M. Engels, "Training sequence versus cyclic prefix - a new look on single carrier communication," *IEEE Commun. Lett.*, vol. 7, no. 5, pp. 292–294, 2001.

[25] G. H. Golub and C. F. van Loan, *Matrix Computations*. The Johns Hopkins University Press, 1989.

[26] Z. Tang, *OFDM Transmission over Rapidly Changing Channels*. Ph.D Disseration at Delft University of Technology, the Netherlands, 2007.

[27] W. C. Jakes, *Microwave Mobile Channels*. New York: Wiley, 1974.

[28] Y. R. Zheng and C. Xiao, "Simulation models with correct statistical properties for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 51, pp. 920–928, June 2003.

Zijian Tang Zijian Tang received the M.Sc. degree in 2003 and the Ph.D. degree in 2007 both in electrical engineering from Delft University of Technology, the Netherlands. In 2003, he took an internship in the R&D center of Philips. From 2003 until 2007, he was with the Circuits & Systems group at Delft University of Technology as a researcher. Since 2007, he has been working with The MathWorks as an application engineer. His research is mainly focused on signal processing for communication, especially in the area of time-varying channels.

Geert Leus Geert Leus was born in Leuven, Belgium, in 1973. He received the electrical engineering degree and the PhD degree in applied sciences from the Katholieke Universiteit Leuven, Belgium, in June 1996 and May 2000, respectively. He has been a Research Assistant and a Postdoctoral Fellow of the Fund for Scientific Research - Flanders, Belgium, from October 1996 till September 2003. During that period, Geert Leus was affiliated with the Electrical Engineering Department of the Katholieke Universiteit Leuven, Belgium. Currently, Geert Leus is an Associate Professor at the Faculty of Electrical Engineering, Mathematics and Computer Science of the Delft University of Technology, The Netherlands. During the summer of 1998, he visited Stanford University, and from March 2001 till May 2002 he was a Visiting Researcher and Lecturer at the University of Minnesota. His research interests are in the area of signal processing for communications. Geert Leus received a 2002 IEEE Signal Processing Society Young Author Best Paper Award and a 2005 IEEE Signal Processing Society Best Paper Award. He is the Vice-Chair of the IEEE Signal Processing for Communications Technical Committee, and an Associate Editor for the IEEE Transactions on Signal Processing and the EURASIP Journal on Applied Signal Processing. In the past, he has served on the Editorial Board of the IEEE Signal Processing Letters and the IEEE Transactions on Wireless Communications.