Hindawi Publishing Corporation
EURASIP Journal on Advances in Signal Processing
Volume 2010, Article ID 974652, 13 pages
doi:10.1155/2010/974652

Research Article

Block Transmissions over Doubly Selective Channels: Iterative Channel Estimation and Turbo Equalization

Kun Fang,[1] Luca Rugini,[2] and Geert Leus[1]

[1] *Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS), Delft University of Technology, 2628 CD Delft, The Netherlands*
[2] *Department of Electronic and Information Engineering, University of Perugia, 06125 Perugia, Italy*

Correspondence should be addressed to Luca Rugini, luca.rugini@diei.unipg.it

Received 7 January 2010; Accepted 8 April 2010

Academic Editor: Xiaoli Ma

Copyright © 2010 Kun Fang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Modern wireless communication systems require high transmission rates, giving rise to frequency selectivity due to multipath propagation. In addition, high-mobility terminals and scatterers induce Doppler shifts that introduce time selectivity. Therefore, advanced techniques are needed to accurately model the time- and frequency-selective (i.e., doubly selective) channels and to counteract the related performance degradation. In this paper, we develop new receivers for orthogonal frequency-division multiplexing (OFDM) systems and single-carrier (SC) systems in doubly selective channels by embedding the channel estimation task within low-complexity block turbo equalizers. Linear minimum mean-squared error (MMSE) pilot-assisted channel estimators are presented, and the soft data estimates from the turbo equalizers are used to improve the quality of the channel estimates.

1. Introduction

Broadband wireless communication systems require high transmission rates, giving rise to frequency selectivity caused by multipath propagation, and consequently to intersymbol interference (ISI). In addition, recent wireless communication standards, such as WiMAX and Long-Term Evolution (LTE), also need to support high mobile speeds, leading to high-mobility terminals and scatterers that introduce Doppler shifts and time selectivity, that is, intercarrier interference (ICI). Due to the concomitant presence of ISI and ICI, specialized techniques are necessary to counteract the related performance degradation. However, with a properly designed transceiver, time- and frequency-selective (i.e., doubly selective) channels can even provide multiplicative delay-Doppler diversity gains [1, 2].

LTE is a major 3GPP step in next generation wireless networks [3]. The LTE physical layer relies on a multiple-access scheme based on orthogonal frequency-division multiplexing (OFDM) in the downlink, and on single-carrier frequency-division multiple access (SC-FDMA) in the uplink [3]. In both cases, the transmission scheme is blockwise, and a cyclic prefix (CP) is included in each data block in order to eliminate the ISI between consecutive data blocks. OFDM and single-carrier (SC) block transmissions share some similarities: since an SC system can be viewed as a discrete Fourier transform (DFT) precoded OFDM system [4], performance and complexity are comparable, but part of the complexity (i.e., an inverse DFT) is moved from the transmitter to the receiver [4]. However, there are also some important differences: with respect to OFDM, SC has a lower peak-to-average power ratio, and hence power-efficient terminals are suitable for the uplink [5]. However, both SC and OFDM systems suffer from doubly selective channels, and call for appropriate ICI mitigation methods.

A possible way to counteract a doubly selective channel is by means of iterative equalizers. The iterative approach, inspired by the turbo equalization principle [6, 7], exchanges soft information between the channel equalizer and the decoder, in an iterative fashion, and greatly improves the system performance. In the last fifteen years, many turbo equalizers have been proposed for time-invariant frequency-selective channels (see [6–11], and the references therein). More recently, the turbo approach has been proposed also

for doubly selective channel equalization, which is more challenging due to the time variation of the channel. For OFDM systems with doubly selective channels, low-complexity minimum mean-squared error (MMSE) turbo equalizers have been proposed in [12, 13]. The turbo equalizers [12, 13], which are based on frequency-domain processing, estimate the data either serialwise, that is, each subcarrier is sequentially equalized [12], or blockwise, that is, all subcarriers are jointly equalized [13]. For SC transmissions over time-invariant frequency-selective channels, time-domain turbo equalization (see [6, 7]) is traditionally more popular than frequency-domain turbo equalization [10]. However, recently, frequency-domain equalization has gained renewed interest, due to its reduced complexity for channels with significant delay spread [5]. For SC systems over doubly selective channels, a low-complexity iterative equalizer has been proposed in [14], which can be regarded as the time-domain counterpart of the iterative frequency-domain equalizer [15]. However, time-domain iterative equalizers are not suitable for channels with significant delay spread, since their complexity is quadratic in the channel length [14]. Besides, maximum likelihood and maximum a posteriori sequence estimators for SC and OFDM systems have been proposed in [16], which models the doubly selective channel using a basis expansion model (BEM).

In this paper, as a first contribution, we apply the block philosophy to design a low-complexity MMSE turbo equalizer for SC systems in doubly selective channels. To the best of our knowledge, all the turbo equalizers proposed so far for SC systems over doubly selective channels employ a serialwise data processing, that is, use a sliding window either in the time domain [14] or in the frequency domain [15, 17]. However, since the presence of the CP makes the transmission scheme blockwise, block equalization becomes a valid alternative. We design our block turbo equalizer for SC systems in the frequency domain, in the same spirit of the turbo equalizers designed for OFDM in [13]. An interesting feature of the proposed block equalizer is its reduced computational complexity, which scales only linearly with the block length. As a result, for doubly selective channels with significant multipath delay spread, our frequency-domain approach is less complex than time-domain equalizers like [14]. To keep the complexity low, some ad-hoc approximations are required, so that the proposed block turbo equalization algorithm for SC turns out to be different from that for OFDM [13]. In this paper, a performance comparison between the proposed algorithm and [13] is also given.

The ICI caused by Doppler spreading also makes the channel estimation problem more difficult. Pilot designs and pilot-assisted channel estimation algorithms have been developed for SC over time-varying flat-fading channels [18], for SC over doubly selective channels [19, 20], and for OFDM over doubly selective channels [21, 22]. All these papers, which employ a BEM for the channel, share the design principle that pilots and data are placed in such a way that they should remain orthogonal after transmission over the fading channel. Indeed, this criterion eliminates the data-to-pilot interference and hence it simplifies the channel estimation task. (The same criterion also eliminates the pilot-to-data interference, whose cancellation is therefore not necessary.) However, in doubly selective channels, orthogonal designs have two drawbacks. First, only approximate-orthogonal designs are really possible, since a doubly selective channel cannot be perfectly diagonalized [23]. Second, a rate loss is introduced by the presence of zero symbols that are necessary to keep the almost orthogonality between data and pilots. On the other hand, nonorthogonal designs are also possible, such as the superimposed training approach developed in [24]. Actually, turbo-inspired iterative channel estimators can handle the data-to-pilot interference by means of reliability-based soft cancellation. In other words, the soft data estimates can be used to improve the quality of channel estimation, as shown by the adaptive iterative channel estimators [25, 26].

As a second contribution of this paper, we present iterative (turbo-like) pilot-assisted channel estimators for both OFDM and SC block transmissions. Differently from the turbo-based channel estimators already proposed for SC transmissions over doubly selective channels [25, 26], the proposed turbo-like channel estimators are nonadaptive and hence more suitable for block transmissions. For both OFDM and SC cases, the proposed iterative channel estimators firstly estimate the time-domain channel exploiting the BEM, and then transform the time-domain channel into the frequency domain for equalization purposes. This strategy is similar to that used for OFDM doubly selective channel estimation in [21, 22]. However, differently from the channel estimators of [21, 22], the proposed channel estimators exploit the reliability of the estimated data and can thus also work in the presence of nonorthogonal pilot designs.

To keep low-complexity channel estimation processing, we assume that the pilot symbols are located in the same domain where the data symbols are placed, that is, we assume frequency-domain pilots for OFDM systems, and time-domain pilots for SC systems. Although this choice is mainly dictated by computational complexity benefits, it is consistent with almost-orthogonal pilot allocation strategies for doubly selective channels, which indeed suggest time-domain pilots for SC systems [19, 20], and frequency-domain pilots for OFDM systems [20].

The rest of this paper is organized as follows. In Section 2, we introduce the system model. Section 3 presents the proposed block turbo MMSE equalizer for SC systems. Section 4 deals with the design of iterative MMSE pilot-assisted channel estimators, for both OFDM and SC systems. In Section 5, we evaluate and compare the performance of the proposed equalizer and of both channel estimators, by means of simulated results. Section 6 concludes the paper.

Notation. We use upper (lower) boldface letters to denote matrices (column vectors). $(\cdot)^T$ and $(\cdot)^H$, and $(\cdot)^\dagger$ represent transpose, complex conjugate transpose (Hermitian), and pseudoinverse, respectively. $[\mathbf{A}]_{m,n}$ indicates the $(m+1, n+1)$th entry of the matrix $\mathbf{A}$. We use the symbol $\circ$ and $\otimes$ to denote the Hadamard (element-wise) product and Kronecker product between matrices, respectively. diag($\mathbf{a}$)

Table 1: QPSK symbol alphabet.

k	1	2	3	4
$(\alpha_{k,1}, \alpha_{k,2})$	(0,0)	(1,0)	(0,1)	(1,1)
α_k	$\frac{(+1+i)}{\sqrt{2}}$	$\frac{(-1+i)}{\sqrt{2}}$	$\frac{(+1-i)}{\sqrt{2}}$	$\frac{(-1-i)}{\sqrt{2}}$

is a diagonal matrix with the vector $\mathbf{a}$ on the diagonal. $\mathrm{E}(\cdot)$ stands for the statistical expectation. The covariance matrix between $\mathbf{x}$ and $\mathbf{y}$ is defined as $\mathrm{Cov}(\mathbf{x}, \mathbf{y}) = \mathrm{E}(\mathbf{x}\mathbf{y}^H) - \mathrm{E}(\mathbf{x})\mathrm{E}(\mathbf{y}^H)$. Finally, $\mathbf{0}_{M\times N}$ and $\mathbf{I}_N$ denote the $M \times N$ all-zero matrix and the $N \times N$ identity matrix, respectively.

2. System Model

We consider a single-user communication system with blockwise transmission, and a channel that is both frequency and time selective. The structure of both transmitter and receiver is shown in Figure 1. At the transmitter, the information bits are encoded with error correction coding, and the coded bits are interleaved and mapped into N_d complex symbols, represented by the $N_d \times 1$ vector $\mathbf{s}^d$, and the data symbols are assumed to be uncorrelated. We define $\mathbf{s}^p$ as the $N_p \times 1$ vector that stands for the pilot symbols, which are multiplexed with $\mathbf{s}^d$ to form a block of $N = N_d + N_p$ transmitted symbols $\mathbf{s}$. For simplicity, we consider unit-energy quaternary phase-shift keying (QPSK) with the symbol alphabet shown in Table 1. However, the equalizers and channel estimators proposed herein can be easily extended to other constellations, like in [8].

As far as the time dispersion of the channel is concerned, we adopt the standard assumption that the maximum channel order is equal to the CP length, both denoted by L, where $L < N$. This way, there is no interference between successive blocks, and the equalizer can be designed separately for each block. As a consequence, we can omit the block index from our notation.

At the receiver, after removing the CP, the $N \times 1$ received vector $\mathbf{y}_t$ can be expressed as

$$\mathbf{y}_t = \mathbf{H}_t\mathbf{P}\mathbf{s} + \mathbf{n}_t, \tag{1}$$

where $\mathbf{H}_t$ is the $N \times N$ time-domain channel matrix, $\mathbf{P}$ denotes the $N \times N$ precoder matrix, $\mathbf{s}$ represents the $N \times 1$ symbol vector consisting of the multiplexed pilot and data symbols, and $\mathbf{n}_t$ stands for the $N \times 1$ noise vector, which is assumed to be uncorrelated with the data symbols. The precoder is set to $\mathbf{P} = \mathbf{I}_N$ for SC systems, and $\mathbf{P} = \mathbf{F}^H$ for OFDM systems, where $\mathbf{F}$ denotes the $N \times N$ unitary DFT matrix. For simplicity, we assume that $\mathbf{n}_t$ is a circularly symmetric complex Gaussian noise vector, with zero mean and covariance matrix $\mathbf{R}_{\mathbf{n}_t} = E(\mathbf{n}_t\mathbf{n}_t^H) = \sigma_n^2\mathbf{I}_N$. At the receiver, a length-N time-domain window can be applied after CP removal and before the DFT operation. In this case, the output vector after the DFT operation can be expressed as

$$\mathbf{y}_f = \mathbf{F}\mathbf{W}\mathbf{H}_t\mathbf{F}^H\mathbf{F}\mathbf{P}\mathbf{s} + \mathbf{F}\mathbf{W}\mathbf{n}_t = \mathbf{H}_f\mathbf{F}\mathbf{P}\mathbf{s} + \mathbf{n}_f, \tag{2}$$

where $\mathbf{y}_f = \mathbf{F}\mathbf{W}\mathbf{y}_t$, $\mathbf{n}_f = \mathbf{F}\mathbf{W}\mathbf{n}_t$, $\mathbf{H}_f = \mathbf{F}\mathbf{W}\mathbf{H}_t\mathbf{F}^H$, and $\mathbf{W} = \mathrm{diag}(\mathbf{w})$, with $\mathbf{w}$ the $N \times 1$ vector denoting the time-domain receiver window. Note that classical systems do not include windowing, that is, $\mathbf{W} = \mathbf{I}_N$.

When the channel is time varying, $\mathbf{H}_t$ is no longer circulant, and the $N \times N$ frequency-domain channel matrix $\mathbf{H}_f$ becomes a nondiagonal matrix, giving rise to ICI, where the ICI coupling is summarized by the nonzero offdiagonal elements of $\mathbf{H}_f$. However, with a proper window design, $\mathbf{H}_f$ is cyclically banded, with the most significant elements around the main diagonal, and on the upper-right and lower-left corners [12]. In this paper, we employ the minimum band approximation error windowing developed in [27], where the window $\mathbf{w}$ is obtained as a sum of complex exponentials. This choice permits the use of low-complexity equalization algorithms specially tailored to banded and cyclically banded matrices, as explained in [12, 28]. Observe that the receiver windowing in [27] only requires some statistical knowledge about the channel time variation, and this knowledge does not even have to be very exact.

To simplify the equalization procedure, the matrix $\mathbf{H}_f$ is further approximated by its cyclically banded version

$$\mathbf{H} = \mathbf{H}_f \circ \mathbf{\Theta}, \tag{3}$$

where $\mathbf{\Theta}$ is the $N \times N$ cyclically banded circulant matrix, which has ones on the main diagonal, on the B_c super- and B_c subdiagonals, and on the upper-right and lower-left B_c-size corners, while the remaining entries are zeros. The *matrix bandwidth* parameter B_c allows for a trade-off between equalization complexity and performance, and it can be chosen according to some rules of thumb [12]. When windowing is included, B_c is usually much smaller than the number of subcarriers N.

It can be observed that the transmitted data block $\mathbf{s}$ represents a time-domain signal in SC systems, while it represents a frequency-domain signal in OFDM systems. This clearly explains why SC systems are more prone to multipath effects, which mix the data due to the associated ISI, while OFDM systems suffer from Doppler effects, which mix the data due to the associated ICI. Our equalizer will be designed in the frequency domain, with the goal of mitigating the interference caused by the offdiagonal elements of $\mathbf{H}$.

3. Low-Complexity Block Turbo Equalization

In order to derive frequency-domain block turbo equalizers for doubly selective channels, let us define s_i as the ith QPSK symbol of $\mathbf{s}$, and $(s_{i,1}, s_{i,2})$ as the related bits. The mean and the variance of the symbol s_i are denoted as $m_i = \mathrm{E}(s_i)$ and $v_i = \mathrm{Cov}(s_i, s_i)$, respectively. Similarly, we have $\mathbf{m} = \mathrm{E}(\mathbf{s}) = [m_1, \ldots, m_N]^T$ and $\mathbf{V} = \mathrm{Cov}(\mathbf{s}, \mathbf{s}) = \mathrm{diag}([v_1, \ldots, v_N]^T)$. As far as the N_d data symbols are concerned, the means and the variances are initialized with zeros and ones, respectively. But in every iteration of the turbo equalizer, they are updated using soft information from the estimated symbols. On the other hand, for each of the N_p pilot symbols, the mean is set

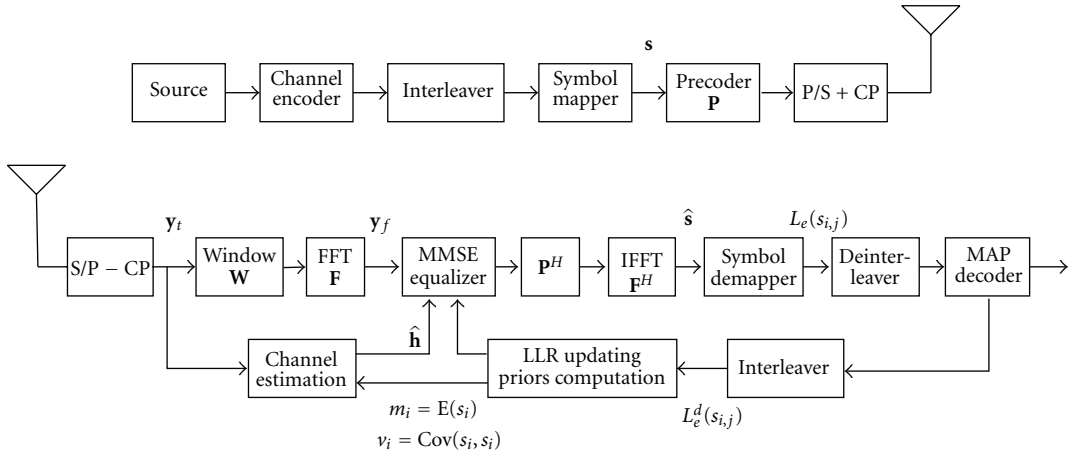

Figure 1: System model.

to the pilot symbol value, while the variance is zero, for all the iterations.

After each iteration of the equalizer, we update the means and the variances using the soft estimated symbols. Specifically, we need to calculate the extrinsic log-likelihood ratio (LLR), $L_e(s_{i,j}) = L(s_{i,j} \mid \hat{s}_i) - L(s_{i,j})$, where $L(s_{i,j})$ is the *a priori* LLR and $L(s_{i,j} \mid \hat{s}_i)$ is the *a posteriori* LLR [7, 13]. To perform this calculation, we should derive the probability density function (PDF) $p(\hat{s}_i \mid s_i = s)$, which can be approximated as Gaussian: $p(\hat{s}_i \mid s_i = s) = (1/\pi\sigma_i^2) \cdot e^{-|\hat{s}_i - \mu_i|^2/\sigma_i^2}$, with mean $\mu_i = \mathrm{E}(\hat{s}_i \mid s_i = s)$ and variance $\sigma_i^2 = \mathrm{Cov}(\hat{s}_i, \hat{s}_i \mid s_i = s)$ [7, 8]. As shown in Figure 1, the extrinsic LLR $L_e(s_{i,j})$ is passed to the decoder to generate a new extrinsic LLR $L_e^d(s_{i,j})$, which is added to the a priori LLR to form the new a posteriori LLR $L_{\mathrm{new}}(s_{i,j})$. The new a posteriori LLR permits to update the means and the variances of the estimated symbols as in [7, 13]:

$$L_{\mathrm{new}}\left(s_{i,j}\right) = L\left(s_{i,j}\right) + L_e^d\left(s_{i,j}\right),$$

$$m_{i,\mathrm{new}} = \frac{\tanh\left(L_{\mathrm{new}}(s_{i,1})/2\right) + \sqrt{-1} \cdot \tanh\left(L_{\mathrm{new}}(s_{i,2})/2\right)}{\sqrt{2}},$$

$$v_{i,\mathrm{new}} = 1 - \left| m_{i,\mathrm{new}} \right|^2. \tag{4}$$

The a posteriori LLR $L_{\mathrm{new}}(s_{i,j})$ of the current iteration then becomes the a priori LLR $L(s_{i,j})$ used in the next iteration. In the first iteration, no prior information is available, and therefore the a priori LLR is zero. The whole procedure described above can then be repeated, depending on the chosen number of iterations.

In the next subsection, we present a block turbo equalizer for SC systems. The proposed equalizer is derived using a similar approach as in [13], which develops three block turbo equalizers for OFDM systems.

3.1. Block Turbo Equalization for SC Systems. In SC systems, the precoder is absent and therefore it is set to $\mathbf{P} = \mathbf{I}_N$. In this case, (2) can be rewritten as

$$\mathbf{y}_f = \mathbf{F}\mathbf{W}\mathbf{H}_t\mathbf{F}^H\mathbf{F}\mathbf{s} + \mathbf{F}\mathbf{W}\mathbf{n}_t \approx \mathbf{H}\mathbf{s}_f + \mathbf{n}_f, \tag{5}$$

where $\mathbf{s}_f = \mathbf{F}\mathbf{s}$. Similarly to our previous notation, we define $s_{f,i}$ as the ith symbol of $\mathbf{s}_f$, $m_{f,i} = \mathrm{E}(s_{f,i})$ and $v_{f,i} = \mathrm{Cov}(s_{f,i}, s_{f,i})$ as the means and the variances of the frequency-domain symbols.

Given $\mathbf{m}$ and $\mathbf{V}$ as prior information, the equalizer exploits the means and the variances of the frequency-domain symbols. Since $\mathbf{s}_f = \mathbf{F}\mathbf{s}$, we have $\mathbf{m}_f = \mathrm{E}(\mathbf{s}_f) = \mathbf{F}\mathbf{m} = [m_{f,1}, \ldots, m_{f,N}]^T$, and $\mathbf{V}_f = \mathrm{Cov}(\mathbf{s}_f, \mathbf{s}_f) = \mathbf{F}\mathbf{V}\mathbf{F}^H$. Since $\mathbf{V}$ is diagonal, $\mathbf{V}_f$ is in general circulant but not diagonal, that is, it can not be written as $\mathrm{diag}([v_{f,1}, \ldots, v_{f,N}]^T)$. However, as it will be explained later, dealing with a diagonal $\mathbf{V}_f$ is crucial for complexity reasons. Therefore, to save complexity, we replace $\mathbf{V}_f$ with its approximated version $\mathbf{V}_f \circ \mathbf{I}_N$ obtained by setting its offdiagonal elements to zero. Since the diagonal elements of $\mathbf{V}_f$ are equal, $\mathbf{V}_f \circ \mathbf{I}_N = \overline{v}_t \cdot \mathbf{I}_N$ is a scaled identity, with $\overline{v}_t = (1/N)\sum_{k=1}^{N} v_k$. A similar approximation is sometimes used also in time-domain equalizers [7, 8].

By using $m_{f,i}$ and $v_{f,i}$ as prior information, a frequency-domain linear MMSE equalizer can be obtained using the approach used in [13] (Equalizer III) for OFDM systems. In SC systems, this approach leads to

$$\hat{s}_{f,i} = \frac{1}{t_i}\mathbf{h}_i^H\mathbf{A}^{-1}\left(\mathbf{y}_f - \mathbf{H}\mathbf{m}_f\right) + m_{f,i}, \tag{6}$$

where $\mathbf{h}_i$ is the ith column of $\mathbf{H}$, $\mathbf{A} = \mathbf{H}\mathbf{V}_f\mathbf{H}^H + \mathbf{R}_{\mathbf{n}_f}$, $t_i = \mathbf{h}_i^H\mathbf{A}^{-1}\mathbf{h}_i$, and $\mathbf{R}_{\mathbf{n}_f} = E(\mathbf{n}_f\mathbf{n}_f^H)$. It should be observed that, when $\mathbf{V}_f$ is approximated as diagonal, the matrix $\mathbf{A}$ is cyclically banded, and hence the computations in (6) can be performed using special algorithms designed for solving cyclically banded linear systems. In this work, we

have used a cyclic band LDL^H factorization obtained by a convenient modification of [28], in the same spirit of the fast Cholesky factorization of [29]. An alternative LU factorization algorithm could be derived using the divide-and-conquer method of [30]. Using the algorithms specifically tailored to cyclically banded matrices, the computational complexity per data block reduces to $\mathcal{O}(B_c^2 N)$, which is linear in the block size N. On the contrary, the complexity of a time-domain equalizer would be $\mathcal{O}(L^2 N)$. Since the Doppler support B_c is usually much lower than the maximum channel order L, our frequency-domain MMSE equalizer is computationally cheaper than the corresponding time-domain MMSE equalizer.

On the other hand, without any approximation on $\mathbf{V}_f$, $\mathbf{A}$ would not be cyclically banded, and therefore the complexity order of MMSE equalizers would be $\mathcal{O}(N^3)$. Since the block size N is by far greater than the Doppler support B_c, the diagonal approximation is necessary for low-complexity MMSE equalizers. However, when moderate complexity is affordable, other approximations are possible. For instance, if $\mathbf{V}_f$ is approximated as cyclically banded with bandwidth B_v, the matrix $\mathbf{A}$ would be cyclically banded too, but the computational complexity would increase to $\mathcal{O}((B_c + B_v)^2 N)$. Alternatively, a low-complexity weighted least-squares (WLS) equalizer that avoids the approximation of $\mathbf{V}_f$ could be employed, by neglecting the noise covariance matrix inside $\mathbf{A}$. However, since doubly selective channels lead to highly ill-conditioned matrices, WLS equalizers produce a very poor performance. Indeed, MMSE equalizers can be interpreted as regularized WLS equalizers.

From (6), the estimated time-domain data-symbol vector is successively obtained by $\hat{\mathbf{s}} = \mathbf{F}^H \hat{\mathbf{s}}_f$, which leads to

$$\begin{aligned}
\hat{s}_i &= \mathbf{i}_i^H \mathbf{F}^H \hat{\mathbf{s}}_f = \mathbf{i}_i^H \mathbf{F}^H \sum_{k=1}^{N} \mathbf{i}_k \hat{s}_{f,k} \\
&= \mathbf{i}_i^H \mathbf{F}^H \sum_{k=1}^{N} \frac{1}{t_k} \mathbf{i}_k \mathbf{h}_k^H \mathbf{A}^{-1} \mathbf{H} \mathbf{F} (\mathbf{s} - \mathbf{m}) + \mathbf{i}_i^H \mathbf{F}^H \sum_{k=1}^{N} \mathbf{i}_k m_{f,k} \\
&\quad + \mathbf{i}_i^H \mathbf{F}^H \sum_{k=1}^{N} \frac{1}{t_k} \mathbf{i}_k \mathbf{h}_k^H \mathbf{A}^{-1} \mathbf{F} \mathbf{n}_t \\
&= \mathbf{i}_i^H \mathbf{F}^H \sum_{k=1}^{N} \frac{1}{t_k} \mathbf{i}_k \mathbf{h}_k^H \mathbf{A}^{-1} \mathbf{H} \mathbf{F} (\mathbf{s} - \mathbf{m}) + \mathbf{i}_i^H \mathbf{F}^H \sum_{k=1}^{N} \mathbf{i}_k \mathbf{i}_k^H \mathbf{F} \mathbf{m} \\
&\quad + \mathbf{i}_i^H \mathbf{F}^H \sum_{k=1}^{N} \frac{1}{t_k} \mathbf{i}_k \mathbf{h}_k^H \mathbf{A}^{-1} \mathbf{F} \mathbf{n}_t \\
&= m_i + \mathbf{i}_i^H \mathbf{F}^H \mathbf{T} \boldsymbol{\Sigma} \mathbf{F} (\mathbf{s} - \mathbf{m}) + \mathbf{i}_i^H \mathbf{F}^H \mathbf{T} \mathbf{H}^H \mathbf{A}^{-1} \mathbf{F} \mathbf{n}_t,
\end{aligned} \tag{7}$$

where $\mathbf{i}_k$ is the $N \times 1$ indicator function, defined as the kth column of $\mathbf{I}_N$, $\mathbf{T} = \sum_{k=1}^{N} (1/t_k) \mathbf{i}_k \mathbf{i}_k^H$ is a diagonal matrix, and $\boldsymbol{\Sigma} = \mathbf{H}^H \mathbf{A}^{-1} \mathbf{H}$.

In order to derive the extrinsic LLR, the mean μ_i and the variance σ_i^2 of the Gaussian PDF $p(\hat{s}_i \mid s_i = s)$ are calculated from (7) as

$$\begin{aligned}
\mu_i &= \mathrm{E}(\hat{s}_i \mid s_i = s) \\
&= \mathbf{i}_i^H \mathbf{F}^H \mathbf{T} \boldsymbol{\Sigma} \mathbf{F} \mathrm{E}(\mathbf{s} - (\mathbf{m} - m_i \mathbf{i}_i) \mid s_i = s) \\
&= \mathbf{i}_i^H \mathbf{F}^H \mathbf{T} \boldsymbol{\Sigma} \mathbf{F} \mathbf{i}_i s \\
&\approx s, \\
\sigma_i^2 &= \mathrm{Cov}(\hat{s}_i, \hat{s}_i \mid s_i = s) \\
&= \mathrm{E}\left(\left| \hat{s}_i - \mu_i \right|^2 \mid s_i = s \right) \\
&= \mathrm{E}\Big(\Big| \mathbf{i}_i^H \mathbf{F}^H \mathbf{T} \boldsymbol{\Sigma} \mathbf{F} [\mathbf{s} - \mathbf{m} - \mathbf{i}_i (s - m_i)] \\
&\qquad + \mathbf{i}_i^H \mathbf{F}^H \mathbf{T} \mathbf{H}^H \mathbf{A}^{-1} \mathbf{F} \mathbf{n}_t \Big|^2 \Big) \\
&= \mathbf{i}_i^H \mathbf{F}^H \mathbf{T} \boldsymbol{\Sigma} \mathbf{F} \left(\mathbf{V} - v_i \mathbf{i}_i \mathbf{i}_i^H \right) \mathbf{F}^H \boldsymbol{\Sigma}^H \mathbf{T}^H \mathbf{F} \mathbf{i}_i \\
&\quad + \sigma_n^2 \mathbf{i}_i^H \mathbf{F}^H \mathbf{T} \boldsymbol{\Sigma} \mathbf{H}^{-1} \mathbf{H}^{-H} \boldsymbol{\Sigma}^H \mathbf{T}^H \mathbf{F} \mathbf{i}_i \\
&\approx \overline{v}_i + e_n,
\end{aligned} \tag{8}$$

where $\overline{v}_i = (1/N)(\sum_{k=1}^{N} v_k - v_i) = \overline{v}_t - v_i/N$ and $e_n = (\sigma_n^2/N) \sum_{k=1}^{N} |\mathbf{h}_k^H \mathbf{h}_k|^{-1}$. In (8), we have approximated the matrices $\mathbf{F}(\mathbf{V} - v_i \mathbf{i}_i \mathbf{i}_i^H) \mathbf{F}^H$, $\mathbf{H}^H \mathbf{H}$, and $\boldsymbol{\Sigma}$ by diagonal matrices, by setting their offdiagonal elements to zero. As it will be explained later, similarly to the diagonal approximation of $\mathbf{V}_f$, these approximations are necessary to maintain a low complexity. We now separately discuss the three approximations. *First*, the approximation of $\mathbf{F}(\mathbf{V} - v_i \mathbf{i}_i \mathbf{i}_i^H) \mathbf{F}^H$ is similar to the approximation of $\mathbf{V}_f$, which has been discussed previously. However, now the obtained diagonal matrix is not a scaled identity. *Second*, since $\mathbf{H}$ is cyclically banded, the offdiagonal elements of $\mathbf{H}^H \mathbf{H}$ decay to zero very rapidly. Hence, we expect that the approximation on $\mathbf{H}^H \mathbf{H}$ will not introduce a significant error. *Third*, the matrix $\boldsymbol{\Sigma} = \mathbf{H}^H \mathbf{A}^{-1} \mathbf{H}$ represents the effect of a linear MMSE equalizer $\mathbf{H}^H \mathbf{A}^{-1}$ applied to the channel matrix $\mathbf{H}$. Since the MMSE equalizer highly mitigates the ICI, $\boldsymbol{\Sigma}$ is already very close to a diagonal matrix. This last approximation also leads to $\mathbf{T}\boldsymbol{\Sigma} \approx \mathbf{I}_N$, which justifies the equalizer unbiasedness $\mu_i \approx s_i$.

We now use the results in (8) to express the extrinsic LLR as [7, 13]

$$\begin{aligned}
L_e(s_{i,1}) &= \ln \frac{p(\hat{s}_i \mid s_i = \alpha_1) P_2(0) + p(\hat{s}_i \mid s_i = \alpha_3) P_2(1)}{p(\hat{s}_i \mid s_i = \alpha_2) P_2(0) + p(\hat{s}_i \mid s_i = \alpha_4) P_2(1)} \\
&= \frac{\sqrt{8}\, \mathrm{Re}(\hat{s}_i)}{\overline{v}_i + e_n}, \\
L_e(s_{i,2}) &= \ln \frac{p(\hat{s}_i \mid s_i = \alpha_1) P_1(0) + p(\hat{s}_i \mid s_i = \alpha_2) P_1(1)}{p(\hat{s}_i \mid s_i = \alpha_3) P_1(0) + p(\hat{s}_i \mid s_i = \alpha_4) P_1(1)} \\
&= \frac{\sqrt{8}\, \mathrm{Im}(\hat{s}_i)}{\overline{v}_i + e_n},
\end{aligned} \tag{9}$$

where $P_j(0) = P(s_{i,j} = 0)$ and $P_j(1) = P(s_{i,j} = 1)$.

It is easy to prove that the calculation of the extrinsic LLR in (9) has complexity $\mathcal{O}(B_c N)$. Therefore, the equalization complexity of (6) dominates over the extrinsic LLR calculation complexity of (9). Taking into account FFT operations, the overall computational complexity per iteration for each block of N symbols is $\mathcal{O}((B_c^2 + \log(N))N)$, which is independent of the channel length L. On the other hand, the complexity of the time-domain equalizer of [14] is $\mathcal{O}(L^2 N)$. Therefore, for multipath channels with a long impulse response, we obtain a significant complexity saving. A more detailed discussion (i.e., flops count) about the computational complexity of banded turbo equalizers can be found in [13].

We highlight that the three diagonal approximations introduced in (8) are fundamental in reducing the computational complexity. For instance, if the full matrix $\mathbf{\Sigma}$ is used, the computation of σ_i^2 in (8) involves full matrices, and therefore the computational complexity would be at least $\mathcal{O}(N^2)$. In this case, the complexity of the extrinsic LLR calculation (9) would dominate. Clearly, the nonapproximated equalizer would be useful only when the block size N is small, which is not feasible in long multipath channels due to the constraint $N > L$. Therefore, if low computational complexity is important, there is no way to avoid diagonal approximations. We also point out that, among the different possible ways to approximate the two matrices $\mathbf{H}^H\mathbf{H}$ and $\mathbf{\Sigma}$ as diagonal, the only reasonable approach is setting their offdiagonal elements to zero. Indeed, as explained after (8), $\mathbf{H}^H\mathbf{H}$ and $\mathbf{\Sigma}$ are almost diagonal. However, for $\mathbf{F}(\mathbf{V} - v_i \mathbf{i}_i \mathbf{i}_i^H)\mathbf{F}^H$, there exist different ways to approximate it as diagonal. Neglecting the offdiagonal elements leads to $\overline{v}_i = (1/N)(\sum_{k=1}^{N} v_k - v_i)$, which could be replaced, for instance, by $\overline{v}_i = \overline{v} = \max\{v_k\}_{k=1}^{N}$. Intuitively, the average approximation assigns to all the symbols the average reliability of all the symbols, whereas the maximum approximation assigns to all the symbols the reliability of the worst symbol estimate. In the simulation section, we compare the performance of both approximations.

4. Iterative Channel Estimation

The turbo equalizers presented in the previous section require the channel-state information (CSI) at the receiver. To acquire the CSI, we propose a modification of a pilot-assisted channel estimator presented in [21] for OFDM. Specifically, we modify the iterative linear MMSE channel estimator of [21] in such a way that it can operate in a turbo fashion. Therefore, besides the pilot symbols, we also use the soft data estimates originating from the turbo equalizer and the decoder. Indeed, after the first iteration, the soft data symbol estimates can be used as auxiliary pilot symbols, in order to improve the quality of the subsequent channel estimates [31]. For both OFDM and SC systems, our channel estimators produce an estimate $\hat{\mathbf{H}}_t$ of the time-domain channel matrix $\mathbf{H}_t$, and then translate $\hat{\mathbf{H}}_t$ into the frequency-domain cyclically banded matrix estimate $\hat{\mathbf{H}}$. The channel estimators are assumed to have perfect knowledge of the channel statistics, that is, the Doppler spectrum and the power-delay profile. We highlight that the channel estimators considered in this paper are nonadaptive, that is, the CSI is newly estimated in each transmitted block, using both pilots and data. This way, severe time variation can be handled.

In pilot-assisted transmissions, there exist various approaches to design the pilot pattern. We can distinguish between two broad categories: multiplexed training and superimposed training [24]. In the multiplexed training case, each element of the transmitted vector contains either a pilot symbol or a data symbol, while in the superimposed case both pilot and data symbols are located in the same positions, typically distributed over the whole transmitted vector. In this paper, we assume multiplexed training, which is also known as periodic training when the pilots are placed in the time domain, and as orthogonal training when the pilots are located in the frequency domain. In particular, we focus on the pilot placement schemes developed in [19, 20], which have been proved to be optimal in the MMSE sense under certain channel conditions. In these schemes, pilot symbols are interleaved with the data symbols to form the transmitted signal vector. For OFDM systems, we employ the frequency-domain Kronecker delta (FDKD) pilot structure [20], while, its dual scheme [19], identified as time-domain Kronecker delta (TDKD), is adopted for SC systems. In both cases, the pilot symbols are grouped into equidistant clusters, each having the same length. Within each cluster, a unique nonzero pilot symbol is located in the middle of the cluster, while null pilot symbols are placed on both sides. Therefore, the FDKD scheme coincides with equispaced pilot tones with guard frequency bands, while the TDKD scheme uses periodic training with guard time intervals.

Suppose that there are M pilot clusters, each containing L_p (odd) pilots, denoted by the vector $\mathbf{s}_m^p$ of size L_p, for $m = 0, \ldots, M-1$. These vectors are interleaved with the data clusters $\mathbf{s}_m^d$, $m = 0, \ldots, M$, to form the transmitted vector $\mathbf{s}$, as expressed by

$$\mathbf{s} = \left[\left(\mathbf{s}_0^d\right)^T, \left(\mathbf{s}_0^p\right)^T, \ldots, \left(\mathbf{s}_{M-1}^d\right)^T, \left(\mathbf{s}_{M-1}^p\right)^T, \left(\mathbf{s}_M^d\right)^T \right]^T. \quad (10)$$

Therefore, by denoting with n_m the starting position of the mth pilot cluster, the mth pilot vector contains the elements of $\mathbf{s}$ with index starting from n_m up to $n_m + L_p - 1$. We also define the aggregate pilot vector and the aggregate data vector as

$$\begin{aligned} \mathbf{s}^p &= \left[\left(\mathbf{s}_0^p\right)^T, \ldots, \left(\mathbf{s}_{M-1}^p\right)^T \right]^T, \\ \mathbf{s}^d &= \left[\left(\mathbf{s}_0^d\right)^T, \ldots, \left(\mathbf{s}_M^d\right)^T \right]^T, \end{aligned} \quad (11)$$

respectively, with size ML_p and $N - ML_p$, respectively.

In addition, we define $h_{n,l}^t$ as the lth channel tap at the nth time instant, where $h_{n,l}^t = 0$ for $l < 0$ or $l > L$, since the maximal channel order is assumed to be L. Thus the elements of $\mathbf{H}_t$ can be expressed as

$$[\mathbf{H}_t]_{p,q} = h_{L+1+p,\,\mathrm{mod}\,(p-q,N)}^t, \quad (12)$$

which means that our channel estimation problem has $N(L+1)$ unknowns. However, these unknowns are correlated in the

time domain. The BEM can be used to reduce the number of unknowns from $N(L+1)$ to $(Q+1)(L+1)$, where $Q+1$ is the number of basis functions [21]. By stacking all the channel taps within the block in a single $N(L+1)\times 1$ vector $\mathbf{h}_t = [h^t_{L+1,0},\ldots,h^t_{L+1,L},\ldots,h^t_{L+N,0},\ldots,h^t_{L+N,L}]^T$, the BEM permits to express this vector as

$$\mathbf{h}_t = (\mathbf{B}\otimes\mathbf{I}_{L+1})\mathbf{h}, \tag{13}$$

where $\mathbf{B} = [\mathbf{b}_0,\ldots,\mathbf{b}_Q]$ is an $N\times(Q+1)$ matrix that has $Q+1$ orthonormal basis functions $\mathbf{b}_q$ as columns, and $\mathbf{h}$ is a $(Q+1)(L+1)$ vector that collects all the BEM coefficients of all the channel taps.

In order to derive our MMSE channel estimator, the following assumptions are made.

Assumption 1. The wireless channel can be regarded as a wide-sense stationary uncorrelated scattering (WSSUS) process, which has the following statistics

$$\mathrm{E}\left(h^t_{t,l}\left(h^t_{t-t',l-l'}\right)^*\right) = \sigma_l^2\delta_{l'}\gamma_{t'}, \tag{14}$$

where σ_l^2 denotes the variance of the lth channel tap, γ_t is the normalized time correlation, and δ_n stands for the Kronecker delta function.

Assumption 2. The data symbols in $\mathbf{s}^d$ are assumed to be uncorrelated with zero mean and variance σ_s^2, while the noise at the receiver is assumed to be uncorrelated with the transmitted symbols, as expressed by

$$\begin{aligned}\mathrm{E}\left(\mathbf{s}^d\left(\mathbf{s}^d\right)^H\right) &= \sigma_s^2\mathbf{I}_{N-ML_p},\\ \mathrm{E}\left(\mathbf{s}\mathbf{n}_t^H\right) &= \mathbf{0},\end{aligned} \tag{15}$$

respectively.

Assumption 3. The BEM coefficients $\mathbf{h}$ are assumed to be uncorrelated with the transmitted signal $\mathbf{s}$ and the noise, respectively, as expressed by

$$\begin{aligned}\mathrm{E}\left(\mathbf{h}\mathbf{s}^H\right) &= \mathbf{0},\\ \mathrm{E}\left(\mathbf{h}\mathbf{n}_t^H\right) &= \mathbf{0}.\end{aligned} \tag{16}$$

Assumption 4. The covariance matrix of the BEM coefficients is assumed known to the receiver, and it is calculated as [21]

$$\mathbf{R}_{\mathbf{h}} = \mathrm{E}\left(\mathbf{h}\mathbf{h}^H\right) = \mathbf{R}_{\mathbf{h},l}\otimes\mathrm{diag}(\sigma_0^2,\ldots,\sigma_L^2), \tag{17}$$

where $\mathbf{R}_{\mathbf{h},l} = \mathrm{E}(\mathbf{h}_l\mathbf{h}_l^H) = \mathbf{B}^\dagger\mathbf{R}^{\{t\}}_{\mathbf{h},l}\mathbf{B}^{\dagger H}$, with $\mathbf{h}_l = [h_{0,l},\ldots,h_{Q,l}]^T$ and $[\mathbf{R}^{\{t\}}_{\mathbf{h},l}]_{p,q} = \mathrm{E}(h^t_{p,l}(h^t_{q,l})^*)$.

Assumption 5. The average power of the pilot symbols is the same as that of the data symbols, as expressed by

$$\sigma_s^2 = \frac{(\mathbf{s}^p)^H\mathbf{s}^p}{ML_p}. \tag{18}$$

4.1. Iterative Channel Estimation for OFDM Systems. For OFDM systems, the pilot and data symbols are interleaved in the frequency domain. Since the frequency-domain channel matrix is cyclically banded only approximately, the received samples used for channel estimation are always contaminated by ICI, independently of the length of the null guard bands inserted. To be precise, the frequency-domain channel matrix is (with high probability) a full matrix, and hence the power of the pilot symbols is spread over all the received samples. While a time-domain receiver window can reduce the ICI to get a better equalization performance, it is still unclear whether the same window can improve the channel estimation quality or not. Thus, to estimate the time-domain channel matrix $\mathbf{H}_t$, we use the frequency-domain received signal *without* applying the time-domain receiver window.

For OFDM systems, the precoder is set to $\mathbf{P} = \mathbf{F}^H$, and (2) can be rewritten as

$$\mathbf{y}_f = \mathbf{F}\mathbf{W}\mathbf{H}_t\mathbf{F}^H\mathbf{s} + \mathbf{F}\mathbf{W}\mathbf{n}_t \approx \mathbf{H}\mathbf{s} + \mathbf{n}_f. \tag{19}$$

By substituting (13) in (19), we can rewrite (19) as

$$\mathbf{y}_f = \sum_{q=0}^{Q}\mathbf{D}_q\mathbf{\Delta}_q\mathbf{s} + \mathbf{n}_f, \tag{20}$$

where $\mathbf{D}_q = \mathbf{F}\,\mathrm{diag}\{\mathbf{b}_q\}\mathbf{F}^H$, $\mathbf{\Delta}_q = \mathrm{diag}\{\mathbf{F}_L[h_{q,0},\ldots,h_{q,L}]\}$, and $\mathbf{F}_L$ represents the first $L+1$ columns of the matrix $\sqrt{N}\mathbf{F}$.

It is noteworthy that, for channel estimation purposes, it is not necessary to process all the received samples. Indeed, the computational complexity of the channel estimator can be highly reduced by extracting a subvector of $\mathbf{y}_f$ before channel estimation [21]. Obviously, this subvector should contain the relevant information given by the pilot symbols. Therefore, with reference to the mth pilot cluster $\mathbf{s}^p_m$, we consider the following observation subvector

$$\mathbf{y}_{f,m} = \left[\left[\mathbf{y}_f\right]_{n_m-\Delta},\ldots,\left[\mathbf{y}_f\right]_{n_m+L_p-1+\Delta}\right]^T, \tag{21}$$

where Δ is a smoothing parameter used to control the amount of interference taken into account for channel estimation. Please observe that Δ can be positive as well as negative, or zero: When Δ is positive, the channel estimator is actually enlarging the observation window, which in this case monitors also the 2Δ data symbol locations closest to the pilot symbols. The received signal in (21) can also be expressed as

$$\mathbf{y}_{f,m} = \sum_{q=0}^{Q}\mathbf{D}_{q,m}\mathbf{\Delta}_q\mathbf{m} + \sum_{q=0}^{Q}\mathbf{D}_{q,m}\mathbf{\Delta}_q(\mathbf{s}-\mathbf{m}) + \mathbf{n}_{f,m}, \tag{22}$$

where $\mathbf{D}_{q,m}$ is an $(L_p+2\Delta)\times N$ matrix consisting of the $L_p+2\Delta$ rows of $\mathbf{D}_q$ with indices from $n_m-\Delta$ to $n_m+L_p-1+\Delta$. It can be observed that the pilot symbols, as well as the soft data estimates $\mathbf{m}$, are used to estimate the CSI, which could help to achieve a better performance than [21], which uses the pilot symbols only. The second term in (22) reflects the uncertainty of the soft data estimates and can be regarded as interference, whose covariance can be taken into account into the channel estimator.

After some tedious manipulations, we can rewrite (22) as a function of $\mathbf{h}$ as

$$\begin{aligned}\mathbf{y}_{f,m} &= \mathbf{D}_m\{\mathbf{I}_{Q+1}\otimes[\mathrm{diag}(\mathbf{m})\mathbf{F}_L]\}\mathbf{h}\\ &\quad+\mathbf{D}_m\{\mathbf{I}_{Q+1}\otimes[\mathrm{diag}(\mathbf{s}-\mathbf{m})\mathbf{F}_L]\}\mathbf{h}+\mathbf{n}_{f,m}\\ &=\mathbf{P}_m\mathbf{h}+\mathbf{d}_{f,m}+\mathbf{n}_{f,m},\end{aligned}\tag{23}$$

where $\mathbf{D}_m=[\mathbf{D}_{0,m},\ldots,\mathbf{D}_{Q,m}]$, and $\mathbf{P}_m=\mathbf{D}_m\{\mathbf{I}_{Q+1}\otimes[\mathrm{diag}(\mathbf{m})\mathbf{F}_L]\}$. Stacking the M observation clusters together, we obtain the reduced-size frequency-domain received signal expressed by

$$\begin{aligned}\mathbf{y}_f^o &= \mathbf{D}^o\{\mathbf{I}_{Q+1}\otimes[\mathrm{diag}(\mathbf{m})\mathbf{F}_L]\}\mathbf{h}\\ &\quad+\mathbf{D}^o\{\mathbf{I}_{Q+1}\otimes[\mathrm{diag}(\mathbf{s}-\mathbf{m})\mathbf{F}_L]\}\mathbf{h}+\mathbf{n}_f^o\\ &=\mathbf{P}^o\mathbf{h}+\mathbf{d}_f^o+\mathbf{n}_f^o,\end{aligned}\tag{24}$$

where $\mathbf{y}_f^o=[\mathbf{y}_{f,0}^T,\ldots,\mathbf{y}_{f,M-1}^T]^T$, $\mathbf{d}_f^o=[\mathbf{d}_{f,0}^T,\ldots,\mathbf{d}_{f,M-1}^T]^T$, and $\mathbf{n}_f^o=[\mathbf{n}_{f,0}^T,\ldots,\mathbf{n}_{f,M-1}^T]^T$ are column vectors of size $M(L_p+2\Delta)$, $\mathbf{D}^o=[\mathbf{D}_0^T,\ldots,\mathbf{D}_{M-1}^T]^T$, and $\mathbf{P}^o=\mathbf{D}^o\{\mathbf{I}_{Q+1}\otimes[\mathrm{diag}(\mathbf{m})\mathbf{F}_L]\}$.

As a result of (24), the linear MMSE estimation of the BEM channel coefficients can be expressed by

$$\hat{\mathbf{h}}=\mathbf{R}_\mathbf{h}\mathbf{P}^{oH}\left(\mathbf{P}^o\mathbf{R}_\mathbf{h}\mathbf{P}^{oH}+\mathbf{R}_{\mathbf{d}_f^o}+\mathbf{R}_{\mathbf{n}_f^o}\right)^{-1}\mathbf{y}_f^o,\tag{25}$$

where $\mathbf{R}_{\mathbf{d}_f^o}=\mathrm{E}(d_f^o d_f^{oH})$, and $\mathbf{R}_{\mathbf{n}_f^o}=\mathrm{E}(\mathbf{n}_f^o\mathbf{n}_f^{oH})$. We can express $\mathbf{R}_{\mathbf{d}_f^o}$ as $\mathbf{R}_{\mathbf{d}_f^o}=\mathbf{D}^o\mathbf{R}_x\mathbf{D}^{oH}$, where

$$\begin{aligned}\mathbf{R}_x &= \{\mathbf{I}_{Q+1}\otimes[\mathrm{diag}(\mathbf{s}-\mathbf{m})\mathbf{F}_L]\}\mathbf{R}_\mathbf{h}\left\{\mathbf{I}_{Q+1}\otimes[\mathrm{diag}(\mathbf{s}-\mathbf{m})\mathbf{F}_L]^H\right\}\\ &=[\mathbf{I}_{Q+1}\otimes\mathrm{diag}(\mathbf{s}-\mathbf{m})]\left\{(\mathbf{I}_{Q+1}\otimes\mathbf{F}_L)\mathbf{R}_\mathbf{h}(\mathbf{I}_{Q+1}\otimes\mathbf{F}_L)^H\right\}\\ &\quad\times\left[\mathbf{I}_{Q+1}\otimes\mathrm{diag}(\mathbf{s}-\mathbf{m})^H\right]\\ &=[\mathbf{I}_{Q+1}\otimes\mathrm{diag}(\mathbf{s}-\mathbf{m})]\mathbf{X}\left[\mathbf{I}_{Q+1}\otimes\mathrm{diag}(\mathbf{s}-\mathbf{m})^H\right],\end{aligned}\tag{26}$$

where $\mathbf{X}=(\mathbf{I}_{Q+1}\otimes\mathbf{F}_L)\mathbf{R}_\mathbf{h}(\mathbf{I}_{Q+1}\otimes\mathbf{F}_L)^H$, which can be further simplified using the assumptions in (15)–(17) as

$$[\mathbf{R}_x]_{m,n}\begin{cases}\nu_{\mathrm{mod}\,(m,N)}[\mathbf{X}]_{m,n}, & \text{if } \mathrm{mod}\,(m-n,N)=0,\\ 0, & \text{otherwise.}\end{cases}\tag{27}$$

After the estimation of the BEM coefficients in $\hat{\mathbf{h}}$, the time-domain channel vector can be reconstructed by (13) as $\hat{\mathbf{h}}_t=(\mathbf{B}\otimes\mathbf{I}_{L+1})\hat{\mathbf{h}}$, whose elements form the estimated time-domain channel matrix $\hat{\mathbf{H}}_t$.

In [21], it has been shown that the BEM-based linear MMSE channel estimator can achieve a better performance by using a larger number of observation samples, that is, when all elements of $\mathbf{y}_f$ are included in the observation vector $\mathbf{y}_f^o$. Obviously, the same behavior is expected in our case: Indeed, our channel estimator additionally includes the reliability of the turbo-equalized data symbols, and hence additional benefit should be obtained by including more data locations into the observation window. However, the main complexity of our channel estimator comes from the matrix inverse in (25), which requires the observation vector length to be small. Thus, the smoothing parameter Δ allows for a trade-off between channel estimation complexity and performance.

4.2. Iterative Channel Estimation for SC Systems. Unlike OFDM systems, where the pilot symbols are inserted in the frequency domain, and the frequency-domain channel matrix is cyclically banded only approximately, in SC systems the pilots are positioned in the time domain, and the time-domain channel matrix is banded, due to the FIR channel assumption. Therefore, using sufficiently long guard intervals, the ISI between pilots and data is completely eliminated [19], thereby simplifying the channel estimation procedure.

Due to the banded structure of the time-domain channel matrix, when the guard intervals are large enough, the mth pilot cluster only contributes to the time-domain received signal $[[\mathbf{y}_t]_{n_m},\ldots,[\mathbf{y}_t]_{n_m+L_p-1+L}]^T$. Similarly to the OFDM case, we consider the mth observation vector

$$\mathbf{y}_{t,m}=\left[[\mathbf{y}_t]_{n_m-\Delta},\ldots,[\mathbf{y}_t]_{n_m+L_p-1+L+\Delta}\right]^T,\tag{28}$$

where the smoothing parameter Δ is defined as in (21), but now operates in the time domain. Using the expressions (12) and (13), we can rewrite the time-domain received signal (1) as [21]

$$\begin{aligned}\mathbf{y}_{t,m} &= \mathbf{H}_{t,m}\mathbf{m}+\mathbf{H}_{t,\mathrm{m}}(\mathbf{s}-\mathbf{m})+\mathbf{n}_{t,m}\\ &=\sum_{q=0}^{Q}\mathrm{diag}\left\{\mathbf{b}_{q,m}\right\}\mathbf{C}_{q,m}\mathbf{m}\\ &\quad+\sum_{q=0}^{Q}\mathrm{diag}\left\{\mathbf{b}_{q,m}\right\}\mathbf{C}_{q,m}(\mathbf{s}-\mathbf{m})+\mathbf{n}_{t,m}\\ &=\mathbf{P}_m\mathbf{h}+\mathbf{d}_{t,m}+\mathbf{n}_{t,m},\end{aligned}\tag{29}$$

where $\mathbf{H}_{t,m}$ consists of the corresponding rows of $\mathbf{H}_t$, with indices from $n_m-\Delta$ to $n_m+L_p-1+L+\Delta$, $\mathbf{b}_{q,m}=[[\mathbf{b}_q]_{n_m-\Delta},\ldots,[\mathbf{b}_q]_{n_m+L_p-1+L+\Delta}]^T$, $\mathbf{C}_{q,m}$ is an $(L_p+L+2\Delta)\times N$ Toeplitz matrix defined as

$$\mathbf{C}_{q,m}=\begin{pmatrix}\mathbf{0}_{1\times(n_m-L-1-\Delta)} & h_{q,L} & \cdots & h_{q,0} & & \mathbf{0}_{1\times(N-n_m+\Delta)}\\ \vdots & & \ddots & & \ddots & \vdots\\ \mathbf{0}_{1\times(n_m+L_p+\Delta-2)} & & & h_{q,L} & \cdots & h_{q,0}\ \mathbf{0}_{1\times[N-(n_m+L_p-1+L+\Delta)]}\end{pmatrix}, \tag{30}$$

and $\mathbf{P}_m = \mathbf{Z}_m\mathbf{U}_m$, where

$$\begin{aligned}\mathbf{Z}_m &= [\mathrm{diag}(\mathbf{b}_{0,m}),\ldots,\mathrm{diag}(\mathbf{b}_{Q,m})],\\ \mathbf{U}_m &= \mathbf{I}_{Q+1}\otimes\mathbf{T}_m,\\ \mathbf{T}_m &= \begin{pmatrix} m_{n_m-\Delta} & \cdots & m_{n_m-L-\Delta}\\ \vdots & \ddots & \vdots\\ m_{n_m+L_p-1+L+\Delta} & \cdots & m_{n_m+L_p-1+\Delta}\end{pmatrix}.\end{aligned} \tag{31}$$

The covariance matrix of the interference term $\mathbf{R}_{\mathbf{d}_{t,m}}$ can be calculated as

$$\begin{aligned}\mathbf{R}_{\mathbf{d}_{t,m}} &= \mathrm{E}\left(\mathbf{d}_{t,m}\mathbf{d}_{t,m}^H\right)\\ &= \mathrm{E}\left[\mathbf{H}_{t,m}(\mathbf{s}-\mathbf{m})(\mathbf{s}-\mathbf{m})^H\mathbf{H}_{t,m}^H\right]\\ &= \mathrm{diag}\left(\sum_{i=0}^{L}\nu_{n_m-\Delta+i}^2\sigma_{L-i}^2,\ldots,\sum_{i=0}^{L}\nu_{n_m+\Delta+L+L_p-1+i}^2\sigma_{L-i}^2\right).\end{aligned} \tag{32}$$

Stacking the M observation clusters together, we get the reduced-size time-domain received signal

$$\mathbf{y}_t^o = \mathbf{P}^o\mathbf{h} + \mathbf{d}_t^o + \mathbf{n}_t^o, \tag{33}$$

where $\mathbf{y}_t^o = [\mathbf{y}_{t,0}^T,\ldots,\mathbf{y}_{t,M-1}^T]^T$, $\mathbf{P}^o = \mathbf{ZU}$, with $\mathbf{Z} = \mathrm{diag}[\mathbf{Z}_0,\ldots,\mathbf{Z}_{M-1}]$ and $\mathbf{U} = [\mathbf{U}_0^T,\ldots,\mathbf{U}_{M-1}^T]^T$, and $\mathbf{d}_t^o$ and $\mathbf{n}_t^o$ are defined similarly to $\mathbf{y}_t^o$.

Like in the OFDM case, the linear MMSE estimate of the BEM channel coefficients can be derived as

$$\widehat{\mathbf{h}} = \mathbf{R}_\mathbf{h}\mathbf{P}^{oH}\left(\mathbf{P}^o\mathbf{R}_\mathbf{h}\mathbf{P}^{oH} + \mathbf{R}_{\mathbf{d}_t^o} + \mathbf{R}_{\mathbf{n}_t^o}\right)^{-1}\mathbf{y}_t^o, \tag{34}$$

where $\mathbf{R}_{\mathbf{d}_t^o}$ and $\mathbf{R}_{\mathbf{n}_t^o}$ are similarly defined as in (25), and $\mathbf{R}_{\mathbf{d}_t^o} = \mathrm{diag}(\mathbf{R}_{\mathbf{d}_{t,0}},\ldots,\mathbf{R}_{\mathbf{d}_{t,M-1}})$. It is easy to understand that, also in this case, a better performance is achieved by including a larger number of observation samples [21], that is, by increasing the smoothing parameter Δ, at the price of increased complexity.

5. Simulation Results

In this section, the proposed algorithms are examined and compared by simulations. We consider a block transmission system with block length $N = 256$. A rate 1/2 convolutional code with generator polynomials (5, 7) (in octal notation) and codeword length of 16384 is used. We employ random interleaving. The maximum channel delay spread and the CP length are equal to $L = 7$. The channel is assumed to be Rayleigh distributed with uniform power-delay profile $\mathrm{E}\{|h_{n,l}|^2\} = 1/(L+1)$, $l = 0,\ldots,L$, and with Jakes' Doppler spectrum [32, 33]. We consider a high-mobility case where the normalized Doppler frequency is $f_dT = 0.15/N$ with f_d the absolute Doppler frequency shift and T the symbol period. It can be interpreted as $f_d/\xi = 0.15$ with ξ the subcarrier spacing in OFDM systems. The time-domain receiver window of [27], as well as the cyclically banded equalizers, are designed for a matrix bandwidth parameter $B_c = 3$, unless otherwise stated. We use the generalized complex-exponential (GCE) BEM to model the time-varying channel at the receiver [21]. Note that the (critically sampled) complex-exponential (CE) BEM would produce a cyclically banded channel matrix estimate, where the number of BEM parameters $Q+1$ coincides with the number of estimated diagonals. On the contrary, the GCE-BEM produces a full channel matrix estimate: Hence, this choice permits to increase the equalizer bandwidth B_c, so that the number of equalizer diagonals $2B_c+1$ can exceed the number of BEM parameters $Q+1$. The channel decoder employs a linear approximation to the log-MAP decoding algorithm.

Figure 2 shows the bit error rate (BER) performance of (the iterative block) Equalizer III [13], equipped with the proposed channel estimator for OFDM systems, as a function of the signal-to-noise ratio (SNR), which is defined as $\mathrm{SNR} = 1/\sigma_n^2$. We insert $ML_p = 10$ frequency-domain pilot symbols, grouped into $M = 10$ clusters, which means $L_p = 1$, that is, there are no guard bands around the nonzero pilot in each cluster. Therefore, the efficiency is $\eta = (N - ML_p)/(N+L) = 0.94$. We use the GCE-BEM with $Q = 4$. The observation length parameter is set to $\Delta = 2$, which leads to a total observation length of $M(L_p + 2\Delta) = 50$ for each block. It is clear that most of the performance gain is obtained when passing from one iteration, which represents the noniterative equalizer, to two iterations. In addition, it is relevant that the performance gain obtained by iterative equalization with respect to the noniterative equalizer is higher in case of estimated CSI: For instance, at $\mathrm{BER} = 10^{-3}$, the performance gain is more than 3 dB. The performance gain with respect to noniterative approaches is confirmed by Figure 3, which displays the normalized mean square error (NMSE) of the iterative MMSE channel estimator, defined as $\mathrm{NMSE} = E\{\|\mathbf{h}_t - (\mathbf{B}\otimes\mathbf{I}_{L+1})\widehat{\mathbf{h}}\|^2/N\}$. Notably, the first iteration of our channel estimator coincides with the linear MMSE channel estimator of [21].

We now compare the BER performance of the iterative block equalizer for OFDM [13], assuming different choices for the basis functions used by the channel estimation algorithm proposed in this paper. Here we assume $B_c = 2$, while the other parameters are the same of Figure 2. Figure 4 explains that GCE-BEM outperforms CE-BEM, for all the

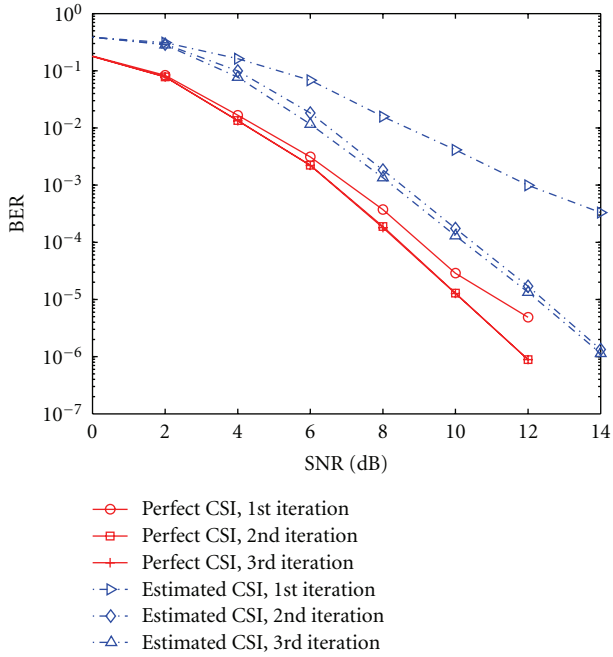

Figure 2: BER performance of OFDM systems.

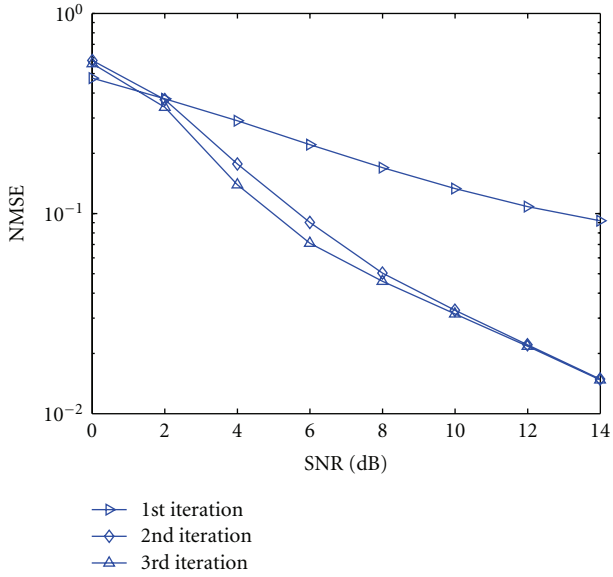

Figure 3: Channel estimation performance of OFDM systems.

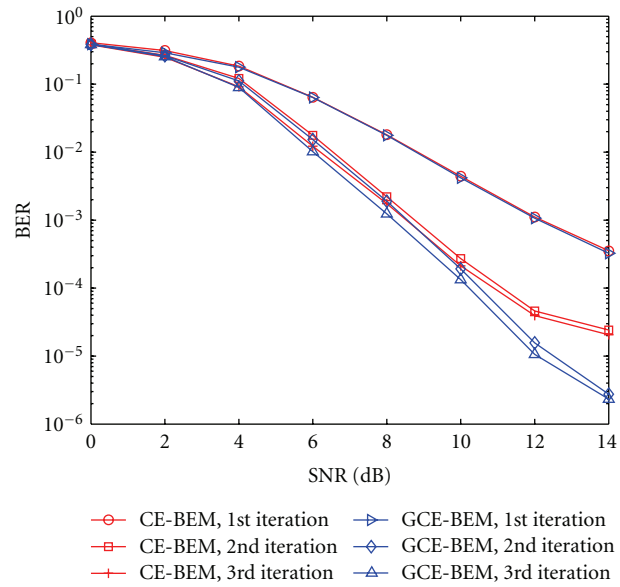

Figure 4: BER performance comparison of different basis functions for OFDM systems.

iterations. These results, obtained for soft-decision data-aided pilot-based channel estimation, are consistent with the results obtained in [21] for nondata-aided and covariance-data-aided pilot-based channel estimation.

Figure 5 shows the BER performance comparison of the proposed iterative frequency-domain block turbo equalizer for SC systems with perfect CSI. We employ $ML_p = 52$ time-domain pilot symbols, grouped into $M = 4$ clusters, which means there are $(L_p - 1)/2 = 6$ guard time symbols on each side of the nonzero pilot in each cluster. In this case, the efficiency is $\eta = (N - ML_p)/(N + \mathrm{L}) = 0.78$. We now use a BEM order $Q = 2$. The observation length parameter is set to $\Delta = -6$, which leads to a total observation length of $M(L_p + L + 2\Delta) = 32$ for each block. It can be observed that the time-domain receiver windowing can improve the BER performance over the system without windowing [34] under doubly selective channels. With time-domain receiver windowing, the BER performances of the average approximation of $\overline{\nu}_i = (1/N)(\sum_{k=1}^{N} \nu_k - \nu_i)$ and the maximum approximation of $\overline{\nu}_i = \overline{\nu} = \max\{\nu_k\}_{k=1}^{N}$ in (8) are almost the same after three iterations.

Figure 6 illustrates the BER performance of the proposed iterative frequency-domain block turbo equalizer for SC systems with average approximation, for both perfect and estimated CSI. It is shown that the second iteration of our block turbo equalizer achieves about 1.5 dB gain with respect to the first iteration, which corresponds to the output of a noniterative equalizer. However, a third iteration does not help. Figure 7 plots the NMSE of the MMSE channel estimator. Note that, differently from the considered SC scenario, in OFDM systems we have not used any guard bands around the nonzero pilot tone. The reason is that in OFDM systems the ICI power has a rapid decay. On the contrary, since the considered channel has a uniform power-delay profile, it generates significant ISI. Thus, the iterative process is not capable of suppressing a lot of interference, and large guard intervals are needed to accurately estimate the CSI.

It is interesting to compare the SC and OFDM systems in doubly selective channels. Previous work has shown some performance comparisons for frequency-selective channels [5, 35]. Figures 8 and 9 illustrate the BER performance comparisons for SC and OFDM systems with different Doppler spreads and different channel lengths, respectively. The BER curves are for the turbo equalizers with perfect

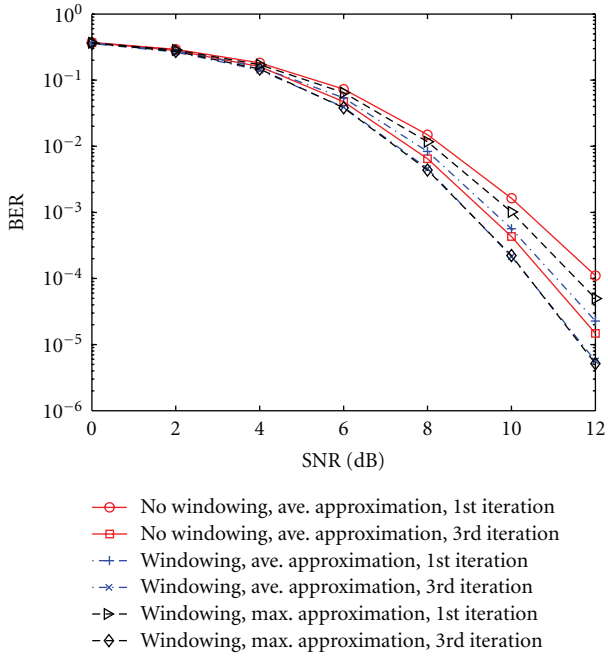

Figure 5: BER performance comparison of different equalization approaches for SC systems.

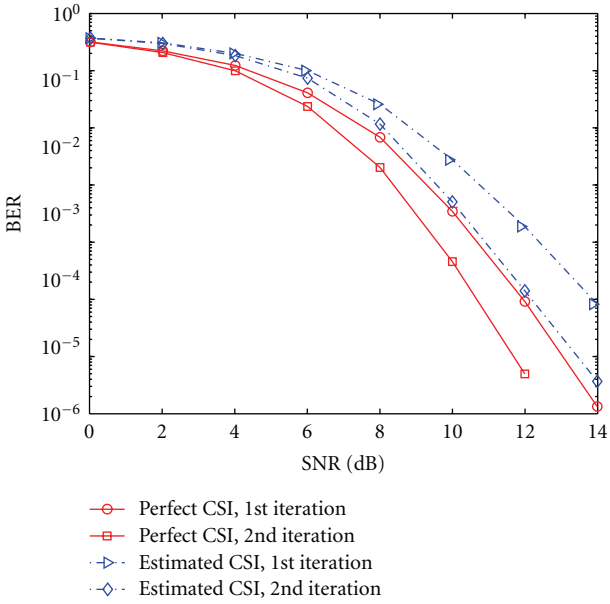

Figure 6: BER performance of SC systems.

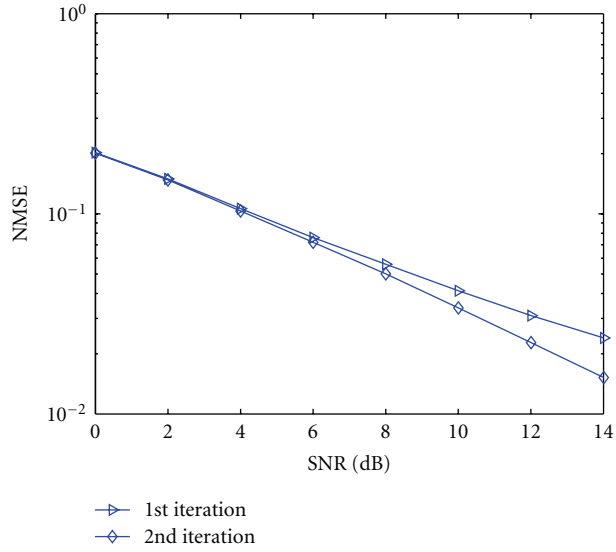

Figure 7: Channel estimation performance of SC systems.

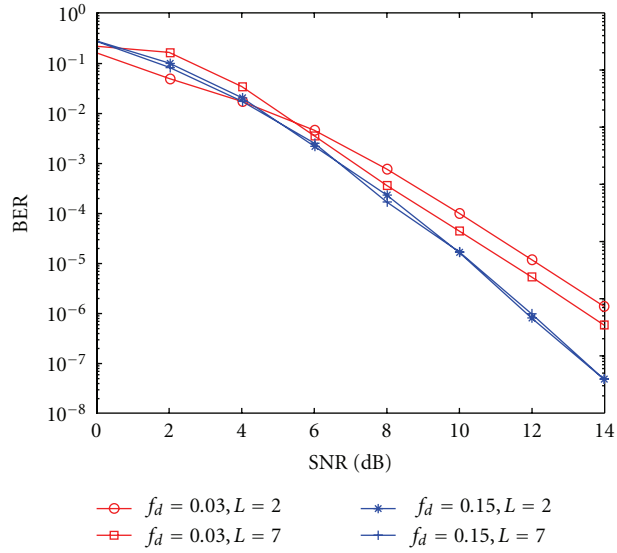

Figure 8: BER comparison of OFDM systems.

CSI after three iterations. The bandwidth parameter $B_c = 3$ is the same for both SC and OFDM systems. It can be shown that, as the values of Doppler spread and channel length increase, a lower BER can be achieved at high SNR. The simulation results confirm that both SC and OFDM systems benefit from channel coding, by obtaining delay and Doppler diversity, respectively, which is not fully exploited in the uncoded case. However, the achievable diversity gain is difficult to analyze, since the cyclic band approximation error impairs the performance at high SNR, and the amount of diversity also depends on the specific error correction code used [5, 35].

6. Conclusions

We have proposed a low-complexity frequency-domain block MMSE turbo equalizer for SC systems in doubly selective channels. We have exploited the cyclically banded structure of the frequency-domain channel matrix, as well as receiver windowing that enforces the cyclically banded structure, to limit the computational complexity, which is linear in the block length. For both OFDM and SC systems, we have developed two iterative MMSE pilot-assisted channel estimators, where the soft data estimates

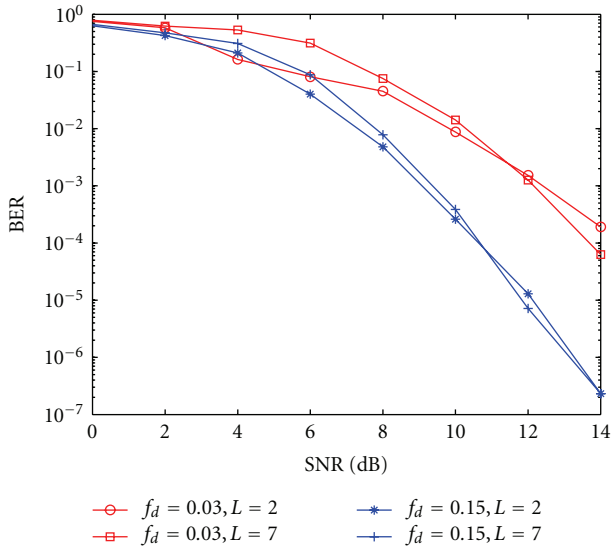

Figure 9: BER comparison of SC systems.

from the turbo equalizers are used to improve the quality of the channel estimates. Combined with error correction coding, both OFDM and SC systems can effectively exploit the delay-Doppler diversity provided by doubly selective channels.

Acknowledgments

The authors would like to thank one of the reviewers, who provided insightful comments and interesting suggestions. This research was supported in part by NWO-STW under the VIDI program (DTC.6577) and VICI program (DTC.5893). This paper was presented in part at the International Conference on Acoustics, Speech and Signal Processing, Las Vegas, NV, April 2008, and International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, April 2009.

References

[1] A. M. Sayeed and B. Aazhang, "Joint multipath-doppler diversity in mobile wireless communications," *IEEE Transactions on Communications*, vol. 47, no. 1, pp. 123–132, 1999.

[2] X. Ma and G. B. Giannakis, "Maximum-diversity transmissions over doubly selective wireless channels," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1832–1840, 2003.

[3] 3GPP, "LTE Physical Layer—General Description," TS 36.201 - v1.0.0, March 2007.

[4] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications: where Fourier meets Shannon," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 29–48, 2000.

[5] F. Pancaldi, G. M. Vitetta, R. Kalbasi, N. Al-Dhahir, M. Uysal, and H. Mheidat, "Single-carrier frequency domain equalization: a focus on wireless applications," *IEEE Signal Processing Magazine*, vol. 25, no. 5, pp. 37–56, 2008.

[6] C. Douillard, M. Jezequel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: turbo-equalization," *European Transactions on Telecommunications and Related Technologies*, vol. 6, no. 5, pp. 507–511, 1995.

[7] M. Tüchler, R. Koetter, and A. C. Singer, "Turbo equalization: principles and new results," *IEEE Transactions on Communications*, vol. 50, no. 5, pp. 754–767, 2002.

[8] M. Tüchler, A. C. Singer, and R. Koetter, "Minimum mean squared error equalization using a priori information," *IEEE Transactions on Signal Processing*, vol. 50, no. 3, pp. 673–683, 2002.

[9] R. R. Lopes and J. R. Barry, "The soft-feedback equalizer for turbo equalization of highly dispersive channels," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 783–788, 2006.

[10] M. Tüchler and J. Hagenauer, "Turbo equalization using frequency domain equalizers," in *Proceedings of the Allerton Conference*, Monticello, Ill, USA, October 2000.

[11] H. Liu and P. Schniter, "Iterative frequency-domain channel estimation and equalization for single-carrier transmissions without cyclic-prefix," *IEEE Transactions on Wireless Communications*, vol. 7, no. 10, pp. 3686–3691, 2008.

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

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

[14] S. Ahmed, M. Sellathurai, S. Lambotharan, and J. A. Chambers, "Low-complexity iterative method of equalization for single carrier with cyclic prefix in doubly selective channels," *IEEE Signal Processing Letters*, vol. 13, no. 1, pp. 5–8, 2006.

[15] P. Schniter and H. Liu, "Iterative frequency-domain equalization for single-carrier systems in doubly-dispersive channels," in *Proceedings of the Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 667–671, Pacific Grove, Calif, USA, November 2004.

[16] I. Barhumi and M. Moonen, "MLSE and MAP equalization for transmission over doubly selective channels," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 8, pp. 4120–4128, 2009.

[17] B. K. Ng and D. Falconer, "A novel frequency domain equalization method for single-carrier wireless transmissions over doubly-selective fading channels," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '04)*, vol. 1, pp. 237–241, Dallas, Tex, USA, November, 2004.

[18] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions: general model, design criteria, and signal processing," *IEEE Signal Processing Magazine*, vol. 21, no. 6, pp. 12–25, 2004.

[19] X. Ma, G. B. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly selective wireless fading channels," *IEEE Transactions on Signal Processing*, vol. 51, no. 5, pp. 1351–1366, 2003.

[20] A. P. Kannu and P. Schniter, "Design and analysis of MMSE pilot-aided cyclic-prefixed block transmissions for doubly selective channels," *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1148–1160, 2008.

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

[22] K. A. D. Teo and S. Ohno, "Pilot-aided channel estimation and viterbi equalization for OFDM over doubly-selective channel," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '06)*, San Francisco, Calif, USA, December 2006.

[23] S. Barbarossa, "Eigenfunctions of underspread linear communication systems," in *Time-Frequency Analysis*, B. Boashash, Ed., Elsevier, Amsterdam, The Netherlands, 2003.

[24] S. He and J. K. Tugnait, "On doubly selective channel estimation using superimposed training and discrete prolate spheroidal sequences," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3214–3228, 2008.

[25] R. Otnes and M. Tüchler, "Iterative channel estimation for turbo equalization of time-varying frequency-selective channels," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 1918–1923, 2004.

[26] X. Li and T. F. Wong, "Turbo equalization with nonlinear Kalman filtering for time-varying frequency-selective fading channels," *IEEE Transactions on Wireless Communications*, vol. 6, no. 2, pp. 691–700, 2007.

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

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

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

[30] M. Hegland, "Divide and conquer for the solution of banded linear systems of equations," in *Proceedings of the IEEE Symposium on Parallel and Distributed Processing (PDP '96)*, pp. 394–401, Washington, DC, USA, January 1996.

[31] J. Bonnet and G. Auer, "Optimized iterative channel estimation for OFDM," in *Proceedings of the IEEE Vehicular Technology Conference (VTC '06)*, pp. 1621–1625, Montreal, Canada, September 2006.

[32] W. C. Jakes, *Microwave Mobile Channels*, John Wiley & Sons, New York, NY, USA, 1974.

[33] Y. R. Zheng and C. Xiao, "Simulation models with correct statistical properties for Rayleigh fading channels," *IEEE Transactions on Communications*, vol. 51, no. 6, pp. 920–928, 2003.

[34] K. Fang, L. Rugini, and G. Leus, "Low-complexity frequency-domain turbo equalization for single-carrier transmissions over doubly-selective channels," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '09)*, pp. 2541–2544, Taipei, Taiwan, April 2009.

[35] Z. Wang, X. Ma, and G. B. Giannakis, "OFDM or single-carrier block transmissions?" *IEEE Transactions on Communications*, vol. 52, no. 3, pp. 380–394, 2004.