AALBORG UNIVERSITY

DENMARK

Channel estimation based on divergence minimization for OFDM systems with co-channel interference

Manchón, Carles Navarro; Fleury, Bernard Henri; Kirkelund, Gunvor Elisabeth; Mogensen, Preben; Deneire, Luc; Sørensen, Troels Bundgaard; Rom, Christian

Channel Estimation Based on Divergence Minimization for OFDM Sytems with Co-Channel Interference

Carles Navarro Manchón[*], Bernard Fleury[*†], Gunvor E. Kirkelund[*], Preben Mogensen[*],
Luc Deneire[*‡], Troels B. Sørensen[*] and Christian Rom[§]

[*]Department of Electronic Systems, Aalborg University
Niels Jernes Vej 12, 9220 Aalborg East
[†]Forschungszentrum Telekommunikation Wien (FTW), Vienna, Austria
[‡]Université de Nice, Sophia Antipolis
Centre National de la Recherche Scientifique
I3S, UMR 6070, France
[§]Infineon Technologies Denmark A/S
Alfred Nobels Vej 25, DK-9220 Aalborg, Denmark

Abstract—In this paper, we present a novel approach for pilot-aided channel estimation in OFDM systems with synchronous co-channel interference. The estimator is derived based on the Kullback-Leibler divergence minimization framework. The obtained solution iteratively updates both the desired user's and the interferer's channels, using a combination of linear minimum mean squared-error (LMMSE) filtering and interference cancellation, avoiding the complex matrix inversions involved in the full LMMSE channel estimation approach. Estimation of the noise variance is also included in the iterative algorithm, accounting for Gaussian noise and residual interference after each iteration. The estimates of both channels are used at the equalizer to reject the interfering signal, thus mitigating the degradation due to co-channel interference. Simulation results show that the receiver using the proposed estimator performs as good as the one employing the full LMMSE estimator and very closely to a receiver having perfect knowledge of the channel coefficients.

I. INTRODUCTION

Orthogonal Frequency Division Multiplexing (OFDM) has become the selected transmission technique for several recent wireless standards, such as the IEEE standard for local and metropolitan area networks (better known as WiMAX) [1], or the 3GPP UTRA Long Term Evolution (LTE) [2]. Its ability to cope with time-dispersive channels while allowing for receivers with low complexity, its ability to easily integrate multiple antenna techniques and its flexibility in terms of bandwidth usage and resource allocation are some of the advantages that have motivated its selection.

In OFDM, the transmission bandwidth is divided into multiple narrowband subcarriers. By the addition of a proper cyclic prefix (CP), these subcarriers become fully orthogonal and experience frequency flat fading conditions in time-invariant channels [3]. This allows for simple equalization of the signal at the receiver, while keeping a high spectral efficiency due to the use of orthogonal overlapping subcarriers. In OFDM systems with frequency re-use, however, the signal transmitted

from other cells may create co-channel interference which, if not correctly treated, can induce a severe degradation of the receiver performance, especially at the cell edge.

Much work has been done in interference cancellation techniques for OFDM, as in [4]–[6]. These methods, however, assume perfect knowledge of the channel at the receiver. In [7], a minimum mean-squared error interference rejection combiner (MMSE-IRC) for OFDM receivers with multiple antennas is proposed. The combiner parameters are estimated using a discrete-Fourier-transform-based robust MMSE instantaneous correlation estimator, which is therefore sensitive to the *leakage* effect [8] when the channel delays are not perfectly aligned with the receiver sampling grid.

In this work, we propose a pilot-aided channel estimator for OFDM systems with severe synchronous co-channel interference in both the data and pilot subcarriers. The estimator is derived by applying the Kullback-Leibler (KL) divergence minimization (DM) approach, which was presented in [9] for multiuser detection in a code-division multiple access system. Our proposed scheme is able to estimate the desired user's and the interferer's channels based on merely the signal received at pilot subcarriers. The estimates are then used in a MMSE-IRC combiner, effectively mitigating the effect of the interference. A similar problem was studied in [10]. The solution proposed there, however, requires a preamble in which no interference is present at the pilot subcarriers. Our estimator, on the contrary, can effectively separate and estimate both channels when the pilot signals of the desired user and the interferer overlap in frequency for every OFDM symbol.

The remainder of the paper is organized as follows. The signal model for our considered system is presented in Section II. In Section III, the DM framework is briefly introduced, and the channel estimator is derived. The performance of the estimator is assessed by means of Monte-Carlo simulations in Section IV. Finally, some concluding remarks are given in
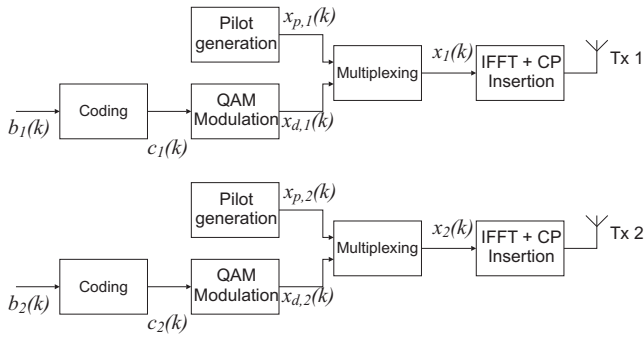
Fig. 1. Block diagram of the transmitters.

Fig. 2. Block diagram of the receiver

Section V.

The following notation will be used throughout the paper. Vectors are represented by boldface lowercase letters, while matrices are denoted as boldface uppercase letters; $(\cdot)^T$ and $(\cdot)^H$ denote respectively the transpose and conjugate transpose of a vector; tr$\{\cdot\}$ denotes the trace operation, and $diag\{\mathbf{x}\}$ represents a diagonal matrix with the elements of vector $\mathbf{x}$; $x \propto y$ denotes direct proportionality, i.e., $x = \alpha y$, and $x \propto^e y$ denotes exponential proportionality, i.e., $\exp[x] = \exp[\beta + y]$, for arbitrary constants α and β; finally, $E_{q_x}\{f(\mathbf{x})\}$ represents the expectation of the function $f(\mathbf{x})$ with respect to the probability distribution $q_x(\mathbf{x})$ of $\mathbf{x}$.

II. SIGNAL MODEL

We consider an OFDM system with single transmit antenna and one interferer, as depicted in Fig. 1. The first transmitter represents the user of interest, while the second transmitter represents a synchronized interferer transmitting in the same time-frequency resources. The information bits $b_m(k)$, $m = 1, 2$, $k = 0, \ldots, N_b - 1$ are coded, yielding a stream of coded bits $c_m(k)$, $k = 0, \ldots, N_c - 1$. These are modulated onto a set of QAM symbols $x_{d,m}(k)$, $k = 0, \ldots, N_d - 1$ to be mapped onto an OFDM block. The number of subcarriers used for data transmission in an OFDM block is $N_d = N_c/M$, N_c being the number of coded bits transmitted in one OFDM block and M being the modulation order. The data symbols are then multiplexed with a sequence of pilot symbols $x_{p,m}(k)$, $k = 0, \ldots, N_p - 1$, N_p being the number of pilot subcarriers per block. We assume that pilot symbols are allocated to the same subcarriers at both transmitters. The resulting sequence of symbols $x_m(k)$, $k = 0, \ldots, N_u - 1$ is then mapped to the $N_u = N_d + N_p$ active subcarriers of the OFDM system and transmitted through the wireless channel after insertion of a cyclic prefix (CP). We assume in this work that the CP is long enough to cope with the time dispersion in both the desired and interfering channels.

The structure of the receiver is shown in Fig. 2. We assume a receiver with two antennas. The extension to a higher number of antennas is straightforward. After FFT and CP removal, the received signal at the kth subcarrier of the nth antenna port is given by

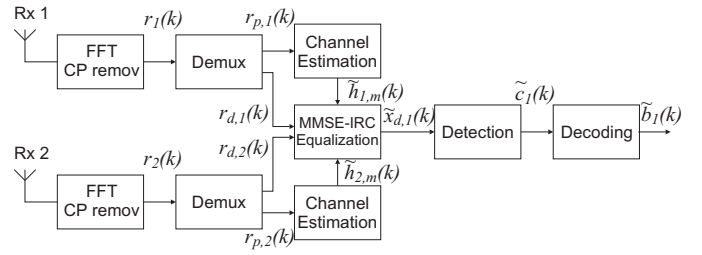$$r_n(k) = x_1(k)h_{n1}(k) + x_2(k)h_{n2}(k) + w_n(k), \quad (1)$$

where $w_n(k)$ is additive white Gaussian noise (AWGN) with variance σ_w^2 and $h_{nm}(k)$ represents the frequency-domain channel gain from transmitter m to receive antenna n at the kth subcarrier. In (1), we assume that the channel response is static during one OFDM block. Hence, full orthogonality between subcarriers is achieved. The received signal at antenna port n for all subcarriers can be re-written in matrix-vector notation as

$$\mathbf{r}_n = \mathbf{X}_1 \mathbf{h}_{n1} + \mathbf{X}_2 \mathbf{h}_{n2} + \mathbf{w}_n \quad (2)$$

with $\mathbf{r}_n = [r_n(0) \cdots r_n(N_u - 1)]^T$, $\mathbf{h}_{n,m} = [h_{nm}(0) \cdots h_{nm}(N_u - 1)]^T$, $\mathbf{w}_n = [w_n(0) \cdots w_n(N_u - 1)]^T$ and $\mathbf{X}_m = diag\{[x_m(0) \cdots x_m(N_u - 1)]\}$ being a diagonal matrix containing the transmitted symbols.

The demultiplexer following the FFT and CP removal block separates the signal received at pilot and data subcarriers. The pilot signals $\mathbf{r}_{p,n} = [r_{p,n}(0) \cdots r_{p,n}(N_p - 1)]^T$ are fed to the respective channel estimator blocks, while the data signals $\mathbf{r}_{d,n} = [r_{d,n}(0) \cdots r_{d,n}(N_d - 1)]^T$ are sent to the equalizer. Based on the signal received on the pilot subcarriers, the channel estimation block (which will be explained in detail in Section III) provides the equalizer with estimates $\tilde{\mathbf{h}}_{n,m}$ of the channel frequency responses of both the desired and interfering channels. Using these estimates and the signal received at data subcarriers, the equalizer performs MMSE-IRC filtering to recover the desired transmitted symbols as

$$\tilde{x}_{d,1}(k) = \tilde{\mathbf{h}}_{d,1}^H(k) \left(\tilde{\mathbf{H}}_d^H(k)\tilde{\mathbf{H}}_d(k) + \sigma_w^2 \mathbf{I} \right)^{-1} \mathbf{r}_d(k). \quad (3)$$

In the above equation, $\mathbf{r}_d(k) = [r_{d,1}(k) r_{d,2}(k)]^T$, $\mathbf{H}_d(k) = [\mathbf{h}_{d,1}(k) \mathbf{h}_{d,2}(k)]$ and $\mathbf{h}_{d,m}(k) = [h_{d,1m}(k) h_{d,2m}(k)]^T$, with $h_{d,nm}(k)$ denoting the gain of the time-frequency response of the nm channel at the kth data subcarrier, and $\mathbf{I}$ denotes the 2x2 identity matrix. Finally, the coded bits of the user of interest $\tilde{c}_1(k)$ are recovered from the equalized symbols in the QAM detector and are fed to the channel decoder which yields the estimates of the information bits $\tilde{b}_1(k)$.

III. CHANNEL ESTIMATOR

In this section, two channel estimation approaches are presented. The first one is the LMMSE estimator, which will be used as a benchmark for the performance evaluation of our estimator. Next, our proposed channel estimator based on the DM framework is introduced. The DM principle is briefly explained, along with the application to our specific scenario.

More details about the DM framework and its relation to other known algorithms can be found in [9].

A. LMMSE Channel Estimator

The linear LMMSE channel estimator aims at minimizing the mean-squared error of the estimate. For the signal model under consideration, the LMMSE transfer function estimate of the $n1$ channel reads:

$$
\begin{aligned}
\tilde{\mathbf{h}}_\mathbf{1} &= \arg\min_{\tilde{\mathbf{h}}_{n1}} \mathrm{E}\{(\mathbf{h}_{n1} - \tilde{\mathbf{h}}_{n1})^H (\mathbf{h}_{n1} - \tilde{\mathbf{h}}_{n1})\} \\
&= \boldsymbol{\Sigma}_{h_{n1}h_{p,n1}} \mathbf{X}_{p,1}^H (\mathbf{X}_{p,1} \boldsymbol{\Sigma}_{h_{p,n1}} \mathbf{X}_{p,1}^H \\
&\quad + \mathbf{X}_{p,2} \boldsymbol{\Sigma}_{h_{p,n2}} \mathbf{X}_{p,2}^H + \boldsymbol{\Sigma}_{w_p})^{-1} \mathbf{r}_{p,n}
\end{aligned}
\tag{4}
$$

where $\boldsymbol{\Sigma}_{w_p} = \mathrm{E}\{\mathbf{w}_p \mathbf{w}_p^H\} = \sigma_w^2 \mathbf{I}$, $\boldsymbol{\Sigma}_{h_{n1}h_{p,n1}} = \mathrm{E}\{\mathbf{h}_{n1}\mathbf{h}_{p,n1}^H\}$ and $\boldsymbol{\Sigma}_{h_{p,nm}} = \mathrm{E}\{\mathbf{h}_{p,nm}\mathbf{h}_{p,nm}^H\}$. Note that we assume no correlation between the channels $n1$ and $n2$. The estimator requires the inversion of an $N_p \times N_p$ matrix every OFDM symbol, which is normally too complex to compute in a mobile receiver for a system with a large number of subcarriers. In the rest of the section, we present an iterative approach which avoids this matrix inversion.

B. Divergence minimization

Let $\boldsymbol{\Phi}$ denote a vector including as components all the unknown parameters to be estimated and $p(\boldsymbol{\Phi}|\mathbf{r})$ be the posterior probability density function (pdf) of $\boldsymbol{\Phi}$ given an observation $\mathbf{r}$. The DM framework approximates $p(\boldsymbol{\Phi}|\mathbf{r})$ by an auxiliary pdf $q(\boldsymbol{\Phi})$ minimizing the KL divergence [11]

$$
D\Big(q(\boldsymbol{\Phi}) \Big\| p(\boldsymbol{\Phi}|\mathbf{r})\Big) \triangleq \int d\boldsymbol{\Phi} \, q(\boldsymbol{\Phi}) \log \frac{q(\boldsymbol{\Phi})}{p(\boldsymbol{\Phi}|\mathbf{r})}
\tag{5}
$$

In our application, the parameters to estimate are the channel responses of the desired and interfering channels, as well as the inverse of the noise covariance matrix, i.e., $\boldsymbol{\Phi} = \{\mathbf{h}_{p,n1}, \mathbf{h}_{p,n2}, \boldsymbol{\Sigma}_{w_{p,n}}^{-1}\}$, where $\boldsymbol{\Sigma}_{w_{p,n}} = \mathrm{E}\{\mathbf{w}_{p,n}\mathbf{w}_{p,n}^H\}$. The index p indicates that only pilot subcarriers are taken into account. In order to get a solution that can be computed with tractable complexity, the auxiliary pdf is assumed to factorize according to

$$
\begin{aligned}
q(\boldsymbol{\Phi}) &= q(\mathbf{h}_{p,n1}, \mathbf{h}_{p,n2}, \boldsymbol{\Sigma}_{w_{p,n}}^{-1}) \\
&= q_{h_{p,n1}}(\mathbf{h}_{p,n1}) q_{h_{p,n2}}(\mathbf{h}_{p,n2}) q_{\Sigma_{w_{p,n}}^{-1}}(\boldsymbol{\Sigma}_{w_{p,n}}^{-1}).
\end{aligned}
\tag{6}
$$

The observation is the received signal at the pilot subcarriers, i.e

$$
\mathbf{r} = \mathbf{r}_{p,n} = \mathbf{X}_{p,1}\mathbf{h}_{p,n1} + \mathbf{X}_{p,2}\mathbf{h}_{p,n2} + \mathbf{w}_{p,n}.
\tag{7}
$$

The algorithm iteratively minimizes the KL divergence with respect to one of the factors in (7), while the other factors are kept fixed, resulting in an iterative scheme.

Note that the channel estimation process is done independently for each of the receive antennas. In the remainder of the section we therefore drop the receive antenna subindex n in order to simplify the notation (e.g. $\mathbf{h}_{p,1}$ denotes $\mathbf{h}_{p,n1}$). The algorithm is started with initial distributions $q_{h_{p,1}}^{[0]}(\mathbf{h}_{p,1})$, $q_{h_{p,2}}^{[0]}(\mathbf{h}_{p,2})$ and $q_{\Sigma_{w_p}^{-1}}^{[0]}(\boldsymbol{\Sigma}_{w_p}^{-1})$, and these distributions are successively updated according to the updating steps detailed in the following two subsections.

C. Update of the channel vectors

In this subsection, the derivation of the updating step for $q_{h_{p,1}}(\mathbf{h}_{p,1})$ is detailed. Due to the symmetry of the problem, the update for $q_{h_{p,2}}(\mathbf{h}_{p,2})$ is analogous.

To update $q_{h_{p,1}}(\mathbf{h}_{p,1})$, the distributions $q_{h_{p,2}}^{[i]}(\mathbf{h}_{p,2})$ and $q_{\Sigma_{w_p}^{-1}}^{[i]}(\boldsymbol{\Sigma}_{w_p}^{-1})$ are kept fixed, and $q_{h_{p,1}}(\mathbf{h}_{p,1})$ is updated by solving the following problem:

$$
\begin{aligned}
\text{minimize} \quad & D\Big(q_{h_{p,1}}(\mathbf{h}_{p,1}) q_{h_{p,2}}^{[i]}(\mathbf{h}_{p,2}) \\
& \cdot q_{\Sigma_{w_p}^{-1}}^{[i]}(\boldsymbol{\Sigma}_{w_p}^{-1}) \Big\| p(\mathbf{h}_{p,1}, \mathbf{h}_{p,2}, \boldsymbol{\Sigma}_{w_p}^{-1}|\mathbf{r}_p)\Big) \\
\text{subject to} \quad & \int q_{h_{p,1}}(\mathbf{h}_{p,1}) d\mathbf{h}_{p,1} = 1 \\
& q_{h_{p,1}}(\mathbf{h}_{p,1}) \geq 0.
\end{aligned}
\tag{8}
$$

The distribution $q_{h_{p,1}}^{[i+1]}(\mathbf{h}_{p,1})$ solving (8) is found to be

$$
\begin{aligned}
q_{h_{p,1}}^{[i+1]}(\mathbf{h}_{p,1}) &\propto p(\mathbf{h}_{p,1}) \\
& \cdot \exp\Big[\mathrm{E}_{q_{h_{p,2}}^{[i]}} \Big\{ \mathrm{E}_{q_{\Sigma_{w_p}^{-1}}^{[i]}} \Big\{ \log p(\mathbf{r}_p|\mathbf{h}_{p,1}, \mathbf{h}_{p,2}, \boldsymbol{\Sigma}_{w_p}^{-1}) \Big\} \Big\} \Big]
\end{aligned}
\tag{9}
$$

where $p(\mathbf{h}_{p,1})$ is the prior pdf of $\mathbf{h}_{p,1}$. The log-likelihood function in (9) reads

$$
\begin{aligned}
& \log p(\mathbf{r}_p|\mathbf{h}_{p,1}, \mathbf{h}_{p,2}, \boldsymbol{\Sigma}_{w_p}^{-1}) \\
& \propto^e \log |\boldsymbol{\Sigma}_{w_p}^{-1}| - \mathrm{tr}\Big\{ \boldsymbol{\Sigma}_{w_p}^{-1} (\mathbf{r}_p - \mathbf{X}_{p,1}\mathbf{h}_{p,1} - \mathbf{X}_{p,2}\mathbf{h}_{p,2}) \\
& \quad \cdot (\mathbf{r}_p - \mathbf{X}_{p,1}\mathbf{h}_{p,1} - \mathbf{X}_{p,2}\mathbf{h}_{p,2})^H \Big\}.
\end{aligned}
\tag{10}
$$

The marginalization of (10) with respect to $q_{h_{p,2}}^{[i]}(\mathbf{h}_{p,2})$ and $q_{\Sigma_{w_p}^{-1}}^{[i]}(\boldsymbol{\Sigma}_{w_p}^{-1})$ yields

$$
\begin{aligned}
& \mathrm{E}_{q_{h_{p,2}}^{[i]}} \Big\{ \mathrm{E}_{q_{\Sigma_{w_p}^{-1}}^{[i]}} \Big\{ \log p(\mathbf{r}_p|\mathbf{h}_{p,1}, \mathbf{h}_{p,2}, \boldsymbol{\Sigma}_{w_p}^{-1}) \Big\} \Big\} \\
& \propto^e -\mathrm{tr}\Big\{ (\boldsymbol{\Omega}_{w,p}^{-1})^{[i]} \mathbf{A}^{[i]} \Big\},
\end{aligned}
\tag{11}
$$

where $(\boldsymbol{\Omega}_{w,p}^{[i]})^{-1} \triangleq \mathrm{E}_{q_{\Sigma_{w_p}^{-1}}^{[i]}}\Big\{ \boldsymbol{\Sigma}_{w_p}^{-1} \Big\}$ and

$$
\begin{aligned}
\mathbf{A}^{[i]} &= (\mathbf{r}_p - \mathbf{X}_{p,1}\mathbf{h}_{p,1} - \mathbf{X}_{p,2}\mathbf{h}_{p,2}^{[i]}) \\
& \cdot (\mathbf{r}_p - \mathbf{X}_{p,1}\mathbf{h}_{p,1} - \mathbf{X}_{p,2}\mathbf{h}_{p,2}^{[i]})^H + \mathbf{X}_{p,2}\boldsymbol{\Sigma}_{h_{p,2}}^{[i]}\mathbf{X}_{p,2}^H.
\end{aligned}
\tag{12}
$$

Details on $(\boldsymbol{\Omega}_{w,p}^{[i]})^{-1}$ and $\boldsymbol{\Sigma}_{h_{p,2}}^{[i]}$ are given later on in this section. Note that terms independent of $\mathbf{h}_{p,1}$ have been neglected in the derivation as they do not affect $q_{h_{p,1}}^{[i+1]}(\mathbf{h}_{p,1})$.

For Rayleigh fading channels, the prior distribution of $\mathbf{h}_{p,1}$ is Gaussian with zero mean and covariance matrix $\boldsymbol{\Sigma}_{h_{p,1}} = \mathrm{E}\Big\{ \mathbf{h}_{p,1}\mathbf{h}_{p,1}^H \Big\}$. Using this prior distribution and (11) in (9), we obtain an updated distribution, which is also Gaussian, with pdf:

$$
\begin{aligned}
& q_{h_{p,1}}^{[i+1]}(\mathbf{h}_{p,1}) \propto \\
& \exp\Big[-(\mathbf{h}_{p,1} - \mathbf{h}_{p,1}^{[i+1]})^H \boldsymbol{\Sigma}_{h_{p,1}}^{[i+1]^{-1}} (\mathbf{h}_{p,1} - \mathbf{h}_{p,1}^{[i+1]}) \Big]
\end{aligned}
\tag{13}
$$

with mean vector

$$\mathbf{h}_{p,1}^{[i+1]} = \boldsymbol{\Sigma}_{h_{p,1}} \left(\boldsymbol{\Sigma}_{h_{p,1}} \mathbf{X}_{p,1}^H \mathbf{X}_{p,1} + \boldsymbol{\Omega}_{w,p}^{[i]} \right)^{-1} \mathbf{X}_{p,1}^H$$
$$\cdot \left(\mathbf{r}_p - \mathbf{X}_{p,2} \mathbf{h}_{p,2}^{[i]} \right) \quad (14)$$

and covariance matrix

$$\boldsymbol{\Sigma}_{h_{p,1}}^{[i+1]} = \left(\boldsymbol{\Sigma}_{h_{p,1}}^{-1} + (\boldsymbol{\Omega}_{w,p}^{[i]})^{-1} \mathbf{X}_{p,1}^H \mathbf{X}_{p,1} \right)^{-1}. \quad (15)$$

By inspecting (14), it is seen that the channel response updating step has the form of a LMMSE or Wiener filter [12], applied to the interference-cancelled received signal at pilot positions. As the interference cancellation is not ideal, the estimate of the noise covariance matrix $\boldsymbol{\Omega}_{w,p}^{[i]}$ takes into account both the noise and the residual interference power, in order to correctly smooth the channel response, as it is shown in the next subsection.

Note that the update of the channel coefficients in (14) does only provide estimates of the channel response at pilot subcarriers. Estimates of the full frequency response at all active subcarriers are obtained by using

$$\tilde{\mathbf{h}}_1 = \boldsymbol{\Sigma}_{h_1 h_{p,1}} \left(\boldsymbol{\Sigma}_{h_{p,1}} \mathbf{X}_{p,1}^H \mathbf{X}_{p,1} + \boldsymbol{\Omega}_{w,p}^{[i]} \right)^{-1} \mathbf{X}_{p,1}^H$$
$$\cdot \left(\mathbf{r}_p - \mathbf{X}_{p,2} \mathbf{h}_{p,2}^{[i]} \right) \quad (16)$$

instead of (14) in the last iteration of the algorithm with $\boldsymbol{\Sigma}_{h_1 h_{p,1}} = \mathrm{E}\left\{ \mathbf{h}_1 \mathbf{h}_{p,1}^H \right\}$.

D. Update of the noise covariance matrix

When updating $q_{\boldsymbol{\Sigma}_{w_p}^{-1}}(\boldsymbol{\Sigma}_{w_p}^{-1})$, the distributions $q_{h_{p,1}}^{[i]}(\mathbf{h}_{p,1})$ and $q_{h_{p,2}}^{[i]}(\mathbf{h}_{p,2})$ are kept fixed, and the optimization problem to solve is the following:

$$\begin{aligned} \text{minimize} \quad & D\Big(q_{h_{p,1}}^{[i]}(\mathbf{h}_{p,1}) q_{h_{p,2}}^{[i]}(\mathbf{h}_{p,2}) \\ & \cdot q_{\boldsymbol{\Sigma}_{w_p}^{-1}}(\boldsymbol{\Sigma}_{w_p}^{-1}) \Big\| p(\mathbf{h}_{p,1}, \mathbf{h}_{p,2}, \boldsymbol{\Sigma}_{w_p}^{-1} | \mathbf{r}_p) \Big) \\ \text{subject to} \quad & \int q_{\boldsymbol{\Sigma}_{w_p}^{-1}}(\boldsymbol{\Sigma}_{w_p}^{-1}) d\boldsymbol{\Sigma}_{w_p}^{-1} = 1 \\ & q_{\boldsymbol{\Sigma}_{w_p}^{-1}}(\boldsymbol{\Sigma}_{w_p}^{-1}) \geq 0. \end{aligned} \quad (17)$$

The solution reads

$$q_{\boldsymbol{\Sigma}_{w_p}^{-1}}^{[i+1]}(\boldsymbol{\Sigma}_{w_p}^{-1}) \propto p(\boldsymbol{\Sigma}_{w_p}^{-1})$$
$$\cdot \exp \left[\mathrm{E}_{q_{h_{p,1}}^{[i]}} \left\{ \mathrm{E}_{q_{h_{p,2}}^{[i]}} \left\{ \log p(\mathbf{r}_p | \mathbf{h}_{p,1}, \mathbf{h}_{p,2}, \boldsymbol{\Sigma}_{w_p}^{-1}) \right\} \right\} \right]. \quad (18)$$

The marginalization of (10) is taken with respect to $\mathbf{h}_{p,1}^{[i]}$ and $\mathbf{h}_{p,2}^{[i]}$, resulting in

$$\mathrm{E}_{q_{h_{p,1}}^{[i]}} \left\{ \mathrm{E}_{q_{h_{p,2}}^{[i]}} \left\{ \log p(\mathbf{r}_p | \mathbf{h}_{p,1}, \mathbf{h}_{p,2}, \boldsymbol{\Sigma}_{w_p}^{-1}) \right\} \right\}$$
$$\propto^e \log |\boldsymbol{\Sigma}_w^{-1}| - \mathrm{tr}\left\{ \boldsymbol{\Sigma}_{w,p}^{-1} \mathbf{B}^{[i]} \right\}, \quad (19)$$

where

$$\mathbf{B}^{[i]} = (\mathbf{r}_p - \mathbf{X}_{p,1} \mathbf{h}_{p,1}^{[i]} - \mathbf{X}_{p,2} \mathbf{h}_{p,2}^{[i]})$$
$$\cdot (\mathbf{r}_p - \mathbf{X}_{p,1} \mathbf{h}_{p,1}^{[i]} - \mathbf{X}_{p,2} \mathbf{h}_{p,2}^{[i]})^H$$
$$+ \mathbf{X}_{p,1} \boldsymbol{\Sigma}_{h_{p,1}}^{[i]} \mathbf{X}_{p,1}^H + \mathbf{X}_{p,2} \boldsymbol{\Sigma}_{h_{p,2}}^{[i]} \mathbf{X}_{p,2}^H. \quad (20)$$

By choosing the prior pdf $\boldsymbol{\Sigma}_{w_p}^{-1}$ to be flat, (18) becomes

$$q_{\boldsymbol{\Sigma}_{w_p}^{-1}}^{[i+1]}(\boldsymbol{\Sigma}_{w_p}^{-1}) \propto |\boldsymbol{\Sigma}_w^{-1}| \exp \left[-\mathrm{tr}\left\{ \boldsymbol{\Sigma}_{w,p}^{-1} \mathbf{B}^{[i]} \right\} \right], \quad (21)$$

which has the form of a Wishart distribution [13] as $\boldsymbol{\Sigma}_{w,p}^{-1} \sim W_{N_p} \left(N_p + 2, \left(\mathbf{B}^{[i]} \right)^{-1} \right)$. The mean of $\boldsymbol{\Sigma}_{w_p}^{-1}$ is therefore

$$(\boldsymbol{\Omega}_{w,p}^{[i+1]})^{-1} \triangleq \mathrm{E}_{q_{\boldsymbol{\Sigma}_{w_p}^{-1}}^{[i+1]}} \left\{ \boldsymbol{\Sigma}_{w_p}^{-1} \right\} = \left(\frac{\mathbf{B}^{[i]}}{N_p + 2} \right)^{-1}. \quad (22)$$

In order to simplify the algorithm, it is assumed that $\boldsymbol{\Sigma}_{w_p}$ represents the covariance matrix of a white Gaussian noise vector with $\boldsymbol{\Sigma}_{w_p}^{-1} = \mathrm{diag}\{\sigma_{w_p}^{-2}, \ldots, \sigma_{w_p}^{-2}\}$. In this case, the updated pdf is given by

$$q_{\sigma_{w_p}^{-2}}(\sigma_{w_p}^{-2}) \propto (\sigma_{w_p}^{-2})^{N_p} \exp \left[-\sigma_{w_p}^{-2} \mathrm{tr}\left\{ \mathbf{B}^{[i]} \right\} \right] \quad (23)$$

which is a chi-square distribution [13]. Specifically $\sigma_{w_p}^{-2} \sim \chi_{N_p+2}^2$, and the expectation of $\sigma_{w_p}^{-2}$ is

$$(\sigma_{w_p}^{-2})^{[i+1]} = \mathrm{E}_{q_{\sigma_{w_p}^{-2}}^{[i+1]}} \left\{ \sigma_{w_p}^{-2} \right\} = \left(\frac{\mathrm{tr}\left\{ \mathbf{B}^{[i]} \right\}}{N_p + 2} \right)^{-1}. \quad (24)$$

E. Implementation Issues

1) Matrix inverse in the update of the channel vectors: As it can be observed in (14), the inversion of a matrix of dimensions $N_p \times N_p$ is still required for the update of the channel coefficients vector. To avoid the matrix inversion, (14) can be rewritten as:

$$\mathbf{h}_{p,1}^{[i+1]} = \mathbf{US} \left(\mathbf{S} + (\sigma_{w_p}^{-2})^{[i]} \mathbf{I}_{N_p} \right)^{-1} \mathbf{U}^H \mathbf{X}_{p,1}^H$$
$$\cdot \left(\mathbf{r}_p - \mathbf{X}_{p,2} \mathbf{h}_{p,2}^{[i]} \right) \quad (25)$$

where $\boldsymbol{\Sigma}_{h_{p,1}} = \mathbf{USU}^H$ is the singular value decomposition (SVD) of the channel covariance matrix. We have also made use of the fact that $\mathbf{X}_{p,1}^H \mathbf{X}_{p,1} = \mathbf{I}_{N_p}$ for constant unit-power pilots, and the simplification of the noise covariance matrix introduced in (23) and (24). Note that the matrix to invert is now a diagonal matrix, which can be inverted with just N_p complex operations. Also, in a wide-sense stationary channel, the prior covariance matrices of the channels will not change over time, and therefore the SVDs need to be computed only once for each channel.

2) Initialization: Details on how to update each of the pdfs have been given previously in this section. An initialization of them for the first iteration of the algorithm, however, is needed. In our proposed implementation, the channel responses are initialized to null vectors, i.e., $\mathbf{h}_{p,m}^{[0]} = [0, \ldots, 0]^T$, and their covariance matrices are initialized to the prior covariance matrices of the channel, $\boldsymbol{\Sigma}_{h_{p,m}}^{[0]} = \boldsymbol{\Sigma}_{h_{p,m}}$. As for the noise variance, it is initialized to the AWGN variance, i.e., $(\sigma_{w_p}^{-2})^{[0]} = \sigma_w^{-2}$. In subsequent iterations, this initialization is updated with the residual interference after the interference cancellation performed in the updates of the channel response vectors.

3) Updating schedule: Another important aspect having an impact on the performance of the algorithm is the order in which the pdfs are updated. So far, no analytical way of determining the optimal updating order has been found. In this article, we evaluate the following updating order: $q_{h_{p,1}}$, $q_{h_{p,2}}$, $q_{\sigma_{w_p}^{-2}}$.

Intuitively, the desired user channel should be as strong or stronger than the interfering channel, thus it is selected to be estimated first. Once a first estimate of the desired channel is available, the interfering channel can be estimated more accurately. Finally an estimate of the residual noise plus interference is obtained to improve the channel estimates in subsequent iterations. Simulation results (which have not been included here due to lack of space) showed no relevant gain by updating the reciprocal of the noise variance between the estimates of the desired and interfering channel. Therefore, this step is not included in the algorithm, yielding a less complex scheme with no appreciable loss in performance.

IV. Performance Evaluation

In this section, we evaluate the performance of the proposed channel estimator by means of Monte-Carlo simulations. In order to do so, we define an OFDM system with parameters inspired by the 3GPP Long Term Evolution (LTE) 5 MHz downlink physical layer parameters [2]. The system operates with an FFT size of 512, with 300 active subcarriers, and a frequency spacing of 15 KHz between them. Pilot subcarriers are transmitted in every OFDM symbol, with a frequency spacing of 6 subcarriers (i.e. 300 KHz) between them. Both the desired and interfering signals have their pilots in the same subcarriers, and perfect synchronization between the transmitters is assumed. Hence, pilots of both transmitted signals overlap in frequency. The pilot sequences are made of random independent and uniformly distributed QPSK symbols. The power of the interfering signal is equal to that of the desired signal, and 16QAM modulation is employed for the data symbols. A convolutional code is used for channel coding, with BCJR [14] decoding at the receiver.

Two different channel models are considered, namely the ITU Indoor Office A channel [15] and the COST 259 Typical Urban channel [16]. The former channel exhibits a low frequency selectivity, with a coherence bandwidth of about 3.2 MHz, while the latter has much less coherence bandwidth of around 467 KHz. Block fading is assumed, with a static channel response over the duration of an OFDM symbol and independent realizations between consecutive OFDM symbols. The same channel profile is assumed for all wireless links (desired and interfering).

In Fig. 3, the mean-squared error (MSE) of the channel estimates of the desired and interfering channel versus the number of iterations of the estimator are shown for the two considered channels. The MSE of the LMMSE estimator is also depicted for comparison's sake. The signal-to-noise ratio (SNR), which is calculated as the ratio between the desired signal power and the noise power for each antenna branch, is fixed to 25 dB. It is observed that the iterative process
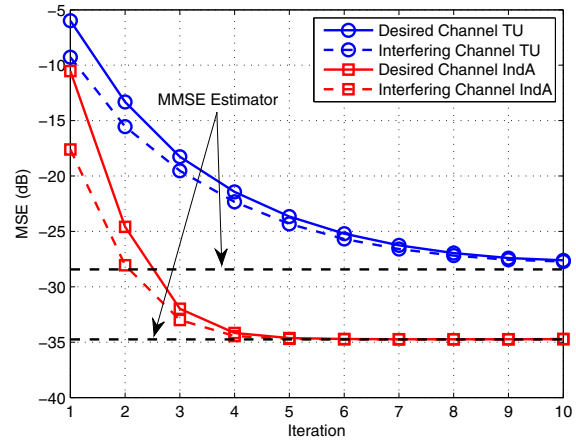

Fig. 3. MSE of the channel estimates versus the number of iterations of the channel estimator at a fixed SNR of 25 dB.
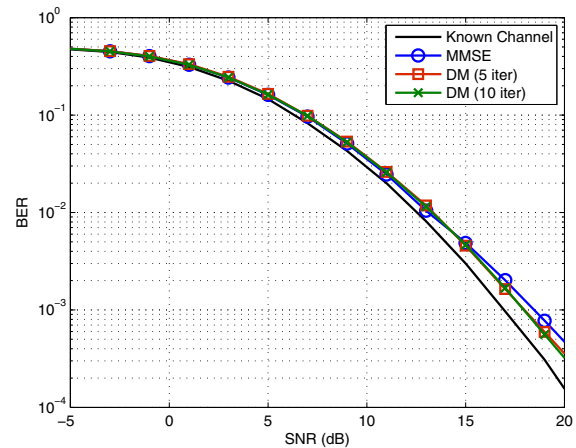

Fig. 4. BER performance for an Indoor A Channel.

improves greatly the quality of the estimates, due to the effectiveness of the interference cancellation and the updating of the noise covariance matrix, which accounts for both the AWGN and the residual interference. A lower MSE (about a 7 dB difference) is achieved in the Indoor A channel. This is a consequence of the lower frequency selectivity, a well-known result from LMMSE channel estimation. It is also noted that the convergence rate of the algorithm depends on the frequency selectivity of the channel as well: while 5 iterations are enough to achieve convergence in the Indoor Channel, around 10 iterations are needed in the Typical Urban Channel. As the results show, the DM channel estimator performance converges to the LMMSE estimator with sufficient number of iterations, and the number of iterations required for convergence depends on the frequency selectivity of the channel.

The receiver's performance is evaluated in terms of bit-error-rate (BER) in Fig. 4 for the Indoor Office A Channel and in Fig. 5 for the Typical Urban Channel. Results are shown for the case where the estimators use 5 and 10 iterations. As a
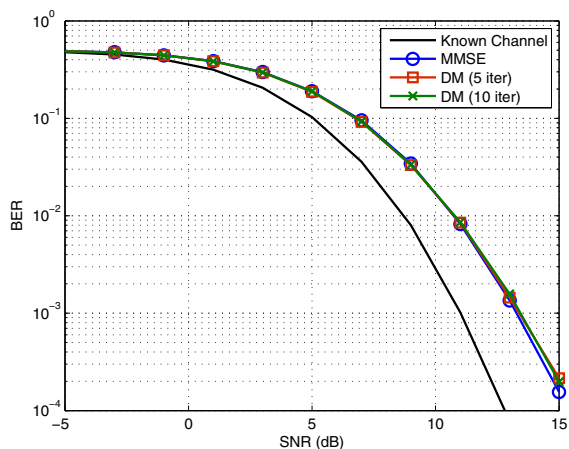
Fig. 5. BER performance for a Typical Urban Channel.

reference, the BER of the receiver with perfect knowledge of the channel is also depicted, as well as the BER of a receiver using the LMMSE estimator. In the Indoor Channel, the DM and LMMSE estimators exhibit the same performance. When compared to a receiver with perfect knowledge of the channel, a very small degradation in the range of 1 dB is observed in the high SNR range. Furthermore, as commented above, the performance of the estimator does not significantly improve after 5 iterations, with only a very marginal gain after 10 iterations. In the Typical Urban Channel, a larger performance deviation from the perfect channel knowledge results is observed. The degradation ranges from 1.7 dB to 2.4 dB at BER of 10% and 0.1% respectively. However, the degradation in the high SNR range is relatively small considering a scenario with such a severe interference. Again, the performance of the DM estimator is very close to that of the LMMSE estimator, and only a very slight gain is observed when the number of iterations of the algorithm is increased from 5 to 10.

V. Conclusion

In this paper, we have presented a novel approach for channel estimation in OFDM systems with synchronized co-channel interferers and overlapped pilot symbols. Based on the KL-divergence minimization principle, an iterative algorithm for estimation of the channel gains based on the signal observed at pilot locations has been derived. The resulting algorithm combines LMMSE channel estimation with successive interference cancellation and estimation of the noise and residual interference power. The effectiveness of the proposed estimator is assessed by Monte-Carlo simulations. The results show that our algorithm performs as good as the full LMMSE channel estimator, with the advantage of avoiding the cumbersome matrix inversion in the latter. An overall receiver performance very close to that of a receiver with perfect knowledge of the channel coefficients is attained, especially in channels with low frequency selectivity.

To conclude, it is worth remarking that although the estimator has been presented and evaluated for an OFDM system with synchronized co-channel interference, application to other scenarios could be very advantageous. For instance, our estimator would allow to reduce the pilot overhead in a MIMO-OFDM system by placing the pilot sequences of all transmit antennas in the same time-frequency locations, instead of having specific locations reserved for each of the antennas as it is common in current wireless standards, e.g. LTE.

References

[1] *IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1*, IEEE Std. 802.16e-2005, 2006.
[2] "Evolved Universal Terrestrial Radio Access (E-UTRA); LTE Physical Layer - General Description (release 8)," 3rd Generation Partnership Project, Tech. Rep. TS 36.211, V8.1.0, Nov. 2007.
[3] L. Hanzo, M. Munster, B. J. Choi, and T. Keller, *OFDM and MC-CDMA for Broadband Multi-User Communications, WLANs and Broadcasting*. West Sussex, England: John Wiley & Sons, 2003.
[4] M. Munster, T. Keller, and L. Hanzo, "Co-channel interference suppression assisted adaptive OFDM in interference limited environments," *Vehicular Technology Conference, 1999. VTC 1999 - Fall. IEEE VTS 50th*, vol. 1, pp. 284–288 vol.1, 1999.
[5] J. Li, K. Ben Letaief, and Z. Cao, "Co-channel interference cancellation for space-time coded OFDM systems," *Wireless Communications, IEEE Transactions on*, vol. 2, no. 1, pp. 41–49, Jan. 2003.
[6] C.-S. Ni and K.-C. Chen, "Cochannel interference suppression for coded OFDM systems over frequency-selective slowly fading channels," *Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th*, vol. 1, pp. 679–683 Vol. 1, Sept. 2004.
[7] Y. Li and N. Sollenberger, "Adaptive antenna arrays for OFDM systems with cochannel interference," *Communications, IEEE Transactions on*, vol. 47, no. 2, pp. 217–229, Feb 1999.
[8] J.-J. van de Beek, O. Edfors, M. Sandell, S. Wilson, and P. Borjesson, "On channel estimation in ofdm systems," *Vehicular Technology Conference, 1995 IEEE 45th*, vol. 2, pp. 815–819 vol.2, Jul 1995.
[9] B. Hu, I. Land, L. Rasmussen, R. Piton, and B. Fleury, "A divergence minimization approach to joint multiuser decoding for coded CDMA," *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 3, pp. 432–445, April 2008.
[10] M. Raghavendra, S. Bhashyam, and K. Giridhar, "Parametric channel estimation in reuse-1 OFDM systems," *Communications, 2007. ICC '07. IEEE International Conference on*, pp. 2999–3004, June 2007.
[11] T. Cover and J. Thomas, *Elements of information theory*. John Wiley & Sons, 1991.
[12] P. Hoeher, S. Kaiser, and P. Robertson, "Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering," *Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on*, vol. 3, pp. 1845–1848 vol.3, Apr 1997.
[13] A. Gupta and D. Nagar, *Matrix Variate Distributions*. Chapman & Hall/CRC, 2000.
[14] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *Information Theory, IEEE Transactions on*, vol. 20, no. 2, pp. 284–287, Mar 1974.
[15] "Guidelines for evaluation of radio transmission technologies for IMT-2000," ITU, Tech. Rep. Recommendation ITU-R M.1225, 1997.
[16] "Deployment aspects," 3rd Generation Partnership Project, Tech. Rep. TS 25.943, Jun. 2002.