Joint Carrier Frequency Offset and Fast Time-varying Channel Estimation for MIMO-OFDM Systems

Eric Pierre Simon, Laurent Ros, Hussein Hijazi, Jin Fang, Davy Gaillot, Marion Berbineau

Joint Carrier Frequency Offset and Fast Time-varying Channel Estimation for MIMO-OFDM Systems

Eric Pierre Simon[1], Laurent Ros[2], Hussein Hijazi[1], Jin Fang[1], Davy Paul Gaillot[1], Marion Berbineau[3]

[1] IEMN/TELICE laboratory, University of Lille - FRANCE
[2] GIPSA-lab, Department Image Signal, BP 46 - 38402 Saint Martin d'Hères - FRANCE
[3] Universite Lille Nord de France, INRETS, LEOST, F-59650 Villeneuve dAscq

e-mail: eric.simon@univ-lille1.fr, hussein.hijazi@hotmail.fr, laurent.ros@gipsa-lab.grenoble-inp.fr, jin.fang@ed.univ-lille1.fr, davy.gaillot@univ-lille1.fr, marion.berbineau@inrets.fr

***Abstract*—[1] In this paper, a novel pilot-aided algorithm is developed for MIMO-OFDM systems operating in fast time-varying environment. The algorithm has been designed to work both with parametric L-path channel model (with known path delays) and equivalent discrete-time channel model to jointly estimate the multi-path Rayleigh channel complex amplitudes (CA) and Carrier Frequency Offset (CFO). Each CA time-variation within one OFDM symbol is approximated by a Basis Expansion Model (BEM) representation. An Auto-Regressive (AR) model is built for the parameters to be estimated. The algorithm performs estimation using Extended Kalman Filtering. The channel matrix is thus easily computed and the data symbol is estimated without Inter-sub-Carrier-Interference (ICI) when the channel matrix is QR-decomposed. It is shown that our algorithm is far more robust to high speed than the conventional algorithm, and the performance approaches that of the ideal case for which the channel response and CFO are known.**

I. INTRODUCTION

Multiple-Input-Multiple-Output (MIMO) antennas with Orthogonal Frequency Division Multiplexing (OFDM) provide high data rates and are robust to multi-path delay in wireless communications. Channel parameters are required for diversity combining, coherent detection and decoding. Therefore, channel estimation is essential to design MIMO-OFDM systems. For MIMO-OFDM systems, most of the channel estimation schemes have focused on pilot-assisted approaches [1][2][3], based on a quasi-static fading model that allows the channel to be invariant within a MIMO-OFDM block. However, in fast-fading channels, the time-variation of the channel within a MIMO-OFDM block results in the loss of subcarrrier orthogonality, and consequently intercarrier interference (ICI) occurs [4][5]. Therefore, the channel time-variation within a block must be considered to support high-speed mobile channels.

On the other hand, similarly to the single-input single-output (SISO) OFDM, one of the disadvantages of MIMO-OFDM lies in its sensitivity to carrier frequency offset (CFO) due to carrier frequency mismatches between transmitter and receiver oscillators. As for the Doppler shift, the CFO produces ICI and attenuates the desired signal. These effects reduce the effective signal-to-noise ratio (SNR) in OFDM reception such that the system performance is degraded [6] [7]. Most of the reported works consider that all the paths present identical Doppler shifts. Hence, they group together the Doppler shift and CFO due to oscillator mismatch to obtain a single offset parameter [8] for each channel branch. However, this model is not sufficiently accurate since separate offset parameters are required for each propagation path given that the Doppler shift depends on the angle of arrival, which is particular to each path. Recently, it has been proposed to directly track channel paths to take into account separate Doppler shifts for each path ([9][10] for SISO and [11] for MIMO). Those works estimate the equivalent discrete-time channel taps ([10]) or the real path Complex Amplitudes (CA) ([9][11]) which are both modeled by a basis expansion model (BEM). The BEM methods include Karhunen-Loeve BEM (KL-BEM), prolate spheroidal BEM (PS-BEM), complex exponential BEM (CE-BEM) and polynomial BEM (P-BEM).

However the CFO due to the mismatch between transmitter and receiver oscillators is not taken into account in those algorithms. The idea of joint channel and CFO estimation has been initially proposed for SISO-OFDM systems in [12] and then extended to MIMO-OFDM systems [13]. The authors proposed an algorithm based on Extended Kalman Filtering (EKF) and on equivalent discrete-time channel model. But the fast time-variation of the channel was not taken into account.

In this paper, we propose a complete algorithm capable of jointly estimating the CFO and the path CA, by taking into account the fast variation of each path CA in MIMO environment.

Generally, it is preferable to directly estimate the physical channel parameters [14] [9] [11] instead of the equivalent discrete-time channel taps [10]. Indeed, as the channel delay spread increases, the number of channel taps also increases and a large number of BEM coefficients have to be estimated. This requires more pilot symbols. Hence, using a parametric channel model rather than an equivalent discrete channel model enables to reduce the signal subspace dimension [14]. Additionally, estimating the physical propagation parameters means estimating path delays and path CA. Note that in Radio-

Frequency transmissions, the delays are quasi-invariant over several MIMO-OFDM blocks [15] [4] (whereas the CA may change significantly, even within one MIMO-OFDM block). In this work, the delays are assumed perfectly estimated and quasi-invariant. It should be noted that an initial, and generally accurate estimation of the number of paths and delays can be obtained by using the MDL (minimum description length) and ESPRIT (estimation of signal parameters by rotational invariance techniques) methods [14][9].

This paper is organized as follows: Section II introduces the MIMO-OFDM system and the BEM modeling. Section III describes the state model and the Extended Kalman Filter. Section IV covers the algorithm for joint channel and CFO estimation together with data recovery. Section V presents the simulations results which validate our technique. Finally, our conclusions are presented in Section VI.

The notations adopted are as follows: Upper (lower) bold face letters denote matrices (column vectors). $[\mathbf{x}]_k$ denotes the kth element of the vector $\mathbf{x}$, and $[\mathbf{X}]_{k,m}$ denotes the $[k,m]$th element of the matrix $\mathbf{X}$. The row and column indices start from 0 (and not 1). We will use the matlab notation $\mathbf{X}_{[k_1:k_2,m_1:m_2]}$ to extract a submatrix within $\mathbf{X}$ from row k_1 to row k_2 and from column m_1 to column m_2. $\mathbf{I}_N$ is a $N \times N$ identity matrix and $\mathbf{0}_N$ is a $N \times N$ matrix of zeros. diag$\{\mathbf{x}\}$ is a diagonal matrix with $\mathbf{x}$ on its main diagonal and blkdiag$\{\mathbf{X}, \mathbf{Y}\}$ is a block diagonal matrix with the matrices $\mathbf{X}$ and $\mathbf{Y}$ on its main diagonal. The superscripts $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$ stand respectively for transpose, conjugate and Hermitian operators. Tr$(\cdot)$ and E$[\cdot]$ are respectively the determinant and expectation operations. $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind. $\nabla_{\mathbf{x}}$ represents the first-order partial derivative operator $i.e.$, $\nabla_{\mathbf{x}} = [\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_N}]^T$.

II. MIMO-OFDM System and Channel Models

A. MIMO-OFDM System Model

Consider a MIMO-OFDM system with N_T transmitter antennas, N_R receiver antennas, N sub-carriers, and a cyclic prefix length N_g. The duration of a MIMO-OFDM block is $T = N_b T_s$, where T_s is the sampling time and $N_b = N + N_g$. Let $\mathbf{x}_n \stackrel{\text{def}}{=} \left[\mathbf{x}_n^{(1)^T}, \mathbf{x}_n^{(2)^T}, ..., \mathbf{x}_n^{(N_T)^T}\right]^T$ be the nth transmitted signal block, where $\mathbf{x}_n^{(t)} \stackrel{\text{def}}{=} \left[x_n^{(t)}[-\frac{N}{2}], x_n^{(t)}[-\frac{N}{2}+1], ..., x_n^{(t)}[\frac{N}{2}-1]\right]^T$ is the nth signal vector transmitted by the tth transmitter antenna and the data symbol $\{x_n^{(t)}[k]\}$ is transmitted on the kth sub-carrier. The data symbol are normalized ($i.e.$, $\mathrm{E}\left[x_n^{(t)}[k] x_n^{(t)*}[k]\right] = 1$). The frequency mismatch between the oscillators used in the radio transmitters and receivers causes a CFO. In multi-antenna systems, each transmitter and receiver typically requires its own Radio Frequency - Intermediate Frequency (RF-IF) chain. Consequently, each transmitter-receiver pair has its own mismatch parameter, yielding separate CFO. In a $N_T \times N_R$ MIMO system this leads to $N_T N_R$ different CFO. However, if transmitter or receiver antennas share RF-IF chains, fewer different CFO occur. The system model describes the general case where it is necessary to compensate for $N_T N_R$ CFO. Assume that the MIMO channel branch between the tth transmit antenna and the rth receive antenna (called (r,t) branch from now on) experiences a normalized frequency shift $\nu^{(r,t)} = \Delta F^{(r,t)} N T_s$, where $\Delta F^{(r,t)}$ is the absolute CFO. All the normalized CFO can be stacked in vector form:

$$\boldsymbol{\nu} \stackrel{\text{def}}{=} \Big[\nu^{(1,1)}, \ldots, \nu^{(1,N_T)}, \ldots, \nu^{(r,1)}, \ldots, \nu^{(r,N_T)}, \ldots, \nu^{(N_R,N_T)}\Big]^T \quad (1)$$

After transmission over a multi-path Rayleigh channel, the nth received MIMO-OFDM block $\mathbf{y}_n \stackrel{\text{def}}{=} \left[\mathbf{y}_n^{(1)^T}, \mathbf{y}_n^{(2)^T}, ..., \mathbf{y}_n^{(N_R)^T}\right]^T$, where $\mathbf{y}_n^{(r)} \stackrel{\text{def}}{=} \left[y_n^{(r)}[-\frac{N}{2}], y_n^{(r)}[-\frac{N}{2}+1], ..., y_n^{(r)}[\frac{N}{2}-1]\right]^T$ is the nth received OFDM symbol by the rth receiver antenna, is given in the frequency domain by [4] [10]:

$$\mathbf{y}_n = \mathbf{H}_n \, \mathbf{x}_n + \mathbf{w}_n \quad (2)$$

where $\mathbf{w}_n \stackrel{\text{def}}{=} \left[\mathbf{w}_n^{(1)^T}, \mathbf{w}_n^{(2)^T}, ..., \mathbf{w}_n^{(N_R)^T}\right]^T$ with $\mathbf{w}_n^{(r)} \stackrel{\text{def}}{=} \left[w_n^{(r)}[-\frac{N}{2}], w_n^{(r)}[-\frac{N}{2}+1], ..., w_n^{(r)}[\frac{N}{2}-1]\right]^T$ a white complex Gaussian noise vector of covariance matrix $N_T \sigma^2 \mathbf{I}_N$. The matrix $\mathbf{H}_n$ is a $N_R N \times N_T N$ MIMO channel matrix given by:

$$\mathbf{H}_n \stackrel{\text{def}}{=} \begin{bmatrix} \mathbf{H}_n^{(1,1)} & \cdots & \mathbf{H}_n^{(1,N_T)} \\ \vdots & \ddots & \vdots \\ \mathbf{H}_n^{(N_R,1)} & \cdots & \mathbf{H}_n^{(N_R,N_T)} \end{bmatrix} \quad (3)$$

where $\mathbf{H}_n^{(r,t)}$ is the (r,t) branch channel matrix. The elements of channel matrix $\mathbf{H}_n^{(r,t)}$ can be written in terms of equivalent channel taps [5] $\left\{g_{l,n}^{(r,t)}(qT_s) = g_l^{(r,t)}(nT + qT_s)\right\}$ or in terms of physical channel parameters [9] ($i.e.$ delays $\left\{\tau_l^{(r,t)}\right\}$ and CA $\left\{\alpha_{l,n}^{(r,t)}(qT_s) = \alpha_l^{(r,t)}(nT + qT_s)\right\}$), yielding Eq. (4) and (5), respectively.

$L'^{(r,t)} < N_g$ is the number of channel taps and $L^{(r,t)}$ the number of paths for the (r,t) branch. The delays are normalized by T_s and not necessarily integers ($\tau_l^{(r,t)} < N_g$). The $L^{(r,t)}$ elements of $\left\{\alpha_{l,n}^{(r,t)}(qT_s)\right\}$ are uncorrelated. However, the $L'^{(r,t)}$ elements of $\left\{g_{l,n}^{(r,t)}(qT_s)\right\}$ are correlated, unless the delays are multiple of T_s as is commonly assumed in the literature. They are wide-sense stationary (WSS), narrow-band zero-mean complex Gaussian processes of variances ${\sigma_{g_l}^{(r,t)}}^2$ and ${\sigma_{\alpha_l}^{(r,t)}}^2$, with the so-called Jakes' power spectrum of maximum Doppler frequency f_d [16]. The average energy of each (r,t) branch is normalized to one, $i.e.$, $\sum_{l=0}^{L'^{(r,t)}-1} {\sigma_{g_l}^{(r,t)}}^2 = 1$ and $\sum_{l=0}^{L^{(r,t)}-1} {\sigma_{\alpha_l}^{(r,t)}}^2 = 1$.

In the next sections, we present the derivations for the second approach (physical channel). The results of the first approach (channel taps) can be deduced by replacing $L^{(r,t)}$ by $L'^{(r,t)}$ and the set of delays $\left\{\tau_l^{(r,t)}\right\}$ by $\left\{l,\ l = 0 : L'^{(r,t)} - 1\right\}$.

B. BEM Channel Model

Let $L \stackrel{\text{def}}{=} \sum_{r=1}^{N_R} \sum_{t=1}^{N_T} L^{(r,t)}$ be the total number of paths for the MIMO channel. There are N_b samples to be estimated

$$[\mathbf{H}_n^{(r,t)}]_{k,m} = \frac{1}{N}\sum_{l=0}^{L'^{(r,t)}-1}\Big[e^{-j2\pi(\frac{m}{N}-\frac{1}{2})\cdot l}\sum_{q=0}^{N-1}e^{j2\pi\frac{\nu^{(r,t)}q}{N}}g_{l,n}^{(r,t)}(qT_s)e^{j2\pi\frac{m-k}{N}q}\Big] \quad (4)$$

$$= \frac{1}{N}\sum_{l=0}^{L^{(r,t)}-1}\Big[e^{-j2\pi(\frac{m}{N}-\frac{1}{2})\tau_l^{(r,t)}}\sum_{q=0}^{N-1}e^{j2\pi\frac{\nu^{(r,t)}q}{N}}\alpha_{l,n}^{(r,t)}(qT_s)e^{j2\pi\frac{m-k}{N}q}\Big] \quad (5)$$

for each path CA due to the fast time-variation of the channel, yielding a total of LN_b samples for the whole MIMO channel. In order to reduce the number of parameters to be estimated, we resort to the Basis Expansion Model (BEM). In this section, our aim is to accurately model the time-variation of $\alpha_{l,n}^{(r,t)}(qT_s)$ from $q=-N_g$ to $N-1$ by using a BEM.

Suppose $\boldsymbol{\alpha}_{l,n}^{(r,t)}$ represents an $N_b\times 1$ vector that collects the time-variation of the lth path of the (r,t) branch within the nth MIMO-OFDM block:

$$\boldsymbol{\alpha}_{l,n}^{(r,t)} \stackrel{\text{def}}{=} \big[\alpha_{l,n}^{(r,t)}(-N_gT_s),...,\alpha_{l,n}^{(r,t)}\big((N-1)T_s\big)\big]^T \quad (6)$$

Then, each $\boldsymbol{\alpha}_{l,n}^{(r,t)}$ can be expressed in terms of a BEM as:

$$\boldsymbol{\alpha}_{l,n}^{(r,t)} = \boldsymbol{\alpha}_{\mathbf{BEM}_{l,n}}^{(r,t)} + \boldsymbol{\xi}_{l,n}^{(r,t)} = \mathbf{B}\,\mathbf{c}_{l,n}^{(r,t)} + \boldsymbol{\xi}_{l,n}^{(r,t)} \quad (7)$$

where the $N_b\times N_c$ matrix $\mathbf{B}$ is defined as: $\mathbf{B}\stackrel{\text{def}}{=}[\mathbf{b}_0,...,\mathbf{b}_{N_c-1}]$. The $N_b\times 1$ vector $\mathbf{b}_d$ is termed as the dth expansion basis. $\mathbf{c}_{l,n}^{(r,t)}\stackrel{\text{def}}{=}\big[c_{l,n}^{(r,t)}[0],...,c_{l,n}^{(r,t)}[N_c-1]\big]^T$ represents the N_c BEM coefficients and $\boldsymbol{\xi}_{l,n}^{(r,t)}$ represents the corresponding BEM modeling error, which is assumed to be minimized in the MSE sense [17]. Under this criterion, the optimal BEM coefficients and the corresponding model error are given by:

$$\mathbf{c}_{l,n}^{(r,t)} = \big(\mathbf{B}^H\mathbf{B}\big)^{-1}\mathbf{B}^H\boldsymbol{\alpha}_{l,n}^{(r,t)} \quad (8)$$

$$\boldsymbol{\xi}_{l,n}^{(r,t)} = (\mathbf{I}_{N_b}-\mathbf{S})\boldsymbol{\alpha}_{l,n}^{(r,t)} \quad (9)$$

where $\mathbf{S}=\mathbf{B}\big(\mathbf{B}^H\mathbf{B}\big)^{-1}\mathbf{B}^H$ is a $N_b\times N_b$ matrix. Then, the MMSE approximation for all BEM with N_c coefficients is given by:

$$\text{MMSE}_l^{(r,t)} \stackrel{\text{def}}{=} \frac{1}{N_b}\text{E}\Big[\boldsymbol{\xi}_{l,n}^{(r,t)}\boldsymbol{\xi}_{l,n}^{(r,t)\,H}\Big] \quad (10)$$

$$= \frac{1}{N_b}\text{Tr}\Big(\big(\mathbf{I}_{N_b}-\mathbf{S}\big)\mathbf{R}_{\boldsymbol{\alpha}_l}^{(r,t)}[0]\big(\mathbf{I}_{N_b}-\mathbf{S}^H\big)\Big) \quad (11)$$

where $\mathbf{R}_{\boldsymbol{\alpha}_l}^{(r,t)}[s]\stackrel{\text{def}}{=}\text{E}\left[\boldsymbol{\alpha}_{l,n}^{(r,t)}\boldsymbol{\alpha}_{l,n-s}^{(r,t)\,H}\right]$ is the $N_b\times N_b$ correlation matrix of $\boldsymbol{\alpha}_{l,n}^{(r,t)}$ with elements given by:

$$[\mathbf{R}_{\boldsymbol{\alpha}_l}^{(r,t)}[s]]_{k,m} = {\sigma_{\alpha_l}^{(r,t)}}^2 J_0\Big(2\pi f_dT_s(k-m+sN_b)\Big) \quad (12)$$

Various traditional BEM designs have been reported to model the channel time-variations, e.g., the Complex Exponential BEM (CE-BEM) $[\mathbf{B}]_{k,m}=e^{j2\pi(\frac{k-N_g}{N_b})(m-\frac{N_c-1}{2})}$ which leads to a strictly banded frequency-domain matrix [18], the Generalized CE-BEM (GCE-BEM) $[\mathbf{B}]_{k,m}=e^{j2\pi(\frac{k-N_g}{aN_b})(m-\frac{N_c-1}{2})}$ with $1<a\leq\frac{N_c-1}{2f_dT}$ which is a set of oversampled complex exponentials [17], the Polynomial BEM (P-BEM) $[\mathbf{B}]_{k,m}=(k-N_g)^m$ [9] and the Discrete Karhuen-Loeve BEM (DKL-BEM) which employs basis sequences that correspond to the most significant eigenvectors of the autocorrelation matrix $\mathbf{R}_{\boldsymbol{\alpha}_l}^{(r,t)}[0]$ [19]. From now on, we can describe the MIMO-OFDM system model derived previously in terms of BEM. Substituting (7) in (2) and neglecting the BEM model error, one obtains after some algebra:

$$\mathbf{y}_n = \boldsymbol{\mathcal{K}}_n(\boldsymbol{\nu})\cdot\mathbf{c}_n + \mathbf{w}_n \quad (13)$$

where the $LN_c\times 1$ vector $\mathbf{c}_n$ and the $N_RN\times LN_c$ matrix $\boldsymbol{\mathcal{K}}_n(\boldsymbol{\nu})$ are given by:

$$\mathbf{c}_n \stackrel{\text{def}}{=} \Big[\mathbf{c}_n^{(1,1)^T},...,\mathbf{c}_n^{(1,N_T)^T},...,\mathbf{c}_n^{(N_R,N_T)^T}\Big]^T \quad (14)$$

$$\mathbf{c}_n^{(r,t)} \stackrel{\text{def}}{=} \Big[\mathbf{c}_{0,n}^{(r,t)^T},...,\mathbf{c}_{L^{(r,t)}-1,n}^{(r,t)^T}\Big]^T$$

$$\boldsymbol{\mathcal{K}}_n(\boldsymbol{\nu}) \stackrel{\text{def}}{=} \text{blkdiag}\Big\{\boldsymbol{\mathcal{K}}_n^{(1)}(\boldsymbol{\nu}^{(1)}),...,\boldsymbol{\mathcal{K}}_n^{(N_R)}(\boldsymbol{\nu}^{(N_R)})\Big\}$$

$$\boldsymbol{\mathcal{K}}_n^{(r)}(\boldsymbol{\nu}^{(r)}) \stackrel{\text{def}}{=} \Big[\boldsymbol{\mathcal{K}}_n^{(r,1)}(\nu^{(r,1)}),...,\boldsymbol{\mathcal{K}}_n^{(r,N_T)}(\nu^{(r,N_T)})\Big]$$

$$\boldsymbol{\mathcal{K}}_n^{(r,t)}(\nu^{(r,t)}) \stackrel{\text{def}}{=} \frac{1}{N}\Big[\mathbf{Z}_{0,n}^{(r,t)}(\nu^{(r,t)}),...,\mathbf{Z}_{L^{(r,t)}-1,n}^{(r,t)}(\nu^{(r,t)})\Big]$$

$$\mathbf{Z}_{l,n}^{(r,t)}(\nu^{(r,t)}) \stackrel{\text{def}}{=} \Big[\mathbf{M}_0^{(r,t)}(\nu^{(r,t)})\ \text{diag}\{\mathbf{x}_n^{(t)}\}\ \mathbf{f}_l^{(r,t)},...,$$
$$\mathbf{M}_{N_c-1}^{(r,t)}(\nu^{(r,t)})\ \text{diag}\{\mathbf{x}_n^{(t)}\}\ \mathbf{f}_l^{(r,t)}\Big]$$

where $\boldsymbol{\nu}^{(r)}\stackrel{\text{def}}{=}\big[\nu^{(r,1)},\ldots,\nu^{(r,N_T)}\big]^T$. Vector $\mathbf{f}_l^{(r,t)}$ is the lth column of the $N\times L^{(r,t)}$ Fourier matrix $\mathbf{F}^{(r,t)}$ whose elements are given by:

$$[\mathbf{F}^{(r,t)}]_{k,l} = e^{-j2\pi(\frac{k}{N}-\frac{1}{2})\tau_l^{(r,t)}}\ , \quad (15)$$

and $\mathbf{M}_d^{(r,t)}$ is a $N\times N$ matrix whose elements are given by:

$$\Big[\mathbf{M}_d^{(r,t)}(\nu^{(r,t)})\Big]_{k,m} = \sum_{q=0}^{N-1}e^{j2\pi\frac{\nu^{(r,t)}q}{N}}[\mathbf{B}]_{q+N_g,d}\ e^{j2\pi\frac{m-k}{N}q}\ . \quad (16)$$

Moreover, the channel matrix of the (r,t) branch can be easily computed by using the BEM coefficients [4]:

$$\mathbf{H}_n^{(r,t)} = \sum_{d=0}^{N_c-1}\mathbf{M}_d^{(r,t)}(\nu^{(r,t)})\text{diag}\{\mathbf{F}^{(r,t)}\boldsymbol{\chi}_{d,n}^{(r,t)}\} \quad (17)$$

where $\boldsymbol{\chi}_{d,n}^{(r,t)}\stackrel{\text{def}}{=}\big[c_{0,n}^{(r,t)}[d],...,c_{L^{(r,t)}-1,n}^{(r,t)}[d]\big]^T$. Eq. (17) will be used in the following to obtain an estimated channel matrix from the estimated CFO and BEM coefficients.

III. AR Model and Extended Kalman Filter

A. The AR Model for $\mathbf{c}_n$

The optimal BEM coefficients $\mathbf{c}_{l,n}^{(r,t)}$ are correlated complex Gaussian variables with zero-means and correlation matrix

given by:

$$\begin{aligned}\mathbf{R}_{\mathbf{c}_l}^{(r,t)}[s] &\stackrel{\text{def}}{=} \mathrm{E}[\mathbf{c}_{l,n}^{(r,t)}{\mathbf{c}_{l,n-s}^{(r,t)}}^H] \\ &= \left(\mathbf{B}^H\mathbf{B}\right)^{-1}\mathbf{B}^H\mathbf{R}_{\boldsymbol{\alpha}_l}^{(r,t)}[s]\mathbf{B}\left(\mathbf{B}^H\mathbf{B}\right)^{-1} \quad (18)\end{aligned}$$

Since the coefficients $\mathbf{c}_{l,n}^{(r,t)}$ are correlated Gaussian variables, their dynamics can be correctly modeled by an auto-regressive (AR) process [20] [21] [9] . A complex AR process of order p can be generated such that:

$$\mathbf{c}_{l,n}^{(r,t)} = \sum_{i=1}^{p}\mathbf{A}^{(i)}\mathbf{c}_{l,n-i}^{(r,t)} + \mathbf{u}_{l,n}^{(r,t)} \quad (19)$$

where $\mathbf{A}^{(1)},...,\mathbf{A}^{(p)}$ are $N_c\times N_c$ matrices and $\mathbf{u}_{l,n}^{(r,t)}$ is a $N_c\times 1$ complex Gaussian vector with covariance matrix $\mathbf{U}_l^{(r,t)}$. From [9], it is sufficient to choose $p=1$ to correctly model the path CA. The matrices $\mathbf{A}^{(1)}=\mathbf{A}$ and $\mathbf{U}_l^{(r,t)}$ are the AR model parameters obtained by solving the set of Yule-Walker equations:

$$\mathbf{A} = \mathbf{R}_{\mathbf{c}_l}^{(r,t)}[1]\left(\mathbf{R}_{\mathbf{c}_l}^{(r,t)}[0]\right)^{-1} \quad (20)$$

$$\mathbf{U}_l^{(r,t)} = \mathbf{R}_{\mathbf{c}_l}^{(r,t)}[0] + \mathbf{A}\mathbf{R}_{\mathbf{c}_l}^{(r,t)}[-1] \quad (21)$$

Using (19), we obtain the first-order AR approximation for the dynamics of $\mathbf{c}_n$:

$$\mathbf{c}_n = \boldsymbol{\mathcal{A}}_{\mathbf{c}}\cdot\mathbf{c}_{n-1} + \mathbf{u}_{\mathbf{c}n} \quad (22)$$

where $\boldsymbol{\mathcal{A}}_{\mathbf{c}} \stackrel{\text{def}}{=} \text{blkdiag}\{\mathbf{A},...,\mathbf{A}\}$ is a $LN_c\times LN_c$ matrix and $\mathbf{u}_{\mathbf{c}n} \stackrel{\text{def}}{=} \left[\mathbf{u}_{0,n}^{(1,1)^T},...,\mathbf{u}_{L^{(N_R,N_T)}-1,n}^{(N_R,N_T)^T}\right]^T$ is a $LN_c\times 1$ zero-mean complex Gaussian vector with covariance matrix $\mathbf{U}_{\mathbf{c}} \stackrel{\text{def}}{=} \text{blkdiag}\left\{\mathbf{U}_0^{(1,1)},...,\mathbf{U}_{L^{(N_R,N_T)}-1}^{(N_R,N_T)}\right\}$.

B. The AR Model for $\boldsymbol{\nu}_n$

Let us write the global first-order AR model for $\boldsymbol{\nu}_n$ as follows:

$$\boldsymbol{\nu}_n = \boldsymbol{\mathcal{A}}_{\boldsymbol{\nu}}\cdot\boldsymbol{\nu}_{n-1} + \mathbf{u}_{\boldsymbol{\nu}n} \quad (23)$$

where the state transition matrix is of size $N_RN_T\times N_RN_T$. Since the CFO can be assumed as constant during the observation interval, $\boldsymbol{\mathcal{A}}_{\boldsymbol{\nu}}$ is considered to be close to the identity matrix $\boldsymbol{\mathcal{A}}_{\boldsymbol{\nu}} = a\mathbf{I}_{N_RN_T}$, where a is typically chosen between 0.99 and 0.9999 [22][13]. The $N_RN_T\times 1$ state noise vector $\mathbf{u}_{\boldsymbol{\nu}n}$ is assumed to be zero-mean complex Gaussian. The state noise covariance matrix is $\mathbf{U}_{\boldsymbol{\nu}} = \sigma_{u_\nu}^2\mathbf{I}_{N_RN_T}$ where $\sigma_{u_\nu}^2$ is the variance of the state noise associated with CFO. The value of the state noise variance depends on the parameter a, as explained in Appendix.

C. State equation

Now, let us write the state-variable model. The state vector at time instance n consists of the BEM coefficients $\mathbf{c}_n$ and the vector of CFO $\boldsymbol{\nu}_n$:

$$\boldsymbol{\mu}_n \stackrel{\text{def}}{=} \left[\mathbf{c}_n^T,\ \boldsymbol{\nu}_n^T\right]^T \quad (24)$$

There are LN_c BEM coefficients and N_TN_R CFO values in the state vector of dimension $LN_c+N_TN_R\times 1$. Then, the linear state equation may be written as follows:

$$\boldsymbol{\mu}_n = \boldsymbol{\mathcal{A}}\cdot\boldsymbol{\mu}_{n-1} + \mathbf{u}_n \quad (25)$$

where the state transition matrix is defined as follows:

$$\boldsymbol{\mathcal{A}} \stackrel{\text{def}}{=} \text{blkdiag}\left\{\boldsymbol{\mathcal{A}}_{\mathbf{c}},\ \boldsymbol{\mathcal{A}}_{\boldsymbol{\nu}}\right\} \quad (26)$$

The $LN_c+N_RN_T\times 1$ noise vector is such that $\mathbf{u}_n \stackrel{\text{def}}{=} \left[\mathbf{u}_{\mathbf{c}n}^T,\ \mathbf{u}_{\boldsymbol{\nu}n}^T\right]^T$ with covariance matrix $\mathbf{U} \stackrel{\text{def}}{=} \text{blkdiag}\left\{\mathbf{U}_{\mathbf{c}},\ \mathbf{U}_{\boldsymbol{\nu}}\right\}$.

D. Extended Kalman Filter (EKF)

The measurement equation (13) can be reformulated as:

$$\mathbf{y}_n = \mathbf{g}\left(\boldsymbol{\mu}_n\right) + \mathbf{w}_n \quad (27)$$

where the nonlinear function $\mathbf{g}$ of the state vector $\boldsymbol{\mu}_n$ is defined as $\mathbf{g}\left(\boldsymbol{\mu}_n\right) = \boldsymbol{\mathcal{K}}_n(\boldsymbol{\nu})\cdot\mathbf{c}_n$. Nonlinearity of the measurement equation (27) is caused by CFO. The BEM coefficients are still linearly related to observations. Since the measurement equation is nonlinear, we use the Extended Kalman filter to adaptively track $\boldsymbol{\mu}_n$. Let $\hat{\boldsymbol{\mu}}_{(n|n-1)}$ be our a priori state estimate at step n given knowledge of the process prior to step n, $\hat{\boldsymbol{\mu}}_{(n|n)}$ be our a posteriori state estimate at step n given measurement $\mathbf{y}_n$ and, $\mathbf{P}_{(n|n-1)}$ and $\mathbf{P}_{(n|n)}$ are respectively the a priori and the a posteriori error estimate covariance matrix of size $LN_c+N_RN_T\times LN_c+N_RN_T$. We initialize the EKF with $\hat{\boldsymbol{\mu}}_{(0|0)} = \mathbf{0}_{LN_c+N_RN_T,1}$ and $\mathbf{P}_{(0|0)}$ given by:

$$\begin{aligned}\mathbf{P}_{(0|0)} &= \text{blkdiag}\left\{\mathbf{R}_{\mathbf{c}}[0],\ \sigma_{u_\nu}^2\mathbf{I}_{N_RN_T}\right\} \quad (28)\\ \mathbf{R}_{\mathbf{c}}[s] &= \text{blkdiag}\left\{\mathbf{R}_{\mathbf{c}}^{(1,1)}[s],...,\mathbf{R}_{\mathbf{c}}^{(N_R,N_T)}[s]\right\}\\ \mathbf{R}_{\mathbf{c}}^{(r,t)}[s] &= \text{blkdiag}\left\{\mathbf{R}_{\mathbf{c}_0}^{(r,t)}[s],...,\mathbf{R}_{\mathbf{c}_{L^{(r,t)}-1}}^{(r,t)}[s]\right\}\end{aligned}$$

where $\mathbf{R}_{\mathbf{c}_l}^{(r,t)}[s]$ is the correlation matrix of $\mathbf{c}_{l,n}^{(r,t)}$ defined in (18). To derive the EKF equations, we need to compute the Jacobian matrix $\mathbf{G}_n$ of $\mathbf{g}\left(\boldsymbol{\mu}_n\right)$ with respect to $\boldsymbol{\mu}_n$ and evaluated at $\hat{\boldsymbol{\mu}}_{(n|n-1)}$:

$$\begin{aligned}&\mathbf{G}_n \stackrel{\text{def}}{=} \nabla_{\boldsymbol{\mu}_n}^T\mathbf{g}\left(\boldsymbol{\mu}_n\right)\big|_{\boldsymbol{\mu}_n=\hat{\boldsymbol{\mu}}_{(n|n-1)}} =\\ &\left[\nabla_{\mathbf{c}_n}^T\mathbf{g}\left(\boldsymbol{\mu}_n\right)\big|_{\boldsymbol{\mu}_n=\hat{\boldsymbol{\mu}}_{(n|n-1)}},\ \nabla_{\boldsymbol{\nu}_n}^T\mathbf{g}\left(\boldsymbol{\mu}_n\right)\big|_{\boldsymbol{\mu}_n=\hat{\boldsymbol{\mu}}_{(n|n-1)}}\right] \quad (29)\end{aligned}$$

Let us define $\boldsymbol{\mu}_n^{(r)} \stackrel{\text{def}}{=} \left[\boldsymbol{\mu}_n^{(r,1)^T},\ldots,\boldsymbol{\mu}_n^{(r,N_T)^T}\right]^T$ and $\boldsymbol{\mu}_n^{(r,t)} \stackrel{\text{def}}{=} \left[\mathbf{c}_n^{(r,t)^T}\nu_n^{(r,t)}\right]^T$. After computation, we find:

$$\mathbf{G}_n = \left[\boldsymbol{\mathcal{K}}_n(\boldsymbol{\nu}_n)|_{\boldsymbol{\nu}_n=\hat{\boldsymbol{\nu}}_{(n|n-1)}},\ \boldsymbol{\mathcal{V}}_n(\boldsymbol{\mu}_n)|_{\boldsymbol{\mu}_n=\hat{\boldsymbol{\mu}}_{(n|n-1)}}\right] \quad (30)$$

where

$$\begin{aligned}&\boldsymbol{\mathcal{V}}_n(\boldsymbol{\mu}_n) \stackrel{\text{def}}{=} \text{blkdiag}\left\{\boldsymbol{\mathcal{V}}_n^{(1)}(\boldsymbol{\mu}_n^{(1)}),...,\boldsymbol{\mathcal{V}}_n^{(N_R)}(\boldsymbol{\mu}_n^{(N_R)})\right\}\\ &\boldsymbol{\mathcal{V}}_n^{(r)}(\boldsymbol{\mu}_n^{(r)}) \stackrel{\text{def}}{=} \left[\mathbf{v}^{(r,1)}(\boldsymbol{\mu}_n^{(r,1)}),\ldots,\mathbf{v}^{(r,N_T)}(\boldsymbol{\mu}_n^{(r,N_T)})\right]\\ &\mathbf{v}^{(r,t)}(\boldsymbol{\mu}_n^{(r,t)}) \stackrel{\text{def}}{=} \boldsymbol{\mathcal{K}}_n'^{(r,t)}(\nu_n^{(r,t)})\cdot\mathbf{c}_n^{(r,t)}\\ &\boldsymbol{\mathcal{K}}_n'^{(r,t)}(\nu_n^{(r,t)}) \stackrel{\text{def}}{=} \frac{1}{N}\left[\mathbf{Z}'^{(r,t)}_{0,n}(\nu_n^{(r,t)}),...,\mathbf{Z}'^{(r,t)}_{L-1,n}(\nu_n^{(r,t)})\right]\\ &\mathbf{Z}'^{(r,t)}_{l,n}(\nu_n^{(r,t)}) \stackrel{\text{def}}{=} \Big[\mathbf{M}'_0(\nu_n^{(r,t)})\ \text{diag}\{\mathbf{x}_n^{(t)}\}\ \mathbf{f}_l^{(r,t)},...,\\ &\qquad\mathbf{M}'_{N_c-1}(\nu_n^{(r,t)})\ \text{diag}\{\mathbf{x}_n^{(t)}\}\ \mathbf{f}_l^{(r,t)}\Big]\end{aligned}$$

The elements of the $N \times N$ matrix $\mathbf{M}'_d(\nu)$ are given by:

$$\left[\mathbf{M}'_d(\nu_n^{(r,t)})\right]_{k,m} = \sum_{q=0}^{N-1} j2\pi \frac{q}{N} e^{j2\pi \frac{\nu_n^{(r,t)} q}{N}} [\mathbf{B}]_{q+N_g,d}\ e^{j2\pi \frac{m-k}{N} q} \tag{31}$$

The EKF is a recursive algorithm composed of two stages: Time Update Equations and Measurement Update Equations, defined as follows:

Time Update Equations:

$$\begin{aligned} \hat{\boldsymbol{\mu}}_{(n|n-1)} &= \boldsymbol{\mathcal{A}} \hat{\boldsymbol{\mu}}_{(n-1|n-1)} \\ \mathbf{P}_{(n|n-1)} &= \boldsymbol{\mathcal{A}} \mathbf{P}_{(n-1|n-1)} \boldsymbol{\mathcal{A}}^H + \mathbf{U} \end{aligned} \tag{32}$$

Measurement Update Equations:

$$\begin{aligned} \mathbf{K}_n &= \mathbf{P}_{(n|n-1)} \mathbf{G}_n^H \left(\mathbf{G}_n \mathbf{P}_{(n|n-1)} \mathbf{G}_n^H + N_T.\sigma^2 \mathbf{I}_{N_R N} \right)^{-1} \\ \hat{\boldsymbol{\mu}}_{(n|n)} &= \hat{\boldsymbol{\mu}}_{(n|n-1)} + \mathbf{K}_n \left(\mathbf{y}_n - \mathbf{g}\left(\hat{\boldsymbol{\mu}}_{(n|n-1)} \right) \right) \\ \mathbf{P}_{(n|n)} &= \mathbf{P}_{(n|n-1)} - \mathbf{K}_n \mathbf{G}_n \mathbf{P}_{(n|n-1)} \end{aligned} \tag{33}$$

where $\mathbf{K}_n$ is the Kalman gain. The Time Update Equations are responsible for projecting forward (in time) the current state and error covariance estimates to obtain the a priori estimates for the next time step. The Measurement Update Equations are responsible for the feedback, $i.e.$, for incorporating a new measurement into the a priori estimate to obtain an improved a posteriori estimate. The Time Update Equations can also be thought of as predictor equations, while the Measurement Update Equations can be thought of as corrector equations.

IV. Joint Data Detection and parameter estimation

A. Proposed algorithm

The algorithm uses N_p pilots subcarriers evenly inserted into the N subcarriers. The pilot positions are the same for all the transmitter antennas, yielding the set of pilot indices $\mathcal{P} = \{\, nL_f + (t-1)N,\ n = 0 \dots N_p - 1,\ t = 1 \dots N_T \}$, where L_f is the distance between two adjacent pilots. The data is detected with a QR-equalizer [9] with free Inter-Carrier-Interference (ICI) thanks to a QR-decomposition.

The general principle is as follows : to detect the data symbols $\mathbf{x}_n$, we need to perform an equalization which requires the knowledge of the channel matrix $\mathbf{H}_n$ (see Eq. (2) for the transmission model and Eq. (3) for the definition of the channel matrix). However, the data symbols $\mathbf{x}_n$ are required to estimate the channel matrix. To alleviate this contradiction, a predicted version of the channel matrix $\hat{\mathbf{H}}_{(n|n-1)}$ obtained with $\mathbf{x}_n$ unknown is computed. $\hat{\mathbf{H}}_{(n|n-1)}$ is subsequently updated into $\hat{\mathbf{H}}_{(n|n)}$ through the EKF measurement update equations (33) with the current received OFDM symbol $\mathbf{y}_n$. The current data symbol $\hat{\mathbf{x}}_{(n|n)}$ is finally retrieved from this updated channel matrix $\hat{\mathbf{H}}_{(n|n)}$.

The algorithm for the nth OFDM symbol is depicted in details in Fig. 1. From the previous OFDM symbol $(n-1)$, we execute the EKF Time Update Equations (32) to obtain the prediction parameters $\hat{\boldsymbol{\mu}}_{(n|n-1)}$. The predicted version of the channel matrix $\hat{\mathbf{H}}_{(n|n-1)}$ is computed from $\hat{\boldsymbol{\mu}}_{(n|n-1)}$ instead of $\boldsymbol{\mu}_n$ with Eq. (17). Therefore, the equalization task is now possible since a version of the channel matrix is available. Before this step, the contribution of the pilots to $\mathbf{y}_n$ is removed:

$$\mathbf{y}'_n = \mathbf{y}_n - \hat{\mathbf{H}}_{(n|n-1)} \cdot \mathbf{x}_{\mathbf{p}_n} \tag{34}$$

where the vector $\mathbf{x}_{\mathbf{p}_n}$ is a $N_T N \times 1$ vector composed of the pilots at the pilot positions and 0 elsewhere. With the assumption that $\hat{\mathbf{H}}_{(n|n-1)} \cdot \mathbf{x}_{\mathbf{p}_n} = \mathbf{H}_n \cdot \mathbf{x}_{\mathbf{p}_n}$, we obtain a new version of the transmission model that only includes the data:

$$\mathbf{y}'_n = \mathbf{H}_n^{\text{data}} \cdot \mathbf{x}_n^{\text{data}} + \mathbf{w}_n^{\text{data}} \tag{35}$$

where the $N_R N \times N_T(N - N_p)$ matrix $\mathbf{H}_n^{\text{data}}$ is obtained by removing the $N_T N_p$ columns of $\mathbf{H}_n$ at the pilot positions $\mathcal{P}$. $\mathbf{x}_n^{\text{data}}$ and $\mathbf{w}_n^{\text{data}}$ are $N_T(N - N_p) \times 1$ vectors built from $\mathbf{x}_n$ and $\mathbf{w}_n$, respectively, by removing the vector elements at pilot positions $\mathcal{P}$.

Equalization is performed on this model, yielding a first version of the detected data symbols $\hat{\mathbf{x}}_{(n|n-1)}$. The Measurement Update Equations (33) are then computed by using $\hat{\mathbf{x}}_{(n|n-1)}$ instead of $\mathbf{x}_n$ in Eq. (30). Finally, a new equalization is performed with the updated parameters $\hat{\boldsymbol{\mu}}_{(n|n)}$ to obtain the updated version of the data symbols $\hat{\mathbf{x}}_{(n|n)}$.

The algorithm is initialized with $\hat{\boldsymbol{\mu}}_{(0|0)} = \mathbf{0}_{LN_c+N_R N_T,1}$, and $\mathbf{P}_{(0|0)}$ computed with Eq. (28).

B. Computational Complexity

The purpose of this section is to determine the implementation complexity in terms of the number of the multiplications needed for our algorithm. The matrices $\mathbf{F}^{(r,t)}$ are pre-computed and stored if the delays are invariant for a great number of OFDM symbols. The computational cost of computing the different terms and processes of the algorithm is given by Table I. The complexity analysis of Time Update Equations and Measurement Update Equations of the Kalman filter in Table I uses the fact that $\boldsymbol{\mathcal{A}}$ is a sparse matrix. In practice, L, N_T, N_R and N_c are much smaller than N, therefore, the computational complexity of our algorithm is $O(N_R^3 N^3)$. So we can say that our proposed algorithm and the algorithm proposed in [13] have asymptotically the same complexity (same order of growth). The algorithm in [13] will be used for performance comparison in Section V.

V. Simulation

In this section, the performance of our recursive algorithm is evaluated in terms of Mean Square Error (MSE) for joint CA and CFO estimation and Bit Error Rate (BER) for data detection. We consider two antennas at the transmitter and two antennas at the receiver ($N_T = N_R = 2$). A normalized 4-QAM MIMO-OFDM system with $N = 128$ subcarriers, $N_g = \frac{N}{8}$, $N_p = \frac{N}{4}$ pilots ($i.e., L_f = 4$), and $\frac{1}{T_s} = 2MHz$ was used.

Both parametric and equivalent discrete channel models are being discussed. We recall that the derivations have been carried out for the parametric model, although the equations for the equivalent discrete-time channel model can also be obtained by substituting the set of delays $\left\{\tau_l^{(r,t)}\right\}$ by the tap indices (see Section II-A).

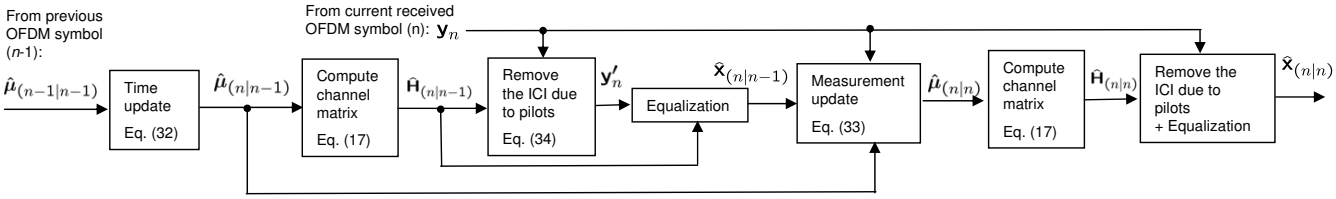

Fig. 1. Joint Data Detection, channel estimation and CFO estimation algorithm.

Term or process	Computational cost (number of multiplications)
$\mathbf{M}_d^{(r,t)}(\nu_n^{(r,t)})$	N^3
$\mathbf{M}_d'(\nu_n^{(r,t)})$	N^3
$\boldsymbol{\mathcal{K}}_n^{(r,t)}(\nu_n^{(r,t)})$	$N(N+1)LN_c$
$\boldsymbol{\mathcal{K}}_n'^{(r,t)}(\nu_n^{(r,t)})$	$N(N+1)LN_c$
$\boldsymbol{\mathcal{V}}_n(\boldsymbol{\mu}_n)$	NN_cL
$\mathbf{H}_n$	$NN_c(NN_TN_R+L)$
Removing ICI	$N_pN_TNN_R$
QR-decomposition	$\frac{2}{3}N_d^3+N_d^2+\frac{1}{3}N_d-2$ with $N_d=N_RN$
data QR-detection	$\frac{1}{2}N'_d(N'_d+1)$ with $N'_d=N_T(N-N_p)$
Time Update Equations	$(LN_c^2+N_TN_R)+(LN_c^2+N_T^2N_R^2)(LN_c+N_TN_R)$
Measurement Update Equations	$2N_RN(LN_c+N_RN_T)^2+N_RN(LN_c+N_RN_T)(2N_RN+1)+N_RN_TN^2+(N_RN)^3$

TABLE I
COMPUTATIONAL COMPLEXITY

In Section V-A, the parametric channel model is being considered with a classical scenario with one base station and one mobile receiver, and one CFO parameter to be estimated. Section V-B deals with the equivalent discrete channel model and considers a more pessimistic scenario where each transmitter and receiver requires its own RF-IF chain. For this scenario, the number of CFO parameters to be estimated ($N_TN_R=4$) is the largest. This scenario could correspond to the area of coordinated base stations or network MIMO. Performance comparisons have been carried out with the algorithm proposed in [13].

A. Parametric channel model

We assume that all the (r,t) channel links, $r=1,\ldots,N_R$, $t=1,\ldots,N_T$ share the same path delays and fading properties ($i.e.$, the same number of paths, of ${\sigma_{\alpha_l}^{(r,t)}}^2$ and $\tau_l^{(r,t)}$) since the antennas are very close to each other, which is typical in practice. The Rayleigh channel model given in [9] [11]($L^{(r,t)}=6$ paths and maximum delay $\tau_{max}=10T_s$, see Table II) was chosen. The MSE will be computed for both path CA and CFO to evaluate the estimation performance. First, let us define:

$$\hat{\boldsymbol{\alpha}}_{(n|n)} \stackrel{\text{def}}{=} \text{blkdiag}\{\underbrace{\mathbf{B},\ldots,\mathbf{B}}_{L \text{ times}}\} \cdot \left(\hat{\boldsymbol{\mu}}_{(n|n)}\big|_{[0:N_cL-1]}\right)$$

$$\boldsymbol{\alpha}_n \stackrel{\text{def}}{=} \Big[{\boldsymbol{\alpha}_{0,n}^{(1,1)}}^T,\ldots,{\boldsymbol{\alpha}_{L^{(1,1)}-1,n}^{(1,1)}}^T,\ldots, {\boldsymbol{\alpha}_{0,n}^{(1,N_T)}}^T,\ldots {\boldsymbol{\alpha}_{L^{(N_R,N_T)}-1,n}^{(N_R,N_T)}}^T\Big]^T$$

$$\hat{\boldsymbol{\nu}}_{(n|n)} \stackrel{\text{def}}{=} \hat{\boldsymbol{\mu}}_{(n|n)}\big|_{[N_cL:N_cL+N_\nu]}$$

Path number	Average Power (dB)	Delay (T_s)
0	-7.219	0
1	-4.219	0.4
2	-6.219	1
3	-10.219	3.2
4	-12.219	4.6
5	-14.219	10

TABLE II
RAYLEIGH CHANNEL PARAMETERS

where N_ν is the number of CFO to be estimated. The MSE of the path CA (denoted MSE_α) and the MSE of the CFO (denoted MSE_ν) are computed as follows (we recall that L is the total number of paths for the MIMO channel, see Section II-B):

$$\text{MSE}_\alpha \stackrel{\text{def}}{=} \frac{1}{K}\sum_{n=0}^{K-1}\frac{1}{N_bL}\left(\hat{\boldsymbol{\alpha}}_{(n|n)}-\boldsymbol{\alpha}_n\right)^H\left(\hat{\boldsymbol{\alpha}}_{(n|n)}-\boldsymbol{\alpha}_n\right) \quad (36)$$

$$\text{MSE}_\nu \stackrel{\text{def}}{=} \frac{1}{K}\sum_{n=0}^{K-1}\frac{1}{N_\nu}\left(\hat{\boldsymbol{\nu}}_{(n|n)}-\boldsymbol{\nu}_n\right)^H\left(\hat{\boldsymbol{\nu}}_{(n|n)}-\boldsymbol{\nu}_n\right) \quad (37)$$

where K is set to 1000 in our simulations. The MSE and the BER were evaluated under a rapid time-varying channel with $f_dT=0.1$ (corresponding to a vehicle speed of $300km/h$ at $f_c=5$ GHz). A GCE-BEM with $N_c=4$ was initially chosen to model the path CA of the channel and $\nu=0.1$.

The tracking capability of our proposed algorithm is first demonstrated as a function of time. Real and imaginary parts of one trajectory example of $\boldsymbol{\alpha}_{l,n}^{(r,t)}$ are plotted in Fig. 2 for $r=1,t=1$ and $l=0,\ldots,5$ at $E_b/N_0=20$ dB. After an initial transient, the algorithm locks on to the true value of the

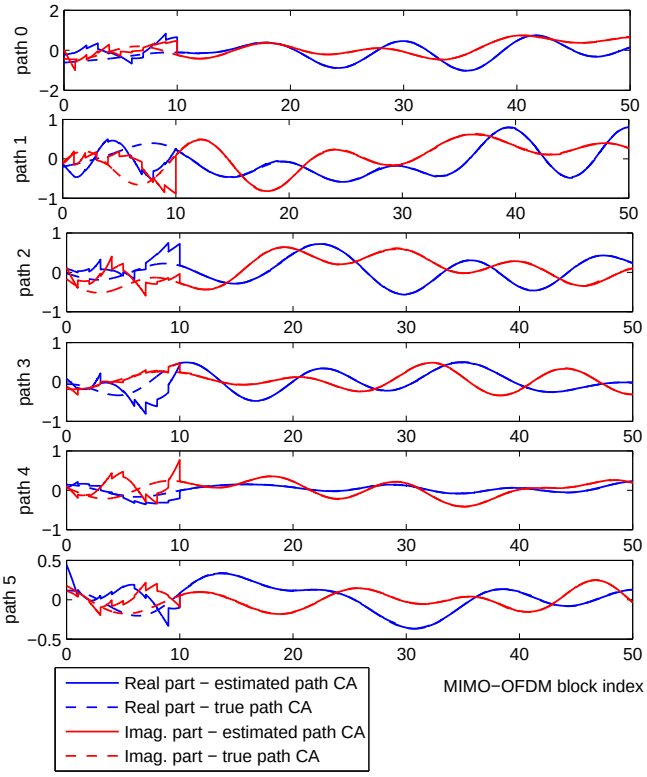

Fig. 2. Time domain tracking of the path CA at $E_b/N_0 = 20$ dB and $f_dT = 0.1$ for $r = 1, t = 1$, $a = 0.99$

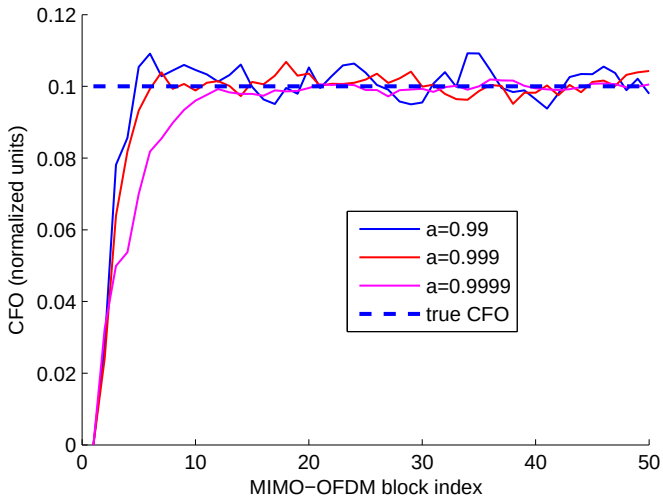

Fig. 3. Time domain tracking of CFO at $E_b/N_0 = 20$ dB and $f_dT = 0.1$ for different values of a

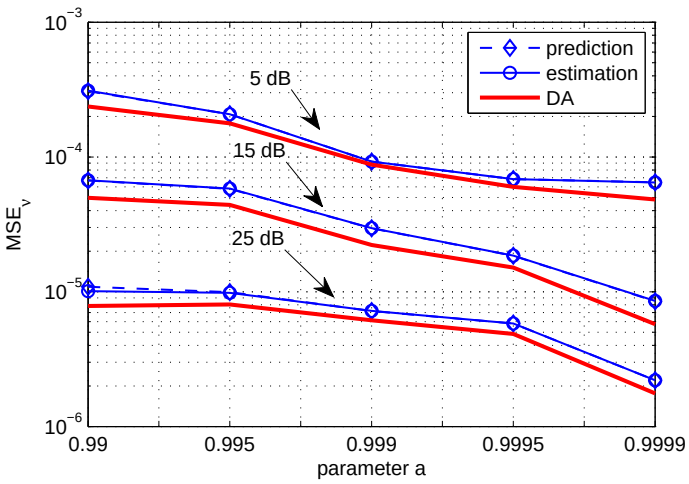

Fig. 4. MSE of the CFO estimation (MSE$_\nu$) as a function of a at $E_b/N_0 = 5,\ 15,\ 25$ dB for $f_dT = 0.1$, $\nu = 0.1$

path CA and tracks them closely, even for the paths with low average power.

The convergence results for the CFO is shown in Fig. 3 for different values of the CFO tracking parameter a (see Section III-B). To emphasize the effect of a, simulations are performed in Data-Aided mode. Classically, a is chosen from 0.99 up to 0.9999 [13][22]. The estimated CFO is initialized to zero (see Section III-D). It is observed that the convergence time increases with a, which is an expected result. On the other hand, the MSE is expected to decrease with increasing values of a, which can be observed in Fig. 4. However, the gain in MSE performance is too small to impact the BER, which remains constant for any values of a (see Fig. 5). So it turns out that our system is relatively independent of a.

Fig. 6 shows the CA MSE as a function of E_b/N_0. For reference, the MSE obtained in Data-Aided (DA) mode (knowledge of the data symbols) is also plotted. In addition to the MSE of the estimated CA (see Eq. (36)), we added the MSE obtained with the predicted CA by substituting $\hat{\boldsymbol{\alpha}}_{(n|n)}$ with $\hat{\boldsymbol{\alpha}}_{(n|n-1)}$ in Eq. (36). As expected, it is observed that both predicted and estimated MSE approach their DA curve when E_b/N_0 is increased. Indeed, for large E_b/N_0 values, the number of detection errors is small. On the other hand, it is seen that the estimated curve is far better than the predicted curve for each E_b/N_0. Hence, it can be concluded that the measurement update task (Eq. (33)) is still efficient, even when the equations are computed with the predicted data symbols $\hat{\mathbf{x}}_{(n|n-1)}$ (see Section IV-A).

Then, to evaluate the performance of our joint algorithm, the curves obtained with the perfect knowledge of the CFO are plotted. It is seen that the performance in terms of CA estimation are unchanged. So, it turns out that the CFO estimation does not impact the CA estimation.

Let us now discuss the CFO estimation. Fig. 7 shows the MSEs for the CFO obtained with the predicted and the estimated parameter. Similarly to the CA MSE, the curves in DA mode and with the perfect knowledge of the CA are shown. First, it is observed that the estimated curve is very close to the predicted one. This is due to the fact that the CFO is constant in our model, and so the AR-model is not very accurate. Unlike for the CA estimation task, the knowledge of the unwanted parameter highly increases the performance of the CFO estimation because the CA rapidly varies in time, yielding high MSE. The impact of their estimation, due to this high MSE, is not negligible on the CFO estimation task.

Figure 8 gives the corresponding BER curve. A lower bound for the BER performance is given by using the ideal channel

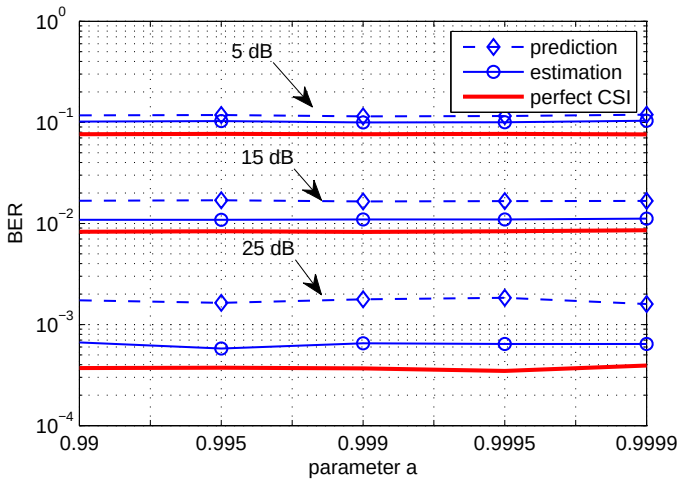

Fig. 5. BER as a function of a at $E_b/N_0 = 5,\ 15,\ 25$ dB for $f_dT = 0.1$

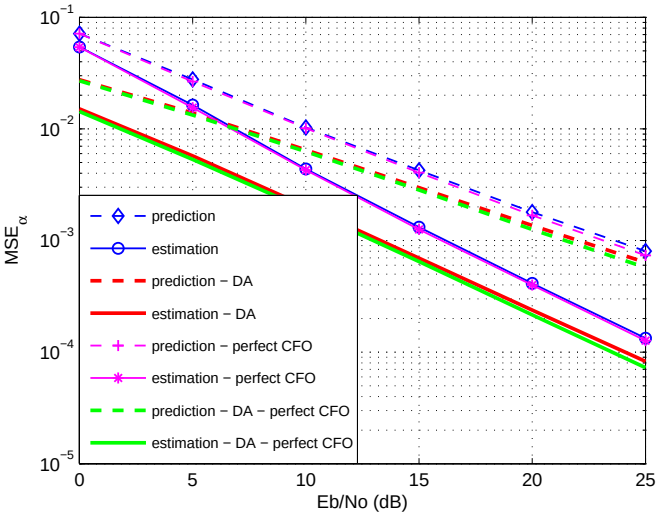

Fig. 6. MSE of the CA estimation (MSE$_\alpha$) as a function of E_b/N_0 for $f_dT = 0.1$, $\nu = 0.1$

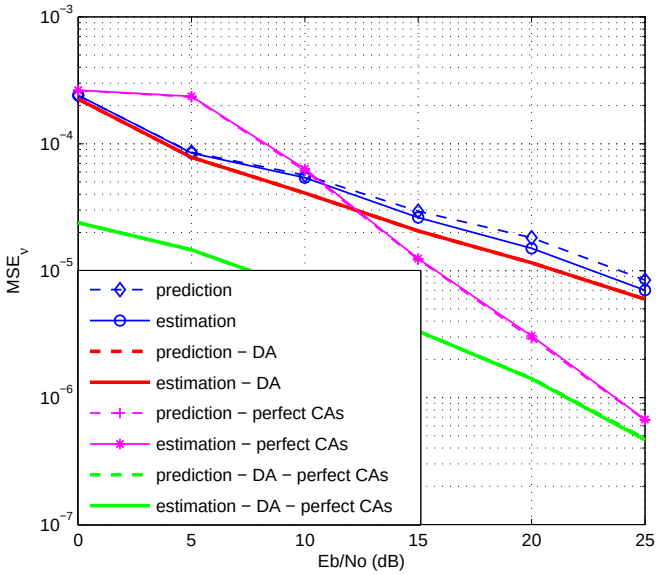

Fig. 7. MSE of the CFO estimation (MSE$_\nu$) as a function of E_b/N_0 for $f_dT = 0.1$, $\nu = 0.1$

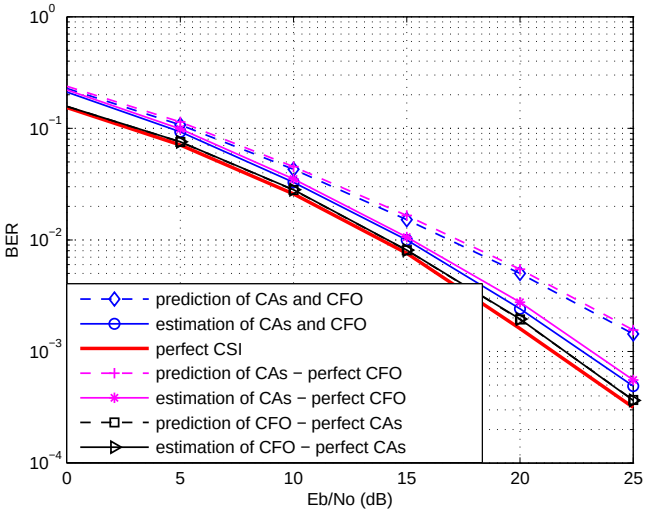

Fig. 8. Bit Error Rate (BER) as a function of E_b/N_0 for $f_dT = 0.1$, $\nu = 0.1$

state information (CSI), i.e. perfectly known CA and CFO at the receiver. Together with this reference curve, we also plotted the BER curves obtained with the perfect knowledge of the CFO only, and the CA only. As expected, the parameter that degrades the most the performance is the CA, due to the high mobility of the channel.

Figures 9 and 10 show the impact of the number of BEM coefficients N_c to the performance for different BEMs. The considered BEM are the P-BEM, the GCE-BEM, and the DKL-BEM (see Section II-B). For low E_b/N_0 values, the P-BEM is the most efficient in terms of MSE, but the gain is negligible on the BER. However, for large E_b/N_0 values, the gain in terms of MSE obtained with the GCE-BEM and DKL-BEM impacts the BER. Hence, it turns out that the best trade-off is to choose $N_c = 3$ and either the GCE-BEM or the DKL-BEM. Nevertheless, these two BEMs require some a-priori information (Doppler frequency f_d for the GCE-BEM and correlation matrix for the DKL-BEM) which is not the case for the P-BEM. It is noteworthy that the BER would be more sensitive to the estimation errors with a higher order modulation (we recall that we used a QPSK modulation).

B. Equivalent discrete channel model - comparison with the algorithm of [13]

Here, we consider the equivalent discrete channel model where 4 CFO have to be estimated (one per sub-channel). This scenario could correspond to the area of coordinated base stations or network MIMO. The CFOs have been arbitrarily fixed to $0.1, 0.07, -0.1, -0.05$. For the sake of comparison, we also show the performance of the algorithm proposed in [13], called the classical algorithm from now on. This

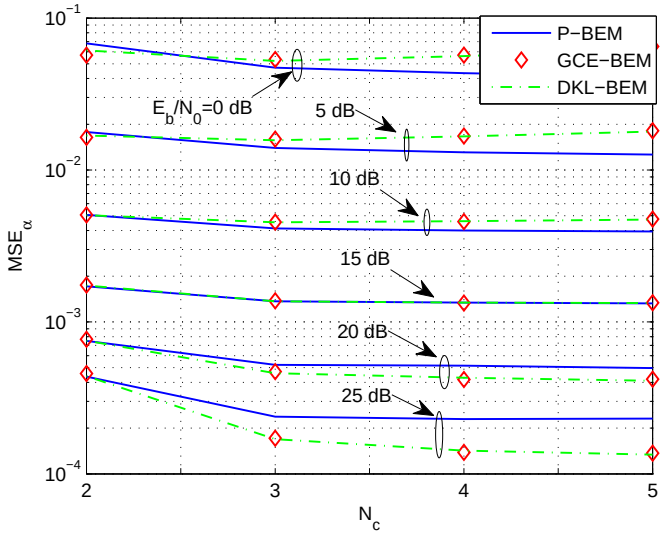

Fig. 9. MSE of the CA estimation (MSE_α) as a function of N_c for different BEM, $f_dT = 0.1$

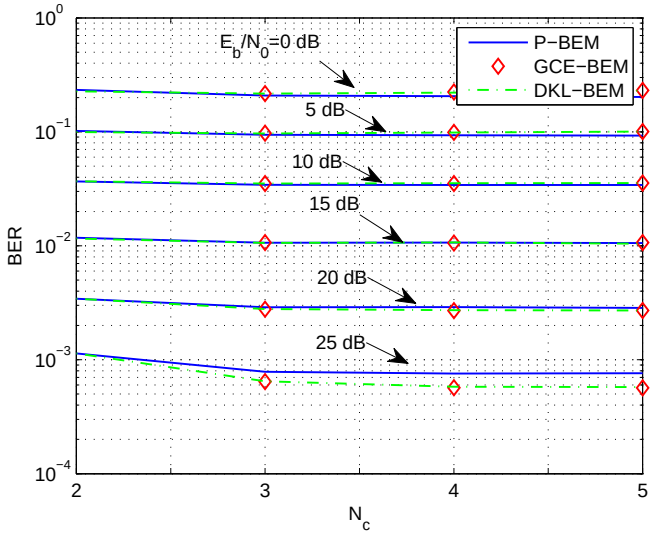

Fig. 10. BER as a function of N_c for different BEM, $f_dT = 0.1$

algorithm is also based on extended Kalman filtering to carry out channel taps and CFO estimation together with data detection. Note that the simulations presented in [13] have been carried out in Decision-Directed (DD) mode only, i.e. only decoded data symbols are used to perform the filtering. However, when introducing their algorithm, the authors also stated that in case of high mobility, pilot signals are also needed [13][23]. So to compare both algorithms, we insert pilots in the algorithm following our pilot scheme (see Section IV-A). The same channel as in [13] has been selected i.e. a power loss $[0, -1, -3, -9][dB]$ and delay profile $[0, 1, 2, 3]\mu s$ ($i.e.$ $[0T_s, 2T_s, 4T_s, 6T_s]$), which corresponds to a urban type of scenario. We also fix the same parameter $a = 0.99$ as in [13].

First, simulations for different speeds ranging from 30 km/h up to 300 km/h have been performed at 20 dB (see Fig. 11).

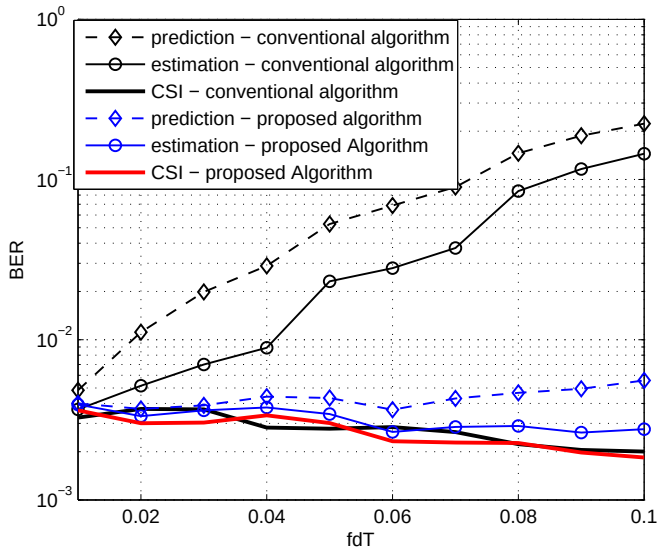

Fig. 11. BER performance for variable terminal velocity ($E_b/N_0 = 20dB$)

For reference, the performance of the algorithm is given by using the ideal channel state information (CSI).

For the classical algorithm, the performance degrades rapidly as the speed is increased. This is expected since this algorithm does not take into account the ICI due to mobility. However, we observe that our algorithm is far more robust to speed. The prediction performance degrades with the speed but is clearly compensated by the estimation.

VI. Conclusion

In this paper, a new algorithm which jointly estimates path Complex Amplitudes (CA) and Carrier Frequency Offsets (CFO) in MIMO environments has been presented. The algorithm is based on a parametric channel model or equivalent discrete channel model. Within one OFDM symbol, each time-varying CA is approximated by a Basis Expansion Model (BEM) representation. The dynamics of the BEM coefficients and that of the CFO parameters are modeled by first-order auto-regressive processes. Parameter estimation is performed by Extended Kalman Filtering and the data recovery is carried out by means of a QR-equalizer. Compared to the conventional algorithm, simulation results show the good robustness of our algorithm to fading rate for normalized Doppler frequency values f_dT up to 0.1. For this very high mobility, the performance of the joint estimation algorithm in terms of Bit Error Rate is close to the performance obtained with perfect knowledge of channel and CFO as long as 3 BEM coefficients are used with either the GCE-BEM or the DKL-BEM.

Appendix

In this section, we detail the computation of the state noise variance $\sigma^2_{u_\nu}$. For the sake of simplicity, only the scalar case is performed. The vectorial case can be easily extended from this. The scalar version of (23) is as follows:

$$\nu_n = a \cdot \nu_{n-1} + u_{\nu_n} \tag{38}$$

First, let us define the correlation function of ν:

$$R_\nu[m] = \mathrm{E}\left[\nu_n \nu_{n-m}\right] \tag{39}$$

Using (38) in (39) yields:

$$R_\nu[m] = a \cdot R_\nu[m-1] + \mathrm{E}\left[u_{\nu_n} \nu_{n-m}\right] \tag{40}$$

Then, we compute (40) for $m = 1$ and $m = 0$, yielding:

$$R_\nu[1] = a \cdot R_\nu[0] \tag{41}$$

$$R_\nu[0] = a \cdot R_\nu[-1] + \sigma^2_{u_\nu} \tag{42}$$

Note that the expectation $\mathrm{E}\left[u_{\nu_n} \nu_{n-m}\right]$ equals zero for $m = 1$ since ν_{n-1} only depends on $u_{\nu_{n-1}}$ (and not on u_{ν_n}) on the one hand, and on the other hand u_{ν_n} is zero-mean white Gaussian noise.

Combining (41) and (42) yields:

$$\sigma^2_{u_\nu} = \left(1 - a^2\right) R_\nu[0] \tag{43}$$

since $R_\nu[-1] = R_\nu[1]$.

ACKOWLEDGEMENT

This work has been carried out in the framework of the CISIT (Campus International sur la Sécurité et Intermodalité des Transports) project and funded by the French Ministry of Research, the Region Nord Pas de Calais and the European Commission (FEDER funds)

REFERENCES

[1] Y. Li, "Simplified Channel Estimation for OFDM Systems with Multiple Transmit Antennas," *IEEE Trans. Wireless Comm.*, vol. 1, 2002.

[2] Z. J. Wang and Z. Han, "A MIMO-OFDM Channel Estimation Approach Using Time of Arrivals," *IEEE Trans. Wirel. Comm.*, vol. 4, 2005.

[3] Z. J. Wang, Z. Han, and K. J. R. Liu, "MIMO-OFDM Channel Estimation via Probabilistic Data Association Based TOAs," in *IEEE Global Telecommunications Conference GLOBECOM '03*, 2003.

[4] H. Hijazi and L. Ros, "Polynomial estimation of time-varying multi-path gains with intercarrier interference mitigation in OFDM systems," *IEEE Trans. Vehic. Techno.*, vol. 57, 2008.

[5] J.-G. Kim and J.-T. Lim, "MAP-Based Channel Estimation for MIMO-OFDM Over Fast Rayleigh Fading Channels," *IEEE Trans. Vehic. Techno.*, vol. 57, 2008.

[6] T. Pollet, M. V. Bladel, and M. Moeneclaey, "BER Sensitivity of OFDM Systems to Carrier Frequency Offset and Wiener Phase Noise," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 191 – 193, February/March/April 1995.

[7] H. Steendam and M. Moeneclaey, "Sensitivity of Orthogonal Frequency-Division Multiplexed Systems to Carrier and Clock Synchronisation Errors," *Signal Processing*, vol. 80, pp. 1217–1229, 2000.

[8] P. H. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Commun*, vol. 42, pp. 2908–2914, 1994.

[9] H. Hijazi and L. Ros, "Joint Data QR-Detection and Kalman Estimation for OFDM Time-varying Rayleigh Channel Complex Gains," *IEEE Trans. Comm.*, vol. 58, pp. 170–178, 2010.

[10] Z. Tang, R. C. Cannizzaro, G. Leus, and P. Banelli, "Pilot-assisted time-varying channel estimation for OFDM systems," *IEEE Trans. Signal Process.*, vol. 55, pp. 2226–2238, 2007.

[11] H. Hijazi, E. P. Simon, M. Lienard, and L. Ros, "Channel Estimation for MIMO-OFDM Systems in Fast Time-Varying Environments," in *international symposium on communications, control and signal processing (ISCCSP)*, 2010, accepted.

[12] T. Roman, M. Enescu, and V. Koivunen, "Joint time-domain tracking of channel and frequency offset for OFDM systems," in *4th IEEE Workshop on Signal Processing Advances in Wireless Communications. SPAWC 2003.*, 15-18 2003, pp. 605 – 609.

[13] ——, "Joint Time-Domain Tracking of Channel and Frequency Offsets for MIMO OFDM Systems," *Wireless Personal Communications*, vol. 31, pp. 181–200, 2004.

[14] B. Yang, K. B. Letaief, R. S. Cheng, and Z. Cao, "Channel estimation for OFDM transmisson in mutipath fading channels based on parametric channel modeling," *IEEE Trans. Commun.*, vol. 49, pp. 467–479, 2001.

[15] E. Simon, L. Ros, and K. Raoof, "Synchronization over Rapidly Time-varying Multipath Channel for CDMA Downlink RAKE Receivers in Time-Division Mode," *IEEE Trans. Vehic. Techno.*, vol. 56, 2007.

[16] W. C. Jakes, *Microwave Mobile Communications*. IEEE Press, 1983.

[17] G. Leus, "On the Estimation of Rapidly Time-Varying Channels," in *Euro. Signal Process. Conf. (EUSIPCO)*, 2004.

[18] K. D. Teo and S. Ohno, "Optimal MMSE Finite Parameter Model for Doubly-selective Channels," in *IEEE Global Telecommunications Conference GLOBECOM '05*, 2005.

[19] A. R. Kannu and P. Schniter, "MSE-optimal Training for Linear Time-varying Channels," in *IEEE ICASSP Conf.*, 2005.

[20] K. E. Baddour and N. C. Beaulieu, "Autoregressive modeling for fading channel simulation," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1650–1662, 2005.

[21] B. Anderson and J. B. Moore, *Optimal filtering*. Prentice-Hall, 1979.

[22] S. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice Hall, 1993.

[23] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Transactions on Signal Processing*, vol. 51, no. 6, pp. 1615 – 1624, june 2003.

Eric P. Simon received the Masters degree in electronics engineering from the Superior School of Electronics (ESCPE), Lyon, France, in 1999, and the Ph.D. degree in signal processing and communications from the National Polytechnic Institute of Grenoble (INPG), France, in 2004. During 2005, he was a Teaching Assistant at the INPG and the following year he joined one of France Telecom R&D Laboratories as a Postdoctoral Fellow. He is currently an Associate Professor at the Institute of Electronics, Microelectronics and Nanotechnology (IEMN), TELICE (Telecommunications, Interference and Electromagnetic Compatibility) Group, University of Lille, France. His main research interests are in mobile communications and carrier and symbol synchronization.

Laurent Ros received the degree in electrical engineering from the "Ecole Supérieure d'Electricité" (Supélec), Paris, France, in 1992 and the Ph.D. degree in signal processing and communications from Grenoble Institute of Technology (Grenoble-INP), Grenoble, France, in 2001. From 1993 to 1995, he was with France-Telecom R&D center, Lannion, France, where he worked in the area of very low frequency transmissions for submarine applications, in collaboration with Direction of Naval Construction, Toulon, France. From 1995 to 1999, he was a Research and Development team manager at Sodielec, Millau, France, where he worked in the design of digital modems and audio codecs for telecommunication applications. Since 1999, he has joined the Gipsa-laboratory/DIS (ex "Laboratory of Image and Signal") and Grenoble-INP where he is currently an Associate Professor. His general research interests include statistical signal processing, synchronization and channel estimation problems for wireless communications.

Hussein Hijazi is an assistant professor in the faculty of engineering at the Department of Telecommunication and Electrical engineering of the Holy-Spirit University of Kaslik (USEK). He received his Diploma in computer and communications engineering in 2004 from the Lebanese University, Beirut, Lebanon. He received his M.S. and Ph.D. degrees in signal processing and communications in 2005 and 2008 respectively from the Institut Polytechnique de Grenoble (Grenoble-INP) , Grenoble, France. From September 2008 to August 2009, he was an assistant professor at the engineering school école d'ingénieurs pour l'énergie, l'eau et l'environnement (ENSE3) of Grenoble-INP. From September 2009 to December 2009, he was a Research engineer in telecommunications at the IEMN laboratory of university of Lille 1. From December 2009 to December 2010, he was a Post-doctoral researcher at France Telecom-orange Labs, RD center, Meylan, France. His research interests are in the areas of signal processing and communications, including synchronisation, channel estimation and equalization problems for wireless digital communications and cooperative communications for wireless sensor network. Dr. Hijazi serves as a reviewer for several international journals and conferences including IEEE transactions on Signal Processing, IEEE transactions on communications, IEEE transactions on Wireless Communications, IEEE transactions on Vehicular Technology, and EURASIP Journal on Wireless Communications and Networking.

Jin Fang received the BS in 2007 at the university of Shanghai and the MS in 2009 at the university of Lille1, France. She is currently at the Institute of Electronics, Microelectronics and Nanotechnology (IEMN), TELICE (Telecommunications, Interference and Electromagnetic Compatibility) Group, University of Lille, France, for a Ph.D. degree. Her main research interests are in mobile communications and carrier and symbol synchronization.

Davy P. Gaillot was born in France. He received the B.S. Degree in Mechanical Engineering from Ecole Nationale d'Ingénieurs de Metz (ENIM), Metz, France, in 2002. Simultaneously, he gained a Masters Degree in Mechanics, Materials, Structures and Processes from the University de Metz, Metz, France. From 2003 to 2007, he joined Prof. Christopher J. Summers' Nanophotonics Group in the Department of Materials Science and Engineering at the Georgia Institute of Technology, Atlanta, USA. After completing his Ph.D. thesis, he was a Post-doctoral fellow in Prof. Didier Lippens' DOME group at the University of Lille. Since 2008, he is an Associate Professor at the University of Lille in the TELICE group. Dr. Gaillot research interests include the development of novel telecommunication applications that are supported by legacy or upcoming WiFi network standards.

Marion Berbineau was born in September 18, 1962, in Toulouse, France. She received the Engineering degree from the Ecole Universitaire d'Ingénieurs de Lille - Informatique, mesures, Automatique, and the Ph.D. degree from the University of Lille, respectively in 1986 and 1989. She joined the Institut national de Recherche sur la Sécurité dans les Transport in 1989. Currently, she is Research Director and Head of the Laboratoire Ondes et Signaux pour les Transports since April 2002. Her fields of interest are electromagnetic, radio propagation in open or confined areas using natural propagation, radiating cables or wave guide, signal processing for wireless telecommunication systems applied to transport, navigation/localization methods. She is currently involved in several national and Europeans projects dealing with communication, navigation and environment perception in the field of terrestrial transports. Dr Berbineau is chairman of the "Intelligent Mobility" pole in the Network of Excellence for Railway Research.