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

Eric Pierre Simon[1], Hussein Hijazi[1], Laurent Ros[2]

[1] IEMN/TELICE laboratory, University of Lille - FRANCE

[2] GIPSA-lab, Department Image Signal, BP 46 - 38402 Saint Martin d'Hères - FRANCE

e-mail: eric.simon@univ-lille1.fr, hussein.hijazi@hotmail.fr, laurent.ros@gipsa-lab.grenoble-inp.fr

***Abstract*—In this paper, a novel pilot-aided iterative algorithm is developed for MIMO-OFDM systems operating in fast time-varying environment. An L-path channel model with known path delays is considered to jointly estimate the multi-path Rayleigh channel complex gains and Carrier Frequency Offset (CFO). Each complex gain 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 recursive estimation using Extended Kalman Filtering. Hence, the channel matrix is easily computed and the data symbol is estimated with free inter-sub-carrier-interference (ICI) when the channel matrix is QR-decomposed. It is shown that only one iteration is sufficient to approach the performance of the ideal case for which the knowledge of the channel response and CFO is available.**

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 critical 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 work consider that all the paths exhibit the same Doppler shift. Hence, they group together the Doppler shift and CFO due to oscillator mismatchs in order to obtain just one offset parameter [8][9] for each channel branch. However, this model is not sufficiently accurate since separate offset parameters are needed for each propagation path given that the Doppler shift depends on the angle of arrival, which is peculiar to each path. Recently, it has been proposed to directly track the channel paths, which permits to take into account separate Doppler shifts for each path ([10][11] for SISO and [12] for MIMO). Those works estimate the equivalent discrete-time channel taps ([11]) or the real path complex gains ([10][12]) which are both modeled by a basis expansion model (BEM). The BEM methods are 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. In this paper, we propose a complete algorithm capable of estimating this CFO jointly with the time-variation of each channel path in MIMO environment.

Generally, it is preferable to directly estimate the physical channel parameters [13] [10] [12] instead of the equivalent discrete-time channel taps [11]. 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. Additionally, estimating the physical propagation parameters means estimating multi-path delays and multi-path complex gains. Note that in Radio-Frequency transmissions, the delays are quasi-invariant over several MIMO-OFDM blocks [14] [4] (whereas the complex gains 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 [13][10]. To further improve the estimation accuracy, our algorithm uses decision feedback. Hence, the accuracy of the channel estimation, frequency offset estimation and symbol detection are simultaneously enhanced. Note also that, since the pilots are used for both channel and frequency offset estimation, the pilot usage efficiency is greatly improved. Our algorithm is a recursive algorithm based on Extended Kalman Filtering

(EKF) combined with QR-equalization for data detection.

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 $\mathbf{X}$. 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 $T = N_s T_s$, where T_s is the sampling time and $N_s = N + N_g$. Let $\mathbf{x}_n = \left[\mathbf{x}_n^{(1)^T}, \mathbf{x}_n^{(2)^T}, ..., \mathbf{x}_n^{(N_T)^T}\right]^T$ be the nth transmitted MIMO-OFDM block, where $\mathbf{x}_n^{(t)} = \left[x_n^{(t)}[-\frac{N}{2}], x_n^{(t)}[-\frac{N}{2}+1], ..., x_n^{(t)}[\frac{N}{2}-1]\right]^T$ is the nth transmitted OFDM symbol by the tth transmit antenna and $\{x_n^{(t)}[b]\}$ are normalized symbols ($i.e.$, $\mathrm{E}\big[x_n^{(t)}[b]x_n^{(t)*}[b]\big] = 1$). The frequency mimatch 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 CFOs. In a $N_T \times N_R$ MIMO system this leads to $N_T N_R$ different CFOs. However, if transmitter or receiver antennas share RF-IF chains, fewer different CFOs occured. The system model describes the general case where it is necessary to compensate for $N_T N_R$ CFOs. 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 CFOs can be stacked in vector form as:

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

After transmission over a multi-path Rayleigh channel, the nth received MIMO-OFDM block $\mathbf{y}_n = \left[\mathbf{y}_n^{(1)^T}, \mathbf{y}_n^{(2)^T}, ..., \mathbf{y}_n^{(N_R)^T}\right]^T$, where $\mathbf{y}_n^{(r)} = \left[y_n^{(r)}[-\frac{N}{2}], y_n^{(r)}[-\frac{N}{2}+1], ..., y_n^{(r)}[\frac{N}{2}-1]\right]^T$ is the nth received OFDM symbol by the rth receive antenna, is given by [4] [11]:

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

where $\mathbf{w}_n = \left[\mathbf{w}_n^{(1)^T}, \mathbf{w}_n^{(2)^T}, ..., \mathbf{w}_n^{(N_R)^T}\right]^T$ with $\mathbf{w}_n^{(r)} = \left[w_n^{(r)}[-\frac{N}{2}], w_n^{(r)}[-\frac{N}{2}+1], ..., w_n^{(r)}[\frac{N}{2}-1]\right]^T$ 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:

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

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 [10] ($i.e.$ delays $\left\{\tau_l^{(r,t)}\right\}$ and complex gains $\left\{\alpha_l^{(n,r,t)}(qT_s) = \alpha_l^{(r,t)}(nT + qT_s)\right\}$), yielding Eq. (4) and (5), respectively.

$L'^{(r,t)} < N_g$ and $L^{(r,t)}$ are respectively the number of channel taps and 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 that the delays are multiple of T_s as mostly 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 [15]. 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 $L = \sum_{r=1}^{N_R} L_{(r)} = \sum_{r=1}^{N_R} \sum_{t=1}^{N_T} L^{(r,t)}$ be the total number of complex gains for the MIMO channel. Since the number of samples to be estimated $L N_s$ is greater than the number of observation equations $N_R N$, it is not efficient to estimate the time-variation of the complex gains, using directly the observation model in (2). Thus, we need to reduce the number of parameters to be estimated. 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 $N_s \times 1$ vector that collects the time-variation of the lth path of the (r,t) branch within the nth MIMO-OFDM block:

$$\boldsymbol{\alpha}_l^{(n,r,t)} = \left[\alpha_l^{(n,r,t)}(-N_g T_s), ..., \alpha_l^{(n,r,t)}\big((N-1)T_s\big)\right]^T \quad (6)$$

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

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

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 $N_s \times N_c$ matrix $\mathbf{B}$ is defined as: $\mathbf{B} = [\mathbf{b}_0, ..., \mathbf{b}_{N_c-1}]$. The $N_s \times 1$ vector $\mathbf{b}_d$ is termed as the dth expansion basis. $\mathbf{c}_l^{(n,r,t)} = \left[c_{(0,l)}^{(n,r,t)}, ..., c_{(N_c-1,l)}^{(n,r,t)}\right]^T$ 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 [16]. Under this criterion, the optimal BEM coefficients and the corresponding model error are given by:

$$\mathbf{c}_l^{(n,r,t)} = \left(\mathbf{B}^H\mathbf{B}\right)^{-1}\mathbf{B}^H\boldsymbol{\alpha}_l^{(n,r,t)} \quad (8)$$

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

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

$$\mathrm{MMSE}_l^{(r,t)} = \frac{1}{N_s}\mathrm{E}\left[\boldsymbol{\xi}_l^{(n,r,t)}\boldsymbol{\xi}_l^{(n,r,t)H}\right] \quad (10)$$

$$= \frac{1}{N_s}\mathrm{Tr}\left(\left(\mathbf{I}_{N_s} - \mathbf{S}\right)\mathbf{R}_{\boldsymbol{\alpha}_l}^{(0,r,t)}\left(\mathbf{I}_{N_s} - \mathbf{S}^H\right)\right) \quad (11)$$

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

$$[\mathbf{R}_{\boldsymbol{\alpha}_l}^{(s,r,t)}]_{k,m} = \sigma_{\alpha_l}^{(r,t)^2}J_0\left(2\pi f_d T_s(k-m+sN_s)\right) \quad (12)$$

Various traditional BEM designs have been reported to model the channel time-variations, e.g., the Complex Exponential BEM (CE-BEM) $[\mathbf{B}]_{k,m} = e^{j2\pi(\frac{k-N_g}{N_s})(m-\frac{N_c-1}{2})}$ which leads to a strictly banded frequency-domain matrix [17], the Generalized CE-BEM (GCE-BEM) $[\mathbf{B}]_{k,m} = e^{j2\pi(\frac{k-N_g}{av})(m-\frac{N_c-1}{2})}$ with $1 < a \le \frac{N_c-1}{2f_dT}$ which is a set of oversampled complex exponentials [16], the Polynomial BEM (P-BEM) $[\mathbf{B}]_{k,m} = (k-Ng)^m$ [10] and the Discrete Karhuen-Loeve BEM (DKL-BEM) which employs basis sequences that corresponds to the most significant eigenvectors of the autocorrelation matrix $\mathbf{R}_{\boldsymbol{\alpha}_l}^{(0,r,t)}$ [18]. 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, we obtain 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:

$$\mathbf{c}_n = \left[\mathbf{c}_n^{(1,1)^T}, ..., \mathbf{c}_n^{(1,N_T)^T}, ..., \mathbf{c}_n^{(N_R,N_T)^T}\right]^T \quad (14)$$

$$\mathbf{c}_n^{(r,t)} = \left[\mathbf{c}_0^{(n,r,t)^T}, ..., \mathbf{c}_{L^{(r,t)}-1}^{(n,r,t)^T}\right]^T \quad (15)$$

$$\boldsymbol{\mathcal{K}}_n(\boldsymbol{\nu}) = \mathrm{blkdiag}\left\{\boldsymbol{\mathcal{K}}_n^{(1)}(\boldsymbol{\nu}^{(1)}), ..., \boldsymbol{\mathcal{K}}_n^{(N_R)}(\boldsymbol{\nu}^{(N_R)})\right\} \quad (16)$$

$$\boldsymbol{\mathcal{K}}_n^{(r)}(\boldsymbol{\nu}^{(r)}) = \left[\boldsymbol{\mathcal{K}}_n^{(r,1)}(\nu^{(r,1)}), ..., \boldsymbol{\mathcal{K}}_n^{(r,N_T)}(\nu^{(r,N_T)})\right] \quad (17)$$

$$\boldsymbol{\mathcal{K}}_n^{(r,t)}(\nu^{(r,t)}) = \frac{1}{N}\left[\mathbf{Z}_0^{(n,r,t)}(\nu^{(r,t)}), ..., \mathbf{Z}_{L^{(r,t)}-1}^{(n,r,t)}(\nu^{(r,t)})\right] \quad (18)$$

$$\mathbf{Z}_l^{(n,r,t)}(\nu^{(r,t)}) = \left[\mathbf{M}_0^{(r,t)}(\nu^{(r,t)})\ \mathrm{diag}\{\mathbf{x}_n^{(t)}\}\ \mathbf{f}_l^{(r,t)}, ..., \mathbf{M}_{N_c-1}^{(r,t)}(\nu^{(r,t)})\ \mathrm{diag}\{\mathbf{x}_n^{(t)}\}\ \mathbf{f}_l^{(r,t)}\right] \quad (19)$$

where $\boldsymbol{\nu}^{(r)} = \left[\nu^{(r,1)}, \ldots, \nu^{(r,N_T)}\right]^T$. 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:

$$[\mathbf{F}^{(r,t)}]_{k,l} = e^{-j2\pi(\frac{k-1}{N}-\frac{1}{2})\tau_l^{(r,t)}}\ , \quad (20)$$

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

$$\left[\mathbf{M}_d^{(r,t)}(\nu^{(r,t)})\right]_{k,m} = \sum_{q=0}^{N-1}e^{j2\pi\frac{\nu^{(r,t)}q}{N}}[\mathbf{B}]_{q+N_g,d}\ e^{j2\pi\frac{m-k}{N}q}\ . \quad (21)$$

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

$$\mathbf{H}_n^{(r,t)} = \sum_{d=0}^{N_c-1}\mathbf{M}_d^{(r,t)}(\nu^{(r,t)})\mathrm{diag}\{\mathbf{F}^{(r,t)}\boldsymbol{\chi}_d^{(n,r,t)}\} \quad (22)$$

where $\boldsymbol{\chi}_d^{(n,r,t)} = \left[c_{(d,0)}^{(n,r,t)}, ..., c_{(d,L^{(r,t)}-1)}^{(n,r,t)}\right]^T$.

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:

$$\mathbf{R}_{\mathbf{c}_l}^{(s,r,t)} = \mathrm{E}[\mathbf{c}_l^{(n,r,t)}\mathbf{c}_l^{(n-s,r,t)H}]$$

$$= \left(\mathbf{B}^H\mathbf{B}\right)^{-1}\mathbf{B}^H\mathbf{R}_{\boldsymbol{\alpha}_l}^{(s,r,t)}\mathbf{B}\left(\mathbf{B}^H\mathbf{B}\right)^{-1} \quad (23)$$

Hence, the dynamics of $\mathbf{c}_l^{(n,r,t)}$ can be well modeled by an auto-regressive (AR) process [19] [20] [10] . A complex AR process of order p can be generated as:

$$\mathbf{c}_l^{(n,r,t)} = \sum_{i=1}^{p}\mathbf{A}^{(i)}\mathbf{c}_l^{(n-i,r,t)} + \mathbf{u}_l^{(n,r,t)} \quad (24)$$

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 [10], it is sufficient to choose $p = 1$ to correctly model the path complex gains. 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 defined as:

$$\mathbf{A} = \mathbf{R}_{\mathbf{c}_l}^{(1,r,t)} \left(\mathbf{R}_{\mathbf{c}_l}^{(0,r,t)}\right)^{-1} \quad (25)$$

$$\mathbf{U}_l^{(r,t)} = \mathbf{R}_{\mathbf{c}_l}^{(0,r,t)} + \mathbf{A}\mathbf{R}_{\mathbf{c}_l}^{(-1,r,t)} \quad (26)$$

Using (24), we obtain the AR model of order 1 for $\mathbf{c}_n$:

$$\mathbf{c}_n = \boldsymbol{\mathcal{A}}_{\mathbf{c}} \cdot \mathbf{c}_{n-1} + \mathbf{u}_{\mathbf{c}n} \quad (27)$$

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

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

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

$$\boldsymbol{\nu}_n = \boldsymbol{\mathcal{A}}_{\boldsymbol{\nu}} \cdot \boldsymbol{\nu}_{n-1} + \mathbf{u}_{\boldsymbol{\nu}n} \quad (28)$$

where the state transition matrix is of size $N_RN_T \times N_RN_T$. Since the CFOs can be assumed as constant during the observation interval, $\boldsymbol{\mathcal{A}}_{\boldsymbol{\nu}}$ is considered to be close to the identity matrix $\boldsymbol{\mathcal{A}}_{\boldsymbol{\nu}} = a\mathbf{I}_{N_RN_T}$, $a = 0.99$. 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 $\mathbf{U}_{\boldsymbol{\nu}} = \sigma_\nu^2 \mathbf{I}_{N_RN_T}$ where σ_ν^2 is the variance of the state noise associated with CFOs.

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 CFOs $\boldsymbol{\nu}_n$:

$$\boldsymbol{\mu}_n = \left[\mathbf{c}_n^T, \ \boldsymbol{\nu}_n^T\right]^T \quad (29)$$

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:

$$\boldsymbol{\mu}_n = \boldsymbol{\mathcal{A}} \cdot \boldsymbol{\mu}_{n-1} + \mathbf{u}_n \quad (30)$$

where the state transition matrix is defined as follows:

$$\boldsymbol{\mathcal{A}} = \text{blkdiag}\{\boldsymbol{\mathcal{A}}_{\mathbf{c}}, \ \boldsymbol{\mathcal{A}}_{\boldsymbol{\nu}}\} \quad (31)$$

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

D. Extended Kalman Filter (EKF)

The measurement equation (13) can be reformulated as:

$$\mathbf{y}_n = \mathbf{g}\left(\boldsymbol{\mu}_n\right) + \mathbf{w}_n \quad (32)$$

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 (32) is caused by CFOs. 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 the a priori and the a posteriori error estimate covariance matrix of size $LN_c + N_RN_T \times LN_c + N_RN_T$, respectively. 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:

$$\begin{aligned} \mathbf{P}_{(0|0)} &= \text{blkdiag}\left\{\mathbf{R}_{\mathbf{c}}^{(0)}, \ b\mathbf{I}_{N_RN_T}\right\} \quad (33) \\ \mathbf{R}_{\mathbf{c}}^{(s)} &= \text{blkdiag}\left\{\mathbf{R}_{\mathbf{c}}^{(s,1,1)}, ..., \mathbf{R}_{\mathbf{c}}^{(s,N_R,N_T)}\right\} \\ \mathbf{R}_{\mathbf{c}}^{(s,r,t)} &= \text{blkdiag}\left\{\mathbf{R}_{\mathbf{c}_0}^{(s,r,t)}, ..., \mathbf{R}_{\mathbf{c}_{L^{(r,t)}-1}}^{(s,r,t)}\right\} \end{aligned}$$

where $\mathbf{R}_{\mathbf{c}_l}^{(s,r,t)}$ is the correlation matrix of $\mathbf{c}_l^{(n,r,t)}$ defined in (23). 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 $\hat{\boldsymbol{\mu}}_{(n|n-1)}$:

$$\begin{aligned} \mathbf{G}_n &= \nabla_{\boldsymbol{\mu}_n}^T \mathbf{g}\left(\boldsymbol{\mu}_n\right)\Big|_{\boldsymbol{\mu}_n=\hat{\boldsymbol{\mu}}_{(n|n-1)}} = \\ &\left[\nabla_{\mathbf{c}_n}^T \mathbf{g}\left(\boldsymbol{\mu}_n\right)\Big|_{\boldsymbol{\mu}_n=\hat{\boldsymbol{\mu}}_{(n|n-1)}}, \ \nabla_{\boldsymbol{\nu}_n}^T \mathbf{g}\left(\boldsymbol{\mu}_n\right)\Big|_{\boldsymbol{\mu}_n=\hat{\boldsymbol{\mu}}_{(n|n-1)}}\right] \quad (34) \end{aligned}$$

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

$$\mathbf{G}_n = \left[\boldsymbol{\mathcal{K}}_n(\boldsymbol{\nu}_n)\big|_{\boldsymbol{\nu}_n=\hat{\boldsymbol{\nu}}_{(n|n-1)}}, \ \boldsymbol{\mathcal{V}}_n(\boldsymbol{\mu}_n)\big|_{\boldsymbol{\mu}_n=\hat{\boldsymbol{\mu}}_{(n|n-1)}}\right] \quad (35)$$

where

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

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

$$\left[\mathbf{M}_d'(\nu_n^{(r,t)})\right]_{k,m} = \sum_{q=0}^{N-1} j2\pi\frac{q}{N} e^{j2\pi\frac{\nu_n^{(r,t)}q}{N}} \left[\mathbf{B}\right]_{q+N_g,d} \ e^{j2\pi\frac{m-k}{N}q} \quad (36)$$

The EKF is a recursive algorithm composed of two stages: Time Update Equations and Measurement Update Equations. These two stages are defined as:

Time Update Equations:

$$\begin{aligned}\hat{\boldsymbol{\mu}}_{(n|n-1)} &= \boldsymbol{\mathcal{A}}\hat{\boldsymbol{\mu}}_{(n-1|n-1)} \\ \mathbf{P}_{(n|n-1)} &= \boldsymbol{\mathcal{A}}\mathbf{P}_{(n-1|n-1)}\boldsymbol{\mathcal{A}}^H + \mathbf{U}\end{aligned} \quad (37)$$

Measurement Update Equations:

$$\begin{aligned}\mathbf{K}_n &= \mathbf{P}_{(n|n-1)}\mathbf{G}_n^H\left(\mathbf{G}_n\mathbf{P}_{(n|n-1)}\mathbf{G}_n^H + N_T.\sigma^2\mathbf{I}_{N_R N}\right)^{-1} \\ \hat{\boldsymbol{\mu}}_{(n|n)} &= \hat{\boldsymbol{\mu}}_{(n|n-1)} + \mathbf{K}_n\left(\mathbf{y}_n - \mathbf{g}\left(\hat{\boldsymbol{\mu}}_{(n|n-1)}\right)\right) \\ \mathbf{P}_{(n|n)} &= \mathbf{P}_{(n|n-1)} - \mathbf{K}_n\mathbf{G}_n\mathbf{P}_{(n|n-1)}\end{aligned} \quad (38)$$

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 a predictor equations, while the Measurement Update Equations can be thought of a corrector equations.

IV. Joint Data Detection and EKF

In the iterative algorithm for joint data detection, channel and CFO extended Kalman estimation, the N_p pilots subcarriers are evently inserted into the N subcarriers at the positions $\mathcal{P} = \{p_r \mid p_r = (r-1)L_f + 1,\ r = 1, ..., N_p\}$, where L_f is the distance between two adjacent pilots. We use the the QR-equalizer [10] for the data detection. The QR-equalizer allows us to estimate the data symbol with free ICI by performing a so-called QR-decomposition. The algorithm proceeds as follows:

initialization:

- $\hat{\boldsymbol{\mu}}_{(0|0)} = \mathbf{0}_{LN_c+N_R N_T,1}$
- *compute* $\mathbf{P}_{(0|0)}$ *as* (33)

$\boldsymbol{n \leftarrow n+1}$ **:**

- *execute the Time Update Equations of EKF* (37)
- *compute the channel matrix by substituting* $\boldsymbol{\mu}_n$ *with the prediction parameters* $\hat{\boldsymbol{\mu}}_{(n|n-1)}$ *in* (22)
- *recursion:*$i \leftarrow 1$
  - *remove the pilot ICI from the received data subcarriers*
  - *Detection of data symbols*
  - *execute the Measurement Update Equations of EKF* (38)
  - *compute the channel matrix using* (22) *with the updated parameters*
  - $i \leftarrow i+1$

where i represents the iteration number.

V. Simulation

In this section, the performance of our recursive algorithm is evaluated in terms of Mean Square Error (MSE) for joint channel and CFO estimation and in terms of Bit Error Rate (BER) for data detection.

We assume that all the (r,t) channel branches, $r = 1, \ldots, N_R,\ t = 1, \ldots, N_T$ have the same path delays and fading properties ($i.e.$, the same number of paths, of ${\sigma_{\alpha_l}^{(r,t)}}^2$ and $\tau_l^{(r,t)}$). This is understood when the antennas are very close to each other, which is typical in practice. The Rayleigh channel model given in [10] [12]($L^{(r,t)} = 6$ paths and maximum delay $\tau_{max} = 10T_s$) was chosen. A normalized 4QAM MIMO-OFDM system, with $N_T = N_R = 2$, $N = 128$ 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. The MSE and the BER were evaluated under a rapid time-varying channel with $f_d T = 0.1$ (corresponding to a vehicle speed of $600km/h$ at $f_c = 2.5GHz$). A GCE-BEM with $N_c = 3$ was chosen to model the path complex gains of the channel. Most advanced technologies have an oscillator frequency tolerance less than 1 ppm ($i.e.$ $\nu = 0.16$ in normalized units with the given parameters). For the simulation, we chose the configuration where each transmitter and receiver requires its own RF-IF chain, which is the configuration discussed in this article. For this scenario, the number of CFO parameters to be estimated ($N_T N_R = 4$) is the largest. Therefore, this is the most pessimistic configuration. The CFO values were arbitrarily chosen as $\nu^{(0,0)} = 0.1$, $\nu^{(0,1)} = -0.1$, $\nu^{(1,0)} = 0.05$ and $\nu^{(1,1)} = -0.07$.

Fig. 1 shows the MSE of the channel complex gain and the MSE of the normalized CFO as a function of E_b/N_0. Seven iterations have been carried out. For reference, the MSEs obtained in Data Aided (DA) mode (knowledge of the data symbols) have been plotted. As expected, the MSEs obtained in Data Aided mode are lower than the MSEs obtained with just the pilots, especialy at low E_b/N_0 where the detection errors are the most important. However, for $E_b/N_0 \geq 20$ dB, the MSE has a floor (especially the channel complex gain MSEs). This is due to the fact that beyond 20 dB, the matrix to be inversed in Eq. (38) becomes badly scaled.

Fig. 2 gives the BER performance of our proposed iterative algorithm. For reference, we also plotted BERs obtained with perfect knowledge of channel response and CFO. It is shown that one iteration is sufficient to approach the reference curve with perfect knowledge of channel response and carrier frequency offset.

VI. Conclusion

A new iterative algorithm which jointly estimates multipath complex gain and CFO in MIMO environment has been presented. The algorithm is based on a parametric channel model. Extended Kalman filtering is used for parameter estimation and the data recovery is carried out by means of a QR-equalizer. Simulation results show that by estimating and removing the ICI at each iteration, the BER is greatly improved, especially after the first iteration. Our algorithm needs only one iteration to approach the performance of the ideal case for which the knowledge of the channel response and CFO is available.

Ackowledgement

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

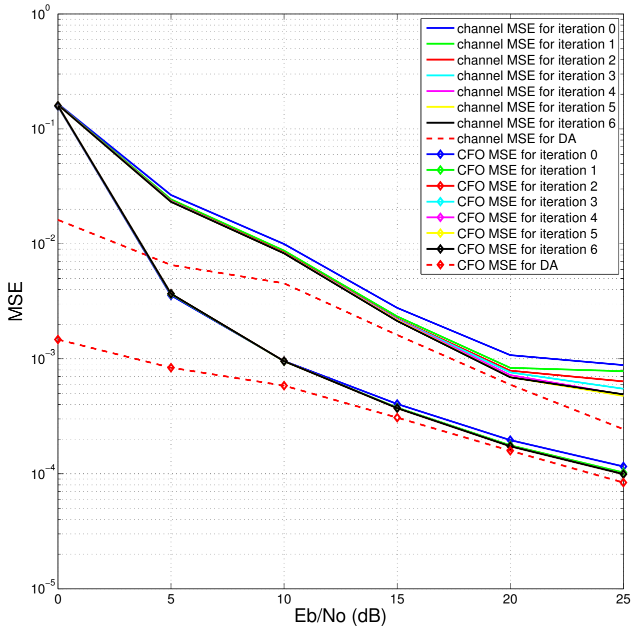

Fig. 1. Mean Square Error (MSE) as a function of E_b/N_0 for $f_dT = 0.1$

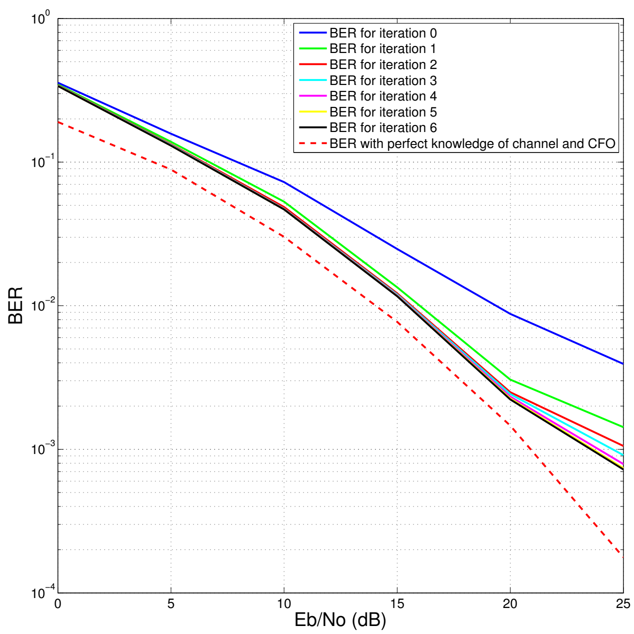

Fig. 2. Bit Error Rate (BER) as a function of E_b/N_0 for $f_dT = 0.1$

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 COMMUNICATIONS Conf.*, 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] T. Roman, M. Enescu, and V. Koivunen, "Joint Time-Domain Tracking of Channel and Frequency Offsets for MIMO OFDM Systems," *Wireless Personal Communications*, vol. 31, pp. 181–200, 2004.

[10] 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.

[11] 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.

[12] H. Hijazi, E. P. Simon, M. Lienard, and L. Ros, "Channel Estimation for MIMO-OFDM Systems in Fast Time-Varying Environments," in *ISCCSP*, 2010, accepted.

[13] 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.

[14] 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.

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

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

[17] K. D. Teo and S. Ohno, "Optimal MMSE Finite Parameter Model for Doubly-selective Channels," in *Proc. IEEE GLOBAL COMMUNICATIONS Conf.*, 2005.

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

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

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