EM-Based ML Estimation of Fast Time-Varying Multipath Channels for SIMO OFDM Systems

Souheib Ben Amor[1], Sofiène Affes[1], and Faouzi Bellili[2]
[1]INRS-EMT, Université du Québec, Montréal, QC, Canada, Emails: {souheib.ben.amor, affes}@emt.inrs.ca
[2]Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB, Canada, Email: faouzi.bellili@umanitoba.ca

***Abstract*—This paper investigates the problem of fast time-varying multipath channel estimation over single-input multiple-output orthogonal frequency-division multiplexing (SIMO OFDM)-type transmissions. We do so by tracking the variations of each complex gain coefficient using a polynomial-in-time expansion. To that end, we derive the log-likelihood function (LLF) in both the data-aided (DA) and non-data-aided (NDA) case. The DA ML estimates are found in closed-form expressions and then used to initialize the expectation maximization (EM) algorithm that is used to iteratively maximize the LLF in the NDA case. We also introduce an alternative initialization procedure that requires less pilot symbols as compared to the DA ML-based solution without incurring a significant performance loss. Simulation results show that the proposed EM-based estimator converges within few iterations providing accurate estimates for all multipath gains, thereby resulting in significant BER gain as compared to the DA least square (LS) technique.**

***Index Terms*—Maximum likelihood (ML), expectation maximization (EM), channel estimation, time-varying channel (TVC), OFDM.**

I. Introduction

Orthogonal frequency-division multiplexing (OFDM) showed its effectiveness in current 4^{th} generation wireless technology (4G). A scalable variety of CP-OFDM is already included in 5^{th} generation (5G) new radio (NR) standards by the 3^{rd} Generation Partnership Project (3GPP) [1]. The adopted waveform will include multiple sub-carrier spacings that depend on the type of deployments and service requirements. Moreover, when coupled with the large-scale antenna technology OFDM is poised to enable the 1000-fold increase in capacity that is required over the next few years. Despite its attractive features such as robustness to frequency selective channels and spatial diversity, OFDM-type radio interface technologies (RITs) are already very sensitive to channel time variations since the latter break the crucial orthogonality between the subcarriers thereby introducing the so-called inter-carrier interference (ICI). Accurate channel estimation, hence, becomes a daunting task at very high mobility [2].

So far, a number of channel estimation techniques have been reported in the literature. They can be categorized in two major categories: $i)$ the data-aided (DA) approaches where the transmitted symbols are assumed to be perfectly known at the receiver. They provide highly-accurate channel estimates performance at a significant cost, however, in terms of overhead; $ii)$ the blind or non-data-aided (NDA) approaches where the receiver has no *a priori* information about the transmitted data. Therefore, NDA techniques do not incur any overhead at the cost, however, of reduced accuracy.

For fast time-varying channels, most of the DA techniques rely on a basis expansion model (BEM) to estimate the equivalent discrete-time channel taps [3-5]. In fact, BEM methods such as Karhunen-Loeve BEM were designed with low mean square error (MSE) [3]. They are, however, sensitive to statistical channel mismatch. The complex-exponential BEM, also proposed in [3], does not make use of the channel statistics but suffers from large modeling errors. The polynomial BEM (P-BEM) investigated in [4] yields accurate channel estimates, but only at low Dopplers. In [6], the complex gain variations of each path was approximated by a polynomial function of time then estimated by least squares (LS) technique. This solution offers accurate performance even at high Doppler. However, it requires that the number of paths to be smaller than the inserted pilot symbols in each OFDM time slot. Moreover, since it was derived for the SISO case, a straightforward extension to SIMO, where the LS technique is applied at the output of each antenna element, does not take advantage of some of the interesting features of a multi-antenna configuration. Indeed, it is possible to combine the copies of the antenna elements to provide a more accurate detection performance.

Under the NDA category, time-varying channel estimation was also investigated in [7]. The authors used the discrete Legendre polynomial BEM along with the space alternating generalized EM-maximum a posteriori probability (SAGE-MAP) technique to estimate the time-domain channel coefficients of OFDM channels. In [8] the EM algorithm was also used to estimate the instantaneous SNR which is obtained by estimating the time-varying channel itself but still for SISO systems only. Both the EM and LS techniques were also leveraged in [9] and [10], respectively, to estimate the SNR over single-carrier SIMO systems. In [11,12], iterative channel estimation using Kalman filter and QR-detection was first investigated under SISO multi-carrier and later generalized to multiple-input multiple-output (MIMO) OFDM systems. Its performance was further enhanced in [13] by exploiting the statistics of the channel estimation errors in the iterative estimation process. However, the kalman filter-based techniques require perfect knowledge of the doppler as well as the power-delay profile. Moreover, a high number of pilots per OFDM block is needed to obtain accurate estimates thereby affecting the overall throughput of the system.

In this paper, we develop an iterative EM-based maximum

Work supported by the DG and CREATE PERSWADE <www.create-perswade.ca> Programs of NSERC, and a Discovery Accelerator Supplement Award from NSERC.

978-1-7281-0962-6/19/$31.00 ©2019 IEEE

likelihood (ML) estimator of fast time-varying channels over SIMO OFDM-type radio interfaces. By relying on the polynomial approximation of the multipath channel gains [6] and resorting to the powerful EM technique, instead of the LS approach, our solution offers a more accurate ML-type acquisition of the polynomial expansion coefficients and hence the resulting time-varying channel gains. To avoid local convergence that is inherent to iterative algorithms, we initialize the EM algorithm with the DA ML technique also developed in this paper. In addition, as a byproduct, the proposed solution yields, MAP-based soft estimates of the unknown symbols. The latter are leveraged to devise a dedicated ICI cancellation scheme that works side by side with the EM-based time-varying estimator according to the turbo principle. Furthermore, we introduce another initialization procedure that also applies when the number of paths exceeds the number of available pilot observations. This renders the proposed solution robust to the rapid changes in propagation environment where the number of paths can change unpredictability due the motion of mobile users.

The rest of the paper is organized as follows: In Section II, we introduce the system model. In Section III, we derive the new NDA EM-base ML solution for the underlying estimation problem. In Section IV, we provide the required details regarding the initialization procedure. In Section V, we use exhaustive computer simulations to assess the performance of the proposed fast time-varying channel estimator. Finally, we draw out some concluding remarks in Section VI.

The notations adopted in this paper are as follows. Vectors and matrices are represented in lower- and upper-case bold fonts, respectively. Moreover, $\{.\}^T$ and $\{.\}^H$ denote the conjugate and Hermitian (i.e., transpose conjugate) operators. The Euclidean norm of any vector is denoted as $\|.\|$. For any matrix $\mathbf{X}$, $[\mathbf{X}]_q$ and $[\mathbf{X}]_{l,k}$ denote its q^{th} column and $(l,k)^{th}$ entry, respectively. For any vector $\mathbf{x}$, diag$\{\mathbf{x}\}$ refers to the diagonal matrix whose elements are those of $\mathbf{x}$. Moreover, $\{.\}^*$, $\angle\{.\}$, and $|.|$ return the conjugate, angle, and modulus of any complex number, respectively. Finally, $\mathbb{E}\{.\}$ stands for the statistical expectation, j is the pure imaginary number (i.e., $j^2=-1$), and the notation $\triangleq$ is used for definitions.

II. System Model

Consider a SIMO OFDM system with N_r receiving antenna elements, N subcarriers, and a cyclic prefix (CP) of a length N_{cp}. The wireless link between the transmitter and the $\{r^{th}\}_{r=1}^{N_r}$ antennas is modeled as a multipath fading channel as follows:

$$h_r(t,\tau)=\sum_{l=1}^{L_r}\alpha_{l,r}(t)\delta(\tau-\tau_{l,r}T_s), \tag{1}$$

where L_r is the number of paths of the r^{th} wireless link. For each path, the delay $\tau_{l,r}$ is normalized by the sampling period T_s and the complex gain $\alpha_{l,r}(t)$ is modeled by a Rayleigh random variable with zero mean and a variance $\sigma^2_{l,r}$. The multipath power profile (i.e., the channel) is assumed to be normalized (i.e., $\sum_{l=1}^{L_r}\sigma^2_{l,r}=1$). For each of the N_r links, we approximate the sampled complex gain of the l^{th} path within the duration of N_c consecutive OFDM blocks, $\boldsymbol{\alpha}_{l,r}=[\alpha_{l,r}(-N_{cp}T_s),\ldots,\alpha_{l,r}(N_bN_c-N_{cp}-1)]^T$, by a polynomial of order N_c-1 as follows [6]:

$$\alpha_{l,r}(pT_s)\approx\sum_{d=1}^{N_c}c_{d,l,r}p^{(d-1)}+\zeta_{l,r}[p], \tag{2}$$

where $p\in[-N_{cp},-N_{cp}+1,\ldots,N_bN_c-N_{cp}-1]$. Moreover, $\mathbf{c}_{l,r}=[c_{1,l,r},c_{2,l,r},\ldots,c_{N_c,l,r}]^T$ gathers the approximating polynomial coefficients corresponding to the l^{th} path between the transmitter and the r^{th} receiving antenna while $\zeta_{l,r}[p]$ is the approximation error. $T=N_bT_s$ denotes the OFDM block duration where $N_b=N+N_{cp}$. At the destination, after removing the CP and applying a $N-$point fast Fourier transform (FFT), the collected OFDM symbols at each local approximation window of N_c OFDM blocks (i.e., $k=1,2,\ldots,N_c$), over the r^{th} antenna element, can be written as follows:

$$\tilde{\mathbf{y}}_{k,r} \;=\; \mathbf{H}_{k,r}\mathbf{a}_k+\mathbf{w}_{k,r}, \tag{3}$$

where $\tilde{\mathbf{y}}_{k,r}=\big[y_{k,r}[1],y_{k,r}[2],\ldots,y_{k,r}[N]\big]^T$ is the received k^{th} OFDM block, and $\mathbf{w}_{k,r}=\big[w_{k,r}[1],w_{k,r}[2],\ldots,w_{k,r}[N]\big]^T$ is the complex white Gaussian noise vector with covariance $\sigma^2\mathbf{I}_N$ where $\mathbf{I}_N$ is the N-dimensional identity matrix. The N transmitted symbols during the k^{th} OFDM block, $\mathbf{a}_k=\big[a_k[1],a_k[2],\ldots,a_k[N]\big]^T$, are generated randomly from a $M-$ary constellation alphabet, denoted $\mathcal{C}^M$, and are assumed equally likely, i.e., $\{P_r(a_m)=\frac{1}{M}\}_{a_m\in\mathcal{C}^M}$. The $N\times N$ matrix, $\mathbf{H}_{k,r}$, is the channel frequency response whose elements are given by:

$$[\mathbf{H}_{k,r}]_{m,n}=\frac{1}{N}\sum_{l=1}^{L_r}\left[e^{-j2\pi\left(\frac{n-1}{N}-\frac{1}{2}\right)\tau_{l,r}}\sum_{q=0}^{N-1}\alpha_{k,l,r}(qT_s)e^{j2\pi\frac{n-m}{N}q}\right], \tag{4}$$

where $\{\alpha_{k,l,r}(qT_s)\}_{q=kN_b}^{kN_b+N-1}$ are the samples corresponding to the l^{th} path within the duration of the k^{th} OFDM block over the r^{th} receiving antenna. As shown in [6], with the above approximation [6], the polynomial coefficients, $\mathbf{c}_{l,r}$ can be obtained using the time average of the channel gain over the effective duration of each OFDM time slot ($\{\bar{\alpha}_{k,l,r}=\frac{1}{N}\sum_{q=kN_b}^{kN_b+N-1}\alpha_{k,l,r}(qT_s)\}_{k=0}^{N_c-1}$) as follows:

$$\mathbf{c}_{l,r}=\mathbf{T}^{-1}\bar{\boldsymbol{\alpha}}_{l,r}, \tag{5}$$

where $\bar{\boldsymbol{\alpha}}_{l,r}=[\bar{\alpha}_{1,l,r},\bar{\alpha}_{2,l,r},\ldots,\bar{\alpha}_{N_c,l,r}]^T$ and $\mathbf{T}$ is a $(N_c\times N_c)$ matrix given by:

$$\mathbf{T}=\begin{pmatrix}1 & \frac{N-1}{2} & \frac{(N-1)(2N-1)}{6}\\ 1 & \frac{N-1}{2}+N_b & \frac{(N-1)(2N-1)}{6}+(N-1)N_b+N_b^2\\ 1 & \frac{N-1}{2}+2N_b & \frac{(N-1)(2N-1)}{6}+2(N-1)N_b+4N_b^2\end{pmatrix}.$$

Using these coefficients, the samples of the complex gain of each channel path over the interval $[-N_{cp},\ldots,N_bN_c-N_{cp}-1]$, $\mathbf{c}_l=[c_{1,l,r},c_{2,l,r},\ldots,c_{N_c,l,r}]$, can be obtained as follows:

$$\boldsymbol{\alpha}_{l,r} \;=\; \mathbf{S}^T\mathbf{c}_{l,r}, \tag{6}$$

where $\mathbf{S}$ is a $(N_c \times N_b N_c)$ matrix whose elements are given by:

$$\Big\{\big\{[\mathbf{S}]_{d,p'} = (p' - N_{cp} - 1)^{d-1}\big\}_{p'=1}^{N_b N_c}\Big\}_{d=1}^{N_c}. \tag{7}$$

The channel gains can be estimated using (6) from the channel coefficient estimates whose estimation in (5) ultimately requires an estimate for the channel gain time averages vector $\bar{\boldsymbol{\alpha}}_{l,r}$.

III. New NDA EM-based ML Path Gains Estimation

We start by stacking the received samples at the output of all the antenna elements, $\Big\{\big\{\{y_{k,r}(n)\}_{n=1}^{N}\big\}_{k=0}^{N_c-1}\Big\}_{r=1}^{N_r}$, into vectors $\Big\{\big\{\mathbf{y}_k(n) = [y_{k,1}(n), y_{k,2}(n), \ldots, y_{k,N_r}(n)]^T\big\}_{n=1}^{N}\Big\}_{k=0}^{N_c-1}$. We also define $\bar{\boldsymbol{\varphi}}_k = [\bar{\boldsymbol{\varphi}}_{k,1}^T, \bar{\boldsymbol{\varphi}}_{k,2}^T, \ldots, \bar{\boldsymbol{\varphi}}_{k,N_r}^T]$ as the vectors containing all the time averages of the channel gains of all $\{L_r\}_{r=1}^{N_r}$ paths with $\{\bar{\boldsymbol{\varphi}}_{k,r} = [\bar{\alpha}_{k,1}, \bar{\alpha}_{k,2}, \ldots, \bar{\alpha}_{k,L_r}]^T\}_{r=1}^{N_r}$. The probability density function (pdf) of the received samples $\{\{\mathbf{y}_k(n)\}_{n=1}^{N}\}_{k=0}^{N_c-1}$ conditioned on the transmitted symbol $a_k[n]$ and parametrized by $\boldsymbol{\psi}_k = \left[\bar{\boldsymbol{\varphi}}_k^T, \sigma^2\right]^T$, is expressed as follows:

$$p(\mathbf{y}_k(n)|a_k[n] = a_m; \boldsymbol{\psi}_k) = \frac{1}{(2\pi\sigma^2)^{N_r}} \exp\left\{\frac{-1}{2\sigma^2}\sum_{r=1}^{N_r}\left|y_{k,r}(n) - a_m[\mathbf{H}_{k,r}]_{n,n}\right|^2\right\}, \tag{8}$$

where:

$$[\mathbf{H}_{k,r}]_{n,n} = \frac{1}{N}\sum_{l=1}^{L_r}\left[e^{-j2\pi\left(\frac{n-1}{N}-\frac{1}{2}\right)\tau_{l,r}}\sum_{q=0}^{N-1}\alpha_{l,k,r}(qT_s)\right], \tag{9}$$

Note that, for the time being, we absorb the effect of the ICI in the additive noise and we also assume that normalized delays,$\{\tau_{l,r}\}_{l=1}^{L_r}$, are perfectly known to the receiver. The n^{th} diagonal element of the matrix $\mathbf{H}_{k,r}$ in (9) can then be written as follows:

$$[\mathbf{H}_{k,r}]_{n,n} = \bar{\boldsymbol{\varphi}}_{k,r}^T \mathbf{F}_{n,r}, \tag{10}$$

where $\mathbf{F}_{n,r}$ is a vector containing the elements of the m^{th} row of the $(N \times L_r)$ matrix $\mathbf{F}_r$ which is defined as:

$$[\mathbf{F}_r]_{m,l} = e^{-j2\pi\left(\frac{m-1}{N}-\frac{1}{2}\right)\tau_{l,r}}. \tag{11}$$

By injecting (10) back into (8) and averaging the result over the alphabet, the pdf of the received samples can be written as follows:

$$p(\mathbf{y}_k(n);\boldsymbol{\psi}_k) = \sum_{m=1}^{M} P_r(a_m)\, p(\mathbf{y}_k(n)|a_k[n] = a_m; \boldsymbol{\psi}_k). \tag{12}$$

As mentioned earlier, the transmitted symbols are generated from a normalized $M-$ary constellation (i.e., PAM, PSK or QAM). It follows that:

$$p(\mathbf{y}_k(n);\boldsymbol{\psi}_k) = \frac{1}{M(2\pi\sigma^2)^{N_r}} \times \sum_{m=1}^{M}\exp\left\{-\frac{1}{2\sigma^2}\sum_{r=1}^{N_r}\left|y_{k,r}(n) - a_m\bar{\boldsymbol{\varphi}}_{k,r}^T\mathbf{F}_{n,r}\right|^2\right\}. \tag{13}$$

It is obvious at this stage that maximizing (13) with respect to $\boldsymbol{\psi}_k$ is analytically intractable. Thus, we will resort to the EM concept to find the maximum of the multidimensional likelihood function (LF). First, we define the log-LF (LLF), $\mathcal{L}(\boldsymbol{\psi}_k|a_k[n] = a_m) \triangleq \ln(\mathbf{y}_k(n)|a_k[n] = a_m; \boldsymbol{\psi}_k)$, of $\mathbf{y}_k(n)$ conditioned on the transmitted symbol $a_k[n]$ for the k^{th} OFDM symbol which can be written as:

$$\mathcal{L}(\boldsymbol{\psi}_k|a_k[n]=a_m) = -N_r\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\left(\sum_{r=1}^{N_r}|y_{k,r}(n)|^2 + \left|a_m\bar{\boldsymbol{\varphi}}_{k,r}^T\mathbf{F}_{n,r}\right|^2 - 2\Re\{y_{k,r}(n)^* a_m\bar{\boldsymbol{\varphi}}_{k,r}^T\mathbf{F}_{n,r}\}\right). \tag{14}$$

During the "expect step (E-STEP)" of the EM algorithm, we compute the expectation of the LLF in (14) over all possible transmitted symbols, $\{a_m\}_{m=1}^{M}$, using the previous estimates of the underlying unknown parameters. Then, the resulting expectation is maximized with respect to the unknown coefficient $\boldsymbol{\psi}_k$ during the "*Maximization step* (M-STEP)". Starting with an initial guess, $\widehat{\boldsymbol{\psi}}_k^{(0)}$, of the channel estimates, the cost function to be maximized during the M-STEP at the i^{th} EM iteration is given by:

$$\boldsymbol{Q}\Big(\boldsymbol{\psi}_k|\widehat{\boldsymbol{\psi}}_k^{(i-1)}\Big) = \textstyle\sum_{n=1}^{N} E_{a_m}\Big\{\mathcal{L}(\boldsymbol{\psi}_k|a_k[n] = a_m)\Big|\mathbf{y}_k(n); \widehat{\boldsymbol{\psi}}_k^{(i-1)}\Big\}, \tag{15}$$

where $E_{a_m}\{.\}$ denotes the expectation over all possible transmitted symbols $\{a_m\}_{m=1}^{M}$ and $\widehat{\boldsymbol{\psi}}_k^{(i-1)} = \left[\widehat{\bar{\boldsymbol{\varphi}}}_k^{(i-1)T}, \widehat{\sigma^2}_k^{(i-1)}\right]^T$ contains the estimates of $\boldsymbol{\psi}_k$ and the noise variance at the $(i-1)^{th}$ EM iteration. The expression in (15) can be further simplified as follows:

$$\boldsymbol{Q}\left(\boldsymbol{\psi}_k|\widehat{\boldsymbol{\psi}}_k^{(i-1)}\right) = -NN_r\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\left(\textstyle\sum_{r=1}^{N_r} Z_{k,r} + \sum_{n=1}^{N}\gamma_{n,k}^{(i-1)}\left|\bar{\boldsymbol{\varphi}}_{k,r}^T\mathbf{F}_{n,r}\right|^2 - 2\beta_{n,k,r}^{(i-1)}\right), \tag{16}$$

where[1]:

$$Z_{k,r} = \textstyle\sum_{n=1}^{N}|y_{k,r}(n)|^2, \tag{17}$$

$$\gamma_{n,k}^{(i-1)} = E_{a_m}\Big\{|a_m|^2\big|\mathbf{y}_k(n); \widehat{\boldsymbol{\psi}}_k^{(i-1)}\Big\}, \tag{18}$$

$$\beta_{n,k,r}^{(i-1)} = E_{a_m}\Big\{\Re\{y_{k,r}(n)^* a_m\bar{\boldsymbol{\varphi}}_{k,r}^T\mathbf{F}_{n,r}\}\big|\mathbf{y}_k(n); \widehat{\boldsymbol{\psi}}_k^{(i-1)}\Big\}. \tag{19}$$

Using the Bayes formula, the a posteriori probability of a_m, $P_{m,n,k}^{(i-1)} = P_r\left(a_m|\mathbf{y}_k(n); \widehat{\boldsymbol{\psi}}_k^{(i-1)}\right)$, at the $(i-1)^{th}$ iteration is given by:

$$P_r\left(a_m|\mathbf{y}_k(n); \widehat{\boldsymbol{\psi}}_k^{(i-1)}\right) = \frac{P_r(a_m)P\left(\mathbf{y}_k(n)|a_m; \widehat{\boldsymbol{\psi}}_k^{(i-1)}\right)}{P\left(\mathbf{y}_k(n); \widehat{\boldsymbol{\psi}}_k^{(i-1)}\right)}. \tag{20}$$

Since the transmitted symbols are equiprobable (i.e., $P_r(a_m) = \frac{1}{M}$), we have the following result:

$$P\left(\mathbf{y}_k(n); \widehat{\boldsymbol{\psi}}_k^{(i-1)}\right) = \frac{1}{M}\textstyle\sum_{n=1}^{N} P\left(\mathbf{y}_k(n)|a_m; \widehat{\boldsymbol{\psi}}_k^{(i-1)}\right). \tag{21}$$

[1]For the particular case of normalized-energy constant-envelope constellations, note that we have $\gamma_{n,k}^{(i-1)} = 1$.

Exploiting the fact that $\bar{\varphi}_{k,r} = \Re\{\bar{\varphi}_{k,r}\} + \Im\{\bar{\varphi}_{k,r}\}$ and $\mathbf{F}_{n,r} = \Re\{\mathbf{F}_{n,r}\} + \Im\{\mathbf{F}_{n,r}\}$, the cost function in (16) can be written as follows:

$$\boldsymbol{Q}\Big(\boldsymbol{\psi}_k|\widehat{\boldsymbol{\psi}}_k^{(i-1)}\Big) = -NN_r\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\Bigg(\sum_{r=1}^{N_r} Z_{k,r} + \sum_{n=1}^{N}\gamma_{n,k}^{(i-1)} \times\Bigg(\mathbf{F}_{n,r}^H\mathbf{G}_{1,k,r}\mathbf{F}_{n,r} + \Im\{\mathbf{F}_{n,r}\}^T\mathbf{G}_{2,k,r}\Re\{\mathbf{F}_{n,r}\} + \Re\{\mathbf{F}_{n,r}\}^T\mathbf{G}_{3,k,r}\Im\{\mathbf{F}_{n,r}\}\Bigg) - 2\sum_{m=1}^{M} P_{m,n,k}^{(i-1)}\eta_{k,n,r}^{(m)}\Bigg), \quad (22)$$

where:

$$\begin{aligned}
\mathbf{G}_{1,k,r} &= \Re\{\bar{\varphi}_{k,r}\}\Re\{\bar{\varphi}_{k,r}\}^T + \Im\{\bar{\varphi}_{k,r}\}\Im\{\bar{\varphi}_{k,r}\}^T,\\
\mathbf{G}_{2,k,r} &= \Re\{\bar{\varphi}_{k,r}\}\Im\{\bar{\varphi}_{k,r}\}^T - \Im\{\bar{\varphi}_{k,r}\}\Re\{\bar{\varphi}_{k,r}\}^T,\\
\mathbf{G}_{3,k,r} &= \Im\{\bar{\varphi}_{k,r}\}\Re\{\bar{\varphi}_{k,r}\}^T - \Re\{\bar{\varphi}_{k,r}\}\Im\{\bar{\varphi}_{k,r}\}^T,\\
\eta_{k,n,r}^{(m)} &= \Re\{y_{k,r}(n)^* a_m \mathbf{F}_{n,r}^T\}\Re\{\bar{\varphi}_{k,r}\}\\
&\quad - \Im\{y_{k,r}(n)^* a_m \mathbf{F}_{n,r}^T\}\Im\{\bar{\varphi}_{k,r}\}. \quad (23)
\end{aligned}$$

As per the M-STEP, we differentiate the cost function in (22) with respect to $\Re\{\bar{\varphi}_{k,r}\}$ and $\Im\{\bar{\varphi}_{k,r}\}$ and set the result to zero to obtain the following results:

$$\sum_{n=1}^{N}\gamma_{n,k}^{(i-1)}\Big(\mathbf{J}_{1,n,r}\Re\{\bar{\varphi}_{k,r}\} - \mathbf{J}_{2,n,r}\Im\{\bar{\varphi}_{k,r}\}\Big) = \sum_{n=1}^{N}\boldsymbol{\mu}_{1,n,k,r},$$

$$\sum_{n=1}^{N}\gamma_{n,k}^{(i-1)}\Big(\mathbf{J}_{1,n,r}\Im\{\bar{\varphi}_{k,r}\} + \mathbf{J}_{2,n,r}\Re\{\bar{\varphi}_{k,r}\}\Big) = -\sum_{n=1}^{N}\boldsymbol{\mu}_{2,n,k,r},$$

where:

$$\begin{aligned}
\mathbf{J}_{1,n,r} &= \Re\{\mathbf{F}_{n,r}\}\Re\{\mathbf{F}_{n,r}\}^T + \Im\{\mathbf{F}_{n,r}\}\Im\{\mathbf{F}_{n,r}\}^T,\\
\mathbf{J}_{2,n,r} &= \Re\{\mathbf{F}_{n,r}\}\Im\{\mathbf{F}_{n,r}\}^T - \Im\{\mathbf{F}_{n,r}\}\Re\{\mathbf{F}_{n,r}\}^T,\\
\boldsymbol{\mu}_{1,n,k,r} &= \textstyle\sum_{m=1}^{M} P_{m,n,k}^{(i-1)}\Re\{y_{k,r}(n)^* a_m \mathbf{F}_{n,r}^T\},\\
\boldsymbol{\mu}_{2,n,k,r} &= \textstyle\sum_{m=1}^{M} P_{m,n,k}^{(i-1)}\Im\{y_{k,r}(n)^* a_m \mathbf{F}_{n,r}^T\}.
\end{aligned}$$

Now, using the identity $\bar{\varphi}_{k,r} = \Re\{\bar{\varphi}_{k,r}\} + j\Im\{\bar{\varphi}_{k,r}\}$ leads to:

$$\sum_{n=1}^{N}(\mathbf{J}_{1,n,r} + j\mathbf{J}_{2,n,r})\gamma_{n,k}^{(i-1)}\bar{\varphi}_{k,r} = \sum_{n=1}^{N}\boldsymbol{\mu}_{1,n,r} - j\boldsymbol{\mu}_{2,n,r}. \quad (24)$$

Hence, the i^{th} EM update for time average of the channel gains at the i^{th} iteration can be obtained as follows:

$$\widehat{\bar{\varphi}}_{k,r}^{(i)} = \Big(\textstyle\sum_{n=1}^{N}\gamma_{n,k}^{(i-1)}(\mathbf{J}_{1,n,r} + j\mathbf{J}_{2,n,r})\Big)^{-1} \times \textstyle\sum_{n=1}^{N}\Big(\sum_{m=1}^{M} P_{m,n,k}^{(i-1)} y_{k,r}^*(n) a_m \mathbf{F}_{n,r}^T\Big)^H. \quad (25)$$

Similarly, by differentiating the cost function in (22) with respect to σ^2 and setting the result to zero, we obtain the following update for the noise variance:

$$\widehat{\sigma^2}^{(i)} = \frac{\sum_{r=1}^{N_r} Z_{k,r} + \sum_{n=1}^{N}\Big|\mathbf{F}_{n,r}^T\widehat{\bar{\varphi}}_{k,r}^{(i-1)}\Big|^2\gamma_{n,k}^{(i-1)} - 2\beta_{n,k,r}^{(i-1)}}{2NN_r}. \quad (26)$$

Finally, after $\mathcal{I}_{\text{EM}}$ iterations of the EM algorithm, the channel estimates, corresponding to N_c consecutive OFDM symbols over the r^{th} antenna element, are obtained as follows:

$$\widehat{\boldsymbol{\alpha}}_{l,r} = \mathbf{S}^T\widehat{\mathbf{c}}_{l,r} = \mathbf{S}^T\mathbf{T}^{-1}\widehat{\bar{\boldsymbol{\alpha}}}_{l,r}^{(\mathcal{I}_{EM})}, \quad (27)$$

where $\widehat{\bar{\boldsymbol{\alpha}}}_{l,r}^{(\mathcal{I}_{\text{EM}})} = [\widehat{\bar{\alpha}}_{1,l,r}^{(\mathcal{I}_{\text{EM}})}, \widehat{\bar{\alpha}}_{2,l,r}^{(\mathcal{I}_{\text{EM}})}, \ldots, \widehat{\bar{\alpha}}_{N_c,l,r}^{(\mathcal{I}_{\text{EM}})}]^T$ is the EM-based ML estimate of the complex channel gain time averages of the l^{th} path over N_c OFDM data symbols. The channel gain estimates in (27) can be further improved by implementing an iterative ICI suppression technique. Indeed, the channel and symbol estimates provided by the EM algorithm can be used to reconstruct then remove the ICI components from the received signal and the resulting samples can be re-injected once again as new inputs to the EM algorithm to enhance accuracy. In this way, the entire process can be repeated $\mathcal{I}_{\text{ICI}}$ iterations until no additional improvements is achieved. The ICI suppression requires decoding the data symbols to be able to reduce the ICI level. Instead of using the successive interference suppression (SIS) at the output of each antenna element, we make use of the symbols' posteriors, $P_{m,n,k}^{(\mathcal{I}_{\text{EM}})}$, already provided by the EM algorithm and decode the data symbols as follows:

$$\widehat{a}_k^{(s)}[n] = \underset{a_m\in\mathcal{C}^M}{\text{argmax}}\Big|a_m - \textstyle\sum_{m=1}^{M} P_{m,n,k}^{(\mathcal{I}_{\text{EM}})} a_m\Big|^2, \quad (28)$$

where $\widehat{a}_k^{(s)}[n]$ is the detected symbol corresponding to the n^{th} subcarrier of each k^{th} OFDM block after s ICI cancellation iterations. At each s^{th} ICI cancellation iteration, the detected symbols are used to remove the ICI component from the original received signal so as to provide the EM algorithm with less-ISI-corrupted observations.

IV. Initialization using the DA ML estimator

As mentioned earlier, the EM technique is iterative in nature and requires a good initial guess in order to return accurate estimates of the channel gains. An intuitive solution for obtaining those initial values is to use the pilot symbols injected at the subcarrier positions $\{p_1, p_2, \ldots, p_{N_p}\}$ within each OFDM block. In the SIMO system, the received N_p subcarriers at each OFDM block , $\mathbf{y}_{k,r}^{(p)} = [y_{k,r}(p_1), y_{k,r}(p_2), \ldots, y_{k,r}(p_{N_p})]^T$, corresponding to the pilot positions (by neglecting the ICI) are given by:

$$\tilde{\mathbf{y}}_{k,r}^{(p)} = \text{diag}\Big\{\mathbf{a}_k^{(p)}\Big\}\mathbf{h}_{k,r}^{(p)} + \mathbf{w}_{k,r}^{(p)}, \quad (29)$$

where $\mathbf{a}_k^{(p)} = [a_k^{(p)}(1), a_k^{(p)}(2), \ldots, a_k^{(p)}(N_p)]^T$ are the transmitted pilots within the k^{th} OFDM block. The channel frequency response and noise component corresponding to the pilot indices are given by $\mathbf{h}_{k,r}^{(p)} = \big[[\mathbf{H}_{k,r}]_{p_1,p_1}, [\mathbf{H}_{k,r}]_{p_2,p_2}, \ldots, [\mathbf{H}_{k,r}]_{p_{N_P},p_{N_p}}\big]^T$ and $\mathbf{w}_{k,r}^{(p)} = [w_{k,r}(p_1), w_{k,r}(p_2), \ldots, w_{k,r}(p_{N_p})]^T$, respectively. By stacking the received pilot samples at the output of the antenna elements into vectors, $\Big\{\mathbf{y}_k^{(p)}(p_n) = [y_{k,1}(p_n), y_{k,2}(p_n), \ldots, y_{k,N_r}(p_n)]^T\Big\}_{n=1}^{N_p}$, we rewrite (29) as follows:

$$\mathbf{y}_k^{(p)} = \mathbf{A}_k^{(p)}\mathbf{F}^{(p)}\bar{\varphi}_k + \mathbf{w}_k^{(p)}, \quad (30)$$

where $\mathbf{w}_k^{(p)} = \left[\mathbf{w}_{k,1}^{(p)\,T}, \mathbf{w}_{k,2}^{(p)\,T}, \ldots, \mathbf{w}_{k,N_r}^{(p)\,T}\right]^T$ and $\mathbf{A}_k^{(p)}$ is a diagonal matrix containing the pilot symbols $\mathbf{a}_k^{(p)}$. The matrix $\mathbf{F}^{(p)}$ is a $(N_r N_p \times L)$ block-diagonal matrix ($L = \sum_{r=1}^{N_r} L_r$) defined as follows:

$$\mathbf{F}^{(p)} \quad = \quad \text{blkdiag}\{\mathbf{F}_1^{(p)}, \mathbf{F}_2^{(p)}, \ldots, \mathbf{F}_{N_r}^{(p)}\}. \tag{31}$$

in which $\mathbf{F}_r^{(p)}$ contains the rows of the matrices $\mathbf{F}_r$ that corresponds to the pilot symbols' indices $\left(\text{i.e., } \left\{ \{[\mathbf{F}_r^{(p)}]_{m,l} = [\mathbf{F}_r]_{p_m,l}\}_{m=1}^{N_p} \right\}_1^{L_r}\right)$. The corresponding LLF is given by:

$$\mathcal{L}(\boldsymbol{\psi}_k) = -N_r N_p \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\left(\mathbf{y}_k - \mathbf{A}_k^{(p)}\mathbf{F}^{(p)}\bar{\boldsymbol{\varphi}}_k\right)^H \left(\mathbf{y}_k - \mathbf{A}_k^{(p)}\mathbf{F}^{(p)}\bar{\boldsymbol{\varphi}}_k\right). \tag{32}$$

By differentiating (32) with respect to $\bar{\boldsymbol{\varphi}}_k$, we obtain the following initial ML-based DA estimates:

$$\widehat{\bar{\boldsymbol{\varphi}}}_k^{(0)} = \left(\mathbf{F}^{(p)\,H}\mathbf{A}_k^{(p)\,H}\mathbf{A}_k^{(p)}\mathbf{F}^{(p)}\right)^{-1}\mathbf{F}^{(p)\,H}\mathbf{A}_k^{(p)\,H}\mathbf{y}_k. \tag{33}$$

The initial estimate of the noise variance is also obtained by differentiating (32) with respect to σ^2 which leads to the following result:

$$\widehat{\sigma^2}^{(0)} = \frac{1}{2N_p N_r}\left\|\mathbf{y}_k - \mathbf{A}_k^{(p)}\mathbf{F}^{(p)}\bar{\boldsymbol{\varphi}}_k^{(0)}\right\|^2. \tag{34}$$

Usually, the solution in (33) requires that $N_p \geq \max\{L_r\}_{r=1}^{N_r}$ otherwise the system of equations is underdetermined and the matrix $\mathbf{F}_r^{(p)\,H}\text{diag}\{\mathbf{a}_k^{(p)}\}\{\mathbf{a}_k^{(p)}\}^H\mathbf{F}_r^{(p)}$ is no longer invertible. In this case, the overall throughput will be strongly dependant on the number of paths $\max\{L_r\}_{r=1}^{N_r}$. Taking into account the fact that the traditional LS solution in (33) corresponds to an ill-posed problem, we opt for a regularization technique to solve this problem. One attracting solution is the Tikhonov regularization which allows us to obtain the initial estimates as follows:

$$\widehat{\bar{\boldsymbol{\varphi}}}_k^{(0)} = \left(\mathbf{F}^{(p)\,H}\mathbf{A}_k^{(p)\,H}\mathbf{A}_k^{(p)}\mathbf{F}^{(p)} + \lambda\mathbf{I}_L\right)^{-1}\mathbf{F}^{(p)\,H}\mathbf{A}_k^{(p)\,H}\mathbf{y}_k. \tag{35}$$

The factor λ is a regularization factor, when set to zero, the solution in (35) becomes equivalent to the one in (33).

V. Simulation Results

In this section, we assess the performance of the new EM-based ML time varying channel estimator $i)$ at the component level in terms of the mean square error (MSE) of the channel gains (averaged over all antennas), and $ii)$ in terms of link-level bit error rate (BER). In all simulations, we consider a SIMO OFDM RIT with $N = 128$ subcarriers, a central frequency, $f_c = 5$ GHz. The sampling period is $T_s = 0.5$ μs. The channel between the transmitter and each r^{th} antenna element is modeled by a multipath Rayleigh fading channel where the individual complex path gains, $\{\alpha_{l,r}(t)\}_{l=1}^{L_r}$, follow a uniform Jake's model. We assume, without loss of generality, that the links between the source and the N_r receiving antennas have the same channel parameters given in Table I.

TABLE I
Channel Parameters

Path number	1	2	3	4	5	6
Average Power [dB]	-7.219	-4.219	-6.219	-10.219	-12.219	-14.219
Normalized Delay	0	0.4	1	3.2	4.6	10

In Fig. 1, we investigate the influence of the number of receiving antenna elements on the estimation performance. We compare our proposed estimator (referred to hereafter as ML-EM) to both the LS technique and the MSE lower bound (LB) derived in [6]. We observe a clear advantage of ML-EM

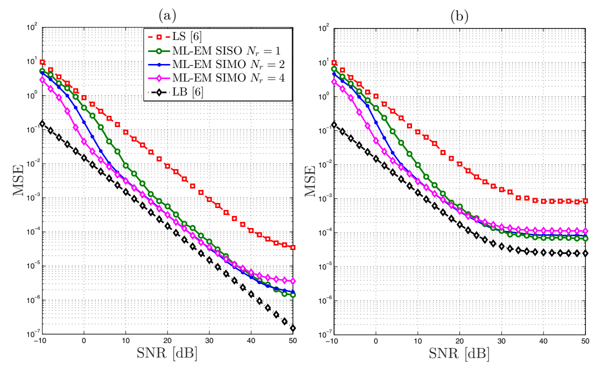

Fig. 1. MSE of ML-EM and LS vs. the SNR using different number of receiving antennas with $N_c = 3$ and $N_p = 8$ at: (a) $F_D T = 0.02$, and (b) $F_D T = 0.1$.

at both low and high Dopplers even in the SISO case. As the number of antenna elements increases, ML-EM exhibits a better estimation accuracy especially at low and medium SNR levels. Since ML-EM takes advantage of the diversity gain of multi-antenna systems, it is able to improve the channel estimates per-antenna. Moreover, the noise variance estimate in (26), provided by ML-EM is a more accurate as it is averaged over many antenna branches. At high SNR, however, we observe that increasing the number of antennas has almost no effect on the estimation accuracy performance. This is due to the noise level being lower than the ICI components. At such SNR levels, the channel estimation accuracy is dictated mainly by ICI cancellation capabilities of the proposed design.

In Fig. 2, we investigate the impact of the regularization factor λ on the performance of the proposed technique. We

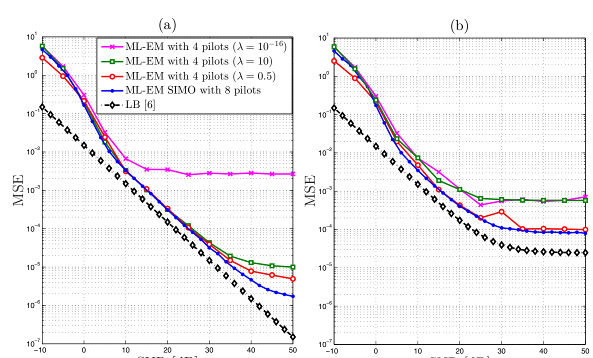

Fig. 2. MSE of ML-EM and LS vs. the SNR for different regularization factors with $N_c = 3$ and $N_p = 4$ at: (a) $F_D T = 0.02$, and (b) $F_D T = 0.1$.

observe that by using an arbitrarily small regularization factor (i.e., $\lambda = 10^{-16}$), the performance of EM-ML deteriorates since its initialization suffers from the same instability issues of the traditional LS technique. By increasing the value of λ, its performance improves and approaches the estimation accuracy achieved with $N_p = 8$ pilot tones. The latter corresponds to an overdetermined problem. However, for higher values of λ, the performance of ML-EM starts to deteriorate again since the regularized LS-based initialization procedure solution departs significantly from the original problem defined in (32) and becomes less sensitive to the received samples.

In Fig. 3, we assess the robustness of the proposed technique to the number of available pilot subcarriers. We see that the gap

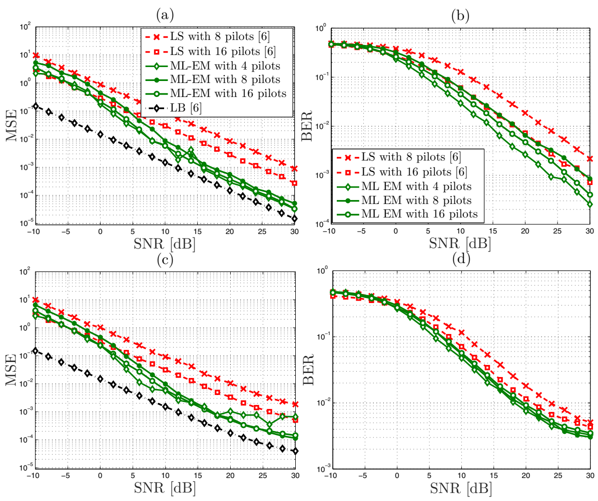

Fig. 3. MSE and BER of ML-EM and LS vs. the SNR with $N_c = 3$ in terms of: (a) MSE at $F_D T = 0.02$, (b) BER at $F_D T = 0.02$, (c) MSE at $F_D T = 0.1$, and (b) BER at $F_D T = 0.1$.

between the two techniques increases by reducing the number of pilots per OFDM block from $N_p = 16$ to $N_p = 8$, more so at high Dopplers. Indeed, the LS technique's performance deteriorates by reducing N_p while ML-EM exhibits exactly the same performance at medium-to-high SNR thresholds. Actually, ML-EM provides approximately the same BER as LS, yet with a lower number of pilots. Consequently, the new technique can achieve a higher throughput since the overhead is reduced by half. The number of pilots can be further reduced to $N_p = 4$ which is lower than the number of paths. At this configuration, the LS technique cannot provide reliable estimates. However, the proposed EM-based solution still works properly when initialized with the regularized LS instead of the traditional ML DA technique and as seen in Fig. 3 (a) and (c), the new technique provides approximately the same performance with slight deterioration at high SNRs. The latter does not affect , however, the BER performance. Indeed, as seen in Fig. 3 (d), ML-EM provides approximately the same BER with the different number of pilots considered there.

VI. Conclusion

In this paper, we addressed the problem of time-varying channel estimation for SIMO OFDM systems in multipath propagation environments. The proposed approach is based on a polynomial approximation of the complex path gains and takes advantage of all the observation (both at pilot and non-pilot positions) enhance the channel estimation capabilities. By using the regularized LS technique, we were able to reduce the number of pilots needed to properly initialize the proposed iterative EM-based solution thereby enhancing the effective throughput of the system. It is shown through exhaustive simulations that the proposed solution outperforms the existing DA-only LS solution. The latter translates into a net advantages in terms of the BER performance, especially at medium-to-high SNR values, more so at relatively higher Doppler frequencies.

References

[1] G.T.38.211, "NR; Physical channels and modulation," 2018.

[2] P. Banelli, S. Buzzi, G. Colavolpe, A. Modenini, F. Rusek, and A. Ugolin, "Modulation formats and waveforms for 5G networks: Who will be the heir of OFDM?: An overview of alternative modulation schemes for improved spectral efficiency," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 80-93, Nov. 2014.

[3] G.L.Z. Tang, R.C. Cannizzaro, and P. Banelli, "Pilot-assisted time-varying channel estimation for OFDM systems," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2226-2238, May 2007.

[4] H.Y.S. Tomasin, A. Gorokhov, and J.P. Linnartz, "Iterative interference cancellation and channel estimation for mobile OFDM," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 238-245, Jan. 2005.

[5] Y. Mostofi and D. Cox, "ICI mitigation for pilot-aided OFDM mobile systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 765-774, Mar. 2005.

[6] H. Hijazi and L. Ros, "Polynomial estimation of time-varying multipath gains with intercarrier interference mitigation in OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 140-151, Jan. 2009.

[7] H. Senol, E. Panayirci, and V. Poor, "Nondata-aided joint channel estimation and equalization for OFDM systems in very rapidly varying mobile channels," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4236-4253, Aug. 2012.

[8] A. Wiesel, J. Goldberg, and H. Messer-Yaron, "NR estimation in time-varying fading channels," *EEE Trans. Commun.*, vol. 54, no. 5, pp. 841-848, May. 2006.

[9] F. Bellili, M. Meftehi, and S. Affes, "Maximum likelihood SNR estimation of linearly-modulated signals over time-varying flat-fading SIMO channels," *IEEE Trans. Signal Process.*, vol. 63, no. 2, pp. 441-456, Jan. 2015.

[10] F. Bellili, A. StÃl'phenne, and S. Affes, "SNR estimation of QAM-modulated transmissions over time-varying SIMO channels," in *Proc. IEEE Int. Symp. Wireless Commun. Syst.*, Reykjavik, Iceland, 2008, pp. 199-203.

[11] H. Hijazi and L. Ros, "Joint data QR-detection and Kalman estimation for OFDM time-varying Rayleigh channel complex gains," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 170-178, Jan. 2010.

[12] H. Hijazi, P. E. Simon, M. Lienard, and L. Ros," Channel estimation for MIMO-OFDM systems in fast time-varying environments," in *Proc. Inte. Sympo. on commun., cont. and sig. proc. (ISCCSP)*, Limassol, Cyprus, 2010, pp. 1-6.

[13] R. Yang, Y. Siyang, S. Pengbo, T. Yinglei, and Z. Yanhua," Iterative channel estimation and detection for high-mobility MIMO-OFDM systems: mitigating error propagation by exploiting error information," *Wireless Pers. Commun.*, vol. 84, no. 3, pp. 1907-1931, May. 2015.