ASYMPTOTICALLY MMSE-OPTIMUM PILOT DESIGN FOR COMB-TYPE OFDM CHANNEL ESTIMATION IN HIGH-MOBILITY SCENARIOS

K.M. Zahidul Islam, Tareq Y. Al-Naffouri and Naofal Al-Dhahir*

Department of Electrical Engineering, The University of Texas at Dallas, Richardson, TX.
*Electrical Engineering Department, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia.

ABSTRACT

Under high mobility, the orthogonality between sub-carriers in an OFDM symbol is destroyed resulting in severe inter-carrier interference (ICI). We present a novel algorithm to estimate the channel and ICI coefficients by exploiting the channel's time and frequency correlations and the (approximately) banded structure of the frequency-domain channel matrix. In addition, we invoke the asymptotic equivalence of Toeplitz and circulant matrices to reduce the dimensionality of the channel estimation problem by retaining the dominant terms only in an offline eigen-decomposition. Furthermore, we show that the asymptotically MMSE-optimum pilot design consists of identical equally-spaced frequency-domain clusters whose size is determined by the channel Doppler spread. Comparisons of our proposed algorithm with a widely-cited recent algorithm demonstrate a significant performance advantage at a comparable real-time complexity.

Index Terms— Channel estimation, Doppler frequency, Model reduction, ICI, OFDM.

I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is widely used in high-rate broadband wireless applications thanks to its multipath resilience and use of low-complexity single-tap frequency-domain equalizers. In highly-mobile scenarios, the channel varies within each OFDM symbol causing the OFDM sub-carriers to lose their orthogonality resulting in inter-carrier interference (ICI) and making signal detection using single-tap equalizers highly sub-optimal. Furthermore, ICI complicates the channel estimation process since for each OFDM symbol, interference caused by the adjacent sub-carriers has to be estimated along with the channel frequency response at each sub-carrier.

A widely-cited hybrid frequency/time domain channel estimation algorithm was proposed by Mostofi and Cox (referred to as MC algorithm in this paper) in [1] based on a linear approximation of the time-variations of each channel impulse response (CIR) coefficient. However, this hybrid algorithm jointly processes 3 consecutive OFDM symbols which increases processing latency significantly. Furthermore, as we show in our simulations, its performance degrades at high Doppler spread.

In most of the existing literature on optimal pilot design for mobile OFDM, the channel is assumed fixed within each OFDM symbol while changing from one OFDM symbol to another. For doubly-selective channels, previously-proposed pilot designs use an impulsive frequency-domain pilot cluster structure made up of a single pilot subcarrier padded with zero subcarriers as guard band on both sides to eliminate ICI. This impulsive pilot design effectively treats the channel frequency response (CFR) matrix as a diagonal instead of a banded matrix even under high-mobility scenarios; thus ignoring useful signal energy dispersed into the adjacent subcarriers.

In this paper, we investigate the optimal pilot design for our pilot-aided OFDM channel estimation algorithm presented in [2]. The novelty of this paper lies in proving that the MMSE-optimum OFDM pilot over doubly-selective channels consists of identical equally-spaced non-impulsive frequency-domain pilot clusters for large FFT sizes. The optimal non-impulsive periodic pilot clusters exploit the banded CFR matrix structure to improve estimation accuracy.

This paper is organized as follows. Section II introduces the system model and assumptions. Section III briefly reviews the channel estimation algorithm in [2]. Optimal pilot design for the proposed algorithm are derived in Section IV. Section V presents the simulation results and performance comparisons. Finally, the paper is concluded in Section VI.

II. MODEL AND ASSUMPTIONS

Assuming perfect synchronization and a cyclic prefix (CP) length equal to or greater than CIR memory length, L, the FFT of the received vector after CP removal is given by

$$\boldsymbol{\mathcal{Y}} = \mathbf{Q}\mathbf{y} = \mathbf{Q}\mathbf{H}\mathbf{Q}^H\boldsymbol{\mathcal{X}} + \mathbf{Q}\mathbf{z} = \mathbf{G}\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{Z}} \tag{1}$$

where $\mathbf{Q}$ is the size-N DFT matrix, $\mathbf{H}$ is the time-domain channel matrix, $\mathbf{G} \triangleq \mathbf{Q}\mathbf{H}\mathbf{Q}^H$ is the CFR matrix and $\boldsymbol{\mathcal{Z}}$ is the frequency-domain noise vector. The vectors $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$ are

This work was supported by King Abdulaziz City for Science and Technology (KACST), Saudi Arabia, Project no. AR 27-98 and Research In Motion (RIM).

the comb-type pilot-data-multiplexed transmit and receive OFDM symbols, respectively, in the frequency domain. For a time-variant channel, $\mathbf{H}$ is not circulant, and $\mathbf{G}$ can no longer be assumed diagonal. Signal energy will be dispersed into the off-diagonal elements of $\mathbf{G}$. When the channel is fast-varying, ICI becomes significant and produces an irreducible error floor.

III. BAYESIAN LINEAR MMSE CHANNEL ESTIMATION

We can decompose $\mathbf{H}$ as the sum of L matrices, each of which corresponds to a CIR tap (and its time evolution), i.e. $\mathbf{H} = \sum_{l=0}^{L-1} \mathbf{A}_l$, where $\mathbf{A}_l$ is the matrix corresponding to the l-th CIR tap given by

$$\mathbf{A}_l = \text{Diag}([h_0(l), h_1(l), \ldots, h_{N-1}(l)])\mathbf{B}^l \quad (2)$$

where $h_n(l)$ is the complex zero-mean unit-variance CIR tap at lag l (for $0 \leq l \leq L-1$) and time instant n and $\mathbf{B}$ is a permutation matrix obtained by cyclic shifts of an identity matrix of size N to the left by one column. Considering Jakes's model for channel time variation, we can write $J(m-n) \triangleq E[h_m(l)h_n(l)^*] = J_0(2\pi f_d(m-n)T_s)$, where f_d and T_s are the Doppler frequency and sampling period, respectively, and $J_0(\cdot)$ is the zero-order Bessel function.

Vectorizing the CFR matrix $\mathbf{G}$ and using the Kronecker product property in Theorem T2.13 in [3] yields

$$\text{vec}(\mathbf{G}) = ((\mathbf{Q}^H)^T \otimes \mathbf{Q})\text{vec}(\mathbf{H}) = (\mathbf{Q}^* \otimes \mathbf{Q})\text{vec}(\mathbf{H}) \quad (3)$$

Now, we calculate the covariance matrix of vec($\mathbf{G}$) in terms of that of vec($\mathbf{H}$) as follows

$$\begin{aligned}\mathbf{R}_G &\triangleq E[\text{vec}(\mathbf{G})\text{vec}(\mathbf{G})^H]\\ &= (\mathbf{Q}^* \otimes \mathbf{Q})E[\text{vec}(\mathbf{H})\text{vec}(\mathbf{H})^H](\mathbf{Q}^* \otimes \mathbf{Q})^H \quad (4)\\ &\triangleq (\mathbf{Q}^* \otimes \mathbf{Q})\mathbf{R}_H(\mathbf{Q}^* \otimes \mathbf{Q})^H\end{aligned}$$

Let $\lambda_1, \ldots, \lambda_N$ be the eigenvalues of $\mathbf{J}$ and let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N$ be the corresponding eigenvectors. Define $\bar{\mathbf{v}}_n$ as an over-sampled version of eigenvector $\mathbf{v}_n$ given by

$$\bar{\mathbf{v}}_n = [v_n(0), \underbrace{0, \ldots, 0}_{N\text{zeros}}, v_n(1), \underbrace{0, \ldots, 0}_{N\text{zeros}}, v_n(2), \ldots, v_n(N-1)]^T$$

Following the eigen-analysis of $\mathbf{R}_H$ and henceforth that of $\mathbf{R}_G$, it has been shown in [2] that the covariance matrix $\mathbf{R}_G$ of the CFR matrix $\mathbf{G}$ has the NL nonzero eigenvalues $\underbrace{\lambda_1, \ldots, \lambda_1}_{L\text{ times}}, \ldots, \underbrace{\lambda_N, \ldots, \lambda_N}_{L\text{ times}}$ with corresponding eigenvectors $\boldsymbol{\mathcal{G}}_p$ ($1 \leq p \leq NL$) given by $(\mathbf{Q}^* \otimes \mathbf{Q})\bar{\mathbf{v}}_1, \ldots, (\mathbf{Q}^* \otimes \mathbf{Q})\mathbf{D}^{L-1}\bar{\mathbf{v}}_1, \ldots, (\mathbf{Q}^* \otimes \mathbf{Q})\bar{\mathbf{v}}_N, \ldots, (\mathbf{Q}^* \otimes \mathbf{Q})\mathbf{D}^{L-1}\bar{\mathbf{v}}_N$. We can write vec($\mathbf{G}$) as vec($\mathbf{G}$) $= \sum_{p=1}^{NL} \alpha_p \boldsymbol{\mathcal{G}}_p$ where the α_p's are independent random variables each with zero mean and variance equal to the eigenvalue λ_p. We reduce the parameter estimation space for $\mathbf{G}$ by retaining only those α_i's with large variance (compared to the noise variance) and considering the rest as modeling noise. Let N_d denote the number of dominant eigenvalues of $\mathbf{J}$. In other words, we can approximate vec($\mathbf{G}$) as follows

$$\text{vec}(\mathbf{G}) = \sum_{p=1}^{N_dL} \alpha_p \boldsymbol{\mathcal{G}}_p + \widetilde{\boldsymbol{\mathcal{Z}}} \approx \sum_{p=1}^{N_dL} \alpha_p \boldsymbol{\mathcal{G}}_p \quad (5)$$

where the term $\widetilde{\boldsymbol{\mathcal{Z}}} = \sum_{p=N_dL+1}^{NL} \alpha_p \boldsymbol{\mathcal{G}}_p$ is ignored. By un-vectorizing (5), we can re-write (1) as

$$\boldsymbol{\mathcal{Y}} = \mathbf{G}\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{Z}} \approx \sum_{p=1}^{N_dL} \alpha_p \underbrace{\mathbf{G}_p\boldsymbol{\mathcal{X}}}_{\boldsymbol{\mathcal{E}}_p} + \boldsymbol{\mathcal{Z}} = \sum_{p=1}^{N_dL} \alpha_p \boldsymbol{\mathcal{E}}_p + \boldsymbol{\mathcal{Z}} \quad (6)$$

Let $\{k_1, k_2, \ldots, k_T\}$ denote the set of indices for training sub-carriers. From (6), we can remove all the sub-carriers that do not belong to the training set, resulting in a linear system of T equations in N_dL unknowns ($T > N_dl$). In matrix form, we can write

$$\underline{\boldsymbol{\mathcal{Y}}} = \underline{\mathbf{E}}_p\boldsymbol{\alpha} + \underline{\boldsymbol{\mathcal{Z}}} \quad (7)$$

where, $\underline{\boldsymbol{\mathcal{Y}}} = [\mathcal{Y}(k_1)\mathcal{Y}(k_2)\cdots\mathcal{Y}(k_T)]^T$, $\underline{\boldsymbol{\mathcal{E}}}_p = [\mathcal{E}_p(k_1)\mathcal{E}_p(k_2)\cdots\mathcal{E}_p(k_T)]^T$, $\underline{\mathbf{E}}_p = [\underline{\boldsymbol{\mathcal{E}}}_1 \cdots \underline{\boldsymbol{\mathcal{E}}}_{N_dL}]$ and $\boldsymbol{\alpha} = [\alpha_1 \cdots \alpha_{N_dL}]^T$. We compute $\boldsymbol{\alpha}$ for a generalized noise covariance matrix $\mathbf{R}_z$ using the following linear minimum mean square error (LMMSE) estimator

$$\begin{aligned}\hat{\boldsymbol{\alpha}} &= \mathbf{R}_\alpha\underline{\mathbf{E}}_p^H\left[\underline{\mathbf{E}}_p\mathbf{R}_\alpha\underline{\mathbf{E}}_p^H + \mathbf{R}_z\right]^{-1}\underline{\boldsymbol{\mathcal{Y}}}\\ &= \underbrace{\left[\mathbf{R}_\alpha^{-1} + \underline{\mathbf{E}}_p^H\mathbf{R}_z^{-1}\underline{\mathbf{E}}_p\right]^{-1}\underline{\mathbf{E}}_p^H\mathbf{R}_z^{-1}}_{\triangleq\mathbf{W}}\underline{\boldsymbol{\mathcal{Y}}} = \mathbf{W}\underline{\boldsymbol{\mathcal{Y}}} \quad (8)\end{aligned}$$

where $\mathbf{R}_\alpha = \text{diag}(\alpha_1, \alpha_2, \cdots, \alpha_{N_dl})$ is a diagonal matrix. We can pre-compute $\mathbf{W}$ in (8) with the knowledge of N, f_d and SNR, hence, the real-time implementation complexity of the channel estimation algorithm can be significantly reduced. The estimation error vector $\boldsymbol{\epsilon} = \boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}$ is zero mean with the following covariance matrix $\mathbf{C}_\epsilon = \left[\mathbf{R}_\alpha^{-1} + \underline{\mathbf{E}}_p^H\mathbf{R}_z^{-1}\underline{\mathbf{E}}_p\right]^{-1}$. Hence, the MSE in estimating α_i is MSE($\hat{\alpha}_i$) $= \mathbf{C}_\epsilon(i,i)$.

IV. OPTIMAL PILOT DESIGN

Assuming that the $\mathcal{Z}(k)$'s are i.i.d samples with zeros mean and variance of σ_z^2, $\mathbf{C}_\epsilon$ can be written as

$$\mathbf{C}_\epsilon = [\mathbf{R}_\alpha^{-1} + \frac{1}{\sigma_z^2}\underbrace{\underline{\mathbf{E}}_p^H\underline{\mathbf{E}}_p}_{\triangleq\mathbf{R}_E}]^{-1} = [\mathbf{R}_\alpha^{-1} + \frac{1}{\sigma_z^2}\mathbf{R}_E]^{-1} \quad (9)$$

Our main objective is to design a frequency-domain pilot structure to minimize the trace of $\mathbf{C}_\epsilon$. The structure of $\mathbf{R}_E$ is given by

$$\begin{aligned}\mathbf{R}_E &= \underline{\mathbf{E}}_p^H\underline{\mathbf{E}}_p\\ &= \begin{bmatrix} \boldsymbol{\mathcal{X}}^H\mathbf{G}_1^H\mathbf{G}_1\boldsymbol{\mathcal{X}} & \boldsymbol{\mathcal{X}}^H\mathbf{G}_1^H\mathbf{G}_2\boldsymbol{\mathcal{X}} & \cdots & \boldsymbol{\mathcal{X}}^H\mathbf{G}_1^H\mathbf{G}_{N_dL}\boldsymbol{\mathcal{X}}\\ \boldsymbol{\mathcal{X}}^H\mathbf{G}_2^H\mathbf{G}_1\boldsymbol{\mathcal{X}} & \boldsymbol{\mathcal{X}}^H\mathbf{G}_2^H\mathbf{G}_2\boldsymbol{\mathcal{X}} & \cdots & \boldsymbol{\mathcal{X}}^H\mathbf{G}_2^H\mathbf{G}_{N_dL}\boldsymbol{\mathcal{X}}\\ \vdots & \vdots & \ddots & \vdots\\ \boldsymbol{\mathcal{X}}^H\mathbf{G}_{N_dL}^H\mathbf{G}_1\boldsymbol{\mathcal{X}} & \boldsymbol{\mathcal{X}}^H\mathbf{G}_{N_dL}^H\mathbf{G}_2\boldsymbol{\mathcal{X}} & \cdots & \boldsymbol{\mathcal{X}}^H\mathbf{G}_{N_dL}^H\mathbf{G}_{N_dL}\boldsymbol{\mathcal{X}} \end{bmatrix}\end{aligned} \quad (10)$$

IV-A. Problem Formulation

Now, $\mathbf{C}_\epsilon^{-1} \triangleq \mathbf{R}_\alpha^{-1} + \frac{1}{\sigma_z^2}\mathbf{R}_E$ is a positive-definite Hermitian matrix of size $N_dL \times N_dL$ for any $\boldsymbol{\mathcal{X}} \in \mathbb{C}^N$. Hence, Jensen's inequality yields following trace inequality

$$\frac{\sum_{i=1}^{N_dL} \lambda_{\mathbf{C}}(i)}{N_dL} \geq \frac{N_dL}{\sum_{i=1}^{N_dL} \frac{1}{\lambda_{\mathbf{C}}(i)}} \tag{11}$$

$$\Rightarrow \quad \text{Trace}(\mathbf{C}_\epsilon)\text{Trace}(\mathbf{C}_\epsilon^{-1}) \geq (N_dL)^2$$

$$\Rightarrow \quad \text{Trace}(\mathbf{C}_\epsilon) \geq \frac{(N_dL)^2}{\text{Trace}\left(\mathbf{R}_\alpha^{-1}\right) + \text{Trace}\left(\mathbf{R}_E\right)}$$

with equality if and only if $\mathbf{C}_\epsilon^{-1} = k\mathbf{I}_{N_dL}$ where k is a real positive constant. It follows from (9) that $\mathbf{R}_E = \sigma_z^2\left(\mathbf{C}_\epsilon^{-1} - \mathbf{R}_\alpha^{-1}\right)$. Since, $\mathbf{R}_\alpha$ is diagonal, to minimize MSE, $\mathbf{R}_E$ must be a diagonal matrix. Therefore, from (10) below, our objective is to determine $\boldsymbol{\mathcal{X}}$ such that

1) $\boldsymbol{\mathcal{X}}^H\mathbf{G}_i^H\mathbf{G}_j\boldsymbol{\mathcal{X}} = c, \quad i = j;\ i,j = 1,2,\cdots,N_dL$
2) $\boldsymbol{\mathcal{X}}^H\mathbf{G}_i^H\mathbf{G}_j\boldsymbol{\mathcal{X}} = 0, \quad i \neq j;\ i,j = 1,2,\cdots,N_dL$

where c is a constant representing pilot power constraint.

IV-B. Asymptotic Analysis

For simplicity, we assume that $\boldsymbol{\mathcal{X}}$ is a comb-type OFDM pilot symbol with data sub-carriers masked out by zeros. We can gain further insight into the pilot optimization problem by approximating the Toeplitz matrix $\mathbf{J}$ by a circulant matrix for large N using Szego's theorem. Hence, the eigenvectors and eigenvalues of $\mathbf{J}$ converge to the FFT columns and FFT transform of the first column of $\mathbf{J}$, respectively. Using the expression $\mathbf{G}_i = \mathbf{Q}\text{diag}(\mathbf{v}_n)\mathbf{B}^l\mathbf{Q}^H$ (see [2] for details) where $\mathbf{v}_n$ are the dominant eigenvectors of $\mathbf{J}$, the (m,n)-th element of $\mathbf{R}_E$ can be rewritten as follows

$$\begin{aligned} R_E(m,n) &= \boldsymbol{\mathcal{X}}^H\mathbf{Q}\mathbf{B}^{H(j1)}\underbrace{\text{diag}(\mathbf{v}_{i1})^H\text{diag}(\mathbf{v}_{i2})}_{\Lambda_{i_1i_2}}\mathbf{B}^{(j2)}\mathbf{Q}^H\boldsymbol{\mathcal{X}} \\ &= \mathbf{x}^H\underbrace{\mathbf{B}^{H(j1)}\Lambda_{i_1i_2}\mathbf{B}^{(j2)}}_{\mathbf{I}_c(i_1,i_2,j_1,j_2)}\bar{\mathbf{x}} = \mathbf{x}^H\mathbf{I}_c(i_1,i_2,j_1,j_2)\mathbf{x} \end{aligned} \tag{12}$$

where $m = (i_1-1)N_d + j_1$, $n = (i_2-1)N_d + j_2$ for $i_1, i_2 = 1,2,\cdots,N_d$ and $j_1, j_2 = 1,2,\cdots,L$. Based on this circulant approximation of $\mathbf{J}$, we have 4 possible values of $R_E(m,n)$ as given in Table I. Now, as long as $j_1 = j_2$, $\mathbf{I}_c(i_1,i_2,j_1,j_2)$ is always a diagonal matrix. Otherwise, $\mathbf{I}_c(i_1,i_2,j_1,j_2)$ will be an upper-shifted or lower-shifted diagonal matrix where the position of the super or sub diagonal depends on the difference between j_1 and j_2. Note that $(L-1)$ is the maximum index of the super or sub diagonal of $\mathbf{I}_c(.)$ that can be non-zero. Hence, by designing $\mathbf{x}$ as a sparse vector with non-zero time-domain samples separated by $N_c - 1$ zeros where $(N_c-1) > (L-1)$, it is possible to zero out all case 2 and 4 non-zero elements in Table I. It is also possible to achieve a diagonal $\mathbf{R}_E$ for large N using such a sparse $\mathbf{x}$. Next, we prove a key result.

Table I: Elements of $\mathbf{R}_E$

Case	$i_1 = i_2$	$j_i = j_2$	$R_E(m,n)$	Comments
1	Yes	Yes	Real, $\Vert\boldsymbol{\mathcal{X}}\Vert^2$	Diagonal elements of $\mathbf{R}_E$
2	Yes	No	0	$\mathbf{I}_c(\cdot)$ is upper/lower shifted diagonal
3	No	Yes	Complex	$\mathbf{I}_c(\cdot)$ is diagonal
4	No	No	0	$\mathbf{I}_c(\cdot)$ is upper/lower shifted diagonal

Proposition: If the frequency-domain pilot vector $\boldsymbol{\mathcal{X}}$ has the periodic clustered structure shown at the top of Fig. 1 with N_p adjacent subcarriers in each pilot cluster and the number N_c and period L_c of the pilot clusters satisfy the relation $N = N_cL_c$, then the time-domain pilot vector $\mathbf{x}$ will be sparse as shown at the bottom of Fig. 1.

Proof: If we use a periodic clustered $\boldsymbol{\mathcal{X}}$ with N_p adjacent sub-carriers in each cluster, the FFT operation will also ensure periodicity in the time-domain signal $\mathbf{x}$. Let L_c and N_c denote the period of the pilot clusters and the total number of pilot clusters in $\boldsymbol{\mathcal{X}}$, respectively. Using the FFT relationship, the m-th element of $\mathbf{x}$ is given by

$$x_m = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\mathcal{X}_n e^{j\frac{2\pi mn}{N}} \tag{13}$$

Using the periodic pilot cluster structure, (13) becomes

$$\begin{aligned} x_m =& \frac{1}{\sqrt{N}}\Big[\mathcal{P}_0 e^{j\frac{2\pi dm}{N}}\left(1 + e^{j\frac{2\pi L_c m}{N}} + \cdots + e^{j\frac{2\pi L_c(N_c-1)m}{N}}\right) \\ &+\mathcal{P}_1 e^{j\frac{2\pi(d+1)m}{N}}\left(1 + e^{j\frac{2\pi L_c m}{N}} + \cdots + e^{j\frac{2\pi L_c(N_c-1)m}{N}}\right) + \cdots \\ &+\mathcal{P}_{N_p-1}\, e^{j\frac{2\pi(d+N_p-1)m}{N}}\left(1 + e^{j\frac{2\pi L_c m}{N}} + \cdots + e^{j\frac{2\pi L_c(N_c-1)m}{N}}\right)\Big] \end{aligned} \tag{14}$$

Using $N = N_cL_c$, (14) can be written as follows

$$x_m = \frac{1}{\sqrt{N}}\sum_{i=0}^{N_p-1}\mathcal{P}_i e^{j\frac{2\pi(d+i)}{N}}\left(\sum_{k=0}^{N_c-1} e^{j\frac{2\pi mk}{N_c}}\right) \tag{15}$$

If m is not an integer multiple of N_c, (15) becomes

$$x_m = \frac{1}{\sqrt{N}}\sum_{i=0}^{N_p-1}\mathcal{P}_i e^{j\frac{2\pi(d+i)}{N}}\left(\sum_{k=0}^{N_c-1} e^{j\frac{2\pi k}{N_c}}\right)^m \tag{16}$$

Now, $\left(\sum_{k=0}^{N_c-1} e^{j\frac{2\pi k}{N_c}}\right)$ is the geometric series sum of all of the N_c-th roots of unity which equals zero. Thus, $\mathbf{x}$ will have only L_c non-zero elements separated by $N_c - 1$ zeros. All we are left with now is the 3rd case in Table I ; i.e. we have to make $\mathbf{x}^H\mathbf{I}_c(i_1,i_2,j_1,j_2)\mathbf{x} = 0$ when $i_1 \neq i_2$ and $j_1 = j_2$. Note that due to the periodic structure of $\boldsymbol{\mathcal{X}}$, all pilot clusters are identical. Hence, we only need to optimize one pilot cluster. Towards this objective, we propose to choose the pilot cluster symbols from standard constellations and then to determine the optimum cluster size. Since the $\mathbf{G}_p$'s are (approximately) banded matrices with M main diagonals, $\boldsymbol{\mathcal{X}}$ should have at least M adjacent non-zero tones in each cluster setting the lower bound of N_p. In addition, $\mathbf{G}_i^H\mathbf{G}_j$'s will also be banded but with $(2M-1)$ diagonals. An upper

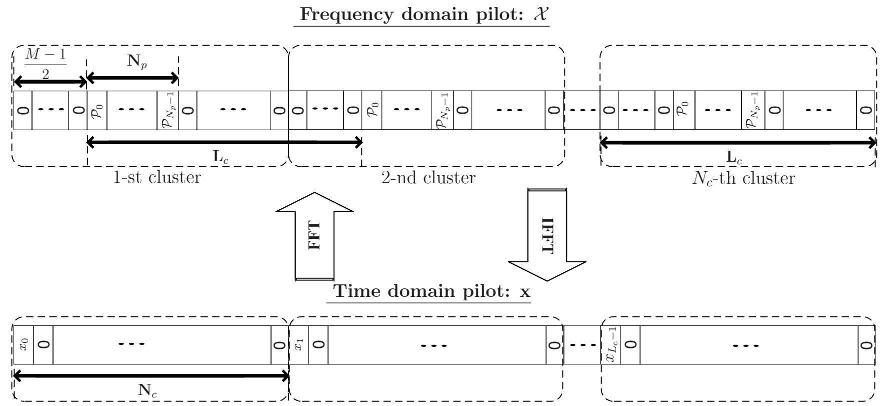

Fig. 1: Optimized pilot structure for our channel estimation algorithm in [2].

bound on N_p is derived by observing that if we choose N_p greater than $(2M-1)$ sub-carriers, we will be only adding zeros to $R_E(m,n)$ since the $\mathbf{G}_i^H\mathbf{G}_j$'s are banded with $(2M-1)$ diagonals. Increasing N_p beyond $(2M-1)$ also decreases the number of pilot clusters given the pilot overhead ratio. Hence, we must have $M \leq N_p \leq 2M-1$; where $M = 3, 5, \cdots$ and N_p odd. In [2] we have shown that for equidistant (but neither identical nor periodic) pilot clusters with random BPSK pilots, cluster size of $(2M-1)$ achieves best BER performance. Hence in the simulations presented in the following section we set the cluster size to $(2M-1)$ pilot sub-carriers. Finally, we propose to use identical periodic clusters with $(2M-1)$ pilot sub-carriers in each cluster modulated with random BPSK signals as pilots up to 20% normalized Doppler spread.

V. SIMULATION RESULTS

In this section, we compare the performance of our proposed algorithm with the MC algorithm [1] and impulsive pilot designs for a mobile coded OFDM system. We assumed the 3-tap SUI-3 channel perturbed by additive white Gaussian noise (AWGN). We found that for a normalized Doppler of 10%, $\mathbf{G}$ can be well approximated by a banded matrix with $M = 3$ diagonals and the first 3 eigenvalues of $\mathbf{J}$ are dominant, i.e. $N_d = 3$. For data detection, we implemented the 3-tap MMSE FEQ designed in [4]. We assumed a 15% pilot training overhead with size-5 periodic pilot clusters.

Figure 2 demonstrates the significant performance gains of our proposed channel estimation algorithm with optimized pilot design over the MC algorithm [1] and over impulsive pilot designs. It can also be seen from the figure that the performance loss due to the complexity-reducing circulant approximation is less than 2 dB for the entire SNR range under consideration.

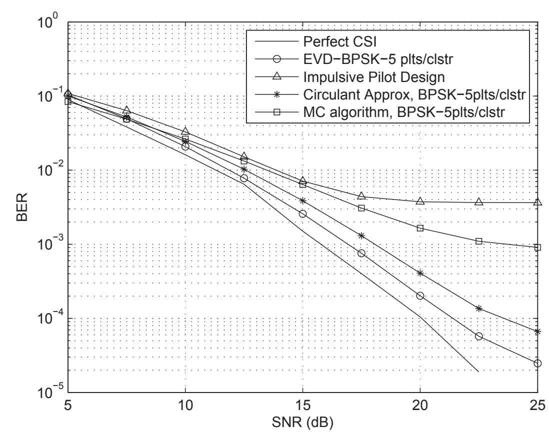

Fig. 2: BER comparison between our proposed algorithm (both exact and approximate versions) and MC algorithm for N=1024, f_d=10% and $M = 3$.

VI. CONCLUSIONS

We proposed a new pilot-aided algorithm for the estimation of fast time-varying channels in OFDM transmission. Unlike many existing OFDM channel estimation algorithms in the literature, we propose to perform channel estimation in the frequency domain, to exploit the structure of the channel response (such as frequency and time correlations and bandedness), optimize the pilot signal structure, and perform most of the computations offline resulting in high performance at substantial complexity reductions.

VII. REFERENCES

[1] Y. Mostofi and D. Cox, "ICI mitigation for pilot-aided OFDM mobile systems," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 765–774, Mar. 2005.

[2] T. Y. Al-Naffouri, K. M. Z. Islam, N. Al-Dhahir, and S. Lu, "A model reduction approach for OFDM channel estimation under high mobility conditions," *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 2181 –2193, April 2010.

[3] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Transactions on Circuits and Systems*, vol. 25, pp. 772–781, Sep 1978.

[4] S. Lu and N. Al-Dhahir, "Coherent and differential ICI cancellation for mobile OFDM with application to DVB-H," *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, pp. 4110–4116, November 2008.