Least Squares End Performance Experiment Design in Multicarrier Systems: The Sparse Preamble Case

Dimitrios Katselis, Cristian R. Rojas and Håkan Hjalmarsson

Abstract— **In this paper, the problem of experiment design for the task of channel identification in cyclic prefixed orthogonal frequency division multiplexing (CP-OFDM) systems is revisited. So far, the optimal input sequences for least squares (LS) channel identification with respect to minimizing the channel mean square error (MSE) under an input energy constraint have been derived. Here, we investigate the same problem for the LS channel estimator, but when the design takes into account an end performance metric of interest, namely, the symbol estimate MSE. Based on some convex approximations, we verify that optimal sparse preambles, i.e, input vectors employing as many pilots as the channel length, for LS channel estimation in its classical context are near optimal in the aforementioned application-oriented context for the symbol estimate MSE.**

I. INTRODUCTION

Cyclic prefixed orthogonal frequency division multiplexing (CP-OFDM) is currently enjoying popularity in both wired and wireless communication systems [1], mainly because of its immunity to multipath fading, which allows for a significant increase in the transmission rate [16]. Using the cyclic prefix (CP) as a guard interval, CP-OFDM can transform a frequency selective channel into a set of parallel flat channels with independent noise disturbances. This greatly simplifies both the estimation of the channel and the recovery of the transmitted data at the receiver. However, these advantages come at the cost of an increased sensitivity to frequency offset and Doppler spread. This is due to the fact that, although the subcarrier functions are perfectly localized in time, they suffer from spectral leakage in the frequency domain. Moreover, the inclusion of the CP entails a loss in spectral efficiency, which in practical systems can become as high as 25% [1]. Nevertheless, despite its aforementioned weaknesses, CP-OFDM is a mainstream system nowadays.

Recently, the application-oriented framework for pilot design in communication systems has been introduced [10], [11]. The origins of this design can be found in pre-existing work in the system identification literature, e.g., in [5] and references therein. This framework is mostly appropriate for such a design since the training sequences are selected to optimize a final performance metric of interest and not some of the classical metrics quantifying the distance between the estimated model and the true one, e.g., the mean square error (MSE). To this end, it is imperative to reexamine all known pilot sequences which are optimal for any communication system and for any estimation task, since it has been already shown that this optimality does *not* carry over to the new framework [10], [11].

The focus of this paper is on revisiting the preamble-based channel estimation task in CP-OFDM systems, when the pilot design is performed with respect to a specific end performance metric of interest. The question of selecting the pilot tones when using a least squares (LS) channel estimator to minimize the symbol estimate MSE subject to a training energy constraint is addressed. The case of a *sparse* preamble, where as many subcarriers as the channel length carry pilots, is examined when the signal-to-noise (SNR) is in usual operating regimes. By approximating the training design problem via a convex formulation, we *verify* that optimal sparse preambles in the classical estimation setup containing equipowered or equal pilots are near optimal in the application-oriented setup.

As far as the LS channel estimation for CP-OFDM in the classical setup is concerned, several results have previously been derived for the case of the channel estimation MSE metric, when the training design goal is to minimize the latter metric subject to a training energy constraint. In [13], it is shown that uniform spacing of L_h pilot tones (L_h being the channel impulse response (CIR) length) is the best choice given that the pilot tones are equipowered[1]. Equispaced and equipowered pilot tones were shown in [2] to be the optimal CP-OFDM preamble for a given training energy that accounts only for the useful signal, excluding the CP, while, including the CP, the pilot tones should be equispaced and equal [7], [8]. Optimal full preambles (all tones carry pilot symbols) with respect to the channel estimation MSE and when the training energy accounts for the CP can contain simply equipowered (not necessarily equal) symbols. A method for constructing such vectors is developed in [8].

Notation. Vectors and matrices are denoted by bold lowercase and uppercase letters, respectively. Superscripts ^{T} and ^{H} stand for transposition and conjugate transposition, respectively. Also, $\jmath = \sqrt{-1}$. $|\cdot|$ is the complex modulus or the absolute value. For a matrix $\boldsymbol{A}$, $(\boldsymbol{A})_{i,j}$ denotes its (i, j)th entry and for a vector $\boldsymbol{a}$, $\boldsymbol{a}_m$ or $[\boldsymbol{a}]_m$ denotes its mth entry. The expectation operator is denoted by $E(\cdot)$. $\boldsymbol{I}_m$ denotes the mth order identity matrix, while $\mathbf{0}_{m\times n}$ and $\mathbf{1}_{m\times n}$ are the all zeros and all ones $m \times n$ matrices, respectively. $C(m, n)$ is the number of n-combinations over a set of cardinality

The authors are with Access Linnaeus Center, School of Electrical Engineering, KTH Royal Institute of Technology, SE 100 44 Stockholm, Sweden dimitrik@kth.se, cristian.rojas@ee.kth.se, hjalmars@kth.se

[1]This is no longer valid if there are suppressed (*virtual*) subcarriers. In such a case, the optimal placement is non-uniform [12].

m. Moreover, $\succeq$ stands for the semidefinite cone partial ordering, while $\geq, \leq$ used with vectors denote elementwise orderings. For a vector $\boldsymbol{a}$, $|\boldsymbol{a}|$ is the vector of moduli or absolute values of its entries. Finally, $\boldsymbol{\mathcal{D}}_{\boldsymbol{a}} = \mathrm{diag}(\boldsymbol{a})$, i.e., $\boldsymbol{\mathcal{D}}_{\boldsymbol{a}}$ is a diagonal matrix having as main diagonal entries the elements of the vector $\boldsymbol{a}$.

II. System Model

Given M subcarriers, the result of the orthogonal frequency division multiplexing (OFDM) modulation of a (complex) $M \times 1$ vector $\boldsymbol{x}$ is

$$\boldsymbol{s} = \frac{1}{\sqrt{M}} \boldsymbol{\mathcal{F}}^H \boldsymbol{x},$$

where $\boldsymbol{\mathcal{F}}$ is the $M \times M$ DFT matrix, with entries $(\boldsymbol{\mathcal{F}})_{i,j} = e^{-\jmath \frac{2\pi}{M} ij}$, $i, j = 0, 1, \ldots, M-1$. Prior to transmission, a CP of length ν is prepended to the previous vector, to yield:

$$\boldsymbol{s}' = \begin{bmatrix} \boldsymbol{0}_{\nu \times (M-\nu)} \quad \boldsymbol{I}_\nu \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \boldsymbol{I}_M \end{bmatrix} \boldsymbol{s}. \tag{1}$$

Assume that the CP length is chosen to be the smallest possible one, namely equal to the channel order: $\nu = L_h - 1$ [13]. Moreover, perfect timing and frequency synchronization are assumed. The channel impulse response (CIR), $\boldsymbol{h} = \begin{bmatrix} \boldsymbol{h}_0 & \boldsymbol{h}_1 & \cdots & \boldsymbol{h}_{L_h-1} \end{bmatrix}^T$, is assumed to be constant over the duration of an OFDM symbol. The input to the OFDM demodulator, after the CP removal, can then be expressed as

$$\boldsymbol{r} = \boldsymbol{\mathcal{H}} \boldsymbol{s} + \boldsymbol{w},$$

where $\boldsymbol{\mathcal{H}}$ is the Toeplitz circulant matrix with its first row given by $\begin{bmatrix} \boldsymbol{h}_0 & \boldsymbol{0}_{1\times(M-L_h)} & \boldsymbol{h}_{L_h-1} & \cdots & \boldsymbol{h}_2 & \boldsymbol{h}_1 \end{bmatrix}$ and $\boldsymbol{w}$ is the noise at the receiver front end, assumed to be white Gaussian with zero mean and variance σ^2. The action of the DFT then results in

$$\boldsymbol{y} = \frac{1}{\sqrt{M}} \boldsymbol{\mathcal{F}} \boldsymbol{r} = \boldsymbol{\mathcal{D}}_{\boldsymbol{H}} \boldsymbol{x} + \boldsymbol{\eta}, \tag{2}$$

where $\boldsymbol{H}_m = \sum_{l=0}^{L_h-1} \boldsymbol{h}_l e^{-\jmath \frac{2\pi}{M} ml}$, $m = 0, 1, \ldots, M-1$ is the M-point channel frequency response (CFR) and $\boldsymbol{\eta} = \frac{1}{\sqrt{M}} \boldsymbol{\mathcal{F}} \boldsymbol{w}$ is the frequency domain noise, with the same statistics as $\boldsymbol{w}$.

In the sparse preamble case, let $\mathcal{I}_{L_h} = \{i_0, i_1, \ldots, i_{L_h-1}\}$ denote the set of subcarrier indices corresponding to nonzero pilots. The LS CIR estimate is

$$\bar{\boldsymbol{h}} = \boldsymbol{F}_{L_h \times L_h}^{-1} \check{\boldsymbol{H}}_{\mathcal{I}_{L_h}}. \tag{3}$$

Denoting as $\boldsymbol{F}_{M \times L_h}$ the submatrix of $\boldsymbol{\mathcal{F}}$ containing its first L_h columns, $\boldsymbol{F}_{L_h \times L_h}$ is the submatrix of $\boldsymbol{F}_{M \times L_h}$ with row indices in $\mathcal{I}_{L_h}$ and $\check{\boldsymbol{H}}_{\mathcal{I}_{L_h}}$ is the subvector of $\check{\boldsymbol{H}}$ with indices in $\mathcal{I}_{L_h}$. Here, $\check{\boldsymbol{H}}$ is the corresponding initially estimated CFR vector computed as

$$\check{\boldsymbol{H}}_m = \boldsymbol{y}_m^{tr} / \boldsymbol{p}_m = \boldsymbol{H}_m + \boldsymbol{\eta}_m^{tr} / \boldsymbol{p}_m,$$

where divisions by the corresponding pilots are performed only on the subcarriers with indices in $\mathcal{I}_{L_h}$. This implies that only noise lies on the rest of the subcarriers. The final LS CFR estimates are given in this case by

$$\check{\boldsymbol{H}} = \boldsymbol{F}_{M \times L_h} \bar{\boldsymbol{h}}. \tag{4}$$

III. Optimizing the Training with Respect to the Symbol Estimate MSE

A desired end performance metric of interest in this paper is the symbol estimate MSE. We assume the use of per subcarrier zero forcing (ZF) symbol estimators. This corresponds to symbol estimates $\hat{\boldsymbol{x}}_m$ equal to $\boldsymbol{y}_m / \check{\boldsymbol{H}}_m$ for all m. Moreover, we assume that the per subcarrier transmitted symbols are independent and identically distributed with zero mean and variance σ_x^2 and uncorrelated with the additive white Gaussian noise (AWGN) at the front end of the receiver. Then, the total symbol estimate MSE is given as follows:

$$\begin{aligned} \mathrm{MSE(ZF)} &= \sum_{m=0}^{M-1} E\left[\left| \frac{\boldsymbol{y}_m}{\check{\boldsymbol{H}}_m} - \boldsymbol{x}_m \right|^2 \right] \\ &= \sum_{m=0}^{M-1} \sigma_x^2 E\left[\left| \frac{\check{\boldsymbol{H}}_m - \boldsymbol{H}_m}{\check{\boldsymbol{H}}_m} \right|^2 \right] + \sigma^2 E\left[\left| \frac{1}{\check{\boldsymbol{H}}_m} \right|^2 \right] \end{aligned} \tag{5}$$

Depending on the probability distribution of $|\check{\boldsymbol{H}}_m|$'s, (5) may fail to exist. The $\mathrm{MSE(ZF)}$ will be finite if and only if the probability distribution function (pdf) of $|\check{\boldsymbol{H}}_m|$ is of order $O(|\check{\boldsymbol{H}}_m|^2)$ for all m as $\check{\boldsymbol{H}}_m \to 0$. Under the Gaussian assumption on $\boldsymbol{\eta}$, (5) is actually infinite, so the LS estimator gives rise to an ill-conditioned problem. In order to obtain well-behaved channel estimators that will be used in conjunction with the actual performance metrics, some sort of regularization is needed. Some ideas for appropriate regularization techniques to use may be obtained by modifying robust estimators (against heavy-tailed distributions), e.g., by trimming a standard estimator, if it gives a value very close to zero. An example of such a per subcarrier trimmed estimator is given as follows:

$$\hat{\boldsymbol{H}}_m = \begin{cases} \check{\boldsymbol{H}}_m, & \text{if } |\check{\boldsymbol{H}}_m| > \chi_m \\ \chi_m \check{\boldsymbol{H}}_m / |\check{\boldsymbol{H}}_m|, & \text{otherwise} \end{cases}, \tag{6}$$

where χ_m is a regularization parameter to be tuned via cross-validation or any other technique. The analysis of such an estimator is beyond the scope of this paper.

Remark: The reader may observe that the definition of the $\hat{\boldsymbol{H}}_m$ preserves the continuity at $|\check{\boldsymbol{H}}_m| = \chi_m$. Additionally, the event $\{\check{\boldsymbol{H}}_m = 0\}$ has zero probability since the distribution of $\check{\boldsymbol{H}}_m$ is continuous. Therefore, $\hat{\boldsymbol{H}}_m$ can be arbitrarily defined when $\check{\boldsymbol{H}}_m = 0$, e.g., $\hat{\boldsymbol{H}}_m = \chi_m$.

Assume a sufficiently small χ_m and a high SNR during training. Then, it can be shown that the following performance metric is a good variation/substitute of (5) in the course of designing the optimal training:

$$[\mathrm{MSE(ZF)}]_0 = \sum_{m=0}^{M-1} \left\{ \sigma_x^2 \frac{E\left[\left| \check{\boldsymbol{H}}_m - \boldsymbol{H}_m \right|^2 \right]}{E\left[\left| \check{\boldsymbol{H}}_m \right|^2 \right]} + \sigma^2 \frac{1}{E\left[\left| \check{\boldsymbol{H}}_m \right|^2 \right]} \right\} \tag{7}$$

We call this performance metric the *zeroth order* symbol estimate MSE. Our analysis in this paper will be based on this performance metric to facilitate the analytical treatment.

Remark: The obtained optimal preambles, as well as other known preambles in the literature will be then numerically compared against the *exact symbol estimate MSE*.

IV. The Sparse Preamble Case

Consider $\check{\boldsymbol{H}}$. This can be expressed as follows:

$$\begin{aligned}\check{\boldsymbol{H}} =& \underbrace{\boldsymbol{F}_{M\times L_h}\boldsymbol{F}^{-1}_{L_h\times L_h}\boldsymbol{H}_{\mathcal{I}_{L_h}}}_{\boldsymbol{H}} + \underbrace{\boldsymbol{F}_{M\times L_h}\boldsymbol{F}^{-1}_{L_h\times L_h}\boldsymbol{\mathcal{D}}^{-1}_{\boldsymbol{p}_s}\boldsymbol{\eta}^{tr}_{\mathcal{I}_{L_h}}}_{\boldsymbol{z}} \\ =& \boldsymbol{H}+\boldsymbol{z}, \end{aligned} \tag{8}$$

where $\boldsymbol{p}_s$ is the $L_h \times 1$ vector containing the nonzero pilot tones and $\boldsymbol{\eta}^{tr}_{\mathcal{I}_{L_h}}$ is the $L_h \times 1$ vector containing the noise components on the subcarriers with nonzero pilots. Using our previous assumptions, we can see that

$$E\left[\left|\check{\boldsymbol{H}}_m\right|^2\right] = \left|\boldsymbol{H}_m\right|^2 + E\left[\left|\boldsymbol{z}_m\right|^2\right].$$

Note that by considering LS channel estimation we implicitly assume that the prior distribution of $\boldsymbol{H}$ is unknown or, equivalently, that $\boldsymbol{H}$ is a deterministic but otherwise unknown quantity. Using this result, we obtain:

$$[\mathrm{MSE(ZF)}]_0 = \sum_{m=0}^{M-1} \frac{\sigma_x^2 E\left[\left|\boldsymbol{z}_m\right|^2\right] + \sigma^2}{\left|\boldsymbol{H}_m\right|^2 + E\left[\left|\boldsymbol{z}_m\right|^2\right]}. \tag{9}$$

Setting $\boldsymbol{\lambda}_m = E\left[\left|\boldsymbol{z}_m\right|^2\right]$ and $\boldsymbol{c}_m = \left|\boldsymbol{H}_m\right|^2$ for all m, we may differentiate the last expression with respect to $\boldsymbol{\lambda}_m$ to obtain:

$$\frac{\partial\,[\mathrm{MSE(ZF)}]_0}{\partial \boldsymbol{\lambda}_m} = \frac{\sigma_x^2\boldsymbol{c}_m - \sigma^2}{(\boldsymbol{c}_m+\boldsymbol{\lambda}_m)^2}, \quad m = 0,1,\ldots,M-1.$$

These partial derivatives will be positive if $\sigma_x^2\boldsymbol{c}_m - \sigma^2 > 0$ for every m. Clearly, this can be guaranteed if we set the SNR during the symbol estimation stage to a sufficiently high value by appropriately selecting σ_x^2. They will be negative in the case of low SNR during data transmission.

We first focus in the case of low SNR defined by the following inequality:

$$\sigma_x^2 < \underline{\sigma_x^2} = \frac{\sigma^2}{\max\{\boldsymbol{c}_0, \boldsymbol{c}_1, \ldots, \boldsymbol{c}_{M-1}\}}. \tag{10}$$

In this case,

$$\frac{\partial^2\,[\mathrm{MSE(ZF)}]_0}{\partial\boldsymbol{\lambda}_m \partial\boldsymbol{\lambda}_j} = 0, \quad m \neq j$$

and

$$\frac{\partial^2\,[\mathrm{MSE(ZF)}]_0}{\partial\boldsymbol{\lambda}_m^2} = -2\frac{\sigma_x^2\boldsymbol{c}_m - \sigma^2}{(\boldsymbol{c}_m+\boldsymbol{\lambda}_m)^3} > 0, \quad \forall m$$

i.e., the Hessian of $[\mathrm{MSE(ZF)}]_0$ is positive definite with respect to the $\boldsymbol{\lambda}_m$'s. Our optimization problem can be formulated as follows:

$$\begin{aligned} \min_{\{|[\boldsymbol{p}_s]_k|\}_{k=0}^{L_h-1}} \quad & [\mathrm{MSE(ZF)}]_0 \\ \text{s.t.} \quad & \sum_{k=0}^{L_h-1} |[\boldsymbol{p}_s]_k|^2 \leq \mathcal{E}, \end{aligned} \tag{11}$$

where in the consideration of the optimization variables we have used the fact that both the objective and the constraints are blind to the pilot phases. Moreover, we can write:

$$\begin{aligned} \boldsymbol{\lambda}_m &= \sigma^2 \sum_{i,j=0}^{L_h-1} e^{\jmath\frac{2\pi(j-i)m}{M}} \sum_{k=0}^{L_h-1} \frac{\left[\boldsymbol{F}^{-1}_{L_h\times L_h}\right]_{i,k}\left[\boldsymbol{F}^{-H}_{L_h\times L_h}\right]_{k,j}}{|[\boldsymbol{p}_s]_k|^2} \\ &= \sum_{k=0}^{L_h-1} \frac{\sigma^2}{|[\boldsymbol{p}_s]_k|^2}\left|\sum_{i=0}^{L_h-1}\left[\boldsymbol{F}^{-1}_{L_h\times L_h}\right]_{i,k} e^{-\jmath\frac{2\pi}{M}im}\right|^2. \end{aligned} \tag{12}$$

Collecting all $\boldsymbol{\lambda}_m$'s in one vector, we obtain the overdetermined system:

$$\boldsymbol{\lambda} = \boldsymbol{A}\boldsymbol{\rho}. \tag{13}$$

Here, $\boldsymbol{\rho}_m = 1/|[\boldsymbol{p}_s]_m|^2, m = 0,1,\ldots,L_h-1$ and $(\boldsymbol{A})_{m,k} = \sigma^2\left|\sum_{i=0}^{L_h-1}\left[\boldsymbol{F}^{-1}_{L_h\times L_h}\right]_{i,k} e^{-\jmath\frac{2\pi}{M}im}\right|^2$ for $m = 0,1,\ldots,M-1$ and $k = 0,1,\ldots,L_h-1$.

We now define the vector $\boldsymbol{\gamma}$ with entries

$$\boldsymbol{\gamma}_m = |[\boldsymbol{p}_s]_m|^2, m = 0,1,\ldots,L_h-1.$$

In order to solve (11) we can pose the following optimization problem:

$$\begin{aligned} \min_{\boldsymbol{b},\boldsymbol{\gamma},\boldsymbol{\lambda},\boldsymbol{\rho}} \quad & \boldsymbol{1}^T\boldsymbol{b} \\ \text{s.t.} \quad & \boldsymbol{1}^T\boldsymbol{\gamma} \leq \mathcal{E}, \\ & \boldsymbol{\gamma} \geq \boldsymbol{0}, \\ & \boldsymbol{\lambda} = \boldsymbol{A}\boldsymbol{\rho}, \\ & \boldsymbol{\rho}_m\boldsymbol{\gamma}_m \geq 1, \quad m \in \mathcal{L}_h \\ & \frac{\sigma_x^2\boldsymbol{\lambda}_m + \sigma^2}{\boldsymbol{c}_m + \boldsymbol{\lambda}_m} \leq \boldsymbol{b}_m, \quad m \in \mathcal{M} \end{aligned} \tag{14}$$

where $\mathcal{L}_h = \{0,1,\ldots,L_h-1\}$ and $\mathcal{M} = \{0,1,\ldots,M-1\}$. The last three constraints, as well as, the cost function are convex. In order to solve this problem efficiently we need to show that the last two sets of constraints can be written as linear matrix inequalities (LMI). To this end, first notice that for $\boldsymbol{\rho}_m, \boldsymbol{\gamma}_m \geq 0$

$$\boldsymbol{\rho}_m\boldsymbol{\gamma}_m \geq 1 \Leftrightarrow \left\|\begin{bmatrix} 2 \\ \boldsymbol{\gamma}_m - \boldsymbol{\rho}_m \end{bmatrix}\right\| \leq \boldsymbol{\gamma}_m + \boldsymbol{\rho}_m \quad [3]$$

$$\Leftrightarrow \begin{bmatrix} \boldsymbol{\gamma}_m+\boldsymbol{\rho}_m & 2 & \boldsymbol{\gamma}_m-\boldsymbol{\rho}_m \\ 2 & \boldsymbol{\gamma}_m+\boldsymbol{\rho}_m & 0 \\ \boldsymbol{\gamma}_m-\boldsymbol{\rho}_m & 0 & \boldsymbol{\gamma}_m+\boldsymbol{\rho}_m \end{bmatrix} \succeq \boldsymbol{0}. \quad [3]$$

The last constraint in (14) is convex for low SNR as shown in this paper. Under condition (10), using the results from

[14] such constraint is equivalent to (for fixed m)

$$\left\{\begin{array}{c} -\sigma_x^2 + \boldsymbol{b}_m + (\sigma_x^2 \boldsymbol{c}_m - \sigma^2)\boldsymbol{z}_m^{(0)} \geq 0 \\ \left[\begin{array}{ccc} 1 & 1-\boldsymbol{z}_m^{(0)} & \boldsymbol{z}_m^{(1)} \\ 1-\boldsymbol{z}_m^{(0)} & -\boldsymbol{c}_m + \boldsymbol{\lambda}_m - \boldsymbol{c}_m \boldsymbol{z}_m^{(0)} & \boldsymbol{b}_m - \boldsymbol{c}_m \boldsymbol{z}_m^{(1)} \\ \boldsymbol{z}_m^{(1)} & \boldsymbol{b}_m - \boldsymbol{c}_m \boldsymbol{z}_m^{(1)} & \boldsymbol{z}_m^{(2)} \end{array}\right] \succeq \mathbf{0} \\ \boldsymbol{z}_m^{(0)} \leq 1 \end{array}\right.$$

for some $\boldsymbol{z}_m^{(0)}, \boldsymbol{z}_m^{(1)}, \boldsymbol{z}_m^{(2)} \in \mathbb{R}$.

Combining these results, we obtain the following semidefinite optimization problem:

$$\begin{array}{ll} \min & \mathbf{1}^T \boldsymbol{b} \\ \text{s.t.} & \mathbf{1}^T \boldsymbol{\gamma} \leq \mathcal{E}, \\ & \boldsymbol{\gamma} \geq \mathbf{0}, \\ & \boldsymbol{\lambda} = \boldsymbol{A}\boldsymbol{\rho}, \\ & \left[\begin{array}{ccc} \boldsymbol{\gamma}_m + \boldsymbol{\rho}_m & 2 & \boldsymbol{\gamma}_m - \boldsymbol{\rho}_m \\ 2 & \boldsymbol{\gamma}_m + \boldsymbol{\rho}_m & 0 \\ \boldsymbol{\gamma}_m - \boldsymbol{\rho}_m & 0 & \boldsymbol{\gamma}_m + \boldsymbol{\rho}_m \end{array}\right] \succeq \mathbf{0}, \quad m \in \mathcal{L}_h \\ & -\sigma_x^2 + \boldsymbol{b}_m + (\sigma_x^2 \boldsymbol{c}_m - \sigma^2)\boldsymbol{z}_m^{(0)} \geq 0, \quad m \in \mathcal{M} \\ & \left[\begin{array}{ccc} 1 & 1-\boldsymbol{z}_m^{(0)} & \boldsymbol{z}_m^{(1)} \\ 1-\boldsymbol{z}_m^{(0)} & -\boldsymbol{c}_m + \boldsymbol{\lambda}_m - \boldsymbol{c}_m \boldsymbol{z}_m^{(0)} & \boldsymbol{b}_m - \boldsymbol{c}_m \boldsymbol{z}_m^{(1)} \\ \boldsymbol{z}_m^{(1)} & \boldsymbol{b}_m - \boldsymbol{c}_m \boldsymbol{z}_m^{(1)} & \boldsymbol{z}_m^{(2)} \end{array}\right] \succeq \mathbf{0}, \\ & m \in \mathcal{M} \\ & \boldsymbol{z}^{(0)} \leq 1 \end{array} \tag{15}$$

where the minimization is with respect to $\boldsymbol{b}, \boldsymbol{\gamma}, \boldsymbol{\lambda}, \boldsymbol{\rho}, \boldsymbol{z}^{(0)}, \boldsymbol{z}^{(1)}, \boldsymbol{z}^{(2)} \in \mathbb{R}^M$.

Remark: In this paper, we focus on the case of the *useful* training energy, i.e., ignoring the CP part. This choice is justified due to the following reasons. First, the analysis will be much simpler and the interest on the application-oriented nature of the proposed design will be better highlighted. Second, we do not gain much in terms of the channel estimation MSE-training energy tradeoff if we consider the CP, as it has been demonstrated both analytically and numerically in [8], [9], while the occurring gain may come at the cost of a high peak-to-average-power ratio (PAPR). Considering only the useful training energy, we do not impose constraints on the phases of the pilots, thus promoting a potentially lower PAPR.

We now examine the case of sufficiently high SNR. This occurs by selecting σ_x^2 as follows:

$$\sigma_x^2 > \overline{\sigma_x^2} = \frac{\sigma^2}{\min\{c_0, c_1, \ldots, c_{M-1}\}}. \tag{16}$$

This choice leads to $\sigma_x^2 c_m - \sigma^2 > 0, \forall m$ and

$$\frac{\partial \left[\text{MSE(ZF)}\right]_0}{\partial \boldsymbol{\lambda}_m} > 0, \quad m \in \mathcal{M} = \{0, 1, \ldots, M-1\}.$$

The corresponding training optimization problem is formulated as in (11). Notice that the objective is concave with respect to the $\boldsymbol{\lambda}_m$'s when (16) holds. To see this, observe that

$$\frac{\partial^2 \left[\text{MSE(ZF)}\right]_0}{\partial \boldsymbol{\lambda}_m \partial \boldsymbol{\lambda}_j} = 0, \quad m \neq j$$

and

$$\frac{\partial^2 \left[\text{MSE(ZF)}\right]_0}{\partial \boldsymbol{\lambda}_m^2} = -2\frac{\sigma_x^2 \boldsymbol{c}_m - \sigma^2}{(\boldsymbol{c}_m + \boldsymbol{\lambda}_m)^3} < 0, \quad \forall m$$

i.e., its Hessian is negative definite. Nevertheless, the training design problem can be written in this case as

$$\begin{array}{ll} \min_{\boldsymbol{b}, \boldsymbol{\lambda} \boldsymbol{\rho}, \boldsymbol{\gamma}} & \mathbf{1}^T \boldsymbol{b} \\ \text{s.t.} & \mathbf{1}^T \boldsymbol{\gamma} \leq \mathcal{E}, \\ & \boldsymbol{\lambda} = \boldsymbol{A}\boldsymbol{\rho}, \\ & \sigma_x^2 \boldsymbol{\lambda}_m + \sigma^2 = \boldsymbol{b}_m \boldsymbol{c}_m + \boldsymbol{b}_m \boldsymbol{\lambda}_m, \quad m \in \mathcal{M} \\ & \boldsymbol{\rho}_m \boldsymbol{\gamma}_m = 1, \quad m \in \mathcal{L}_h \\ & \boldsymbol{\gamma} \geq \mathbf{0}. \end{array} \tag{17}$$

To convexify the last formulation, we have to appropriately handle the products $\boldsymbol{\rho}_m \boldsymbol{\gamma}_m$ and $\boldsymbol{b}_m \boldsymbol{\lambda}_m$. Relaxing the equality $\boldsymbol{\rho}_m \boldsymbol{\gamma}_m = 1$ to $\boldsymbol{\rho}_m \boldsymbol{\gamma}_m \leq 1$, we may use the Schur complement to write:

$$\left[\begin{array}{cc} 1 & \boldsymbol{\rho}_m \\ \boldsymbol{\gamma}_m & 1 \end{array}\right] \succeq \mathbf{0}.$$

The problem is that the left hand side matrix is not symmetric. We may therefore consider its symmetric part, i.e.,

$$\frac{1}{2}\left(\left[\begin{array}{cc} 1 & \boldsymbol{\rho}_m \\ \boldsymbol{\gamma}_m & 1 \end{array}\right] + \left[\begin{array}{cc} 1 & \boldsymbol{\rho}_m \\ \boldsymbol{\gamma}_m & 1 \end{array}\right]^T\right) \succeq \mathbf{0}.$$

This approximation can be easily seen to correspond to bounding $\boldsymbol{\rho}_m \boldsymbol{\gamma}_m$ by $(\boldsymbol{\rho}_m + \boldsymbol{\gamma}_m)^2/4$ and then requiring that $(\boldsymbol{\rho}_m + \boldsymbol{\gamma}_m)^2/4 \leq 1$. As far as $\boldsymbol{b}_m \boldsymbol{\lambda}_m$ is concerned, we may replace it by the auxiliary variable $\boldsymbol{z}_m$, while we may relax this equality to the inequality $\boldsymbol{z}_m \geq \boldsymbol{b}_m \boldsymbol{\lambda}_m$. Using again the same Schur-complement treatment as before, we obtain the constraint

$$\frac{1}{2}\left(\left[\begin{array}{cc} \boldsymbol{z}_m & \boldsymbol{b}_m \\ \boldsymbol{\lambda}_m & 1 \end{array}\right] + \left[\begin{array}{cc} \boldsymbol{z}_m & \boldsymbol{b}_m \\ \boldsymbol{\lambda}_m & 1 \end{array}\right]^T\right) \succeq \mathbf{0}.$$

Given all the above approximations, Problem (17) can be approximated by the following convex program:

$$\begin{array}{ll} \min\limits_{\boldsymbol{\gamma}} & \mathbf{1}^T \boldsymbol{b} \\ \text{s.t.} & \mathbf{1}^T \boldsymbol{\gamma} \leq \mathcal{E}, \\ & \boldsymbol{\lambda} = \boldsymbol{A}\boldsymbol{\rho}, \\ & \sigma_x^2 \boldsymbol{\lambda}_m + \sigma^2 = \boldsymbol{b}_m \boldsymbol{c}_m + \boldsymbol{z}_m, \quad m \in \mathcal{M} \\ & \frac{1}{2}\left(\left[\begin{array}{cc} 1 & \boldsymbol{\rho}_m \\ \boldsymbol{\gamma}_m & 1 \end{array}\right] + \left[\begin{array}{cc} 1 & \boldsymbol{\rho}_m \\ \boldsymbol{\gamma}_m & 1 \end{array}\right]^T\right) \succeq \mathbf{0}, m \in \mathcal{L}_h, \\ & \frac{1}{2}\left(\left[\begin{array}{cc} \boldsymbol{z}_m & \boldsymbol{b}_m \\ \boldsymbol{\lambda}_m & 1 \end{array}\right] + \left[\begin{array}{cc} \boldsymbol{z}_m & \boldsymbol{b}_m \\ \boldsymbol{\lambda}_m & 1 \end{array}\right]^T\right) \succeq \mathbf{0}, m \in \mathcal{M}, \\ & \boldsymbol{\gamma} \geq \mathbf{0}, \quad \boldsymbol{b} \geq \mathbf{0}, \quad \boldsymbol{z} \geq \mathbf{0}, \quad \boldsymbol{\rho} \geq \mathbf{0} \end{array} \tag{18}$$

Note that the constraints $\boldsymbol{\gamma} \geq \mathbf{0}, \boldsymbol{b} \geq \mathbf{0}, \boldsymbol{z} \geq \mathbf{0}$ and $\boldsymbol{\rho} \geq \mathbf{0}$ have been added to keep the solution meaningful due to all the approximations we encountered in the last formulation.

Remark: The constraints $\boldsymbol{\rho}_m \boldsymbol{\gamma}_m \leq 1$ are nonconvex. The corresponding semidefinite constraints that we introduced in their position are convex approximations of them. The constraints $\boldsymbol{\rho}_m \boldsymbol{\gamma}_m \geq 1$ are instead convex. We could replace them by $\ln(\boldsymbol{\rho}_m) + \ln(\boldsymbol{\gamma}_m) \geq 0$. Having both the semidefinite and the $\ln$ constraints in the last formulation, the problem may become infeasible. We have tested that using either the semidefinite or the $\ln$ constraints, the final symbol estimate MSE performance is not affected, i.e., the corresponding solutions produced by each of these programs behave in approximately the same fashion.

Finally, the remaining SNR regime is defined by the interval

$$\underline{\sigma_x^2} \leq \sigma_x^2 \leq \overline{\sigma_x^2}.$$

In this case, the Hessian of $[\mathrm{MSE(ZF)}]_0$ is indefinite. If for some m, $\sigma_x^2 \boldsymbol{c}_m - \sigma^2 > 0$, then the unconstrained minimum value of the corresponding $\boldsymbol{b}_m$ is $\sigma^2/\boldsymbol{c}_m$. In the opposite case, the minimum value is σ_x^2. Nevertheless, formulation (18) can be used in this case as well.

To determine the optimal pilot placement, a sequence of $C(M, L_h)$ optimization problems has to be solved, each one corresponding to a different pilot placement and the best solution to be adopted. This task implies high computational complexity and hence, potentially, large delay. To resolve this problem, we can follow two approaches. The first is to fix the pilot placement to one for which we know its optimality in the classical channel estimation setup, i.e., the equidistant one and then solve the corresponding optimization problem to determine the optimal energy allocation to the nonzero pilots. The second leads to a closed form solution. We impose a *fairness* condition, i.e., that all $\boldsymbol{\lambda}_m$'s are equal to $\overline{\lambda}$. It then follows that the equidistant placement in combination with equipowered pilot tones is an optimal solution. The optimal pilot moduli in this case are given by

$$|[\boldsymbol{p}_s]_k| = \sigma\sqrt{\frac{1}{\overline{\lambda}}}, k = 0, 1, \ldots, L_h - 1. \qquad (19)$$

Remarks:

1) The choice of equidistant placement of the tones is not only justified by the fairness condition. In practice, there might be invertibility problems of $\boldsymbol{F}_{L_h \times L_h}$ based on any other pilot placement.

2) The most important remark in this section refers to the implementation of the corresponding convex approximations. These approximations depend on the true CFR coefficients. Nevertheless, their use is only to verify that the sparse preamble with equidistant and equipowered pilots are near optimal. Assuming that we knew the true second order statistics of the CFR coefficients, we could replace $|\boldsymbol{H}_m|^2$ by $E[|\boldsymbol{H}_m|^2]$ in the corresponding formulations. We follow this approach in the simulation section. This strengthens even more the near optimality of the aforementioned preambles, since their performance is compared against a "genie-aided" scheme.

V. Simulations

In this section we present numerical results to verify our analysis. In all figures, $\boldsymbol{h} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{C})$, where $\boldsymbol{C}_{i,j} = r^{j-i}, \quad j \geq i$, with $r = 0.9$ and i.i.d. QPSK symbols are assigned to all subcarriers. The energy during training highlights how good the channel estimate is. The parameter χ_m has been empirically selected to be 0.2. All schemes in Figs. 1, 2 and 3 use the same χ_m for all m.

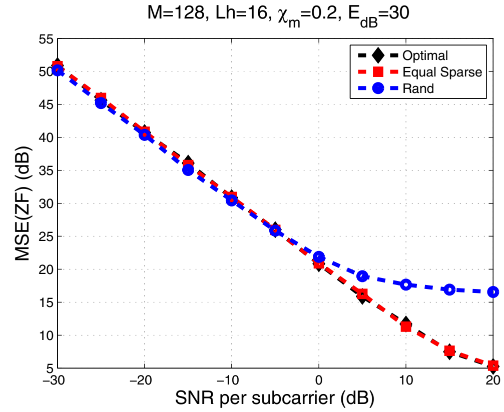

Fig. 1. Sparse Preamble $M = 128$, $L_h = 16$ and $\chi_m = 0.2$, $\forall m$: [MSE(ZF)] with training energy equal to 30 dB.

In Figs. 1 and 2, $[\mathrm{MSE(ZF)}]$ versus the SNR per subcarrier during data transmission is presented for energy during training equal to 30 dB, when $M = 128, L_h = 16$ and $M = 256, L_h = 64$ respectively. "Optimal" is the preamble vector produced by the formulations in this paper. The sparse preamble with equidistant and equipowered pilots, i.e., the "Equal Sparse", is better than a sparse preamble with equidistant and random pilots, i.e., the "Rand", after the SNR value per subcarrier equal to 0 (Fig. 1) and -5 dB (Fig. 2), while it coincides with the "Optimal" preamble. Note that the performance of the "Rand" preamble seems to be better in the low SNR regime. Nevertheless, notice that the y-axis corresponds to the true symbol estimate MSE and not to $[\mathrm{MSE(ZF)}]_0$. Additionally, the performance of the random preamble quickly reaches a floor value, the existence of which can be justified based on (9) as $\sigma \to 0$.

Finally, Fig. 3 aims at highlighting a crucial point: the preambles produced by the optimization problems in this paper do not correspond to one with equal pilot moduli. Instead, ignoring the corner effects (corner moduli), the moduli produced are around the nominal line of equal moduli. The mapped performance of such preambles through the ZF symbol estimate MSE prism is approximately the same with that of the equal pilot moduli preamble. In this figure, $M = 64$, $L_h = 16$ and $\chi_m = 0.2, \quad \forall m$, while the instantiation has been produced from a high SNR regime solution due to formulation (18).

VI. Conclusions

In this paper, application-oriented preamble selection for CP-OFDM systems has been investigated, when the employed channel estimator is the LS. We have highlighted the fact that the application-oriented preamble selection should be the appropriate way to perform training sequence design in practice. Additionally, we have verified that for the

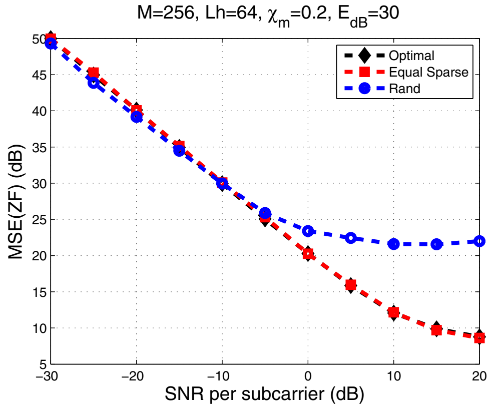

Fig. 2. Sparse Preamble $M = 256$, $L_h = 64$ and $\chi_m = 0.2$, $\forall m$: [MSE(ZF)] with training energy equal to 30 dB.

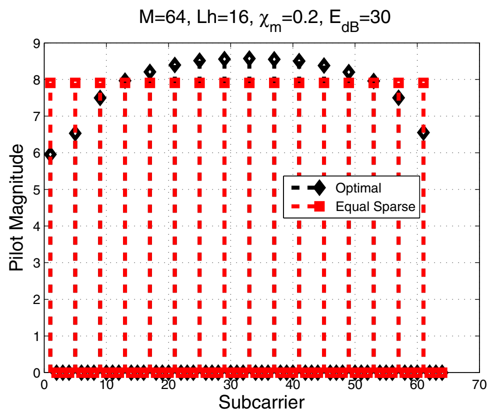

Fig. 3. Sparse Preamble $M = 64$, $L_h = 16$ and $\chi_m = 0.2$, $\forall m$: Pilot modulus vs. subcarrier index with training energy equal to 30 dB.

symbol estimate MSE based on per subcarrier ZF equalizers performance metric, a class of near optimal sparse preambles corresponds to equipowered pilots. This result is quite satisfactory, since the aforementioned class of optimal preambles is optimal even in the classical LS training design setup based on the channel estimation MSE subject to a training energy constraint. Nevertheless, using other types of equalizers or channel estimators at the receiver or other assumptions in the system, e.g., concerning the employed performance metrics or the correlation of transmitted symbols, we may obtain different classes of optimal preambles. This strengthens even more the main contribution of this paper, i.e., the observation that training sequence design should be performed with respect to an end performance metric of interest rather than in the classical channel estimation MSE setup.

References

[1] J. G. Andrews, A. Ghosh, and R. Muhamed, *Fundamentals of WiMAX: Understanding Broadband Wireless Networking*, Prentice-Hall, 2007.

[2] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels", *IEEE Trans. Signal Processing*, volume 51, no. 6, pages 1615–1624, June 2003.

[3] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[4] S. Coleri, M. Ergen, A. Puri, A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems", *IEEE Trans. Broadcasting*, volume 48, no. 3, pages 223–229, Sept. 2002.

[5] H. Hjalmarsson, "System Identification of Complex and Structured Systems", *Plenary address Euroepan Control Conference/European Journal Control*, volume 15, no. 4, pages 275–310, 2009.

[6] E. Jorswieck, H. Boche, *Majorization and Matrix-Monototne Functions in Wireless Communications*, Foundations and Trends in Communications and Information Theory, 2006.

[7] D. Katselis, E. Kofidis, A. Rontogiannis, S. Theodoridis, "Preamble-based channel estimation for CP-OFDM and OFDM/OQAM systems: A comparative study", *IEEE Trans. Signal Processing*, volume 58, no. 5, pages 2911–2916, May 2010.

[8] D. Katselis, E. Kofidis, A. Rontogiannis, S. Theodoridis, "Preamble-Based Channel Estimation for CP-OFDM and OFDM/OQAM Systems: A Comparative study", http://arxiv.org/abs/0910.3928, extended version of [7].

[9] D. Katselis, "Some Preamble Design Aspects in CP-OFDM Systems", *IEEE Communications Letters*, volume 16, no. 3, pages 356-359, Mars 2012.

[10] D. Katselis, C. R. Rojas, H. Hjalmarsson, M. Bengtsson, "A Chernoff Convexification for Chance Constrained MIMO Training Sequence Design", *Proc. SPAWC 2012*, Cesme, Turkey, June 2012.

[11] D. Katselis, C. R. Rojas, H. Hjalmarsson, M. Bengtsson, "Application-Oriented Finite Sample Experiment Design: A Semidefinite Relaxation Approach", *Proc. SYSID 2012*, Brussels, Belgium, July 2012.

[12] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems", *IEEE Trans. Signal Processing,* volume 49, no. 12, pages 3065–3073, Dec. 2001.

[13] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system", *IEEE Trans. Consumer Electronics,* volume 44, no. 3, pages 1122–1128, Aug. 1998.

[14] J. Nie, "First Order Conditions for Semidefinite Representations of Convex Sets defined by Rational or Singular Polynomials", *Math. Program., Ser. A*, vol. 131: 1–36, 2012.

[15] J. G. Proakis, *Digital Communications*, McGraw-Hill, 3rd edition, 1995.

[16] R. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*, Artech House Publ., 2000.