

Application-Oriented Least Squares Experiment Design in Multicarrier Communication Systems^{*}

Dimitrios Katselis^{*} Cristian R. Rojas^{*} Håkan Hjalmarsson^{*}

^{*} ACCESS Linnaeus Center, School of Electrical Engineering, KTH
Royal Institute of Technology, SE 100-44, Stockholm, Sweden. (e-mail:
dimitrik@kth.se, cristian.rojas@ee.kth.se, hjalmar@kth.se)

Abstract: Recently, the application-oriented framework for pilot design in communication systems has been introduced. 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). In this perspective, the known pilot sequences that are optimal for any communication system and for any estimation task have to be reexamined. In this paper, the problem of training pilot design for the task of channel estimation in cyclic prefixed orthogonal frequency division multiplexing (CP-OFDM) systems is revisited. So far, the optimal training sequences for least squares (LS) channel estimation with respect to minimizing the channel MSE under a training 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 the optimal full preamble, i.e, the preamble employing pilots on all subcarriers, for LS channel estimation in its classical context are near optimal in the aforementioned application-oriented context for the symbol estimate MSE in certain target signal-to-noise ratio (SNR) operating intervals.

1. INTRODUCTION

Cyclic prefixed orthogonal frequency division multiplexing (CP-OFDM) is currently enjoying popularity in both wired and wireless communication systems [Andrews, 2007], mainly because of its immunity to multipath fading, which allows for a significant increase in the transmission rate [Van Nee et al., 2000]. 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% [Andrews, 2007]. 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 in [Katselis et al., 2012a,b]. The origins of this design can be found in pre-existing work in the system identification literature, e.g., in [Hjalmarsson, 2009] 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 [Katselis et al., 2012a,b].

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 *full* preamble, where *all* subcarriers carry pilots, is examined. By approximating the training design problem via a convex formulation, we *verify* that optimal full preambles in the classical estimation setup containing equipowered or equal pilots are near optimal in the application-oriented setup for high signal-to-noise ratios (SNR).

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 [Negi et al., 1998], it is shown that uniform spacing of L_h pilot tones (L_h being the channel impulse response (CIR)

^{*} This work was partially supported by the Swedish Research Council under contract 621-2010-5819.

length) is the best choice given that the pilot tones are equipowered¹. Equispaced and equipowered pilot tones were shown in [Barhumi et al., 2003] 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 [Katselis et al., 2010, 2009]. Optimal full preambles 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 [Katselis et al., 2009].

The rest of the paper is organized as follows: Section 2 presents the system model for CP-OFDM and some results and definitions related to LS channel estimation in that system. Section 3 introduces some useful observations with respect to the symbol estimate MSE, while Section 4 investigates the optimal pilot design problem for the symbol estimate MSE when the employed preamble is full. Finally, simulations supporting the analysis in this paper are presented in Section 5, while Section 6 concludes this paper.

Notation. Vectors and matrices are denoted by bold lowercase and uppercase letters, respectively. Superscripts T and H stand for transposition and conjugate transposition. The complex conjugate of a complex number z is denoted by z^* . Also, $j = \sqrt{-1}$. $\|\cdot\|$ is the Euclidean norm and $|\cdot|$ is the complex modulus or the absolute value. For a matrix \mathbf{A} , $(\mathbf{A})_{i,j}$ denotes its (i,j) th entry and for a vector \mathbf{a} , \mathbf{a}_m or $[\mathbf{a}]_m$ denotes its m th entry. The expectation operator is denoted by $E(\cdot)$. \mathbf{I}_m denotes the m th 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. Moreover, \succeq stands for the semidefinite cone partial ordering, while \geq, \leq used with vectors denote elementwise orderings. For a vector \mathbf{a} , $|\mathbf{a}|$ is the vector of moduli or absolute values of its entries. Finally, $\mathcal{D}_{\mathbf{a}} = \text{diag}(\mathbf{a})$, i.e., $\mathcal{D}_{\mathbf{a}}$ is a diagonal matrix having as main diagonal entries the elements of the vector \mathbf{a} .

2. SYSTEM MODEL

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

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

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

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

Assume that the CP length is chosen to be the smallest possible one, namely equal to the channel order: $\nu = L_h - 1$ [Negi et al., 1998]. Moreover, perfect timing and frequency synchronization are assumed. The channel impulse response (CIR), $\mathbf{h} = [\mathbf{h}_0 \ \mathbf{h}_1 \ \dots \ \mathbf{h}_{L_h-1}]^T$, is assumed to

¹ This is no longer valid if there are suppressed (*virtual*) subcarriers. In such a case, the optimal placement is non-uniform [Morelli et al., 2001].

be constant over the duration of an OFDM symbol. The input to the OFDM demodulator, after the CP removal, can then be expressed as

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

where \mathcal{H} is the Toeplitz circulant matrix with its first row given by $[\mathbf{h}_0 \ \mathbf{0}_{1 \times (M-L_h)} \ \mathbf{h}_{L_h-1} \ \dots \ \mathbf{h}_2 \ \mathbf{h}_1]$ and \mathbf{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

$$\mathbf{y} = \frac{1}{\sqrt{M}} \mathcal{F} \mathbf{r} = \mathcal{D}_H \mathbf{x} + \boldsymbol{\eta}, \quad (2)$$

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

We focus now on the full preamble case. Denoting the received vector during training by \mathbf{y}^{tr} , the preamble vector by \mathbf{p} and the corresponding noise vector by $\boldsymbol{\eta}^{tr}$, the CIR estimate, in the LS sense, can then be computed as

$$\hat{\mathbf{h}} = (1/M) \mathbf{F}_{M \times L_h}^H \tilde{\mathbf{H}}, \quad (3)$$

where $\tilde{\mathbf{H}}_m = \mathbf{y}_m^{tr} / \mathbf{p}_m = \mathbf{H}_m + \boldsymbol{\eta}_m^{tr} / \mathbf{p}_m$ and $\mathbf{F}_{M \times L_h}$ the $M \times L_h$ left submatrix of \mathcal{F} . Moreover, the final LS CFR estimates are given by

$$\hat{\mathbf{H}} = \mathbf{F}_{M \times L_h} \hat{\mathbf{h}}. \quad (4)$$

3. 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 $\hat{\mathbf{x}}_m$ equal to $\mathbf{y}_m / \hat{\mathbf{H}}_m$ 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:

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

Depending on the probability distribution of $|\hat{\mathbf{H}}_m|$'s, (5) may fail to exist. The MSE(ZF) will be finite if and only if the probability distribution function (pdf) of $|\hat{\mathbf{H}}_m|$ is of order $O(|\hat{\mathbf{H}}_m|^2)$ for all m as $\hat{\mathbf{H}}_m \rightarrow 0$. To make our analysis possible, we may use the following lemma:

Lemma 1. Given two random variables m, n and assuming that n has support in $[\epsilon, +\infty)$ for some $\epsilon > 0$, the mean value of their ratio can be expressed as follows:

$$(i) \quad E \left[\frac{m}{n} \right] = \frac{E[m]}{E[n]} \left(1 + \sum_{k=1}^{\infty} (-1)^k \frac{E[(n - E[n])^k]}{(E[n])^k} \right), \quad (6)$$

if m, n are independent.

(ii)

$$E\left[\frac{m}{n}\right] = \frac{E[m]}{E[n]} \left(1 + \sum_{k=1}^{\infty} (-1)^k \frac{E[(n - E[n])^k]}{(E[n])^k} + \sum_{k=1}^{\infty} (-1)^k \frac{E[(n - E[n])^k (m - E(m))]}{(E[n])^k E(m)} \right), \quad (7)$$

if m, n are not independent.

Proof See the Appendix A. ■

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:

$$\hat{\mathbf{H}}_m = \begin{cases} \hat{\mathbf{H}}_m, & \text{if } |\hat{\mathbf{H}}_m| > \chi_m \\ \chi_m \hat{\mathbf{H}}_m / |\hat{\mathbf{H}}_m|, & \text{otherwise} \end{cases}, \quad (8)$$

where $\hat{\mathbf{H}}_m$ can be any estimator and χ_m 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 $\hat{\mathbf{H}}_m$ preserves the continuity at $|\hat{\mathbf{H}}_m| = \chi_m$. Additionally, the event $\{\hat{\mathbf{H}}_m = 0\}$ has zero probability since the distribution of $\hat{\mathbf{H}}_m$ is continuous. Therefore, $\hat{\mathbf{H}}_m$ can be arbitrarily defined when $\hat{\mathbf{H}}_m = 0$, e.g., $\hat{\mathbf{H}}_m = 1$.

Assume that for given \mathbf{H}_m 's we choose $\chi_m \ll |\mathbf{H}_m|, \forall m$ and consider a high SNR during training. Then, applying Lemma 1 to (5), it can be shown that the following performance metric is a good approximation of (5):

$$[\text{MSE}(\text{ZF})]_0 = \sum_{m=0}^{M-1} \left\{ \sigma_x^2 \frac{E\left[|\hat{\mathbf{H}}_m - \mathbf{H}_m|^2\right]}{E\left[|\hat{\mathbf{H}}_m|^2\right]} + \sigma^2 \frac{1}{E\left[|\hat{\mathbf{H}}_m|^2\right]} \right\} \quad (9)$$

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 preamble, as well as other known preambles in the literature will be then numerically compared against the *exact symbol estimate MSE*.

4. THE FULL PREAMBLE CASE

Consider the full preamble case. Then, (4) can be written as follows:

$$\hat{\mathbf{H}} = \underbrace{\frac{1}{M} \mathbf{F}_{M \times L_h} \mathbf{F}_{M \times L_h}^H \mathbf{H}}_{\mathbf{H}} + \underbrace{\frac{1}{M} \mathbf{F}_{M \times L_h} \mathbf{F}_{M \times L_h}^H \mathcal{D}_{\mathbf{p}}^{-1} \boldsymbol{\eta}^{tr}}_{\mathbf{z}} = \mathbf{H} + \mathbf{z}, \quad (10)$$

where in the case of \mathbf{H} we have used the fact that the channel length is known. Using our previous assumptions, we can see that

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

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

$$[\text{MSE}(\text{ZF})]_0 = \sum_{m=0}^{M-1} \frac{\sigma_x^2 E\left[|\mathbf{z}_m|^2\right] + \sigma^2}{|\mathbf{H}_m|^2 + E\left[|\mathbf{z}_m|^2\right]}. \quad (11)$$

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

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

These partial derivatives will be positive if $\sigma_x^2 \mathbf{c}_m - \sigma^2 > 0$ for every m . Clearly, this can be guaranteed if we set the signal-to-noise ratio (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:

$$\sigma_x^2 < \frac{\sigma^2}{\max\{\mathbf{c}_0, \mathbf{c}_1, \dots, \mathbf{c}_{M-1}\}}. \quad (12)$$

In this case,

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

and

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

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

$$\begin{aligned} & \min_{\{\mathbf{p}_k\}_{k=0}^{M-1}} [\text{MSE}(\text{ZF})]_0 \\ & \text{s.t.} \quad \sum_{k=0}^{M-1} |\mathbf{p}_k|^2 \leq \mathcal{E}, \end{aligned} \quad (13)$$

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, by (10) we can write:

$$\begin{aligned} \boldsymbol{\lambda}_m &= \frac{\sigma^2}{M^2} \sum_{k=0}^{M-1} \frac{1}{|\mathbf{p}_k|^2} \sum_{i=0}^{L_h-1} \sum_{j=0}^{L_h-1} e^{j \frac{2\pi}{M} (i-j)(k-m)} \\ &= \frac{\sigma^2}{M^2} \sum_{k=0}^{M-1} \frac{1}{|\mathbf{p}_k|^2} \left| \sum_{i=0}^{L_h-1} e^{j \frac{2\pi}{M} i(k-m)} \right|^2. \end{aligned} \quad (14)$$

Collecting all λ_m 's in one vector, we obtain:

$$\lambda = \mathbf{A}\rho, \quad (15)$$

where $\rho_m = 1/|\mathbf{p}_m|^2$, $m = 0, 1, \dots, M-1$ and $(\mathbf{A})_{m,k} = (\sigma^2/M) \left| \sum_{i=0}^{L_h-1} e^{j\frac{2\pi}{M}i(k-m)} \right|^2$ for $m, k = 0, 1, \dots, M-1$. Notice that \mathbf{A} is a Gramian matrix formed by inner products from the set of vectors $(\sigma/M)[1, e^{j(2\pi/M)k}, \dots, e^{j(2\pi/M)(L_h-1)k}]^T$, $k = 0, 1, \dots, M-1$.

We now focus on formulating an appropriate training optimization problem. We define the vector γ with entries $\gamma_m = |\mathbf{p}_m|^2$, $m = 0, 1, \dots, M-1$. In order to solve (13) we can pose the following optimization problem:

$$\begin{aligned} \min_{\mathbf{b}, \gamma, \lambda, \rho} \quad & \mathbf{1}^T \mathbf{b} \\ \text{s.t.} \quad & \mathbf{1}^T \gamma \leq \mathcal{E}, \\ & \gamma \geq \mathbf{0}, \\ & \lambda = \mathbf{A}\rho, \\ & \rho_m \gamma_m \geq 1, \quad m \in \mathcal{M} \\ & \frac{\sigma_x^2 \lambda_m + \sigma^2}{\mathbf{c}_m + \lambda_m} \leq \mathbf{b}_m, \quad m \in \mathcal{M} \end{aligned} \quad (16)$$

where $\mathcal{M} = \{0, 1, \dots, M-1\}$. 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 $\rho_m, \gamma_m \geq 0$

$$\begin{aligned} \rho_m \gamma_m \geq 1 & \Leftrightarrow \left\| \begin{bmatrix} 2 \\ \gamma_m - \rho_m \end{bmatrix} \right\| \leq \gamma_m + \rho_m \\ & \Leftrightarrow \begin{bmatrix} \gamma_m + \rho_m & 2 & \gamma_m - \rho_m \\ 2 & \gamma_m + \rho_m & 0 \\ \gamma_m - \rho_m & 0 & \gamma_m + \rho_m \end{bmatrix} \succeq \mathbf{0}. \end{aligned}$$

The last constraint in (16) is convex for low SNR as shown in this paper. Under condition (12), using the results from [Nie, 2012] such constraint is equivalent to (for fixed m)

$$\left\{ \begin{array}{l} -\sigma_x^2 + \mathbf{b}_m + (\sigma_x^2 \mathbf{c}_m - \sigma^2) \mathbf{z}_m^{(0)} \geq 0 \\ \begin{bmatrix} 1 & 1 - \mathbf{z}_m^{(0)} & \mathbf{z}_m^{(1)} \\ 1 - \mathbf{z}_m^{(0)} & -\mathbf{c}_m + \lambda_m - \mathbf{c}_m \mathbf{z}_m^{(0)} & \mathbf{b}_m - \mathbf{c}_m \mathbf{z}_m^{(1)} \\ \mathbf{z}_m^{(1)} & \mathbf{b}_m - \mathbf{c}_m \mathbf{z}_m^{(1)} & \mathbf{z}_m^{(2)} \end{bmatrix} \succeq \mathbf{0} \\ \mathbf{z}_m^{(0)} \leq 1 \end{array} \right.$$

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

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

$$\begin{aligned} \min \quad & \mathbf{1}^T \mathbf{b} \\ \text{s.t.} \quad & \mathbf{1}^T \gamma \leq \mathcal{E}, \\ & \gamma \geq \mathbf{0}, \\ & \lambda = \mathbf{A}\rho, \\ & \begin{bmatrix} \gamma_m + \rho_m & 2 & \gamma_m - \rho_m \\ 2 & \gamma_m + \rho_m & 0 \\ \gamma_m - \rho_m & 0 & \gamma_m + \rho_m \end{bmatrix} \succeq \mathbf{0}, \quad m \in \mathcal{M} \\ & -\sigma_x^2 + \mathbf{b}_m + (\sigma_x^2 \mathbf{c}_m - \sigma^2) \mathbf{z}_m^{(0)} \geq 0, \quad m \in \mathcal{M} \\ & \begin{bmatrix} 1 & 1 - \mathbf{z}_m^{(0)} & \mathbf{z}_m^{(1)} \\ 1 - \mathbf{z}_m^{(0)} & -\mathbf{c}_m + \lambda_m - \mathbf{c}_m \mathbf{z}_m^{(0)} & \mathbf{b}_m - \mathbf{c}_m \mathbf{z}_m^{(1)} \\ \mathbf{z}_m^{(1)} & \mathbf{b}_m - \mathbf{c}_m \mathbf{z}_m^{(1)} & \mathbf{z}_m^{(2)} \end{bmatrix} \succeq \mathbf{0}, \\ & m \in \mathcal{M} \\ & \mathbf{z}_m^{(0)} \leq 1 \end{aligned} \quad (17)$$

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

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

$$\sigma_x^2 > \overline{\sigma_x^2} = \frac{\sigma^2}{\min\{c_0, c_1, \dots, c_{M-1}\}}. \quad (18)$$

This choice leads to $\sigma_x^2 \mathbf{c}_m - \sigma^2 > 0, \forall m$ and

$$\frac{\partial [\text{MSE}(\text{ZF})]_0}{\partial \lambda_m} > 0, \quad m \in \mathcal{M}.$$

Therefore, the value of $[\text{MSE}(\text{ZF})]_0$ is minimized when the λ_m 's take on their lowest possible values. These values are directly related to the available energy during training. Letting the available training energy become arbitrarily large is equivalent to letting $\lambda_m \rightarrow 0, \forall m$. Thus, we obtain:

$$[\text{MSE}(\text{ZF})]_0 \geq \sigma^2 \sum_{m=0}^{M-1} \frac{1}{\mathbf{c}_m}. \quad (19)$$

Our optimization problem can be formulated as Problem (13). Notice that the objective is concave with respect to the λ_m 's when (18) holds. To see this, observe that

$$\frac{\partial^2 [\text{MSE}(\text{ZF})]_0}{\partial \lambda_m \partial \lambda_j} = 0, \quad m \neq j$$

and

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

i.e., its Hessian is negative definite. Nevertheless, the training design problem (13) can be written in this case as

$$\begin{aligned} \min_{\mathbf{b}, \lambda, \rho, \gamma} \quad & \mathbf{1}^T \mathbf{b} \\ \text{s.t.} \quad & \mathbf{1}^T \gamma \leq \mathcal{E}, \\ & \lambda = \mathbf{A}\rho, \\ & \sigma_x^2 \lambda_m + \sigma^2 = \mathbf{b}_m \mathbf{c}_m + \mathbf{b}_m \lambda_m, \quad m \in \mathcal{M} \\ & \rho_m \gamma_m = 1, \quad m \in \mathcal{M} \\ & \gamma \geq \mathbf{0}, \end{aligned} \quad (20)$$

To convexify the last formulation, we have to appropriately handle the products $\rho_m \gamma_m$ and $\mathbf{b}_m \lambda_m$. Relaxing the equality $\rho_m \gamma_m = 1$ to $\rho_m \gamma_m \leq 1$, we may use the Schur complement to write:

$$\begin{bmatrix} 1 & \rho_m \\ \gamma_m & 1 \end{bmatrix} \succeq \mathbf{0}.$$

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

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

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 $\mathbf{b}_m \boldsymbol{\lambda}_m$ is concerned, we may replace it by the auxiliary variable \mathbf{z}_m , while we may relax this equality to the inequality $\mathbf{z}_m \geq \mathbf{b}_m \boldsymbol{\lambda}_m$. Using again the same Schur-complement treatment as before, we obtain the constraint

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

Given all the above approximations, Problem (20) can be approximated by the following convex program:

$$\begin{aligned} & \min_{\boldsymbol{\gamma}} \mathbf{1}^T \mathbf{b} \\ & \text{s.t. } \mathbf{1}^T \boldsymbol{\gamma} \leq \mathcal{E}, \\ & \quad \boldsymbol{\lambda} = \mathbf{A} \boldsymbol{\rho}, \\ & \quad \sigma_x^2 \boldsymbol{\lambda}_m + \sigma^2 = \mathbf{b}_m \mathbf{c}_m + \mathbf{z}_m, \quad m \in \mathcal{M} \\ & \quad \frac{1}{2} \left(\begin{bmatrix} 1 & \boldsymbol{\rho}_m \\ \boldsymbol{\gamma}_m & 1 \end{bmatrix} + \begin{bmatrix} 1 & \boldsymbol{\rho}_m \\ \boldsymbol{\gamma}_m & 1 \end{bmatrix}^T \right) \succeq \mathbf{0}, \quad m \in \mathcal{M}, \\ & \quad \frac{1}{2} \left(\begin{bmatrix} \mathbf{z}_m & \mathbf{b}_m \\ \boldsymbol{\lambda}_m & 1 \end{bmatrix} + \begin{bmatrix} \mathbf{z}_m & \mathbf{b}_m \\ \boldsymbol{\lambda}_m & 1 \end{bmatrix}^T \right) \succeq \mathbf{0}, \quad m \in \mathcal{M}, \\ & \quad \boldsymbol{\gamma} \geq \mathbf{0}, \quad \mathbf{b} \geq \mathbf{0}, \quad \mathbf{z} \geq \mathbf{0}, \quad \boldsymbol{\rho} \geq \mathbf{0} \end{aligned} \quad (21)$$

Note that the constraints $\boldsymbol{\gamma} \geq \mathbf{0}, \mathbf{b} \geq \mathbf{0}, \mathbf{z} \geq \mathbf{0}$ 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 $\ln(\boldsymbol{\rho}_m) + \ln(\boldsymbol{\gamma}_m) \geq 0$. We cannot have both the semidefinite and the ln constraints in the last formulation because 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 $[\text{MSE}(\text{ZF})]_0$ is indefinite. If for some m , $\sigma_x^2 \mathbf{c}_m - \sigma^2 > 0$, then the unconstrained minimum value of the corresponding \mathbf{b}_m is σ^2/\mathbf{c}_m . In the opposite case, the minimum value is σ_x^2 . Nevertheless, formulation (21) can be used in this case as well.

Assume now that the minimum $[\text{MSE}(\text{ZF})]_0$ achieved by any of the above convex approximation is denoted by $\overline{[\text{MSE}(\text{ZF})]_0}$ and the corresponding vector of $\boldsymbol{\lambda}_m$'s by $\bar{\boldsymbol{\lambda}}$. To achieve this value or a lower value with some closed form solutions, we note that any $\overline{[\text{MSE}(\text{ZF})]_0}$ can be achieved at

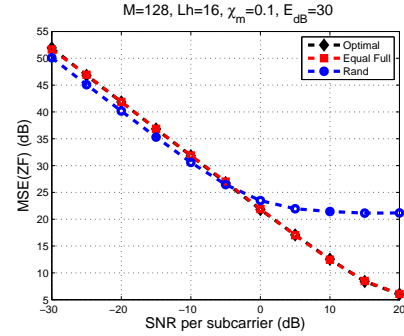


Fig. 1. Full Preamble with $M = 128$, $L_h = 16$ and $\chi_m = 0.1$, $\forall m$: $[\text{MSE}(\text{ZF})]$ with training energy equal to 30 dB.

the cost of an increased training energy by simply setting $\bar{\boldsymbol{\lambda}} = \min\{\bar{\boldsymbol{\lambda}}_0, \bar{\boldsymbol{\lambda}}_1, \dots, \bar{\boldsymbol{\lambda}}_{M-1}\}$ and requiring $\boldsymbol{\lambda}_m = \bar{\boldsymbol{\lambda}}, \forall m$. This can equivalently be thought as imposing a *fairness* condition on the channel estimates, i.e., as requiring that all channel estimates are constrained to suffer from the same MSE. Then by (14) it follows that one possible optimal choice of the pilot moduli corresponds to equipowered pilots. The common modulus equals $\sigma \sqrt{L_h/(M\bar{\boldsymbol{\lambda}})}$. Additionally, if we would like to minimize the CP energy, then the pilots should be chosen to be equal, i.e., to have the same phases or to be constructed by an algorithm presented in [Katselis et al., 2009, 2012b]. Nevertheless, the elimination of the CP energy comes at the cost of a high PAPR [Katselis et al., 2012b].

5. SIMULATIONS

In this section we present numerical results to verify our analysis. In all figures, $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C})$, where $\mathbf{C}_{i,j} = r^{j-i}$, $j \geq i$, 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 0.1. All schemes in Figs. 1 and 2 use the same χ_m for all m .

In Fig. 1, $[\text{MSE}(\text{ZF})]$ versus the SNR per subcarrier during data transmission is presented for energy during training equal to 30 dB, when $M = 128$ and $L_h = 16$. The full preamble with equipowered pilots is better than a full preamble with random pilots after SNR per subcarrier equal to -5 dB, while it coincides with the preamble produced by formulation (17). Note that the performance of the rand preamble seems to be better in the low SNR regime. Nevertheless, the y -axis corresponds to the true symbol estimate MSE and not to $[\text{MSE}(\text{ZF})]_0$. Additionally, the performance of the random preamble quickly reaches a floor value, the existence of which can be justified based on (11) as $\sigma \rightarrow 0$. Fig. 1 is a first numerical verification that performing our theoretical analysis based on $[\text{MSE}(\text{ZF})]_0$ instead of $[\text{MSE}(\text{ZF})]$ can actually yield valid results and conclusions.

Fig. 2 presents the corresponding results for $M = 256$ and $L_h = 64$. The conclusions coincide with the those following Fig. 1. An interesting observation is that the point that all curves coincide has now been shifted to the left.

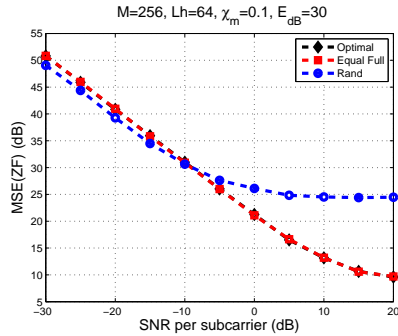


Fig. 2. Full Preamble $M = 256$, $L_h = 64$ and $\chi_m = 0.1$, $\forall m$: [MSE(ZF)] with training energy equal to 30 dB.

6. 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 symbol estimate MSE based on per subcarrier ZF equalizers performance metrics, a class of near optimal full 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, may lead to different classes of optimal preambles. This strengthens even further 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.

Appendix A. PROOF OF LEMMA 1

We first prove (ii). Assume that m, n are two random variables and n has either no mass at zero or it has support $[0, +\infty)$. Clearly, we can write:

$$m = E[m] + (m - E[m]) \quad (\text{A.1})$$

$$n = E[n] + (n - E[n]) \quad (\text{A.2})$$

Then,

$$E \left[\frac{m}{n} \right] = \frac{E[m]}{E[n]} E \left[\left(1 + \frac{m - E[m]}{E[m]} \right) \left(1 + \frac{n - E[n]}{E[n]} \right)^{-1} \right]$$

Using the Taylor series expansion theorem, we obtain:

$$\left(1 + \frac{n - E[n]}{E[n]} \right)^{-1} = 1 + \sum_{k=1}^{+\infty} (-1)^k \frac{(n - E[n])^k}{(E[n])^k}$$

Combining the above results, the desired expression follows.

Moreover, (i) follows from (ii) by assuming independence of the random variables.

REFERENCES

- J. G. Andrews, A. Ghosh, and R. Muhamed. *Fundamentals of WiMAX: Understanding Broadband Wireless Networking*. Prentice-Hall, 2007.
- 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.
- S. Boyd, L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- 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.
- H. Hjalmarsson. System Identification of Complex and Structured Systems. *Plenary address European Control Conference/European Journal Control*, volume 15, no. 4, pages 275–310, 2009.
- E. Jorswieck, H. Boche. *Majorization and Matrix-Monotone Functions in Wireless Communications*. Foundations and Trends in Communications and Information Theory, 2006.
- 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.
- 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 [Katselis et al., 2010].
- D. Katselis. Some Preamble Design Aspects in CP-OFDM Systems. *IEEE Communications Letters*, volume 16, no. 3, pages 356–359, Mars 2012.
- 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.
- 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.
- 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.
- 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.
- 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.
- J. G. Proakis. *Digital Communications*, McGraw-Hill, 3rd edition, 1995.
- R. van Nee and R. Prasad. *OFDM for Wireless Multimedia Communications*. Artech House Publ., 2000.