Training Symbol Design for Channel Estimation and IQ Imbalance Compensation in OFDM Systems

Emmanuel Manasseh, Shuichi Ohno and Masayoshi Nakamoto
Dept. of Artificial Complex Systems Engineering, Hiroshima University
1-4-1 Kagamiyama Higashi-Hiroshima, 739-8527, JAPAN
E-mail: {manassehjc,ohno,msy}@hiroshima-u.ac.jp

***Abstract*—In this paper, training symbol designs for estimation of frequency selective channels and compensation of in-phase (I) and quadrature (Q) imbalances on OFDM transmitters and receivers are studied. We utilize cross entropy (CE) optimization techniques together with convex optimization to design training sequence that minimizes the channel estimate mean squared error (MSE) as well as estimating the effect of I/Q mismatch while lowering the peak power of the training signals. The proposed design provide better channel estimate MSE and bit error rate (BER) performances. The efficacies of the proposed designs are corroborated by analysis and simulation results.**

I. Introduction

In recent years, the growth of multimedia services and applications of digital data transmissions has led to ever greater demands on throughput capacity of wired as well as wireless communication systems. The implementation of cost and power efficient transmission and reception systems are challenging following the impairments associated with the analog components. One such impairment is the imbalance between the inphase (I) and quadrature (Q) branches when the received radio-frequency (RF) signal is down-converted to baseband [1].

Direct conversion receivers (DCRs) are prominent candidates for physical layer applications due to its low-cost and low-power implementation on silicon, as compared to the heterodyne receivers. DCRs can convert the RF signals directly into baseband without any intermediate frequencies (IF). However, a problem with DCRs over heterodyne receivers is that the baseband signals are more severely distorted by imbalances within the I and Q branches [1], [2].

The I/Q imbalance is a critical impairment as it causes intercarrier interferences (ICI) in orthogonal frequency division multiplexing (OFDM) systems that results to severe degradation of the system's performance. The I/Q imbalances in general exist at both transmitter and receiver, and may vary over time (e.g., due to sudden changes in temperature or the dynamics of the cooperating/relaying users) [3]–[5].

The frequency selective nature of the channels together with I/Q imbalance are the main factors that degrade the performance of OFDM systems. Since I/Q imbalance problem is originated from the unavoidable differences in the relative amplitudes and phases of the physical analog I and Q signal paths, the effects of I/Q imbalances in OFDM systems cannot be rectified by increasing signal power, thus, to obtain better quality of the high rate communications systems, efficient channel estimation and I/Q imbalance compensation techniques are crucial [4], [5].

In the literature, training symbol designs for channel estimation and I/Q imbalance compensation for OFDM systems have been predominantly developed (see [2], [4]–[6] and the references there in). However, there is no existing method that optimally select phase information of the training (or pilot) symbols to reduce the peak to average power ratio (PAPR) effect. In [5] training symbol design for joint channel and I/Q imbalance estimation is presented. To minimize PAPR, phase information of the training symbols are randomly selected which does not guarantee a substantial PAPR reduction.

In [1], an estimator that utilizes special pilot pattern to estimate channel and I/Q imbalance is presented. The scheme effectively estimate I/Q imbalance by setting some of the pilot subcarriers to zero. Even though I/Q imbalance is efficiently estimated and compensated using the special pilot pattern in [1], but the channel estimate MSE is poor since channel is estimated by pilot symbols allocated on either one side of the active subcarrier band and other subcarriers are nulled.

This paper presents efficient training symbol design for the estimation of frequency-selective channels and I/Q imbalances in OFDM system while lowering the PAPR. We utilize cross entropy (CE) optimization techniques together with convex optimization to design training sequences that minimizes the channel mean squared error (MSE) as well as the the peak power of the training signals while suppressing the effect of I/Q imbalance. We propose a CE based algorithm for selecting position of training symbols when some active subcarriers are nulled. Unlike [1], where special pilot patterns are obtained by setting active subcarriers on the lower or upper side of the active band to zero, here we propose a scheme for selecting the position of training symbols to obtain better channel estimate MSE. The proposed designs provide better channel estimate MSE as well as the bit error rate (BER) performances for OFDM systems having different channel length. Simulation results are provided to substantiate the effectiveness of the proposed designs.

The following notations are used in the description of the system. The superscripts $*$, T and $\mathcal{H}$ represent conjugate, transpose and the conjugate transpose (hermitian) respectively. The operator $\otimes$, denotes the convolution operation. Other operators that are used will be defined whenever used.

978-1-4673-0990-5/12/$31.00

II. System Model and Channel estimation

We present a system model that captures the effects of frequency independent and frequency dependent I/Q imbalances at both transmitter and receiver side, for an OFDM system with a single transmit and a single receive antenna. The frequency-independent gain and phase offsets of the I and Q branches at the transmitter side are denoted as $\{a_t^I, a_t^Q\}$ and $\{\theta_t^I, \theta_t^Q\}$ respectively. The corresponding impulse shaping filters, which includes amplifiers, digital to analog converters (DAC), pulse shapers and frequency-dependent imbalances for the I and Q branches of the transmitter are denoted as $\{g_t^I(t), g_t^Q(t)\}$. The equivalent receiver side parameters are denoted in the same manner with subscript t replaced by r, i.e., a_r^I,a_r^Q, θ_r^I, θ_r^Q, $g_r^I(t)$ and $g_r^Q(t)$. The low pass equivalent channel is denoted as $h(t)$. The system with I/Q imbalances is composed of the direct and the mirror system. The impulse responses of the direct and mirror system at the transmitter and receiver are denoted as $g_T^D(t)$, $g_T^M(t)$ and $g_R^D(t)$, $g_R^M(t)$ respectively and are related to the I/Q imbalance parameters as [2]

$$g_T^D(t) = \frac{1}{2}\left[a_t^I e^{j\theta_t^I} g_t^I(t) + a_t^Q e^{j\theta_t^Q} g_t^Q(t)\right], \quad (1)$$

$$g_T^M(t) = \frac{1}{2}\left[a_t^I e^{j\theta_t^I} g_t^I(t) - a_t^Q e^{j\theta_t^Q} g_t^Q(t)\right], \quad (2)$$

$$g_R^D(t) = \frac{1}{2}\left[a_r^I e^{-j\theta_r^I} g_r^I(t) + a_r^Q e^{-j\theta_r^Q} g_r^Q(t)\right], \quad (3)$$

$$g_R^M(t) = \frac{1}{2}\left[a_r^I e^{j\theta_r^I} g_r^I(t) - a_r^Q e^{j\theta_r^Q} g_r^Q(t)\right]. \quad (4)$$

The equivalent direct and mirror channel including the transmit and receive filters as well as the channel $h(t)$ are given by

$$\xi(t) = g_T^D(t) \otimes h(t) \otimes g_R^D(t) + (g_T^M)^*(t) \otimes h^*(t) \otimes g_R^M(t), \quad (5)$$
$$\chi(t) = g_T^M(t) \otimes h(t) \otimes g_R^D(t) + (g_T^D)^*(t) \otimes h^*(t) \otimes g_R^M(t). \quad (6)$$

For an OFDM block with N number of subcarriers, we denote the frequency domain transmitted block of data as $\boldsymbol{X} = [X(1), \dots, X(N)]$. Each block is passed through the IFFT operation to obtain the time domain signal given by $\boldsymbol{x} = \boldsymbol{F}^{\mathcal{H}}\boldsymbol{X}$, where

$[F]_{m,n} = \frac{1}{\sqrt{N}} e^{\frac{-j2\pi mn}{N}}, \qquad m,n = \{0, 1, \cdots N-1\}$. The discrete time versions of the direct and the mirror channels $\xi(t)$ and $\chi(t)$ are denoted by $\boldsymbol{\xi}$ and $\boldsymbol{\chi}$ respectively and consist of maximum L taps. We assume that the number of the cyclic prefix samples N_{cp} is greater than channel length L (i.e., $N_{cp} \geq L$) in order to preserve the orthogonality of the tones. For a system with I/Q imbalance, if the direct channel sees an input signal $x(t)$, the mirror channel sees an input $x^*(t)$. Note that, the DFT of the complex conjugate of a sequence is related to the DFT of the original sequence through a mirrored relation (for $1 \leq n \leq N$ and $1 \leq k \leq N$) [1], that is

$$\begin{aligned} x(n) &\xrightarrow{DFT} X(k), \\ x^*(n) &\xrightarrow{DFT} X^*(N-k+2). \end{aligned} \quad (7)$$

The time-domain received signal vector $\boldsymbol{y}$ for one OFDM symbol after CP removal can be written as

$$\boldsymbol{y} = \sqrt{N}\boldsymbol{F}^{\mathcal{H}}\left(D(\boldsymbol{X})\boldsymbol{F}_L\boldsymbol{\xi} + D(\boldsymbol{X}^{\#})\boldsymbol{F}_L^*\boldsymbol{\chi}\right) + \boldsymbol{v}, \quad (8)$$

where $D(\cdot)$ denotes the diagonal matrix of the vector and $\boldsymbol{F}_L$ is an $N \times L$ sub-matrix of the DFT matrix $\boldsymbol{F}$. The elements of the mirror vector $\boldsymbol{X}^{\#}$ for $\boldsymbol{X} = \{X(1), X(2), \dots, X(N)\}$ are given by $X^{\#} = \{X^*(1), X^*(N), \dots, X^*(2)\}$.

Define $\boldsymbol{P} \triangleq \boldsymbol{F}\boldsymbol{F}^T$, then $\boldsymbol{P}\boldsymbol{P}^T = \boldsymbol{I}_N$ and $\boldsymbol{P}\boldsymbol{F}_L^* = \boldsymbol{F}_L$, then from this definitions,

$$\begin{aligned} D(\boldsymbol{X}^{\#}) &= D(\boldsymbol{P}\boldsymbol{X}^*) = \boldsymbol{P}D^{\mathcal{H}}(\boldsymbol{X})\boldsymbol{P} \\ &= \mathrm{diag}\{X^*(1), X^*(N), \dots, X^*(2)\}. \end{aligned} \quad (9)$$

Suppose that there are some null subcarriers, let $\mathcal{K}_a$ and $N_a = |\mathcal{K}_a|$ be a set and the number of active subcarriers respectively. Then, given $\boldsymbol{X}_a$ as a transmitted OFDM block in frequency domain at the active subcarriers, using the definition in (9) the received signal can be expressed as

$$\boldsymbol{y} = \sqrt{N}\boldsymbol{F}_a^{\mathcal{H}}\left(D(\boldsymbol{X}_a)\boldsymbol{F}_{L,a}\boldsymbol{\xi} + \boldsymbol{P}_a D^{\mathcal{H}}(\boldsymbol{X}_a)\boldsymbol{F}_{L,a}\boldsymbol{\chi}\right) + \boldsymbol{v}, \quad (10)$$

where $\boldsymbol{F}_a$ is an $N_a \times N$ and $\boldsymbol{F}_{L,a}$ is an $N_a \times L$ sub-matrices of $\boldsymbol{F}$ and $\boldsymbol{F}_L$ respectively corresponding to N_a number of active subcarriers. Note that, $\boldsymbol{P}_a = \boldsymbol{F}_a\boldsymbol{F}_a^T$.

The frequency-domain representation of the received time signals in (10) is given by

$$\boldsymbol{Y} = D(\boldsymbol{X_a})\boldsymbol{F}_{L,a}\boldsymbol{\xi} + \boldsymbol{P}_a D^{\mathcal{H}}(\boldsymbol{X}_a)\boldsymbol{F}_{L,a}\boldsymbol{\chi} + \boldsymbol{w}, \quad (11)$$

where $\boldsymbol{w} = \boldsymbol{F}_a\boldsymbol{v}$.

For coherent detection, the direct channel $\boldsymbol{\xi}$ and the mirror channel $\boldsymbol{\chi}$ need to be estimated at the receiver. Since in practical system the statistics of the channel and the imperfections are unknown, we consider the least square channel estimators. The estimates of the direct and mirror channel impulse response vectors are given by

$$\hat{\boldsymbol{\xi}} = \underbrace{\left(\boldsymbol{F}_{L,a}^{\mathcal{H}} D^{\mathcal{H}}(\boldsymbol{X}_a) D(\boldsymbol{X}_a)\boldsymbol{F}_{L,a}\right)^{-1} \boldsymbol{F}_{L,a}^{\mathcal{H}} D^{\mathcal{H}}(\boldsymbol{X}_a)}_{\mathcal{M}} \boldsymbol{Y}, \quad (12)$$

$$\hat{\boldsymbol{\chi}} = \underbrace{\left(\boldsymbol{F}_{L,a}^{\mathcal{H}} D(\boldsymbol{X}_a) D^{\mathcal{H}}(\boldsymbol{X}_a)\boldsymbol{F}_{L,a}\right)^{-1} \boldsymbol{F}_{L,a}^{\mathcal{H}} D(\boldsymbol{X}_a)\boldsymbol{P}_a^{\mathcal{H}}}_{\mathcal{N}} \boldsymbol{Y}. \quad (13)$$

Substitution of (11) to (12) and (13) gives

$$\hat{\boldsymbol{\xi}} = \boldsymbol{\xi} + \mathcal{M}\boldsymbol{P}_a D^{\mathcal{H}}(\boldsymbol{X}_a)\boldsymbol{F}_{L,a}\boldsymbol{\chi} + \mathcal{M}\boldsymbol{w}, \quad (14)$$
$$\hat{\boldsymbol{\chi}} = \boldsymbol{\chi} + \mathcal{N} D(\boldsymbol{X}_a)\boldsymbol{F}_{L,a}\boldsymbol{\xi} + \mathcal{N}\boldsymbol{w}. \quad (15)$$

To measure the impact of the channel noise and IQ mismatch, the MSE caused by both effects is derived.

$$\begin{aligned} ||\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}||_2^2 = &\underbrace{\sigma_\chi^2 \,\mathrm{trace}\left(\mathcal{M}\boldsymbol{P}_a D^{\mathcal{H}}(\boldsymbol{X}_a)\boldsymbol{F}_{L,a}\boldsymbol{F}_{L,a}^{\mathcal{H}} D(\boldsymbol{X}_a)\boldsymbol{P}_a^{\mathcal{H}}\mathcal{M}^{\mathcal{H}}\right)}_{MSE_{\boldsymbol{\chi}}} \\ &+ \underbrace{\sigma_w^2 \,\mathrm{trace}\left(\mathcal{M}\mathcal{M}^{\mathcal{H}}\right)}_{MSE_w}. \end{aligned} \quad (16)$$

Let us define the matrix $\boldsymbol{\Lambda}_a = D(\boldsymbol{X}_a)D^{\mathcal{H}}(\boldsymbol{X}_a)$, then the

MSE caused by gaussian noise can be written as

$$MSE_w = \sigma_w^2 \,\text{trace}\left[\boldsymbol{F}_{L,a}\left(\boldsymbol{F}_{L,a}^{\mathcal{H}}\Lambda_a\boldsymbol{F}_{L,a}\right)^{-1}\boldsymbol{F}_{L,a}^{\mathcal{H}}\right]. \quad (17)$$

III. Training Symbol Design

We utilize some criteria to design training symbols that minimizes the channel estimate MSE while suppressing the interference replica caused by I/Q imbalances. From equation (14), it can be seen that, the second and third terms are interferences and their contribution to channel estimation errors is reduced if they are reduced to zero. The last term is due to noise and is considered to be white. Thus we need to design training symbols that suppress the interference caused by the second term. Substituting $\mathcal{M}$ in (14), the second term of (14) can be represented as

$$\left(\boldsymbol{F}_{L,a}^{\mathcal{H}}D^{\mathcal{H}}(\boldsymbol{X}_a)D(\boldsymbol{X}_a)\boldsymbol{F}_{L,a}\right)^{-1}\boldsymbol{F}_{L,a}^{\mathcal{H}}D^{\mathcal{H}}(\boldsymbol{X}_a)\boldsymbol{P}_aD^{\mathcal{H}}(\boldsymbol{X}_a)\boldsymbol{F}_{L,a}\boldsymbol{\chi}$$

To suppress the interference replica, the training symbols should be designed in such away that

$$\boldsymbol{F}_{L,a}^{\mathcal{H}}D^{\mathcal{H}}(\boldsymbol{X}_a)\boldsymbol{P}_aD^{\mathcal{H}}(\boldsymbol{X}_a)\boldsymbol{F}_{L,a} = 0. \quad (18)$$

The easiest way of meeting this condition is to set the active subcarriers in the lower or upper band of the central DC subcarrier to zero and allocate power to the subcarriers on one sideband which is not nulled as in [1]. However, this will lead to poor estimate of the channel since channel is only estimated by either upper or lower subcarriers. To ensure better MSE performance, both training allocation and power distribution need to be careful considered. In this paper we propose an algorithm for selecting training position that gives better channel estimate MSE performance.

Let us represent the training symbol with phase information as $\boldsymbol{X}_a(\boldsymbol{\phi}_a) = \left[X_1e^{j\phi_1}, X_2e^{j\phi_2}, \ldots, X_{N_a}e^{j\phi_{N_a}}\right]$, where $\boldsymbol{X}_a$ and $\boldsymbol{\phi}$ are the amplitudes and phase information of the training symbol. To obtain the approximate of the continuos time signal, we oversample the training signal at a sampling rate $\mathcal{L}$. The signal obtained by $\mathcal{L}N$-points IFFT can be expressed as

$$\boldsymbol{x}(\boldsymbol{\phi}) = \boldsymbol{\Gamma}\boldsymbol{X}(\boldsymbol{\phi}), \quad (19)$$

where $\boldsymbol{\Gamma}$ is an $\mathcal{L}N \times N$ DFT matrix with

$$\Gamma_{t,k} = \frac{1}{\sqrt{\mathcal{L}N}}e^{j\frac{2\pi kt}{\mathcal{L}N}}, \quad t \in [0, \mathcal{L}N-1], \quad k \in [0, N-1]. \quad (20)$$

For a given energy to be utilized for channel estimation, we normalize the sum of training symbol power such that

$$\sum_{k\in\mathcal{K}_a}\Lambda_k = \sum_{k=1}^{N_a}\lambda_k = 1. \quad (21)$$

To minimize the channel estimate MSE, all we need is to determine the optimal $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_{N_a}]^T$, since the channel estimate MSE does not depend on the phase of the training symbols. However, to minimize PAPR requires careful selection of phase information of the trainings. The optimization problem under these constraints can be stated as

$$\begin{aligned}
\underset{\boldsymbol{\lambda}}{\text{minimize}} \quad & \text{trace}\left[\boldsymbol{F}_{L,a}^{\mathcal{H}}\left(\frac{1}{\sigma_w^2}\boldsymbol{F}_{L,a}^{\mathcal{H}}\Lambda_a\boldsymbol{F}_{L,a}\right)^{-1}\boldsymbol{F}_{L,a}\right] \\
\text{subject to} \quad & [1,\ldots,1]\,\boldsymbol{\lambda} \le 1, \quad \boldsymbol{\lambda} \succeq 0, \\
& \boldsymbol{F}_{L,a}^{\mathcal{H}}D^{\mathcal{H}}(\boldsymbol{X}_a)\boldsymbol{P}_aD^{\mathcal{H}}(\boldsymbol{X}_a)\boldsymbol{F}_{L,a} = 0, \\
& \min_{\boldsymbol{\phi}\in 0,\pi} |\boldsymbol{x}(\boldsymbol{\phi})|_\infty \le \Upsilon.
\end{aligned} \quad (22)$$

The problem in (22) above is a non-convex optimization problem and can not be solved easily. However we can split it in two parts. First, power distribution to minimize the channel estimate mean squared error (MSE) under white gaussian noise, then design of phase information with a potential of reducing the PAPR.

We adopt the method presented in [7], to optimally allocate power to the training symbols in order to minimize the MSE caused by the gaussian noises.

To design the special training sequences, we need to set some subcarriers to zero as proposed in [1]. Thus, once power distribution to all training symbol is obtained, we utilize cross entropy method to select training symbol that minimizes the channel estimate MSE and eliminate the interference replica.

A. Training symbol selection with cross entropy optimization

The most straightforward approach to obtain the optimal position of the training symbols, is by exhaustive search. However, because of its high computational complexity, it becomes prohibitive for systems with large number of training symbols. Inspired by the efficiency of the cross-entropy (CE) method for finding near-optimal solutions in huge search spaces, this paper proposes the application of the CE method to search for the optimal position of the training symbol that minimizes the channel estimate MSE.

In [1], it has been shown that, training symbol that utilizes some null subcarriers can effectively estimate the I/Q imbalance. However there is no proposed algorithm for selecting optimal training symbols.

Assume $\mathcal{K}_p$ to be a selected set from the $\mathcal{K}_a$. Our optimal training sequence design as a combinatorial optimization problem can be formulated as

$$\mathcal{K}_p^\star = \arg\min_{\mathcal{K}_p^m\in\boldsymbol{\Omega}}\mathcal{C}_{sel}(\mathcal{K}_p^m), \quad (23)$$

where

$$\mathcal{C}_{sel}(\mathcal{K}_p^m) = \text{trace}\left[\boldsymbol{F}_{L,p}\left(\boldsymbol{F}_{L,p}^{\mathcal{H}}\Lambda_p\boldsymbol{F}_{L,p}\right)^{-1}\boldsymbol{F}_{L,p}^{\mathcal{H}}\right], \quad (24)$$

represents the channel estimate MSE of the training set $\mathcal{K}_p^m$ and $\mathcal{K}_p^\star$ is the global optimal set of the objective function. The set $\mathcal{K}_p^m$ is defined by

$$\mathcal{K}_p^m = \mathcal{K}_a(\{I_k\}_{k=1}^{|\mathcal{K}_a|} = 1), \quad I_k \in \{0,1\}, \quad m = 1,\ldots,M, \quad (25)$$

where the indicator function I_k shows whether a subcarrier at the kth position is selected. The set of all $M = \binom{|\mathcal{K}_a|}{|\mathcal{K}_p|}$ possible subsets is denoted by $\Omega = \{\mathcal{K}_p^1, \dots \mathcal{K}_p^M\}$, where $\binom{a}{b}$ denotes the possible combination set.

Applying the CE to solve (23), the first step is transforming the deterministic optimization problem (23) into a family of stochastic sampling problems. Since the considered problem is on a discrete case, a family of Bernoulli probability density functions associated with the training symbol selection vector, $\boldsymbol{\omega} = \left[\omega_1, \omega_2, \dots \omega_{|\mathcal{K}_a|}\right]$, $\omega_k \in \{0, 1\}$, is given by

$$f(\boldsymbol{\omega}, \boldsymbol{p}) = \prod_{k=1}^{|\mathcal{K}_a|} p_k^{1_k(\boldsymbol{\omega})}(1 - p_k^{1-1_k(\boldsymbol{\omega})}), \tag{26}$$

where $\boldsymbol{p} = [p_1, p_2, \dots p_{|\mathcal{K}_a|}]$ is a probability vector whose p_k entry indicates the probability of selecting the kth subcarrier, and the indicator function $1_k(\boldsymbol{\omega}) \in \{0, 1\}$ indicates whether the kth element of ω_k(the kth tone) is selected. If the ω_k is selected, then $1_k(\omega_k) = 1$. Each element is modeled as an independent Bernoulli random variable with probability mass function $p(\omega_k = 1) = p_k$, and $p(\omega_k = 0) = 1 - p_k$, for $k = 1, \dots |\mathcal{K}_a|$.

The CE method aims to find an optimal distribution $\boldsymbol{p}^\star$ that generates an optimal solution $\boldsymbol{\omega}^\star$ with minimum channel estimate MSE. However, $\boldsymbol{\omega}^\star$ occurs with a very small probability. In this case, (23), is associated with the problem of estimating the probability $Pr[\mathcal{C}_{sel}(\boldsymbol{\omega}) \leq \gamma]$ for a given threshold γ. To estimate the rare event, CE iteratively updates the probability vector $\boldsymbol{p}$ so that most samples generated by $f(\boldsymbol{\omega}; \boldsymbol{p})$ satisfy $\mathcal{C}_{sel} \leq \gamma$. By iteratively improving γ , $f(\boldsymbol{\omega}; \boldsymbol{p})$ eventually converges to an optimum probability density function $f(\boldsymbol{\omega}; \boldsymbol{p}^\star)$ and optimal $\boldsymbol{\omega}^\star$ can be obtained from $\boldsymbol{p}^\star$ by $f(\boldsymbol{\omega}; \boldsymbol{p}^\star)$.

A standard CE procedure for solving combinatorial problems contains two stages [8].

1) Adaptive updating of γ^t: For a given p^{t-1} generate $\mathcal{U}$ random samples $\{\boldsymbol{\omega}^{(t,u)}\}_{u=1}^{\mathcal{U}}$ from $f(.; \boldsymbol{p}^{t-1})$, where t denotes the iteration index of CE. Then, calculate the channel estimate MSE according to (23) to obtain a set of performance values $\{\mathcal{C}_{sel}(\boldsymbol{\omega}^{(t,u)})\}_{u=1}^{\mathcal{U}}$ and rank them in ascending order so that $\mathcal{C}_{sel}^1 \leq \dots \leq \mathcal{C}_{sel}^{\mathcal{U}}$. Finally, assign
$$\gamma^t = \mathcal{C}_{sel}^{\rho\mathcal{U}}, \tag{27}$$
where ρ denotes the fraction of the best samples and $\lceil \cdot \rceil$ is the ceiling operation.
2) Adaptive updating of $\boldsymbol{p}^t$: For a given $\gamma^{(t)}$ and $\boldsymbol{p}^{(t-1)}$, use the same samples $\{\mathcal{C}_{sel}(\boldsymbol{\omega}^{(t,u)})\}_{u=1}^{\mathcal{U}}$ to update the parameter $\boldsymbol{p}^{(t)} = \left[p_0^{(t)}, p_1^{(t)}, \dots, p_{|\mathcal{K}_a|}^{(t)}\right]$ via
$$p_k^{(t)} = \frac{\sum_{u=1}^{\mathcal{U}} 1_{\{\mathcal{C}_{sel}(\boldsymbol{\omega}^{(t,u)}) \leq \gamma^{(t)}\}} 1_k(\boldsymbol{\omega}^{(t,u)})}{\sum_{u=1}^{\mathcal{U}} \boldsymbol{I}_{\{\mathcal{C}_{sel}(\boldsymbol{\omega}^{(t,u)}) \leq \gamma^{(t)}\}}}, \tag{28}$$
where $\{\mathcal{C}_{sel}(\boldsymbol{\omega}^{(t,u)}) \leq \gamma^{(t)}\}$ is a variable defined by
$$1_{\{\mathcal{C}_{sel}(\boldsymbol{\omega}^{(t,u)}) \leq \gamma^{(t)}\}} = \begin{cases} 1, & \text{if } \mathcal{C}_{sel}(\boldsymbol{\omega}^{(t,u)}) \leq \gamma^{(t)}, \\ 0, & \text{otherwise.} \end{cases} \tag{29}$$

Note that in order to prevent fast convergence to a local optimum, parameter $\boldsymbol{p}^{t-1}$ is not updated to $\boldsymbol{p}^t$ directly; a smoothing factor, α , $0 \leq \alpha \leq 1$ was suggested by [8], [9] to update (29) into

$$\boldsymbol{p}^t = \alpha \times p^t + (1 - \alpha) \times \boldsymbol{p}^{t-1}. \tag{30}$$

When , $\alpha = 1$ the original updating formulation is achieved.

It should be noted that $\boldsymbol{\omega}$ is a binary selection vector subject to constraint (i.e., the restricted location of the training subcarrier and its corresponding mirror subcarrier), there is no guarantee that the samples generated from (26) satisfy the condition in (18). Therefore an additional operation to rearrange the necessary 1s to meet the constraint (18) is used.

The following algorithm summarizes our proposed design. $\mathcal{J}$ is the predefined total number of iteration and $\mathcal{Z}$ is number of iterations without improvement.

Algorithm 1: Training symbol selection

1. Optimize power of the active subcarriers, using convex optimization.
2. Initialize probability vector $\boldsymbol{p}^{(0)} = p_0^{(0)}, p_1^{(0)}, \dots, p_{|\mathcal{K}_a|}^{(0)}$ with $p_k^{(0)} = 0.5$.
3. Set the iteration counter $t = 0$ and $t' = 0$
4. **while** ($t' < \mathcal{Z}$ and $t < \mathcal{J}$) **do**
5. Generate $\mathcal{U}$ random samples $\{\boldsymbol{\omega}^{(t,u)}\}_{u=1}^{\mathcal{U}}$ from the density function $f(\boldsymbol{\omega}; \boldsymbol{p}^{(t-1)})$, under the constraints in (18).
6. Compute the performance functions $\mathcal{C}_{sel}(\boldsymbol{\omega}^{(t,u)})$ for $u = 1, 2, \dots, \mathcal{U}$, and rank them in ascending order so that $\mathcal{C}_{sel}^1 \leq \mathcal{C}_{sel}^2 \leq \dots \leq \mathcal{C}_{sel}^{\mathcal{U}}$ and use (27) to find γ^t.
7. Calculate the parameter $\boldsymbol{p}^{(t)}$ using (28).
8. Update $\boldsymbol{p}^{(t)} = \alpha \times \boldsymbol{p}^t + (1 - \alpha) \times \boldsymbol{p}^{(t-1)}$.
9. **if** $\mathcal{C}_{sel}^*) \leq \mathcal{C}_{sel}^1$ **then**
10. $t' \leftarrow t' + 1$
11. **else**
12. $\mathcal{C}_{sel}^* = \mathcal{C}_{sel}^1$ and $t' = 0$
13. **end if**
14. Increment $t \leftarrow t + 1$
15. **end while**

Note that, to design phase information to the symbols, we utilize CE technique as in selection of training positions.

IV. Design Examples

We conduct computer simulations to demonstrate the effectiveness of our proposed schemes. The efficacy of the proposed design is evaluated by the mean squared error (MSE) as well as bit error rate (BER) performances. An OFDM transmission frame with $N = 64$ and $N_a = 52$ is considered. The remaining 12 subcarriers, 6 are null in the lower frequency guard band while 5 are nulled in the upper frequency guard band and one is the central DC null subcarrier. To estimate and compensate for the I/Q imbalances, we adopt the compensator proposed in [1] that utilizes special training pattern.

Fig. 1 depicts the proposed training symbols and the equal powered training symbols allocated on one side about the

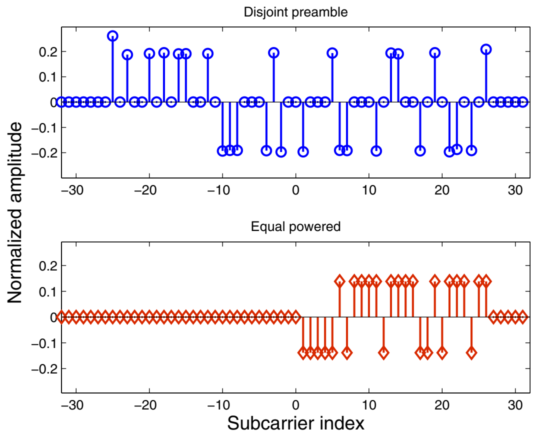

Fig. 1. Comparison between our proposed training symbols and the conventional equal power trainings; for the two designs, phase information is obtained by using cross entropy optimization.

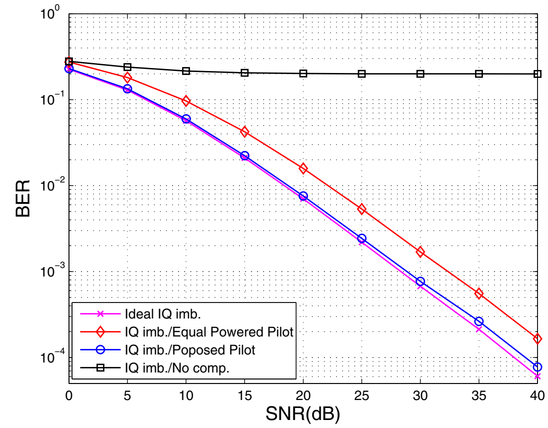

Fig. 3. BER versus SNR results for 64QAM constellation, phase imbalance of 2°, amplitude imbalance of 1dB and $L = 4$

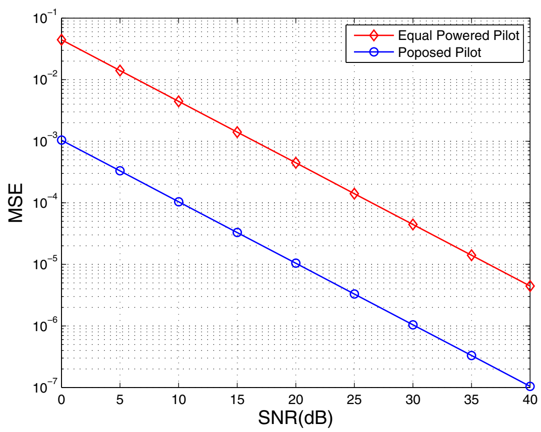

Fig. 2. Comparison of the Channel estimate MSE between the proposed training symbols and the equal powered trainings

central DC subcarrier. The proposed design distribute the subcarriers to minimize the channel MSE. In the following, the efficacy of the two designs is compared by the MSE and BER performances.

Fig. 2 shows the MSE performance of the two designs for different signal to noise ratio (SNR). From the plot it is clear that our proposed design outperforms equal powered training allocated on the upper sideband due to the fact that, for the equal powered trainings channel is only estimated by upper subcarriers. This suggests that to ensure better MSE performance, both training allocation and power distribution need to be careful considered.

Next, we demonstrate the performance of our training symbols by considering the BER performance . It is well known that, to obtain better BER performance, proper compensation of I/Q imbalance and accurate channel estimates are of primary important. Fig. 3 depicts the BER performance of the training symbols in Fig. 1 together with the results of the known channel state information without any IQ imbalance. From the results it is clear that the BER performance of our proposed design is comparable to that of the known channel state information. This demonstrates the efficiency of our designed training symbols. Although total power allocated to the training symbols is the same for the two designs, our proposed design outperforms the equal powered training symbols. This is because the proposed design optimally allocate the training symbols to reduce the channel estimate MSE. To obtain better BER as well as MSE performances require optimal allocation of training symbols.

References

[1] A. Tarighat, R. Bagheri, and A. H. Sayed, "Compensation schemes and performance analysis of IQ imbalances in OFDM receivers," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 3257–3268, 2005.

[2] H. Minn and D. Munoz, "Pilot designs for channel estimation of OFDM systems with frequency-dependent I/Q imbalances," in *Proc. IEEE Wireless Communications and Networking Conf. WCNC 2009*, 2009, pp. 1–6.

[3] H. Lin, X. Zhu, and K. Yamashita, "Low-complexity pilot-aided compensation for carrier frequency offset and i/q imbalance," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 448–452, 2010.

[4] B. Narasimhan, S. Narayanan, M. Hlaing, and N. Al-Dhahir, "Reduced-complexity baseband compensation of joint Tx/Rx I/Q imbalance in mobile MIMO-OFDM," *IEEE Journal on Wireless Comms.*, vol. 9, no. 5, pp. 1720–1728, 2010.

[5] E. Lopez-Estraviz, S. De Rore, F. Horlin, and A. Bourdoux, "Pilot design for Joint Channel and Frequency-Dependent Transmit/Receive IQ Imbalance Estimation and Compensation in OFDM-Based Transceivers," in *Proc. IEEE Int. Conf. Communications ICC '07*, 2007, pp. 4861–4866.

[6] Z. Yaning, V. Mikko, and R. Markku, "Analysis and compensation of transmitter and receiver I/Q imbalances in space-time coded multiantenna ofdm systems," *EURASIP Journal onWireless Communications and Networking*, p. 16, 2008.

[7] S. Ohno, E. Manasseh, and M. Nakamoto, "Preamble and pilot symbol design for channel estimation in OFDM systems with null subcarriers," *EURASIP Journal on Wireless Communications and Networking*, vol. 2011, pp. 1–17, June 2011.

[8] R. Rubinstein and D. Kroese, *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning*. Springer-Verlag, 2004.

[9] W. Maliki, Y. Zhang, D. O'Brien, and D. Edwards, "Cross entropy optimization of MIMO capacity by transmit antenna selection," *IET Microwave Antennas Propagation*, vol. 1, no. 6, pp. 1131–1136, Dec. 2007.