[31] M. B. Shenouda and T. N. Davidson, "Tomlinson-Harashima precoding for broadcast channels with uncertainty," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1380–1389, Sep. 2007.
[32] A. Pascual-Iserte, D. P. Palomar, A. I. Perez-Neira, and M. A. Lagunas, "A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 346–360, May 2006.
[33] M. D. Nisar, "Minimax robustness in signal processing for communications," Ph.D. dissertation, TUM, Munich, Germany, Aug. 2001.
[34] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Belmont, MA, USA: Athena Scientific, 2003.

Pilot Design for Sparse Channel Estimation in OFDM-Based Cognitive Radio Systems

Chenhao Qi, *Member, IEEE*, Guosen Yue, *Senior Member, IEEE*, Lenan Wu, and A. Nallanathan, *Senior Member, IEEE*

***Abstract*—In this correspondence, sparse channel estimation is first introduced in orthogonal frequency-division multiplexing (OFDM)-based cognitive radio systems. Based on the results of spectrum sensing, the pilot design is studied by minimizing the coherence of the dictionary matrix used for sparse recovery. Then, it is formulated as an optimal column selection problem where a table is generated and the indexes of the selected columns of the table form a pilot pattern. A novel scheme using constrained cross-entropy optimization is proposed to obtain an optimized pilot pattern, where it is modeled as an independent Bernoulli random process. The updating rule for the probability of each active subcarrier selected as a pilot subcarrier is derived. A projection method is proposed so that the number of pilots during the optimization is fixed. Simulation results verify the effectiveness of the proposed scheme and show that it can achieve 11.5% improvement in spectrum efficiency with the same channel estimation performance compared with the least squares (LS) channel estimation.**

***Index Terms*—Cognitive radio (CR), compressed sensing (CS), orthogonal frequency-division multiplexing (OFDM), pilot design, sparse channel estimation.**

I. Introduction

Traditionally, every wireless system is required to have an exclusive spectrum license to avoid interference from other systems or users. However, recent studies have shown that a large portion of the licensed spectrum is underutilized. This has motivated studies on cognitive radio (CR), which allows secondary users (SUs) to utilize the licensed spectrum without interfering with licensed users or primary users (PUs) and also improves spectrum utilization without allocating a new spectrum resource [1], [2]. Orthogonal frequency-division multiplexing (OFDM), which has been considered one of the best candidates for the physical layer of CR systems, can efficiently avoid interference by dynamically nulling corresponding subcarriers. Hence, the subcarriers may be noncontiguous in OFDM-based CR systems, and the efficient selection of pilot tones is crucial to the performance of pilot-assisted channel estimation. In [3], the pilot design is formulated as an optimization problem minimizing an upper bound related to the mean square error (MSE), where the pilot indexes are obtained by solving a series of 1-D low-complexity subproblems. In [4], a pilot design scheme using convex optimization together with the cross-entropy optimization is proposed to minimize the MSE. In [5], parameter adaptation for wireless multicarrier-based CR systems is investigated, where the cross-entropy method is demonstrated to outperform the genetic algorithm and particle swarm optimization. However, all of them are based on the least squares (LS) channel estimation.

Manuscript received March 25, 2013; revised June 18, 2013; accepted August 6, 2013. Date of publication September 4, 2013; date of current version February 12, 2014. This work was supported in part by the National Natural Science Foundation of China under Grant 61302097, by the Ph.D. Programs Foundation of the Ministry of Education of China under Grant 20120092120014, and by the Huawei Innovative Research Plan. The review of this paper was coordinated by Prof. X. Wang.

C. Qi and L. Wu are with the School of Information Science and Engineering, Southeast University, Nanjing 210096, China (e-mail: qch@seu.edu.cn; wuln@seu.edu.cn).
G. Yue is with NEC Laboratories America, Inc., Princeton, NJ 08540 USA (e-mail: yueg@nec-labs.com).
A. Nallanathan is with the Center for Telecommunications, King's College London, London WC2R 2LS, U.K. (e-mail: nallanathan@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.
Digital Object Identifier 10.1109/TVT.2013.2280655

Recently, applications of compressed sensing (CS) to channel estimation, i.e., sparse channel estimation, have shown that improved channel estimation performance and reduced pilot overhead can be achieved by exploring the sparse nature of wireless multipath channels. The sparse channel estimation for OFDM systems has been intensively studied [6], [7], and many CS algorithms, including orthogonal matching pursuit (OMP), compressive sampling matching pursuit, and basis pursuit, have been applied. Therefore, it is natural to extend this technique to OFDM-based CR systems, which can further improve the data rate and flexibility of SUs. However, it also brings new challenges to the pilot design. To the authors' best knowledge, so far, there has been no study focused on the pilot design for sparse channel estimation in OFDM-based CR systems. Although we can continue to use the same pilot design schemes as LS, e.g., predesigning pilot tones and deactivating those tones occupied by PUs and using the nearest available subcarriers instead, apparently, it is not optimal since it does not benefit from the sparse channel estimation.

In this correspondence, we first introduce sparse channel estimation in OFDM-based CR systems. After spectrum sensing, we explore the pilot design by minimizing the coherence of the dictionary matrix used for sparse recovery. We then formulate it as an optimal column selection problem where a table is generated and the indexes of the selected columns of the table form a pilot pattern. A novel scheme using constrained cross-entropy optimization is proposed to obtain an optimized pilot pattern, where we model it as an independent Bernoulli random process. The updating rule for the probability of each active subcarrier being selected as a pilot subcarrier is derived. Moreover, a projection method is proposed so that the number of pilot subcarriers during optimization is fixed.

The remainder of this correspondence is organized as follows. Section II formulates the pilot-assisted channel estimation in OFDM-based CR systems as a sparse recovery problem. Section III proposes a pilot design scheme using constrained cross-entropy optimization. Simulation results are provided in Section IV, and finally, Section V concludes this correspondence.

The notations used in this paper are defined as follows. Symbols for matrices (uppercase) and vectors (lowercase) are in boldface. $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$, $\mathrm{diag}\{\cdot\}$, $\boldsymbol{I}_L$, $\mathcal{CN}$, $|\cdot|$, and $\lceil\cdot\rceil$ denote the matrix transpose, the conjugate transpose (Hermitian), the diagonal matrix, the identity matrix of size L, the complex Gaussian distribution, the absolute value, and the ceiling function, respectively.

II. Problem Formulation

The OFDM-based CR system under consideration is shown in Fig. 1, where we employ sparse channel estimation instead of the

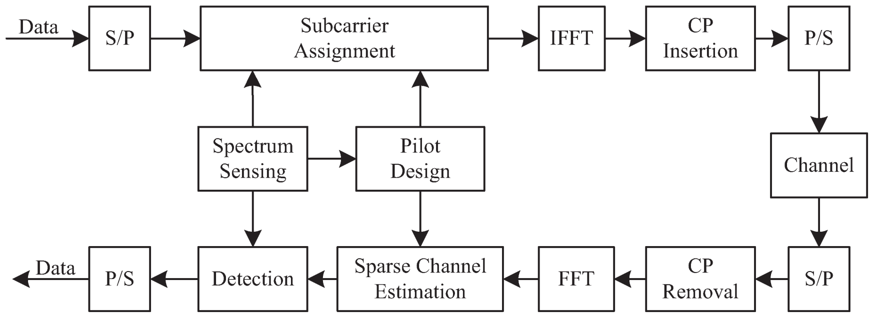

Fig. 1. OFDM-based CR system using sparse channel estimation.

traditional channel estimation methods. Correspondingly, we have to reconsider the pilot design since the sparse channel estimation using CS techniques is essentially different with traditional methods. We perform the pilot design based on the results of spectrum sensing. Note that here, we assume ideal spectrum sensing without false alarm or missing detection. After spectrum sensing, the OFDM subcarriers occupied by PUs are first deactivated. From the remaining active subcarriers, we select some to transmit pilot symbols and the others to transmit data symbols for SUs. The resulting sequence is transformed into the time-domain signal via inverse fast Fourier transform (IFFT). To avoid intersymbol interference, we add a cyclic prefix (CP) whose length is usually larger than the maximum channel delay spread L. The wireless multipath channel is modeled as a finite-impulse response filter with the channel impulse response (CIR) to be $\boldsymbol{h} = [h(1), h(2), \ldots, h(L)]^{\mathrm{T}}$. At the receiver, after CP removal and fast Fourier transform (FFT), we perform the sparse channel estimation using the pilot symbols. Since OFDM transforms the frequency-selective multipath fading channel into parallel flat-fading channels, we may use the single-tap zero-forcing channel equalization with very low complexity for data detection. We assume the transmitter broadcasts the results of pilot design to receivers through control signaling.

After deactivating those subcarriers occupied by PUs, we suppose there are M remaining OFDM subcarriers, denoted by $\mathcal{C} = \{c_1, c_2, \ldots, c_M\}$, which is a noncontigous integer set representing the indexes of active subcarriers. Without loss of generality, we suppose $1 \leq c_1 < c_2 < \cdots < c_M$. Now, we select $K (K \leq M)$ pilot subcarriers indicated by $c_{p_1}, c_{p_2}, \ldots, c_{p_K} (1 \leq p_1 < p_2 < \cdots < p_K \leq M)$ from M subcarriers for frequency-domain pilot-assisted channel estimation. The transmit and receive pilot symbols are denoted by $x(c_{p_1}), x(c_{p_2}), \ldots, x(c_{p_K})$ and $y(c_{p_1}), y(c_{p_2}), \ldots, y(c_{p_K})$, respectively. For traditional least squares (LS) channel estimation, we first acquire channel frequency response (CFR) at pilot positions as $\{y(c_{p_i})/x(c_{p_i}),\ i = 1, 2, \ldots, K\}$, and then make interpolations for the rest of the subcarriers. However, it usually requires a large number of pilots, i.e., $K > L$, so that the interpolations can approximate the true value of CFR.

The relation between the transmit pilots and the receive pilots can be written as

$$\begin{bmatrix} y(c_{p_1}) \\ y(c_{p_2}) \\ \vdots \\ y(c_{p_K}) \end{bmatrix} = \begin{bmatrix} x(c_{p_1}) & 0 & 0 & 0 \\ 0 & x(c_{p_2}) & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & x(c_{p_K}) \end{bmatrix} \cdot \boldsymbol{F}_{K \times L} \cdot \begin{bmatrix} h(1) \\ h(2) \\ \vdots \\ h(L) \end{bmatrix} + \begin{bmatrix} \eta(1) \\ \eta(2) \\ \vdots \\ \eta(K) \end{bmatrix} \quad (1)$$

where $\eta(i) \sim \mathcal{CN}(0, \sigma^2),\ i = 1, 2, \ldots, K$ is the independently and identically distributed (i.i.d.) additive white Gaussian noise, and $\boldsymbol{F}_{K \times L}$ is a discrete Fourier transform submatrix, given by

$$\boldsymbol{F}_{K \times L} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & \omega^{c_{p_1}} & \cdots & \omega^{c_{p_1} \cdot (L-1)} \\ 1 & \omega^{c_{p_2}} & \cdots & \omega^{c_{p_2} \cdot (L-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{c_{p_K}} & \cdots & \omega^{c_{p_K} \cdot (L-1)} \end{bmatrix}$$

where $N (N \geq M)$ is the number of points for IFFT and FFT in Fig. 1, and $\omega = e^{-j 2\pi / N}$. We denote

$$\begin{aligned} \boldsymbol{X} &= \mathrm{diag}\,\{x(c_{p_1}), x(c_{p_2}), \ldots, x(c_{p_K})\} \\ \boldsymbol{y} &= [y(c_{p_1}), y(c_{p_2}), \ldots, y(c_{p_K})]^{\mathrm{T}} \\ \boldsymbol{\eta} &= [\eta(1), \eta(2), \ldots, \eta(K)]^{\mathrm{T}} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \boldsymbol{I}_K). \end{aligned}$$

Moreover, we let $\boldsymbol{A} = \boldsymbol{X} \boldsymbol{F}_{K \times L}$. Then, (1) can be rewritten as

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{h} + \boldsymbol{\eta}. \quad (2)$$

If $L \leq K \leq M$ and $\boldsymbol{A}$ has full-column rank, (2) can be solved by the LS method, which essentially employs the FFT interpolations with the estimated CIR given by

$$\hat{\boldsymbol{h}}_{\mathrm{LS}} = (\boldsymbol{A}^{\mathrm{H}} \boldsymbol{A})^{-1} \boldsymbol{A}^{\mathrm{H}} \boldsymbol{y}. \quad (3)$$

However, if we can further reduce the pilot overhead, i.e., $K < L$, improved spectrum efficiency and flexibility can be achieved for SUs. In this case, we have to explore the sparse property of the wireless multipath channel and employ sparse recovery instead of LS. In practice, since the sampling period is usually much smaller than the channel delay spread [8], particularly for OFDM systems with oversampling, most components of $\boldsymbol{h}$ are either zero or nearly zero, implying that $\boldsymbol{h}$ is sparse. With this *a priori* condition, we can apply CS algorithms to estimate $\boldsymbol{h}$. Many works have already demonstrated that these algorithms outperform LS for channel estimation [6].

The theory of restrict isometry property (RIP) shows that $\boldsymbol{h}$ in (2) can be recovered from measurement $\boldsymbol{y}$ with a high probability when dictionary matrix $\boldsymbol{A}$ satisfies the RIP. However, it is difficult to check whether a given matrix satisfies the RIP. Alternatively, we can minimize the coherence of $\boldsymbol{A}$, which is known as the mutual incoherence property (MIP) and is more practical than the RIP. In fact, the MIP condition is stronger than the RIP. The MIP implies RIP, but the inverse is not true [9].

Given a pilot pattern, i.e.,

$$\boldsymbol{p} = \{c_{p_1}, c_{p_2}, \ldots, c_{p_K}\} \quad (4)$$

we define the coherence of $\boldsymbol{A}$ as the maximum absolute correlation between any two different columns of $\boldsymbol{A}$, i.e.,

$$\begin{aligned} g(\boldsymbol{p}) &\triangleq \max_{0 \leq m < n \leq L-1} |\langle A(m), A(n) \rangle| \\ &= \max_{0 \leq m < n \leq L-1} \left| \sum_{i=1}^{K} |x(c_{p_i})|^2 \, \omega^{c_{p_i}(n-m)} \right| \end{aligned} \quad (5)$$

where $\langle A(m), A(n)\rangle$ denotes the inner product of $A(m)$ and $A(n)$, i.e., $\langle A(m), A(n)\rangle = A^{\mathrm{H}}(m)\cdot A(n)$. One straightforward approach to obtain $g(\boldsymbol{p})$ is to compute the ℓ_∞ norm of $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$.

The objective function for the pilot design is to minimize the coherence of $\boldsymbol{A}$, i.e.,

$$Q = \min_{\boldsymbol{p}} g(\boldsymbol{p}). \tag{6}$$

The optimal pilot pattern is then given by

$$\boldsymbol{p}_{\mathrm{opt}} = \arg\min_{\boldsymbol{p}} g(\boldsymbol{p}). \tag{7}$$

We assume equal transmit power among all OFDM pilot symbols, i.e., $|x(c_{p_1})|^2 = |x(c_{p_2})|^2 = \cdots = |x(c_{p_K})|^2 = E$, which can be implemented with low complexity. Let $d = n-m$ and $\Lambda = \{1, 2, \ldots, L-1\}$. Then, (5) can be rewritten as

$$g(\boldsymbol{p}) = E\cdot\max_{d\in\Lambda}\left|\sum_{i=1}^{K}\omega^{c_{p_i}d}\right|. \tag{8}$$

For simplicity, we assume $E = 1$. We can generate a table, i.e.,

$$\boldsymbol{G} = \begin{bmatrix} \omega^{c_1} & \omega^{c_2} & \cdots & \omega^{c_M} \\ \omega^{2c_1} & \omega^{2c_2} & \cdots & \omega^{2c_M} \\ \vdots & \vdots & \ddots & \vdots \\ \omega^{(L-1)c_1} & \omega^{(L-1)c_2} & \cdots & \omega^{(L-1)c_M} \end{bmatrix} \tag{9}$$

with the dimension being $(L-1)\times M$. Once $\boldsymbol{p}$ is given, we look up $\boldsymbol{G}$ and find the corresponding K columns, composing an $(L-1)\times K$ submatrix $\tilde{\boldsymbol{G}}$. We make summations for each row of $\tilde{\boldsymbol{G}}$, resulting in an $(L-1)$-dimensional vector, where we obtain its ℓ_∞ norm as the final value of $g(\boldsymbol{p})$. Compared with the approach that always computes the ℓ_∞ norm of $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ involving matrix multiplications, it can be much faster implemented with $\boldsymbol{G}$ generated beforehand.

III. Pilot Design Scheme

Although we may exhaustively search all possible pilot patterns for the best pilot pattern with the minimum objective, it is computationally very inefficient to implement for SUs equipped with power-constrained mobile devices. For example, if $M = 200$ and $K = 12$, we have a huge search space with $\binom{200}{12} = 6.1\times 10^{18}$ candidates. For this combinatorial optimization problem, we now propose a scheme using stochastic optimization algorithms that are iterative in nature and converge to the global optimum in probability [10]. The scheme is based on the use of constrained cross-entropy optimization [11] and is discussed as follows.

We first define a binary sequence, i.e.,

$$\boldsymbol{z} = \{z(i), \qquad i = 1, 2, \ldots, M\}, \qquad z(i)\in\{0, 1\}. \tag{10}$$

If c_i is selected as a pilot subcarrier, $z(i) = 1$; otherwise, $z(i) = 0$, $i = 1, 2, \ldots, M$. The restriction is

$$\sum_{i=1}^{M} z(i) = K \tag{11}$$

as the number of pilots is explicitly given as K. Hence, $\boldsymbol{z}$ and $\boldsymbol{p}$ are equivalent. The optimization for $\boldsymbol{p}$ is therefore converted to the optimization for $\boldsymbol{z}$. We model each entry of $\boldsymbol{z}$ as an independent Bernoulli random variable with the probability to be

$$\Pr\left(z(i)\right) = \begin{cases} q_i, & z(i) = 1 \\ 1 - q_i, & z(i) = 0 \end{cases} \tag{12}$$

for $i = 1, 2, \ldots, M$, where q_i is the probability of c_i being selected as a pilot subcarrier. We denote $\boldsymbol{q} = \{q_i,\ i = 1, 2, \ldots, M\}$. Then, the probability density function (pdf) associated with the pilot design problem is

$$f(\boldsymbol{z}, \boldsymbol{q}) = \prod_{i=1}^{M} q_i^{z(i)}(1-q_i)^{1-z(i)}. \tag{13}$$

Since from each of $\boldsymbol{p}$ and $\boldsymbol{z}$ we can decide the other, $\boldsymbol{p}$ and $\boldsymbol{z}$ are equivalent. We define $s(\boldsymbol{z}) \triangleq -g(\boldsymbol{p})$. Correspondingly, (6) is converted into a maximization problem that is typical of cross-entropy optimization, i.e.,

$$Q = -\max_{\boldsymbol{z}} s(\boldsymbol{z}). \tag{14}$$

Meanwhile, (7) is converted into

$$\boldsymbol{z}_{\mathrm{opt}} = \arg\max_{\boldsymbol{z}} s(\boldsymbol{z}) \tag{15}$$

where $\boldsymbol{z}_{\mathrm{opt}}$ is equivalent to $\boldsymbol{p}_{\mathrm{opt}}$. Then, we define an indicator function $I(e)$, which returns 1 if event e is true and returns 0 otherwise. For some threshold r, the probability of $s(\boldsymbol{z})$ being greater than r is

$$\psi(r) \triangleq \Pr\left(s(\boldsymbol{z})\geq r\right) = \int I\left(s(\boldsymbol{z})\geq r\right) f(\boldsymbol{z}, \boldsymbol{q}) d\boldsymbol{z} \tag{16}$$

with respect to pdf (13) given parameter $\boldsymbol{q}$. For a large value of r, (16) is a rare event simulation. In fact, $\boldsymbol{q}$ reflects the relationship between (15) and (16). It is anticipated that $\boldsymbol{q}$ contains either 1's or 0's so that the generation of $\boldsymbol{z}$ is determined and $\boldsymbol{q}$ can equal $\boldsymbol{z}_{\mathrm{opt}}$ at the optimum point. Therefore, we aim to improve the estimation for $\boldsymbol{q}$.

A simple way to estimate $\psi(r)$ is using the crude Monte Carlo simulation, which draws J random samples $\{\boldsymbol{z}_{[i]}, i = 1, 2, \ldots, J\}$ with each sample $\boldsymbol{z}_{[i]}$ being a binary sequence defined in (10), producing an unbiased estimator of $\psi(r)$ as

$$\hat{\psi}(r) = \frac{1}{J}\sum_{i=1}^{J} I\left(s\left(\boldsymbol{z}_{[i]}\right)\geq r\right). \tag{17}$$

However, it requires a large number of samples to have an accurate estimation, as most samples are not effective in learning $\hat{\psi}(r)$. Instead, we may use importance sampling, which can greatly reduce the complexity, i.e., finding another importance distribution $\xi(\boldsymbol{z})$ for generating samples $\boldsymbol{z}$ such that $s(\boldsymbol{z})\geq r$ occurs more often. Then, the unbiased estimator can be obtained from

$$\hat{\psi}(r) = \frac{1}{J}\sum_{i=1}^{J} I\left(s\left(\boldsymbol{z}_{[i]}\right)\geq r\right)\frac{f\left(\boldsymbol{z}_{[i]}, \boldsymbol{q}\right)}{\xi\left(\boldsymbol{z}_{[i]}\right)}. \tag{18}$$

In the sense of minimizing the variance of $\hat{\psi}(r)$, the optimal biased importance distribution is given by

$$\xi^*(\boldsymbol{z}) = \frac{I\left(s(\boldsymbol{z})\geq r\right) f(\boldsymbol{z}, \boldsymbol{q})}{\psi(r)}. \tag{19}$$

Cross-entropy optimization minimizes the Kullback–Leibler divergence, which is also referred to as the cross-entropy distance, between

$\xi^*(\boldsymbol{z})$ and $f(\boldsymbol{z}, \boldsymbol{q})$ by solving

$$\max_{\boldsymbol{q}} \int \xi^*(\boldsymbol{z}) \ln f(\boldsymbol{z}, \boldsymbol{q}) d\boldsymbol{z}. \tag{20}$$

Substituting $\xi^*(\boldsymbol{z})$ in (19), we then have

$$\max_{\boldsymbol{q}} \int \frac{I\left(s(\boldsymbol{z}) \geq r\right) f(\boldsymbol{z}, \boldsymbol{q})}{\psi(r)} \ln f(\boldsymbol{z}, \boldsymbol{q}) d\boldsymbol{z}. \tag{21}$$

Although it is intractable to obtain a closed-form solution for (21), we may estimate it by finding

$$\boldsymbol{q}^* = \arg\max_{\boldsymbol{q}} \Gamma(\boldsymbol{q}) \tag{22}$$

where

$$\Gamma(\boldsymbol{q}) = \frac{1}{J} \sum_{i=1}^{J} I\left(s\left(\boldsymbol{z}_{[i]}\right) \geq r\right) \ln f\left(\boldsymbol{z}_{[i]}, \boldsymbol{q}\right). \tag{23}$$

We set $\partial \Gamma(\boldsymbol{q}) / \partial \boldsymbol{q} = 0$ and obtain the updating rule as

$$q_l = \frac{\sum_{i=1}^{J} I\left(s\left(\boldsymbol{z}_{[i]}\right) \geq r\right) \boldsymbol{z}_{[i]}(l)}{\sum_{i=1}^{J} I\left(s\left(\boldsymbol{z}_{[i]}\right) \geq r\right)} \tag{24}$$

for $l = 1, 2, \ldots, M$, where q_l and $\boldsymbol{z}_{[i]}(l)$ denote the lth entry of $\boldsymbol{q}$ and the lth entry of $\boldsymbol{z}_{[i]}$, respectively. The rule is iteratively performed until r converges to an optimum value r^*.

At the tth iteration, we employ $\boldsymbol{q}^{(t)}$ to generate a set of new samples, i.e., $\{\boldsymbol{z}_{[i]}^{(t)}, i = 1, 2, \ldots, J\}$, according to (12). Then, based on (24), we obtain

$$\tilde{q}_l^{(t+1)} = \frac{\sum_{i=1}^{J} I\left(s\left(\boldsymbol{z}_{[i]}^{(t)}\right) \geq r^{(t)}\right) \boldsymbol{z}_{[i]}^{(t)}(l)}{\sum_{i=1}^{J} I\left(s\left(\boldsymbol{z}_{[i]}^{(t)}\right) \geq r^{(t)}\right)} \tag{25}$$

for $l = 1, 2, \ldots, M$. To prevent the occurrences of all 0's or all 1's in $\tilde{\boldsymbol{q}}^{(t+1)}$, which could terminate the iteration unexpectedly, a smoothing factor, i.e., $\lambda \in (0, 1)$, is used so that

$$\boldsymbol{q}^{(t+1)} = (1 - \lambda) \tilde{\boldsymbol{q}}^{(t+1)} + \lambda \boldsymbol{q}^{(t)}. \tag{26}$$

As an initial for the iterations, we may set

$$\boldsymbol{q}^{(1)} = \underbrace{\left\{\frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2}\right\}}_{M} \tag{27}$$

supposing equal probability for each subcarrier being selected as pilots.

In practice, we usually set $r^{(t)}$ to be the $(1 - \rho)$ sample quantile after we sort the objective values $s(\boldsymbol{z}_{[1]}^{(t)}), s(\boldsymbol{z}_{[2]}^{(t)}), \ldots, s(\boldsymbol{z}_{[J]}^{(t)})$ from the smallest to the biggest, where $0 < \rho < 1$. Suppose the results after sorting are $s_1, s_2, \ldots, s_J$. We set

$$r^{(t)} = s_\alpha \tag{28}$$

where $\alpha = \lceil (1 - \rho) J \rceil$.

In addition, note that to satisfy the following restriction condition:

$$\sum_{l=1}^{M} z_{[i]}^{(t)}(l) = K \tag{29}$$

which is essentially (11), an additional projection method is needed. Although we may simply perform the projection as randomly adding or deleting 1's from $\boldsymbol{z}_{[i]}^{(t)}$, so that (29) can be satisfied, it essentially changes the distribution of $\boldsymbol{q}^{(t)}$. Therefore, we propose a projection method as follows, if (29) is not satisfied.

1) If $\sum_{l=1}^{M} z_{[i]}^{(t)}(l) > K$, from vector $\boldsymbol{z}_{[i]}^{(t)}$, we sequentially remove the nonzero entry corresponding to the smallest probability in $\boldsymbol{q}^{(t)}$ until (29) is satisfied.
2) If $\sum_{l=1}^{M} z_{[i]}^{(t)}(l) < K$, from vector $\boldsymbol{z}_{[i]}^{(t)}$, we sequentially change the zero entry corresponding to the largest probability in $\boldsymbol{q}^{(t)}$ to 1 until (29) is satisfied.

The procedures for the pilot design using constrained cross-entropy optimization are summarized in Algorithm 1. Given the maximum number of iterations T, we iteratively draw random samples using $\boldsymbol{q}^{(t)}$ and figure out $\boldsymbol{q}^{(t+1)}$ for the next iteration. For some positive integer u, if $r^{(t)} = r^{(t-1)} = \cdots = r^{(t-u)}$ holds, which means $r^{(t-u)}$ does not change for a number of subsequent iterations, we break from the loop and output $\boldsymbol{q}^{(t)}$. The positions of 1's in $\boldsymbol{q}^{(t)}$ make up an optimized pilot pattern $\boldsymbol{p}_o$, which has been demonstrated to converge to $\boldsymbol{p}_{\text{opt}}$. Moreover, $r^{(t)}$ converges to $g(\boldsymbol{p}_o)$ [10].

Algorithm 1-Pilot Design Algorithm using Constrained Cross-Entropy Optimization

```
Input: C, M, K, L, T, J, λ, ρ, and u.
Generate G according to (9).
Initialize q^(1) according to (27).
for t = 1, 2, ..., T
    Draw J random samples {z_i^(t), i = 1, 2, ..., J} from the
    probability function (12) with q^(t).
    Check the restriction condition (29) and make projections
    if needed.
    Obtain the objective values {s(z_i^(t)), i = 1, 2, ..., J}.
    Compute r^(t) according to (28).
    if t ≥ u and r^(t) = r^(t-1) = ··· = r^(t-u)
        Break;
    end if
    Obtain q^(t+1) using (25) and (26).
end for
output: q^(t).
```

IV. Simulation Results

An OFDM-based CR system with $N = 1024$ subcarriers is considered. Note that we assume ideal spectrum sensing without false alarm or missing detection.[1] After spectrum sensing and deactivating those subcarriers occupied by PUs, we assume that there are $M = 512$ remaining OFDM subcarriers for SUs, including three noncontiguous subcarrier blocks, i.e., $\{1, 2, \ldots, 256\}$, $\{513, 514, \ldots, 640\}$ and $\{897, 898, \ldots, 1024\}$, with the number of subcarriers in each block being 256, 128, and 128, respectively. From $\mathcal{C} = \{1, 2, \ldots, 256, 513, 514, \ldots, 640, 897, 898, \ldots, 1024\}$, which can be equivalently treated as the union of all available CR subbands, we want to pick up $K = 16$ pilot subcarriers for frequency-domain pilot-assisted channel estimation. A sparse multipath channel $\boldsymbol{h}$ is generated with $L = 60$ taps, where $V = 5$ dominant nonzero channel taps are

[1] With false alarm, some idle subcarriers will be falsely alarmed as unavailable subcarriers occupied by PUs. However, the set of the remaining subcarriers is still large enough for us to obtain an optimized pilot pattern using the proposed scheme. With missing detection, some unavailable subcarriers occupied by PUs will be falsely treated as ideal subcarriers. Once they happen to be the entries of an optimized pilot pattern, the received pilots at SUs will be corrupted due to the interference between PUs and SUs, leading to unreliable channel estimation. To deal with the interference caused by missing detection, we will introduce the cooperation strategies [2], [12] to our work in the future.

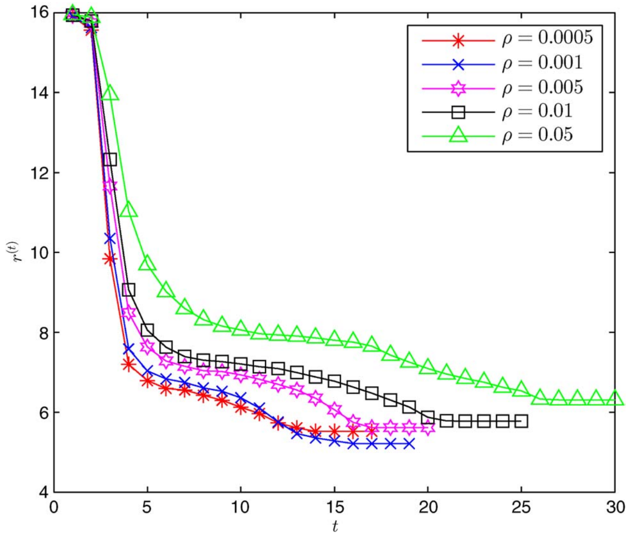

Fig. 2. Convergence of $r^{(t)}$ with t for different values of ρ.

TABLE I
COMPARISONS OF PILOT PATTERNS USING DIFFERENT PILOT DESIGN SCHEMES AND PARAMETERS

Type	$g(\boldsymbol{p}_o)$	$\boldsymbol{p}_o$
$\rho = 0.0005$	5.5220	76, 91, 144, 171, 202, 225, 249, 514 526, 622, 636, 910, 949, 976, 1012, 1023
$\rho = 0.001$	5.2164	47, 95, 110, 162, 180, 193, 246, 513 524, 627, 640, 897, 910, 939, 976, 1019
$\rho = 0.005$	5.6151	77, 101, 171, 186, 217, 237, 248, 514 525, 627, 640, 943, 970, 996, 1007, 1024
$\rho = 0.01$	5.7806	65, 97, 166, 179, 191, 202, 228, 513 524, 620, 640, 897, 945, 978, 1023, 1024
$\rho = 0.05$	6.3118	120, 134, 164, 197, 530, 544, 628, 640 897, 911, 939, 957, 978, 992, 1023, 1024
Exhaustive search	6.7889	70, 105, 133, 147, 197, 255, 523, 531 588, 625, 937, 948, 958, 963, 977, 1009
Equally spaced pilots	10.9241	1, 37, 73, 109, 145, 181, 217, 253, 513, 555, 597, 639, 897, 939, 981, 1023

randomly placed among L taps. The channel gain of each path is i.i.d. complex Gaussian distributed with zero mean and unit variance, i.e., $\mathcal{CN}(0, 1)$. Quaternary phase-shift keying modulation is employed in the simulations.

Since it is computationally very inefficient to exhaustively search all $\binom{512}{16} = 8.41 \times 10^{29}$ pilot patterns, we use the proposed scheme based on constrained cross-entropy optimization to design the pilots. As shown in Fig. 2, we compare the convergence[2] of $r^{(t)}$ for different values of ρ. The other parameters are fixed to be $T = 50$, $J = 100\,000$, $u = 3$, and $\lambda = 0.3$. It is seen that with larger ρ, more iterations are needed for convergence. However, we cannot always decrease ρ. Although $\rho = 0.0005$ achieves faster convergence than $\rho = 0.001$ with two less iterations, the latter obtains a smaller objective value than the former, as indicated by $g(\boldsymbol{p}_o)$ in Table I. The reason is that we have to guarantee that the number of elite samples $J \times \rho$ is sufficiently large for the rare event simulation. For a very small ρ, we have to increase J, which, in turn, takes more running time for each iteration. From Table I, we observe that $\rho = 0.001$ achieves the smallest $g(\boldsymbol{p}_o)$. Both $\rho = 0.0005$ and $\rho = 0.005$ have greater $g(\boldsymbol{p}_o)$ than $\rho = 0.001$. For comparisons, we also give the results using the exhaustive search, where we generate the pilot pattern randomly, and the best result obtained during a predetermined maximum running time, i.e., 540 s, which is the running time for $\rho = 0.05$, is the final output of the exhaustive search.[3] We also list the equally spaced pilots that prove best for the traditional LS channel estimation, where we use eight, four, and four pilot subcarriers for contiguous subcarrier blocks $\{1, 2, \ldots, 256\}$, $\{513, 514, \ldots, 640\}$, and $\{897, 898, \ldots, 1024\}$, respectively.

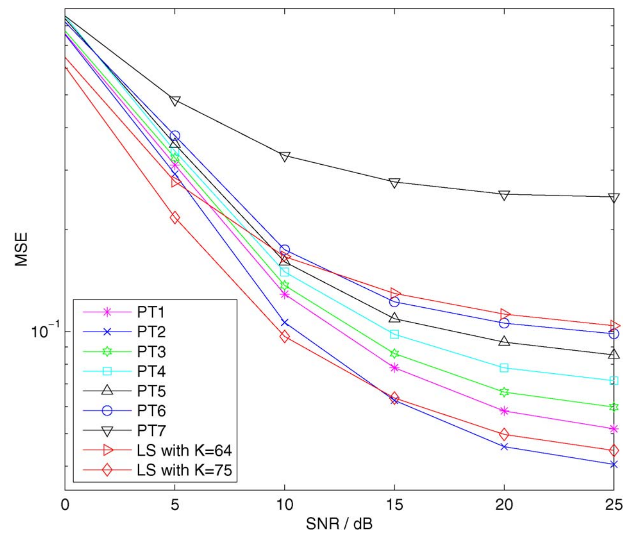

Fig. 3. MSE performance comparisons of channel estimation for different pilot design schemes.

We now evaluate the channel estimation performance using the designed pilot patterns. For ease of notation in the figure, we use PT1, PT2, . . ., PT7 to represent the seven different pilot patterns in Table I. The MSE performance for channel estimation and the bit error rate (BER) performance for data detection are shown in Figs. 3 and 4, respectively. Both the MSE and BER are averaged over 10 000 sparse channel realizations. The popular OMP algorithm is employed for sparse channel estimation given the pilot patterns PT1, PT2, . . ., PT7. For comparisons, the performance of LS channel estimation using $K = 64$ and $K = 75$ equally spaced pilots, with pilot interval being 8 and 7, respectively, is also provided. We observe that the proposed scheme with $\rho = 0.001$ (PT2) performs slightly better than LS with $K = 75$. In other words, it can save 59 pilot subcarriers under the same channel estimation performance, thus leading to 11.5% improvement in spectrum efficiency. Even for the exhaustive search, it reaches the performance of LS with $K = 64$, resulting in 9.4% improvement in spectrum efficiency. However, for equally spaced pilots with $K = 16$ (PT7), it is much worse than the others, indicating that the pilot pattern without optimization is improper for sparse channel estimation. The error floor observed in Figs. 3 and 4 can be overcome by using more pilots, however, at the expense of reduced spectrum efficiency and increased pilot design complexity. With a fixed number of pilots, we can improve the performance by finely tuning the parameters, i.e., ρ, for the proposed scheme. It is seen that the proposed scheme with $\rho = 0.001$ (PT2) outperforms that with other values of ρ and achieves a 12-dB signal-to-noise ratio gain at BER $= 0.014$ compared with the exhaustive search. With channel encoding and decoding, a BER of around 0.01 can be completely removed in practical systems.

[2]Here, we compare the convergence speed with respect to the number of the iterations instead of the running time, because the running time of each iteration is almost the same for different ρ's. The simulations are performed using MATLAB v7.14 (R2012a) on a laptop equipped with an Intel Core 2 Duo CPU at 2.5 GHz and 4 GB of memory.

[3]In general, the exhaustive search is a brute-force method that exhaustively examines all possibilities. However, since the generation of all $\binom{512}{16} = 8.41 \times 10^{29}$ candidates is almost impossible, we define the exhaustive search with consistency to most literatures [13].

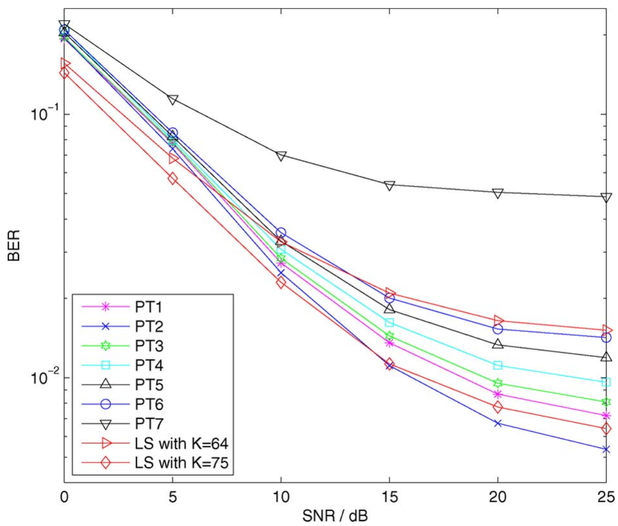

Fig. 4. BER performance comparisons of channel estimation for different pilot design schemes.

V. Conclusion

In this correspondence, we have investigated the sparse channel estimation in OFDM-based CR systems. Based on the results of spectrum sensing, we have considered the pilot design by minimizing the coherence of the dictionary matrix. We have formulated it as an optimal column selection problem where a table has been generated and the indexes of the selected columns of the table form a pilot pattern. A novel scheme using constrained cross-entropy optimization has been proposed to obtain an optimized pilot pattern, where we have modeled it as an independent Bernoulli random process. The updating rule for the probability of each active subcarrier being selected as a pilot subcarrier has been derived. A projection method has been proposed so that the number of pilot subcarriers during optimization is fixed. Simulation results have verified the effectiveness of the proposed scheme and shown that it can achieve 11.5% improvement in spectrum efficiency with the same channel estimation performance compared with the LS channel estimation.

References

[1] Y. Liang, K. C. Chen, G. Y. Li, and P. Mahonen, "Cognitive radio networking and communications: An overview," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3386–3407, Sep. 2011.

[2] Y. Zou, Y. Yao, and B. Zheng, "Cooperative relay techniques for cognitive radio systems: Spectrum sensing and secondary user transmissions," *IEEE Commun. Mag.*, vol. 50, no. 4, pp. 98–103, Apr. 2012.

[3] D. Hu, L. He, and X. Wang, "An efficient pilot design method for OFDM-based cognitive radio systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1252–1259, Apr. 2011.

[4] E. Manasseh, S. Ohno, and M. Nakamoto, "Pilot design for non-contiguous spectrum usage in OFDM-based cognitive radio networks," in *Proc. EUSIPCO*, Bucharest, Romania, Aug. 2012, pp. 465–469.

[5] J. C. Chen and C. K. Wen, "A novel cognitive radio adaptation for wireless multicarrier systems," *IEEE Commun. Lett.*, vol. 14, no. 7, pp. 629–631, Jul. 2010.

[6] C. R. Berger, Z. Wang, J. Huang, and S. Zhou, "Application of compressive sensing to sparse channel estimation," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 164–174, Nov. 2010.

[7] F. Wan, W. P. Zhu, and M. N. S. Swamy, "Semiblind sparse channel estimation for MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2569–2582, Jul. 2011.

[8] C. Qi and L. Wu, "Application of compressed sensing to DRM channel estimation," in *Proc. 73rd IEEE VTC-Spring*, Budapest, Hungary, May 2011, pp. 1–5.

[9] T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4680–4688, Jul. 2011.

[10] L. Margolin, "On the convergence of the cross-entropy method," *Ann. Oper. Res.*, vol. 134, no. 1, pp. 201–214, Feb. 2005.

[11] R. Y. Rubinstein and D. P. Kroese, *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation, and Machine Learning*. New York, NY, USA: Springer-Verlag, 2004.

[12] Y. Zou, Y. Yao, and B. Zheng, "Cognitive transmissions with multiple relays in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 2, pp. 648–659, Feb. 2011.

[13] C. Qi and L. Wu, "A study of deterministic pilot allocation for sparse channel estimation in OFDM systems," *IEEE Comm. Lett.*, vol. 16, no. 5, pp. 742–744, May 2012.

Optimal Correlative Coding for Discrete-Time OFDM Systems

Miaowen Wen, Xiang Cheng, *Senior Member, IEEE*, Jianjun Wu, *Member, IEEE*, Liuqing Yang, *Senior Member, IEEE*, and Bingli Jiao, *Senior Member, IEEE*

***Abstract*—Frequency-domain correlative coding (CC) has been widely considered to be effective in intercarrier interference (ICI) mitigation in orthogonal frequency-division multiplexing (OFDM) systems due to its potential in the considerable enhancement of the carrier-to-interference power ratio (CIR). Existing work has focused on searching for the optimal coding coefficients in minimizing the ICI power. However, the fundamental analysis is largely based on the continuous-time signal model (CTSM) and the assumption of an infinite number of subcarriers. In practice, the OFDM signal always involves a finite bandwidth and is sampled for digital processing purposes. Attentive to the given needs of practical systems, we derive the optimal coding coefficients in the maximization of the CIR or the minimization of the ICI power based on the discrete-time signal model (DTSM) with a finite number of subcarriers. Results show that the optimal coding coefficients for the DTSM differ from the existing coefficients developed for the CTSM and that our performance analysis leads to a much improved coding gain prediction for practical systems.**

***Index Terms*—Correlative coding (CC), Doppler shift, intercarrier interference (ICI), orthogonal frequency-division multiplexing (OFDM).**

Manuscript received November 13, 2012; revised March 30, 2013 and June 3, 2013; accepted July 11, 2013. Date of publication August 7, 2013; date of current version February 12, 2014. This work was supported in part by the National Natural Science Foundation of China under Grant 61101079 and Grant 61071083, by the National 973 project under Grant 2013CB336700, by the Science Foundation for the Youth Scholar of the Ministry of Education of China under Grant 20110001120129, by the State Key Laboratory of Integrated Service Networks of Xidian University under Grant ISN13-05 and Grant ISN-ZR02, by the National Science Foundation under Grant CNS-1129043, by the Ministry of Transport of China under Grant 2012-364-X03-104, and by the China Scholarship Council. The review of this paper was coordinated by Dr. I. Kridikis. *(Corresponding author: X. Cheng.)*

M. Wen, X. Cheng, and B. Jiao are with the State Key Laboratory of Advanced Optical Communication Systems and Networks, School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China, and also with the State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China (e-mail: wenmiaowen@pku.edu.cn; xiangcheng@pku.edu.cn; jiaobl@pku.edu.cn).

J. Wu is with the School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China (e-mail: just@pku.edu.cn).

L. Yang is with Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523 USA (e-mail: liuqing.yang@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2013.2277575