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 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 the sparse recovery. Then it's formulated as an optimal column selection problem where a table is generated and the indices 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's modeled as an independent Bernoulli random process. The updating rule for the probability of each active subcarrier selected as 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 to the LS channel estimation.

Index Terms—Cognitive radio (CR), OFDM, pilot design, sparse channel estimation, compressed sensing (CS).

I. INTRODUCTION

Traditionally every wireless system is required to have an exclusive spectrum license in order to avoid interference from other systems or users. However, recent studies have shown that a large portion of the licensed spectrum is underutilized. It then motivates the studies on cognitive radio (CR), which allows the secondary users (SUs) to utilize the licensed spectrum without making any interference to the licensed users or the primary users (PUs) and aims to improve the spectrum utilization without allocating new spectrum resource [1], [2]. Orthogonal frequency division multiplexing (OFDM), which has been considered as 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 non-contiguous 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 indices are obtained by solving a series of one-dimensional 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 cross entropy method is demonstrated to outperform genetic algorithm (GA) and particle swarm optimization (PSO). However, all of them are based on the least squares (LS) channel estimation.

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 (CoSaMP) and basis pursuit (BP) have been applied. Therefore it’s nature to extend this technique to OFDM-based CR systems, which can further improve the data rate and flexibility of SUs. But it also brings new challenges to the pilot design. To the authors’ best knowledge, so far there have been no literature studying 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’s not optimal since it does not benefit from the sparse channel estimation.

In this correspondence, we first introduce the 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 the sparse recovery. We then formulate it as an optimal column selection problem where a table is generated and the indices 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 pilot subcarrier is derived. Moreover, a projection method is proposed so that the number of pilot subcarriers during the optimization is fixed.

The remainder of the 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 the correspondence.

The notations used in this paper are defined as follows. Symbols for matrices (upper case) and vectors (lower case) are in boldface. $(\cdot)^T$, $(\cdot)^H$, diag{}, $I_L$, $CN$, $|\cdot|$ and $\lceil \cdot \rceil$ denote the matrix transpose, 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 Figure 1, where we employ sparse channel estimation instead of the 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 firstly 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 (ISI), 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 FIR filter with the channel impulse response (CIR) to be $h = [h(1), h(2), \ldots, h(L)]^T$. At the receiver, after CP removal and 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 a 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 as $\mathcal{C} = \{c_1, c_2, \ldots, c_M\}$, which is a non-contiguous integer set representing the indices 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 pilot symbols and the receive pilot symbols are denoted as $x(c_{p_1}), x(c_{p_2}), \ldots, x(c_{p_K})$ and $y(c_{p_1}), y(c_{p_2}), \ldots, y(c_{p_K})$, respectively. For traditional least square (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

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

where $\eta(i) \sim CN(0, \sigma^2), i = 1, 2, \ldots, K$ is the independently and identically distributed (i.i.d.) additive white Gaussian noise (AWGN), and $F_{K \times L}$ is a DFT submatrix, given by

$$
F_{K \times L} = \frac{1}{\sqrt{N}}
\begin{bmatrix}
1 & \omega^{f_{p_1}} & \ldots & \omega^{f_{p_1}(L-1)} \\
1 & \omega^{f_{p_2}} & \ldots & \omega^{f_{p_2}(L-1)} \\
\vdots & \vdots & \ddots & \vdots \\
1 & \omega^{f_{p_K}} & \ldots & \omega^{f_{p_K}(L-1)}
\end{bmatrix}
$$

where $N (N \geq M)$ is the number of points for IFFT and FFT in Figure 1, and $\omega = e^{-j2\pi/N}$. We denote

$$
X = \text{diag}\{x(c_{p_1}), x(c_{p_2}), \ldots, x(c_{p_K})\},
$$

$$
y = [y(c_{p_1}), y(c_{p_2}), \ldots, y(c_{p_K})]^T,
$$

$$
\eta = [\eta(1), \eta(2), \ldots, \eta(K)]^T \sim CN(0, \sigma^2 I_K).
$$

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

$$
y = Ah + \eta.
$$

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

$$
\hat{h}_{LS} = (A^HA)^{-1}A^Hy.
$$

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 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 \( h \) are either zero or nearly zero, implying \( h \) is sparse. With this a-priori condition, we can apply CS algorithms to estimate \( h \). Many works have already demonstrated that these algorithms outperform LS for channel estimation [6].

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

Given a pilot pattern
\[
\mathbf{p} = \{c_{p_1}, c_{p_2}, \ldots, c_{p_K}\},
\]
we define the coherence of \( A \) as the maximum absolute correlation between any two different columns of \( A \), i.e.,
\[
g(\mathbf{p}) = \max_{0 \leq m < n \leq L-1} |\langle A(m), A(n) \rangle| = \max_{0 \leq m < n \leq L-1} \left| \sum_{i=1}^{K} x(c_{p_i})^2 \right|
\]
where \( \langle A(m), A(n) \rangle \) denotes the inner product of \( A(m) \) and \( A(n) \), i.e., \( \langle A(m), A(n) \rangle = A^H(m) \cdot A(n) \). One straightforward approach to obtain \( g(\mathbf{p}) \) is to compute the \( \ell_\infty \) norm of \( A^H A \). The objective function for the pilot design is to minimize the coherence of \( A \), i.e.,
\[
Q = \min_{\mathbf{p}} g(\mathbf{p}).
\]
The optimal pilot pattern is then given by
\[
\mathbf{p}_{\text{opt}} = \arg \min_{\mathbf{p}} g(\mathbf{p}).
\]
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 complexities. Let \( d = n - m \) and \( \Lambda = \{1, 2, \ldots, L-1\} \). Then (5) can be rewritten as
\[
g(\mathbf{p}) = E \cdot \max_{d \in \Lambda} \left| \sum_{i=1}^{K} \omega^{c_{p_i}d} \right|.
\]
For simplicity, we assume \( E = 1 \). We can generate a table
\[
G = \begin{bmatrix}
\omega^1 & \omega^2 & \cdots & \omega^M \\
\omega^{21} & \omega^{22} & \cdots & \omega^{2M} \\
\vdots & \vdots & \ddots & \vdots \\
\omega^{(L-1)c_1} & \omega^{(L-1)c_2} & \cdots & \omega^{(L-1)c_M}
\end{bmatrix}
\]
with the dimension being \((L-1) \times M\). Once \( \mathbf{p} \) is given, we look up \( G \) and find the corresponding \( K \) columns, composing a \((L-1) \times K\) submatrix \( \hat{G} \). We make summations for each row of \( \hat{G} \), resulting in a \((L-1) \times 1\) dimensional vector, where we obtain its \( \ell_\infty \) norm as the final value of \( g(\mathbf{p}) \). Compared to the approach that always computes the \( \ell_\infty \) norm of \( A^H A \) involving matrix multiplications, it can be much faster implemented with \( G \) generated beforehand.

### III. Pilot Design Scheme

Although we may exhaustively search all possible pilot patterns for the best one with the minimum objective, it’s 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, now we propose a scheme using stochastic optimization algorithms which 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
\[
z = [z(i), \ i = 1, 2, \ldots, M], \ z(i) \in \{0, 1\}.
\]
If \( c_i \) is selected as pilot subcarrier, \( z(i) = 1 \); otherwise, \( z(i) = 0 \), \( i = 1, 2, \ldots, M \). The restriction is
\[
\sum_{i=1}^{M} z(i) = K
\]
as the number of pilots is explicitly given as \( K \). So \( z \) and \( p \) are equivalent. The optimization for \( p \) is therefore converted to the optimization for \( z \). We model each entry of \( z \) as an independent Bernoulli random variable with the probability to be
\[
Pr(z(i) = 1) = q_i, \ z(i) = 0
\]
for \( i = 1, 2, \ldots, M \), where \( q_i \) is the probability of \( c_i \) being selected as pilot subcarrier. We denote \( \mathbf{q} = \{q_i, \ i = 1, 2, \ldots, M\} \).

Then the probability density function (p.d.f.) associated with the pilot design problem is
\[
f(z, \mathbf{q}) = \prod_{i=1}^{M} q_i^{z(i)} (1 - q_i)^{1-z(i)}.
\]
Since from each of \( p \) and \( z \) we can decide the other, \( p \) and \( z \) are equivalent. We define \( s(z) \equiv -g(\mathbf{p}) \). Correspondingly, (6) is converted into a maximization problem which is typical for cross entropy optimization, i.e.,
\[
Q = -\max_{z} s(z).
\]
Meanwhile, (7) is converted into
\[
z_{\text{opt}} = \arg \max_{z} s(z)
\]
where \( z_{\text{opt}} \) is equivalent to \( p_{\text{opt}} \). Then we define an indicator function \( I(e) \) which returns one if the event \( e \) is true and returns zero otherwise. For some threshold \( r \), the probability of \( s(z) \) being greater than \( r \) is
\[
\psi(r) = Pr(s(z) \geq r) = \int I(s(z) \geq r) f(z, \mathbf{q}) dz
\]
with respect to the p.d.f. (13) given the parameter \( \mathbf{q} \). For a large value of \( r \), (16) is a rare event simulation. In fact, \( \mathbf{q} \) reflects the relationship between (15) and (16). It’s anticipated that \( \mathbf{q} \) contains either ones or zeros so that the generation of \( z \) is determined and \( \mathbf{q} \) can equal \( z_{\text{opt}} \) at the optimum point. Therefore we aim to improve the estimation for \( \mathbf{q} \).
A simple way to estimate \( \psi(r) \) is using the crude Monte-Carlo simulation, which draws \( J \) random samples \( \{z_{[i]}\}, i = 1, 2, \ldots, J \) with each sample \( z_{[i]} \) being a binary sequence defined in (10), producing an unbiased estimator of \( \psi(r) \) as

\[
\hat{\psi}(r) = \frac{1}{J} \sum_{j=1}^{J} I\left(s(z_{[j]}), q_{[j]} \right) f(z_{[j]}). (17)
\]

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

\[
\hat{\psi}(r) = \frac{1}{J} \sum_{i=1}^{J} I\left(s(z_{[i]}), q_{[i]} \right) f(z_{[i]}, q_{[i]}). (18)
\]

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

\[
\xi^*(z) = \frac{I(s(z) \geq r) f(z, q)}{\psi(r)}. (19)
\]

Cross entropy optimization minimizes the Kullback-Leibler divergence, also referred to as the cross entropy distance, between \( \xi^*(z) \) and \( f(z, q) \) by solving

\[
\max_q \int \xi^*(z) \ln f(z, q)dz. (20)
\]

Substituting \( \xi^*(z) \) in (19), we then have

\[
\max_q \int \frac{I(s(z) \geq r) f(z, q)}{\psi(r)} \ln f(z, q)dz. (21)
\]

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

\[
q^* = \arg \max_q \Gamma(q) (22)
\]

where

\[
\Gamma(q) = \frac{1}{J} \sum_{l=1}^{J} I\left(s(z_{[l]}), q_{[l]} \right) \ln f(z_{[l]}, q_{[l]}). (23)
\]

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

\[
q_l = \frac{\sum_{s(z_{[l]}) \geq r} I(s(z_{[l]}), q_{[l]}(l))}{\sum_{s(z_{[l]}) \geq r}} (24)
\]

for \( l = 1, 2, \ldots, M \), where \( q_l \) and \( z_{[l]}(l) \) denote the \( l \)th entry of \( q \) and the \( l \)th entry of \( z_{[l]} \), respectively. The rule is iteratively performed until \( r \) converges to an optimum value \( r^\ast \).

At the \( k \)th iteration, we employ \( q^{(k)} \) to generate a set of new samples, \( \{z_{[i]}^{(k)}, i = 1, 2, \ldots, J\} \), according to (12). Then based on (24) we obtain

\[
q_l^{(k+1)} = \frac{\sum_{l=1}^{J} I(s(z_{[l]}^{(k)}) \geq r^{(k)}(l))z_{[l]}^{(k)}(l)}{\sum_{l=1}^{J} I(s(z_{[l]}^{(k)}) \geq r^{(k)}(l))} (25)
\]

for \( l = 1, 2, \ldots, M \). In order to prevent the occurrences of all zeros or all ones in \( q^{(k+1)} \) which could terminate the iteration unexpectedly, a smoothing factor, \( \lambda \in (0, 1) \), is used so that

\[
q^{(k+1)} = (1 - \lambda)q^{(k+1)} + \lambda q^{(k)}. (26)
\]

As an initial for the iterations, we may set

\[
q^{(1)} = \left\{ \frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2} \right\} (27)
\]

supposing equal probability for each subcarrier being selected as pilots.

In practice, we usually set \( r^{(k)} \) to be the \( (1 - \rho) \) sample quantile after we sort the objective values \( s(z_{[1]}^{(k)}), s(z_{[2]}^{(k)}), \ldots, s(z_{[M]}^{(k)}) \) 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

\[
r^{(k)} = s_\alpha (28)
\]

where \( \alpha = \lfloor (1 - \rho)J \rfloor \).

Also note that in order to satisfy the following restriction condition

\[
\sum_{l=1}^{M} z_{[l]}^{(k)}(l) = K (29)
\]

which is essentially (11), an additional projection method is needed. Although we may simply perform the projection as randomly adding or deleting ones from \( z_{[l]}^{(k)} \), so that (29) can be satisfied, it essentially changes the distribution of \( q^{(k)} \). Therefore we propose a projection method as follows, if (29) is not satisfied.

- If \( \sum_{l=1}^{M} z_{[l]}^{(k)}(l) > K \), from the vector \( z_{[l]}^{(k)} \) we sequentially remove the nonzero entry corresponding to the smallest probability in \( q^{(k)} \) until (29) is satisfied.
- If \( \sum_{l=1}^{M} z_{[l]}^{(k)}(l) < K \), from the vector \( z_{[l]}^{(k)} \) we sequentially change the zero entry corresponding to the largest probability in \( q^{(k)} \) to one 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 \( q^{(k)} \) and figure out \( q^{(k+1)} \) for the next iteration. For some positive integer \( u \), if \( r^{(k)} = r^{(k-u)} = \cdots = r^{(k-u)} \), which means \( r^{(k-u)} \) does not change for a number of subsequent iterations, we break from the loop and output \( q^{(k)} \). The positions of ones in \( q^{(k)} \) make up an optimized pilot pattern \( p_{opt} \), which has been demonstrated to converge to \( p_{opt} \) [10].

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. After spectrum sensing and deactivating those subcarriers occupied by PUs, we assume there are \( M = 512 \) remaining OFDM subcarriers for SUs, including three non-contiguous subcarriers.

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. In order to deal with the interference caused by missing detection, we will introduce the cooperation strategies [2], [12] to our work in the future.
Algorithm 1 - Pilot Design Algorithm using Constrained Cross Entropy Optimization

1: Input: \( C, M, K, L, T, J, \lambda, \rho \) and \( u \).
2: Generate \( G \) according to (9).
3: Initialize \( q^{(1)} \) according to (27).
4: for \( t = 1, 2, \ldots, T \) do
5:     Draw \( J \) random samples \( \{z_t^{(i)} \}, i = 1, 2, \ldots, J \) from the probability function (28) with \( q^{(t)} \).
6:     Check the restriction condition (29) and make projections if needed.
7:     Obtain the objective values \( s(z_t^{(i)}), i = 1, 2, \ldots, J \).
8:     Compute \( r^{(t)} \) according to (28).
9:     if \( t \geq u \) and \( r^{(t)} = r^{(t-1)} = \cdots = r^{(t-u)} \) then
10:        Break;
11:     end if
12:     Obtain \( q^{(t+1)} \) using (25) and (26).
13: end for
14: output: \( q^{(T)} \).

Fig. 2. Convergence of \( r^{(t)} \) with \( t \) for different value of \( \rho \).

subcarrier blocks, \{1, 2, ..., 256\}, \{513, 514, ..., 640\} and \{897, 898, ..., 1024\}, with the number of subcarriers in each block being 256, 128, and 128, respectively. From \( C = \{1, 2, ..., 256, 513, 514, ..., 640, 897, 898, ..., 1024\}\), 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 \( h \) is generated with \( L = 60 \) taps where \( V = 5 \) dominant non-zero channel taps are 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., \( \mathcal{CN}(0, 1) \). QPSK modulation is employed in the simulations.

Since it’s computationally very inefficient to exhaustively search all \( 2^{512} = 8.41 \times 10^{159} \) pilot patterns, we use the proposed scheme based on constrained cross entropy optimization to design the pilots. As shown in Figure 2, we compare the convergence\(^2\) of \( r^{(t)} \) for different value of \( \rho \). The other parameters are fixed to be \( T = 50, J = 100000, u = 3 \) and \( \lambda = 0.3 \).

It’s seen that with larger \( \rho \), more iterations are needed for the convergence. However, we cannot always decrease \( \rho \). Although \( \rho = 0.0005 \) achieves faster convergence than \( \rho = 0.001 \) with 2 less iterations, the latter obtains smaller objective value than the former, as indicated by \( g(p_\rho) \) in Table I. The reason is that we have to guarantee the number of elite samples \( J \times \rho \) is sufficiently large for the rare event simulation. For a very small \( \rho \), 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 \( g(p_\rho) \). Both \( \rho = 0.0005 \) and \( \rho = 0.005 \) have greater \( g(p_\rho) \) 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 seconds 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 8, 4, and 4 pilot subcarriers for contiguous subcarrier block \{1, 2, ..., 256\}, \{513, 514, ..., 640\} and \{897, 898, ..., 1024\}, respectively.

We now evaluate the channel estimation performance using the designed pilot patterns. To ease the notation on 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 illustrated in Figure 3 and Figure 4, respectively. Both of MSE and BER are averaged over 10000 sparse channel realizations. The popular OMP algorithm is employed for sparse channel estimation given the pilot pattern PT1, PT2, ..., PT7. For comparisons, the performance of LS channel estimation using \( K = 64 \) and \( K = 75 \) equally spaced pilots,

\(^2\)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 \( \rho \). 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 4GB of memory.

\(^3\)In general, the exhaustive search is a brute-force method which exhaustively examines all possibilities. However, since the generation of all \( 2^{512} = 8.41 \times 10^{159} \) candidates is almost impossible, we define the exhaustive search with consistency to most literatures [13].

<table>
<thead>
<tr>
<th>Type</th>
<th>( g(p_\rho) )</th>
<th>( p_\rho )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho = 0.0005 )</td>
<td>5.5220</td>
<td>76, 91, 144, 171, 202, 225, 249, 514</td>
</tr>
<tr>
<td>( \rho = 0.001 )</td>
<td>5.2164</td>
<td>47, 95, 110, 162, 180, 193, 246, 513</td>
</tr>
<tr>
<td>( \rho = 0.005 )</td>
<td>5.6151</td>
<td>524, 627, 640, 897, 910, 939, 976, 1019</td>
</tr>
<tr>
<td>( \rho = 0.01 )</td>
<td>5.7806</td>
<td>77, 101, 171, 217, 237, 248, 514</td>
</tr>
<tr>
<td>( \rho = 0.05 )</td>
<td>6.3118</td>
<td>525, 627, 640, 943, 970, 996, 1007, 1024</td>
</tr>
<tr>
<td>Exhaustive search</td>
<td>6.7889</td>
<td>65, 97, 166, 179, 191, 202, 228, 513</td>
</tr>
<tr>
<td>Equally spaced pilots</td>
<td>10.9241</td>
<td>1, 37, 73, 109, 145, 181, 217, 253, 513, 555, 597, 639, 897, 939, 981, 1023</td>
</tr>
</tbody>
</table>

TABLE I
Comparisons of pilot patterns using different pilot design schemes and parameters.
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 compared to the exhaustive search. With channel encoding and decoding, BER around 0.01 can be completely removed in practical systems.

V. CONCLUSIONS

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 indices 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 pilot subcarrier have been derived. A projection method has been proposed so that the number of pilot subcarriers during the 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 to the LS channel estimation.

REFERENCES