A Study of Deterministic Pilot Allocation for Sparse Channel Estimation in OFDM Systems

Chenhao Qi, *Member, IEEE,* and Lenan Wu

***Abstract*—In this letter, we investigate the deterministic pilot allocation for sparse channel estimation in OFDM systems. Based on the rule of minimizing the coherence of the DFT submatrix, we derive that the pilot design according to the cyclic different set (CDS) is optimal. However, the CDS only exists for some specific number of OFDM subcarriers. For those cases where the CDS is unavailable, we propose a scheme using discrete stochastic approximation to obtain a near-optimal pilot pattern. Simulation results demonstrate that our scheme is much faster convergent and more efficient than the exhaustive search; and it has been shown that substantial improvement for channel estimation can be achieved.**

***Index Terms*—Pilot allocation, channel estimation, discrete stochastic approximation, compressed sensing.**

I. Introduction

OFDM transforms the frequency-selective wireless channel into several parallel flat-fading narrowband subchannels. Each subchannel only needs a single-tap equalizer, and therefore the high complexity associate with the long equalizer to combat inter-symbol interference (ISI) is mitigated. Nevertheless, this approach hinges on accurate channel estimation. Recent developments in compressed sensing (CS) has motivated the extensive research on the application of sparse recovery algorithms to channel estimation, which needs less pilots and demonstrates to be more accurate than the standard least squares (LS) [1]. Many CS algorithms including matching pursuit (MP), orthogonal matching pursuit (OMP) and basis pursuit (BP) have been employed for pilot-assisted channel estimation in OFDM systems [2]. But few works discuss the OFDM pilot allocation for sparse channel estimation.

The well-known restricted isometry property (RIP) indicates that the measurement using random matrices guarantees the sparse recovery with high probability [3], which implies the randomly-generated pilot pattern is theoretically optimal. However, it has drawbacks of high complexity, large storage and low efficiency in real applications. In [4], a deterministic pilot selection scheme is proposed for sparse channel estimation using Dantzig selector. In [5], a clustered pilot design is presented in the study of underwater acoustic (UWA) channel estimation. And a scheme using channel data to offline train the pilots and search the optimized pilot placements at the transmitter is proposed in [6].

Manuscript received December 17, 2011. The associate editor coordinating the review of this letter and approving it for publication was S. Kotsopoulos.

The authors are with the School of Information Science and Engineering, Southeast University, Nanjing 210096, China (e-mail: qch@seu.edu.cn).

The work is supported by the National Natural Science Foundation of China (NSFC) under the Grant 60872075, the Scientific Research Foundation of Southeast University under the Grant Seucx201116, Huawei Innovative Research Plan and CAST8804009011.

Digital Object Identifier 10.1109/LCOMM.2012.032612.112553

In this letter, we investigate the deterministic pilot allocation for sparse channel estimation in OFDM systems. Based on the rule of minimizing the coherence of the DFT submatrix, we derive that the pilot design according to the cyclic different set (CDS) is optimal. For those cases where the CDS is unavailable, we propose a scheme using discrete stochastic approximation to obtain a near-optimal pilot pattern via offline training.

Notationwise, symbols for matrices (upper case) and vectors (lower case) are in boldface. $(\cdot)^T$, $(\cdot)^H$, $diag\{\cdot\}$, $\boldsymbol{I}_L$, $\mathbb{R}^M$ and $\mathcal{CN}$ denote transpose, conjugate transpose (Hermitian), the diagonal matrix, the identity matrix with dimension L, the set of real-valued vector with dimension M and the complex Gaussian distribution, respectively. $A(l)$ denotes the l-th column of $\boldsymbol{A}$.

II. System Model

We treat the comb-type pilot assisted channel estimation in OFDM systems. Assume the number of OFDM subcarriers to be N, we use N_p pilot subcarriers as $k_1, k_2, \ldots, k_{N_p}$ ($1 \le k_1 < k_2 < \ldots < k_{N_p} \le N$) for frequency-domain channel estimation. The transmit pilot symbols and the receive pilot symbols are denoted as $X(k_1), X(k_2), \ldots, X(k_{N_p})$ and $Y(k_1), Y(k_2), \ldots, Y(k_{N_p})$, respectively. Then the problem is formulated as

$$\begin{bmatrix} Y(k_1) \\ Y(k_2) \\ \vdots \\ Y(k_{N_p}) \end{bmatrix} = \begin{bmatrix} X(k_1) & 0 & 0 & 0 \\ 0 & X(k_2) & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & X(k_{N_p}) \end{bmatrix} \cdot \boldsymbol{F}_{N_p \times L} \cdot \begin{bmatrix} h(1) \\ h(2) \\ \vdots \\ h(L) \end{bmatrix} + \begin{bmatrix} \eta(1) \\ \eta(2) \\ \vdots \\ \eta(N_p) \end{bmatrix} \quad (1)$$

where $\boldsymbol{h} = [h(1), h(2), \cdots, h(L)]^{\mathrm{T}}$ is the equivalent sampled channel impulse response (CIR) with length L, $\boldsymbol{\eta} = [\eta(1), \eta(2), \cdots, \eta(N_p)]^{\mathrm{T}} \sim \mathcal{CN}(\boldsymbol{0}, \sigma_\eta^2 \boldsymbol{I}_{N_p})$ is an additive white Gaussian noise (AWGN) term, and

$$\boldsymbol{F}_{N_p \times L} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & \omega^{k_1} & \cdots & \omega^{k_1 \cdot (L-1)} \\ 1 & \omega^{k_2} & \cdots & \omega^{k_2 \cdot (L-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{k_{N_p}} & \cdots & \omega^{k_{N_p} \cdot (L-1)} \end{bmatrix}$$

is a DFT submatrix where $\omega = e^{-j2\pi/N}$. We denote

$$\begin{aligned} \boldsymbol{X} &= diag\{X(k_1), X(k_2), \cdots, X(k_{N_p})\} \\ \boldsymbol{y} &= [Y(k_1), Y(k_2), \cdots, Y(k_{N_p})]^{\mathrm{T}} \end{aligned}$$

and

$$\boldsymbol{A} = \boldsymbol{X}\boldsymbol{F}_{N_p \times L} \tag{2}$$

then (1) is reformulated as

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{h} + \boldsymbol{\eta} \tag{3}$$

Since the system sampling interval is usually much smaller than the channel delay spread, most components of $\boldsymbol{h}$ are either zero or nearly zero, which means that $\boldsymbol{h}$ is a sparse vector. In this case, we can improve the data rate by using less pilots than the unknown channel coefficients, i.e., $N_p < L$, where CS algorithms are employed to reconstruct $\boldsymbol{h}$ instead of the standard LS channel estimation. It's also noticed that L is usually no larger than the length of OFDM guard interval, i.e., $N/4$.

III. Pilot Allocation

Recent advances in CS show that under noiseless condition $\boldsymbol{h}$ can be reconstructed from the measurement $\boldsymbol{y}$ with high probability when dictionary matrix $\boldsymbol{A}$ satisfies RIP [3]. However, there is no known method to test in polynomial time whether a given matrix satisfies RIP. An alternative approach we adopt here is to minimize the coherence of $\boldsymbol{A}$ [7]. We define the coherence to be the maximum absolute correlation between two different columns, denoted as

$$\begin{aligned} g(\boldsymbol{p}) &= \max_{0 \le m < n \le L-1} \left| \langle A(m), A(n) \rangle \right| \\ &= \max_{0 \le m < n \le L-1} \left| \sum_{i=1}^{N_p} \left| X(k_i) \right|^2 \omega^{k_i(n-m)} \right| \end{aligned} \tag{4}$$

and therefore the objective function is

$$Q = \min_{\boldsymbol{p}} g(\boldsymbol{p}) \tag{5}$$

with respect to the pilot allocation $\boldsymbol{p} = [k_1, k_2, \ldots, k_{N_p}]$ because different $\boldsymbol{p}$ results in different $\boldsymbol{A}$ according to (2). The optimal pilot pattern is

$$\boldsymbol{p}_{opt} = \arg\min_{\boldsymbol{p}} g(\boldsymbol{p}) \tag{6}$$

Suppose all OFDM pilot symbols are equipowered to be

$$E = \left| X(k_i) \right|^2, \ i = 1, 2, \ldots, N_p \tag{7}$$

and denote $c = n - m$, (4) is simplified as

$$g(\boldsymbol{p}) = E \cdot \max_{1 \le c \le L-1} \left| \sum_{i=1}^{N_p} \omega^{k_i c} \right| \tag{8}$$

We define

$$f(c) = \left| \sum_{i=1}^{N_p} \omega^{k_i c} \right|^2 = \sum_{i=1}^{N_p} \sum_{l=1}^{N_p} \omega^{c(k_i - k_l)} \tag{9}$$

For all possible

$$d = (k_i - k_l) \bmod N, \ i \ne l \tag{10}$$

let a_d denote the number of occurrence of d ($d = 1, 2, \ldots, N-1$), then

$$f(c) = N_p + \sum_{d=1}^{N-1} a_d \cdot \omega^{cd} \tag{11}$$

Since $\omega^d \ne 1$, we have

$$\sum_{c=1}^{L-1} f(c) = N_p(L-1) + \sum_{d=1}^{N-1} a_d \frac{\omega^d - \omega^{Ld}}{1 - \omega^d} \tag{12}$$

Therefore

$$\max_{1 \le c \le L-1} f(c) \ge N_p + \frac{1}{L-1} \sum_{d=1}^{N-1} a_d \frac{\omega^d - \omega^{Ld}}{1 - \omega^d} \tag{13}$$

and the equality holds only with

$$f(1) = f(2) = \cdots = f(L-1) = N_p + \frac{1}{L-1} \sum_{d=1}^{N-1} a_d \frac{\omega^d - \omega^{Ld}}{1 - \omega^d} \tag{14}$$

which means

$$a_1 = a_2 = \cdots = a_{N-1} = \frac{\sum_{d=1}^{N-1} a_d}{N-1} = \frac{N_p(N_p - 1)}{N-1} \tag{15}$$

We plug (15) into (11) and get

$$Q = \sqrt{f(1)} = \sqrt{f(2)} = \cdots = \sqrt{f(L-1)} = \sqrt{\frac{N_p(N - N_p)}{N-1}} \tag{16}$$

It's known that the CDS satisfies the above Welch bound [8] and therefore minimizes the coherence of $\boldsymbol{A}$. So the pilot allocation according to the CDS is optimal. However, the CDS only exists for some specific N. For those N where the CDS is unavailable, we propose a scheme using discrete stochastic approximation to search for a near-optimal pilot allocation. The procedures for implementation are summarized in Algorithm 1. It's a modification of the algorithm in [6].

Algorithm 1 - Pilot Allocation

Randomly generate a $\boldsymbol{p}_0$, $\hat{\boldsymbol{p}}_0 \Leftarrow \boldsymbol{p}_0$.
$\boldsymbol{\pi}[0] \Leftarrow \boldsymbol{0}_{N_x}$, $\pi[0,0] \Leftarrow 1$, $u \Leftarrow 0$, $v \Leftarrow 0$.
for $n = 0, 1, \ldots, M-1$
  for $k = 0, 1, \ldots, N_p - 1$
    $m \Leftarrow n * N_p + k$.
    generate $\tilde{\boldsymbol{p}}_m \backslash \boldsymbol{p}_m$.
    if $g(\tilde{\boldsymbol{p}}_m) < g(\boldsymbol{p}_m)$
      $\boldsymbol{p}_{m+1} \Leftarrow \tilde{\boldsymbol{p}}_m$, $u \Leftarrow m+1$.
    else
      $\boldsymbol{p}_{m+1} \Leftarrow \boldsymbol{p}_m$.
    end if
    $\boldsymbol{\pi}[m+1] \Leftarrow \boldsymbol{\pi}[m] + (\boldsymbol{r}[m+1] - \boldsymbol{\pi}[m])/(m+1)$.
    if $\pi[m+1, u] > \pi[m+1, v]$
      $\hat{\boldsymbol{p}}_{m+1} \Leftarrow \boldsymbol{p}_{m+1}$, $v \Leftarrow u$.
    else
      $\hat{\boldsymbol{p}}_{m+1} \Leftarrow \hat{\boldsymbol{p}}_m$.
    end if
  end for(k)
end for(n)

We define $\boldsymbol{p}_m$, $\hat{\boldsymbol{p}}_m$ and $\tilde{\boldsymbol{p}}_m$ ($m = 0, 1, 2, \ldots$) as different pilot patterns, i.e., $\boldsymbol{p}_m = [k_1, k_2, \ldots, k_{N_p}]$. We denote $N_x = MN_p$ and define $\boldsymbol{\pi}[m] \in \mathbb{R}^{N_x}$ ($m = 0, 1, 2, \ldots$) as an occupation-probability vector which indicates an estimate of the occupation probability of one pilot pattern. The ith component of $\boldsymbol{\pi}[m]$ is denoted as $\boldsymbol{\pi}[m, i]$. $\boldsymbol{r}[m+1] \in R^{N_x}$ is defined as a zero vector except for its $(m+1)$th component to be 1.

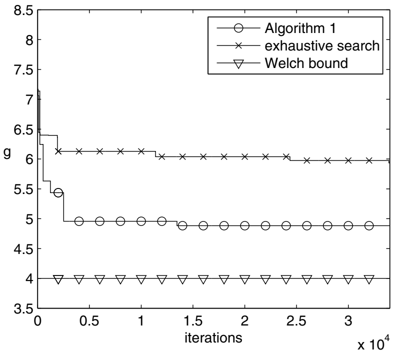

Fig. 1. Comparisons of Algorithm 1 with the exhaustive search and the Welch bound for N = 273.

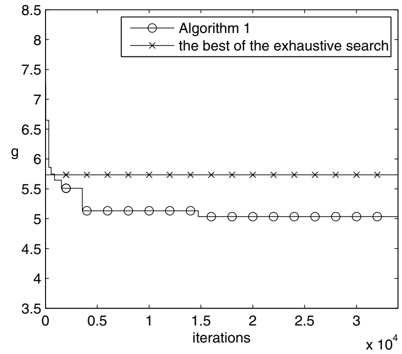

Fig. 2. Comparisons of Algorithm 1 with the best of the exhaustive search for N = 256.

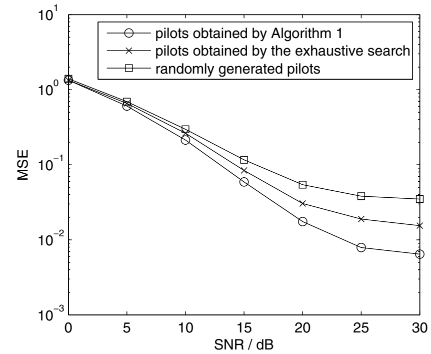

Fig. 3. Comparisons of channel estimation for different pilot allocation schemes.

Starting from a random $\boldsymbol{p}_0$, we initialize $\hat{\boldsymbol{p}}_0$ and generate a zero $\boldsymbol{\pi}[0]$ where $\boldsymbol{\pi}[0, 0]$ is then set to be 1. In each iteration, we obtain a new $\tilde{\boldsymbol{p}}_m$ by sequentially changing one pilot position of $\boldsymbol{p}_m$. We compare $g(\tilde{\boldsymbol{p}}_m)$ with $g(\boldsymbol{p}_m)$ and move a small step. The algorithm will converge to the optimal pilot pattern which has the largest occupation probability [6]. The main difference with [6] is that we minimize the coherence of the measurement matrix $\boldsymbol{A}$ rather than using channel data to minimize the mean square error (MSE) of channel estimate, which indicates its applicability to the transmitter without *a-prior* knowledge of the channel, and leads to the simplicity in updating the occupation probability. In practice, the calculation of $g(\boldsymbol{p})$ is equivalent to find the largest off-diagonal upper-triangular component of $\boldsymbol{A}^H\boldsymbol{A}$. Moreover, Algorithm 1 also works for different-powered pilot symbols in contrast to (7). The final choice of $\hat{\boldsymbol{p}}_{m+1}$ is a near-optimal pilot pattern and believed to converge to the global optimum that is the best of all possible $\binom{N}{N_p}$ pilot patterns.

IV. Simulation Results

As shown in Figure 1, we compare Algorithm 1 with the exhaustive search and the Welch bound for $N = 273$ and $N_p = 17$, where the CDS exists and achieves the Welch low bound $Q = 4$ according to (16). Algorithm 1 is demonstrated to converge much faster than the exhaustive search that exhaustively searches for the best pilots from all possible $\binom{N}{N_p}$ pilot patterns in terms of (5). However, it's difficult to implement fast FFT operations for $N = 273$ OFDM subcarriers. In practice, the number of subcarriers is usually designed to be 64, 256, 512 or 1024, where the CDS does not exist. So in Figure 2, we compare Algorithm 1 with the exhaustive search for $N = 256$. It's observed that Algorithm 1 exceeds the best of 32000 exhaustive search in no more than 900 iterations, and proves much faster convergent and more efficient. So for those cases CDS unavailable, we can use Algorithm 1 to offline search a near-optimal pilot pattern.

Comparisons of sparse channel estimation for different pilot allocation are illustrated in Figure 3. A sparse multipath channel is generated as a zero CIR vector $\boldsymbol{h}$ with $L = 50$, where $S = 5$ positions are randomly selected as nonzero channel taps. The attenuation of each taps satisfies the independent and identically distributed (i.i.d.) $\mathcal{CN}(0, 1)$. The system is designed that $N_p = 16$ pilots of the total $N = 256$ OFDM subcarriers are employed for frequency-domain channel estimation using OMP. As shown in Figure 3, we compare the MSE of channel estimation for different pilot allocation schemes. Algorithm 1 outperforms the exhaustive search and randomly generated pilots, which verifies the effectiveness of our proposed scheme.

Acknowledgment

The authors would like to thank Prof. Xiaodong Wang of Columbia University for his helpful suggestions.

V. Conclusion

Based on the rule of minimizing the coherence of the DFT submatrix, we have derived that the pilot design according to CDS is optimal. For those cases where the CDS is unavailable, we have proposed a scheme using discrete stochastic approximation to obtain a near-optimal pilot pattern via offline training. Simulation results have validated the effectiveness of the proposed scheme, which is demonstrated to be much faster convergent and more efficient than the exhaustive search.

References

[1] 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.

[2] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: a new approach to estimating sparse multipath channels," *Proc. IEEE*, vol. 98, no. 6, pp. 1058–1076, June 2010.

[3] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[4] L. Applebaum, W. U. Bajwa, A. R. Calderbank, J. Haupt, and R. Nowak, "Deterministic pilot sequences for sparse channel estimation in OFDM systems," in *Proc. 2011 International Conference on Digital Signal Processing*, pp. 1–7.

[5] C. R. Berger, J. Gomes, and J. M. F. Moura, "Study of pilot designs for cyclic-prefix OFDM on time-varying and sparse underwater acoustic channels," in *Proc. 2011 OCEANS*, pp. 1–8.

[6] C. Qi and L. Wu, "Optimized pilot placement for sparse channel estimation in OFDM systems," *IEEE Signal Process. Lett.*, vol. 18, no. 12, pp. 749–752, Dec. 2011.

[7] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[8] P. Xia, S. Zhou, and G. B. Giannakis, "Achieving the Welch bound with difference sets," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1900–1907, May 2005.