Optimized Pilot Placement for Sparse Channel Estimation in OFDM Systems

Chenhao Qi, Member; IEEE, and Lenan Wu

Abstract—Compressed sensing (CS) has recently been applied for pilot-aided sparse channel estimation. However, the design of the pilot placement has not been considered. In this letter, we propose a scheme using the modified discrete stochastic approximation to optimize the pilot placement in OFDM systems. The channel data is employed to offline search the near-optimal pilot placement before the transmission. Meanwhile we also get a criterion to select CS algorithms based on the mean squared error (MSE) minimization. Simulations using a sparse wireless channel model have validated the effectiveness of the proposed scheme, which is demonstrated to be much faster convergent and more efficient than the exhaustive search. It has been shown that substantial performance improvement can be achieved for OMP and YALL1 based channel estimation, where YALL1 is preferred.

Index Terms—Channel estimation, compressed sensing, discrete stochastic approximation, pilot placement.

I. INTRODUCTION

collection of sparse recovery algorithms has recently emerged with the name compressed sensing (CS) [1], which enables efficient reconstruction of sparse signals from relatively few linear measurements. And more recently, CS techniques have been applied for sparse channel estimation [2]. Sparse recovery algorithms including matching pursuit (MP), orthogonal matching pursuit (OMP) and basis pursuit (BP) are employed for the pilot-assisted channel estimation rather than using traditional least squares (LS) methods [3]. Furthermore, convex optimization solvers as SpaRSA [4], SPGL1, YALL1, GPSR and ℓ_1 -LS [5] can also be applied. However, few works concern on the design of the pilot placement. Although the best pilot placement for the LS channel estimation in OFDM systems is equipowered, equispaced and phase shift orthogonal [6], there is no general theory on the optimized pilot placement for the channel estimation using sparse recovery algorithms.

In this letter, we investigate the pilot optimization for sparse channel estimation in OFDM systems and propose a scheme using modified discrete stochastic approximation to offline

Manuscript received August 13, 2011; revised September 24, 2011; accepted September 26, 2011. Date publication October 13, 2011. Date of current version November 03, 2011. The work was supported by the National Natural Science Foundation of China (NSFC) under the Grant 60872075 and the Scientific Research Foundation of Southeast University under the Grant Seucx201116. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Tan F. Wong.

C. Qi is with the School of Information Science and Engineering, Southeast University, Nanjing 210096, China, and also with the Key Laboratory of Underwater Acoustic Signal Processing of Ministry of Education, Nanjing, China (e-mail: qch@seu.edu.cn).

L. Wu is with the School of Information Science and Engineering, Southeast University, Nanjing 210096, China (e-mail: wuln@seu.edu.cn).

Digital Object Identifier 10.1109/LSP.2011.2170834

search the pilot placements before transmission. Since well-established channel models are usually available in wireless communications, it's beneficial to take advantage of them for pilot optimization at the transmitter. Even for underwater acoustic communications [7], a collection of sampled real channel data of specific sea area is accessible. Besides, our scheme also presents a criterion to select CS algorithms based on the mean-squared error (MSE) minimization.

The remainder of the letter is organized as follows. Section II briefly describes the model of OFDM frequency-domain channel estimation and formulates as a sparse recovery problem. Section III proposes a scheme using modified discrete stochastic approximation to optimize the pilot placement. Simulation results are given in Section IV. Finally, Section V concludes the letter.

The notation used in this letter is according to the convention. Symbols for matrices (upper case) and vectors (lower case) are in boldface. $(\cdot)^T$, $(\cdot)^H$, $\|\cdot\|_1 \|\cdot\|_2$, \mathbb{R} , \mathbb{Z} , $diag\{\cdot\}$, I_L , $\mathbf{0}_{M \times N}$ and \mathcal{CN} denote transpose, conjugate transpose (Hermitian), ℓ_1 -norm, ℓ_2 -norm, the set of real number, the set of integers, the diagonal matrix, the identity matrix with dimension L, the M by N zero matrix and the complex Gaussian distribution, respectively. $\mathcal{O}(\cdot)$ means the order. $\hat{\phi}$ denotes the estimate of the parameter of interest ϕ .

II. SYSTEM MODEL

Considering an OFDM system with N subcarriers, among which N_p subcarriers are selected as pilots with positions represented by $k_1, k_2, \ldots, k_{N_p}$ $(1 \le k_1 < k_2 < \cdots < k_{N_p} \le N)$ and N_d $(N_d = N - N_p)$ subcarriers are used for data transfer, we denote the transmit pilot symbols and the receive pilot symbols 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 OFDM frequency-domain channel estimation is formulated as [8]

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{h} + \boldsymbol{\eta} \tag{1}$$

where $\mathbf{A} = \mathbf{X} \mathbf{F}_{N_p \times L}$, $\mathbf{X} = diag \{X(k_1), X(k_2), \dots, X(k_{N_p})\}$, $\mathbf{F}_{N_p \times L}$ is a submatrix selected by the row indices $[k_1, k_2, \dots, k_{N_p}]$ and column indices $[0, 1, \dots, L - 1]$ from the standard $N \times N$ Fourier matrix, $\mathbf{y} = [Y(k_1), Y(k_2), \dots, Y(k_{N_p})]^T$, $\mathbf{h} = [h(0), h(1), \dots, h(L - 1)]^T$ is the equivalent sampled channel impulse response (CIR) with length L, and $\boldsymbol{\eta} = [\eta(0), \eta(1), \dots, \eta(N_p - 1)]^T \sim \mathcal{CN}(\mathbf{0}, \sigma_{\eta}^2 \mathbf{I}_{N_p})$ is the noise term. Since the system sampling interval is usually much smaller compared to the channel delay spread, most channel coefficients are either zero or nearly zero, which means that \mathbf{h} is a sparse vector. If \mathbf{A} has more rows than columns, i.e., $N_p > L$, then (1) is a standard LS problem with the estimated CIR

$$\hat{\boldsymbol{h}} = (\boldsymbol{A}^H \boldsymbol{A})^{-1} \boldsymbol{A}^H \boldsymbol{y}.$$
(2)

Nevertheless, we are more interested in the sparse case with $N_p < L$, i.e., the number of pilots is less than the number of channel coefficients. Considering that $\mathbf{h} \in \mathbb{R}^L$ has S nonzero components, with $S \ll L$, we can recover \mathbf{h} from (1) by solving the following ℓ_0 -norm minimization problem

$$\min_{\boldsymbol{h}} \|\boldsymbol{h}\|_0 \text{ s.t. } \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{h}\|_2 \le \sigma_{\eta}$$
(3)

where $\|\boldsymbol{h}\|_0$ counts the number of nonzero elements of \boldsymbol{h} . This is an NP-hard combinatorial problem. However, it can be replaced by the following ℓ_1 -norm optimization problem

$$\min_{\boldsymbol{h}} \|\boldsymbol{h}\|_1 \text{ s.t. } \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{h}\|_2 \le \sigma_{\eta}$$
(4)

which can be solved by quadratic programming (QP) and some other convex optimization algorithms. Recent advances in CS show that under noiseless condition h can be recovered from the measurement y with high accuracy when dictionary matrix Asatisfies the restricted isometry property (RIP) [9]. It's also observed that the pilot design can be regarded as the measurement design in sparse recovery. However, so far there is no known method to test in polynomial time whether a given matrix satisfies RIP.

Considering the fact that the channel data is usually available at the transmitter side, intuitively we can use the channel data to offline train the pilots and exhaustively search the best pilot from all possible pilot placements. It makes no change to the sparse channel estimation at the receiver side. However, such an exhaustive search method is computationally prohibitive.

III. PILOT OPTIMIZATION

Suppose selecting N_p pilot subcarriers from a total of N subcarriers, there are $\binom{N}{N_p}$ possible pilot placements. For example, if $N_p = 12$ and N = 256, we have $\binom{256}{12} = 1.27 \times 10^{21}$ possible pilot placements. It's computationally prohibitive to exhaustively search all the pilot placements for the best one. Here we propose a scheme using modified discrete stochastic approximation to sequentially search a near-optimal pilot placement, which typically exhibits fast convergence [10]. The basic idea is to generate a sequence of the estimates of the optimal pilot subset where the new estimate is based on the previous one by moving a small step in a good direction towards the global optimizer. Through the iterations, the global optimizer can be found by means of maintaining an occupation probability vector π which indicates an estimate of the occupation probability of one state (i.e., pilot subset). Under certain conditions, the algorithm converges to the state which has the largest occupation probability in π . Compared with the exhaustive search, more computations of this algorithm are performed on the promising candidates and the better candidates will be evaluated more than the others.

We define the normalized MSE corresponding to a pilot placement p and channel vector h for a specific channel estimation algorithm alg (i.e., OMP, YALL1) as

$$f(\boldsymbol{p}, \boldsymbol{h}, alg) = \frac{E\left\{\|\boldsymbol{h} - \hat{\boldsymbol{h}}\|_{2}^{2}\right\}}{\|\boldsymbol{h}\|_{2}^{2}}$$
(5)

where the expectation is with respect to the channel noise, $\boldsymbol{p} = \{k_1, k_2, \ldots, k_{N_p}\}$, and $\hat{\boldsymbol{h}}$ denotes the estimated channel using *alg*. Once \boldsymbol{p} is given, we can produce \boldsymbol{A} and \boldsymbol{y} . Then we get an $\hat{\boldsymbol{h}}$ from \boldsymbol{y} and \boldsymbol{A} through *alg*. So the averaged MSE over all possible channel data available at the transmitter is

$$g(\boldsymbol{p}, alg) = E_{\boldsymbol{h}} \left\{ f(\boldsymbol{p}, \boldsymbol{h}, alg) \right\}$$
(6)

where we approximate this ensemble average over h by the sample average, which means to average over all available channel realizations of h. Our objective is to minimize the averaged MSE of channel estimation for different CS algorithms. The near-optimal pilot placement corresponding to alg is

$$\min_{\boldsymbol{p}} g(\boldsymbol{p}, alg). \tag{7}$$

Meanwhile we also get a criterion as

$$\arg\min_{\text{alg},\boldsymbol{p}} g(\boldsymbol{p}, alg) \tag{8}$$

to select CS algorithms. After the same number of iterations, the CS algorithm with the minimum MSE is selected.

The procedure to optimize the pilot placement is summarized in Algorithm 1. Let \boldsymbol{p}_m , $\hat{\boldsymbol{p}}_m$ and $\tilde{\boldsymbol{p}}_m$ denote pilot placements at the *m*th iteration and suppose the number of all possible pilot placements to be $N_x = \binom{N}{N_p}$. The auxiliary matrix $\boldsymbol{P} \in \mathbb{Z}^{N_x \times N_p}$ is constructed with each row

$$\boldsymbol{P}[i] = \{k_1, k_2, \dots, k_{N_p}\}, \ i \in \{1, 2, \dots, N_x\}$$
(9)

to store a different pilot placement. The state occupation probability at the mth iteration is represented by

$$\boldsymbol{\pi}[m] = \left\{ \left[\pi[m, 1], \pi[m, 2], \dots, \pi[m, N_x] \right]^T \in \mathbb{R}^{N_x} \mid \sum_i \pi[m, i] = 1, \pi[m, i] \in [0, 1], i \in \{1, 2, \dots, N_x\} \right\}.$$
 (10)

Algorithm 1 is divided into five steps. We start from a random pilot placement p_0 and initialize both P and $\pi[0]$ to be zero, as described in Step 0. Then given p_m and $g(p_m)$ from the last iteration, we obtain a new pilot placement \tilde{p}_m by sequentially changing one pilot position of p_m . Unlike randomly generating \tilde{p}_m in aggressive and conservative discrete stochastic approximation [11], every time we only change one pilot position of previously generated pilot placement, which reduces the computational complexity from $\mathcal{O}(2^{N_p}M)$ to $\mathcal{O}(MN_p)$ [12]. We run *alg* for all available channel data to obtain a new averaged MSE $g(\tilde{p}_m)$ and compare it with $g(p_m)$. We make a choice for the next move towards the global optimizer. At each iteration, the frequency of the selected pilot placement is updated. We update $\pi[m + 1]$ by

$$\pi[m+1] = \pi[m] + \frac{r[q_m] - \pi[m]}{(m+1)}$$
(11)

where $\mathbf{r}[q_m] \in \mathbb{R}^{N_x}$ is defined as a zero vector except for its q_m th element to be 1. The final choice $\tilde{\mathbf{p}}_m$ is the optimized pilot placement and believed to converge to the global optimum [11]. Therefore we can perform the same procedure for different *alg* (i.e., OMP, YALL1), from which we choose the CS algorithm with the smallest $g(\tilde{\mathbf{p}}_{N_x})$.

Algorithm 1 - Pilot Placement Optimization

Step 0: Initialization randomly generate a pilot placement p_0 ; set $\tilde{p}_0 = p_0$, $\hat{p}_0 = p_0$; set $P = \mathbf{0}_{N_x \times N_p}$, $P[1] = p_0$; set $\pi[0] = \mathbf{0}_{N_x}$, $\pi[0, 0] = 1$; set u = 0, v = 0.

FOR $n = 0, 1, \dots, M-1$ FOR $k = 0, 1, \dots, N_p - 1$

Step 1: Sampling and evaluation $m = n * N_p + k$; given p_m and $g(p_m)$ from the last iteration, generate another $\tilde{p}_m \setminus p_m$ uniformly, where $\tilde{p}_m = p_m$ except for the *k*-th pilot position; calculate $g(\tilde{p}_m)$.

Step 2: Acceptance

if $g(\tilde{p}_m) < g(p_m)$ $p_{m+1} = \tilde{p}_m, u = m + 1$ else $p_{m+1} = p_m$ end if

```
Step 3: Updating state occupation probabilities searching p_{m+1} in P;
```

if found q_m = the found row index in Pelse $q_m = m + 1$, store p_{m+1} into P[m + 1]end if $\pi[m+1] = \pi[m] + (r[q_m] - \pi[m])/(m+1)$

Step 4: Selection if $\pi[m+1, u] > \pi[m+1, v]$ $\hat{p}_{m+1} = p_{m+1}, v = u$ else $\hat{p}_{m+1} = \hat{p}_m$ end if

EndFOR (k) EndFOR (n)

In theory Algorithm 1 requires a large memory size such as $\binom{256}{12} = 1.27 \times 10^{21}$ floating-point units to store the occupancy probabilities of all possible pilot placements. However in the actual implementation, the memory requirement is much smaller. Noticing that most pilot placement will not be selected, we can

TABLE I PARAMETERS OF THE SIMULATION

Number of total subcarriers	<i>N</i> = 256
Number of pilot subcarriers	$N_{p} = 12$
Number of channel realizations	$N_r = 2000$
Number of multipaths	S = 5
Length of CIR	<i>L</i> = 50
Modulation	QPSK

shrink the row dimension of P and $\pi[m]$ to $N_x = MN_p$. Instead of keeping records for all candidates, we dynamically allocate and maintain record in P for the new pilot placement found in each iteration. If a pilot placement already has a record in P, the corresponding occupancy probability will be updated. Otherwise, a new element is appended in P with its index and occupation probability. Such a dynamic scheme avoids the huge memory requirement since typically in practice only a small fraction of all possible subsets is visited. Therefore we only have to store each selected pilot placement in P. Once entering Step 3, we first search in P to check whether p_{m+1} exists. If so, we store the found row index of P into q_m . Otherwise, we store p_{m+1} into P[m+1] and set $q_m = m + 1$. In this way we guarantee the pilot placement in P is unique.

IV. SIMULATION RESULTS

We consider the following sparse multipath channel model

$$c(\tau) = \sum_{i=1}^{S} a_i \delta(\tau - \tau_i)$$
(12)

to obtain a set of channel data at the transmitter. $\{a_i\}$ is a set of independent and identically distributed (i.i.d.) random variables which satisfy $a_i \sim C\mathcal{N}(0, e^{-b\tau_i})$. b = 1/16 is the exponential power delay profile and τ_i is the delay spread for the *i*th path [13]. The parameters used in our simulations are listed in Table I. A zero CIR vector with the length L = 50 is first generated, where S = 5 positions are randomly selected as channel taps. Then we produce $\{a_i\}$ as the attenuation for each path. $N_r = 2000$ channel realizations with the sparsity S/L = 10%are generated. Considering the oversampling of most OFDM systems, the actual sparity is even smaller [14]. In practice, we may replace the above simulated channel realizations with the real channel data, which cannot be infinite since it's collected by countable field measurements.

We run Algorithm 1 for M = 200 iterations to optimize the pilot placement for OMP and YALL1 [15], respectively. The stop condition is

$$\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{h}\|_2 \le \sigma_\eta \tag{13}$$

according to (4). At each iteration, we generate pilot placements by one-by-one sequentially changing the pilot position, which corresponds to the exhaustive search for totally $N_x = 2400$ iterations. As shown in Fig. 1, we compare the MSE of OMP and YALL1 using Algorithm 1 and the exhaustive search, respectively. Algorithm 1 for OMP exceeds the best of 2400 exhaustive search by no more than 48 iterations, which proves much faster convergent and more efficient. It's demonstrated in [10] that with the growing number of iterations, it will converge to the global optimum. Considering the computational complexity of Algorithm 1, it's roughly $\mathcal{O}(MN_r 2^{N_p})$ [12]. And since we



Fig. 1. Convergence of the pilot placement optimization using Algorithm 1.



Fig. 2. Performance improvement with the optimized pilot placement.

always offline train the pilot placement before the transmission, it won't bring any additional complexity to the channel estimation at the receiver. According to Fig. 1 and the criterion (8), YALL1 is preferred since its MSE is lower than OMP at the stop point.

In Fig. 2, we compare the MSE of random pilot and optimized pilot for both OMP and YALL1. We run 1000 iterations for each SNR point, where 1000 OFDM symbols are used for each iteration. The improvement with the pilot optimization is more than one order of magnitude at high SNR conditions, i.e., $SNR = 35 \, dB$. In particular, OMP with the optimized pilot even outperforms YALL1 with the random pilot. The reason is that the MSE is averaged over all channel data at the transmitter, so we can not guarantee the sparse recovery to be always successful for the random pilot placement. While for the optimized pilot placement, the sparse recovery is improved since Algorithm 1 has already eliminated the cases that the CS algorithm fails or obtains large MSE. Therefore, the pilot optimization at the transmitter is beneficial, especially for the heuristic CS algorithms. Considering the fact that the best pilot placement for LS channel estimation is equispaced [6], we also simulate the cubic-spline-interpolation based LS channel estimation with $N_p = 64$. However, the spectral efficiency has reduced from 95.3% to 75%.

V. CONCLUSION

In this letter, we have investigated the pilot placement for the sparse channel estimation in OFDM systems. We have proposed a scheme using modified discrete stochastic approximation to offline search the pilot placements before transmission. Mean-while a criterion to select CS algorithms based on the MSE minimization is also presented. Simulations using a sparse wireless channel model have validated the effectiveness of the proposed scheme, which is demonstrated to be much faster convergent and more efficient than the exhaustive search. It has been shown that substantial performance improvement can be achieved for OMP and YALL1 based channel estimation, where YALL1 is preferred. Future work will continue to explore effective approaches for pilot design and low complexity CS algorithms, as well as the possible tradeoff between the data rate and the channel estimation accuracy.

ACKNOWLEDGMENT

The authors would like to thank Prof. X. Wang of Columbia University for his valuable comments and suggestions.

REFERENCES

- D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [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, Jun. 2010.
- [3] 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.
- [4] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2479–2493, Jul. 2009.
 [5] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An in-
- [5] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interiorpoint method for large-scale l₁-regularized least squares," *IEEE Trans. Signal Process.*, vol. 1, no. 4, pp. 606–617, Dec. 2007.
- [6] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1615–1624, Jun. 2003.
- [7] C. R. Berger, S. Zhou, J. C. Preisig, and P. Willett, "Sparse channel estimation for multicarrier underwater acoustic communication: From subspace methods to compressed sensing," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1708–1721, Mar. 2010.
- [8] C. Qi and L. Wu, "A hybrid compressed sensing algorithm for sparse channel estimation in MIMO OFDM systems," in *Proc. ICASSP*, Prague, Czech Republic, May 2011, pp. 3488–3491.
- [9] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [10] V. B. Tadic, "On the almost sure rate of convergence of linear stochastic approximation algorithms," *IEEE Trans. Inf. Theory*, vol. 50, no. 2, pp. 401–409, Feb. 2004.
- [11] I. Berenguer, X. Wang, and V. Krishnamurthy, "Adaptive MIMO antenna selection via discrete stochastic optimization," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4315–4329, Nov. 2005.
- [12] I. Jang, X. Wang, and V. Krishnamurthy, "Discrete stochastic approximation algorithms for design of optimal sensor fusion rules," *Int. J. Sensor Netw.*, vol. 2, no. 3, pp. 211–217, Sep. 2007.
- [13] M. R. Raghavendra and K. Giridhar, "Improving channel estimation in OFDM systems for sparse multipath channels," *IEEE Signal Process. Lett.*, vol. 12, no. 1, pp. 52–55, Jan. 2005.
- [14] C. Qi and L. Wu, "Application of compressed sensing to DRM channel estimation," in *Proc. VTC-Spring*, Budapest, Hungary, May 2011, pp. 1–5.
- [15] J. Z. Huang, C. R. Berger, S. Zhou, and J. Huang, "Comparison of basis pursuit algorithms for sparse channel estimation in underwater acoustic OFDM," in *Proc. MTS/IEEE OCEANS*, Sydney, Australia, May 2010, pp. 1–6.