A HYBRID COMPRESSED SENSING ALGORITHM FOR SPARSE CHANNEL ESTIMATION IN MIMO OFDM SYSTEMS

Chenhao Qi and Lenan Wu

School of Information Science and Engineering
Southeast University, Nanjing 210096, China
Email: {qch,wuln}@seu.edu.cn

ABSTRACT

Due to multipath delay spread and relatively high sampling rate in OFDM systems, the channel estimation is formulated as a sparse recovery problem, where a hybrid compressed sensing algorithm as subspace orthogonal matching pursuit (SOMP) is proposed. SOMP first identifies the channel sparsity and then iteratively refines the sparse recovery result, which essentially combines the advantages of orthogonal matching pursuit (OMP) and subspace pursuit (SP). Since SOMP still belongs to greedy algorithms, its computational complexity is in the same order as OMP. With frequency orthogonal random pilot placement, the technique is also extend to MIMO OFDM systems. Simulation results based on 3GPP spatial channel model (SCM) demonstrate that SOMP performs better than OMP, SP and interpolated least square (LS) in terms of normalized mean square error (NMSE).

***Index Terms*—** compressed sensing, MIMO, OFDM, channel estimation

1. INTRODUCTION

In traditional least square (LS) OFDM channel estimation, we must acquire channel frequency response at pilot positions and then use these observations to interpolate the rest of the subcarriers. Generally, accurate channel estimation requires more pilots than unknown channel coefficients. When the channel has large delay spread and contains abundant multipaths, the pilot number raises rapidly. For MIMO, the overhead of pilot symbols becomes considerable as transmit antennas increase. Therefore, one possible solution is to assume the channel sparsity as *a priori*. Wireless channels in practice are typically sparse. Channel impulse response usually presents to be a large number of taps with very few of them nonzero. With sparse recovery algorithms, the number of pilots can be substantially reduced. Some published work has already shown progress in this field [1, 2]. Matching pursuit (MP) [3] and orthogonal matching pursuit (OMP) [4] are commonly employed, which sequentially identifies a small subset of nonzero taps. Although the algorithms are suboptimal and greedy in nature, they are efficient in terms of performance and complexity. They have been proved to be more accurate than LS approach.

In this paper, we formulate OFDM frequency domain channel estimation as a sparse recovery problem and apply MP, OMP, and subspace pursuit (SP) algorithms. After that, we propose a hybrid compressed sensing algorithm as subspace orthogonal matching pursuit (SOMP) combining the advantages of OMP and SP. Random pilot placement is adopted according to restricted isometry property (RIP) [5]. With frequency orthogonal pilot placement, we extend our work to MIMO OFDM. 3GPP spatial channel model (SCM) [6] is applied in our simulations.

The remainder of the paper is organized as follows. Section 2 establishes the system model and formulates the estimation problem of channel impulse response. Section 3 introduces SOMP. In section 4, we make simulations with 3GPP SCM. And finally section 5 concludes the paper.

2. SYSTEM MODEL

We consider a multipath environment with S clusters or scatters. The channel impulse response between the i-th transmitter and the j-th receiver is modeled as

$$h_{ji}(\tau, t) = \sum_{p=1}^{S} \alpha_p^{ji}(t)\delta(\tau - \tau_p(t)) \tag{1}$$

where $\alpha_p^{ji}(t) \in \mathbb{C}$ and $\tau_p(t) \in \mathbb{R}^+$ are complex-valued magnitude and real-valued delay spread for path p, respectively. With block-fading channel assumption where the channel parameters are constant over each block and assuming perfect symbol synchronization, the equivalent discrete impulse response of the channel can be modeled as

$$h_{ji}(m) = \sum_{p=1}^{S} \alpha_p^{ji}\delta((m - \tau_p)T_s) \tag{2}$$

where T_s is the sampling interval of the system. We notice that in high data rate communication systems where T_s is very small compared to the maximum delay spread, (2) results in a channel with relatively few nonzero taps. Assuming total channel taps to be L and S of them nonzero ($S \ll L$), we call it S-sparse channel.

Considering an OFDM system with N subcarriers, among which N_p subcarriers are selected as pilots with positions represented by $k_1, k_2, \ldots, k_{N_p}$ ($1 \le k_1 < k_2 < \ldots < k_{N_p} \le N$) 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. The estimated transfer function on pilot subcarriers is

$$\hat{H}(m) = \frac{Y(m)}{X(m)}, \quad m = k_1, k_2, \ldots, k_{N_p} \tag{3}$$

Then we make interpolations between each two neighboring pilot subcarriers and get the full channel transfer function $\hat{H}(m)$ ($m = 1, 2, \ldots, N$), which approximates discrete fourier transform (DFT) of the channel impulse response as defined in (2). In order to make use of channel sparsity, we formulate the problem as

$$\boldsymbol{y} = \boldsymbol{X} \cdot \boldsymbol{F}_{N_p \times L} \cdot \boldsymbol{h} + \boldsymbol{n} \tag{4}$$

where $\boldsymbol{X} = diag\Big\{X(k_1), X(k_2), \cdots, X(k_{N_p})\Big\}$, $\boldsymbol{h} = [h(1), h(2), \cdots, h(L)]^{\mathrm{T}}$ and $\boldsymbol{n} = [n(1), n(2), \cdots, n(N_p)]^{\mathrm{T}}$ are the diagonal matrix, the channel impulse response, and the noise vector with each element to be an AWGN variable, respectively. $\boldsymbol{y} = [Y(k_1), Y(k_2), \cdots, Y(k_{N_p})]^{\mathrm{T}}$ and

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

where $\omega = e^{-j2\pi/N}$. Actually $\boldsymbol{F}_{N_p \times L}$ is a submatrix selected by row index $[k_1, k_2, \cdots, k_{N_p}]$ and column index $[0, 1, \cdots, L-1]$ from a standard $N \times N$ Fourier matrix. We can further assume $\boldsymbol{A} = \boldsymbol{X} \cdot \boldsymbol{F}_{N_p \times L}$ and get

$$\boldsymbol{y} = \boldsymbol{A} \cdot \boldsymbol{h} + \boldsymbol{n} \tag{5}$$

It's observed that the purpose of channel estimation is to obtain $\boldsymbol{h}$ from $\boldsymbol{y}$ and $\boldsymbol{A}$. If rows of $\boldsymbol{A}$ is more than its columns ($N_p > L$), equation(5) is a standard LS problem with its solution

$$\hat{\boldsymbol{h}}_{LS} = (\boldsymbol{A}^H \boldsymbol{A})^{-1} \boldsymbol{A}^H \cdot \boldsymbol{y} \tag{6}$$

Obviously, we are more interested in the case when the pilots are less than the channel coefficients ($N_p < L$). It's significantly appealing in reducing pilots and thus improving the spectral efficiency. Theoretically, there's feasible solution for sparse recovery problem [5] if most elements of vector $\boldsymbol{h}$ are zero ($S \ll L$).

Since MIMO OFDM channel estimation can be decomposed into simultaneously estimating of several SISO OFDM channels where we employ frequency orthogonal pilot placement for different transmitters, we will mainly focus on sparse recovery algorithms for each SISO OFDM channel in the following section.

3. COMPRESSED SENSING ALGORITHMS

A collection of sparse recovery algorithms has recently emerged with the name compressed sensing [7, 8], which enables efficient reconstruction of sparse signals from relatively few linear measurements. A S-sparse vector $\boldsymbol{h} \in \mathbb{R}^L$ can be recovered from equation (5) with deliberately designed $\boldsymbol{A} \in \mathbb{R}^{N_p \times L}$ by solving ℓ_0-norm minimization problem

$$\min_{\boldsymbol{h} \in \mathbb{R}^L} \|\boldsymbol{h}\|_0 \ \text{ s.t. } \ \|\boldsymbol{y} - \boldsymbol{A} \cdot \boldsymbol{h}\|_2 \le \sigma \tag{7}$$

where $\|\boldsymbol{h}\|_0$ counts the number of nonzero components of $\boldsymbol{h}$ and $S \le N_p \le L$. σ is the variance of noise $\boldsymbol{n}$. This problem is combinatorial and NP hard. However, Candes, Tao, Donoho, Tropp and their colleagues have shown that it can be replaced by a convex optimization problem [7, 9]

$$\min_{\boldsymbol{h} \in \mathbb{R}^L} \|\boldsymbol{h}\|_1 \ \text{ s.t. } \ \|\boldsymbol{y} - \boldsymbol{A} \cdot \boldsymbol{h}\|_2 \le \sigma \tag{8}$$

Methods for solving above and closely related problems can be roughly divided into two classes, including greedy algorithms and convex optimization algorithms [10]. Here we mainly focus on greedy algorithms due to its low complexity. In application of time-varying channel where channel estimation is frequently carried out, it's inappropriate to choose high computational convex optimization algorithms.

MP is one greedy algorithm that constructs a linear combination of matrix columns closest to the signal [3]. Although MP can rapidly find an approximation with asymptotic convergence, its shortcoming lies in the fact that it may select the same columns several times which lowers the efficiency. Hence, OMP has been proposed as a revised MP by only using residue's orthogonal component for the next iteration [4]. Only the component that is orthogonal with the space spanned by the previous selected columns is preserved. The shortcoming of OMP lies in its unidirectional adding new columns without removing out-dated columns. When a selection error occurs, the iteration will continue to the end without correcting them adaptively.

The idea of SP is to iteratively refine S columns selection from the dictionary matrix through LS method until the stop condition is satisfied [11]. At each step, it selects S columns rather than only one column as in MP and OMP. The subspace spanned by S columns is thus tracked down. The weak point of SP is that we should know S before the start of the algorithm. So it's necessary to extend SP to the occasion where the sparsity is unknown.

The stop condition for OMP employs the threshold that equals to the noise variance, while the counterpart for SP only relies on previous iterative result. Apparently the latter is more appealing since it can iteratively refine the result. Besides, SP allows the columns to enter into as well as leave the selection set, which is the chief drawback for OMP. At each iteration, OMP always greedily selects one column vector, while SP selects several columns in batch. The possibility to correctly find one column with one selection is much lower than with batch selection. As a result, we combine their advantages and propose a hybrid compressed sensing algorithm as SOMP.

Definition: If matrix $\boldsymbol{M}$ satisfies that $\boldsymbol{M}^H\boldsymbol{M}$ is invertible, we define the orthogonal part of $\boldsymbol{y}$ on $\boldsymbol{M}$ to be

$$\text{orthg}(\boldsymbol{y}, \boldsymbol{M}) \triangleq \boldsymbol{y} - \boldsymbol{M}\boldsymbol{M}^{\dagger}\boldsymbol{y} \tag{9}$$

where

$$\boldsymbol{M}^{\dagger} = (\boldsymbol{M}^H\boldsymbol{M})^{-1}\boldsymbol{M}^H \tag{10}$$

is the pseudo inverse of $\boldsymbol{M}$.

Algorithm 1 SOMP

Input: $\boldsymbol{A}, \boldsymbol{y}, \sigma$

1. Normalize columns of $\boldsymbol{A}$:
   Normalize each column of matrix $\boldsymbol{A}$ with a coefficient diagonal matrix $\boldsymbol{C}$ so that $\boldsymbol{A} = \boldsymbol{D} \cdot \boldsymbol{C}$
2. Identify the sparsity:
   Initialization:
   $\boldsymbol{r}_1 = \boldsymbol{y}, I_1 = \varnothing, I_1^c = \{1, \ldots, L\}$
   Iteration: $k = 1, 2, \ldots$
   $m_k = \arg\max_{i \in I_k^c} | < \boldsymbol{D}(i), \boldsymbol{r}_k > |$
   $\boldsymbol{u}_k = \boldsymbol{D}(m_k) - \sum_{i \in I_k} < \boldsymbol{D}(m_k), \boldsymbol{u}_i > \frac{\boldsymbol{u}_i}{\|\boldsymbol{u}_i\|_2^2}$
   $\boldsymbol{r}_k = < \boldsymbol{r}_k, \boldsymbol{u}_k > \frac{\boldsymbol{u}_k}{\|\boldsymbol{u}_k\|_2^2} + \boldsymbol{r}_{k+1}$
   if $\|\boldsymbol{r}_{k+1}\|_2 \leq \sigma$, break
   $I_{k+1} = \{I_k, m_k\}, I_{k+1}^c = I_k^c \backslash \{m_k\}$
   Store I_{k+1} in $\hat{I}$
3. Get the sparse solution:
   Initialization:
   $S = \|\hat{I}\|_0, \boldsymbol{d}_1 = \text{orthg}(\boldsymbol{y}, \boldsymbol{D}(\hat{I}))$
   Iteration: $k = 1, 2, \ldots$
   If $\boldsymbol{d}_k = 0$, break
   $\boldsymbol{z} = \boldsymbol{D}^H\boldsymbol{d}_k$
   $I_p = \left\{(l_1, \ldots, l_S) : |\boldsymbol{z}(l_1)| \geq \cdots \geq |\boldsymbol{z}(l_S)| \geq \cdots \geq |\boldsymbol{z}(l_L)| \right\}$
   $I' = \hat{I} \bigcup I_p, \boldsymbol{w} = \boldsymbol{D}^{\dagger}(I')\boldsymbol{y}$
   $I_q = \left\{(l_1, \ldots, l_S) : |\boldsymbol{w}(l_1)| \geq \cdots \geq |\boldsymbol{w}(l_S)| \geq \cdots \geq |\boldsymbol{w}(l_L)| \right\}$
   $\boldsymbol{d}_{k+1} = \text{orthg}(\boldsymbol{y}, \boldsymbol{D}(I_q))$
   If $\|\boldsymbol{d}_{k+1}\|_2 > \|\boldsymbol{d}_k\|_2$, break
   $\hat{I} = I_q$
   Store $\boldsymbol{x}$: $\boldsymbol{x}(\hat{I}) = \boldsymbol{D}^{\dagger}(\hat{I})\boldsymbol{y}, \quad \boldsymbol{x}(\hat{I}^c) = \boldsymbol{0}$
4. Output:
   $\hat{\boldsymbol{h}}_{somp} = \boldsymbol{C}^{-1}\boldsymbol{x}$

We describe the SOMP algorithm as follows. First we normalize each columns of $\boldsymbol{A}$ and get a coefficient diagonal matrix $\boldsymbol{C}$. Then we start to identify the sparsity of the solution. Index set I_1 indicating current selected columns is initialized to be empty while its complementary set I_1^c is $\{1, 2, \ldots, L\}$. Current residue $\boldsymbol{r}_1$ is initialize to be $\boldsymbol{y}$. At k-th iterative step, we select a column index m_k from I_k^c so that $\boldsymbol{D}(m_k)$ has the largest inner product with current residue $\boldsymbol{r}_k$. Gram-Schmidt orthogonalization is implemented on $\boldsymbol{D}(m_k)$ to remove the component inside the column space spanned by I_k. $\{\boldsymbol{u}_k\}$ is an iteratively generated set which can be regarded as unnormalized base vectors for the space spanned by I_k. Then we update $\boldsymbol{r}_k$ by projecting it on this space. If the stop condition $\|\boldsymbol{r}_{k+1}\|_2 \leq \sigma$ is satisfied, we break the iteration and store I_{k+1} in $\hat{I}$. Otherwise we update I_{k+1} by adding m_k into I_k; meanwhile its complementary set I_{k+1}^c is also updated. After that, we enter into the stage for sparse solution. The identified sparsity S is initialized to be the size of $\hat{I}$. $\boldsymbol{d}_1$ is initialized as the orthogonal part of $\boldsymbol{y}$ on $\boldsymbol{D}(\hat{I})$, where $\boldsymbol{D}(\hat{I})$ is defined as the submatrix from $\boldsymbol{D}$ with its columns indexed by $\hat{I}$. At k-th iterative step, we first check whether $\boldsymbol{d}_k$ is zero. If so, we break the iteration. Otherwise, we project $\boldsymbol{d}_k$ onto $\boldsymbol{D}$, from which we pick up S largest components and store their indices in I_p. The union of $\hat{I}$ and I_p is denoted as I'. From $\boldsymbol{D}^{\dagger}(I')\boldsymbol{y}$ we pick up S largest components and refine I_q. Let $\boldsymbol{d}_{k+1}$ denote the orthogonal part of $\boldsymbol{y}$ on $\boldsymbol{D}(I_q)$. If $\|\boldsymbol{d}_{k+1}\|_2$ appears to be greater than the last step, it means the orthogonal part can't be smaller. We break the iteration. Otherwise we update $\hat{I}$ by I_q. When out of iteration, $\boldsymbol{x}$ is yielded with nonzero components indexed by $\hat{I}$ satisfying $\boldsymbol{x}(\hat{I}) = \boldsymbol{D}^{\dagger}(\hat{I})\boldsymbol{y}$ and $\boldsymbol{C}^{-1}\boldsymbol{x}$ is output as the final solution.

In optimization research field [12], it's common to use two steps for sparse recovery, with the first to identify the sparsity and the second to decide corresponding value [13]. The difference is that SOMP uses a hybrid greedy algorithm instead of high computational optimization algorithm. In terms of the computational complexity, OMP is roughly in the order of $O(SN_pL)$, which is also the upper bound of SP [11]. Therefore SOMP is in the same order as OMP.

4. SIMULATION RESULTS

In our simulations, we consider a MIMO system with two transmit and two receive antennas. The channels are generated using 3GPP SCM [6], where we apply MP, OMP, SP and SOMP for channel estimation. According to [5], we randomly place pilots among all OFDM subcarriers so that the matrix $\boldsymbol{A}$ in (5) satisfies RIP. Meanwhile, with the purpose to simplify MIMO channel estimation problem, we employ frequency orthogonal pilot placement for different transmit antennas. For example, we place 12 pilots among 256 subcarriers for two different transmit antennas. First we randomly select 24 positions from 256 candidates. Then we randomly choose 12 positions from the selected 24-position subset for one transmit antenna. The remaining 12 positions are used for the other transmit antenna. Consequently pilots for different antennas are un-overlapped and do not interfere with each other. 2×2 MIMO OFDM channel estimation problem is decoupled into 4 SISO ones. Hence we can concentrate on sparse channel estimation for each SISO OFDM system.

System parameters in our simulations are listed in Table 1. Pilot positions used in simulations are [6, 20, 36, 58, 70, 90, 118, 169, 182, 202, 223, 240]. Performance comparisons of MP, OMP, SP, SOMP and cubic spline interpolated LS in terms of normalized mean square error (NMSE) are illustrated in Figure 1. We define NMSE to be $\|\hat{\boldsymbol{h}} - \boldsymbol{h}\|_2^2 / \|\boldsymbol{h}\|_2^2$, where $\hat{\boldsymbol{h}}$

Table 1. System parameters

Number of transmitting antennas	$N_t = 2$
Number of receiving antennas	$N_r = 2$
Number of total subcarriers	$N = 256$
Number of pilot subcarriers	$N_p = 12$
Number of cyclic prefix	$N_G = 64$
Number of channel multipaths	$S = 5$
Length of channel impulse response	$L = 40$
Modulation	QPSK

is the estimation of $\boldsymbol{h}$.

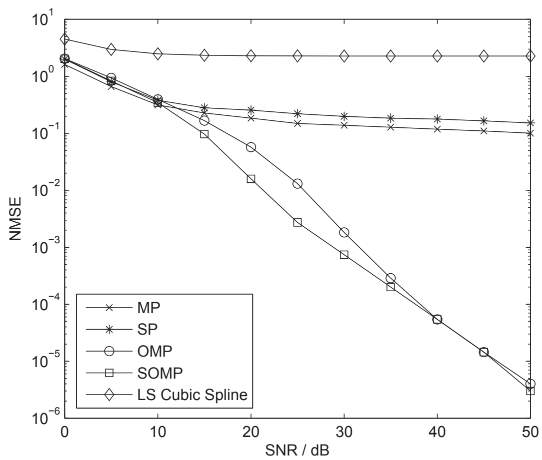

Fig. 1. NMSE vs SNR with unknown sparsity

It's observed from Figure 1 that LS cubic spline which does not take channel sparsity as *a prori* performs much worse than sparse recovery algorithms. SOMP has the best performance especially for SNR from 15dB to 30dB. In this range, the noise is unnegligible. It disturbs the correct selection for OMP due to its one selection manner as we analyzed in Section 3. When SNR goes beyond 35dB, OMP can also do the right thing as SOMP because the noise is much smaller compared to the signal. For SNR lower than 15dB, OMP and SOMP are approximated because the noise deteriorates both algorithms. Consequently, SOMP combines the advantages of OMP to identify the sparsity and SP to refine the best group selection. It has been demonstrated to be an appropriate candidate for low complexity sparse recovery.

5. CONCLUSIONS

This paper studied the sparse recovery algorithms for pilot assisted MIMO OFDM channel estimation, where SOMP is proposed and proved to be better than MP, OMP and SP. Further work will continue on sparse recovery algorithms. Complexity reduction will be emphasized.

6. REFERENCES

[1] M. Sharp and A. Scaglione, "Application of Sparse Signal Recovery to Pilot-Assisted Channel Estimation," in *Proc. ICASSP*, Las Vegas, NV, USA, Mar. 2008, pp. 3469–3472.

[2] S. F. Cotter and B. D. Rao, "Sparse Channel Estimation via Matching Pursuit with Application to Equalization," *IEEE Trans. Commun.*, vol. 50, no. 3, pp. 374–377, Mar. 2002.

[3] S. G. Mallat and Z. Zhang, "Matching Pursuits With Time-Frequency Dictionaries," *IEEE Trans. Sig. Proc.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[4] J. A. Tropp and A. C. Gilbert, "Signal Recovery From Random Measurements via Orthogonal Matching Pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[5] E. J. Candes, J. Romberg, and T. Tao, "Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[6] J. Salo, G. Galdo, J. Salmi, P. Kyosti, M. Milojevic, D. Laselva, and C. Schneider, "MATLAB implementation of the 3GPP Spatial Channel Model," *3GPP TR 25.996*, Nov. 2005.

[7] D. L. Donoho, "Compressed Sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[8] R. G. Baraniuk, "Compressive Sensing," *IEEE Sig. Proc. Mag.*, vol. 24, no. 4, pp. 118–121, July 2007.

[9] J. A. Tropp, "Just Relax: Convex Programming Methods for Identifying Sparse Signals in Noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1030–1051, Mar. 2006.

[10] J. A. Tropp and S. J. Wright, "Computational Methods for Sparse Solution of Linear Inverse Problems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 948–958, June 2010.

[11] W. Dai and O. Milenkovic, "Subspace Pursuit for Compressive Sensing Signal Reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, Apr. 2009.

[12] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, U.K., Dec. 2004.

[13] S. S. Chen, D. L. Donoho, and M. A Saunders, "Atomic Decomposition by Basis Pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, 2001.