

Compressed Channel Estimation with Joint Pilot Symbol and Placement Design for High Mobility OFDM Systems

Xiang Ren, Wen Chen, Meixia Tao, and Xiaofei Shao

Department of Electronic Engineering,
Shanghai Jiao Tong University, China

Email: {renx; wenchen; mxtao; simon_sho}@sjtu.edu.cn

Abstract—High mobility introduces significant Doppler frequency spread and hence imposes great challenge to channel estimation in orthogonal frequency division multiplexing (OFDM) systems. In this paper, we propose a new low coherence compressed channel estimation method for OFDM systems in the high mobility environment. Based on the compressed sensing (CS) coherence minimization criterion, we propose a new pilot design algorithm using the discrete stochastic optimization method to improve the estimation performance. The pilot symbol and pilot placement are jointly designed to obtain the optimal pilot. Numerical results demonstrate that the proposed method has a fast convergency property and achieves better MSE performance than existing channel estimation methods over high mobility channels.

Index Terms—High mobility; channel estimation; compressed sensing (CS); orthogonal frequency division multiplexing (OFDM).

I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is an attractive technique for high-speed data transmission in wireless broadband communication systems. In OFDM systems, each subcarrier has a narrow bandwidth, which makes the signal robust against the frequency selectivity caused by the multipath effect. However, OFDM is sensitive to the time selectivity, which is induced by rapid time channel variations. In high mobility environments, wireless channels are both rapidly time-varying and frequency-selective. The large Doppler spreads caused by rapid time variations would destroy the orthogonality among subcarriers and hence can highly degrade the performance of OFDM. Moreover, due to the rapid time variation, the channel estimation has been a long-standing challenge. Many previous works have focused on the advances of new channel estimation methods in high mobility channels [1]–[4].

Among the various channel estimation methods, compressed sensing (CS)-based channel estimation is promising as it can utilize the inherent sparsity in high-mobility wireless channels [3]. In CS technique, coherence is a key factor and a lower coherence leads to a better CS performance [5]. Therefore, to further enhance the performance of compressed channel estimation, researchers [6]–[8] have proposed different pilot design methods to minimize the CS coherence. In particular,

the works [6]–[7] proposed pilot placement methods with large iterations. The work [8] designed the pilot symbol by shrinking the entries in the Gram matrix of the pilot matrix.

In this paper, we introduce a new low coherence compressed channel estimation method for OFDM systems in the high mobility environment, in which the pilot symbol and the pilot placement are jointly designed. Based on the CS coherence minimization criterion, a pilot design algorithm using the discrete stochastic optimization method is proposed with the high mobility channel model basis. The pilot symbol and pilot placement are jointly designed to obtain the optimal pilot. Numerical results demonstrate that the proposed method has a good convergency property and satisfies estimation performance over high mobility channels.

Throughout this paper, $\|\cdot\|_{\ell_0}$ denotes the number of non-zero entries in a matrix, $\|\cdot\|_{\ell_2}$ denotes the Euclidean norm, $(\cdot)^T$ denotes the transpose of a matrix, $(\cdot)^H$ denotes Hermitian of a matrix, \otimes denotes the Kronecker product, $\text{diag}(\cdot)$ denotes a diagonal matrix and $\mathbf{a} = \text{vec}(\mathbf{A})$ denotes the vector obtained by stacking columns of matrix \mathbf{A} . Finally, \mathbb{R} denotes the field of reals and \mathbb{Z} denotes the set of integers.

II. SYSTEM MODEL

We consider an OFDM system with K subcarriers in a high mobility environment. Denote $X^n(k)$ is the n th OFDM symbol in the frequency domain at the k th subcarrier, in which $n = 1, 2, \dots, N$ denotes the index of the OFDM symbol and $k = 1, 2, \dots, K$ denotes the subcarrier index. Then the inverse discrete Fourier transform (IDFT) is performed at the transmitter. After the parallel to serial conversion, the cyclic prefix (CP) is inserted to avoid the intersymbol interference (ISI). After removing CP at the receiver and passing the discrete Fourier transform (DFT) operator, the received signal in the frequency domain can be represented as

$$\mathbf{Y}^n = \mathbf{H}^n \mathbf{X}^n + \mathbf{W}^n, \quad (1)$$

where \mathbf{Y}^n is the received signal vector at each subcarrier, \mathbf{H}^n is a $K \times K$ channel matrix in the frequency domain, \mathbf{X}^n is the transmitted signal vector at each subcarrier, and \mathbf{W}^n is the additive white Gaussian noise (AWGN) vector in the frequency domain consisting of $w^n(k)$ with a zero mean

and σ_w^2 variance. If the channel is time-invariant or the ICI mitigation is performed perfectly, then the off-diagonal term $H^n(k, d)$ ($k \neq d$) is ignored, and the diagonal term $H^n(k, d)$ ($k = d$) alone represents the channel in the frequency domain.

For the sake of convenience, we call the diagonal term as the ICI-free channel and the off-diagonal term as the ICI channel. Therefore, the channel matrix \mathbf{H}^n can be divided into two parts, the ICI-free channel matrix $\mathbf{H}_{free}^n \triangleq \text{diag}\{[H^n(1, 1), H^n(2, 2), \dots, H^n(K, K)]\}$ and the ICI channel matrix $\mathbf{H}_{ICI}^n \triangleq \mathbf{H}^n - \mathbf{H}_{free}^n$. Then (1) can be rewritten as

$$\mathbf{Y}^n = \mathbf{H}_{free}^n \mathbf{X}^n + \mathbf{H}_{ICI}^n \mathbf{X}^n + \mathbf{W}^n, \quad (2)$$

$$= \mathbf{X}_d^n \mathbf{H}_{vec}^n + \mathbf{H}_{ICI}^n \mathbf{X}^n + \mathbf{W}^n, \quad (3)$$

where $\mathbf{X}_d^n = \text{diag}\{[X^n(1), X^n(2), \dots, X^n(K)]^T\}$ is a diagonal matrix of the transmitted signal vector \mathbf{X}^n , $\mathbf{H}_{vec}^n = \text{vec}\{\mathbf{H}_{free}^n\}$ is the stacking vector of the diagonal of \mathbf{H}_{free}^n , and \mathbf{H}_{ICI}^n denotes the ICI channel matrix, respectively.

A. High Mobility Channels

Wireless communication in a mobility environment is unavoidable impaired by both the delay and the Doppler spread which are caused by the multipath effect and fast time-variant channels, respectively. Therefore, high mobility channels can be considered as time and frequency doubly-selective channels [1][3].

Let τ_{\max} be the delay spread, v_{\max} be the maximum Doppler frequency spread, T_d be the packet duration, W be the bandwidth, T_0 be the OFDM symbol duration, W_0 be the bandwidth of each subcarrier, $N_t = T_d/T_0$ and $N_f = W/W_0$. The high mobility channel between the transmitter and the receiver in the delay-Doppler domain can be modeled as

$$H(n, f) = \sum_{l=0}^{L-1} \sum_{m=-M}^M \beta_{l,m} e^{j2\pi \frac{m}{N_t} n} e^{-j2\pi \frac{l}{N_f} f}, \quad (4)$$

where $L = \lceil W\tau_{\max} \rceil + 1$ represents the maximum number of the resolvable delays and $M = \lceil T_d v_{\max} \rceil$ represents the maximum number of the resolvable Doppler spreads of the high mobility channel, f is the subcarrier frequency and n is the time slot. For the sake of writing convenience, we define two vectors $\mathbf{u}_f = [1, e^{-j2\pi \frac{1}{N_f} f}, \dots, e^{-j2\pi \frac{(L-1)}{N_f} f}]^T$

and $\mathbf{u}_n = [e^{j2\pi \frac{M}{N_t} n}, e^{j2\pi \frac{(M-1)}{N_t} n}, \dots, e^{j2\pi \frac{0}{N_t} n}]^T$. Then the channel model can be rewritten as a matrix form:

$$H(n, f) = \mathbf{u}_f \mathbf{B} \mathbf{u}_n^T = (\mathbf{u}_n \otimes \mathbf{u}_f) \mathbf{b}, \quad (5)$$

where \mathbf{B} is an $L \times (2M+1)$ channel coefficient matrix in the delay-Doppler domain of the high mobility channel, i.e.,

$$\mathbf{B} = \begin{bmatrix} \beta_{0,-M} & \cdots & \beta_{0,M} \\ \vdots & \ddots & \vdots \\ \beta_{L-1,-M} & \cdots & \beta_{L-1,M} \end{bmatrix}, \quad (6)$$

and we define that $\mathbf{b} \triangleq \text{vec}(\mathbf{B})$ is the stacking vector of the channel coefficient matrix, i.e.,

$$\mathbf{b} = [\beta_{0,-M}, \dots, \beta_{L-1,-M}, \dots, \beta_{0,M}, \dots, \beta_{L-1,M}]^T. \quad (7)$$

Each $\beta_{l,m}$ is the sum of the complex gains of all physical paths lying in the unit sampling subspace in the delay-Doppler domain. The coefficients are considered constant in each OFDM symbol and different between two symbols. For some appropriately chosen threshold $\gamma \geq 0$, the channel is said to be S -sparse, if $|\beta_{l,m}|^2 > \gamma$ and $\|\mathbf{b}\|_{\ell_0} = S \ll L(2M+1)$. Recent researches [1] and [3] have shown that the doubly-selective channels can be modeled accurately with sparse \mathbf{b} in the delay-Doppler domain. Therefore, CS is introduced in the following section to utilize the sparsity of high mobility channels.

B. Pilot-Assisted Channel Estimation

Considering the high mobility channel is fast time-varying, here we propose a pilot-assisted method based on the comb-type pilot. Assume that there are P pilots and the set of pilot subcarriers placement is $\mathbf{p} = [k_1, k_2, \dots, k_P]$, and $P \leq K$. Let $\Phi^n = [\mathbf{u}_n \otimes \mathbf{u}_{k_1}; \mathbf{u}_n \otimes \mathbf{u}_{k_2}; \dots; \mathbf{u}_n \otimes \mathbf{u}_{k_P}]$ be the $P \times L(2M+1)$ channel model dictionary matrix in the n th OFDM symbol, in which $\mathbf{u}_{k_p} = \mathbf{u}_f|_{f=k_p}$ and $P < L(2M+1)$. Since we only consider the estimation process in one OFDM symbol, the superscripts n in the rest of the paper are omitted for compactness. Then the received pilot vector can be represented as the following matrix form:

$$\mathbf{Y}(\mathbf{p}) = \mathbf{X}_d(\mathbf{p}) \mathbf{H}_{vec}(\mathbf{p}) + \mathbf{H}_{ICI}(\mathbf{p}, :) \mathbf{X} + \mathbf{W}(\mathbf{p}), \quad (8)$$

$$= \mathbf{X}_d(\mathbf{p}) \Phi \mathbf{b} + \mathbf{H}_{ICI}(\mathbf{p}, :) \mathbf{X} + \mathbf{W}(\mathbf{p}), \quad (9)$$

where $\mathbf{Y}(\mathbf{p})$, $\mathbf{X}(\mathbf{p})$, and $\mathbf{W}(\mathbf{p})$ denote the received pilots, the transmitted pilots, and AWGN at the pilot placement \mathbf{p} , respectively. So far, the task of estimating the high mobility channel \mathbf{H} in the frequency domain converts to estimating the channel coefficients \mathbf{b} in the delay-Doppler domain, which fortunately are sparse in practice [1][3].

III. PILOT DESIGNED COMPRESSED ESTIMATION

CS is an innovative and revolutionary idea that can utilize the inherent sparsity of the wireless channel, which is known as the compressed channel estimation [3]. Now we briefly introduce CS theory for better explanation. Let signal $\mathbf{x} \in \mathbb{R}^m$ be an $m \times 1$ vector and has the sparsity of S under the dictionary basis $\mathbf{D} \in \mathbb{R}^{m \times U}$ ($m < U$), which means $\mathbf{x} = \mathbf{D}\mathbf{a}$ and only S elements in vector \mathbf{a} are non-zero. Then \mathbf{x} can projects to \mathbf{y} by $\mathbf{y} = \mathbf{P}\mathbf{x} = \mathbf{P}\mathbf{D}\mathbf{a}$ with a measurement matrix $\mathbf{P} \in \mathbb{R}^{p \times m}$, which is not related to the dictionary basis \mathbf{D} . If $\mathbf{P}\mathbf{D}$ satisfies the restricted isometric property (RIP) [9], then CS reconstruction methods such as basis pursuit (BP) [10] can reconstruct \mathbf{x} from \mathbf{y} .

A. Coherence of CS

In this subsection, we review the definition of the coherence [9], which is a fundamental concept of CS, and give an useful theorem.

Definition 1 (Coherence [9]): For a matrix \mathbf{M} with the i th column of \mathbf{d}_i , its coherence is defined as the largest absolute value of the normalized inner product between different

columns in \mathbf{M} and can be written as follows:

$$\mu\{\mathbf{M}\} = \max_{i \neq j} \frac{|\mathbf{d}_i^H \mathbf{d}_j|}{\|\mathbf{d}_i\| \cdot \|\mathbf{d}_j\|}. \quad (10)$$

Previous works [12] and [13] established that both BP and orthogonal greedy algorithms (OGA) [11] are valid if the following theorem [8] is satisfied.

Theorem 1 ([8]): For a dictionary matrix \mathbf{D} and measurement matrix \mathbf{P} , assume that \mathbf{PD} satisfies the RIP, if the representation $\mathbf{y} = \mathbf{P}\mathbf{x} = \mathbf{PD}\mathbf{a}$ satisfies the requirement

$$S = \|\mathbf{a}\|_{\ell_0} < \frac{1}{2} \left(1 + \frac{1}{\mu\{\mathbf{PD}\}} \right), \quad (11)$$

then a) \mathbf{a} is the unique sparsest representation of \mathbf{x} ; b) the deviation of the reconstructed $\hat{\mathbf{a}}$ from \mathbf{a} by BP or OGA can be bounded by

$$\|\hat{\mathbf{a}} - \mathbf{a}\|_{\ell_2}^2 \leq \frac{\epsilon^2}{1 - \mu\{\mathbf{PD}\}(2S - 1)}, \quad (12)$$

for some constant $\epsilon > 0$.

From Theorem 1, we can find that if \mathbf{P} is designed with a fixed \mathbf{D} such that $\mu\{\mathbf{PD}\}$ is as small as possible, then a large number of candidate signals are able to reside under the umbrella of successful CS behavior and lead to better CS performance. This is reasonable for lower $\mu\{\mathbf{PD}\}$ resulting in a higher upper-bound of S and introducing a lower reconstruction error bound which leads to better CS recovery performance.

B. Problem Statement

As we have already known that a lower μ leads to a better CS performance, we are going to reduce the coherence in our system to get better estimation performance. In this subsection, the transmitted pilot matrix is optimized to reduce the system coherence. Both the pilot symbol and the pilot placement of the transmitted pilot matrix $\mathbf{X}_d(\mathbf{p})$ are considered with the channel model dictionary Φ . Our objective is to minimize the coherence $\mu\{\mathbf{X}_d(\mathbf{p})\Phi\}$. This optimization problem can be formulated as

$$\min_{|\mathbf{X}|, \mathbf{p}} \mu\{\mathbf{X}_d(\mathbf{p})\Phi\}, \quad (13)$$

where $|\mathbf{X}|$ denotes the pilot symbols in $\mathbf{X}_d(\mathbf{p})$ and \mathbf{p} denotes the set of pilot subcarrier placement. According to Definition 1, the objective function can be represented as

$$\mu\{\mathbf{X}_d(\mathbf{p})\Phi\} = \max_{m \neq r} \left| \sum_{k_i \in \mathbf{p}} |X(k_i)|^2 \phi(k_i, m)^H \phi(k_i, r) \right|, \quad (14)$$

where $\phi(k_i, m)$ is the entry of the channel model dictionary Φ and $0 \leq m < r \leq L(2M + 1) - 1$. Suppose all pilot and data symbols are modulated with two symbol powers (such as the star 16-QAM), then there are two pilot powers E_1 and E_2 corresponding to pilot placement subsets \mathbf{t}_1 and \mathbf{t}_2 ($\mathbf{t}_1 \cup \mathbf{t}_2 = \mathbf{p}$), i.e.,

$$E_1 = |X(k_{j_1})|^2, \quad k_{j_1} \in \mathbf{t}_1, \quad (15)$$

and

$$E_2 = |X(k_{j_2})|^2, \quad k_{j_2} \in \mathbf{t}_2. \quad (16)$$

Taking pilot powers into consideration, then the objective function can be represented as (19). Hence the coherence optimized pilot matrix is given as

$$\hat{\mathbf{X}}_d = \arg \min_{|\mathbf{X}|, \mathbf{p}} \mu\{\mathbf{X}_d(\mathbf{p})\Phi\}. \quad (20)$$

C. Pilot Design Algorithm

The optimization problem (20) is not easy to solve. An intuitive method is to do exhaustive search by $2^P \binom{K}{P}$ combinations to achieve the global optimal solution. Obviously, the complexity is too large to be practical for an OFDM communication system. Fortunately, the problem (20) can be considered as a discrete stochastic optimization problem [14].

In this subsection, we propose an iterative algorithm that resembles a stochastic optimization method. The algorithm is aiming to reduce $\mu\{\mathbf{X}_d(\mathbf{p})\Phi\}$ and get the optimized pilot matrix $\hat{\mathbf{X}}_d$ with the channel model Φ . The key idea of our algorithm is to generate a sequence of pilot sets, where each new set is obtained from the previous one by taking a step towards the global optimum. Define \mathbf{p}_m , $\tilde{\mathbf{p}}_m$ and $\hat{\mathbf{p}}_m$ as different pilot placement sets at the m th iteration. *Iter* denotes the iteration times, and N_x denotes the number of total pilot sets. The pilot placement set occupation probability vector $\mathbf{I}[m] = [I[m, 1], I[m, 2], \dots, I[m, N_x]]^T$ indicates the occupation probability of each pilot placement set at the m th iteration, in which $I[m, i] \in [0, 1]$ and $\sum_i I[m, i] = 1$. The details are given in Algorithm 1. The global convergence property of Algorithm 1 is proven in [14]. Numerical results in the following section will show the convergence property of Algorithm 1.

D. Practical Applicability

Here we discuss the practical applicability of the proposed scheme. The flow chart of a practical scheme is given as Fig. 1. Firstly, in some practical high mobility systems, such as a high speed train (HST) system in [15], the system parameters (such as the maximum delay spread τ_{max} , the maximum Doppler

$$\mu\{\mathbf{X}_d(\mathbf{p})\Phi\} = \max_{0 \leq m < r \leq L(2M+1)-1} \left| \sum_{k_i \in \mathbf{p}} |X(k_i)|^2 \phi(k_i, m)^H \phi(k_i, r) \right|, \quad (18)$$

$$= \max_{0 \leq m < r \leq L(2M+1)-1} \left| \sum_{k_{j_1} \in \mathbf{t}_1} E_1 \cdot \phi(k_{j_1}, m)^H \phi(k_{j_1}, r) + \sum_{k_{j_2} \in \mathbf{t}_2} E_2 \cdot \phi(k_{j_2}, m)^H \phi(k_{j_2}, r) \right|. \quad (19)$$

Algorithm 1 : Pilot Design Algorithm**Input:** Random pilot matrix \mathbf{X}_0 .**Output:** Optimized pilot matrix $\hat{\mathbf{X}}_d$.

```

1: Initialization: Set  $\mathbf{I}[0] = \mathbf{0}_{N_x}$ ,  $I[0,0] = 1$ ,  $u = 0, v = 0$ 
   and  $\hat{\mathbf{p}}_0 = \mathbf{p}_0$ , set  $Iter$ .
2: for  $n = 0, 1, \dots, Iter - 1$  do
3:   for  $k = 0, 1, \dots, P - 1$  do
4:      $m \leftarrow n \times P + k$ ;
5:     generate  $\tilde{\mathbf{p}}_m$  with operation  $\tilde{\mathbf{p}}_m \leftarrow \mathbf{p}_m$ ;
6:     if  $\mu\{\mathbf{X}_d(\tilde{\mathbf{p}}_m)\Phi\} < \mu\{\mathbf{X}_d(\mathbf{p}_m)\Phi\}$  then
7:       if  $\mu\{\mathbf{X}_d(\tilde{\mathbf{p}}_m)\Phi\} < \mu\{\mathbf{X}_d(\tilde{\mathbf{p}}_m^{E_2})\Phi\}$  then
8:          $\mathbf{X}_{m+1} \leftarrow \mathbf{X}_d(\tilde{\mathbf{p}}_m^{E_1})$ ;
9:       else
10:         $\mathbf{X}_{m+1} \leftarrow \mathbf{X}_d(\tilde{\mathbf{p}}_m^{E_2})$ ;
11:       end if
12:     else
13:       if  $\mu\{\mathbf{X}_d(\mathbf{p}_m^{E_1})\Phi\} < \mu\{\mathbf{X}_d(\mathbf{p}_m^{E_2})\Phi\}$  then
14:         $\mathbf{X}_{m+1} \leftarrow \mathbf{X}_d(\mathbf{p}_m^{E_1})$ ;
15:       else
16:         $\mathbf{X}_{m+1} \leftarrow \mathbf{X}_d(\mathbf{p}_m^{E_2})$ ;
17:       end if
18:     end if
19:      $u \leftarrow m + 1$ ;
20:      $\mathbf{I}[m+1] \leftarrow \mathbf{I}[m] + (\mathbf{D}[m+1] - \mathbf{I}[m])/(m+1)$ ;
21:     if  $I[m+1, u] > I[m+1, v]$  then
22:        $\hat{\mathbf{X}}_{m+1} \leftarrow \mathbf{X}_{m+1}$ ;  $v \leftarrow u$ ;
23:     else
24:        $\hat{\mathbf{X}}_{m+1} \leftarrow \hat{\mathbf{X}}_m$ ;
25:     end if
26:   end for (k)
27: end for (n)

```

frequency spread v_{max} and etc.) can always be estimated in advance. Thus, we can get the channel model dictionary Φ and model the high mobility channel as $\mathbf{H} = \Phi\mathbf{b}$. Secondly, the optimal pilot can be obtained by Algorithm 1 and sent to the transmitter and the receiver, which is an off-line process. When the system runs, the transmitter sends the designed pilot $\hat{\mathbf{X}}_d$ to estimate the channel. After $\hat{\mathbf{X}}_d$ passes the high mobility channel, the received signal at the receiver is represented as Eq. (9). Then, CS reconstruction algorithms such as BP and OMP can reconstruct the estimated coefficients $\tilde{\mathbf{b}}$ in the delay-Doppler domain. After that, the CSI is recovered by $\tilde{\mathbf{H}} = \Phi\tilde{\mathbf{b}}$ at the receiver. Finally, the estimated CSI $\tilde{\mathbf{H}}$ is used to recover the received symbol.

Moreover, except for the system parameters needed by conventional pilot-assisted channel estimation methods, the only necessarily priori information of the proposed scheme is the maximum Doppler frequency spread v_{max} , which makes the proposed scheme feasible for implementation in other OFDM communication systems.

IV. NUMERICAL RESULTS

In this section, under the high mobility environment, we compare the mean square error (MSE) performances of two typical compressed channel estimators BP [10] and OMP [11]

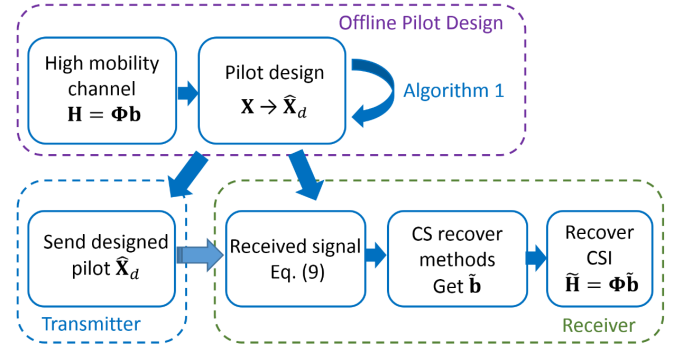


Fig. 1. Flow chart of a practical channel estimation scheme.

with different pilot design methods to show the performance of the proposed method Algorithm 1. In addition, the performances of the conventional LS estimator, the linear minimum mean square error (LMMSE) [2] estimator and the best linear unbiased estimator (BLUE) [2] are also compared.

Here we consider an OFDM system in the high mobility environment. Assumed that there are 512 subcarriers and 12.5% are pilot subcarriers. The bandwidth is 7.68MHz, the packet duration is $T_d = 0.45\text{ms}$ and the carrier frequency $f_c = 2.1\text{GHz}$, according to the Long Term Evolution (LTE) standard. The additive noise is a Gaussian and white random process. The high mobility channel is modeled as Eq. (4). The maximum Doppler frequency spread is $v_{max} = 972.2\text{Hz}$, which means that the maximum speed of the receiver is 500km/h. In our experiment, we assumed that there are only 12% of the channel coefficients are dominant. The pilots and symbols are modulated by the star 16-QAM, which means they have two symbol powers.

Fig. 2 presents the comparison of the MSE performances of different estimators with different designed pilots at 500km/h, which means $v_{max} = 972.2\text{Hz}$. Linear estimators LS, LMMSE and BLUE are considered as [2] with Algorithm 1. The equidistant method is the equidistant pilot placement with random pilot symbols, which is claimed in [4] as the optimal pilot placement to the doubly selective channels. The exhaustive method is proposed in [6] which operates an exhaustive search from a designed optimal pilot subset. The CDS method is considered as the pilot placement designed in [7] with random pilot symbols. Our method Algorithm 1 is operated with $Iter = 7$, which means we calculate 7 set of pilots and get the optimized pilot. It can be observed that the BP channel estimators significantly improves the performances by utilizing the inherent sparsity of the high mobility channels. On the other hand, LS, LMMSE and BLUE need more pilots to obtain better CSI, and we can find that Algorithm 1 has few effect on linear estimators for their not utilizing the coherence of CS. As expected, Algorithm 1 gets better performance than other methods with the same iteration times. It means that Algorithm 1 effectively reduces the system coherence by jointly design the pilot symbol and placement, which improves the system performance.

Fig. 3 presents the comparison of the MSE performances of

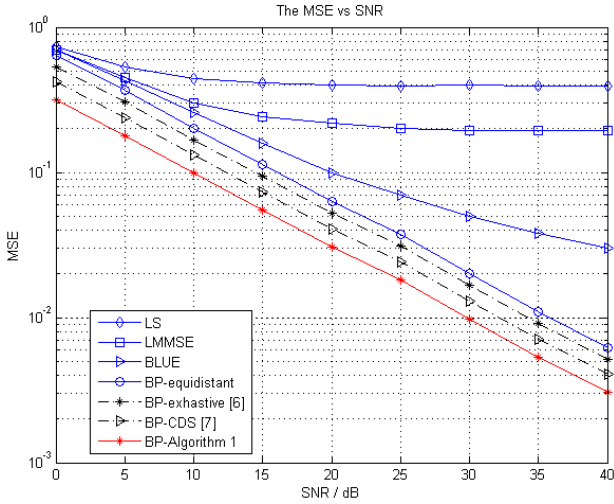


Fig. 2. MSE performance of the linear estimators and BP estimators with different pilots in an OFDM system at 500km/h, in which there are 512 subcarriers and 12.5% are pilots.

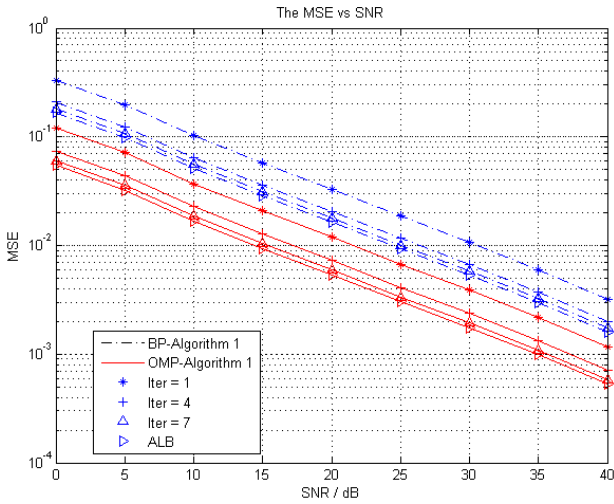


Fig. 3. MSE performance of BP and OMP estimators with Algorithm 1 by different iterations in an OFDM system at 300km/h, in which there are 512 subcarriers and 12.5% are pilots.

BP and OMP estimators at 300km/h, that generates the maximum Doppler shift frequency spread as $v_{max} = 583.3\text{Hz}$. The performances of BP and OMP with Algorithm 1 by different iterations are given in this figure. As a reference, $Iter = 20$ is given to show the approximate lower bound (ALB) of Algorithm 1, which means the performance improves extremely little with increasing iterations. It can be found that the gains become smaller and smaller with growing iterations, which shows the convergence property of Algorithm 1. Numerical results show that the optimized pilot with $Iter = 7$ is good enough for a practical system. Note that the proposed pilot design scheme is an off-line process, thus its complexity can be ignored in a practical system. As can be seen, the proposed algorithm improves the performances of both BP and OMP estimators in the high mobility environment.

V. CONCLUSION

In this paper, a low coherence compressed channel estimation method is proposed for an OFDM system in high mobility environments, in which the pilot symbol and the pilot placement are jointly designed. Numerical results demonstrated that the proposed algorithm converges fast and efficiently improves the system performance in the high mobility environments. Furthermore, except for the system parameters needed by conventional pilot-assisted channel estimation methods, the only necessarily priori information of the proposed scheme is the maximum Doppler frequency v_{max} . This makes the proposed scheme feasible for implementation in future wireless OFDM communication systems.

ACKNOWLEDGMENT

This work is supported by the National 973 Project #2012CB316106, by NSF China #61161130529, #61328101, and #61322102, by STCSM Science and Technology Innovation Program #13510711200, and by SEU National Key Laboratory #2013D11.

REFERENCES

- [1] W. U. Bajwa, A. M. Sayeed, and R. Nowak, "Sparse multipath channels: modeling and estimation," *IEEE Digital Signal Processing Workshop*, vol. 1, pp. 1-6, 2009.
- [2] Z. Tang, R. C. Cannizzaro, G. Leus, and P. Banelli, "Pilot-assisted time-varying channel estimation for OFDM systems," *IEEE Transactions on Signal Processing*, vol. 55, pp. 2226-2238, 2007.
- [3] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: a new approach to estimating sparse multipath channels," *IEEE Signal Processing Mag.*, vol. 98, no. 6, pp. 1058-1076, June 2010.
- [4] X. Ma, G. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly-selective fading channels," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1351-1366, May 2003.
- [5] D. L. Donoho, M. Elad, and V. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 6-18, Jan. 2006.
- [6] X. He and R. Song, "Pilot pattern optimization for compressed sensing based sparse channel estimation in OFDM systems," *International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1-5, Oct. 2010.
- [7] C. Qi and L. Wu, "Optimized pilot placement for sparse channel estimation in OFDM systems," *IEEE Signal Processing Letters*, vol. 18, no. 12, pp. 749-752, Dec. 2010.
- [8] X. Ren, W. Chen, and Z. Wang, "Low coherence compressed channel estimation for high mobility MIMO OFDM systems," *Conference of Global Communications (GLOBECOM)*, Dec. 2013.
- [9] J. C. Emmanuel and B. W. Michael, "An introduction to compressive sampling," *IEEE Signal Processing Mag.*, vol. 25, no. 2, Mar. 2008.
- [10] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, pp. 129-159, 2001.
- [11] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," *Proceedings of the 27th Annual Asilomar Conference on Signals, Systems, and Computers*, 1993.
- [12] M. Elad, "Optimized projections for compressed sensing," *IEEE Transactions on Signal Processing*, vol. 55, no. 12, pp. 5695-5702, Dec. 2007.
- [13] E. J. Candes, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207-1223, 2006.
- [14] 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.
- [15] J. Ma, P. Orlik, J. Zhang, and G. Y. Li, "Statistics-based ICI mitigation in OFDM over high-mobility channels with line-of-sight components," *IEEE Transactions on Wireless Communications*, vol. 10, no. 11, pp. 3577-3582, Nov. 2011.