This is an author produced version of Weighted Compressive Sensing Based Uplink Channel Estimation for TDD Massive MIMO Systems.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/109523/

**Article:**

https://doi.org/10.1049/iet-com.2016.0625

© 2016 Institution of Engineering and Technology. This paper is a postprint of a paper submitted to and accepted for publication in IET Communications and is subject to Institution of Engineering and Technology Copyright. The copy of record is available at IET Digital Library. Uploaded in accordance with the publisher's self-archiving policy.
Weighted Compressive Sensing Based Uplink Channel Estimation for TDD Massive MIMO Systems

Yang Nan¹*, Li Zhang², Xin Sun¹

¹School of Electrical and Information Engineering, Beijing Jiaotong University, Beijing, 100044
²School of Electronic and Electrical Engineering, University of Leeds, Leeds, LS2 9JT
*yananbjtu@hotmail.com

Abstract: In this paper, the channel estimation problem for the uplink massive multi-input multi-output (MIMO) system is considered. Motivated by the observations that the channels in massive MIMO systems may exhibit sparsity and the channel support changes slowly over time, we propose one efficient channel estimation method under the framework of compressive sensing. By exploiting the channel impulse response (CIR) estimated from the previous OFDM symbol, we firstly estimate the probabilities that the elements in the current CIR are nonzero. Then, we propose the probability-weighted subspace pursuit (PWSP) algorithm exploiting these probability information to efficiently reconstruct the uplink massive MIMO channel. Moreover, noting that the massive MIMO systems also share a common support within one channel matrix due to the shared local scatterers in the physical propagation environment, an antenna collaborating method is exploited for the proposed method to further enhance the channel estimation performance. Simulation results show that compared to the existing compressive sensing methods, the proposed methods could achieve higher spectral efficiency as well as more reliable performance over time-varying channel.

1. Introduction

As a promising technology for future communication networks, massive multiple-input multiple-output (MIMO) systems where the base station (BS) is equipped with a large number of antennas have gained much interest in the research community [1]. In order to benefit from the advantages of massive MIMO systems, there are still many issues that need to be properly addressed. For example, the channel state information (CSI) is crucial for both of the uplink and downlink. But in typical massive MIMO systems, reliable CSI acquisition is a challenging problem, especially for the downlink where each user has to estimate the channels from large number of BS antennas. To solve this problem, lots of works today have considered massive MIMO of time division duplex (TDD) systems, where CSI can be acquired only at the uplink, and then utilised at both transmission directions based on the assumption of channel reciprocity [2][3]. The uplink channel estimation is therefore critical for TDD massive MIMO since it affects the signal detection at both transmission directions.

Recently, there has been a growing interest in compressive sensing (CS) based channel estimation algorithms [4][5]. By exploiting the inherent sparsity of the MIMO channels, sparse channel estimation can give better estimation performance than conventional schemes such as least square (LS) and minimum mean square error (MMSE) [6]. In [4], a structured subspace pursuit (SSP) algorithm is proposed to estimate the massive MIMO channel through superimposed pilots. In
[5], the block based orthogonal matching pursuit scheme is proposed for massive multiple-input single-output (MISO) systems. Noting that consecutive frames tend to share some common multi-paths even in time-varying propagation environment, some recent schemes exploit the previously estimated channel support to enhance current channel estimation [7]-[11]. [7]-[9] proposed the modified basis pursuit (MBP) utilizing the a priori signal support. In [10], the Auxiliary information based Block Subspace Pursuit (ABSP) is proposed, which utilizes the previous channel support to initialize the estimated support for subspace pursuit algorithm. However, using the a priori support information directly may lead to even worse performance when the channel is fast varying, since the path delays may change from their previous locations. In [11], the authors have further considered the incorrect indices in the previous support, and exclude them adaptively. However, this method does not take into account delay variance and the relationships between the channel supports in adjacent frames.

In this paper, we propose an efficient channel estimation scheme in uplink massive MIMO system. Specifically, inspired by the observation that in massive MIMO systems, the path delays change slowly although the path gains may change quickly [15][16], we first explore the temporal correlation of channel support in uplink TDD massive MIMO systems. And then we propose a method that easily estimates the probabilities of the nonzero path delays in current channel impulse response (CIR) based on the knowledge of the previous CIR. After that we utilize the probabilities as a priori information in the subspace pursuit (SP) algorithm [17][18], and propose the probability-weighted SP (PWSP) algorithm to improve the uplink massive MIMO channel estimation. Moreover, noting that the massive MIMO system may also share common support within one channel matrix [19]. Based on this assumption, an antenna collaborating method is exploited for the proposed channel estimation approach, so that the antennas could share information with their neighbouring antennas to strengthen their beliefs about the locations of the path delays, and thus to further improve the channel estimation performance. Compared with the conventional SP method, the proposed method reduces the pilot overhead significantly while improving the channel estimation performance.

The rest of the paper is organized as follows. We first describe the massive MIMO OFDM system model in Section II. Then the PWSP algorithm is proposed in Section III. Section IV discusses the simulation results while Section V concludes the paper.

Notations: Throughout this paper, boldface lower and upper case symbols represent vectors and matrices, respectively. Operators $^T$, $^H$ and $^\dagger$ represent transpose, Hermite and Moore-Penrose matrix inversion, respectively. $\text{diag}\{x\}$ is the diagonal matrix with $x$ at its main diagonal. $\mathcal{P}(x)$, $|x|$, $\|x\|_p$, $\text{supp}(x)$ and $\text{supp}_K(x)$ denote the probability, cardinality, $\ell_p$-norm, support and largest $K$ elements in the support of $x$, respectively. $x\rangle_k$ denotes the $k$-sparse vector of $x$ which retains the $k$ largest elements of $x$ while setting the rest of the elements to zero.

2. MASSIVE MIMO SYSTEM MODEL

Consider the uplink TDD massive MIMO OFDM system where the BS with $M$ antennas is serving a large number $U$ autonomous single-antenna user terminals (UTs) ($M > U$). During the $i$th OFDM symbol, the CIR between the $m$th BS antenna and one certain UT can be denoted as $h_{i,m} = [h_{i,m}(0), h_{i,m}(1), \ldots, h_{i,m}(L-1)]^T$ with $1 \leq m \leq M$, where $L$ is the maximum delay spread of the CIR. Under the assumption of channel sparsity, only $K$ elements are nonzero in $h_{i,m}$, satisfying $K \ll L$ [4].

Suppose the total number of OFDM subcarriers is $N$, among which $N_p$ subcarriers are randomly
employed to transmit pilot symbols. Thus we can denote the pilot sequence transmitted from one
certain UT to the BS as $x = [x_{p_1}, \ldots, x_{p_j}, \ldots, x_{p_{N_p}}]$, where $p = [P_1, \ldots, P_j, \ldots, P_{N_p}]$ is the
corresponding subcarrier indices of $N_p$ pilots. At the BS side, the pilot vectors received by different
BS antennas are distinct due to the different path gains in each uplink channel. We denote the $N_p \times 1$
received pilot vector as

$$y_{i,m} = \text{diag}(x)Fh_{i,m} + n_{i,m},$$

(1)

where $F$ is a $N_p \times L$ submatrix comprising the $N_p$ rows and the first $L$ columns of the standard
$N \times N$ discrete Fourier transform matrix, and $n_{i,m}$ is the additive white Gaussian noise with zero
mean and unit variance. Moreover, let $Y_i = [y_{i,1}, y_{i,2}, \ldots, y_{i,M}]$ represent the $N_p \times M$ received
pilots for all $M$ BS antennas, $H_i = [h_{i,1}, h_{i,2}, \ldots, h_{i,M}]$ be the $L \times M$ CIR matrix and $N_i =
[n_{i,1}, n_{i,2}, \ldots, n_{i,M}]$ be the $N_p \times M$ channel’s AWGN, the signal model (1) can be equivalently
written as

$$Y_i = \Phi H_i + N_i,$$

(2)

where $\Phi \triangleq \text{diag}(x)F$.

Based on many practical measurements of the channel matrix of massive MIMO systems [11]-
[14], we have the following two important observations:

**Observation I** (Temporal Correlation Between Channel Matrices) In [4], the authors exploit the
temporal correlation of wireless fading channels, whereby the channels exhibit identical common
support in $R$ adjacent OFDM symbols, indicating that $S_{1,m} = \cdots = S_{i,m} = \cdots = S_{R,m}$, where
$S_{i,m} = \text{supp}(h_{i,m}) = \{l : |h_{i,m}(l)| > 0\}_{l=0}^{L-1}$. However, this method suffers from performance
deterioration even when there is only slight changes in the support. In the Fig.3 of [18], the authors
illustrate the CIRs of four adjacent OFDM symbols over the International Telecommunications U-
ion Vehicular B (ITU-VB) channel with 120km/h receiver velocity and signal bandwidth $B=7.56$
MHz [15][16]. As shown in these figures, the path delays are not completely the same, although
they are nearly invariant, since the support of CIR changes over time when there is a relative
movement between the UT and the BS. Nevertheless, note that the path delays just vary within
their vicinity, we can still extract some useful information from the CIR of the previous OFDM
symbol (i.e., $S_{i-1,m}$) to assist current channel estimation.

**Observation II** (Common Support within Channel Matrix) Some existing researches [5]-[11]
indicate that the massive MIMO channels are usually correlated at the BS side. That is to say,
the CIRs of different uplink channels share a common support since the distance between the BS
and UT is far larger than the antenna spacing at the BS side. This can also be proved by the fact
that two channel taps are not resolvable if the time interval of arrival is smaller than $\frac{1}{10B}$ [19],
i.e., $\frac{d_{\text{max}}}{c} < \frac{1}{10B}$, where $d_{\text{max}}$ is the maximum resolvable distance, and $c$ is the speed of light.
Consider a $16 \times 16$ massive MIMO system where $B = 20$MHz and the signal wavelength is
$\lambda = 0.116m$ [22]. The distance between two adjacent antennas is $d = \frac{\lambda}{2} = 0.058m$. Hence,
the maximum distance between two BS antennas is $D = 15d = 0.87m$ ($D$ is a function of the
antenna spacing $d$ and the number of antennas in the array), while $d_{\text{max}} = \frac{c}{10B} = 3.96m$ which
is much larger than $D$. As a result, the BS uplink channels share the common support $S_i$, i.e.,
$S_i = S_{i,1} = \cdots = S_{i,m} = \cdots = S_{i,M}$ [10].

Note that Observation I refers to the temporal shared support between different channel matrices,
while Observation II refers to the spatial common support within one channel matrix. In the
next section, we shall propose an efficient channel estimation method based on these observations.
The relationship between the signal bandwidth $B$ and channel taps in two successive OFDM symbols.

3. **A PRIORI INFORMATION AIDED CHANNEL ESTIMATION ALGORITHMS**

3.1. Estimation of probabilities

Note that the spatial resolution of channel taps is determined by the inverse bandwidth $\frac{1}{B}$ [19], in Fig.1 we demonstrate the relationship between $B$ and the channel taps in two successive OFDM symbols. From Fig.1(a) we can see that the channel tap with delay $\tau$ can be recognized as $h_{i-1,m}(j)$ (i.e., $j \in S_{i-1}$) if $\frac{j}{B} - \frac{1}{2B} \leq \tau \leq \frac{j}{B} + \frac{1}{2B}$. Assume $B = 7.56$MHz [18], then we can obtain the minimum resolvable interval of the channel taps as $\frac{1}{2B} \approx 0.06\mu$s. Let the maximum delay spread $\tau_{\text{max}} = 20\mu$s according to the ITU-VB channel with 120km/h receiver velocity, and the maximum variation rate of the delays between two adjacent OFDM symbols is $\nu = \pm 0.5\%$§, then we can acquire the variation of the delays by $|\text{var}| \leq |\nu \times \tau_{\text{max}}| = 0.1\mu$s [20]. Noting that $|\text{var}| < \frac{1}{B}$, from Fig.1(b) we can see that in the following OFDM symbol, this path tap will (i) be invariant (i.e., $j \in S_i$) if $\frac{j}{B} - \frac{1}{2B} \leq \tau + \text{var} \leq \frac{j}{B} + \frac{1}{2B}$; (ii) move to $j-1$ (i.e., $j-1 \in S_i$) if $\tau + \text{var} < \frac{j}{B} - \frac{1}{2B}$; or (iii) move to $j+1$ (i.e., $j+1 \in S_i$) if $\tau + \text{var} > \frac{j}{B} + \frac{1}{2B}$.

In addition, by assuming that the value of $\text{var}$ occurs uniformly within $|\text{var}| \leq 0.1\mu$s, we have $P(|\text{var}| \leq \frac{1}{2B}) = \frac{1/2B}{\nu \tau_{\text{max}}} = \frac{3}{5}$ and $P(|\text{var}| > \frac{1}{2B}) = \frac{\nu \tau_{\text{max}} - 1/2B}{\nu \tau_{\text{max}}} = \frac{2}{5}$. Therefore, the conditional probabilities of the nonzero channel taps can be expressed as

$$P(h_{i,m}(l) \neq 0 | l \in S_{i-1}) = \frac{3}{5}, \quad P(h_{i,m}(l \pm 1) \neq 0 | l \in S_{i-1}) = \frac{2}{5}. \quad (3)$$

Assume $\varsigma$ is a small integer, indicating the offset of delays from the $(i-1)$th symbol to the $i$th symbol. Based on the previous analysis, we can easily derive the relationship between the maximum $\varsigma$ (denoted as $\varsigma_{\text{max}}$), the signal bandwidth $B$, the maximum delay spread $\tau_{\text{max}}$ and the variation $\nu$ as

$$\frac{\varsigma_{\text{max}}}{B} = \frac{1}{2B} < \nu \tau_{\text{max}} < \frac{\varsigma_{\text{max}}}{B} + \frac{1}{2B}. \quad (4)$$

§Since the channel support depends on the large scale properties of the scattering environment and changes slowly over a very long timescale in practice, we can obtain the knowledge of $\nu$ easily based on the *a priori* knowledge of the propagation environment [11].
Fig. 2. The probability $t_l$ based on the previous channel support $S_{i-1}$.

Thus, we have

$$s_{\text{max}} = \left\lceil \frac{2B\nu \tau_{\text{max}} - 1}{2} \right\rceil,$$

where $\lceil \cdot \rceil$ denotes the ceil function. Moreover, we can further exploit the relationship between $B$, $\tau_{\text{max}}$, $\nu$ and the probability $P(h_{i,m}(l + \varsigma) \neq 0 | l \in S_{i-1})$. It is obvious that if $s_{\text{max}} = 0$, the path delays are invariant, that is

$$\begin{align*}
P(h_{i,m}(l) \neq 0 | l \in S_{i-1}) &= 1, \\
\quad P(h_{i,m}(l) \neq 0 | l \notin S_{i-1}) &= 0.
\end{align*}$$

When $s_{\text{max}} > 0$, we have

$$\begin{align*}
\begin{cases}
P(h_{i,m}(l) \neq 0 | l \in S_{i-1}) = P(|\text{var}| \leq \frac{1}{2B}) \\
\quad P(h_{i,m}(l \pm \varsigma) \neq 0 | l \in S_{i-1}) = P\left(\frac{2\varsigma - 1}{2B} < |\text{var}| \leq \frac{2\varsigma + 1}{2B}\right), \\
\quad P(h_{i,m}(l \pm s_{\text{max}}) \neq 0 | l \in S_{i-1}) = P\left(\frac{2s_{\text{max}} - 1}{2B} < |\text{var}|\right)
\end{cases}
\end{align*}$$

We can then obtain the probabilities that the $l$th delay in $(i - 1)$th symbol is shifted to the position $l \pm \varsigma$ in the $i$th symbol as,

$$P(h_{i,m}(l \pm \varsigma) \neq 0 | l \in S_{i-1}) = \begin{cases}
\frac{1}{2B\tau_{\text{max}}}, & \varsigma = 0 \\
\frac{B\tau_{\text{max}}}{2B\tau_{\text{max}}}, & 0 < \varsigma < s_{\text{max}} \\
1 - \frac{B\tau_{\text{max}} - 1}{2B\tau_{\text{max}}}, & \varsigma = s_{\text{max}} \\
0, & \text{else}
\end{cases}$$

Next, let $T = \text{diag}(t_0, \ldots, t_l, \ldots, t_{L-1})$ where $t_l$ denotes the probability that the $l$th element in $h_{i,m}$ is nonzero, i.e. $t_l = P(h_{i,m}(l) \neq 0)$. Therefore, we have

$$t_l = \sum_{b=0}^{L-1} P(h_{i,m}(l) \neq 0 | b \in S_{i-1}) P(b \in S_{i-1}),$$

where $P(b \in S_{i-1})$ equals 1 or 0 is known a priori according to the previous CIR.

In Fig.2, we illustrate three examples of the probabilities $t_l$ with (a) $s_{\text{max}} = 0, S_{i-1} = \{10\}$, (b) $s_{\text{max}} = 1, S_{i-1} = \{10\}$ and (c) $s_{\text{max}} = 1, S_{i-1} = \{5, 10, 12\}$, respectively. Clearly, in Fig.2(a) the channel is static where the path delays are invariable during the successive symbols. On the other hand, in Fig.2(b) and Fig.2(c) the channels are time-varying, where the path delay may change but only within its neighbourhood due to $s_{\text{max}} = 1$. It is also worth noting that $t_{11}$ is larger than $t_0$ in Fig.2(c) since both of its neighbours are in the support of $h_{i-1}$, which increases the probability that the path delay at $l = 11$ is nonzero.
3.2. Uplink channel estimation with the aid of the probabilities

In order to incorporate the probability $t_l$ into the CS algorithm, we assign a diagonal matrix $W = \text{diag}([w_0, \cdots, w_l, \cdots, w_{L-1}])$ based on the probability $T$. In this paper, we first initial the weights by

$$w_l = 1 + \alpha t_l, \ 0 \leq l \leq L - 1.$$  \hspace{1cm} (10)

where $\alpha \geq 0$ is a user-selected parameter.

Note that the weights $W$ are exploited based on not only the probabilities $T$ but also the parameter $\alpha$, where the value of $\alpha$ determines the degree that the probabilities $T$ may impact the algorithm. Obviously, in the slow time-varying channel where the variance rate $\nu$ is small, a larger number of indices in $\text{supp}(T)$ is correct, which indicates a more reliable $T$. Therefore, by using a larger $\alpha$, we encourage the PWSP algorithm to select indices from $\text{supp}(T)$ to form the support set $\Omega$. On the contrary, in the fast time-varying channel where $\nu$ is large, less indices in $\text{supp}(T)$ are correct. In such cases, we reduce the impact of the a priori information on the support detection by using a relatively smaller $\alpha$. In this way, the a priori support information $\text{supp}(T)$ is utilized adaptively based on the variance rate $\nu$, and hence better recovery performance can be achieved.

Then we propose the PWSP approach designed by exploiting the a priori information and the antenna collaboration based on the conventional subspace pursuit [4]. The details of PWSP is presented in Algorithm 1.

Algorithm 1

1: Input: Received pilot matrix $Y_i$, sensing matrix $\Phi$, previous support $S_{i-1}$, approximated channel sparsity $K = \|S_{i-1}\|_0$.
2: Initialization:
3: The initial residual $V_0 = Y_i$, $W = I + T/\nu$, $\alpha = \frac{1}{100|\nu|}$, $\Gamma = S_{i-1}$ and $k = 1$;
4: while $\|V_k\|_F < \|V_{k-1}\|_F$ do
5: $\Omega = \arg \max_{|\Omega'|=K} \| (W\Phi^H V_k)^{\Omega'} \|_F$;
6: $\Gamma \leftarrow \Gamma \cup \Omega$;
7: $\Gamma \leftarrow \arg \max_{|\Omega'|=K} \| (W^{\Gamma}\Phi^{\Gamma}_i Y_i)^{\Omega'} \|_F$;
8: Set $\hat{H}_i = \Phi^{\Gamma}_i Y_i$ and $\hat{H}_i^{\{(0,\cdots,L-1)\backslash \Gamma\}} = 0$;
9: $k \leftarrow k + 1$;
10: $V_k \leftarrow Y_i - \Phi^H \hat{H}_i$;
end while
12: Output: The estimated CIR vector $\hat{H}_i$.

In Algorithm 1, $\Gamma$ is the estimated support set, $V_k$ is the residual at the $k$th iteration. $\Phi_\Gamma$ and $\Phi^{\Gamma}$ denote the sub-matrix formed by collecting the column vectors and row vectors of $\Phi$ respectively, whose indices belong to $\Gamma$.

Note that the PWSP is developed based on the classical SP algorithm but with two major changes:

1) **Antenna Collaborating.** In the SP algorithm, the $K$ significant path delays are captured by maximizing the correlations between the column vectors of $\Phi$ and the residual measurement, i.e.

---

\footnote{Note that when $M = 1$ and $\nu \to \infty$, Algorithm 1 will reduce to conventional SP [17].}
\[ \Gamma \leftarrow \arg \max_{|\Omega'|=K} \| (\Phi_1^i y_{1,m})_{|\Omega'|} \|, \] where \( y_{1,m} \) is the \( i \)th received pilot sequence on the \( m \)th antenna. As a result, the estimated support may be inaccurate due to the imperfect \( \Phi \) with only approximate orthogonal columns \([17]\). In the proposed PWSP method, the receive antennas collaborate with each other to take advantage of Observation II and estimate the path delays jointly, i.e., \( \Gamma \leftarrow \arg \max_{|\Omega'|=K} \| (\Phi_1^j y_i)_{|\Omega'|} \|_F \). Therefore, according to the law of large numbers, it is more likely to capture the correct path delays since the antennas share information with each other to reach a decision on the most probable support (see more details in our previous work \([10]\)).

**2) Weights Aided.** Suppose that \( j \in S_{i-1}, j = j - 1 \in S_i \) and \( j' \notin S_{i-1} \cup S_i \). For the classical SP algorithm with antenna collaborating, consider an ideal case that all the columns of \( \Phi \) are orthogonal, the \( K \) significant path delays can be easily captured by calculating \( \arg \max_{|\Omega'|=K} \| (\Phi_1^j Y_i)_{|\Omega'|} \|_F \), since \( \| (\Phi_1^j Y_i)_{j} \|_F \gg \| (\Phi_1^j Y_i)_{j'} \|_F \). However, in the practical case, the columns of \( \Phi \) are only locally near-orthogonal. As a result, it may happen that the value of \( \| (\Phi_1^j Y_i)_{|\Omega'|} \|_F \) is smaller than the value of \( \| (\Phi_1^j Y_i)_{j} \|_F \), and this leads to incorrect support detection. On the other hand, in the modified SP (MSP) algorithm \([7]\), the previous channel support (i.e., the \( j \)th path) is used to enhance the channel estimation directly, which is equivalent to adding a large weight on the \( j \)th path. Consequently, MSP could improve the channel estimation performance if the channel is invariant (i.e., \( j \in S_{i-1} \cap S_i \)). However, in the time-varying channel where \( j \notin S_i \), this scheme may lead to bad performance.

Compared with these conventional algorithms, the proposed PWSP exploits more information from the previous channel support. Specifically, according to the analyses in Section 3.1, we have \( w_j > w_{j'} \). Hence, it is more likely that \( \| (W \Phi_1^j Y_i)_{j} \|_F > \| (W \Phi_1^j Y_i)_{j'} \|_F \) although \( \| (\Phi_1^j Y_i)_{j} \|_F \leq \| (\Phi_1^j Y_i)_{j'} \|_F \). Consequently, PWSP is more probable to obtain the correct support. To illustrate this point, we give a simple example in Fig. 3. Assume \( j \in S_{i-1}, j = j - 1 \in S_i \) and \( j' \notin S_{i-1} \cup S_i \), it can be seen from Fig. 3(a) that \( \| (\Phi_1^j Y_i)_{j} \|_F < \| (\Phi_1^j Y_i)_{j'} \|_F < \| (\Phi_1^j Y_i)_{j'} \|_F \). As a result, the conventional SP algorithm will capture the incorrect path delay \( j' \), which leads to performance deterioration. For the MSP scheme in Fig. 3(b), we can see that \( \| (\Phi_1^j Y_i)_{j} \|_F < \| (\Phi_1^j Y_i)_{j'} \|_F < \| (\Phi_1^j Y_i)_{j'} \|_F \). Hence, the MSP will capture the incorrect index \( j \). On the contrary, with the PWSP scheme in Fig. 3(c), we have \( \| (W \Phi_1^j Y_i)_{j} \|_F < \| (W \Phi_1^j Y_i)_{j'} \|_F < \| (W \Phi_1^j Y_i)_{j'} \|_F \). Therefore, the proposed PWSP will still capture the correct index \( j \) into the support set.

![Fig. 3. The impact of the weights on support detection.](image-url)
4. SIMULATION RESULTS

In this section, we conduct simulation studies to evaluate the performance of the proposed channel estimation algorithms. The signal bandwidth is 7.56 MHz locating at a central frequency of 760 MHz. For the uplink transmission, the system parameters are set as: OFDM subcarriers $N = 2048$, number of BS antennas $M = 64$, sparsity of channel $K = 6$, maximum variation $\nu = \pm 0.5\%$ and maximum delay spread $L = 151$ [20]. We assume the path delays of the previous CIR is known a priori*. In order to verify the effectiveness of the proposed scheme, we compare the performance of PWSP with the conventional SP [17], modified SP (MSP), Basis Pursuit (BP) [17], modified Basis Pursuit (MBP) [7], and exact Least Square (exact LS) with perfectly known support of the sparse channel as the performance benchmark [4]. The MSP and MBP are developed based on the conventional SP and BP respectively, with exploitation of the previous support information.

4.1. Channel Estimation Performance Versus Overhead $N_p$

Firstly, in Fig.4 we compare the mean square error (MSE) of the estimated channel versus the number of pilots, with signal-to-noise ratio (SNR) $SNR = 25dB$. From the figure, we observe that the channel estimation improves with the increase of $N_p$, and the proposed PWSP outperforms other conventional methods. Specifically, PWSP achieves a substantial performance gain over the MSP. This is because the a priori channel support $S_{i-1,m}$ is not always correct (i.e., $S_{i-1,m}$ is not identical to $S_{i,m}$). In the MSP, $S_{i-1,m}$ is used directly to aid current channel estimation. As a result, the incorrect a priori information may lead to performance deterioration. On the other hand, PWSP also exploits how reliable the a priori channel support is as illustrated in Section 3.1. Therefore, it could provide more stable performance in time varying channel. Moreover, we observe that PWSP performs closely to the exact LS when $N_p > 23$, which is caused by the fact that $\Phi$ has imperfect but approximate orthogonal columns.

In order to evaluate the stability, we compare the success rate of channel recovery versus the number of pilots $N_p$ in Fig.5. The success rate is defined as the ratio of the number of success trails to the number of total trails, where a trail is recognized to be successful when the MSE of channel estimation is better than $10^{-1}$ [22]. The number of pilots $N_p$ is varied from 1 to 35, while 100 independent trails are implemented for each $N_p$. It can be seen from the figure that PWSP evidently outperforms other conventional methods. Specifically, when 15 pilots are adopted, PWSP can achieve a success rate of 66% while the MSP, the best among the conventional methods, can only reach the success rate of 26%. In addition, PWSP requires 23 pilots to achieve a success rate of 100%, while MSP requires at least 27 pilots, which indicates that the proposed PWSP could lead to a significant improvement (about 17%) in spectral efficiency.

4.2. Channel Estimation Performance Versus Transmit SNRs

In Fig.6 we present the MSE performance comparison of the estimated channel versus the transmit SNR with $N_p = 23$. From the figure, we can see that the proposed PWSP can achieve enormous MSE gains over other conventional methods, and approach the performance bound (i.e., exact LS) in higher SNR regions.

*This assumption is reasonable since we could always employ more pilots in the first OFDM symbol to obtain accurate CIR, and then reduce the number of pilots by using PWSP.
4.3 Channel Estimation Performance With Different Number of BS Antennas $M$

In Fig. 7, we set $K = 12$ and compare the MSE performance of the PWSP scheme with different number of BS antennas for both $N_p = 16$ and $N_p = 32$. It is evident from both Fig. 7(a) and Fig.
Fig. 7. MSE performance comparisons of PWSP with different number of BS antennas $M$.

Fig. 8. MSE performance comparisons of PWSP versus the variation rate $|\nu|$.

7(b) that the channel estimation performance improves as $M$ increases. This is because that the BS antennas could coordinate with each other and reach a decision on the most probable channel support for channel estimation (i.e., Line 5 and Line 7 of Algorithm 1) [10]. Thus, according to the law of large numbers, more coordinated BS antennas could provide more accurate support detection, as long as Observation II is satisfied.

4.4. Channel Estimation Performance Versus the Variation Rate $|\nu|$

In order to evaluate the impact of $\nu$ on the proposed PWSP, we set $N_p = 23$, $M = 64$ and illustrate the MSE comparisons with different $|\nu|$ in Fig.8. It is obvious that both of MSP and PWSP achieve a substantial performance gain over the SP when the channel is invariant ($|\nu| = 0$). However, the performance of the MSP degrades severely when $|\nu| > 0.$ On the contrary, the PWSP is relatively stable and still provides performance gains over the SP and MSP, demonstrating the robustness of the proposed PWSP scheme in fast time-varying channel$^8$.

5. CONCLUDING REMARKS

This paper considers the uplink channel estimation for massive MIMO system. By extracting the $a$ priori information from the previous CIRs and exploiting the spatial correlation of massive MIMO

$^8$Note that PWSP is reduced to the conventional SP when $|\nu| = \infty$. 

10
channels, the PWSP algorithm is proposed to efficiently perform the channel estimation with only few pilots. Simulation results have shown that the proposed scheme could achieve higher spectral efficiency as well as more reliable performance. It is also demonstrated that PWSP could still achieve performance gain in fast time-varying channel.

6. Acknowledgements

The authors would like to thank the support by “National Natural Science Foundation of China (61201199)”.

7. References


