# Estimation of NON-WSSUS Channel for OFDM System: Exploiting Support Correlations through a Novel Adaptive Weighted Predict-Re-Estimate L1 Minimization Approach

Chen Wang[1], Yong Fang[1], and Jin-Yan Li [2]
[1] Key Laboratory of Specialty Fiber Optics and Optical Access Networks, School of Communication and Information Engineering, Shanghai University, Shanghai 200444, Shanghai, China
[2] Advanced Analytics Institute, Faculty of Engineering & IT, University of Technology, Sydney, 81 Broadway, Broadway NSW 2007, Australia
Email: cwangsh@shu.edu.cn; yfang@staff.shu.edu.cn; Jinyan.Li@uts.edu.au

***Abstract*** —It is challenging to estimate the wireless channel of the Orthogonal Frequency-Division Multiplexing (OFDM) broadband system under a changing communication environment. The difficulty is mainly attributed to this wireless channel's Non Wide Sense Stationary Uncorrelated Scattering (Non-WSSUS) which has an implication that the delay and Doppler shift of such a channel are non-stationary and correlated. A Non-WSSUS channel is very different from the classical time-varying channel with constant delay and Doppler shift. In this paper, we propose an estimation method for the Non-WSSUS Channel Impulse Response (CIR) of the OFDM system. Based on the sparsity property of the delay-Doppler spread function, the delay and Doppler shift of Non-WSSUS channel can be extracted through a Compressive Sensing (CS) approach. Then a novel CS algorithm referred as Pre-Re L1 is proposed. The proposed CS algorithm exploits the correlations of the sparse supports to obtain adaptive weights for $L_1$ minimization. Numerical Simulation results show that the proposed CS method improves the performance of the Non-WSSUS wireless channel estimation.

***Index Terms***—OFDM, Non-WSSUS channel estimation, Compressive Sensing (CS), adaptive weighed $L_1$ minimization

## I. INTRODUCTION

Applying the Compressive Sensing (CS) idea to the estimation of wireless channel has been gaining wide interests in recent years [1]-[4]. In practice, the wireless channels in several scenarios have the property of intrinsic sparsity, i.e., only a few channel gains are dominant. CS algorithms have been proposed to improve the estimation performance for wireless channels with specific characteristics, such as block fading, time varying, and shallow-water acoustic channels. In this work, we explore the CS approach for estimating and tracking the Non-WSSUS channel.

Non-WSSUS wireless channel can be observed in many modern commutation scenarios, such as vehicle to vehicle (V2V) communications [5], and 5G massive MIMO channels [6]. The essential cause of the Non-WSSUS channel is the dynamic change of the communication environment, especially the movement of the transmitter and receiver [7]. This dynamic condition gives rise to that the channel scatterers change over time in the sense that the delay, and Doppler shift of each scatterer all change dynamically and follow some certain statistical characteristics, which leads to the Non-WSSUS channel. OFDM as an outstanding technique of the modern wireless communication system has been widely used in 3G and 4G. But the Non-WSSUS characteristic affects each OFDM symbol at the physical layer, which greatly impacts the performance of OFDM system. So the estimation of the channel variation for each OFDM symbol is very crucial [8]. Typically, the delay-Doppler spread function follows a sparse structure [9] due to the large bandwidths of an OFDM broadband system. Moreover, the sparse delay-Doppler spread function of Non-WSSUS channel exhibits non-stationary support transitions due to the movement of the transmitter and receiver. Luckily, this non-stationary transition is somewhat predictable. Because the support of the delay-Doppler spread function is correlated in this scenario. Combing the sparsity with the Non-WSSUS characteristic, the Non-WSSUS channel estimation for the OFDM system becomes a problem of recovering the non-stationary sparse vector with correlated support transitions.

The work presented here mainly focus on the formulation of an adaptive weighted $L_1$ minimization algorithm for Non-WSSUS CIR estimation in the OFDM system, which exploits the correlations of the support transitions in the sparse delay-Doppler domain. The major contributions are summarized as follows:

*1) Estimation of the CIR of Non-WSSUS channel*

The problem of estimating the CIR of Non-WSSUS channel is solved by utilizing the sparse delay-Doppler delay-Doppler-spread function. While the in work of [10]-[11], only the basic definition of Non-WSSUS

Manuscript received July 21, 2015; revised January 29, 2016.
This work was supported by National Science Foundation (NSF) of China (No. 61271213) and the Ph.D. Programs Foundation of Ministry of Education of China (No. 20133108110014)
Corresponding author email: yfang@staff.shu.edu.cn.
doi: 10.12720/jcm.11.2.149-156

channel was proposed, and [11] only suggested that the wireless channel can be analyzed by a time-frequency approach. The works in [7], [12], [13] mainly focused on modeling the Non-WSSUS channel from measurement.

*2) Algorithm designs to exploit correlated support transitions of the sparse channel*

In this work, a novel adaptive weighted $L_1$ minimization approach is proposed to estimate the sparse Non-WSSUS channel by exploiting the correlated support transitions. By utilizing the correlation of the supports, the weights of $L_1$ minimization at time $t$ are predicted by the information of supports at time $t-1$. In the work of [1], [2], channel is assumed to be constant. In the work of [3], the time-varying channel estimation was performed by using the joint PDF of the $K$ received OFDM signals, and the block-sparsity channel was recovered by exploiting inner-correlations. In [14], the dynamic sparse signal is estimated by approximate message passing, and the model parameters are learned by using the available prior knowledge. However, in our scenario, the channel at next time $t+1$ is unknown to the estimator, which implies that the previous approaches in [3], [14] are inapplicable here. In [4], the correlated change of the channel is not concerned.

*3) Adaptive designs to add algorithm robust*

While the $L_1$ minimization algorithm proposed here is related to [15]-[20], it capitalizes the bias of predicted supports to obtain the appropriate weights over time. Then the appropriate weights are assigned to the predicted supports, followed by a decision step to make the weights be adaptive to the estimation result. The re-estimation is then performed with the adaptive weights. The proposed algorithm formulates a predict-re-estimate $L_1$ minimization framework. By introducing the Re-estimation phase, robust is added to the estimation result with the adaptive weights. This is a new approach, and it has not been considered at earlier studies.

Important notations used in this paper include: $\|\mathbf{x}\|_1$ and $\|\mathbf{x}\|_2$ denote the $L_1$ and $L_2$ norm of vector $\mathbf{x}$, respectively. Bold symbols are reserved for vectors and matrices. Particularly, $\mathbf{I}_L$ denotes the identity matrix with the size $L\times L$. Term $\text{vec}(\mathbf{x})$ denotes the vector of $\mathbf{x}$, $\mathbf{x}^\dagger$ is the pseudo inverse of $\mathbf{x}$, for a vector $\mathbf{x}$, its $i$th coordinate is denoted by $x_i$. The support of $\mathbf{x}$, $T(\mathbf{x})$ is the set of indices at which $\mathbf{x}$ is nonzero. For a subset $T\subseteq\{1,2,\cdots n\}$, $|T|$ denotes its cardinality, and $T^c$ denotes its complement. The symbols $\bigcup$, $\bigcap$, and $\backslash$ denote set union, set intersection, and set difference respectively (recall $T_1\backslash T_2 := T_1\bigcap T_2^c$). $\otimes$ denotes the Kronecker product. $x(t\,|\,t-1)$ means the predicted $x(t)$, given the measurement at time $t-1$. While, $x(t\,|\,t)$ stands for $x(t)$, given the measurement at time $t$.

## II. Problem Formulation

### A. A sparse Channel Model for OFDM

The orthogonal short-time Fourier (STF) basis waveforms of OFDM signaling to counteract the time selectivity of doubly-selective channels are used in the OFDM system, and $L$ delays, $K$ (one-sided) Doppler shifts are considered here. With these considerations, the corresponding discrete OFDM channel via "virtual representation" [9] can be written as:

$$\mathbf{H}_{n,m}=\sum_{l=0}^{L-1}\sum_{k=-K}^{K}h(l,k)e^{j2\pi\frac{k}{N_t}n}e^{-j2\pi\frac{l}{N_f}m}=\mathbf{u}'_{f,m}\mathbf{H}_\alpha\mathbf{u}_{t,n} \quad (1)$$

where

$$\mathbf{u}_{t,n}\triangleq\left\{\frac{1}{\sqrt{N_t}}(W_{N_t}^{-kn}\quad W_{N_t}^{(-k+1)n}\cdots\cdots W_{N_t}^{kn})\right\},\quad n=0,\cdots,N_t-1,$$

$$\mathbf{u}'_{f,m}\triangleq\left\{\frac{1}{\sqrt{N_f}}(1\quad W_{N_f}^{1m}\cdots\cdots W_{N_f}^{(L-1)m})\right\}\qquad m=0,\cdots,N_f-1$$

with $W_{N_0}=e^{-j\frac{2\pi}{N_0}}$, $N_0$ is the total number of STF basis waveforms, $N_t$ is the number of the time separation of the STF basis, and $N_f$ is the number of the frequency separation of the STF basis. Readers are referred to [21, Section Ⅳ] for the details.

According to (1), the sparse delay-Doppler-spread function $H_\alpha(l,k)$ can be written in the matrix form as $\mathbf{H}_\alpha$ [21]:

$$\mathbf{H}_\alpha=\begin{bmatrix} h_{0,-K} & h_{0,-K+1} & \cdots & h_{0,K}\\ h_{1,-K} & h_{1,-K+1} & \cdots & h_{1,K}\\ \vdots & \vdots & \vdots & \vdots\\ h_{L-1,-K} & h_{L-1,-K+1} & \cdots & h_{L-1,K}\end{bmatrix} \quad (2)$$

**Remark 1:** $\mathbf{H}_\alpha$ is the key to double-selective channel. The knowledge of the channel delays and Doppler shifts is completely determined by this matrix. The value of the entry of matrix $\mathbf{H}_\alpha$ is the channel gain at discrete delay $l$ and Doppler shift $k$. The support of this matrix is the "position" of the discrete channel delays and Doppler shifts with a dominant (or non-zero) channel gain.

**Remark 2:** Note that $\mathbf{H}_\alpha$ is obtained by uniformly sampling the delay-Doppler space at the Nyquist rate [9]. So though the supports $L$ and $K$ are integer, they are corresponding to real-valued delay and Doppler shift of the double- selective channel. Readers are referred to [9, Section Ⅱ.B] for the details

Let $y_{n,m}(t)$ is the received pilot signal at time $t$. By uniformly sampling pilot symbols at random (without replacement) [9], we can obtain the transmitted pilot signal, i.e., $\mathbf{X}=\left\{\sqrt{\varepsilon/N_r}(\mathbf{u}'_{t,n}\otimes\mathbf{u}'_{f,m}):(n,m)\in S_r\right\}$, where the total number of pilots is $N_r$, $S_r$ is set of indices of pilots, and $\varepsilon$ is the system transmit energy budget. Finally, the sparse channel delay-Doppler-spread function estimation follows equation:

$$\mathbf{y}(t) = \mathbf{X}\mathbf{H}(t) + \mathbf{W}_t \tag{3}$$

where $\mathbf{y}(t)$ is obtained by stacking the received training symbols $\{y_{n,m}(t)\}$ into an $N_r$ -dimensional vector, $\mathbf{H}(t) = \mathrm{vec}(\mathbf{H}_\alpha(t))$ , and $\mathbf{W}_t \sim \mathcal{CN}(0, \sigma_{obs}^2 \mathrm{I}_{N_r})$ is the channel observation noise.

Due to the sparsity of delay-Doppler spread function, $\mathbf{H}(t)$ is an $s(t)-sparse$ vector, which means that the number of dominant (or non-zero) elements of $\mathbf{H}(t)$ is $s(t)$ . Note that though $s(t)$ changes over time $t$ , the dimension of $\mathbf{H}(t)$ is bounded by $L \bullet (2K+1)$ .

*B. Non-Stationary Support Transition with Correlation*

As described in [6], the non-stationary process of a Non-WSSUS channel is due to the updates of geometry relationships of scatterers from $t$ to $t + \Delta t$ . Specially in [7], the Non-WSSUS characteristic is formed by the two scatterers from the same reflection. Within the time interval $\Delta t$ , the maximum change in delay is given by $\delta_\tau = \frac{v}{c} \bullet \Delta t$ . Assuming that the moving receiver passes by a reflecting surface at distance $d$ , the maximum change in Doppler shift is given by $\delta_v = -\frac{v^2 f}{cd} \bullet \Delta t$ , where $c$ is the speed of light, $f$ is the carrier frequency, $v$ is the motion velocity. Then the transition of the survived scatterers' delay $\tau(t)$ and Doppler shift $f_D(t)$ can be approximated as a first-order model:

$$\begin{bmatrix} \tau(t+\Delta t) \\ f_D(t+\Delta t) \\ \delta_\tau(t+\Delta t) \\ \delta_v(t+\Delta t) \end{bmatrix} \approx \begin{bmatrix} \mathbf{I}_2 & \mathbf{I}_2 \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix} \begin{bmatrix} \tau(t) \\ f_D(t) \\ \delta_\tau \\ \delta_v \end{bmatrix} + \mathbf{n} \tag{4}$$

where $\mathbf{n}$ is the small random change on $\delta_\tau$ and $\delta_v$ . It is worth noting that the value change of the delay and Doppler shift is equivalent to the support change of matrix (2), e.g., the dominant (or non-zero) channel gain change from discrete delay $l$ to delay $l+1$ (or change from discrete Doppler shift $k$ to $k+1$ ). The newly generated scatterers are not concerned here.

Note that the time interval for the channel estimator is constant. Combing this condition with the OFDM sparse channel model (3), the corresponding discrete form of (4) is:

$$\begin{bmatrix} T(\mathbf{H}(t)) \\ \mathbf{V}(t) \end{bmatrix} = \mathbf{A} \begin{bmatrix} T(\mathbf{H}(t-1)) \\ \mathbf{V}(t-1) \end{bmatrix} + \mathbf{n} \tag{5}$$

where $\mathbf{V}(t-1) = \begin{bmatrix} \boldsymbol{\delta}_l \\ \boldsymbol{\delta}_k \end{bmatrix} \in \mathbb{R}^{s(t)\times 1}$ is the vector of the maximum change of the discrete delay and Doppler shift, related to the speed of motion of the scatterers, at time $t-1$. $\mathbf{A} = \begin{bmatrix} \mathbf{I}_{s(t)} & \mathbf{I}_{s(t)} \\ \mathbf{0} & \mathbf{I}_{s(t)} \end{bmatrix}$, $\begin{bmatrix} T(\mathbf{H}(t-1)) \\ \mathbf{V}(t-1) \end{bmatrix} \in \mathbb{R}^{2s(t)\times 1}$. $\mathbf{n} = \begin{bmatrix} \mathbf{0} \\ \mathbf{Q} \end{bmatrix}$ is the small random change on $\mathbf{V}$ , $\mathbf{Q}$ is Gaussian distribution with expectation zero and covariance $Q$ . Here we bound the transition of the supports in the size of $L \bullet (2K+1)$ .

Combing equation (5) with equation (3), the sparse non-stationary channel model with support transitions for OFDM system can be formulated as:

$$\begin{cases} \mathbf{y}(t) = \mathbf{X}\mathbf{H}(t) + \mathbf{W}_t \\ \begin{bmatrix} T(\mathbf{H}(t)) \\ \mathbf{V}(t) \end{bmatrix} = \mathbf{A} \begin{bmatrix} T(\mathbf{H}(t-1)) \\ \mathbf{V}(t-1) \end{bmatrix} + \mathbf{n} \end{cases} \tag{6}$$

The upper equation of (6) is the measure model, and the lower equation of (6) is the state transition model. In the next section, we propose a CS-based algorithm to estimate this non-stationary sparse vector by exploiting the correlations of support transitions.

## III. Non-WSSUS Channel Estimation Algorithm

Our goal is to estimate the sparse $\mathbf{H}(t)$ through a small number of observations ( $N_r \ll L \bullet (2k+1)$ ). As a method to promote sparsity, the weighted $L_1$ minimization problem can be formulated as:

$$\underset{\mathbf{H}(t)}{\text{minimize}} \sum_i w_i(t) \left\| H_i(t) \right\|_1 \quad \text{s.t.} \mathbf{X}\mathbf{H}(t) = \mathbf{y}(t) \tag{7}$$

where $i$ is the coordinate.

It has been recently observed [22], [23] that the Alternating Direction Method of Multipliers (ADMM) is a powerful tool to tackle the $L_1$ norm problem. In this paper, the impact of the weights to the solution of (7) is generated through the ADMM method.

*A. Weighted $L_1$ Minimization via AMDD*

In ADMM the weighted $L_1$ minimization problem can be written as:

$$\begin{aligned} &\underset{\mathbf{H}(t)}{\text{minimize}} \ \sum_i w_i(t) \left\| Z_i(t) \right\|_1 + f(\mathbf{H}(t)) \\ &\text{subject to} \ \ \mathbf{H}(t) - \mathbf{Z}(t) = 0 \end{aligned} \tag{8}$$

where $f$ is the indicator of $\{\mathbf{H}(t) \in \mathbf{R}^n \,|\, \mathbf{X}\mathbf{H}(t) = \mathbf{y}(t)\}$ .

The augmented Lagrangian [23, Section 5] corresponding to the optimization problem (using the scaled dual variable) (8) is given by:

$$\begin{aligned} L_\rho(\mathbf{H}(t), \mathbf{Z}(t), \mathbf{u}(t)) = &\sum_i w_i(t) \left\| Z_i(t) \right\|_1 \\ &+ f(\mathbf{H}(t)) \\ &+ (\rho/2) \left\| \mathbf{H}(t) - \mathbf{Z}(t) + \mathbf{u}(t) \right\|_2^2 \end{aligned} \tag{9}$$

To find the minimum of the constrained problem (8), the ADMM algorithm uses a sequence of iterations:

$$\mathbf{H}^{k+1}(t) = \arg\min L_\rho(\mathbf{H}(t), \mathbf{Z}^k(t), \mathbf{u}^k(t)) \tag{10}$$

$$\mathbf{Z}^{k+1}(t) = \arg\min L_\rho(\mathbf{H}^{k+1}(t), \mathbf{Z}(t), \mathbf{u}^k(t)) \quad (11)$$

$$\mathbf{u}^{k+1}(t) = \mathbf{u}^k(t) + \mathbf{H}^{k+1}(t) - \mathbf{Z}^{k+1}(t) \quad (12)$$

Until $\|\mathbf{H}^{k+1}(t) - \mathbf{Z}^{k+1}(t)\|_2 \le \varepsilon_s$, and $\|\mathbf{Z}^{k+1}(t) - \mathbf{Z}^k(t)\|_2 \le \varepsilon_s$, where $\varepsilon_s = 10^{-3}$ is stopping criterion.

The rationale behind using ADMM is that we could effectively split the original non-differentiable problem into a "$H-$ minimization step" (10) and a "$Z-$minimization step" (11).

*1) $H$ -minimization step*

Completing the squares with respect to $\mathbf{H}(t)$ in the augmented Lagrangian (9), the $H-$minimization step is to:

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{H}(t)) \\ & +(\rho/2)\|\mathbf{H}(t) - \mathbf{Z}^k(t) + \mathbf{u}^k(t)\|_2^2 \end{aligned} \quad (13)$$

The answer to (13) is:

$$\begin{aligned} \mathbf{H}^{k+1}(t) := & (\mathbf{I} - \mathbf{X}^T(\mathbf{X}\mathbf{X}^T)^{-1}\mathbf{X})(\mathbf{Z}^k(t) - \mathbf{u}^k(t)) \\ & + \mathbf{X}^T(\mathbf{X}\ \mathbf{X}^T)^{-1}\mathbf{H}\ (t) \end{aligned} \quad (14)$$

*2) $Z$ -minimization step*

Completing the squares with respect to $\mathbf{Z}$ in the augmented Lagrangian (9), the $Z$ -minimization step is to:

$$\text{minimize} \sum_i w_i(t)\|Z_i(t)\|_1 + (\rho/2)\|Z_i(t) - V_i^k(t)\|_2^2 \quad (15)$$

where $V_i^k = H_i^{k+1} + u_i^k$. We note that (11) is decomposed into sub-problems expressed in terms of the individual element of $\mathbf{Z}(t)^k$ via (15). The unique solution to (15) is given by the soft thresholding:

$$Z_i^*(t) = \begin{cases} (1 - \frac{a}{|V_i|})V_i & |V_i(t)| > a \\ 0 & |V_i(t)| < a \end{cases} \quad (16)$$

where $a = w_i(t)/\rho$. In particular, $Z_i^*(t)$ is set to zero if $|V_i(t)| < w_i(t)/\rho$, implying that a more aggressive scheme for driving $Z_i^*(t)$ to zero can be obtained by increasing $w_i(t)$, vice versa.

### B. Exploiting Correlations of the Support Transitions by Choosing $w_i$

Through (15), it can be discovered that $w_i(t)$ is essentially the bias towards driving $Z_i^*(t)$ to zero in the sense that if $w_i(t) = 1$, $Z_i^*(t)$ are more likely to be zero (note that when $w_i(t) = 1$, the problem (8) is the $L_1$ minimization), if $w_i(t) = 0$, $Z_i^*(t)$ are more likely to be the dominant element (when $w_i(t) = 0$, the $L_1$ norm penalty is not concerned).

Let the predicted support of $\mathbf{H}(t)$ at time $t-1$ be $T(\mathbf{H}(t|t-1))_p$, which can be obtained through (5). Here we make the assumption that $T(\mathbf{H}(t|t-1))_p$ is highly accurate to the true supports of $\mathbf{H}(t)$. Since the support transitions are correlated here, this assumption is valid. Therefore, the proposed predict-re-estimate algorithm, referred as **Pre-Re L1,** mainly includes two phases:

1) Prediction phase: At time $t-1$, given the $T(\mathbf{H}(t-1))$, we first obtain the predicted support of $\mathbf{H}(t)$, i.e., $T(\mathbf{H}(t|t-1))_p$, through (5). Solve (7) using:

$$w_i(t|t-1) = \begin{cases} w_i = 0 & if \quad i \in T(\mathbf{H}(t|t-1))_P \\ 1 & if \quad i \notin T(\mathbf{H}(t|t-1))_P \end{cases} \quad (17)$$

Since the predicted support $T(\mathbf{H}(t|t-1))_p$ is assumed to be highly accurate, here we set $w_i = 0$.

**Remark 3:** By introducing th**e** prediction phase, the correlation of the non-stationary supports are exploited by choosing the value of $w_i(t|t-1)$.

2) Add robust by Re-estimation phase with adaptive weights: At time $t$, we solve (7) using $w_i(t|t-1)$, and obtain the estimate of the support, i.e., $T(\mathbf{H}(t|t))_E$. Compute the ratio $\alpha(t)$ by: $\alpha(t) = \dfrac{|T(\mathbf{H}(t|t))_E \cap T(\mathbf{H}(t|t-1))_P|}{|T(\mathbf{H}(t|t))_E|}$. As shown in [15], if $\alpha(t) \ge 0.5$, i.e., the estimated supports are highly accurate, then the weighted $L_1$ minimization (7) using weights (17) achieves the smallest error bound constants. While, if $\alpha(t) < 0.5$, [15] proposes to set $w_i$=0.5 to add robustness to the $L_1$ weighted problem. So if $\alpha(t) < 0.5$, we re-solve (7) as the way of (18):

$$w_i(t|t) = \begin{cases} w_i = 0.5 & if \quad i \in T(\mathbf{H}(t|t-1))_P \\ 1 & if \quad i \notin T(\mathbf{H}(t|t-1))_P \end{cases} \quad (18)$$

**Remark 4:** By introducing th**e** Re-estimation phase, the signal information at time $t$ is used. The weights are adaptive to the estimation result at time $t$. The adaptive weights at time $t$ change the bias towards the predicted support obtained at time $t-1$. Since the predicted supports are not good enough to be the non-zero supports, we should drive them to zero by increasing $w_i$, and here we propose to set $w_i = 0.5$.

In [18], a time-varying weight is also used. By using the similar idea, we can set $w_i = \dfrac{|T(\mathbf{H}(t|t))_E \setminus T(\mathbf{H}(t|t-1))_P|}{|T(\mathbf{H}(t|t))_E|}$, when $\alpha(t) < 0.5$. This approach is intuitive, since the more $T(\mathbf{H}(t|t))_E$ are in the true support, i.e., $T(\mathbf{H}(t|t-1))_P$, the smaller $w_i$ will be. But this choice is lack of theoretical analysis, and it is not robust to the noisy environment, which will be shown in Section IV.

### C. Improved Support Estimation

The estimation of the supports can be improved by the Add-LS procedure [16], [18]. The Add-LS procedure first sets a thresh to detect additions to the support, i.e., $T \leftarrow Thresh(\mathbf{x}, \lambda)$ which means $T = \{i : |x_i| \geq \lambda\}$; then compute an Least Squares (LS) estimate on that support, i.e., $\tilde{\mathbf{x}} \leftarrow LS(\mathbf{y}, \mathbf{A}, T)$ which means $\tilde{\mathbf{x}}_T = (\mathbf{A}_T)^\dagger \mathbf{y}, \tilde{\mathbf{x}}_{T^c} = 0$. The choice of add threshold parameter $\alpha_{add}$ in the proposed **Pre-Re L1** is given by $0.25\sqrt{\|\mathbf{H}_r\|^2 / L(2K+1)}$, which is also used in [18]. In addition to Add-LS procedure, when $\alpha(t) < 0.5$, an expansion of supports size $k$, i.e., $k = \beta |T(\mathbf{H}(t\,|\,t-1))_P|$ can be implemented to make it more likely to take supports in the complement of set $T(\mathbf{H}(t\,|\,t-1))_P$ into set $T(\mathbf{H}(t\,|\,t))_E$. The expansion procedure is denoted as $T \leftarrow Expend(\mathbf{x}, k)$ which means $T$ contains the $k$ largest magnitude element of $\mathbf{x}$, as a result, the number of support of $\mathbf{x}$ is expanded to $k$. Since a small increase is good enough, a $\beta \in (1, 1.5]$ is suggested ($\beta = 1.4$ in our experiment). We note that the expansion step also utilizes the correlation of the supports by taking $T(\mathbf{H}(t\,|\,t-1))_P$ as a reference support to the true support at time $t$.

This predict-re-estimate $L_1$ minimization referred as **Pre-Re L1** for the sparse Non-WSSUS channel by exploiting its support correlation is detailed in **Algorithm 1**.

---

**Algorithm 1: Pre-Re L1**

---

**Input**: $\mathbf{y}(t)$, $\mathbf{X}$ ;**Output**: $\mathbf{H}(t)$

Definition: $T \leftarrow Thresh(\mathbf{x}, \lambda)$ means $T = \{i : |x_i| \geq \lambda\}$

$T \leftarrow Expend(\mathbf{x}, k)$ means $T$ contains the $k$ largest magnitude elements of $\mathbf{x}$

$\tilde{\mathbf{x}} \leftarrow LS(\mathbf{y}, \mathbf{A}, T)$ means $\tilde{\mathbf{x}}_T = (\mathbf{A}_T)^\dagger \mathbf{y}, \tilde{\mathbf{x}}_{T^c} = 0$

**1. Initialization**:

**1.1.** Solve (8) with $w_i(t) = 1$ at time $t = 1$.

**1.2.** Add-LS procedure:

(a) $\tilde{T}(1) \leftarrow Thresh((\mathbf{H}(1))_E, \alpha_{add})$

(b) $\mathbf{H}(1) \leftarrow LS(\mathbf{H}_r(t), \mathbf{U}_r, T(\mathbf{H}(1))_E)$

**For** $t = 2, 3 \ldots$ do

2. **Prediction**:

**2.1** Use (5) to get the predicted supports $T(\mathbf{H}(t\,|\,t-1))_P$

**2.2** $w_i(t\,|\,t-1) = \begin{cases} w_i = 0 & if \quad i \in T(\mathbf{H}(t\,|\,t-1))_P \\ 1 & if \quad i \notin T(\mathbf{H}(t\,|\,t-1))_P \end{cases}$

**3. Re-estimation:**

**3.1** solve (7) using $w_i(t\,|\,t-1)$ to obtain $\mathbf{H}(t\,|\,t)$

**3.2** calculate $\alpha(t) = \dfrac{|T(\mathbf{H}(t\,|\,t))_E \cap T(\mathbf{H}(t\,|\,t-1))_P|}{|T(\mathbf{H}(t\,|\,t))_E|}$

**3.3 If** $\alpha(t) > 0.5$

(a) $T(\mathbf{H}(t))_E \leftarrow Thresh(\mathbf{H}_r(t), \alpha_{add})$

(b) $\mathbf{H}(t) \leftarrow LS(\mathbf{H}_r(t), \mathbf{U}_r, T(\mathbf{H}(t))_E)$

**Else If** $\alpha(t) < 0.5$

(a) Re-solve (7) using $w_i(t\,|\,t) = \begin{cases} w_i = 0.5 & if \quad i \in T(\mathbf{H}(t\,|\,t-1))_P \\ 1 & if \quad i \notin T(\mathbf{H}(t\,|\,t-1))_P \end{cases}$, to obtain $\mathbf{H}(t\,|\,t)$

(b) $T(\mathbf{H}(t\,|\,t))_E \leftarrow Expend(\mathbf{H}(t\,|\,t))_E, \beta |T(\mathbf{H}(t\,|\,t))_P|)$

(c) $\mathbf{H}(t) \leftarrow LS(\mathbf{H}_r(t), \mathbf{U}_r, T(\mathbf{H}(t\,|\,t))_E)$

(d) $\mathbf{H}(t\,|\,t) \leftarrow Thresh(\mathbf{H}(t), \alpha_{add})$

**End If**

**End for**

---

## IV. Simulation Results

Our experimental results are obtained by applying the proposed **Pre-Re L1** algorithm to the estimation of the sparse Non-WSSUS wireless channel. To assess the estimation performance of the **Pre-Re L1** algorithm, a comparison against a number of popular sparse optimization techniques for time-varying systems has been conducted. These techniques include Kalman filtering by compressive sensing (**KF-CS**) [19], Compressive Sensing on the least squares residual (**LS-CS**) [20], ADMM-based L1 minimization algorithm without knowing the prior information of supports (**L1-with no prior**), and a modified Compressive Sensing with partially known support (**modified-CS**) [16].

We consider a sparse wireless channel vector of 180 dimensions. Only 12 elements are dominant (non-zero), having arbitrary support sets. The number of pilots is 45. We assume that the maximum change of the discrete delay and Doppler shift is 0.25, and the covariance of the small random change is $1 \times 10^{-3}$. The time length is 50. Gaussian noise is added to the channel, whose variance is adjusted according to the SNR level of each experiment. Here, $\mathrm{SNR} = 10\log 10(\dfrac{\mathrm{signal}}{\mathrm{noise}})$. The estimation performance of the algorithms is measured in terms of the Normalized Mean Square Error (NMSE), and its unit is dB. When come across the non-integer value of supports, we make them to their nearest integer.

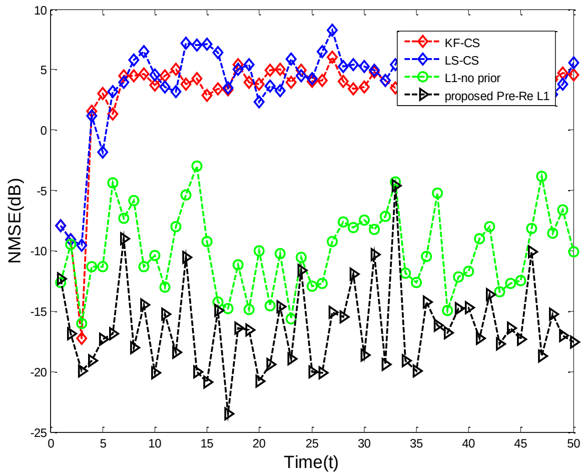


Fig. 1. NMSE of recovering $\mathbf{H}(t)$.

The first experiment is used to demonstrate the estimation performance of the Pre-Re L1 algorithm. The SNR is set to 20dB. As illustrated in Fig. 1, it can be seen that the proposed algorithm has the smallest error, and it is stable. **KF-CS** and **LS-CS** give larger errors in this scenario. This may be caused by the relatively quick support change of $\mathbf{H}(t)$. The comparison between **Pre-Re L1** and **L1-with no prior** illustrates that the proposed algorithm which exploits the correlation of the support transition indeed increases the performance of sparse non-stationary correlated vector estimation.

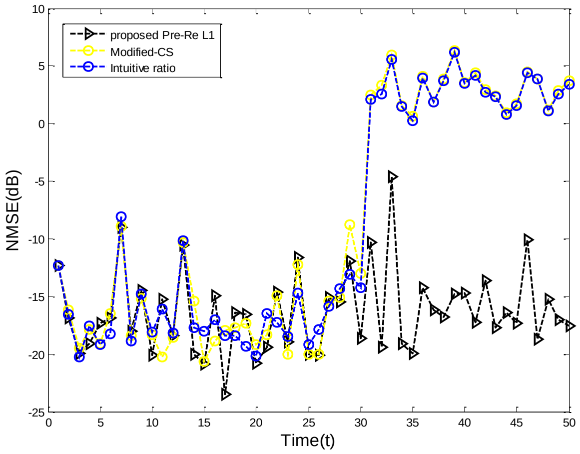


Fig. 2. NMSE of different weight schemes, SNR=20dB .

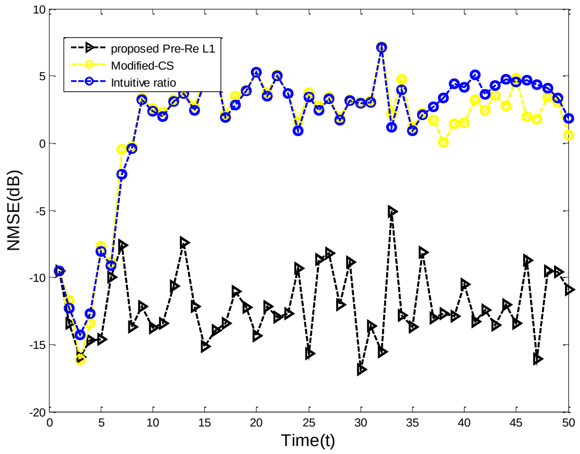


Fig. 3. NMSE of different weight schemes, SNR=16dB .

The next experiments are used to illustrate that the proposed **Pre-Re L1** adds robustness to the $L_1$ weighted problem. We compare the NMSE of recovering $\mathbf{H}(t)$ with different schemes of choosing weight $w_i(t\,|\,t)$. The experiments are under SNR=20dB , SNR $\approx$ 16dB, and SNR $\approx$ 15dB. The simulation result is shown in Fig. 2, Fig. 3, and Fig. 4. Note that the weighted L1 minimization with $w_i(t\,|\,t)=0$ is essentially the **modified-CS** [17]. The intuitive weight scheme mentioned in end of section III.B is denoted as **Intuitive ratio**. It can be seen that our proposed Pre-Re L1 with adaptive weights averagely achieves the best performance, and the other weight schemes fail to track the sparse Non-WSSUS vector as time increasing. This simulation results is consistent with the discussion provided above.

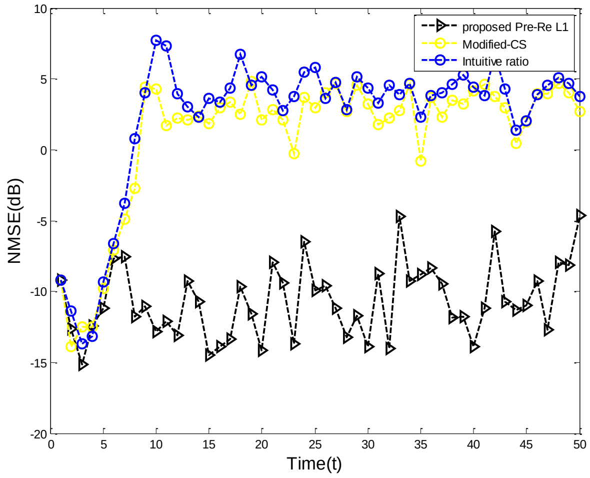


Fig. 4. NMSE of different weight schemes, SNR=15dB .

To show that when $\alpha(t)>0.5$, $w_i=0$ gain better performance than $w_i=0.5$. NMSE under deferent SNR are shown in Table I. It can be seen that in the case that when $\alpha(t)>0.5$, $w_i=0$ outperferned $w_i=0.5$. So just set $w_i=0.5$ without justifying $\alpha(t)$ will decrease the estimaiton performance. This result shows the effectiveness of the proposed adaptive weights approach.

TABLE I: NMSE OF WEIGHTS=0 AND WEIGHTS=0.5 WHEN $\alpha(t)>0.5$

| Weights | SNR | | |
|---|---|---|---|
| | 20 | 18 | 16 |
| $w_i=0$ | 0.0206 | 0.0199 | 0.0451 |
| $w_i=0.5$ | 0.0206 | 0.0371 | 0.0836 |

## V. DISCUSSION

In this section, we mainly discuss the issue of using previous time instants to decrease the computational burden. In our proposed algorithm, if $\alpha(t)<0.5$ then a re-estimation is needed, which increases the computational burden. It is seems that the first estimation to obtain $T(\mathbf{H}(t\,|\,t))_E$ can be replaced by using previous two instants, i.e., $t-1$, $t-2$, which is used in [18]. However, we find that this method has poorer performance than our proposed Pre-Re L1. This is mainly because the change of support is assumed to be slow in [18], while in our scenario the change of support is relatively quick. Using the method in [18] leads to the result that the threshold is often below 0.5, as a result, $w_i(t\,|\,t)=w_i=0.5$ is used. However the true situation may be that $\alpha(t)>0.5$, and when $\alpha(t)>0.5$, $w_i=0$ outperforms $w_i=0.5$ [15]. In fact, the meaning of the threshold in [18] is very different from our $\alpha(t)$. The threshold defined in [18] is used to check whether previous support estimate is a good predictor of the current support, while in our proposed Pre-Re L1 the

$\alpha(t)$ is used to measure the difference between the true support of $\mathbf{H}(t)$ and the estimated support of $\mathbf{H}(t)$. Moreover, since an accurate estimation of CIR in the OFDM system is crucial [8], it is worthy to trade some computational burden for accuracy.

## VI. Conclusion

In this paper a weighted $L_1$ minimization algorithm is presented to estimate the sparse Non-WSSUS channel vector. By exploiting the correlation of the non-stationary supports, a prediction-re-estimation procedure is established in the proposed algorithm. Compared with other CS techniques for wireless channel estimation, the proposed **Pre-Re L1** exploits the correlation of the non-stationary supports. Furthermore, its weights are adapted to the result which is estimated by the aid of prediction. Simulation results have shown that the proposed prediction-re-estimation $L_1$ minimization technique improves the estimation performance for the non-stationary correlated vector, and the adaptive weights add robustness to the $L_1$ minimization technique.

## Acknowledgment

The authors wish to thank the anonymous reviewers for their constructive advice. This work was supported in part by a National Science Foundation (NSF) of China (No. 61271213) and the Ph.D. Programs Foundation of Ministry of Education of China (No. 20133108110014).

## References

[1] W. C. Li and J. C. Preisig, "Estimation of rapidly time-varying sparse channels," *IEEE J. Ocean. Eng.*, vol. 32, no. 4, pp. 927-939, October 2007.

[2] G. Taubock, F. Hlawatsch, D. Eiwen, and H. Rauhut, "Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 255-271, April 2010.

[3] R. Prasad, C. R. Murthy, and B. D. Rao, "Joint approximately sparse channel estimation and data detection in OFDM systems using sparse bayesian learning," *IEEE Trans. Signal Process.*, vol. 62, no. 14, pp. 3591-3603, July 2014.

[4] A. S. Gupta and J. Preisig, "A geometric mixed norm approach to shallow water acoustic channel estimation and tracking," *Physical Communication*, vol. 5, no. 2, pp. 119-128, 2014.

[5] U. A. K. C. Okonkwo, S. Z. M. Hashim, R. Ngah, N. F. Nanyan, and T. A. Rahman, "Time-Scale domain characterization of nonstationary wideband vehicle-to-vehicle propagation channel," in *Proc. IEEE Asia-Pacific Conference on Applied Electromagnetics*, Tiara Beach Resort, Port Dickson, Nov. 2010, pp. 1-6.

[6] S. B. Wu, C. X. Wang, E. H. M. Aggoune, M. M. Alwakeel, and Y. J. He, "A Non-Stationary 3-D wideband twin-cluster model for 5G massive MIMO channels," *IEEE J. Sel. Topics Signal Process.*, vol. 32, no. 6, pp. 1207-1218, June 2014.

[7] S. Gligorevic, "Stochastic modeling of non-stationary channels," in *Proc. 7th European Conference on Antennas and Propagation (EuCAP)*, Gothenburg, Apr. 2013, pp. 1677-1681.

[8] F. Pena-Campos, R. Carrasco-Alvarez, O. Longoria-Gandara, and R. Parra-Michel, "Estimation of fast time-varying channels in OFDM systems using two-dimensional prolate," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 898-907, February 2013.

[9] W. U. Bajwa, *et al*., "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proc. IEEE*, vol. 98, no. 6, pp. 1058-1076, 2010.

[10] G. Matz, "On Non-WSSUS wireless fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2465-2478, 2005.

[11] G. Matz, H. Bolcskei, and F. Hlawatsch, "Time-Frequency foundations of communications: Concepts and tools," *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 87-96, November 2013.

[12] D. W. Matolak, I. Sen, and W. H. Xiong, "The 5-GHz airport surface area Channel—Part I: Measurement and modeling results for large airports," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2014-2026, July 2008.

[13] L. Bernadó, *et al*., "Delay and doppler spreads of nonstationary vehicular channels for safety-relevant scenarios," *IEEE Trans. Veh. Technol.*, vol. 63, no. 1, pp. 82-93, 2014.

[14] J. Ziniel and P. Schniter, "Dynamic compressive sensing of time-varying signals via approximate message passing," *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5270-5284, November 2013.

[15] M. P. Friedlander, H. Mansour, R. Saab, and O. Yilmaz, "Recovering compressively sampled signals using partial support information," *IEEE Trans. Info. Th.*, vol. 58, no. 2, pp. 1122-1134, February 2012.

[16] N. Vaswani and W. Lu, "Modified-CS: Modifying compressive sensing for problems with partially known support," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4595-4607, September 2010.

[17] C. L. Qiu and N. Vaswani. ReProCS: A missing link between recursive robust PCA and recursive sparse recovery in large but correlated noise. [Online]. Available: http://arxiv.org/abs/1106.3286

[18] H. Guo, C. Qiu, and N. Vaswani, "An online algorithm for separating sparse and low-dimensional signal sequences from their sum," *IEEE Trans. Signal Process.*, vol. 62, no. 16, pp. 4284-4297, August 2014.

[19] N. Vaswani, "Kalman filtered Compressed Sensing," in *Proc. 15th IEEE International Conference on Image Processing*, San Diego, Oct. 2008, pp. 893-896.

[20] N. Vaswani, "LS-CS-Residual (LS-CS): Compressive sensing on least squares residual," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4108-4120, August 2010.

[21] W. U. Bajwa, A. M. Sayeed, and R. Nowak, "Learning sparse doubly-selective channels," in *Proc. 46th Annual Allerton Conference on Communication, Control, and Computing*, USA, Sept. 2008, pp. 575-582.

[22] J. F. Yang and Y. Zhang, “Alternating direction algorithms for $ell_1$-Problems in compressive sensing,” *Sci. Comput, SIAMJ*, vol. 33 no. 1, pp. 250-278, 2011.
[23] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

**Chen Wang** is currently working toward the Ph.D. degree with the School of Communication and Information Engineering, Shanghai University, Shanghai, China. His research interests include Signal processing for wireless communication systems, Compressive Sensing and its application.

**Yong Fang** received the Ph.D. degree from Electronic Engineering, City University of Hong Kong, Hong Kong, in 1999. He is now a professor in School of Communication and Information Engineering, Shanghai University, Shanghai. His research interests include Communication Signal Processing, Blind Signal Processing, and Adaptive Information System.

**Jin-Yan Li** received his PhD degree (computer science) from the University of Melbourne (Australia). He is now an associate professor in Advanced Analytics Institute, University of Technology, Sydney, Australia. His research interests include data mining, information theory, and machine learning.