Zhou *et al. EURASIP Journal on Advances in Signal Processing* 2014, **2014**:29
http://asp.eurasipjournals.com/content/2014/1/29

EURASIP Journal on Advances in Signal Processing
a SpringerOpen Journal

RESEARCH **Open Access**

A novel method for sparse channel estimation using super-resolution dictionary

Fei Zhou*, Jing Tan, Xinyue Fan and Liang Zhang

Abstract

Due to the sparse distribution of reflectors in space, wireless channels are commonly sparse. Thus, utilizing the sparsity of channels in the delay-Doppler domain, a channel estimation method based on compressed sensing (CS) theory can reduce the number of pilots. However, because of discrete truncation in the time domain and limited bandwidth, the time delay and frequency shift of non-integer multiple samples can cause energy leakage in the delay and Doppler domain, which seriously reduce the delay-Doppler sparsity of the equivalent channel, thus affecting the accuracy of channel estimation. In this paper, we use an over-complete dictionary based on super-resolution to enhance the sparsity of the equivalent channel and reconstruct a doubly selective channel with greater accuracy. Simulation results demonstrate that the equivalent channel frequency response in the dictionary is sparser than that in the delay-Doppler domain. Compared with the traditional algorithm, the method proposed in this paper can effectively improve the performance of channel estimation.

Keywords: Sparse channel estimation; Doubly selective channel; Energy leakage; Super-resolution dictionary

1 Introduction

Traditional channel estimation methods on orthogonal frequency division multiplexing (OFDM) systems commonly assume that the channel has rich multipath and require a large number of pilots to obtain more accurate state information of the channel, which seriously reduce the utilization efficiency of the channel. Meanwhile, traditional methods of linear channel estimation already attained optimal estimation performance such as utilization efficiency, so it is difficult to improve them further. To overcome the bottleneck, we need to explore the own characteristic of the channel. More and more experimental evidences show that the sparse distribution of reflectors in space makes the transmission channel sparse.

The research on sparse wireless channel has already begun since the 1990s. Cotter and Rao utilize matching pursuit (MP) algorithm to estimate a small amount of non-zero channel taps in a single-carrier selective channel [1]. MP algorithm on decision feedback equalizer can effectively improve the performance of channel estimation. Compared with MP algorithm, the method proposed by Raghavendra and Giridhar estimates the tap position in a frequency-selective channel based on generalized Akaike information criterion and least squares (LS) algorithm, which greatly reduces the calculation burden of LS algorithm [2]. However, the above results were basically obtained by simulation and lacked relating theory analysis. In recent years, Donoho and Candes et al. propose a novel theory, i.e., compressed sensing, on the basis of functional analysis and approximation theory. The theory suggests that if the signal is sparse in a certain domain, it can be accurately reconstructed by a small amount sampling signal with high probability [3,4]. Bajwa et al. firstly applied compressed sensing (CS) theory for channel estimation and proposed the concept of compressive channel estimation (CCE) [5]. In the literature [5,6], Bajwa made a feasibility analysis of CCE and extended it to a doubly selective channel. The literature [7] gives a virtual channel model by Nyquist sampling for physical transmission environment in the angle-delay-Doppler domain and makes a comparison between CCE and LS. Meanwhile, aiming at acoustic OFDM systems, Berger et al. proved that CS channel estimation is superior over the traditional linear channel estimation method, which is reflected in the experimental data, like root-MUSIC and

*Correspondence: zhoufei@cqupt.edu.cn
Chongqing Key Lab of Mobile Communications Technology, Chongqing University of Posts and Telecommunications, Chongwen Road, Chongqing 400065, China

ESPRIT algorithms [8,9]. Taubock and Hlawatsch transform a doubly selective channel model to a solvable basis pursuit inequality constraint model, utilize the pilot signal as the key measurement that CS reconstruction requires, and analyze the sparsity of the channel parameter in the delay-Doppler domain [10]. However, it was concluded from the analysis that the energy leakage problem caused by discrete truncation of time domain and limited bandwidth obviously deteriorates the channel's sparsity which limits the performance improvement of CS-based channel estimation methods. To solve the problem, Taubock et al. propose an iterative basis optimization procedure that aims to maximize sparsity [11,12]. Although the method proposed by Taubock achieves significant performance gains, its basis optimization, which adds additional complexity, has to be performed before data transmission. Aimed at common problems caused by time delay and Dopler frequency shift of non-integer multiple samples, we use a super-resolution over-complete dictionary to improve the performance of channel estimation. The dictionary only increases the run time of the sparse reconstruction procedure with the increase of basis. We find that the over-complete dictionary representation of the channel is much sparser than the classical delay-Doppler representation in most cases, and it can effectively reduce the usage of pilots and improve the estimation performance.

The rest of this paper is organized as follows. Section 1 introduces CS theory and Section 2 introduces the OFDM system model. Section 3 analyzes the energy leakage of the channel in the delay-Doppler domain firstly, then uses the super-resolution dictionary instead of Fourier basis to enhance the sparsity of the channel, and next presents the CS-based channel estimation method. In Section 4, we present numerical results. Finally, Section 5 concludes the paper.

2 Compressed sensing theory

CS is a novel and highly promising theory that combines applied mathematics and signal processing. It breaks through the limitation of traditional Nyquist sampling and greatly reduces sampling frequency, data storage, and transmission burden. In CS, if a vector $\mathbf{x} \in \mathbb{R}^N$ is K-sparsity, or approximate K-sparsity, it can be represented by using $K(\leq N)$ non-zero coefficients [3]. Then, a linear measurement value about $\mathbf{x}$ can be obtained by selecting appropriate measurement matrix $\Phi \in \mathbb{C}^{M\times N}(M < N)$, as shown in (1). Only M measurements within $\mathbf{y}$ can be utilized to reconstruct the original signal with very high probability.

$$\mathbf{y} = \Phi\mathbf{x} + \mathbf{z} \tag{1}$$

The dimension number of $\mathbf{y}$ is far less than that of $\mathbf{x}$, so equation array (1) is underdetermined. However, in view of $\mathbf{x}$ which is K-sparse, it is only required to obtain K non-zero coefficients and their position. Candes et al. have proved that if the number of measurement $M = O\left(K\log(N)\right)$, and the measurement matrix satisfies the constraint of restricted isometry property (RIP), signal $\mathbf{x}$ can be reconstructed by solving the l_0-norm minimization in (2) [3].

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{y} = \Phi\mathbf{x} \tag{2}$$

Tao et al. already proved that Gaussian random measurement matrix, random partly Fourier measurement matrix, and Toeplize random matrix can satisfy the RIP criterion with very high probability [4], i.e., any N dimension K-sparsity vectors $\mathbf{a}$ all satisfy the following rule:

$$(1-\delta_k)\,\|\mathbf{a}\|_2^2 \leq \|\Phi\mathbf{a}\|_2^2 \leq (1+\delta_k)\,\|\mathbf{a}\|_2^2 \tag{3}$$

where $\delta_k \in (0,1)$ is a constant.

Unfortunately, l_0-norm is not convex. Actually, this problem is NP hard and therefore cannot be solved in a reasonable amount of time. By now, there are many different algorithms to solve it. Orthogonal matching pursuit (OMP) becomes a popular way in CS theory because it is simple for computation and easy for implementation [13]. OMP algorithm transforms the problem, l_0-norm minimization, to a relative simple problem shown in (4):

$$\mathbf{x} = \min_{\mathbf{x}'\in\mathbb{R}^N} \left\|\mathbf{x}'\right\|_0 \quad \text{subject to} \quad \left\|\Phi\mathbf{x}'-\mathbf{y}\right\|_2 \leq \varepsilon \tag{4}$$

where ε is the upper bound of the noise level. The basic idea of OMP algorithm is how to select the column vector of measurement matrix Φ by utilizing the greedy iteration way and reconstruct the signal by computing the support vector set on iterative algorithm of parameter $\mathbf{x}$ [13].

3 OFDM system model

3.1 System model

We describe a generalized cyclic prefix (CP) OFDM system shown in Figure 1. The discrete-time transmission can be written as

$$x[n] = \frac{1}{\sqrt{K}} \sum_{l=0}^{L-1} \sum_{k=0}^{K-1} x_{l,k} e^{j2\pi kn/K} g[n-lN] \tag{5}$$

where K is the number of subcarriers, L is the number of transmitted symbol periods, and N denotes the symbol duration. $N_{\mathrm{CP}} = N - K$ is the guard interval for the CP which is used to avoid the intersymbol interference (ISI). $x_{l,k}$ denotes the lth symbol transmitted at subcarrier k, and discrete transmit pulse $g[n]$ is 1 on $[0, N]$ and 0 otherwise.

The baseband-equivalent doubly selective channel $h(t,\tau)$ includes physical channel $h_{\mathrm{ch}}(t,\tau)$, transmitter filter $f_{\mathrm{tr}}(t)$, and received filter $f_{\mathrm{rec}}(t)$, so we have

$$h(t,\tau) = \int\int f_{\mathrm{rec}}(s) f_{\mathrm{tr}}(\tau - s - \theta)\, h_{\mathrm{ch}}(t - s, \theta)\, ds d\theta \tag{6}$$

In the receiver, the received signal after being sampled with period T_s is given by

$$r[n] = \sum_{\theta \in \mathbb{R}} h[n,\theta]\, x[n-\theta] + z[n] \tag{7}$$

where $h[n,\theta] = h(nT_s, \theta T_s)$ and $z[n] = z(nT_s)$ is discrete-time noise.

Assuming that the receiver is synchronous, if signal $r[n]$ is demodulated, we can obtain

$$r_{l,k} = \frac{1}{\sqrt{K}} \sum_{n=-\infty}^{\infty} r[n]\, e^{-j2\pi k(n-lN)/K} \gamma[n - lN] \tag{8}$$

where $l = 0, 1, \ldots, L-1$, $k = 0, 1, \ldots, K-1$, and $\gamma[n]$ is only 1 in $[N-K, N-1]$ and 0 otherwise. Combining (5), (7), and (8), we have

$$r_{l.k} = H_{l,k} x_{l,k} + z_{l,k} \tag{9}$$

where $z_{l,k} = \frac{1}{\sqrt{K}} \sum_{n=-\infty}^{\infty} z[n]\, e^{-j2\pi k(n-lN)/K} \gamma[n-lN]$ denotes the noise or the interference terms. $H_{l,k}$ is the system channel coefficients which will be analyzed in the following sections.

3.2 Doubly selective fading channel

According to the wide-sense stationary uncorrelated scattering (WSSUS) model, the time-varying multipath channel is expressed as [14]

$$h_{\mathrm{ch}}(t,\tau) = \sum_{q=1}^{P} \eta_q \delta(\tau - \tau_q)\, e^{j2\pi \nu_q t} \tag{10}$$

where P is the number of multipath components and η_q,τ_q, and ν_q are the attenuation coefficient, the delay, and the Doppler shift of path qth, respectively. δ denotes the Dirac delta function. We obtain the delay-Doppler spreading function $S(\nu,\tau)$ via Fourier transform.

$$\begin{aligned} S(\nu,\tau) &= \int h(t,\tau)\, e^{-j2\pi\nu t} dt \\ &= \sum_{q=1}^{P} \eta_q \delta(\tau - \tau_q)\, \delta(\nu - \nu_q) \end{aligned} \tag{11}$$

Assuming that physical channel $h(t,\tau)$ does not vary in the area of received filter $f_{\mathrm{rec}}(t)$, Equation 12 can be derived from Equation 6.

$$\begin{aligned} h(t,\tau) &= \iint f_{\mathrm{rec}}(s) f_{\mathrm{tr}}(\tau - s - \theta)\, h_{\mathrm{ch}}(t,\theta)\, ds d\theta \\ &= \int \psi(\tau - \theta)\, h_{\mathrm{ch}}(t,\theta)\, d\theta \\ &= \sum_{q=1}^{P} \eta_q \psi(\tau - \tau_q)\, e^{j2\pi\nu_q t} \end{aligned} \tag{12}$$

where $\psi(\tau - \theta) = f_{\mathrm{tr}}(t) \otimes f_{\mathrm{rec}}(t)$. After discretizing Equation 12, we have

$$h[n,\theta] = \sum_{q=1}^{P} \eta_q \psi(\theta T_s - \tau_q)\, e^{j2\pi\nu_q n T_s} \tag{13}$$

Due to the non-linear relation between $h[n,\theta]$ and channel parameters $[\eta_q, \nu_q, \tau_q]$, it is difficult to analyze the channel by utilizing Equation 13. It caused us to search for a new model with less parameters.

4 Sparse channel estimation using a dictionary

4.1 The effect of sparsity caused by energy leakage

Because the basis expansion model (BEM) [15] is simple for calculation and independent on statistical characteristics of the channel, it is widely used in time-varying multipath channel estimation. To compute $r_{l,k}$ in (8) for all $l = 0, 1, \ldots L-1$, the discrete-time received signal $r[n]$ has to be known for $n = 0, \ldots, N_0 - 1$, where $N_0 = LN$. The discrete time channel impulse response $h[n,\theta]$ in Equation 13 can be represented by BEM with a period of N_0 [16], so we have

$$h[n,\theta] = \sum_{i=-J}^{J} S_h[i,\theta]\, e^{j2\pi in/N_0} \tag{14}$$

where J satisfies $J/(N_0 T_s) \geq \nu_{\max}/2$ and $\nu_{\max}/2$ denotes the single-sided maximum Doppler shift. $S_h[i,\theta]$ is the discrete delay-Doppler spread function:

$$\begin{aligned} S_h[i,\theta] &= \frac{1}{N_0} \sum_{n=0}^{N_0-1} h[n,\theta]\, e^{-j2\pi in/N_0} \\ &= \sum_{q=1}^{P} \eta'_q \psi(\theta T_s - \tau_q)\, \mathrm{dir}_{N_0}(i - \nu_q T_s N_0) \end{aligned} \tag{15}$$

where $\mathrm{dir}_N(x) = \sin(\pi x)/(N \sin(\pi x/N))$, $\eta'_q = \eta_q e^{j\pi(\nu_q T_s - i/N_0)(N_0 - 1)}$, and $\theta = 0, \ldots, D-1$. $D \geq$

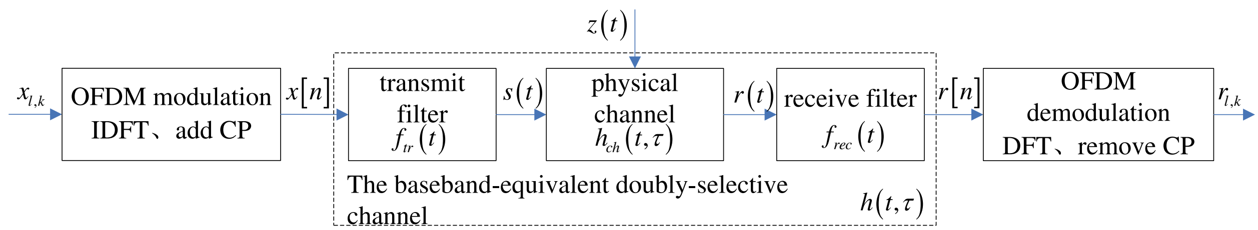

Figure 1 OFDM system model. The baseband-equivalent doubly selective channel $h(t,\tau)$ includes physical channel $h_{\text{ch}}(t,\tau)$, transmitter filter $f_{\text{tr}}(t)$, and received filter $f_{\text{rec}}(t)$.

$\tau_{\max}/T_s$ denotes the number of discrete time-delay sampling points. It is obvious that if the maximum time-delay satisfies $\tau_{\max} \leq N_{\text{CP}}$, the intersymbol interference can be eliminated. And if the ideal filter $f_{\text{tr}}(t) = f_{\text{rec}}(t) = \sqrt{1/T_s}\sin c(t/T_s)$ is applied, where $\sin c(x) = \sin(\pi x)/(\pi x)$, $\psi(\theta T_s - \tau_q) \approx \sin c(\theta - \tau_q/T_s)$ can be obtained and applied in Equation 15. So we have

$$S_h[i,\theta] = \sum_{q=1}^{P} \eta'_q \Lambda_q[i,\theta] \tag{16}$$

where

$$\Lambda_q[i,\theta] = \sin c\left(\theta - \frac{\tau_q}{T_s}\right) \text{dir}_{N_0}\left(i - \nu_q T_s N_0\right) \tag{17}$$

Combining (7), (9), and (14), we can derive the channel coefficient $H_{l,k}$

$$H_{l,k} = \sum_{\theta=0}^{D-1}\sum_{i=-J}^{J-1} S_h[i,\theta] e^{-j2\pi\left(\frac{k\theta}{K}-\frac{il}{L}\right)} A_{\gamma,g}\left(\theta, \frac{i}{N_0}\right) \tag{18}$$

where $A_{\gamma,g}(\theta, i/N_0) = \sum_n \gamma[n]\, g[n-\theta]\, e^{j2\pi in/N_0}$ is the cross-ambiguity function. Therefore, we can analyze the sparsity of $\Lambda_q[i,\theta]$ instead of $S_h[i,\theta]$. From Equation 11, the delay-Doppler function $S(\nu,\tau)$ consisted of the Dirac function in delay-Doppler point (τ_q, ν_q) which corresponds to reflecting path q and is supposed to be sparse. However, the Dirac function is replaced by $\sin c(x)$ and $\text{dir}_N(x)$ in $S_h[i,\theta]$. Only when τ_q/T_s and $\nu_q T_s N_0$ are all integers, $\Lambda_q[i,\theta]$ can be simplified to the Dirac function. The reason is that $\sin c(x)$ and $\text{dir}_N(x)$ are equal to 1 on $x = 0$ and 0 otherwise, as shown in Figure 2. Under any other conditions, $\Lambda_q[i,\theta]$ is not equal to 0 for any i and θ. In other words, the peak energy of the discrete delay-Doppler function may have leakage to the near delay-Doppler area. Figure 3 demonstrates the leakage effect of $S_h[i,\theta]$ where $P = 10$, $N_{\text{CP}} = 16$, $K = 25$, and $L = 16$. η_q, τ_q/T_s, and $\nu_q T_s N_0$ are random variables.

From Figure 3, the peak energy of $S_h[i,\theta]$ leaks to the near delay-Doppler area, and then its value would fade and approximate to 0 gradually. Therefore, $S_h[i,\theta]$ is approximate sparse, i.e., the channel coefficient $H_{l,k}$ is approximate sparse in the delay-Doppler domain. However, the fading of $\sin c(x)$ and $\text{dir}_N(x)$ is slow, and a lot of values in $S_h[i,\theta]$ cannot be neglected. It seriously influences the sparsity of the channel and the estimation performance of the channel.

4.2 Sparsity enhancement using the super-resolution dictionary

In real wireless channels, the time delay and Doppler frequency shift of non-integer times sampling points exist generally. They seriously influence the sparsity of channel coefficient $H_{l,k}$ in the time delay and Doppler domain and do not satisfy the prerequisite condition. Hence, how to avoid the energy leakage is a very important issue in channel estimation. Essentially, the channel energy leakage is actually introduced by channel discrete characterization, and the sparsity of the channel itself does not disappear.

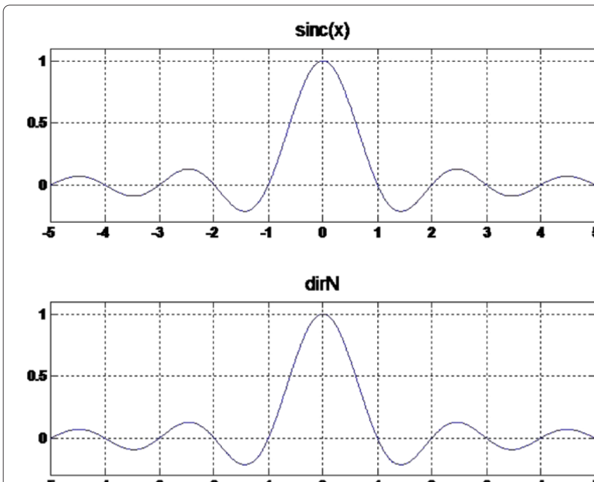

Figure 2 Functions $\sin c(x)$ and $\text{dir}_N(x)$. $\sin c(x)$ and $\text{dir}_N(x)$ are equal to 1 on $x = 0$ and 0 otherwise, but their intensity decreases due to the decay.

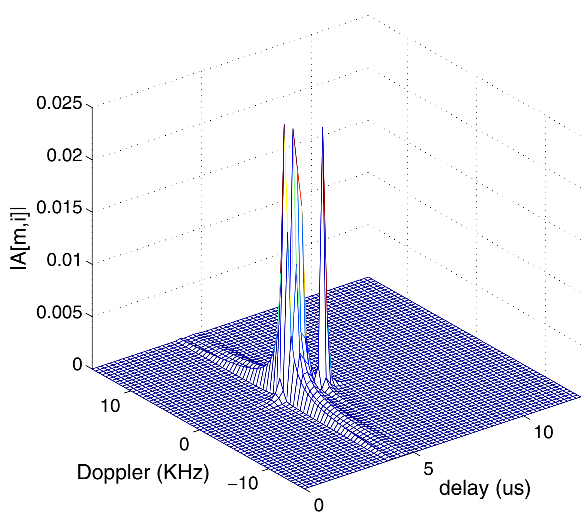

Figure 3 Modulus of the expansion coefficients for DFT basis. Since the fading of $\sin c\,(x)$ and $\mathrm{dir}_N\,(x)$ is slow, a lot of coefficients should not be neglected which decrease the sparsity of the channel. The parameters are $P = 10$, $N_{CP} = 16$, $K = 256$, and $L = 16$. η_q, τ/T_s, and $v_q T_s N_0$ are random variables.

Therefore, if we can improve the accuracy of channel discrete characterization, it would greatly reduce the energy leakage. Therefore, improving the discrete accuracy of the channel impulse response reduces the energy leakage [17].

Assuming that $h\,[n,\theta]$ can be represented by BEM with a period of λN_0, we have

$$h\,[n,\theta] = \sum_{i=-J}^{J} S^{(\lambda)}\,[i,\theta]\, e^{j2\pi in/(\lambda N_0)} \tag{19}$$

where $J/(\lambda N_0 T_s) \geq v_{\max}/2$.

Let $\mathbf{h}_\theta^{(\lambda)} = [h\,[0,\theta]\,,\ldots,h\,[\lambda N_0-1,\theta]]^T$ and $\mathbf{S}_\theta^{(\lambda)} = \left[S^{(\lambda)}\,[-J,\theta]\,,\ldots,S^{(\lambda)}\,[J-1,\theta]\right]$, we can derive $\mathbf{S}_\theta^{(\lambda)}$ based on LS and obtain

$$\min_{\mathbf{S}_\theta^{(\lambda)}} \left\| \mathbf{h}_\theta^{(\lambda)} - \frac{1}{\lambda N_0}\mathbf{F}^{(\lambda)}\mathbf{S}_\theta^{(\lambda)} \right\|^2 \tag{20}$$

where $\mathbf{F}^{(\lambda)} = \begin{bmatrix} 1 & \cdots & 1 \\ e^{-\frac{j2\pi J}{\lambda N_0}} & \cdots & e^{\frac{j2\pi (J-1)}{\lambda N_0}} \\ \vdots & \cdots & \vdots \\ e^{-\frac{j2\pi J(\lambda N_0-1)}{\lambda N_0}} & \cdots & e^{\frac{j2\pi (J-1)(\lambda N_0-1)}{\lambda N_0}} \end{bmatrix}$.

Then,

$$\mathbf{S}_\theta^{(\lambda)} = \left(\mathbf{F}^{(\lambda)}\right)^{\dagger}\mathbf{h}_\theta^{(\lambda)} = \frac{1}{\lambda N_0}\left(\mathbf{F}^{(\lambda)}\right)^{H}\mathbf{h}_\theta^{(\lambda)} \tag{21}$$

where $\mathbf{S}_\theta^{(\lambda)}$ corresponds to $2J+1$ Doppler sample point under λ times over-sampling.

We may define $\mathbf{S}_{\theta,m}^{(\lambda)} = \left[S^{(\lambda)}\,[-a_m\lambda+m,\theta]\,,\ldots,\ S^{(\lambda)}\,[b_m\lambda+m,\theta]\right]$ and $\mathbf{h}_\theta = [h\,[0,\theta]\,,\ldots,h\,[N_0-1,\theta]]^T$, where $a_m = (J+m) \bmod \lambda$ and $b_m = (J-m) \bmod \lambda$. It is easy to prove Equation 22:

$$\mathbf{S}_{\theta,m}^{(\lambda)} = \min_{\mathbf{S}_{\theta,m}^{(\lambda)}} \left\| \mathbf{D}_m^{(\lambda)}\mathbf{h}_\theta - \frac{1}{N_0}\mathbf{F}_m^{(\lambda)}\mathbf{S}_{\theta,m}^{(\lambda)} \right\|^2 \tag{22}$$

where $\mathbf{D}_m^{(\lambda)} = \mathrm{diag}\left\{\left[1, e^{-\frac{j2\pi m}{\lambda N_0}},\ldots, e^{-\frac{j2\pi m(N_0-1)}{\lambda N_0}}\right]^T\right\}$ and

$$\mathbf{F}_m^{(\lambda)} = \begin{bmatrix} 1 & \cdots & 1 \\ e^{-j2\pi a_m/N_0} & \cdots & e^{j2\pi b_m/N_0} \\ \vdots & \cdots & \vdots \\ e^{-j2\pi a_m(N_0-1)/N_0} & \cdots & e^{j2\pi b_m(N_0-1)/N_0} \end{bmatrix}.$$

Then,

$$\mathbf{S}_{\theta,m}^{(\lambda)} = \left(\mathbf{F}_m^{(\lambda)}\right)^{\dagger}\mathbf{D}_m^{(\lambda)}\mathbf{h}_\theta^{(\lambda)} = \frac{1}{N_0}\left(\mathbf{F}_m^{(\lambda)}\right)^{H}\mathbf{D}_m^{(\lambda)}\mathbf{h}_\theta \tag{23}$$

Equation 23 corresponds to taking the (a_m+b_m+1) samples around zero from the critically sampled Doppler spectrum of the $m/(\lambda N)$ frequency-shifted version of $\mathbf{h}_\theta$. So all samples $(m = 0, 1, \ldots, \lambda-1)$ can deduce $2J+1$ Doppler frequency shift point. By replacing $\mathbf{D}_m^{(\lambda)}$, $\mathbf{F}_m^{(\lambda)}$, and $\mathbf{h}_\theta$ into Equation 23, we have

$$S^{(\lambda)}\,[i,\theta] = \sum_{q=1}^{P} \eta_q \psi\left(\theta T_s - \tau_q\right) e^{j\pi\left(v_q T_s - \frac{i}{\lambda N_0}\right)(N_0-1)} \times \mathrm{dir}_{N_0}\left(\pi\left(i - \lambda v_q T_s N_0\right)\right) \tag{24}$$

From Equations 24 and 16, we find that $\mathrm{dir}_{N_0}\left(\pi\left(i - v_q T_s N_0\right)\right)$ in Equation 16 is replaced into $\mathrm{dir}_{N_0}\left(\pi\left(i - \lambda v_q T_s N_0\right)\right)$. So if $v_q T_s N_0 = n/\lambda$, $\mathrm{dir}_{N_0}\left(\pi\left(i - \lambda v_q T_s N_0\right)\right)$ can be transformed into a Dirac function. The higher the parameter λ is, the more the sample of the Doppler spectrum is. And when the position of the sample is closer to the real position, the problem of energy leakage would greatly be reduced. Admittedly, a part of the Doppler spectrum still results in a certain energy leakage; however, its value is very small. Letting $\mathcal{I}$ denote all integer set which satisfied $\left|i - \lambda v_q T_s N_0\right| \leq \Delta i$ in the area of $i \in \{-J,\ldots,J\}$, the energy sum of samples in which the distance $v_q T_s N_0$ in

$\text{dir}_{N_0}\left(\pi\left(i-\lambda v_q T_s N_0\right)\right)$ is larger than $\Delta i \in \{2,3,\ldots\}$ can be given.

$$\begin{aligned}\sum_{i\notin\mathcal{I}}\left|\text{dir}_{N_0}\left(\pi\left(i-\lambda v_q T_s N_0\right)\right)\right|^2 &= \sum_{i\in\mathcal{I}}\left|\frac{\sin\left(\pi\left(i-\lambda v_q T_s N_0\right)\right)}{N_0\sin\left(\pi\left(i-\lambda v_q T_s N_0\right)/N_0\right)}\right|^2 \\ &\le \sum_{i\in\mathcal{I}}\frac{1}{\left|N_0\sin\left(\pi\left(i-\lambda v_q T_s N_0\right)/N_0\right)\right|^2} \\ &\le \frac{2}{N_0^2}\int_{\Delta i-1}^{\infty}\frac{dx}{\sin^2\left(\pi x/N_0\right)} = \frac{2}{N_0\pi}\cot\left(\frac{\pi}{N_0}\left(\Delta i-1\right)\right) \\ &\le \frac{1}{\pi\left(\Delta i-1\right)}\end{aligned} \tag{25}$$

Figure 4 shows the leakage effect of $\text{dir}_{N_0}\left(\pi\left(i-v_q T_s N_0\right)\right)$ and $\text{dir}_{N_0}\left(\pi\left(i-\lambda v_q T_s N_0\right)\right)$, where $v_q T_s N_0 = 0.4$ and $v_q T_s N_0 = 0.5$.

Similarly, we can apply the same method to solve the energy leakage problem in the time-delay domain [18,19]. Lastly, equivalent discrete-time baseband channel frequency response is given by

$$H_{l,k} = \sum_{\theta=0}^{D-1}\sum_{i=-J}^{J-1} S_D\left[i,\theta\right]e^{-j2\pi k\theta/\lambda_{\text{delay}}K}e^{j2\pi il/\lambda_{\text{Doppler}}L} \times A_{\gamma,g}\left(\theta,i/N_0\right) \tag{26}$$

where $DT_s/\lambda_{\text{delay}} \ge \tau_{\max}$, λ_{delay} and λ_{Doppler} are the over multiple number of delay and Doppler, respectively, and $A_{\gamma,g}\left(\theta,i/N_0\right)$ is the same as that in Equation 17.

Redundant dictionary $\mathbf{U}$ is defined by $[\mathbf{U}]_{kL+l,\,(i+J)D+\theta} = e^{-j2\pi(k\theta/\lambda K - ni/\lambda N_0)}$, $[\mathbf{h}]_{kL+l} = H_{l,k}$, and $\left[\mathbf{g}\right]_{(i+J)D+\theta} = S_D\left[i,\theta\right]A'_{\gamma,g}\left(\theta,i/N_0\right)$. Obviously, $\mathbf{g}$ is sparse according to the above analysis. So we can obtain the vector form of Equation 26:

$$\mathbf{h} = \mathbf{U}\mathbf{g} \tag{27}$$

Especially, if $\lambda = 1$, $\mathbf{U}$ is the 2-D Fourier basis as used in Equation 18. Under the same conditions as those in Figure 3, Figure 5 shows the distribution of coefficient $\mathbf{g}$ in the dictionary domain.

Obviously, the effective value $\mathbf{g}$ in Figure 5 is less than that in Figure 3. So the channel coefficient corresponding to $\lambda = 2$ has a higher sparsity. To analyze the sparsity in the dictionary domain better, we utilize OMP algorithm to solve S-sparse approximation and obtain the most S maximum value in $\mathbf{g}$. From Figure 6, with the increase of S, the mean square error $E[\,|\mathbf{h}-\hat{\mathbf{h}}_S|^2]$ would decrease gradually and would be close to 10^{-1}, i.e., we can obtain 90% channel energy based on S strongest atoms. If λ is higher, the atoms that satisfy the required mean error are fewer. So we can conclude that if λ is higher, the sparsity of frequency response $\mathbf{h}$ is higher in the over-complete dictionary.

4.3 The estimation of the sparse channel based on the dictionary domain

Assuming that $(l,k) \in \mathcal{P}$, where $\mathcal{P}$ is the pilot set, the total number of pilot point is $Q = |\mathcal{P}|$. The number of pilots must satisfy the lowest demand of compressive

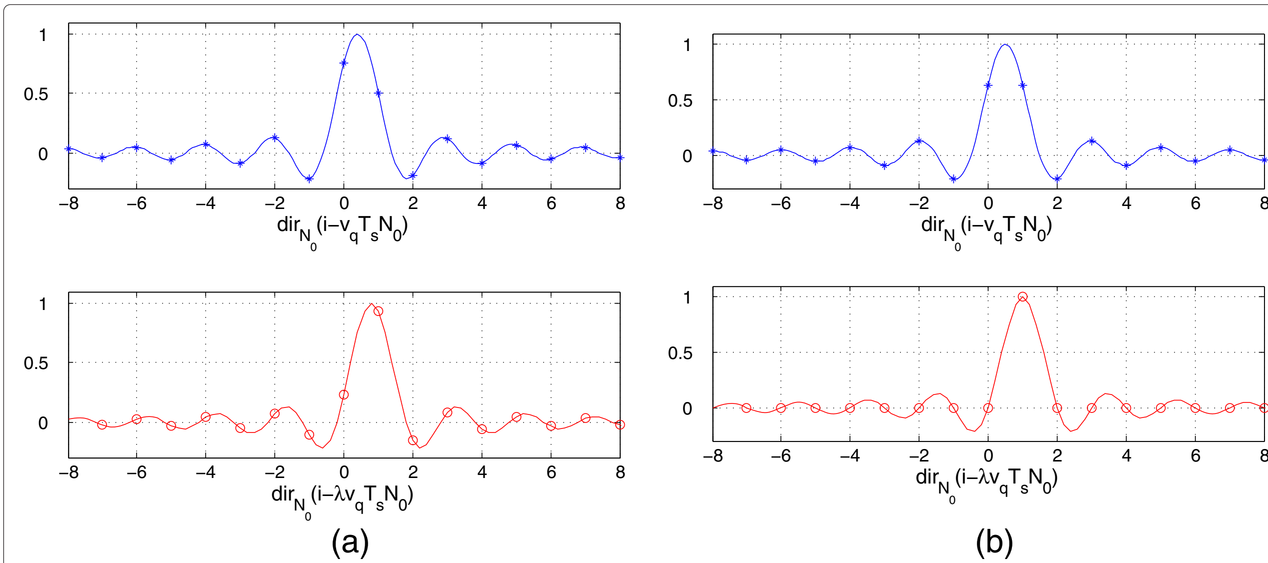

Figure 4 Comparison between $\text{dir}_{N_0}\left(\pi\left(i-\lambda v_q T_s N_0\right)\right)$ and $\text{dir}_{N_0}\left(i-v_q T_s N_0\right)$. $\text{dir}_{N_0}\left(\pi\left(i-\lambda v_q T_s N_0\right)\right)$ has a quicker fade speed of side lobe and a smaller value of side lobe than $\text{dir}_{N_0}\left(i-v_q T_s N_0\right)$. **(a)** $v_q T_s N_0 = 0.4$, $\lambda = 2$. **(b)** $v_q T_s N_0 = 0.5$, $\lambda = 2$.

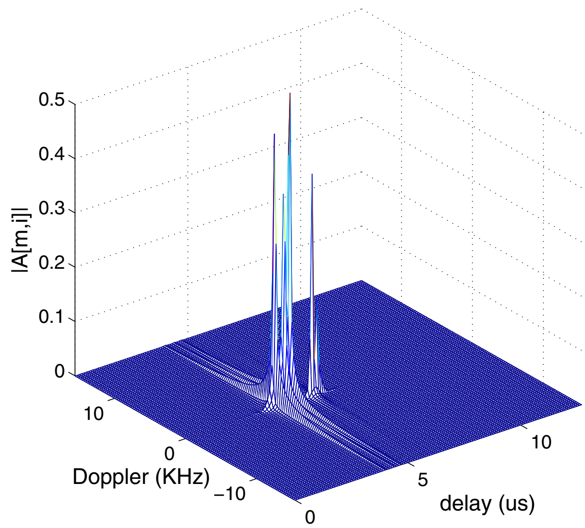

Figure 5 Sparsity enhancement obtained with the over-complete dictionary corresponding to $\lambda = 2$. The effective value is much less than that in Figure 3; the parameters are the same as those in Figure 3 except the basis.

sensing theory to represent the measurement signal without distortion. The literature [14] gives the limitation of the number of measurement samples required by OMP. $N_m \geq KS \ln(N_r/\delta)$, where S denotes sparsity. Commonly, $K \leq 20$ is reasonable. If S is too large, we may also set $K \approx 4$ and $\delta \in (0, 0.36)$. OMP may reconstruct a signal with $1 - 2\delta$ probability. So we can select the suitable number of pilots, $Q \geq N_m$.

According to (9), the estimation of the channel coefficient in the pilot is given by

$$\tilde{H}_{l,k} = \frac{r_{l,k}}{x_{l,k}} = H_{l,k} + z_{l,k} \tag{28}$$

Let $\mathbf{h}_\Delta = \mathbf{h}\left|_{(l,k)\in\mathcal{P}}\right., \mathbf{U}_\Delta = \mathbf{U}\left|_{(l,k)\in\mathcal{P}}\right.$ be the matrix corresponding to the pilot point, and $\mathbf{z}_\Delta$ be the set of $\tilde{z}_{l,k}$ in $(l,k) \in \mathcal{P}$. We have

$$\mathbf{h}_\Delta = \mathbf{U}_\Delta \mathbf{g} + \mathbf{z}_\Delta \tag{29}$$

According to the above analysis, we can conclude that $\mathbf{g}$ is sparse. So Equation 27 is a standard equation of CS. Measurement matrix $\mathbf{U}_\Delta$ is a structured random matrix which is formed by selecting a row vector of unitary matrix corresponding to the pilot point. If we select the position of the pilot uniformly and randomly and Q is large enough, the normalized matrix $\sqrt{1/Q}\mathbf{U}_\Delta$ has a small constraint isometric constant with very high probability. In other words, we can obtain very good reconstruction performance. So we can obtain Equation 30 from Equation 29.

$$\mathbf{h}_\Delta = \Phi\mathbf{x} + \mathbf{z}_\Delta \tag{30}$$

where $\Phi = \sqrt{KL/Q}\mathbf{U}_\Delta$ and $\mathbf{x} = \sqrt{Q/KL}\mathbf{g}$. Therefore, we may realize channel equalization based on CS theory. The following are the detailed steps:

1. Obtain the estimation $\tilde{H}_{l,k}$ of the channel coefficient according to Equation 28, and then combine all $\tilde{H}_{l,k}$ to form $\tilde{\mathbf{h}}_\Delta$.
2. Utilize OMP to obtain an estimation $\tilde{\mathbf{x}}$ of $\mathbf{x}$ based on known $\tilde{\mathbf{h}}_\Delta$ and Φ according to Equation 30. The detail of realization refers to Algorithm 1. After rescaling $\tilde{\mathbf{x}}$, we can obtain estimation value $\tilde{\mathbf{g}}$, as well the spread coefficient $S_D[i, \theta]$ corresponding to the dictionary.
3. Compute all channel coefficient $\tilde{\mathbf{h}}$ by utilizing Equation 27 based on known $\tilde{\mathbf{g}}$ and $\mathbf{U}$.

The run time of OMP mainly depends on the index set Λ_k selected in the iteration process. We need to select the optimal atom $\mathrm{O}(D(2J+1))$ from $D \times (2J+1)$ atom. So with the increase of λ, the number of time-delay sample D and the number of Doppler sample J would increase exponentially. However, with the increase of λ, the channel sparsity would be enhanced and the iteration times required by OMP would reduce. Lastly, it can save a certain run time.

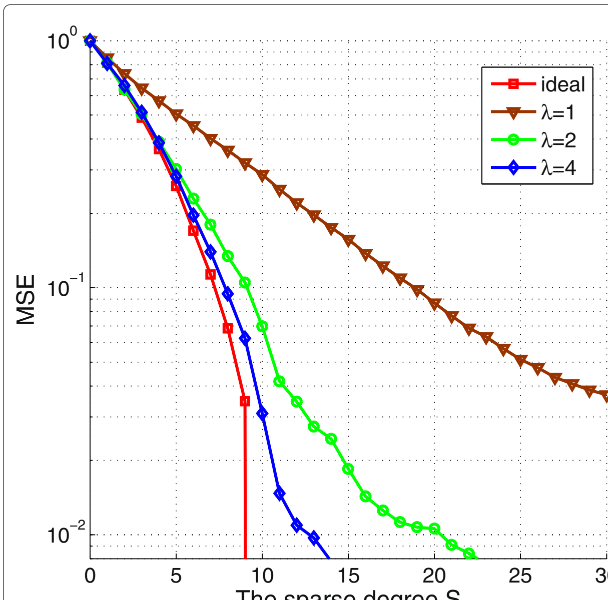

Figure 6 MSE of sparse approximation of h in Equation 27. We utilize OMP algorithm to solve *S*-sparse approximation and obtain the most *S* maximum value in **g**; a redundant basis leads to significantly fewer terms than Fourier basis.

Algorithm 1 Steps of reconstitution

Input:
Data vector $\tilde{\mathbf{h}}_\Delta$, measurement matrix Φ, noise variance σ^2

Output:
The proposed solution is $\hat{\mathbf{x}}_k$ obtained after k iterations.

1: **initialize:** Residual $\mathbf{r}_0 = \tilde{\mathbf{h}}_\Delta$, index set $\Lambda_0 = \emptyset$, iteration counter $k = 0$
2: **while** ($\|\mathbf{r}_k\|_2 \geq \sigma^2$) **do**
3: $k = k + 1$
4: Find the index λ_k: $\lambda_k = \arg\max_j \left|\langle \mathbf{r}_{k-1}, \varphi_j \rangle\right|, \varphi_j \in \Phi$
5: Augment the index set and the matrix of chosen atoms: $\Lambda_k = \Lambda_{k-1} \cup \{\lambda_k\}$ and $\Phi_k = \left[\Phi_{k-1}, \varphi_{\lambda_k}\right]$
6: Solve a least squares problem to obtain a new signal estimation: $\hat{\mathbf{x}}_k = \arg\min_\chi \left\|\mathbf{y} - \Phi_{\Lambda_k}\chi\right\|_2$
7: Calculate the new residual: $\mathbf{r}_k = \mathbf{y} - \Phi\hat{\mathbf{x}}_k$
8: **end while**

5 Simulation and analysis

In this section, we present numerical results to analyze the performance of the CS-based channel estimation algorithm using the over-complete dictionary. The following are the relating simulation parameters: carrier frequency $f_c = 2$ GHz, bandwidth $B = 10.24$ MHz, subcarrier number $K = 1,024$, the length of cyclic prefix $N_{\text{CP}} = 128$, and sample period $T_s = 0.1$ ms. We may utilize the channel simulation tool IlmProp [20] based on the geometrical structure of space to simulate a doubly selective fading channel. The simulated frame for the OFDM block has eight symbols, i.e., $L = 8$. In the simulated environment, the distance between the transmitter and the receiver is 2,000 m, and 10 reflectors, in which 2 reflectors are distributed within 150 m from the transmitter, form 10 multipath clusters, which satisfy the Gaussian distribution. The random speed of each path is less than 100 m/s and the acceleration is less than 20 m/s^2. Assume that the noise $z[n]$ is additive white Gaussian noise (AWGN), in which the mean is 0 and the variance is σ^2. So the signal-noise ratio (SNR) of the symbol block is given by

$$\text{SNR} = \sum_{n=0}^{N_0-1} \mathrm{E}\left\{|r[n] - z[n]|^2\right\} \bigg/ \sum_{n=0}^{N_0-1} \mathrm{E}\left\{|z[n]|^2\right\} \tag{31}$$

In addition, simulation tools are MATLAB 2009 and an Intel Core PC with a 2.8-G processor and 1.5-G RAM.

To begin with, we give the performance comparison for a variety of algorithms under different SNR conditions. The SNR varies within -10 ~20 dB. For LS channel estimation, we used two different rectangular pilot constellations, i.e., selected uniformly 6.5% and 12.5% of all symbols for pilots, respectively. For CS-based estimation, we select randomly 6.25% of all symbols for pilots and three different basis, i.e., Fourier basis (DFT) (i.e., $\lambda = 1$) [10], iterative optimize basis [12], and over-complete dictionary (DIC) ($\lambda_{\text{delay}} = \lambda_{\text{Doppler}} = 2, \lambda_{\text{delay}} = \lambda_{\text{Doppler}} = 4$). Figure 7 gives the MSE comparison of different algorithms under SNR. Figure 8 shows the bit error ratio (BER) comparison of equalization decoding. Both figures suggest that when we only apply 6.25% resource as pilots, the performance of LS estimation is much bad; the reason is that the distribution of the pilot cannot satisfy the Nyquist sampling criterion. However, when OMP algorithm also applies 6.25% resource for the pilot, the performance of the channel estimation proposed in this paper is better than that of LS estimation that applies 12.5% resource for the pilot. So the utilization efficiency of the spectrum is improved obviously. In addition, OMP algorithm on the over-complete dictionary domain has a better performance than the traditional algorithm on the delay-Doppler domain. When $\lambda = 2$, the performance on the over-complete dictionary domain approaches that on the iterative optimize basis. And when $\lambda = 4$, the performance on the over-complete dictionary domain is better than that on the iterative optimize basis. With the increase

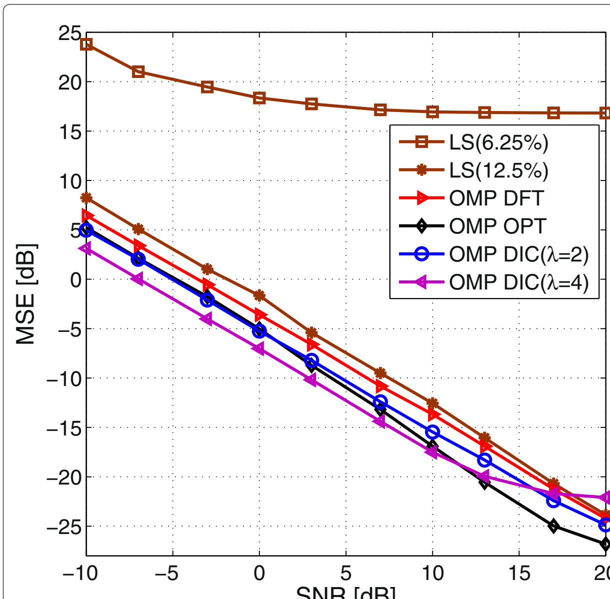

Figure 7 MSE performance in different SNR environments. The compressed sensing methods can increase their performance significantly by using dictionaries with a finer resolution (for OMP, $\lambda = 2, 4$); OMP_DFT, OMP_OPT, and OMP_DIC represent Fourier basis, iterative optimize basis, and dictionary used in OMP algorithm, respectively.

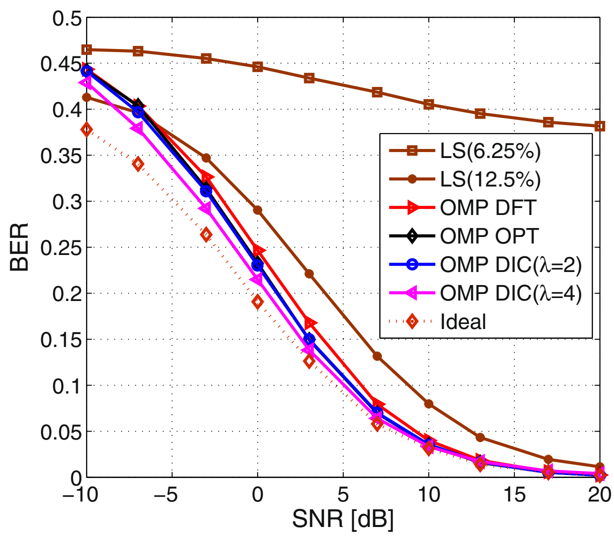

Figure 8 BER performance in different SNR environments. With the increase of λ, the bit error rate is gradually reduced and reaches the ideal channel estimation.

of the λ, the BER performance of OMP algorithm using over-complete dictionary is close to that of ideal channel estimation. So the higher is λ, the better is the reconstruction performance of OMP and the higher is the estimation accuracy.

Although the performance of the dictionary on $\lambda = 4$ is the best, its complexity is also the highest, as shown in Table 1. Considering the reconstruction time only, iterative optimize basis is the optimal except for LS estimation. However, the method based on iterative optimize basis needs extra 163.8541 s to compute the optimal basis. But, in the channel estimation algorithm proposed in this paper, we can produce the dictionary by FFT. The method only enlarges the size of the reconstructed atomic set and does not need extra computation. By comparing the required time between $\lambda = 4$ and $\lambda = 2$, it can be concluded that the sparsity of the channel coefficient in the dictionary domain would be sparer with the increase of parameter λ. Meanwhile, the sparsity of the channel coefficient can also be influenced by the physical path and lower than the number of reflectors. Therefore, when $\lambda = 2$, the channel estimation method based on the super-resolution dictionary domain has a certain advantage on the algorithm performance and computation complexity.

Then, we present the performance comparison of different algorithms under the different numbers of pilot symbols. Here, SNR is assumed to be 0 dB, the number of pilot symbols varies within 3%∼10%, and the other parameters are the same as those in the above simulation. Figure 9 shows that the performance is improved with the increase of the pilot number. Under the same accuracy condition, the higher the solution in the dictionary domain is, the less the required number of pilot is. For example, when $\text{MSN} = -5$ dB and $\lambda = 4$, the required resource of pilots is only about 4%. If $\lambda = 2$, the percentage is about 5%. However, the method based on Fourier basis needs about 7.5% resource. Figure 10 shows the comparison under the different numbers of pilots. With the increase of pilot number, BER would be close to the performance of ideal channel estimation. In other words, under the same performance of MSE or BER, the pilot number required by DIC is less than that by DFT and OPT. Hence, the sparse channel estimation based on the dictionary domain can effectively reduce the number of pilots and improve the spectrum efficiency.

Lastly, to reduce the computation complexity, we may apply different over-sampling times in the time-delay domain or the Doppler domain. In a wireless channel, we assumed that the random speed of each path cluster is less than 10 m/s and the acceleration is less

Table 1 The run time of different algorithms

Algorithm	Reconstruction (s)	Extra time (s)
LS (6.25%)	0.0455	0
LS (12.5%)	0.0372	0
OMP_DFT	0.4282	0
OMP_OPT	0.3017	163.8541
OMP_DIC ($\lambda = 2$)	0.9661	0
OMP_DIC ($\lambda = 4$)	2.9548	0

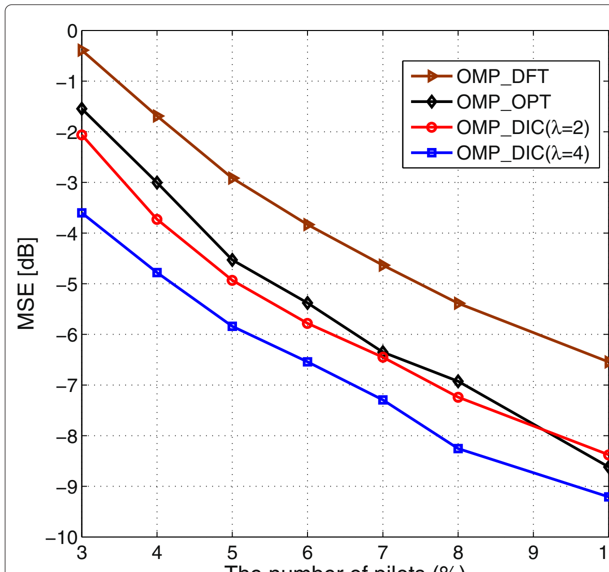

Figure 9 MSE performance under the different numbers of pilot symbols. Dictionaries on $\lambda = 2, 4$ need less pilots than the optimized basis (OPT) and Fourier basis (DFT) in compressive channel estimation.

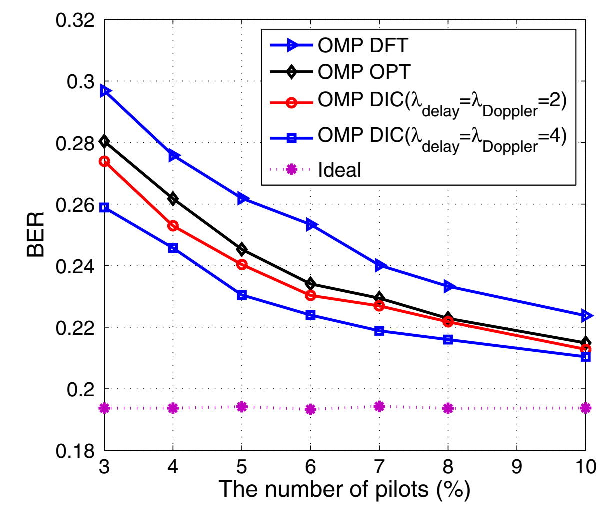

Figure 10 BER performance under the different numbers of pilot symbols. An increase in either the measure values or lambda can reduce BER.

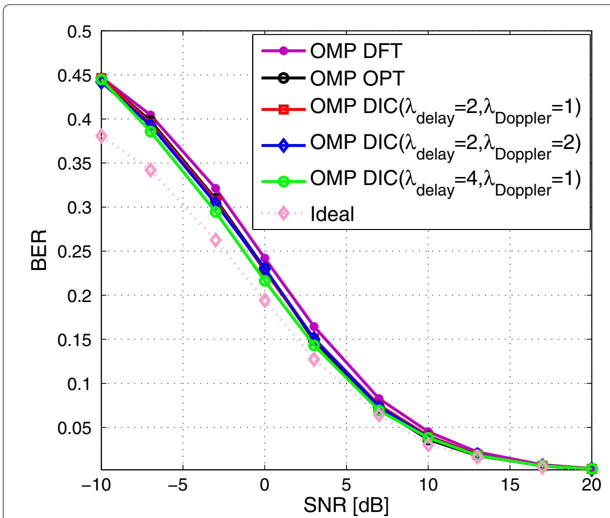

Figure 12 BER performance using different resolution dictionaries. The more sparse the channel is, the higher the degree of accuracy and the lower the BER we get.

than 1 m/s^2, so the Doppler influence is not much serious. The other parameters are the same as those in the first simulation. Figures 11 and 12 show that when $\lambda_{\text{delay}} = 4$, $\lambda_{\text{Doppler}} = 1$, the performance of channel estimation is better than that of others. The accuracy performance on $\lambda_{\text{delay}} = \lambda_{\text{Doppler}} = 2$ is almost the same as that on $\lambda_{\text{delay}} = 2$, $\lambda_{\text{Doppler}} = 1$; however, the former run time is about twice the latter, as shown in Table 2. Compared with that of the DFT and OPT, the run time of dictionary basis on $\lambda_{\text{delay}} = 2$, $\lambda_{\text{Doppler}} = 1$ is only slightly more than that of Fourier basis; however, the former performance is obviously better than the latter. Considering that the Doppler influence is not much serious, we only over-sample in the time-delay domain, so we can obtain an optimal selection in both computation complexity and performance. Similarly, we can also apply the same method in the Doppler domain.

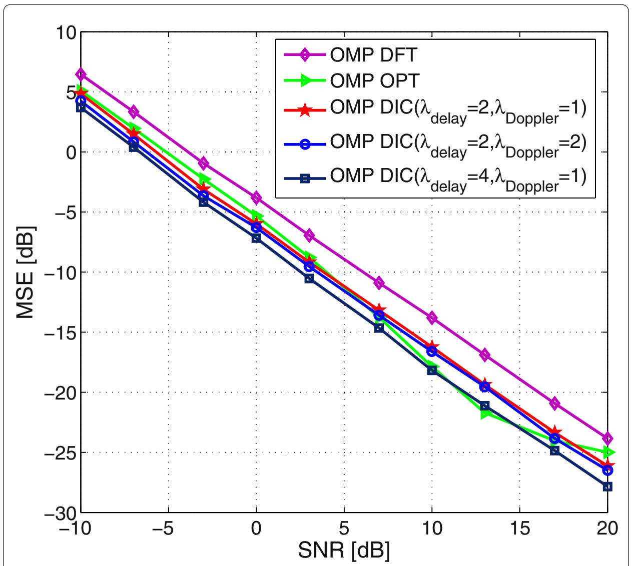

Figure 11 MSE performance using different resolution dictionaries. When $\lambda_{\text{delay}} = 4$, $\lambda_{\text{Doppler}} = 1$, the performance of channel estimation is the best, while the accuracy performance on $\lambda_{\text{delay}} = \lambda_{\text{Doppler}} = 2$ is almost the same as that on $\lambda_{\text{delay}} = 2$, $\lambda_{\text{Doppler}} = 1$; the over-sampling in the Doppler domain cannot improve the performance very much; the channel has a mild Doppler spread.

6 Conclusions

This paper proposes a novel estimation method of a sparse and doubly selective channel based on CS theory. The method can reduce the problem of energy leakage caused by discrete truncation and the limited bandwidth by over-sampling in the delay-Doppler domain, enhance the sparsity of the equivalent channel in the dictionary domain, and then improve the performance of channel estimation. The results show that although the method based on the over-complete dictionary needs more computation, the estimation accuracy is improved obviously and the pilot resource is reduced very much. Lastly, compared with

Table 2 The required run time about different algorithms

Algorithm	Reconstruction (s)	Extra time (s)
DFT	0.4543	0
OPT	0.3210	164.4662
DIC ($\lambda_{\text{delay}} = 2, \lambda_{\text{Doppler}} = 1$)	0.5343	0
DIC ($\lambda_{\text{delay}} = 2, \lambda_{\text{Doppler}} = 2$)	0.9310	0
DIC ($\lambda_{\text{delay}} = 4, \lambda_{\text{Doppler}} = 1$)	0.8603	0

the increase of spectrum utilization, it is worth for more complexity.

Competing interests
The authors declare that they have no competing interests.

Acknowledgements
This work was supported by the Fundamental and Frontier Research Project of Chongqing (cstc2013jcyjA40034), National Natural Science Foundation of China (61301126), Program for Changjiang Scholars and Innovative Research Team in University (IRT1299), and special fund of Chongqing key laboratory (CSTC).

Received: 31 October 2013 Accepted: 12 February 2014
Published: 6 March 2014

References
1. SF Cotter, BD Rao, Matching pursuit based decision-feedback equalizers. IEEE. Acoustics Signal. **5**, 2713–2716 (2000)
2. MR Raghavendra, K Giridhar, Improving channel estimation in OFDM systems for sparse multipath channels. IEEE Signal Process. Lett. **12**(1), 52–55 (2005)
3. DL Donoho, Compressed sensing. IEEE Trans. Inform. Theory. **52**(4), 1289–1306 (2006)
4. EJ Candes, J Romberg, T Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inform. Theory. **52**(2), 489–509 (2006)
5. WU Bajwa, J Haupt, G Raz, R Nowak, Compressed channel sensing. IEEE, Inform. Sci. Syst. **42**, 5–10 (2008). doi:10.1109/CISS.2008.4558485
6. WU Bajwa, AM Sayeed, R Nowak, Learning sparse doubly-selective channels. Commun. Control, Comput. **46**, 575–582 (2008). doi:10.1109/ALLERTON.2008.4797610
7. WU Bajwa, J Haupt, AM Sayeed, R Nowak, Compressed channel sensing: a new approach to estimating sparse multipath channels. Proc. IEEE. **98**(6), 1058–1076 (2010)
8. CR Berger, SL Zhou, W Chen, P Willett, Sparse channel estimation for OFDM: over-complete dictionaries and super-resolution. IEEE, Signal Process. Adv. Wireless Commun. **42**, 196–200 (2009). doi:10.1109/SPAWC.2009.5161774
9. CR Berger, SL Zhou, W Chen, JC Preisig, P Willett, Sparse channel estimation for multicarrier underwater acoustic communication: from subspace methods to compressed sensing. IEEE Trans. Signal Process. **58**(3), 1708–1721 (2010)
10. G Taubock, F Hlawatsch, A compressed sensing technique for OFDM channel estimation in mobile environments: exploiting channel sparsity for reducing pilots. IEEE Acoustics Speech Signal Process. ICASSP-08, 2885–2888 (2008). doi:10.1109/ICASSP.2008.4518252
11. G Taubock, F Hlawatsch, Compressed sensing based estimation of doubly selective channels using a sparsity-optimized basis expansion. In Processing European Signal Processing Conference (Lausanne, Switzerland, 2008), pp. 3029–3032
12. G Taubock, F Hlawatsch, D Eiwen, H Rauhut, Compressive estimation of doubly selective channels in multicarrier systems: leakage effects and sparsity-enhancing processing. IEEE J. Select. Topics Signal Process. **4**(2), 255–271 (2010)
13. JA Tropp, AC Gilbert, Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans. Inform. Theory. **53**(12), 267–288 (2007)
14. AM Sayeed, B Aazhang, Joint multipath-Doppler diversity in mobile wireless communications. IEEE Trans. Commun. **47**(1), 123–132 (1999)
15. GB Giannakis, C Tepedelenlioglu, Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels. Proc. IEEE. **86**(10), 1969–1986 (1998)
16. I Barhumi, G Leus, M Moonen, Estimation and direct equalization of doubly selective channels. EURASIP J. Appl. Signal Process., Article ID 62831, 1–15 (2006)
17. G Leus, On the estimation of rapidly time-varying channels. European Signal Processing Conference (EUSIPCO), Vienna, Austria, **12**, 2227–2230 (2004)
18. CR Berger, Z Wang, J Huang, S Zhou, Application of compressive sensing to sparse channel estimation. IEEE Commun. Mag. **48**(11), 164–174 (2010)
19. QH Shi, Y Karasawa, Frequency-domain oversampling for OFDM systems: exploiting inter-carrier interference and multipath diversity. Int. Symp. Commun. Inform. Technol. **9**, 1097–1101 (2009). doi:10.1109/ISCIT.2009.5341017
20. GD Galdo, M Haardt, IlmProp: a flexible geometry-based simulation environment for multiuser MIMO communication. European Cooperation in Science and Technology (EURO-COST), COST 273 TD(03) 188, Prague, Czech Republic (2003)

doi:10.1186/1687-6180-2014-29
Cite this article as: Zhou *et al.*: **A novel method for sparse channel estimation using super-resolution dictionary.** *EURASIP Journal on Advances in Signal Processing* 2014 **2014**:29.