Estimation of Sparse Time Dispersive SIMO Channels with Common Support in Pilot Aided OFDM Systems Using Atomic Norm

Slavche Pejoski[(✉)] and Venceslav Kafedziski

Faculty of Electrical Engineering and Information Technologies,
University Saints Cyril and Methodious Skopje, Skopje, Republic of Macedonia
{slavchep,kafedzi}@feit.ukim.edu.mk

Abstract. We consider the problem of estimation of sparse time dispersive channels in pilot aided OFDM systems on Single Input Multiple Output (SIMO) channels, i.e. with a single transmit and multiple receive antennas. In such systems the channels are inherently continuous-time and sparse, and there is a common support of the channel coefficients of channels associated with different antennas, resulting from the same scatterer. To exploit these properties, we propose a new channel estimation algorithm that combines the atomic norm minimization of the Multiple Measurement Vector (MMV) model, the MUSIC and the least squares (LS) methods. The atomic norm minimization of the MMV model allows to exploit the common support assumption and the continuous-time nature of the channels, MUSIC allows for simple joint estimation of the delays corresponding to the same scatterer, and LS allows for estimation of the path gains. To evaluate the proposed algorithm, we compare its performance with the case when the common support assumption is not used.

Keywords: Channel estimation · Joint atomic norm minimization · SIMO channel · Pilot aided OFDM

1 Introduction

Channel estimation is essential in contemporary communication systems. Many of these systems are OFDM based and use pilot subcarriers for channel estimation. Additionally, most systems are designed to work with multiple antennas at the receive end. Here we consider such systems in a scenario where the channels between the transmit and the receive antennas are time dispersive and sparse in the sense that there are few strong channel paths that are resolvable. With the introduction of compressed sensing (CS), sparse channels have received significant attention [1–4], but CS, in its ordinary form, is not well suited for channel estimation since the delays of different paths can be arbitrary within a certain range (due to the channel continuous-time nature) and CS requires the delays to be on a predefined grid of delays. The recent framework of atomic norm

minimization [5,6] allows for estimation of sparse quantities not falling on a pre-defined grid. In [7], we used this framework to propose an algorithm for channel estimation of sparse time dispersive channel in pilot aided OFDM system with single antennas at both the transmit and receive ends. Here we extend the algorithm to the SIMO case. It has been shown in [8,9] that when the signal to noise ratio is not too high, the delays of the paths corresponding to the same scatterer, received on different receive antennas, are not clearly distinguishable and can be treated as being the same. This leads to the common support assumption of the coefficients in the channel impulse responses, which we exploit to jointly estimate the delays corresponding to the same scatterer, by modifying the algorithm from [7]. In fact, we replace the atomic norm minimization from [7] with the atomic norm minimization for the Multiple Measurement Vector (MMV) model [10,11]. The algorithms presented in [8,9] require that the pilot subcarriers are equidistant, and in the proposed algorithm we place them at random positions [7], resulting in the ordinary CS subsampling.

The novelty of this paper comes from introducing the atomic norm minimization in the estimation of sparse time dispersive channels with common support in pilot aided OFDM SIMO systems.

The paper is organized as follows. In Sect. 2 we describe the OFDM channel estimation problem in time dispersive SIMO channels, its connection to the atomic norm minimization of the MMV model and the new algorithm. In Sect. 3 we show simulation results. Section 4 concludes the paper.

2   Problem and Algorithm Description

We investigate an OFDM system with a single transmit antenna and N_r receive antennas. The system uses N subcarriers to transmit pilot or data symbols. The OFDM signal between the transmit antenna and the r-th receive antenna is sent through a time dispersive channel whose baseband channel impulse response, when I specular (point) scatterers are present, during a single frame transmission (assuming a block fading model), is [2,4]:

$$h_r(\tau) = \sum_{i=0}^{I-1} h_{r,i}\delta(\tau - \tau_{r,i}) \tag{1}$$

where $h_{r,i}$ and $\tau_{r,i}$ are the complex gain and the delay associated with the i-th path. For this channel model, the channel estimation is carried out for each block and the pilots are used in a single OFDM symbol in each block. At the receiver, after processing the signal at antenna r, $r = 1, ..., N_r$ we obtain:

$$Y_{0_r}(n) = H_r(n)X(n) + W_r(n), \quad n = 0, ..., N-1 \ \ r = 1, ..., N_r \tag{2}$$

where $X(n)$ is the pilot/data symbol sent on the n-th subcarrier, $W_r(n)$ is the noise sample at the n-th subcarrier and r-th antenna (noise samples are

independent in both the frequency and space coordinates and are modeled as zero mean circularly symmetric complex Gaussian variables with variance $\sigma_{n_0}^2$)

$$H_r(n) = \sum_{i=0}^{I-1} h_{r,i} e^{-j2\pi \frac{n}{N} \frac{\tau_{r,i}}{T_s}}, \quad n = 0, ..., N-1 \ \ r = 1, ..., N_r \tag{3}$$

and T_s is the sampling interval. We introduce a constant L, such that $0 \leq \tau_{r,i} \leq L \leq L_{cp}$ (where L_{cp} is the length of the cyclic prefix), and, assume that L is an integer of the form $L = N/D$ for an integer D. In the rest of the paper, the estimation of $H_r(n)$ for $n = 0, ..., N-1$ and $r = 1, ..., N_r$ is termed the channel estimation problem. We set $\mathbf{H}_r = [H_r(0), H_r(1), \ldots, H_r(N-1)]^T$.

Since the true $\tau_{r,i}$'s are continuous, the estimation of $H_r(n)$ can be formulated as a CS problem based on the atomic norm minimization [5,6,10,11] (such formulation for the SISO channel is proposed in [7] but it does not exploit the common support assumption, and, thus, it can be used only for independent channel estimation for each r). To decrease the complexity of the solution for such formulation, as explained in [7], the pilot subcarriers are allocated at P positions (since the channel is sparse $P < L$) n_p, $p = 0, \ldots, P-1$, that create a randomly chosen subset of the set of the equidistant L positions which are at integer multiples of N/L i.e. $n_p \in \{0, \frac{N}{L}, ..., \frac{(L-1)N}{L}\}$. With such a pilot allocation scheme, setting $n_p^{\grave{}} = n_p \frac{L}{N}$, $p = 0, ..., P-1$, using equi-powered (constant amplitude and random phase) pilot symbols and dividing the received samples $Y_{0_r}(n_p)$ with $X(n_p)$ we obtain:

$$Y_r(n_p \frac{N}{L}) = \sum_{i=0}^{I-1} h_{r,i} e^{-j2\pi \frac{\tau_{r,i}}{LT_s} n_p^{\grave{}}} + W_{1_r}(n_p \frac{N}{L}) \tag{4}$$

where $W_{1_r}(n_p \frac{N}{L})$ are zero mean circularly symmetric complex Gaussian variables with variance σ_n^2, independent in both the frequency and space coordinates. Equation (4) represents a model with subsampled measurements of a signal in noise. It should be noted that the subcarriers not carrying pilot symbols can be used to transmit data symbols which increases the system capacity.

For clarity, we first introduce the atomic norm and an algorithm for independent channel estimation (based on [7]) and then extend it to the MMV model. The atomic norm of $\mathbf{x} \in \mathbb{C}^{U \times 1}$ [5] is:

$$||\mathbf{x}||_{\mathcal{A}} = \inf_{\substack{c_i \geq 0 \\ \phi_i \in [0, 2\pi) \\ f_i \in [0,1)}} \{\sum_i c_i : \mathbf{x} = \sum_i c_i \mathbf{a}(f_i, \phi_i)\} \tag{5}$$

where $\mathcal{A} = \{\mathbf{a}(f, \phi) : f \in [0, 1), \phi \in [0, 2\pi)\}$ is the set of atom vectors $\mathbf{a}(f, \phi)$ whose components are $a_u(f, \phi) = e^{-j(2\pi f u + \phi)}$ for $u = 0, ..., U-1$, and $c_i \in \mathbb{R}$. As explained in [5], the SDP form of $||\mathbf{x}||_{\mathcal{A}}$ is:

$$||\mathbf{x}||_{\mathcal{A}} = \inf_{\mathbf{x}, \mathbf{u}, t} \{\frac{\text{trace}(\text{Toep}(\mathbf{u}))}{2U} + \frac{t}{2} : \begin{bmatrix} \text{Toep}(\mathbf{u}) & \mathbf{x} \\ \mathbf{x}^* & t \end{bmatrix} \succeq 0\} \tag{6}$$

where $\succeq 0$ stands for a positive semidefinite matrix, Toep($\mathbf{u}$) is a Toeplitz matrix created using the elements of $\mathbf{u}$ and trace($\cdot$) is the matrix trace.

In the approach with independent channel estimation for each r, $r = 1, \ldots, N_r$, we estimate each $\mathbf{H}_r$ as in [7]. We use the mapping $f_i = \frac{\tau_{r,i}}{LT_s} \in [0,1)$, $c_i e^{-j\phi_i} = h_{r,i}$, $U = L$, introduce the notation $\mathbf{y}_r = [Y_r(n_0'\frac{N}{L}), \ldots, Y_r(n_{P-1}'\frac{N}{L})]^T$, $\mathbf{H}_{L,r} = [H_r(0), \ldots, H_r(\frac{(L-1)N}{L})]^T$ and $\mathbf{H}_{P,r} = [H_r(n_0'\frac{N}{L}), \ldots, H_r(n_{P-1}'\frac{N}{L})]^T$, and estimate $\mathbf{H}_{L,r}$ for the subsampled signal in noise scenario as [7]:

$$\min_{\mathbf{H}_{L,r}} \mu\|\mathbf{H}_{L,r}\|_{\mathcal{A}} + \frac{1}{2}\|\mathbf{y}_r - \mathbf{H}_{P,r}\|_2^2 \tag{7}$$

where μ is a constant. Using the SDP form of the atomic norm (6) and the reconstruction algorithm in the subsampled noisy case (7), the estimates of $\mathbf{H}_{L,r}$ and $\mathbf{u}_r$ can be obtained as:

$$[\hat{\mathbf{H}}_{L,r}, \hat{\mathbf{u}}_r] = \arg\min_{\mathbf{H}_{L,r}, \mathbf{u}_r, t_r} \mu\left(\frac{\text{trace}(\text{Toep}(\mathbf{u}_r))}{2L} + \frac{t_r}{2}\right)$$
$$+ \frac{1}{2}\|\mathbf{y}_r - \mathbf{H}_{P,r}\|_2^2$$
$$\text{subject to} \begin{bmatrix} \text{Toep}(\mathbf{u}_r) & \mathbf{H}_{L,r} \\ \mathbf{H}_{L,r}^* & t_r \end{bmatrix} \succeq 0 \tag{8}$$

where μ can be estimated as $\frac{\rho}{\rho-1}\sqrt{L(\ln \overline{M} + \ln(\pi\rho) + 1)}\sigma_n$ with $\rho = \lim_{k\to\infty} \rho_k$, $\rho_{k+1} = 2\ln\rho_k + 2\ln(\pi\overline{M}) + 3$ for $\rho_0 > 2$ and $\overline{M} = n_{P-1}' - n_0' + 1$ [6].

As explained in [7], the estimates $\hat{\tau}_{r,i}$ of $\tau_{r,i}$ can be obtained from the Toeplitz matrix Toep($\hat{\mathbf{u}}_r$) using root MUSIC. As in [7] we assume that I is known (otherwise, it can be estimated as explained in [6]). Having obtained $\hat{\tau}_{r,i}$'s, we estimate the $h_{r,i}$'s using the LS method:

$$\hat{\mathbf{h}}_{I,r} = \arg\min_{\mathbf{h}_{I,r}} \|\mathbf{D}_r\mathbf{h}_{I,r} - \mathbf{y}_r\|_2^2 \tag{9}$$

where $\mathbf{D}_r$ is a $P \times I$ matrix with elements $[\mathbf{D}_r]_{p,i} = e^{-j2\pi\frac{\hat{\tau}_{r,i}}{LT_s}n_p'}$, $i = 0, \ldots, I-1$, $p = 0, \ldots, P-1$ and $\mathbf{h}_{I,r}$ is an $I \times 1$ vector that contains the $h_{r,i}$'s. The estimated $\hat{\tau}_{r,i}$'s and $\hat{h}_{r,i}$'s are used in (3) instead of the $\tau_{r,i}$'s and the $h_{r,i}$'s to estimate $\mathbf{H}_r$. To obtain the estimates of $\mathbf{H}_r$ of all N_r channels the procedure is repeated N_r times ($r = 1, \ldots, N_r$). It should be noted that if the correlations among the channels are available at the receiver, $\hat{\mathbf{h}}_{I,r}$ can be obtained as an MMSE estimate instead of using the LS method, which may improve the method performance, but the evaluation of such improvement is left for future work.

If the signal to noise ratio is not high, then it can be assumed that [8,9]:

$$\tau_{1,i} \approx \tau_{2,i} \approx \ldots \approx \tau_{N_r,i} \quad \text{for} \quad i = 1, \ldots, I \tag{10}$$
$$h_{1,i} \neq h_{2,i} \neq \ldots \neq h_{N_r,i} \quad \text{for} \quad i = 1, \ldots, I \tag{11}$$

which means that the different channels have common support. It should be noted that the correlation between $h_{r,i}$ for different r and given i depends on the

system parameters (carrier frequency and distance between receive antennas). The expressions for calculating its value can be found in [9], and, for current communication systems its value is below 0.5 [8].

Based on (10) and (11) we reformulate the channel estimation problem as an atomic norm minimization of the MMV model. Namely, the definition of the atomic norm of matrix $\mathbf{X}$ of dimensions $U \times N_r$ is [10,11]:

$$||\mathbf{X}||_{\mathcal{A}} = \inf_{\substack{b_i \geq 0, b_i \in \mathbb{R} \\ \boldsymbol{\psi}_i \in \mathbb{C}^{1 \times N_r}, ||\boldsymbol{\psi}_i||_2 = 1 \\ f_i \in [0,1)}} \{\sum_i b_i : \mathbf{X} = \sum_i b_i \mathbf{a}(f_i, 0) \boldsymbol{\psi}_i\} \tag{12}$$

When $b_i \neq 0$ in (12), $\mathbf{a}(f_i, 0)$ can be included in the creation of each column of $\mathbf{X}$ with arbitrary phase and gain that are included in the value of the corresponding element of $\boldsymbol{\psi}_i$ (the influence of the gains is partially contained in $\boldsymbol{\psi}_i$ and in b_i). So, each column of $\mathbf{X}$ is constructed of the same atoms $\mathbf{a}(f_i, 0)$ but mixed with different complex gains which means that each column of $\mathbf{X}$ can be considered as a different measurement vector. By constructing $\mathbf{H}_L = [\mathbf{H}_{L,1}, \mathbf{H}_{L,2}, ..., \mathbf{H}_{L,N_r}]$, $\mathbf{H}_P = [\mathbf{H}_{P,1}, \mathbf{H}_{P,2}, ..., \mathbf{H}_{P,N_r}]$ and $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_{N_r}]$ the joint estimation of the channel delays can be carried out using an SDP program for the atomic norm minimization of the MMV model [11]:

$$[\hat{\mathbf{H}}_L, \hat{\mathbf{u}}] = \arg \min_{\mathbf{H}_L, \mathbf{u}, \mathbf{W}} \mu_X \left(\frac{\text{trace}(\text{Toep}(\mathbf{u}))}{2} + \frac{\text{trace}(\mathbf{W})}{2} \right)$$
$$+ \frac{1}{2} ||\mathbf{Y} - \mathbf{H}_P||_2^2$$
$$\text{subject to } \begin{bmatrix} \text{Toep}(\mathbf{u}) & \mathbf{H}_L \\ \mathbf{H}_L^* & \mathbf{W} \end{bmatrix} \succeq 0 \tag{13}$$

where $\text{Toep}(\mathbf{u})$ is of the form $\text{Toep}(\mathbf{u}) = \mathbf{A}\mathbf{K}\mathbf{A}^H$ for $\mathbf{A} = [\mathbf{a}(f_0, 0) ... \mathbf{a}(f_{I-1}, 0)]$, $\mathbf{K} = \text{diag}([k_0 ... k_{I-1}])$ ($k_i, i = 0, ..., I-1$ are positive real numbers). Using $\text{Toep}(\hat{\mathbf{u}})$ from (13), the estimation of $\tau_{r,i}$ is jointly carried out for all $r = 1, ..., N_r$ using root MUSIC [7]. Having obtained $\hat{\tau}_{r,i}$, the $h_{r,i}$'s are estimated using (9). In (13), μ_X can be estimated as $\sigma_n \sqrt{1 + \frac{1}{L}} \sqrt{N_r + \ln \alpha N_r + \sqrt{2N_r \ln \alpha N_r} + \sqrt{\frac{\pi N_r}{2}} + 1}$ for $\alpha = 4\pi \overline{M} \ln L$, which is obtained by combining the results from [6] and [11].

3   Numerical Results

To evaluate the performance of the proposed algorithm we carried out Matlab simulations. We assumed $N = 512$ and $L = 64$ in the OFDM system and used CVX [12] for solving the SDP programs. In each simulation run we used different realizations of the channel impulse responses. To generate the $\tau_{r,i}$'s we first generated I values τ_i from a uniform distribution on $[2T_s, (L-2)T_s)$ such that $|\tau_i - \tau_j| \geq 1.5T_s$ for $i \neq j$ (see [5–7]), then generated $\Delta\tau_{r,i}$'s for $r = 1, ..., N_r$ and $i = 1, ..., I$ to have a uniform distribution on $[-\frac{T_s}{50}, \frac{T_s}{50}]$, and

obtained the $\tau_{r,i}$'s as $\tau_{r,i} = \tau_i + \Delta\tau_{r,i}$. We generated the channel gains $h_{r,i}$ as zero mean and unit variance circularly symmetric complex Gaussian variables, independent in both the frequency and space variable. We chose the P pilot positions from the L available using equal probability for each position. We called the algorithm that uses (13) 'JAtomSR' and we called the algorithm that uses (8) 'AtomSR'. Additionally, when appropriate, we showed the MSE of the LS estimates obtained if all N subcarriers were used as pilot subcarriers transmitting equi-powered symbols and the estimation was carried out as $H_r(n) = Y_{0_r}(n)/X(n), n = 0, \ldots, N-1, r = 1, ..., N_r$ (termed 'Full LS'). In (8) we used μ obtained by scaling the value below (8) by $1/1.4$ and in (13) for μ_X we used the value below (13) scaled by $1/1.7$. To obtain each point in each plot we averaged the results from 500 simulation runs. As a performance criterion we used the average per sample and per antenna mean squared error defined as:

$$MSE = \mathbb{E}\left[\frac{1}{NN_r} \sum_{r=0}^{N_r} \sum_{n=0}^{N-1} |\hat{H}_r(n) - H_r(n)|^2 \right] \tag{14}$$

A performance comparison in terms of P and I is shown in Fig. 1. The lower bound, assuming perfect knowledge of $\tau_{r,i}$'s prior to the LS step (termed 'Known delays') is also shown. JAtomSR shows significant performance gain over AtomSR due to the utilization of the common support assumption, and its performance is very close to the lower bound. The improvement is highest in regions where P is small and/or I is high, and as P increases or I decreases the improvement decreases. By comparing Full LS and JAtomSR in Fig. 1(b) we conclude that exploiting the channel sparsity leads to reduced MSE even when low number of pilot subcarriers are used. It should be noted that the Full LS algorithm has low complexity, negligible compared to JAtomSR.

Performance comparison in terms of N_r and the inverse noise power $10\log_{10}\frac{1}{\sigma_n^2}$ is shown in Fig. 2. The improvement of JAtomSR over AtomSR, shown in Fig. 2(a), increases as N_r increases, but the improvement slope is highest for $N_r = 2$ and
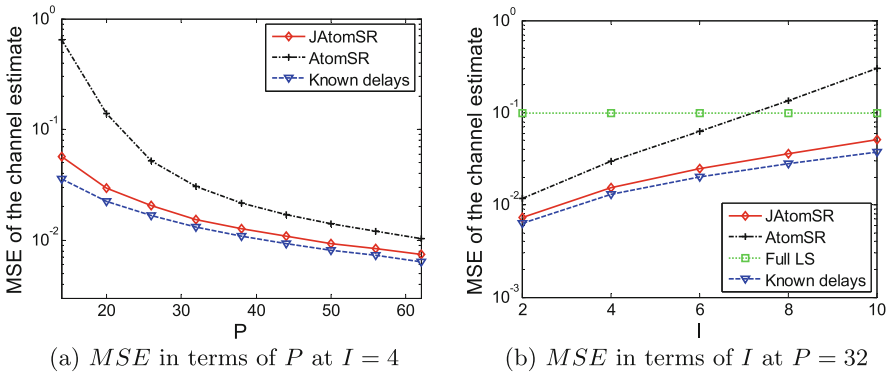

(a) MSE in terms of P at $I = 4$ (b) MSE in terms of I at $P = 32$

Fig. 1. The MSE in terms of P and I at $10\log_{10}\frac{1}{\sigma_n^2} = 10dB$, $N_r = 5$

(a) MSE in terms of N_r at $10 \log_{10} \frac{1}{\sigma_n^2} =$ 10 dB and $I = 8$

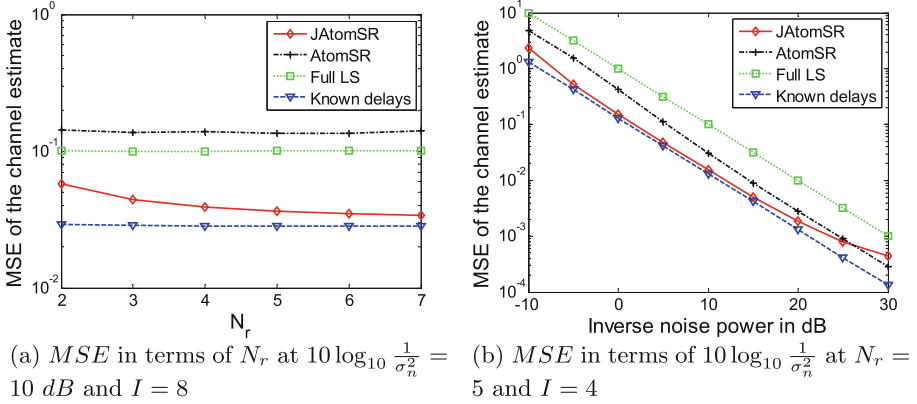(b) MSE in terms of $10 \log_{10} \frac{1}{\sigma_n^2}$ at $N_r =$ 5 and $I = 4$

Fig. 2. The MSE in terms of N_r and the inverse noise power at $P = 32$

decreases with the increase of N_r. Figure 2(b) shows that the improvement of JAtomSR compared to AtomSR is almost constant for broad range of inverse noise powers. However, in the high inverse noise power region the performance of AtomSR becomes comparable to and even better than the performance of JAtomSR. In this region, using the common support assumption does not help anymore. Namely, here the noise power is very low and AtomSR is capable of estimating the $\tau_{r,i}$ with such a precision that the evaluated different delays at different antennas, associated with the same scatterer, significantly decrease the MSE. The complexity analysis of JAtomSR is left for future work.

4   Conclusion

We proposed the use of atomic norm in the estimation of sparse time dispersive SIMO channels with common support in pilot aided OFDM systems. The proposed algorithm uses specific subsampled subcarrier pilot allocation scheme and is based on the atomic norm minimization of the Multiple Measurement Vector model to jointly estimate the delays of the paths at different antennas, corresponding to the same scatterer. The simulation results show that the performance improvement compared to the algorithm that does not use the common support assumption, increases with the increase of the number of receive antennas and is significant in the low to medium SNR region. At high SNR the common support assumption does not hold and the performance of the proposed algorithm degrades. The proposed algorithm can be also used in a MIMO OFDM system, where specific orthogonal pilot patterns are used in the different OFDM symbols of the training sequence, which allows the estimation of the channels between any transmit and each receive antenna.

References

1. Berger, C.R., Wang, Z., Huang, J., Zhou, S.: Application of compressive sensing to sparse channel estimation. IEEE Commun. Mag. **48**, 164–174 (2010)
2. Taubock, G., Hlawatsch, F., Eiwen, D., Rauhut, H.: Compressive estimation of doubly selective channels in multicarrier systems: leakage effects and sparsity-enhancing processing. IEEE J. Sel. Top. Sig. Process. **4**(2), 255–271 (2010)
3. Cheng, P., Chen, Z., Rui, Y., Guo, Y.J., Gui, L., Tao, M., Zhang, Q.T.: Channel estimation for OFDM systems over doubly selective channels: a distributed compressive sensing based approach. IEEE Trans. Comm. **61**, 4173–4185 (2013)
4. Hu, D., Wang, X., He, L.: A new sparse channel estimation and tracking method for time-varying OFDM systems. IEEE Tran. Veh. Tech. **62**, 4848–4653 (2013)
5. Tang, G., Bhaskar, B.N., Shah, P., Recht, B.: Compressed sensing off the grid. IEEE Trans. Inf. Theor. **59**, 7465–7490 (2013)
6. Yang, Z., Xie, L.: On Gridless sparse methods for line spectral estimation from complete and incomplete data. IEEE Trans. Sig. Process. **63**, 3139–3153 (2015)
7. Pejoski, S., Kafedziski, V.: Estimation of sparse time dispersive channels in pilot aided OFDM using atomic norm. IEEE Wirel. Commun. Lett. **4**, 397–400 (2015)
8. Barbotin, Y., Hormati, A., Rangan, S., Vetterli, M.: Estimation of sparse MIMO channels with common support. IEEE Trans. Comm. **60**, 3705–3716 (2012)
9. Barbotin, Y., Vetterli, M.: Fast and robust parametric estimation of jointly sparse channels. IEEE J. Emerg. Sel. Top. Circ. Syst. **2**, 402–412 (2012)
10. Yang, Z., Xie, L.: Exact Joint Sparse Frequency Recovery via Optimization Methods (2014). arXiv:1405.6585
11. Li, Y., Chi, Y.: Off-the-Grid Line Spectrum Denoising and Estimation with Multiple Measurement Vectors (2014). arXiv:1408.2242
12. Grant, M., Boyd, S.: CVX: Matlab Software for Disciplined Convex Programming, version 2.0 beta (2013). http://cvxr.com/cvx