A Single Source Point Detection Algorithm for Underdetermined Blind Source Separation Problem

Yu Zhang[1], Zhaoyue Zhang[2], Hongxu Tao[1], and Yun Lin[1](✉)

[1] College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China
linyun_phd@hrbeu.edu.cn

[2] College of Air Traffic Management, Civil Aviation University of China, Tianjin 300300, China

Abstract. To overcome the traditional disadvantages of single source points detection methods in underdetermined blind source separation problem, this paper proposes a novel algorithm to detect single source points for the linear instantaneous mixed model. First, the algorithm utilizes a certain relationship between the time-frequency coefficients and the complex conjugate factors of the observation signal to realize single source points detection. Then, the algorithm finds more time-frequency points that meets the requirements automatically and cluster them by utilizing a clustering algorithm based on the improved potential function. Finally, the estimation of the mixed matrix is achieved by clustering the re-selected single source points. Simulation experiments on linear mixture model demonstrates the efficiency and feasibility for estimating the mixing matrix.

Keywords: Time-frequency domain · Mixing matrix estimation · Single source points detection

1 Introduction

Blind source separation (BSS) aims at recovering N source signals from P observation signals without any prior information. So far, research in this field has been widely applied to mechanical equipment fault diagnosis [8], speech signals [7], communication systems [9, 15, 16], etc. Our research is based on the underdetermined case, i.e. the number of source signals is greater than the number of observed signals $(N > P)$.

At present, sparse component analysis (SCA) [12] is the most commonly used method to solve the problem of Underdetermined Blind Source Separation (UBSS). SCA usually adopts a two-step method that includes mixing matrix estimation and source signal recovery. The accuracy of the former directly affects the result of the latter, so research on the former is quite meaningful. In this paper, we aim at estimating the mixing matrix. The BSS algorithm based on single source points detection usually has high requirements on the sparsity of signals. In fact, signals have satisfactory sparsity in the time-frequency (TF) domain than in the time domain. Short-time Fourier transform (STFT) [11] is usually used to make signals get better sparsity.

S. Liu and G. Yang (Eds.): ADHIP 2018, LNICST 279, pp. 68–76, 2019.
https://doi.org/10.1007/978-3-030-19086-6_8

Only one source exists or plays a major role in TF domain are called single source points. If the source signal is sparse in the TF domain, then the observed signal will exhibit directional clustering property so that the mixed matrix can be estimated by utilizing the corresponding clustering algorithm. The direction corresponds to one column of the mixed matrix. In other words, the mixing matrix can be estimated if the direction of the single source points is estimated. Many scholars researched different approaches in this field. Some scholars [1, 3, 6, 14] made strategies to search for single source regions and then each element in the mixing matrix is estimated from the region. Some scholars [2–5, 10, 13] proposed various algorithm to realize the detection of single source points and finally the mixing matrix is estimated. This paper proposes a novel algorithm to estimate the mixing matrix based on single source points detection. The algorithm gets satisfactory performance than other algorithms.

This paper is organized as follows. In Sect. 2, we introduce the basic linear instantaneous mixed model and the basic theory of UBSS problem. Section 3 shows the process of our algorithm. We then give the simulation experiment results in Sect. 4 and draw conclusion in Sect. 5.

2 Problem Formulation

The linear instantaneous mixed model of BSS problems in the noiseless case can be expressed as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) = \sum_{m=1}^{M} \mathbf{a}_m s_m(t) \tag{1}$$

Where $M > N$, $\mathbf{x}(t) = [x_1(t), x_2(t), \ldots, x_N(t)]^T$ is the observation signal vector, $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_M] \in \mathbb{R}^{N \times M}$ is the mixed matrix, $\mathbf{s}(t) = [s_1(t), s_2(t), \ldots, S_M(t)]^T$ is the source signal vector, $\mathbf{a}_m$ is the mth column of the mixed matrix and $s_m(t)$ is the mth source signal. If only the mth source signal presents at t, Eq. (1) can be simplified as

$$\mathbf{x}(t) = \mathbf{a}_m s_m(t) \tag{2}$$

Under the condition of neglecting the amplitude, estimating the direction of the mixing signal vector also realizes the estimation of the first column vector of the mixed matrix. If the observation signal is sufficiently sparse, then all similar direction vectors can be obtained by clustering and the mixing matrix can be successfully estimated.

In the UBSS method, the necessary assumptions need to be satisfied. On one hand, the mixed matrix should be full column rank, On the other hand, there should be some single source points exists in the TF domain.

3 The Proposed Algorithm

We usually adopt STFT before estimating the mixing matrix to make signal sparser, we can obtain representations of the mixture signals

$$\mathbf{X}(t,f) = \mathbf{AS}(t,f) \tag{3}$$

Where $\mathbf{X}(t,f) = [X_1(t,f), X_2(t,f), \ldots, X_N(t,f)]^T$ and $\mathbf{S}(t,f) = [S_1(t,f), S_2(t,f), \ldots, S_M(t,f)]^T$ are the STFT coefficients of observation signals and source signals, respectively. The paper takes two observation signals and four source signals for example, so Eq. (3) can be written as

$$\begin{bmatrix} X_1(t,f) \\ X_2(t,f) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1M} \\ a_{21} & a_{22} & \cdots & a_{2M} \end{bmatrix} \begin{bmatrix} S_1(t,f) \\ S_2(t,f) \\ \vdots \\ S_M(t,f) \end{bmatrix} \tag{4}$$

Assuming that there only source s_1 occurs at one TF point (t_p, f_p), we can obtain the following two formulas

$$X_1(t_p,f_p) = a_{11}S_1(t_p,f_p) = a_{11}[\mathrm{Re}(S_1(t_p,f_p)) + j\,\mathrm{Im}(S_1(t_p,f_p))] \tag{5}$$

$$X_2(t_p,f_p) = a_{21}S_1(t_p,f_p) = a_{21}[\mathrm{Re}(S_1(t_p,f_p)) + j\,\mathrm{Im}(S_1(t_p,f_p))] \tag{6}$$

Based on Eqs. (5) and (6), we have

$$X_1^*(t_p,f_p) = a_{11}\{\mathrm{Re}[S_1(t_p,f_p)] - j\,\mathrm{Im}[S_1(t_p,f_p)]\} \tag{7}$$

$$X_2^*(t_p,f_p) = a_{21}\{\mathrm{Re}[S_1(t_p,f_p)] - j\,\mathrm{Im}[S_1(t_p,f_p)]\} \tag{8}$$

where $X_1^*(t_p,f_p)$ and $X_2^*(t_p,f_p)$ are complex conjugates of $X_1(t_p,f_p)$ and $X_2(t_p,f_p)$, respectively. Based on Eqs. (5)–(8), we have

$$\begin{aligned} &\frac{X_1(t_p,f_p)X_2^*(t_p,f_p)}{X_2(t_p,f_p)X_1^*(t_p,f_p)} \\ &= \frac{a_{11}a_{21}\left[\mathrm{Re}\left(S_1\left(t_p,f_p\right)\right) + j\,\mathrm{Im}\left(S_1\left(t_p,f_p\right)\right)\right]\left[\mathrm{Re}\left(S_1\left(t_p,f_p\right)\right) - j\,\mathrm{Im}\left(S_1\left(t_p,f_p\right)\right)\right]}{a_{11}a_{21}\left[\mathrm{Re}\left(S_1\left(t_p,f_p\right)\right) + j\,\mathrm{Im}\left(S_1\left(t_p,f_p\right)\right)\right]\left[\mathrm{Re}\left(S_1\left(t_p,f_p\right)\right) - j\,\mathrm{Im}\left(S_1\left(t_p,f_p\right)\right)\right]} \\ &= 1 \end{aligned} \tag{9}$$

If two signals s_1 and s_2 are assumed to exist at some TF point (t_q, f_q), If we simplify $S_1(t_q,f_q)$ and $S_2(t_q,f_q)$ as S_1 and S_2, $X_1(t_q,f_q)$ and $X_2(t_q,f_q)$ can be simplified as X_1 and X_2

$$X_1 = [a_{11}\mathrm{Re}(S_1) + a_{12}\mathrm{Re}(S_2)] + j[a_{11}\mathrm{Im}(S_1) + a_{12}\mathrm{Im}(S_2)] \tag{10}$$

$$X_2 = [a_{21}\mathrm{Re}(S_1) + a_{22}\mathrm{Re}(S_2)] + j[a_{21}\mathrm{Im}(S_1) + a_{22}\mathrm{Im}(S_2)] \tag{11}$$

Similarly, $X_1^*(t_p,f_p)$ and $X_2^*(t_p,f_p)$ can be defined as

$$X_1^* = [a_{11}\mathrm{Re}(S_1) + a_{12}\mathrm{Re}(S_2)] - j[a_{11}\mathrm{Im}(S_1) + a_{12}\mathrm{Im}(S_2)] \tag{12}$$

$$X_2^* = [a_{21}\mathrm{Re}(S_1) + a_{22}\mathrm{Re}(S_2)] - j[a_{21}\mathrm{Im}(S_1) + a_{22}\mathrm{Im}(S_2)] \tag{13}$$

Based on Eqs. (10)–(13), we can obtain

$$\begin{aligned} X_1X_2^* = {} & [a_{11}\mathrm{Re}(S_1) + a_{12}\mathrm{Re}(S_2)][a_{21}\mathrm{Re}(S_1) + a_{22}\mathrm{Re}(S_2)] \\ & + [a_{11}\mathrm{Im}(S_1) + a_{12}\mathrm{Im}(S_2)][a_{21}\mathrm{Im}(S_1) + a_{22}\mathrm{Im}(S_2)] \\ & + j[a_{11}\mathrm{Im}(S_1) + a_{12}\mathrm{Im}(S_2)][a_{21}\mathrm{Re}(S_1) + a_{22}\mathrm{Re}(S_2)] \\ & - j[a_{11}\mathrm{Re}(S_1) + a_{12}\mathrm{Re}(S_2)][a_{21}\mathrm{Im}(S_1) + a_{22}\mathrm{Im}(S_2)] \end{aligned} \tag{14}$$

$$\begin{aligned} X_2X_1^* = {} & [a_{11}\mathrm{Re}(S_1) + a_{12}\mathrm{Re}(S_2)][a_{21}\mathrm{Re}(S_1) + a_{22}\mathrm{Re}(S_2)] \\ & + [a_{11}\mathrm{Im}(S_1) + a_{12}\mathrm{Im}(S_2)][a_{21}\mathrm{Im}(S_1) + a_{22}\mathrm{Im}(S_2)] \\ & - j[a_{11}\mathrm{Im}(S_1) + a_{12}\mathrm{Im}(S_2)][a_{21}\mathrm{Re}(S_1) + a_{22}\mathrm{Re}(S_2)] \\ & + j[a_{11}\mathrm{Re}(S_1) + a_{12}\mathrm{Re}(S_2)][a_{21}\mathrm{Im}(S_1) + a_{22}\mathrm{Im}(S_2)] \end{aligned} \tag{15}$$

The following two variables are assumed

$$\begin{aligned} T_1 = {} & [a_{11}\mathrm{Re}(S_1) + a_{12}\mathrm{Re}(S_2)][a_{21}\mathrm{Re}(S_1) + a_{22}\mathrm{Re}(S_2)] \\ & + [a_{11}\mathrm{Im}(S_1) + a_{12}\mathrm{Im}(S_2)][a_{21}\mathrm{Im}(S_1) + a_{22}\mathrm{Im}(S_2)] \end{aligned} \tag{16}$$

$$\begin{aligned} T_2 = {} & [a_{11}\mathrm{Re}(S_1) + a_{12}\mathrm{Re}(S_2)][a_{21}\mathrm{Re}(S_1) + a_{22}\mathrm{Re}(S_2)] \\ & - [a_{11}\mathrm{Im}(S_1) + a_{12}\mathrm{Im}(S_2)][a_{21}\mathrm{Im}(S_1) + a_{22}\mathrm{Im}(S_2)] \end{aligned} \tag{17}$$

Equations (14) and (15) can be simplified as

$$X_1X_2^* = T_1 + jT_2 \tag{18}$$

$$X_2X_1^* = T_1 - jT_2 \tag{19}$$

Then, we can obtain

$$\frac{X_1X_2^*}{X_2X_1^*} = \frac{T_1 + jT_2}{T_1 - jT_2} = \frac{T_1^2 - T_2^2}{T_1^2 + T_2^2} + j\frac{2T_1T_2}{T_1^2 + T_2^2} \tag{20}$$

If we want the Eq. (20) is equal to Eq. (9), we can get anyone of the following two conditions through setting T_2 as 0.

$$\frac{a_{11}}{a_{21}} = \frac{a_{12}}{a_{22}} \tag{21}$$

$$\frac{\mathrm{Re}(S_1)}{\mathrm{Re}(S_2)} = \frac{\mathrm{Im}(S_1)}{\mathrm{Im}(S_2)} \tag{22}$$

Given the assumption that the mixing matrix should be full column rank, we don't consider Eq. (21). Therefore, only when Eq. (22) is satisfied can Eq. (20) achieve the same consequence in Eq. (9). However, the probability of this situation is very low. Therefore, we set the following standard to detect single source points.

$$\frac{X_1(t,f)X_2^*(t,f)}{X_2(t,f)X_1^*(t,f)} = 1 \tag{23}$$

In practical applications, this condition is very demanding and difficult to achieve, so the relaxation condition is

$$\left|\mathrm{Re}\left(\frac{X_1(t,f)X_2^*(t,f)}{X_2(t,f)X_1^*(t,f)}\right) - 1\right| < \varepsilon_1 \tag{24}$$

where ε_1 is a positive number that is close to 0.

After selecting the corresponding single source point, there are still some time-frequency points with low energy, which seriously affects the later estimation result. We set the following rule to remove low energy points to get better performance

$$\frac{\|\mathrm{Re}(\mathbf{X}(t,f))\|}{\max(\|\mathrm{Re}(\mathbf{X}(t,f))\|)} < \varepsilon_3 \tag{25}$$

where ε_3 is a number close to 1.

We cluster these selected points and get corresponding clustering centers through utilizing clustering algorithm. The number of the selected points is K and they are denoted as $(Y_k, Z_k) = (k = 1, 2, \ldots, K)$. Now we define the potential function as follows

$$J(\mathbf{b}_k) = \sum_{i=1}^{T} \{\exp[\beta \cos(\theta_{\mathbf{b}_k \mathbf{b}_i})]\}^{\gamma} \tag{26}$$

where $\mathbf{b}_k$ and $\mathbf{b}_i$ are single source points, and they are parameters that adjust the degree of attenuation of this function at non-extreme points. The potential function values at different points can be calculated by the above formula, and then a three-dimensional diagram about $\mathbf{b}_{k1}$, $\mathbf{b}_{k2}$ and $J(\mathbf{b}_k)$ is obtained. In this three-dimensional diagram, there are some significant peaks appearing, and the number of peaks is equal to the number of source signals. Assume that the amplitude of each point in the three-dimensional diagram is $P(k)(k = 1, 2, \ldots, K)$. In order to eliminate the interference term, we set the following smoothing function to

$$\hat{P}(k) = P(\mathrm{k})/\max(P(\mathrm{k})) \tag{27}$$

$$\begin{aligned} p_k = \big[&\hat{P}(k-h) + 2\hat{P}(k-h+1) + \ldots + 2^{h-1}\hat{P}(k-1) + 2^h\hat{P}(k) \\ &+ 2^{h-1}\hat{P}(k+1) + \ldots + 2\hat{P}(k+h-1) + \hat{P}(k+h)\big]/\big(3 \cdot 2^h - 2\big) \end{aligned} \tag{28}$$

where h is an integer that is >1, and p_k is the new peak amplitude. We set the following rule to get the correct peak position.

$$\begin{cases} p_{k-1} < p_k \text{ and } p_{k+1} < p_k \\ p_{k-2} < p_k \text{ and } p_{k+2} < p_k \end{cases} \tag{29}$$

Through this method, the subinterval position $\mathbf{b}_k$ corresponding to the peak and the initial clustering centers $(A_m, B_m) = (m = 1, 2, \ldots, M)$ can be obtained. Single source points close to the initial cluster center can be re-selected through following rules

$$\frac{A_m Y_k + B_m Z_k}{\sqrt{A_m^2 + B_m^2}\sqrt{Y_k^2 + Z_k^2}} > \varepsilon_4 \tag{30}$$

where ε_4 is a threshold between 0 and 1. The mixed matrix can be estimated through these re-selected single source points.

4 Simulation Results and Analysis

We consider four speech signals in [10] to test the practicality of the proposed algorithm. The sampling number is 160000, STFT size is 1024, Overlapping is 512, Weighting function is Hanning Window. $\varepsilon_1 = 0.999$, $\varepsilon_3 = 0.02$, $\varepsilon_4 = 0.997$. The mixed matrix **A** is defined as

$$\mathrm{A} = \begin{bmatrix} 0.763 & 0.658 & 0.328 & 0.442 \\ 0.313 & 0.360 & 0.766 & 0.540 \end{bmatrix}$$

We consider the scatter plot of two time-domain observation signals under noiseless conditions. We reduce the number of points to reduce the amount of calculation. First, the descending order of the real parts at different frequency points after the time-frequency conversion of the first observation signal is performed in descending order. According to the order of the first observation signals, the order of the second observation signals is adjusted, and the time-frequency observation signals at the frequency points with large variances are selected. In this chapter, the corresponding observation signals at the first 50 points are selected before the single source point is detected. The scatter plot of two observation signals before detecting is shown in Fig. 1.

Figure 1 present obvious linear clustering characteristics, but some stray points affect this property. The existence of spurious points makes direct clustering will produce large estimation errors. At the same time, it can be found that a large number of scatter points are accumulated near the origin. However, the amplitudes of these scatter points are small, and the directions of the straight lines in the scatter plot are far less effective than the scatters far from the origin. We eliminate these points for better performance. The scatter plot of the two observation signals in the TF domain after detecting the single source point and removing the low energy point is shown in Fig. 2.

Figure 2 shows that the linear clustering characteristics of the two observation signals are more obvious. Finally, the mixing matrix is estimated.

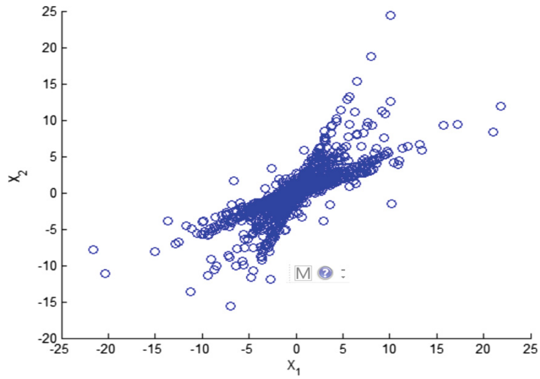

Fig. 1. The scatter plot of two observed signals in TF domain.

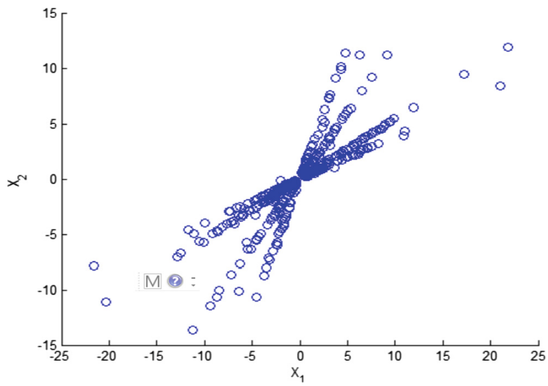

Fig. 2. The scatter plot of the two observation signals in the TF domain after detecting the single source point and removing the low energy point.

$$\hat{\mathbf{A}} = \begin{bmatrix} 0.7628 & 0.6568 & 0.3272 & 0.4407 \\ 0.3087 & 0.3611 & 0.7649 & 0.5405 \end{bmatrix}$$

We take the normalized mean square error (NMSE) to measure the performance of algorithms. It can be written as

$$\mathrm{NMSE} = 10 \log \left[\frac{\sum_{i,j} (\tilde{a}_{i,j} - a_{i,j})^2}{\sum_{i,j} a_{i,j}^2} \right] (\mathrm{dB}) \tag{31}$$

Where a_{ij} is the (i,j)th element of $\mathbf{A}$ and $\hat{a}_{ij}$ is the (i,j)th element of $\hat{\mathbf{A}}$. This parameter gets a lower value when the estimated mixed matrix is more similar to the real mixed matrix.

The result of the different algorithms are shown in Table 1.

Table 1. The NMSE comparison of different algorithms

The TIFROM algorithm	Dong's algorithm	Reju's algorithm	Our algorithm
−45.7654	−50.4102	−48.9157	−55.2361

From Table 1, we can find that our algorithm has lower NMSE, which means a better performance.

5 Conclusion

A novel algorithm is proposed to solve the problem of mixed matrix estimation in UBSS under linear instantaneous mixed model. First, a new method is proposed for detecting single source points. Then, the algorithm clusters them by utilizing a method based on the improved potential function. Finally, the mixing matrix is obtained. The detection algorithm is feasible and efficient, which lays the foundation for post-processing.

Acknowledgment. This work is supported by the National Natural Science Foundation of China (61771154) and the Fundamental Research Funds for the Central Universities (HEUCFG201830).

This paper is also funded by the International Exchange Program of Harbin Engineering University for Innovation-oriented Talents Cultivation.

Meantime, all the authors declare that there is no conflict of interests regarding the publication of this article.

We gratefully thank of very useful discussions of reviewers.

References

1. Abrard, F., Deville, Y.: A time-frequency blind signal separation method applicable to underdetermined mixtures of dependent sources. Sig. Process. **85**(7), 1389–1403 (2005)
2. Aissaelbey, A., et al.: Underdetermined blind separation of nondisjoint sources in the time-frequency domain. IEEE Trans. Sig. Process. **55**(3), 897–907 (2007)
3. Arberet, S., Gribonval, R., Bimbot, F.: A robust method to count and locate audio sources in a stereophonic linear instantaneous mixture. In: Rosca, J., Erdogmus, D., Príncipe, José C., Haykin, S. (eds.) ICA 2006. LNCS, vol. 3889, pp. 536–543. Springer, Heidelberg (2006). https://doi.org/10.1007/11679363_67
4. Dong, T., Lei, Y., Yang, J.: An algorithm for underdetermined mixing matrix estimation. Neurocomputing **104**, 26 (2013)

5. Kim, S.G., Chang, D.Y.: Blind separation of speech and sub-Gaussian signals in underdetermined case. In: INTERSPEECH 2004 - ICSLP, International Conference on Spoken Language Processing, Jeju Island, Korea, October 2004. DBLP (2004)
6. Xie, S., Yang, L., Yang, J.M., et al.: Time-frequency approach to underdetermined blind source separation. IEEE Trans. Neural Netw. Learn. Syst. **23**(2), 306–316 (2012)
7. Pedersen, M.S., Wang, D.L., Larsen, J., et al.: Two-microphone separation of speech mixtures. IEEE Trans. Neural Netw. **19**(3), 475–492 (2008)
8. Poncelet, F., Kerschen, G., Golinval, J.C., et al.: Output-only modal analysis using blind source separation techniques. Mech. Syst. Sig. Process. **21**(6), 2335–2358 (2007)
9. Zhou, J.T., Zhao, H., Peng, X., Fang, M., Qin, Z., Goh, R.S.M.: Transfer hashing: from shallow to deep. IEEE Trans. Neural Netw. Learn. Syst. https://doi.org/10.1109/tnnls.2018.2827036
10. Reju, V.G., Koh, S.N., Soon, I.Y.: An algorithm for mixing matrix estimation in instantaneous blind source separation. Sig. Process. **89**(9), 1762–1773 (2009)
11. Ming, X., Xie, S.L., Yu-Li, F.U.: Underdetermined blind delayed source separation based on single source intervals in frequency domain. Acta Electronica Sinica **35**(12), 2279–2283 (2007)
12. Yu, K., Yang, K., Bai, Y.: Estimation of modal parameters using the sparse component analysis based underdetermined blind source separation. Mech. Syst. Signal Process. **45**(2), 302–316 (2014)
13. Tu, Y., Lin, Y., Wang, J., et al.: Semi-supervised learning with generative adversarial networks on digital signal modulation classification. CMC Comput. Mater. Continua **55**(2), 243–254 (2018)
14. Li, Y., et al.: Underdetermined blind source separation based on sparse representation. IEEE Trans. Signal Process. **54**(2), 423–437 (2006)
15. Zheng, Z., Kumar Sangaiah, A., Wang, T.: Adaptive communication protocols in flying ad-hoc network. IEEE Commun. Mag. **56**(1), 136–142 (2018)
16. Zhao, N., Richard Yu, F., Sun, H., Li, M.: Adaptive power allocation schemes for spectrum sharing in interference-alignment-based cognitive radio networks. IEEE Trans. Veh. Technol. **65**(5), 3700–3714 (2016)