DOA ESTIMATION OF SPARSELY SAMPLED NONSTATIONARY SIGNALS

Liang Guo [†‡], *Yimin D. Zhang* [‡], *Qisong Wu* [‡], *and Moeness G. Amin* [‡]

[†] School of Physics and Optoelectronic Engineering, Xidian University, Xi'an 710071, China
[‡] Center for Advanced Communications, Villanova University, Villanova, PA 19085, USA

ABSTRACT

The paper deals with sparsely sampled nonstationary signals in a multi-sensor array platform. We examine direction-of-arrival (DOA) estimation using sparsity-based time-frequency signal representation (TFSR). While conventional time-frequency analysis techniques suffer from noise-like artifacts due to missing data samples, high-fidelity time-frequency signatures can be obtained by applying kernelled processing and sparse reconstruction. Since the signals received at different sensors occupy the same time-frequency regions and share a common nonzero support, the recovery of TFSRs can be cast as a group sparse reconstruction problem. The reconstructed auto- and cross-sensor TFSRs enable the formation of the spatial time-frequency distribution (STFD) matrix, which is used, in turn, to propose the sparse time-frequency MUSIC (STF-MUSIC). The proposed STF-MUSIC method achieves effective source discrimination capability, leading to improved DOA estimation performance.

Index Terms— Time-frequency analysis, DOA estimation, compressive sensing, sparse sampling

1. INTRODUCTION

A large number of nonstationary signals encountered in radar, sonar, communications, and biomedical applications are frequency modulated (FM) and can be characterized by their instantaneous frequencies (IFs) [1–3]. Time-frequency signal representations (TFSRs) of such signals enable exploitation of signal sparsity and local signal behaviors. The Cohen's class defines a general bilinear TFSR framework [4], within which reduced interference distributions employ time-frequency (TF) kernels to mitigate the effect of undesirable cross-terms [5–7].

For multi-sensor array applications, incorporation of the signal TF signatures into parameter estimation and signal recovery was accomplished within the framework of spatial time-frequency distributions (STFDs). The STFD has effectively been used to serve both problems of direction-of-arrival (DOA) estimation [8–13] and blind source separation

[14–17] for nonstationary signals. Traditional TF analysis and STFD applications assume that the nonstationary signals are uniformly sampled with a rate satisfying the Nyquist criterion. In many real-world applications, however, nonstationary signals are often observed with missing samples due to fading, obstruction, and/or impulsive noise. The emerging compressive sensing (CS) techniques also allow intentional random undersampling for reduced system complexity [18]. Missing samples generally yield noise-like artifacts that spread over the entire TF domain [19]. It is shown that effective artifact mitigation can be achieved by applying TF smoothing kernels [19, 20]. In multi-sensor systems, artifact mitigation can also be accomplished by averaging the TFSRs or ambiguity functions (AFs) corresponding to the different available sensors [21–23].

In this paper, inspired by the TF-MUSIC developed in the STFD context [8, 9, 11–13] and the sparsity-based TFSR reconstruction as recently introduced in [22–25], we examine the sparsity-based STFD reconstruction of sparsely sampled nonstationary signals in a multi-sensor platform, and propose the sparse time-frequency MUSIC (STF-MUSIC) for effective DOA estimation. We exploit the fact that the auto- and cross-sensor TFSRs share the same nonzero TF support to cast their sparse reconstruction as a block-sparse or group-sparse compressive sensing problem [26–30]. To mitigate the effect of the artifacts due to missing samples in TF kernel optimizations, the well-known data-dependent adaptive optimal kernel (AOK) [6] is modified such that the kernel operates on multi-sensor observations and applied to the AF averaged over all available sensors [23].

Once the kernel is designed, the corresponding TFSRs can be obtained using two-dimensional (2-D) Fourier transform of the kernelled AF. CS techniques may, however, replace the Fourier transform with sparse parametric estimation and achieve enhanced resolution in the TF domain [24]. By utilizing the one-dimensional (1-D) Fourier transform relationship between the instantaneous auto-correlation function (IAF) and the TF distribution (TFD), in lieu of the 2-D Fourier transform relationship between the AF and TF domains, CS reconstruction not only reduces the computation complexity, but also improves the performance as a result of the presence of local sparsity over each time instant. Considering that the TFSRs across different array sensors share the same nonzero

support, a group sparse compressive sensing algorithm is adopted. The reconstructed auto- and cross-sensor TFSRs enable formation of the STFD matrix, which is used to propose the STF-MUSIC algorithm for effective DOA estimation of sparsely sampled nonstationary signals. The proposed STF-MUSIC method achieves effective source discrimination capability, leading to improved DOA estimation performance.

Notations. We use lower-case (upper-case) bold characters to denote vectors (matrices). $\mathcal{F}_x$ and $\mathcal{F}_x^{-1}$ respectively represent the discrete Fourier Transform (DFT) and inverse DFT (IDFT) with respect to x. In particular, $\mathbf{I}_N$ denotes the $N \times N$ identity matrix. $(\cdot)^*$ denotes complex conjugate of a complex value, and $(\cdot)^T$ and $(\cdot)^H$ respectively denote the transpose and conjugate transpose of a matrix or vector. In addition, $\mathbb{C}^{N \times K}$ denotes the complete set of $N \times K$ complex entries.

2. PROBLEM FORMULATION

2.1. Signal Model

In narrowband array processing, consider K nonstationary signals impinging on an array consisting of N sensors. The $N \times 1$ received data vector $\mathbf{x}(t)$ and the $K \times 1$ source signal vector $\mathbf{d}(t)$ are related by

$$\mathbf{x}(t) = \mathbf{y}(t) + \mathbf{n}(t) = \mathbf{H}(\boldsymbol{\theta})\mathbf{d}(t) + \mathbf{n}(t), \quad (1)$$

where $t \in [1, \cdots, T]$ is the time index of a total number of T data samples, $\mathbf{H}(\boldsymbol{\theta}) = [\mathbf{h}(\theta_1), \mathbf{h}(\theta_2), \cdots, \mathbf{h}(\theta_K)] \in \mathbb{C}^{N \times K}$ is the steering matrix with $\mathbf{h}(\theta_q)$ denoting the steering vector corresponding to the DOA θ_q of the q-th source. $\mathbf{n}(t) \in \mathbb{C}^{N \times 1}$ is an additive noise vector which consists of independent and identically distributed (i.i.d.) zero-mean, white and complex Gaussian distributed processes with variance $\sigma_n^2 \mathbf{I}_N$. For simplicity but without loss of generality, we assume σ_n to be the same for all sensors, i.e., $\sigma_n = \sigma$, for $n = 1, \cdots, N$. The noise elements are assumed to be independent of the signals, which are assumed to be deterministic. Here, each element of vector $\mathbf{d}(t) = [d_1(t), d_2(t), \cdots, d_K(t)]^T$ is a single component FM signal. Consider the thinned sampling of the array observations with a random pattern applied to each array sensor, where the number of missing samples is M_q satisfying $M_q < T$ for $q = 1, \cdots, N$. As such, for the q-th array sensor, the thinned observation, $x_q(t)$, can be expressed as the product of $y_q(t)$, expressed in (1), and the following observation mask,

$$b_q(t) = \begin{cases} 1, & \text{if } t \in \mathbb{S}_q, \\ 0, & \text{if } t \notin \mathbb{S}_q, \end{cases} \quad (2)$$

where $\mathbb{S}_q \subset \{1, \cdots, T\}$ is the set of observed time instants and its cardinality is $|\mathbb{S}_q| = T - M_q$. Note that $b_q(t)$ may or may not be the same for each sensor.

2.2. The MUSIC Algorithm

Conventional MUSIC algorithm uses the correlation matrix of $\mathbf{x}(t)$, defined as

$$\mathbf{R_{xx}} = E[\mathbf{x}(t)\mathbf{x}^H(t)]. \quad (3)$$

In practice, it is commonly estimated from the measured snapshots as

$$\hat{\mathbf{R}}_{\mathbf{xx}} = \frac{1}{N} \sum_{t=1}^{N} \mathbf{x}(t)\mathbf{x}^H(t). \quad (4)$$

Let $\hat{\mathbf{G}}$ denote the noise subspace of matrix $\hat{\mathbf{R}}_{\mathbf{xx}}$. The MUSIC technique estimates the DOAs by determining the values of ϑ for which the following spatial spectrum is maximized [31]:

$$P(\vartheta) = \frac{1}{\mathbf{h^H}(\vartheta)\hat{\mathbf{G}}\hat{\mathbf{G}}^{\mathbf{H}}\mathbf{h}(\vartheta)}, \quad (5)$$

where $\mathbf{h}(\vartheta)$ is the steering vector corresponding to DOA ϑ.

It is noted that missing data in each antenna will result in a missing entry in the estimation of $\hat{\mathbf{R}}_{\mathbf{xx}}$.

3. CS-BASED SPARSE TIME-FREQUENCY MUSIC

3.1. Spatial Time-Frequency Distributions

The Cohen's class of bilinear TFD for signals $x_1(t)$ and $x_2(t)$ is given by [4]

$$D_{x_1 x_2}(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A_{x_1 x_2}(\theta, \tau)\phi(\theta, \tau)e^{-j2\pi\theta t - j2\pi f\tau}d\theta d\tau, \quad (6)$$

where $A_{x_1 x_2}(\theta, \tau)$ is referred to as the ambiguity function, which is given by

$$A_{x_1 x_2}(\theta, \tau) = \int_{-\infty}^{\infty} x_1(t + \frac{\tau}{2})x_2^*(t - \frac{\tau}{2})e^{-j2\pi\theta t}dt. \quad (7)$$

In the above expressions, f, θ and τ, respectively, denote the frequency index, the frequency shift, and the time lag. The kernel $\phi(\theta, \tau)$ characterizes the distribution and will be discussed later. $D_{x_1 x_2}(t, f)$ is referred to as auto-terms when $x_1(t) = x_2(t)$, and as cross-terms when $x_1(t) \neq x_2(t)$. Expressions (6) can be used to defined the data STFD matrix, $\mathbf{D}(t, f)$, whose (i, k)th element is $D_{x_i x_k}(t, f)$.

For the linear data model of (1), the STFD matrix can be decomposed as [11, 14]

$$\mathbf{D_{xx}}(t, f) = \mathbf{D_{yy}}(t, f) + \mathbf{D_{yn}}(t, f) + \mathbf{D_{ny}}(t, f) + \mathbf{D_{nn}}(t, f). \quad (8)$$

Under the uncorrelated signal and noise assumption and the zero-mean noise property, the expectation of the cross-term

STFD matrices between the signal and noise vectors is zero, i.e., $E[\mathbf{D_{yn}}(t,f)] = E[\mathbf{D_{ny}}(t,f)] = 0$, and it follows that

$$E[\mathbf{D_{xx}}(t,f)] = \mathbf{D_{yy}}(t,f) + E[\mathbf{D_{nn}}(t,f)]$$
$$= \mathbf{H}\mathbf{D_{dd}}(t,f)\mathbf{H}^H + E[\mathbf{D_{nn}}(t,f)], \quad (9)$$

where $\mathbf{D_{dd}}(t,f)$ is the source TFD matrix whose entries are the auto- and cross-source TFDs of the source. For narrowband array signal processing applications, matrix $\mathbf{H}$ holds the spatial information and maps the auto- and cross-source TFDs into auto- and cross-sensor TFDs [14]. Expression (9) is similar to (4) which used in DOA estimation problems relating the signal correlation matrix to the data spatial correlation matrix.

3.2. Multi-sensor Adaptive Optimal Kernel

Missing measurements produce artifacts in the TFD and AF domains which resemble additive noise in the sense that they spread over the entire respective domains of joint-variable representations [19]. These artifacts can be mitigated by applying proper TF kernels. In practice, to simplify the selection of auto-terms and suppress the cross-terms and artifacts, we apply a smoothing kernel that significantly decreases the contribution of the cross-terms in the TF plane. As discussed in [19, 23], the well-known data-dependent adaptive optimal kernel (AOK) is uesd for this purpose. AOK is obtained by solving the following optimization problem defined in the polar coordinates [6]:

$$\max_{\Phi} \int_0^{2\pi} \int_0^\infty |A(r,\psi)\Phi(r,\psi)|^2 r\, dr\, d\psi$$

$$\text{subject to} \quad \Phi(r,\psi) = \exp\left(-\frac{r^2}{2\sigma(\psi)}\right), \quad (10)$$

$$\frac{1}{4\pi^2}\int_0^{2\pi}\sigma(\psi)d\psi \le \alpha,$$

where $\alpha \ge 0$. In this expression, $A(r,\psi)$ is the AF, and $\Phi(r,\psi)$ is the kernel function, both defined in the polar coordinates.

In the underlying multi-sensor system, as discussed in [23], averaging the AF obtained from each array sensor, $A_q(r,\psi), q = 1, ..., N$, can enhance the auto-terms and reduce cross-terms. Meanwhile, the missing data samples yield artifacts that randomly spread over the entire AF domain, and the overall variance increases with more missing data samples. As different sampling patterns in each sensor generate distinct artifact distributions, averaging the AFs over all sensors effectively reduces the presence of such artifacts induced from missing samples.

Let $A_q(r,\psi)$ denote the auto-sensor AF in the qth sensor, $q = 1, ..., N$. The averaged AF over all sensors is given by

$$A_\Sigma(r,\psi) = \frac{1}{N}\sum_{q=1}^N A_q(r,\psi). \quad (11)$$

Then, an improved kernel in the multi-sensor platform is obtained by replacing $A(r,\psi)$ in (10) by $A_\Sigma(r,\psi)$ in (11), yielding better auto-term enhancement and artifact reduction.

3.3. Compressive Sensing Model

The auto- and cross-sensor TFDs can be computed as the 2-D Fourier transform of the corresponding kernelled auto- and cross-sensor AFs, i.e., $A_{x_i x_k}(r,\psi)\Phi(r,\psi)$, for $i,k = 1,\cdots,N$. By converting the kernelled and averaged AFs to the rectangular coordinate system, and denoting the result as $\tilde{A}_{x_i x_k}(\theta,\tau)$, the corresponding TFDs is expressed as

$$D_{x_i x_k}(t,f) = \mathcal{F}_\theta^{-1}\{\mathcal{F}_\tau[\tilde{A}_{x_i x_k}(\theta,\tau)]\}. \quad (12)$$

Alternatively, we can obtain the TFDs through a group sparse reconstruction from the same kernelled and averaged AFs. While earlier sparse TFD reconstructions were based on the 2-D Fourier transform relationship between the AF and the TFD [22, 24], it is shown in [19, 25] that the 1-D Fourier transform relationship between the IAF and the TFD yields simpler computations and enables the exploitation of local sparsity in the TF domain.

For each pair of AF and IAF, the 1-D inverse Fourier transform of $A_{xx}(\theta,\tau)$ with respect to θ yielding the kernelled IAF,

$$C_{x_i x_k}(t,\tau) = \mathcal{F}_\theta^{-1}[\tilde{A}_{x_i x_k}(\theta,\tau)]. \quad (13)$$

For each pair of IAF and TFSR, we denote $\mathbf{c}_{x_i x_k}^{[t]}$ as a vector that consists of all IAF entries between $x_i(t)$ and $x_k(t)$ along the τ dimension corresponding to time t, and $\mathbf{w}_{x_i x_k}^{[t]}$ as a vector contains all the corresponding TFD entries with respect to the frequency for the same time t. According to the Fourier relationship between the IAF and the TFD, we obtain

$$\mathbf{c}_{x_i x_k}^{[t]} = \mathbf{\Phi}\mathbf{w}_{x_i x_k}^{[t]} + \boldsymbol{\epsilon}_{x_i x_k}^{[t]}, \quad (14)$$

for $1 \le t \le T$ and $i,k = 1,\cdots,N$. Note in the above expression that all the $\mathbf{w}_{x_i x_k}^{[t]}$ entries have the sparsity support for different sensor pairs, i.e., the respective positions of the nonzero entries of $\mathbf{w}_{x_i x_k}^{[t]}$ are the same for different pairs of i and j. In addition, since the same $\mathbf{\Phi}$ is shared for all sensor pairs, the above problem is referred to as a multiple measurement vector (MMV) model. Such problem can be solved using block OMP (BOMP) [26], group Lasso (GLasso) [27], or multi-task BCS [29, 30]. In this paper, the BOMP algorithm is used as it is simple and allows specification of the sparsity, but other techniques may also be used.

3.4. Sparse Time-Frequency MUSIC

To perform the STF-MUSIC, we form an STFD matrix corresponding to a TF region $\mathcal{S}$, expressed as

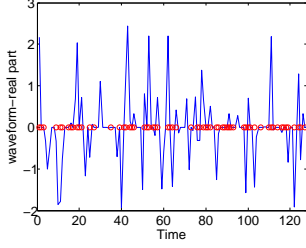$$\hat{\mathbf{D}}_{\mathbf{xx}}^{(\mathcal{S})} = \sum_{(t,f)\in\mathcal{S}} \mathbf{D_{xx}}(t,f). \quad (15)$$

Fig. 1. Waveform with missing samples marked in red circles.

(a) PWVD

(b) TFD from CS reconstruction

Fig. 2. PWVD and CS-based TFD of the sparsely samples signals.

Depending on the separability of the TFSRs of the sources, the TF region $\mathcal{S}$ may include only a single source, a subset of sources, or all the sources. Generally, the selection of fewer sources in the TFSRs yields better DOA estimation performance [11].

Let $\hat{\mathbf{G}}^{(\mathcal{S})}$ denote the noise subspace of matrix $\hat{\mathbf{D}}_{\mathbf{xx}}^{(\mathcal{S})}$. The proposed STF-MUSIC technique estimates the DOAs of K_0 (where $1 \leq K_0 \leq K$) selected sources by determining the values of ϑ for which the following spatial spectrum is maximized,

$$P^{(\mathcal{S})}(\vartheta) = \frac{1}{\mathbf{h}^H(\vartheta)\hat{\mathbf{G}}^{(\mathcal{S})}(\hat{\mathbf{G}}^{(\mathcal{S})})^H\mathbf{h}(\vartheta)}. \quad (16)$$

When $K_0 < K$, the above procedure should be repeated for other TF regions until the DOA of all sources are estimated. It is also noted that, with the source selection capability, the total number of sources, K, may exceed the total number of sensors, N, as long as $K_0 \leq N - 1$ is satisfied in each TF region in which the STF-MUSIC algorithm is performed.

4. SIMULATION RESULTS

In this section, we consider a uniform linear array (ULA) of $N = 4$ sensors with a half-wavelength interelement spacing. Two nonstationary signals ($K = 2$) impinge on the array with their respective arbitrary spatial signatures. The signals emitted from the two sources are polynomial-phase FM signals with closely separated signatures. Their IF laws are expressed as

$$\begin{aligned} f_1(t) &= 0.05 + 0.1t/T + 0.1t^2/T^2, \\ f_2(t) &= 0.15 + 0.1t/T + 0.1t^2/T^2, \end{aligned} \quad (17)$$

where $t = [1, \cdots, T]$ and $T = 128$. The two sources imping from respective angles of $\theta_1 = -3°$ and $\theta_2 = 3°$. The input signal-to-noise ratio (SNR) is 0dB for both signals, and 50% of the data samples are randomly missing in each array sensor with different missing patterns. The real part of the waveform is shown in Fig. 4.

Fig. 2(a) shows the pseudo Wigner-Ville distribution (PWVD), as used in [11], of the signal received at the first array sensor. It is clear that, due to the 50% missing data samples, the TFDs are highly cluttered by the artifacts, and
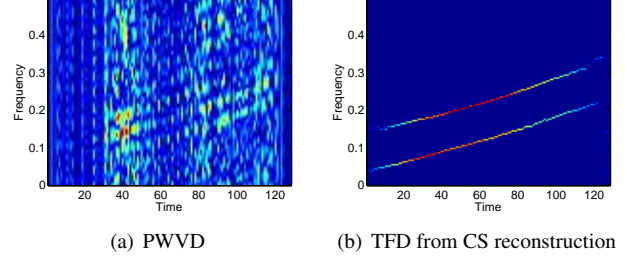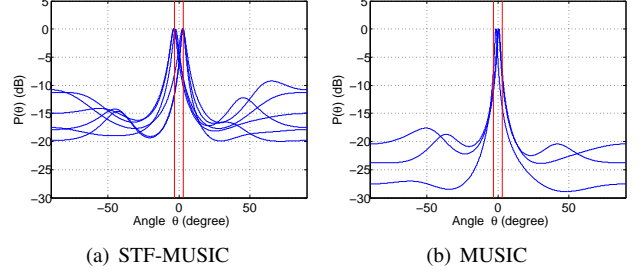

(a) STF-MUSIC

(b) MUSIC

Fig. 3. DOA estimation results.

strong cross-terms are present between the auto-terms. As such, it is difficult to perform instantaneous frequency estimation and source selection from this result. In comparison, Fig. 2(b) shows the TFD obtained from the CS technique for the first sensor, where the AOK function is obtained based on array averaged AF. The TFD shown in Fig. 2(b) clearly reveal the true instantaneous frequencies of the two impinging sources for individual source selection. As a result, we can separately perform the proposed STF-MUSIC on each of the two sources.

The estimated STF-MUSIC pseudo spectra, overlaid for separated obtained results of both sources, are shown in Fig. 3(a) for three trials for each source, where the estimated spectrum peaks respectively point to each DOA. On the other hand, conventional MUSIC algorithm, which does not have the capability for source selection, fails to separate the two sources in its estimated pseudo spectra, as shown in Fig. 3(b).

5. CONCLUSION

Compressive sensing techniques, combined with proper data-dependent kernel design, enable effective time-frequency signature reconstruction from randomly sampled nonstationary signals due to the sparsity of such signals in their time-frequency signal representations (TFSRs). The group sparsity of the TFSRs of the signals observed in different sensors allows auto- and cross-sensor time-frequency distribution reconstruction, which leads to the formation of the spatial time-frequency distribution matrix. The letter was used to define the sparse time-frequency MUSIC (STF-MUSIC) algorithm which demonstrated effective source selection and direction-of-arrival (DOA) estimation capabilities.

6. REFERENCES

[1] M. Skolnik, *Radar Systems*. McGraw-Hill, 2001.

[2] V. C. Chen and H. Ling, *Time-Frequency Transforms for Radar Imaging and Signal Analysis*. Artech House, 2001.

[3] B. Boashash, *Time Frequency Signal Analysis and Processing*. Elsevier, 2003.

[4] L. Cohen, *Time-Frequency Analysis: Theory and Applications*. Prentic-Hall, 1995.

[5] H. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 862–871, 1989.

[6] D. L. Jones and R. G. Baraniuk, "An adaptive optimal-kernel time-frequency representation," *IEEE Trans. Signal Process.*, vol. 43, no. 10, pp. 2361–2371, 1995.

[7] M. G. Amin and W. J. Williams, "High spectral resolution time-frequency distribution kernels," *IEEE Trans. Signal Process.*, vol. 46, no. 10, pp. 2796–2804, 1998.

[8] A. Belouchrani and M. G. Amin, "Time-frequency MUSIC," *IEEE Signal Process. Lett.*, vol. 6, no. 5, pp. 109–110, 1999.

[9] M. G. Amin and Y. D. Zhang, "Direction finding based on spatial time-frequency distribution matrices," *Digital Signal Process.*, vol. 10, no. 4, pp. 325–359, 2000.

[10] Y. D. Zhang, W. Mu, and M. G. Amin, "Time-frequency maximum likelihood methods for direction finding," *J. Franklin Inst.*, vol. 337, no. 4, pp. 483–497, 2000.

[11] Y. D. Zhang, W. Mu, and M. G. Amin, "Subspace analysis of spatial time-frequency distribution matrices," *IEEE Trans. Signal Process.*, vol. 49, no. 4, pp. 747–759, 2001.

[12] M. G. Amin, Y. D. Zhang, G. J. Frazer, and A. R. Lindsey, "Spatial time-frequency distributions: Theory and applications," in L. Debnath (ed.), *Wavelets and Signal Processing*, Birkhauser 2003.

[13] Y. D. Zhang, M. G. Amin, and B. Himed, "Joint DOD/DOA estimation in MIMO radar exploiting time-frequency signal representations," *EURASIP J. Adv. Signal Process.*, vol. 2012, no. 1, pp. 102–112, 2012.

[14] A. Belouchrani and M. G. Amin, "Blind source separation based on time-frequency signal representation," *IEEE Trans. Signal Process.*, vol. 46, no. 11, pp. 2888–2898, 1998.

[15] Y. D. Zhang and M. G. Amin, "Spatial averaging of time-frequency distributions for signal recovery in uniform linear arrays," *IEEE Trans. Signal Process.*, vol. 48, no. 10, pp. 2892–2902, 2000.

[16] L. Cirillo, A. M. Zoubir, N. Ma, and M. Amin, "Automatic classification of auto- and cross-terms of time-frequency distributions in antenna arrays," in *Proc. IEEE ICASSP*, Orlando, FL, May 2002.

[17] Y. D. Zhang and M. G. Amin, "Blind separation of nonsta-

tionary sources based on spatial time-frequency distributions," *EURASIP J. Applied Signal Process.*, vol. 2006, article ID 64785, 13 pages, 2006.

[18] L. Stankovic, S. Stankovic, I. Orovic, and Y. D. Zhang, "Time-frequency analysis of micro-Doppler signals based on compressive sensing," in M. G. Amin (ed.), *Compressive Sensing for Urban Radars,* CRC Press, 2014.

[19] Y. D. Zhang, M. G. Amin, and B. Himed, "Reduced interference time-frequency representations and sparse reconstruction of undersampled data," in *Proc. European Signal Process. Conf.,* Marrakech, Morocco, Sept. 2013.

[20] B. Jokanovic, M. G. Amin, Y. D. Zhang, and F. Ahmad, "Adaptive time-frequency kernel design for sparse joint-variable signal representations," in *Proc. European Signal Process. Conf.,* Lisbon, Portuga, Sept. 2014.

[21] W. Mu, M. G. Amin, and Y. D. Zhang, "Bilinear signal synthesis in array processing," *IEEE Trans. Signal Process.*, vol. 51, no. 4, pp. 90–100, 2003.

[22] Y. D. Zhang and M. G. Amin, "Compressive sensing in nonstationary array processing using bilinear transforms," in *Proc. IEEE Sensor Array and Multichannel Signal Process. Workshop,* Hoboken, NJ, Jun. 2012.

[23] Y. D. Zhang, L. Guo, Q. Wu, and M. G. Amin, "Multi-sensor kernel design for time-frequency analysis of sparsely sampled non-stationary signals," in *Proc. IEEE Radar Conf.,* Arlington, VA, May 2015.

[24] P. Flandrin and P. Borgnat, "Time-frequency energy distributions meet compressed sensing," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 2974–2982, 2010.

[25] Q. Wu, Y. D. Zhang, and M. G. Amin, "Continuous structure based bayesian compressive sensing for sparse reconstruction of time-frequency distribution," in *Proc. Int. Conf. Digital Signal Process.,* Hong Kong, China, Aug. 2014.

[26] L. Jacob, G. Obozinski, and J. Vert, "Group Lasso with overlap and graph Lasso," in *Proc. Int. Conf. Machine Learning*, Montreal, Canada, June 2009.

[27] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *Royal Statistical Society Series B*, vol. 68, no. 1, pp. 49–67, 2006.

[28] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signal: uncertainty relations and efficient recovery," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3042–3054, 2010.

[29] S. Ji, D. Dunson, and L. Carin, "Multitask compressive sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 92–106, 2009.

[30] Q. Wu, Y. D. Zhang, M. G. Amin, and B. Himed, "Complex multitask Bayesian compressive sensing," in *Proc. IEEE ICASSP,* Florence, Italy, May 2014.

[31] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagat.*, vol. AP-34, pp. 276–280, 1986.