High-Resolution Radar Imaging with Unknown Noise

Xueru Bai
National Lab of Radar Signal Processing
Xidian University
Xi'an, China
xrbai@xidian.edu.cn

Yu Zhang
National Lab of Radar Signal Processing
Xidian University
Xi'an, China

Feng Zhou
The Ministry Key Laboratory of Electronic Information Countermeasure and Simulation
Xidian University
Xi'an, China

***Abstract*—This paper addresses the problem of high-resolution radar imaging in complex environments with unknown noise in a Bayesian framework. In the new statistical model, the noise obeys Gaussian mixture distribution, and the weights are governed by the sparsity-promoting Gamma-Gaussian hierarchical prior. Then, the weights are estimated via the maximum a posterior-expectation maximization (MAP-EM) technique. Experiments have shown that the proposed method provides an effective way of high-resolution radar imaging in complex environments such as barrage jamming.**

Keywords—Radar imaging; Bayesian learning; GMM; MAP-EM; jamming suppression

I. INTRODUCTION

High-resolution radar imaging plays significant roles in air surveillance, reconnaissance, and space situation awareness [1]-[2]. For ideal observation environments, i.e. high signal-to-noise (SNR) ratio, jamming-free, and steady rigid-body motion, etc., well-focused imaging can be achieved by mature Fourier techniques, such as the range-Doppler algorithm (RDA) and the polar formatting algorithm (PFA) [3]. Sometimes, however, imaging radar may encounter non-cooperative or even unfriendly environments with unknown noise. Whether intentional or not, they may offer numerous characteristics and impeding techniques to diminish the effectiveness of available imaging methods [4].

Generally, barrage jamming, strong clutter, and the inconsistency of noise statistics, etc. constitute the complex imaging environments. To settle the issue of barrage jamming or strong clutter, one may resort to hardware upgrading, which includes the development of new antennas (which have better cross-polarization rejection), high-power transmitters, and robust waveform generators, or includes constructions of cognitive imaging radars [5]. On the other hand, such strategies suffer from tremendous cost and design limitations, and become invalid in the case of noise inconsistency among radar channels. To obtain well-focused imaging without incurring significant hardware cost, this paper exploits statistical characteristics of the complex environments mentioned above, and research into robust high-resolution imaging algorithms based on Bayesian learning.

A. Prior Work

Considering the sparse nature of scattering centers, high-resolution radar imaging can be cast as a sparse signal reconstruction problem, i.e.

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

where $\mathbf{y} \in \mathbb{C}^N$ is the vector of target echoes; $\mathbf{\Phi} \in \mathbb{C}^{N \times D}$ is the overcomplete dictionary with $N \ll D$; the weights $\mathbf{x} \in \mathbb{C}^D$ represents the location and amplitude of scattering centers, which is sparse under $\mathbf{\Phi}$; and $\boldsymbol{\varepsilon} \in \mathbb{C}^N$ is the noise vector. Particularly, the over-complete dictionary $\mathbf{\Phi}$ could be represented by basis functions (or atoms).

Because $\mathbf{x}$ is sparse, one can find the optimal solution to $\mathbf{x}$ by assuming that most of its entries are exactly zero. Making use of this property, the problem of estimating $\mathbf{x}$ can be reformulated as l_0-norm optimization and solved by greedy algorithms such as matching pursuit (MP) and orthogonal matching pursuit (OMP) [6]. Although they can be conveniently implemented, the greedy algorithms generally result in a local optimum [7]; and it is difficult to set their stopping threshold with low SNR.

As a substitution, two types of strategies have been proposed to obtain sparse solutions. The first type adds a penalty term to the target function $\min \|\mathbf{y} - \mathbf{\Phi x}\|_2^2$, which is proportional to the l_1-norm of $\mathbf{x}$ [8]. Such strategies require careful tuning of the regularization parameter, which is still an open problem [7]. Additionally, their performance degrades rapidly with low SNR. The second type of strategies assumes sparsity-promoting priors on $\mathbf{x}$, and estimates the model parameters using the maximum posterior, expectation maximization, variational inference, or sampling techniques in a Bayesian framework [9]-[11]. By introducing sparse priors, the posterior of most coefficients is strongly peaked at zero, while only a few of the coefficients have a large posterior being significantly larger than zero. Different sparsity-promoting priors have been proposed in machine learning and statistic literature, and typical ones include the Laplace, Gamma-Gaussian, and spike-and-slab, etc.

The sparse signal reconstruction techniques introduced above have found successful applications in short or sparse aperture imaging [12]-[13], and in ultra-wideband synthesizing [14]. Although they have exhibited certain robustness to receiver noise, a significant limitation of the state-of-the-art methods is that they usually assume the noise to be

This work was supported in part by the National Natural Science Foundation of China under Grant 61522114, Grant 61631019, in part by the NSAF under Grant U1430123, in part by the Foundation for the Author of National Excellent Doctoral Dissertation of P. R. China under Grant 201448, and in part by the Young Scientist Award of Shaanxi Province under Grant 2016KJXX-82.

independent, zero-mean, and Gaussian distributed. As discussed above, measured datasets gathered in complex environments are often corrupted by unknown noise, e.g. outliers and time or spatial-varying jamming, whose distribution is unlikely to be purely Gaussian. Accordingly, the model mismatch will invalidate the available techniques and generate poor estimates of $\mathbf{x}$.

B. Contribution

In this paper, we propose a new statistical model for high-resolution radar imaging in complex environments when the noise is no longer single Gaussian. The key development is that, we use Gaussian mixture models (GMM) to describe the unknown noise $\boldsymbol{\varepsilon}$. By using a sufficient number of Gaussians, and by adjusting their means, covariances and mixing coefficients, almost any continuous density can be approximated to arbitrary accuracy [15]. In this way, we achieve flexible modeling and increase the robustness to complex environments.

In the new model, we introduce a Gamma-Gaussian hierarchical prior for $\mathbf{x}$, which is sparsity-promoting and conjugate. The model parameters are then estimated using a MAP-EM framework, which converges quickly and generates closed-form solutions. The model order, i.e. the number of Gaussian mixtures, is determined automatically by keeping the class-label with maximum posterior during iteration. Experiments have shown that in complex environments such as strong barrage jamming, the proposed method obtains better performance than traditional imaging techniques.

The remainder of this paper is organized as follows. Section II presents the signal model for high-resolution imaging in complex environments with unknown noise. Section III constructs the probabilistic model for sparse imaging with GMM noise, derives the MAP-EM solution to model parameters, and discusses the algorithm initialization and computational burden. Section IV tests the performance in various scenarios, and applies the proposed method to high-resolution radar imaging from sparse apertures with strong jamming. Results show that although the performance of traditional methods degrades, the proposed method obtains satisfying imaging results. Section V concludes the paper and discusses future work.

II. Signal Model

In this paper, the imaging radars are operating in the high-frequency regime and the point-scattering model holds [3]. With plane wave, small accumulation angle, and fixed turntable model assumptions, the radar echoes collected in complex environments could be described by a unified model similar to (1), in which $\boldsymbol{\varepsilon}$ denotes the unknown noise. In the following discussions, we will present the signal model for high-resolution azimuth imaging. Still, this model can be readily extended to other imaging applications.

For the convenience of discussion, we model the radar echoes after matched-filtering in the range-slow time domain. For the target that rotates a small angle during the imaging interval, or for echoes after migration through range cell correction [3], we have $\Delta R_p(t_m) \approx x_p + y_p \omega_{rot} t_m$ [12], where ω_{rot} is the angular rotation frequency of the turntable model. Then, the range-compressed echoes satisfy,

$$s(r,t_m) = \sum_{p=1}^{P} A_{p_a} \mathrm{a}_r(x_p) \exp\left(j2\pi\omega_{p_a} m\right) + UN(r,t_m) \quad (2)$$

where r is the range dimension, t_m is the slow-time, and $\mathrm{a}_r(x_p)$ is a sinc-like function whose main-lobe width is inversely proportional to the radar bandwidth B; $A_{p_a} = A_p \exp\left(-j4\pi x_p/\lambda\right)$, $\mathrm{a}_r(x_p)$, and $\omega_{p_a} = -2y_p\omega_{rot}t_m/\lambda$ determine the amplitude, range coordinate, and Doppler of the p th scattering center, respectively. Additionally, $UN(r,t_m)$ is the unknown noise such as barrage jamming.

Let m_a be the index of observable samples in sparse or short aperture observation with $card\{m_a\} = M_0$, then in the r th range bin, the sparse signal reconstruction problem is formulated to be,

$$\mathbf{y}_a = \boldsymbol{\Phi}_a \mathbf{x}_a + \mathbf{UN} \quad (3)$$

where $\mathbf{y}_a = \left[s_r(m_{a1}), \cdots s_r(m_{aM_0})\right]^T$ is the signal vector and $\mathbf{UN}$ is the unknown noise vector; $\mathbf{x}_a$ is the sparse azimuth image and $\boldsymbol{\Phi}_a \in \mathbb{C}^{M_0 \times D_a}$ is the Doppler dictionary. Particularly, the row vectors of $\boldsymbol{\Phi}_a$ satisfy

$$\boldsymbol{\Phi}_a(m_a,\cdot) = \exp\left(j2\pi\omega_{d_a} m_a\right), d_a \in [1, D_a] \quad (4)$$

where $\{\omega_{d_a}\}$ constitutes the Doppler grid, and $D_a \gg M_0$.

III. Bayesian Modeling and the MAP-EM Approach

Analysis in the previous section indicates that the problem of high-resolution radar imaging in complex environments with unknown noise could be described by a unified model similar to (1), while the main difference is the non-Gaussian noise $\boldsymbol{\varepsilon}$. Considering complex-valued radar echoes, we generalize (1) as

$$\begin{bmatrix} \mathrm{Re}(\mathbf{y}) \\ \mathrm{Im}(\mathbf{y}) \end{bmatrix} = \begin{bmatrix} \mathrm{Re}(\boldsymbol{\Phi}) & -\mathrm{Im}(\boldsymbol{\Phi}) \\ \mathrm{Im}(\boldsymbol{\Phi}) & \mathrm{Re}(\boldsymbol{\Phi}) \end{bmatrix} \begin{bmatrix} \mathrm{Re}(\mathbf{x}) \\ \mathrm{Im}(\mathbf{x}) \end{bmatrix} + \begin{bmatrix} \mathrm{Re}(\boldsymbol{\varepsilon}) \\ \mathrm{Im}(\boldsymbol{\varepsilon}) \end{bmatrix} \quad (5)$$

where $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ denote taking the real and imaginary parts, respectively. For simplicity and consistency, we redefine the signal model in the real domain, i.e.

$$\mathbf{y} = \boldsymbol{\Phi}\mathbf{x} + \boldsymbol{\varepsilon} \quad (6)$$

where

$$\mathbf{y} = \left[\mathrm{Re}(\mathbf{y}), \mathrm{Im}(\mathbf{y})\right]^T, \ \boldsymbol{\Phi} = \begin{bmatrix} \mathrm{Re}(\boldsymbol{\Phi}) & -\mathrm{Im}(\boldsymbol{\Phi}) \\ \mathrm{Im}(\boldsymbol{\Phi}) & \mathrm{Re}(\boldsymbol{\Phi}) \end{bmatrix}$$

$$\mathbf{x} = \left[\mathrm{Re}(\mathbf{x}), \mathrm{Im}(\mathbf{x})\right]^T, \ \boldsymbol{\varepsilon} = \left[\mathrm{Re}(\boldsymbol{\varepsilon}), \mathrm{Im}(\boldsymbol{\varepsilon})\right]^T$$

Then, $\mathbf{y} \in \Re^{2N}$ is the vector of available samples; $\mathbf{\Phi} \in \Re^{2N \times 2D}$ is the overcomplete dictionary; and $\mathbf{x} \in \Re^{2D}$ are weights. Particularly, $\boldsymbol{\varepsilon} \in \Re^{2N}$ is the vector of unknown noise, which cannot be represented linearly by the basis function of $\mathbf{\Phi}$.

A. Bayesian Modeling

According to (6), the i th observation satisfies

$$y_i = \boldsymbol{\varphi}_i \mathbf{x} + \varepsilon_i \tag{7}$$

where $\boldsymbol{\varphi}_i$ is the i th row of $\mathbf{\Phi}$ with $i \in [1, 2N]$. For unknown noise or jamming, we model ε_i by GMM, i.e.

$$p(\varepsilon_i) = \sum_{j=1}^{J} \pi_j N\left(\mu_j, \alpha_j^{-1}\right) \tag{8}$$

where j is the cluster index, J is the number of Gaussian clusters; π_j is the mixing coefficients with $\sum \pi_j = 1$, $\pi_j \geq 0$; and (μ_j, α_j) is the mean and precision of the j th Gaussian cluster. Particularly, α_j obeys the Gamma distribution,

$$\mathrm{Gam}(\alpha_j | a, b) = \frac{1}{\Gamma(a)} b^a \alpha_j^{a-1} e^{-b\alpha_j} \tag{9}$$

The model defined in (8) can be expanded as

$$p(c_i) = \mathrm{Discrete}(\boldsymbol{\pi}),\ p(\varepsilon_i \mid c_i) = N\left(\mu_{c_i}, \alpha_{c_i}^{-1}\right) \tag{10}$$

where the class label of the i th input c_i is a discrete distribution of $\boldsymbol{\pi} = [\pi_1, \pi_2, \cdots \pi_J]^T$, and the prior of $\boldsymbol{\pi}$ obeys the Dirichlet distribution, i.e.

$$\mathrm{Dir}(\boldsymbol{\pi} | \boldsymbol{\gamma}) = \frac{\Gamma(\gamma_0)}{\Gamma(\gamma_1) \cdots \Gamma(\gamma_J)} \prod_{j=1}^{J} \pi_j^{\gamma_j - 1} \tag{11}$$

with $\gamma_0 = \sum_{j=1}^{J} \gamma_j$.

Because the number of observations is usually smaller than the dimension of weights, we assume that $\mathbf{x}$ is sparse and impose Gamma-Gaussian hierarchical model, i.e.

$$\mathbf{x} = N\left(0, \mathbf{\Lambda}^{-1}\right) \tag{12}$$

where $\mathbf{\Lambda} = \mathrm{diag}\{\lambda_1, \cdots \lambda_d, \cdots, \lambda_{2D}\}$ is the precision matrix. To increase the model flexibility, we assume that the λ_d s are different with $\lambda_d = \mathrm{Gam}(e_d, f_d)$, i.e.

$$\mathrm{Gam}(\lambda_d | \mathrm{e}_d, f_d) = \frac{1}{\Gamma(\mathrm{e}_d)} f_d^{e_d} \lambda_d^{e_d - 1} e^{-f_d \lambda_d} \tag{13}$$

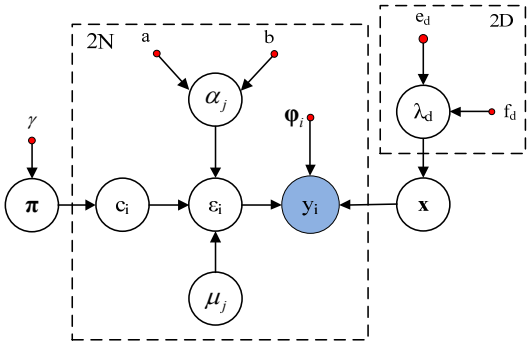

Fig. 1. Graphical model of the proposed method.

The graphical model is shown in Fig. 1, and the log-likelihood is derived to be

$$\begin{aligned} &\ln p(\mathbf{y}, \mathbf{c}, \boldsymbol{\lambda}, \mathbf{x}, \boldsymbol{\pi}, \boldsymbol{\alpha}, \boldsymbol{\mu}) \\ &= \sum_{i=1}^{2N} \sum_{j=1}^{J} \mathbf{1}(c_i = j) \left(\ln \pi_j + \frac{1}{2} \ln \alpha_j - \frac{1}{2} \alpha_j \left(y_i - \boldsymbol{\varphi}_i \mathbf{x} - \mu_j \right)^2 \right) \\ &\quad - \frac{1}{2} \mathbf{x}^T \mathbf{\Lambda} \mathbf{x} + \sum_{j=1}^{J} \left((a-1) \ln \alpha_j - b\alpha_j + (\gamma_j - 1) \ln \pi_j \right) + \text{const.} \end{aligned} \tag{14}$$

where $\mathbf{c} = [c_1, \cdots, c_{2N}]^T$, $c_i \in \mathbb{Z}$, is the vector of cluster-labels; and the indicator $\mathbf{1}(c_i = j)$ equals to one if $c_i = j$; or zero, otherwise.

Obviously, we cannot find closed-form maximum likelihood solutions to model parameters. As a substitution, we employ the MAP-EM framework for effective parameter inference.

B. MAP-EM Inference of the Model Parameters

In MAP-EM framework, the following equation holds,

$$\begin{aligned} \ln p(X, \theta) &= \int q(\phi) \ln \frac{p(X, \phi, \theta)}{q(\phi)} d\phi \\ &\quad + \int q(\phi) \ln \frac{q(\phi)}{p(\phi | X, \theta)} d\phi \end{aligned} \tag{15}$$

where X denotes the observation; θ denotes model parameters; ϕ denotes the hidden variable; and $q(\phi)$ denotes the posterior. The first term on the right hand side is the lower bound $\mathcal{L}(X, \theta)$, and the second term is the Kullback-Leibler divergence [16].

In our problem, we set $X = \{y_i\}$; $\phi = \{\mathbf{c}, \boldsymbol{\lambda}\}$; $\theta = \{\mathbf{x}, \boldsymbol{\pi}, \boldsymbol{\alpha}, \boldsymbol{\mu}\}$, and define $\phi_i(j) \triangleq q(c_i)$. Then, the proposed MAP-EM method for high-resolution imaging with unknown noise is summarized as follows:

Step1: $k = 0$, initialize $\phi_i^{(k)}(j)$, $\mathbf{\Lambda}_1^{(k)}$, $\mu_j^{(k)}$, and $\alpha_j^{(k)}$ randomly, and set the number of mixtures J;

Step2: Let $k = k + 1$, for $i = 1, \cdots, 2N$, $j = 1, \cdots, J$, do

M-Step:

$$n_j^{(k)} = \sum_{i=1}^{2N} \phi_i^{(k-1)}(j) \tag{16}$$

$$\pi_j^{(k)} = \frac{n_j^{(k)} + \gamma_j - 1}{2N + \sum_{j=1}^{J} \gamma_j - J} \tag{17}$$

$$\mathbf{x}^{(k)} = \left(\sum_{i=1}^{2N} \left(\sum_{j=1}^{J} \phi_i^{(k-1)}(j) \alpha_j^{(k-1)} \right) \boldsymbol{\varphi}_i^T \boldsymbol{\varphi}_i + \boldsymbol{\Lambda}_1^{(k-1)} \right)^{-1} \cdot \sum_{i=1}^{2N} \left(\sum_{j=1}^{J} \phi_i^{(k-1)}(j) \alpha_j^{(k-1)} \right) \left(y_i - \mu_j^{(k-1)} \right) \boldsymbol{\varphi}_i^T \tag{18}$$

$$\alpha_j^{(k)} = \frac{\frac{1}{2} n_j + a - 1}{\frac{1}{2} \sum_{i=1}^{2N} \phi_i^{(k-1)}(j) \left(y_i - \boldsymbol{\varphi}_i^T \mathbf{x}^{(k)} - \mu_j^{(k-1)} \right)^2 + b} \tag{19}$$

$$\mu_j^{(k)} = \frac{\sum_{i=1}^{2N} \phi_i^{(k-1)}(j) \left(y_i - \boldsymbol{\varphi}_i \mathbf{x}^{(k)} \right)}{n_j^{(k)}} \tag{20}$$

E-Step:

$$\phi_i^{(k)}(j) = \frac{\pi_j^{(k)} N\left(\boldsymbol{\varphi}_i \mathbf{x}^{(k)} + \mu_j^{(k)}, \alpha_j^{(k)-1} \right)}{\sum_{j=1}^{J} \pi_j^{(k)} N\left(\boldsymbol{\varphi}_i \mathbf{x}^{(k)} + \mu_j^{(k)}, \alpha_j^{(k)-1} \right)} \tag{21}$$

$$\lambda_d^{(k)} = \frac{e_d + \frac{1}{2}}{f_d + \frac{1}{2}\left(x_d^{(k)} \right)^2} \tag{22}$$

$$\boldsymbol{\Lambda}_1^{(k)} = \mathrm{diag}\left\{ \lambda_1^{(k)}, \cdots, \lambda_{2D}^{(k)} \right\} \tag{23}$$

Then, keep the columns corresponding to the maximum probability for the i th input, update J, and do normalization

$$\phi_i^{(k)}(j) = \phi_i^{(k)}(j) / \sum_j \phi_i^{(k)}(j) \tag{24}$$

Step3: Repeat Step2 until convergence.

C. Discussions

Given sufficient prior knowledge, we could determine the number of clusters and noise labels; otherwise, we choose large J for initialization. To obtain good results, we usually assume uninformative distributions by setting a, b, e_d, and f_d to very small values, e.g. $10^{-5} \sim 10^{-4}$.

It is observed that for $i = 1, \cdots, 2N$, $j = 1, \cdots, J$, $\phi_i^{(k)}(j)$ in the k th iteration corresponds to the entry of the cluster label matrix $\mathbf{L}_{ij}$, thus we could implement key steps of the proposed method efficiently by matrix operation. According to Section III-B, the main operation is matrix production and matrix inversion, and the complexity is $o\left(ND^2\right)$ and $o\left(D^3\right)$, respectively.

IV. Experiments

In this section, we will analyse the performance of the proposed method and compare it with traditional imaging techniques.

A. Illustrative Example

Below, we compare performance of the proposed method with relevance vector machine (RVM) [10], which has some robustness to low SNR. The dictionary $\boldsymbol{\Phi}$ obeys Gaussian distribution $N(0,1)$; the number of available samples is 1024 and the dimension of $\mathbf{x}$ is 1707. As shown in Fig. 2(a), the weights $\mathbf{x}$ have 30 spikes. The samples are contaminated by GMM noise drawn from three noise clusters, $N(-10,2)$, $N(0,1.5)$, $N(10,2)$, and the cluster label is given in Fig. 2(b). For the proposed method, the estimate of $\mathbf{x}$ is given in Fig. 2(c), where the root-mean-square (RMS) error is 0.006. Initially, $J = 10$ and finally it reduced to 3. As shown in Fig. 2(d), the cluster label matrix $\mathbf{L}_{ij}$ gives correct estimates of $\phi_i(j)$. As the noise distribution violates single Gaussian assumption, the RVM method fails with large RMS error, as shown in Fig. 2(e). Comparisons of the estimated probability density functions (PDF) are given in Fig. 2(f), where the RVM method is unable to describe the noise distribution properly.

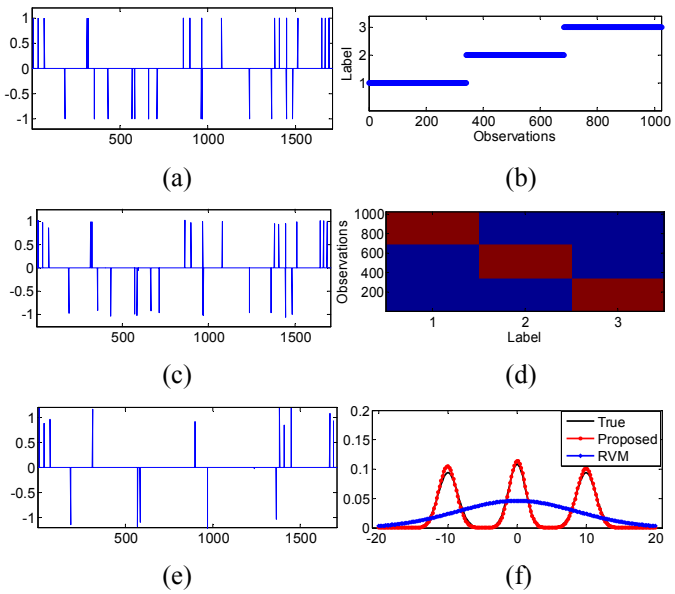

Fig. 2. Estimates of $\mathbf{x}$ in the illustrative example: (a) original weights; (b) noise label; (c) estimated $\mathbf{x}$ using the proposed method; (d) cluster label matrix; (e) estimated $\mathbf{x}$ using RVM; (f) comparison of estimated noise PDFs.

B. Experimental Results for Jamming Suppression

Below, the proposed method is applied to high-resolution imaging of gapped data contaminated by strong jamming. In the experiment, measured data of the Yak-24 plane is

contanimated by time-varying wideband jamming. To make matters worse, some echoes are missing when continuous measurements are not possible or measurements during certain periods are invalid, as shown in Fig. 3(b). The data missing rate is 50%.

For the complete data shown in Fig. 3(a), the original RD image in Fig. 3(c) is well-focused. For contaminated echoes, however, strong jamming exists in the OMP image, as shown in Fig. 3(d). Similarly, residual jamming exists in the RVM image in Fig. 3(e). The image obtained by the proposed method is shown in Fig. 3(f), which is well-focused with clear background. Compared with Fig. 3(c), strong scattering centers in Fig. 3(f) are enhanced because of the sparsity-promoting prior.

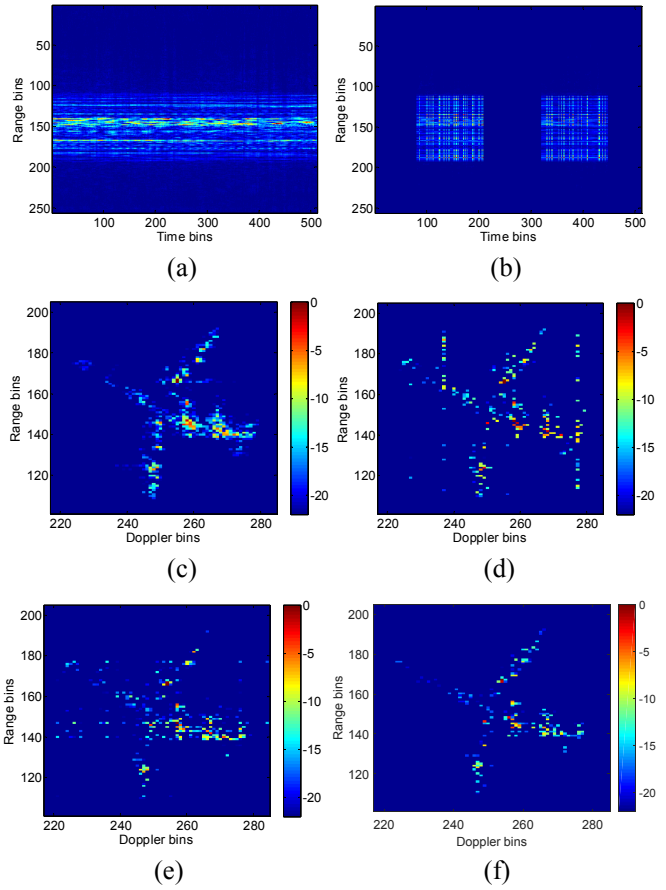

Fig. 3. Jamming suppression of Yak-42 plane: (a) range-compressed echoes; (b) gapped echoes contaminated by jamming; (c) RD image of (a); (d) OMP image; (e) RVM image; (f) image of the proposed method.

V. CONCLUSION

This paper treated high-resolution radar imaging in complex environments with unknown noise as Bayesian inference in a statistical model of weights and noise. The former were governed by Gamma-Gaussian sparsity-promoting prior; and the latter was modelled by GMM, which turned the jamming suppression problem into parameter estimation of each Gaussian cluster. To obtain weights, we resorted to the MAP-EM framework, in which parameters are estimated iteratively by maximizing the lower bound.

Contrary to related works which assumes independent, zero-mean, single Gaussian noise, we considered noise or model error with unknown distribution. We showed that the proposed model had closed-form solution, and can be robust to complex environments with satisfying performance. Experimental results demonstrated that the proposed method provide an effective way to high-resolution radar imaging with barrage jamming.

In the future, we would find surrogation for the Gamma-Gaussian prior which is selective in enforcing the sparsity of weights, and find effective ways to deal with high-resolution imaging of non-steadily moving targets in complex environments.

REFERENCES

[1] J. J. Stambaugh, R. K. Lee, and W. H. Cantrell, "The 4 GHz bandwidth millimeter-wave radar," *Lincoln Laboratory Journal*, vol. 19, no. 2, pp. 64–76, 2012.

[2] M. G. Czerwinski and J. M. Usoff, "Development of the Haystack ultrawideband satellite imaging radar," *Lincoln Laboratory Journal*, vol. 21, no. 1, pp. 28–44, 2014.

[3] W. G. Carrara, R. S. Goodman, and R. M. Majewski, *Spotlight Synthetic Aperture Radar-Signal Processing and Algorithms*. Boston, MA, USA: Artech House, 1995, ch. 2.

[4] A. W. Doerry, "Comments on radar inference sources and mitigation techniques," *SPIE 2015 Defense & Security Symposium*, Vol. 9461, pp. 1–8.

[5] Y. Chen, Q. Zhang, N. Yuan, Y. Luo, and H. Lou, "An adaptive ISAR-imaging-considered task scheduling algorithm for multi-function phased array radars," *IEEE Trans. Signal Process.*, vol. 63, no. 19, pp. 5096–5110, Oct. 2015.

[6] J. Tropp. "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[7] L. Zhao, L. Wang, L. Yang, A. M. Zoubir, and G. Bi, "The race to improve radar imagery," *IEEE Signal Process. Mag.*, vol. 33, no. 6, pp. 85–102, Nov. 2016.

[8] D. Donoho, M. Elad, and V. Temlyakov. "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 6–18, 2006.

[9] T. Moon, "The EM algorithm in signal processing," *IEEE Signal Process. Mag.*, vol. 13, no. 6, pp. 47–60, 1996.

[10] D. Tzikas, A. Likas, and N. Galatsanos, "The variational approximation for Bayesian inference," *IEEE Signal Processing Mag.*, vol. 25, no. 6, pp. 131–146, Nov. 2008.

[11] C. Andrieu, N. De Freitas, A. Doucet, and M. Jordan, "An introduction to MCMC for machine learning," *Mach. Learn.*, vol. 50, no. 1-2, pp. 5–43, 2003.

[12] X. Bai, F. Zhou, M. Xing, and Z. Bao, "High-resolution radar imaging of air-targets from sparse azimuth data," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 2, pp. 1643–1655, Apr. 2012.

[13] L. Wang, L. Zhao, G. Bi, C. Wan, and L. Yang, "Enhanced ISAR imaging by exploiting the continuity of the target scene," *IEEE Trans. Geosci. Remote Sensing*, vol. 52, no. 9, pp. 5736–5750, 2014.

[14] H. Zhang and R. Chen, "Coherent processing and superresolution technique of multi-band radar data based on fast sparse Bayesian learning algorithm," *IEEE Trans. Antennas Propag.*, vol. 62, no. 12, pp. 6217–6227, Dec. 2014.

[15] D. Meng, F. Torre, "Robust matrix factorization with unknown noise," *Proc. IEEE International Conference on Computer Vision (ICCV)*, Dec. 2013.

[16] C. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2006, ch. 9.