The Race to Improve Radar Imagery: An overview of recent progress in statistical sparsity-based techniques

Lifan Zhao, Lu Wang, Lei Yang, *Member, IEEE*, Abdelhak M. Zoubir, *Fellow, IEEE*, Guoan Bi, *Senior Member, IEEE*

Abstract

The exploitation of sparsity has significantly advanced the field of radar imaging over the last few decades, leading to substantial improvements in resolution and quality of the processed images. More recent developments of compressed sensing (CS) suggest that statistical sparsity can lead to further performance benefits by imposing sparsity as a statistical prior on the considered signal. In this article, a comprehensive survey is made on recent progress of statistical sparsity based techniques for various radar imagery applications. Firstly, we give a brief review of state-of-the-art sparsity based techniques for radar imagery. Then, we present a comprehensive review of statistical sparse models and algorithms as a basis for subsequent sections where various applications are treated. In particular, we stress the differences with sparse regularized models and highlight the advantages of statistical approaches. Subsequently, we cover super-resolution imaging, enhanced target imaging, auto-focusing and moving target imaging, all of which are treated with statistical sparsity based methods. We demonstrate how to achieve desirable improvements by adequately manipulating the statistical sparse model as compared with conventional sparsity based methods. Finally, open questions and potential directions are identified to motivate future research in the area of radar imagery.

Lifan Zhao, Lei Yang and Guoan Bi are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore (email: zhao0145@e.ntu.edu.sg; yanglei@ntu.edu.sg; egbi@ntu.edu.sg).

Lu Wang is with the School of Marine Science and Technology, Northwestern Polytechnical University, China (email:wanglu@nwpu.edu.cn).

Abdelhak M. Zoubir is with the Signal Processing Group, Darmstadt University of Technology, Darmstadt, Germany (email: zoubir@spg.tu-darmstadt.de).

Index Terms

Sparsity based radar imagery, statistical sparsity, high resolution, compressed sensing

I. INTRODUCTION

The capability of operating in all-day-night and all-weather conditions has allowed high-resolution radar imagery to play an important role in both civilian and military remote sensing applications. The principle of radar imagery is to utilize wideband transmitting signals and aperture synthesis to obtain the desirable slant-range and cross-range resolution, respectively [1]. However, the processing of wideband signals requires high-speed analog-to-digital converters at the receiver. Also, a long coherent processing interval (CPI) is required to obtain high cross-range resolution, which would inevitably introduce undesirable higher-order Doppler effects. To overcome the above limitations, sparsity based techniques [2] have become increasingly important in radar imagery. Successful applications in recent years can be found in [3]–[7] and references therein. In addition to achieve high-resolution radar images with limited data, sparsity based techniques can offer additional advantages over conventional spectral analysis based techniques, such as de-noising and side-lobe suppression [3], [4].

In a nutshell, the theory of CS states that a high-dimensional signal can be accurately and robustly recovered from its low-dimensional projections if the signal is sparse or can be sparsely represented [8]. The success of CS techniques is the proper exploitation and utilization of sparsity. One obvious way to solve the problem is to enforce sparsity, i.e., to minimize the number of non-zero entries in the signal. However, the main difficulty in solving an ℓ_0 minimization problem, i.e. minimizing the number of its non-zero elements, is generally NP-hard, which requires an exhaustive search with intractable computational cost. Various algorithms have been developed in CS to obtain an approximate solution to this problem. In many radar imagery applications, CS theory can be conveniently applied due to the natural presence of sparsity. In particular, the target scene of interest is often parsimonious or can be parsimoniously represented by an appropriately chosen linear basis [3], [4]. Despite the diverse applications of sparsity based techniques in radar imaging, methodologically, these techniques can be summarized into three main categories.

1) The first category is based on the use of greedy algorithms that lead to a sparse solution in an iterative and greedy manner. Often, these algorithms require approximation of the signal's support and amplitude by refining the sparse signal estimation based on evaluation of the difference between the recovered signal and the observations. In [9], for example,

greedy algorithm based CS is used to operate wide-swath modes for radar imaging with reduced measurements. In [10], the authors investigate the uniform and non-uniform target imaging problem with greedy approaches. Although greedy algorithms can be conveniently implemented and have desirable guarantees under some conditions, they generally result in a local optimum, which does not coincide with the sparsest solution.

2) Another major category of algorithms is a generalization of an ℓ_1 regularized optimization method, which can be considered as the tightest convex relaxation of the ℓ_0 minimization problem. Basis Pursuit (BP) and Basis Pursuit Denoising (BPDN) [11] are convex formulations to recover sparse signals in noiseless and noisy environments, respectively. Various approaches have been proposed to solve the ℓ_1 minimization problem, such as linear programming and the interior point method. Since the objective functions are convex, it is guaranteed that these algorithms lead to a global optimum. However, the solution does not necessarily coincide with the maximally sparse solution, except that the problem satisfies some specific conditions. In [12], [13], the ℓ_1 regularized optimization is applied for radar image reconstruction. In [14], [15], phase error correction and imaging are formulated within a convex optimization framework. The synthetic aperture radar (SAR) ground moving target problem is also formulated in a sparsity-driven manner and solved by convex optimization techniques [5], [16]. Remarkably, empirical results suggest that the ℓ_1 regularized optimization has substantial improvements over greedy approaches for some radar imaging problems [3]. However, the regularization parameter should be carefully tuned to obtain a desirable performance, but finding an optimal selection rule is still an open problem.
3) The third category of statistical sparsity based methods [17] provides remarkable statistical advantages over conventional ones. Under certain conditions, the resulting algorithm guarantees that its global optimum coincides with the maximally sparse solution and smooths the shallow local minimum [18]. Theoretical and empirical results show that enhanced performance can be achieved from Bayesian inference over conventional ℓ_1 regularized optimization [17], [19]–[21]. This technique is particularly useful in overcoming some limitations of the above two categories. The advantages for radar imagery applications include incorporation of flexible prior knowledge, estimation accuracy improvement, as well as estimation of error bars.

The main objective in radar imagery applications is to properly utilize one of these methods

to obtain enhanced imaging performance, which is particularly useful in situations where the number of measurements is limited and the signal-to-noise ratio (SNR) is low. Our objective in this review article is to present the motivation and ways of utilizing statistical sparsity based radar imaging techniques.

Recent overview articles [3]–[5] focus on sparsity based radar imagery techniques from either a greedy or regularized perspective. These articles demonstrate the importance and effectiveness of sparsity in radar imagery applications. With the recent development of sparse Bayesian methods [17], [18], statistical sparsity based techniques have become a more promising research area for radar imagery applications. Compared to the deterministic sparsity-inducing framework, statistical sparsity based techniques provide new opportunities to significantly improve the performance of radar imagery. This is mostly due to the capability of avoiding regularization parameter tuning, providing desirable statistical information and allowing flexible modeling. These advantages are respectively due to the inherent advantage of the statistical framework, the desirable statistical information obtained from the estimation of the full posterior distribution, and the inherent flexibility of statistical sparse based model. In order to benefit from these advantages, sophisticated design is required. As a matter of fact, this article is a good companion of recent tutorial articles on sparsity based radar imagery [3]–[5], but with particular emphasis on sparsity based radar imagery from a statistical perspective, which is missing in the recent literature.

We show how this design is to be performed and demonstrate how the various radar imagery problems can be formulated within a sparse Bayesian framework to exploit sparsity. We illustrate in detail why the statistical formulation greatly enhances the radar imaging performance in various practical problems. The introduced framework has much promise for future radar imaging systems as it provides substantial improvements as well as new opportunities. The notations used in this article are summarized in Table I.

II. Statistical Sparsity Formulation Of Radar Imagery

We begin our treatment by reviewing the fundamentals of statistical sparsity-inducing models in radar imagery. We formulate the statistical sparsity based framework and highlight where the advantages arise from, along with the limitations of statistical sparsity based methods.

A. Data Modeling

In high-resolution radar imagery, the scattering response of a target of interest is often expressed as a sum of scatterers' responses. Without loss of generality, assuming that the radar emits

TABLE I

NOTATION SUMMARY

Notations	
$\mathbb{C}^{M\times N}$	The set of a complex $M \times N$ matrix
a, $\mathbf{a}$	Scalar and vector
$\mathbf{A}$, $\mathbf{A}_{i\cdot}$, $\mathbf{A}_{\cdot j}$, $\mathbf{A}_{n,m}$	Matrix, the i-th row, the j-th column and the (n,m)-th entry of a matrix
$(\cdot)^T$ and $(\cdot)^H$	Matrix or vector transpose and conjugate transpose
$\mathbf{A}^{-1}$	Matrix inverse
$\lvert\cdot\rvert$	The absolute value of a scalar
$\lVert\cdot\rVert_p$	The ℓ_p norm of a vector
$\lVert\cdot\rVert_F$	The Frobenius norm of a matrix
$\exp(\cdot)$	The exponential function
$\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	The complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$
$\mathrm{Beta(a, b)}$	The Beta distribution with parameters a and b
$\Gamma(a, b)$	The Gamma distribution with parameters a and b

successive pulses with time interval T_r and that there exists K strong scatterers in an imaging scene, the received radar signal can be given by

$$s_r(t, t_n) = \sum_{k=1}^{K} \sigma_k \cdot s\left(t - \frac{2R_k(t_n)}{c}\right) + n(t, t_n) \tag{1}$$

where σ_k represents the amplitude of the k-th scatterer, c is the speed of light, and $R_k(t_n)$ represents the range from the radar to scatterer k in slow time t_n. The fast time and slow time of pulse n are denoted by $0 \leq t \leq T_r$ and $t_n = nT_r$, respectively. To achieve high range resolution, the emitted series of pulses $s(t, t_n)$ are often chosen as linear frequency modulated (LFM) signals, but other waveforms such as sparse stepped-frequency signals [22], sparse probing-frequency signals [23], and adaptively optimized signals [24] can also be used for the purpose of improved imaging performance. To achieve high cross-range resolution, a large aspect angle between the radar and the target is required during the CPI. Note that sparsity based methods go beyond the convention in the sense that high range resolution can be obtained with less bandwidth and high cross-range resolution can be obtained with a reduced coherent processing time.

After pre-processing and arranging the cross-range measurements column-wise, a linear model is obtained as

$$\mathbf{Y} = \mathbf{AXB} + \mathbf{N} \tag{2}$$

where $\mathbf{Y} \in \mathbb{C}^{M\times N}$ is the pre-processed radar echoes, $\mathbf{X} \in \mathbb{C}^{M\times N}$ is the unknown scattering

coefficient, and $\mathbf{A} \in \mathbb{C}^{M\times M}$ and $\mathbf{B} \in \mathbb{C}^{N\times N}$ are the measurement matrices constructed from (1) for cross-range and range, respectively. In general, $\mathbf{A}$ and $\mathbf{B}$ are Fourier matrices. There exist, however, other ways to construct the dictionary other than simply employing the Fourier matrix. These include the frame based matrix [25] and the matched filter based matrix [26]. Note that the model in (2) does not yet include the case of under-sampled or an incomplete measurement of $\mathbf{Y}$. Modifications to capture these are straightforward. A more detailed discussion on this issue will be presented in Section III.

In (2), $\mathbf{N}$ is assumed to be independently circularly-symmetric complex Gaussian distributed so that the received signal $\mathbf{Y}$ follows a complex Gaussian distribution with a likelihood function given by

$$p(\mathbf{Y}|\mathbf{X},\alpha_0) = \prod_{i=1}^{M}\prod_{j=1}^{N} \mathcal{CN}(\mathbf{Y}_{ij}|\mathbf{A}_{i\cdot}\mathbf{X}\mathbf{B}_{\cdot j}, \alpha_0^{-1}\mathbf{I}) \tag{3}$$

where α_0 is the noise precision or the reciprocal of the variance.

For the sake of convenient inference for the noise level, we model the noise precision with a Gamma distribution,

$$p(\alpha_0|v_1,v_2) = \Gamma(\alpha_0|v_1,v_2)$$

where v_1 and v_2 are hyper-parameters, often set as trivial values to impose a non-informative prior on the noise level. The rationale of the Gamma prior for the noise precision is due to the likelihood and prior conjugacy [27]. By this modeling, noise level estimation can be naturally incorporated into the signal estimation task.

In order to obtain a solution for the linear equation in (2), sparsity is often imposed to constrain the solution space. Unlike convex based approach constraining the solution space by regularization, statistical approaches seek posterior estimation to robustly estimate the signal $\mathbf{X}$ by properly imposing sparse priors on the signals to be estimated.

B. Probabilistic Sparse Modeling

In sparsity based radar imaging applications, we impose sparsity on target scattering coefficients $\mathbf{X}$, i.e., some strong scatterers in the target scene. In other words, most of the coefficients in $\mathbf{X}$ are zeros or nearly zeros due to the fact that the scatterers are sparsely distributed in the imaging scene. We show statistical ways of imposing sparsity on various radar imaging applications. In order to impose sparse priors and allow inference, the models need to be carefully designed. Essentially, the model should have two key characteristics of sparsity and model conjugacy. In

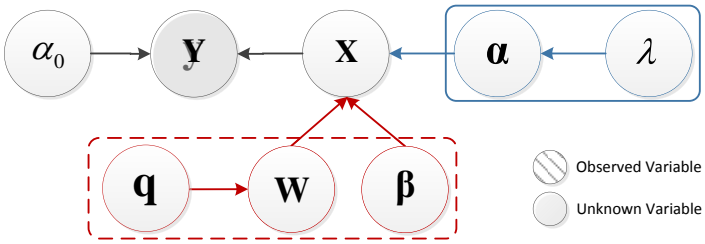

Fig. 1. Graphical representation of the scaled Gaussian mixture model (in the frame of solid line) and the spike-and-slab model (in the frame of broken line).

other words, the constructed model should not only induce sparsity, but also allow convenient inference of the unknown parameters. Towards this end, a hierarchical model is often utilized instead of a single layer model. In the sequel, two main classical models, the so-called scaled Gaussian mixture model and the spike-and-slab model, are briefly reviewed, which is shown in the frame of the solid line and the broken line in Fig. 1, respectively. In the scaled Gaussian mixture model, the signal is assumed to follow a Gaussian distribution and its variance is assumed to follow a particular distribution to induce sparsity and convenient inference. In contrast, in the spike-and-slab model, the signal is assumed to be a dot product of the support and the amplitude coefficients, where the support coefficient is hierarchically modeled for the sake of sparsity achievement.

1) ***The Scaled Gaussian Mixture Model****:* Although there exist various models within this class, we choose to briefly review the three-stage hierarchical model [28] as an example. As shown in Fig. 1, the sparse signal $\mathbf{X}$ is hierarchically constructed.

- In the first stage, the sparse signal $\mathbf{X}$ is modeled with a complex Gaussian distribution,

$$p(\mathbf{X}|\boldsymbol{\alpha}) = \prod_{i=1}^{M}\prod_{\mathrm{j}=1}^{\mathrm{N}} \mathcal{CN}(\mathbf{X}_{ij}|\mathbf{0}, \alpha_{ij}). \tag{4}$$

- In the second stage, we find the distribution of the variance $\boldsymbol{\alpha}$ of the scattering coefficient $\mathbf{X}$. It is assumed to follow an independent Gamma distribution since it is the conjugate prior of a Gaussian distribution, and thus, it makes inference tractable [27],

$$\mathrm{p}(\boldsymbol{\alpha}_{ij}|\boldsymbol{\lambda}) = \prod_{i=1}^{M}\prod_{j=1}^{N} \Gamma(\alpha_{ij}|\eta, \lambda) \tag{5}$$

  Combining (4) and (5), it can be shown that the marginalized distribution of $\mathbf{X}_{ij}$ is a complex Laplace distribution, which is known to be a suitable model for sparsity [2].

TABLE II

SUMMARY OF THE SCALED GAUSSIAN MIXTURE FOR SPARSE MODELING

Scaled Gaussian Mixture Model		
Number of Stages	Model Specification	Marginalized Distribution
Two Stages	Gaussian-Jeffery [29]	No Closed-form Representation
	Gaussian-Gamma [30]	Laplace Distribution
	Gaussian-Inverse-Gamma [17]	Student's t Distribution
	Gaussian-Exponential [31]	Double Exponential Distribution
	Gaussian-Half-Cauchy [32]	No Closed-form Representation
Three Stages	Gaussian-Gamma-Gamma [33]	Laplace Distribution

- In the third stage, we choose the Gamma distribution

$$p(\lambda|v_3, v_4) = \Gamma(\lambda|v_3, v_4) \tag{6}$$

in order to infer $\boldsymbol{\lambda}$ that controls sparsity of the prior during the learning from the data.

There are many variants of the hierarchical model, which can all be summarized as scaled Gaussian mixture models in Table II.

*2) **The Spike-and-Slab Model**:* In contrast, the spike-and-slab model is another popular approach to model the support and amplitude of sparse signals. In particular, the prior of the sparse signal is expressed as a mixture of a point probability distribution (spike) and a Gaussian distribution (slab),

$$p(\mathbf{X}|\mathbf{W}, \beta) = \prod_{i=1}^{M}\prod_{j=1}^{N}[(\mathbf{1} - \mathbf{W}_{ij}) \cdot \delta(\mathbf{X}_{ij}) + \mathbf{W}_{ij} \cdot \mathcal{CN}(\mathbf{X}_{ij}|\mathbf{0}, \beta_{ij})] \tag{7}$$

where δ is the point probability mass centered at 0, $\mathbf{W}$ is the support coefficient, and $\boldsymbol{\beta}$ controls the amplitude of coefficient $\mathbf{X}$. The support coefficient $\mathbf{W}$ determines the sparsity profile of the signal and the amplitude coefficient $\boldsymbol{\beta}$ controls the amplitude of the signal.

- In the first stage, the support coefficient $\mathbf{W}$ is modeled by a Bernoulli distribution,

$$p(\mathbf{W}|\mathbf{q}) = \prod_{i=1}^{M}\prod_{j=1}^{N} q_{ij}^{\mathbf{W}_{ij}}(1 - q_{ij})^{(1-\mathbf{W}_{ij})}$$

where q_{ij} is the probability of $\mathbf{W}_{ij} = 1$ and $1 - q_{ij}$ is the probability of $\mathbf{W}_{ij} = 0$. Note that in the spike-and-slab model, sparsity can be obtained by imposing a prior such that most of the entries in $\mathbf{W}$ are zeros. In other words, the probability of an entry being zero should

be larger than the probability of an entry being non-zero. In order to conveniently estimate parameter q_{ij}, a Beta distribution is imposed on q_{ij},

$$p(q_{ij}) = \text{Beta}(q_{ij}|e, f)$$

where e and f are hyper-parameters to be set as trivial values. The reason of selecting Beta distribution is for its conjugacy to Bernoulli distribution [27]. In some applications, their values can be specified by some prior information.

- In the second stage, the coefficient $\boldsymbol{\beta}$ that controls the amplitude follows a complex Gaussian distribution,

$$p(\boldsymbol{\beta}) = \prod_{i=1}^{M}\prod_{j=1}^{N} \mathcal{CN}(\beta_{ij}|\mathbf{0}, v_0).$$

To allow for inference of v_0, it can be modeled by a Gamma distribution. In particular, unlike in the Gaussian mixture model, a single variance parameter v_0 is assumed. This is because sparsity has already been captured by the Bernoulli-Beta model in the first stage. Based on the two stage model, this spike-and-slab model can therefore impose sparsity on the signal.

In summary, these two classical models are frequently utilized and modified in statistical sparsity based radar imagery. In the latter part of this article, we will demonstrate ways to fully utilize these two models for desirable improvements in specific applications.

*3) **Connections With Convex Optimization**:* The conventional sparsity regularization based methods can be interpreted from a Bayesian perspective. The Laplace distribution, which is a popular choice as a sparse prior [2], is imposed on the signal $\mathbf{X}$. Then, the maximum $a\ posterior$ (MAP) technique is utilized for parameter estimation. It can be shown that the ℓ_1 regularized method corresponds to the MAP estimation with a Gaussian likelihood and a Laplace prior [18],

$$\begin{aligned}\hat{\mathbf{X}} &= \arg\max_{\mathbf{X}} \frac{p(\mathbf{Y}|\mathbf{X})p(\mathbf{X})}{p(\mathbf{Y})} \\ &= \arg\min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{AXB}\|_F^2 + \lambda \sum_{i=1}^{N} \|\mathbf{X}_{\cdot i}\|_1. \end{aligned} \quad (8)$$

This strategy, however, can only provide point estimation of $\mathbf{X}$ without any higher-order statistical information. In contrast, the full posterior, including higher-order statistical information, can be obtained in a statistical sparse framework due to the hierarchical model. This main difference allows statistical sparsity based methods to perform better in many tasks. For example, incorporating a more sophisticated prior on the signals provides flexibility of the hierarchical Bayesian model. This is not the case in the regularized framework.

In summary, the statistical sparsity based models, such as the scaled Gaussian mixture and the spike-and-slab model, avoid the laborious tuning of the regularization parameter λ in (8). These methods are also flexible in view of choosing different priors and provide higher-order statistical information in the posterior distribution (due to the Bayesian inference). In the latter sections of this article, we will show ways to properly manipulate the statistical sparse model so as to make use of these desirable properties.

C. Bayesian Inference

Based on the formulated probabilistic model, the remaining task is to infer the parameters. We recall the graphical representation in Fig. 1, where all the unknown variables are required to be estimated. Based on the likelihood of $\mathbf{Y}$ and a scaled Gaussian or a spike-and-slab prior, the posterior distribution can be expressed according to the Bayesian theorem,

$$p(\mathbf{\Theta}|\mathbf{Y}) = \frac{p(\mathbf{\Theta})p(\mathbf{Y}|\mathbf{\Theta})}{p(\mathbf{Y})} \tag{9}$$

where $\mathbf{\Theta}$ is a set of all the parameters to be estimated, i.e., the unknown parameters in Fig. 1. However, one major difficulty is that the marginalized distribution cannot be explicitly calculated due to the intractability of the integral $p(\mathbf{Y}) = \int_{\mathbf{\Theta}} p(\mathbf{\Theta})p(\mathbf{Y}|\mathbf{\Theta})d\mathbf{\Theta}$ and thus the posterior distribution in (9) is not attainable. Although MAP estimation can be obtained from this model, the full posterior is more desirable, as it provides a more accurate description of the estimated parameter.

Because the exact inference is not attainable from this model, two classical inference techniques known as the Markov chain Monte Carlo (MCMC) method and the variational Bayesian (VB) method are often used to approximate the posterior from sampling and optimization, respectively. In this way, the approximated posterior can be obtained at a cost of increased computational complexity compared with other sparse signal recovery methods. The MCMC method is accurate when the number of samples becomes large, while the VB method provides a desirable approximation with a reasonable computational complexity.

*1) **The Markov Chain Monte Carlo Method**:* This method is a popular approach to approximate the posterior by sampling. MCMC is a strategy for generating samples, while the equilibrium distribution of the Markov chain is the same as the desired probability distribution [34]. The most widely used MCMC algorithms are the Metropolis-Hastings and the Gibbs sampling algorithms [27]. Under the assumption that all the conditional distributions are available, Gibbs sampling is

easily applicable. In fact, the Gibbs sampling method can be considered as a special case of the Metropolis-Hastings algorithm if the conditional distributions are provided [34].

Since the conditional distribution is available in our graphical model as shown in Fig. 1, we will briefly review the Gibbs sampling method. In this approach, sequential sampling of the conditional distribution, expressed as

$$\mathbf{\Theta}_i \sim p(\mathbf{\Theta}_i|\mathbf{\Theta}_{k\neq i}, \mathbf{Y}), \tag{10}$$

is performed. Therefore, the algorithm iterates until the desirable posterior is obtained. A more detailed description of the algorithm can be found in [34] and references therein.

*2) **The Variational Bayesian Method**:* The main idea of this method is to approximate the true posterior with proper assumption. The assumption herein is that the approximated posterior has a factorisable form,

$$q(\mathbf{\Theta}) = \prod_{i=1}^{k} q(\mathbf{\Theta}_i). \tag{11}$$

This is known as the mean-field assumption [27]. The objective is to find a factorisable $q(\mathbf{\Theta})$ which is as close as the true posterior $p(\mathbf{\Theta}|\mathbf{Y})$. The closeness of the estimated posterior to the true one in the variational Bayesian (VB) method is measured by the Kullback-Leibler (KL) divergence, and thus the optimal approximated posterior is obtained by minimizing the following KL divergence,

$$q^*(\mathbf{\Theta}) = \arg\min_{q(\mathbf{\Theta})} \int q(\mathbf{\Theta}) \ln \frac{q(\mathbf{\Theta})}{p(\mathbf{\Theta}|\mathbf{Y})} d\mathbf{\Theta}. \tag{12}$$

Based on (11) and (12), it can be shown that the approximated posterior for each of the variables can be calculated as [27], [35],

$$q^*(\mathbf{\Theta}_i) = \exp\left\{ \langle \ln p(\mathbf{\Theta}, \mathbf{Y}) \rangle_{q(\mathbf{\Theta}\setminus\mathbf{\Theta}_i)} \right\} \tag{13}$$

where $\langle\cdot\rangle_{q(\cdot)}$ represents expectation with respect to the probability density function $q(\cdot)$.

In the MCMC implemented statistical sparsity based methods, sampling procedure is required for each step during iterations. In contrast, the VB implemented statistical sparsity based methods require calculating matrix inversion in each step during iterations. Notably, the sampling procedure and matrix inversion calculation would generally induce high computational complexity for MCMC and VB, respectively. Due to these reasons, the statistical sparsity based methods generally cost more computations than the non-statistical approaches. The MCMC method can achieve better estimation accuracy than the VB method, but at a higher computational expense. In a practical application, one should choose the method according to its computational cost tolerance.

To conclude this section, we summarize the key advantages of statistical sparsity based methods as follows:

- avoiding regularization parameter tuning. Parameter tuning is not required in statistical sparsity based methods, which will be demonstrated in all the applications reviewed in this article.
- providing full posterior. With this capability, desirable improvements can be achieved by properly manipulating the statistical model, particularly as shown in Sections V and VI.
- flexible modeling. Since the model is constructed probabilistically, encoding the prior can be carried out in a rather flexible way, which will be demonstrated in Section IV.

Despite their remarkable advantages, the key limitations of the statistical sparsity based methods lie primarily in:

- high computational complexity. The generally required computational cost of statistical sparsity based methods is higher than that required by greedy or regularized methods.
- sparsity assumption. The success of almost all sparsity based methods depends on the existence of sparsity or compressibility. If the radar target scene does not exhibit sparsity, modifications should be made to allow a sparse representation [3], [5].

III. Super-Resolution Radar Imagery

In conventional Fourier based radar imagery, the resolutions in cross-range and slant-range are bounded by the Rayleigh limit, which can be overcome by super-resolution techniques. In general, super-resolution radar imaging can be well formulated as an inverse problem, where the scattering field is required to be inversely estimated from the received radar echoes. This problem is ill-posed since it requires estimation of high-dimensional signals from low-dimensional observations. Recent advances in super-resolution radar imagery consider this problem from a sparsity perspective and obtain its solution by either greedy or regularized methods, see a review of state-of-the-art works in [3], [5]. Although empirical results have demonstrated superior performance of sparsity based algorithms over conventional spectral analysis based methods, their success often requires careful selection of key parameters.

To alleviate this drawback, super-resolution imagery techniques have been developed more recently in a statistical sparsity based framework. Herein, the radar returns and the target scattering coefficient are modeled probabilistically, and the formulated models have the advantage to be free of laborious parameter tuning. Assuming pre-processing procedures have been carried out, the

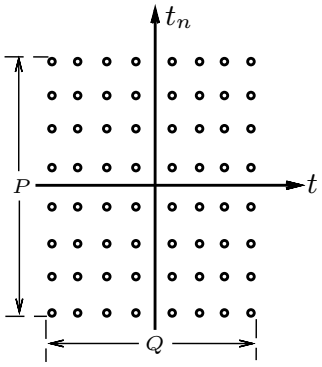

Fig. 2. The obtained data arrangement.

super-resolution radar imaging problem can be explicitly formulated to be under-determined,

$$\mathbf{Y} = \mathbf{\Phi}_1\mathbf{A}\mathbf{X}\mathbf{B}\mathbf{\Phi}_2 + \mathbf{N} \tag{14}$$

where $\mathbf{Y} \in \mathbb{C}^{P\times Q}$ is the pre-processed data, $\mathbf{X} \in \mathbb{C}^{N\times M}$ is the unknown sparse scattering coefficient, $\mathbf{N} \in \mathbb{C}^{P\times Q}$ is the noise, and $\mathbf{\Phi}_1 \in \mathbb{C}^{P\times N}$ and $\mathbf{\Phi}_2 \in \mathbb{C}^{M\times Q}$ represent the cross-range and the slant-range under-sampling matrices, respectively. The data arrangement is shown in Fig. 2 with $P \leq N$ and $Q \leq M$, where each column represents the accumulated echoes from each range cell. An important message herein is that the matrices $\mathbf{\Phi}_1\mathbf{A}$ and $\mathbf{B}\mathbf{\Phi}_2$ should be carefully designed to achieve desirable properties, such as a low mutual coherence or a certain restricted isometry property (RIP) [2]. To achieve these goals, the emitting signal waveform or the under-sampling patterns should be appropriately designed, where examples can be found in [13], [36]. More specifically, the so-called *Alltop* sequences are utilized in [36] as emitting signals, which is proved to practically achieve the lower bound of maximum mutual coherence. In contrast, the authors in [13] investigated different under-sampling patterns to obtain low maximal mutual coherence and achieve radar data compression purpose.

1) In an attempt to obtaining super cross-range resolution for radar imaging, a scheme was proposed in [37]. By setting the under-sampling matrix $\mathbf{\Phi}_2$ to identity and carrying out proper pre-processing, a degenerated model of (14) has been proposed,

$$\mathbf{Y} = \mathbf{\Phi}_1\mathbf{A}\mathbf{X} + \mathbf{N}$$

where $\mathbf{Y} \in \mathbb{C}^{P\times M}$ is the cross-range under-sampled data. The matrix $\mathbf{\Phi}_1\mathbf{A}$ is constructed as a partial Fourier matrix, whose property was investigated empirically in [38].

2) Similarly, in an attempt to obtain super range resolution, probing frequency based signals have been proposed to obtain under-sampled data in range dimension. The mathematical model can be expressed as

$$\mathbf{Y} = \mathbf{X}\mathbf{B}\mathbf{\Phi}_2 + \mathbf{N}$$

where $\mathbf{Y} \in \mathbb{C}^{N\times Q}$ is the range under-sampled data. In particular, the maximum coherence condition of $\mathbf{B\Phi_2}$ is shown in [39]. Apart from the advantage of data compression, this strategy can greatly simplify the radar system design by avoiding emitting wideband signals.

3) In the scenario of obtaining both super cross-range and range resolution, the under-sampling matrices for cross-range and range are $\mathbf{\Phi}_1$ and $\mathbf{\Phi}_2$, respectively. The authors in [40] proposed a random sampling scheme for both slant-range and cross-range dimensions, where random selection is carried out in each dimension. This scheme is particularly useful for reducing data storage costs in various radar imagery applications.

To impose sparsity on the target scene, the above-reviewed works employed a scaled Gaussian mixture model to induce sparsity as well as to avoid parameter tuning. Empirical experimental results found in [37], [40], [41] demonstrate that the statistical sparsity model obtains cleaner images without parameter tuning, as compared to other sparse regularized based methods [40]. Additionally, the statistical sparsity based methods have been empirically shown to be less sensitive to noise and clutter in radar imaging.

Remark: In [37], [40], [41] discussed above, multiplicative speckle noise in radar imagery has not been considered. In presence of speckle noise, the performance of sparsity based methods would be compromised. For this particular problem, the authors in [3], [13] proposed to used ℓ_1 norm of the scene's gradient in addition to ℓ_1 norm of the scene to obtain de-speckled radar image. To the best of our knowledge, the de-speckling problem in radar imaging has not yet been specifically considered under the statistical sparsity based framework. However, similar idea of exploiting gradient sparsity along with sparsity has been proposed in a statistical framework, where example can be found in [42]. An immediate advantage would be that the formulated statistical sparsity based method could be free of regularized parameter tuning process, where good performances could be readily attainable.

In summary, there are two important points to be noted here. Firstly, the success of statistical sparsity based methods depends on the proper selection of sampling patterns, which is also the case for conventional sparsity based methods. In fact, statistical sparsity based methods are less sensitive to highly correlated measurement matrices induced by some sampling patterns than

sparsity regularized methods. A detailed analysis can be found in [43] and references therein. Secondly, a trade-off between design convenience and performance is to be properly balanced, depending on performance requirements for super-resolution in the problem at hand.

IV. Enhanced Target Imagery by Exploiting Structured Sparsity

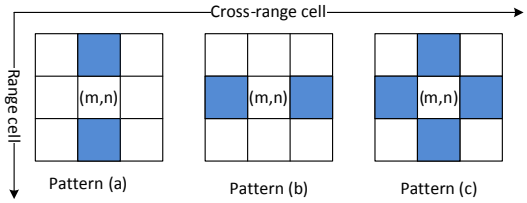

Fig. 3. First-order continuity patterns.

In the previous section, we have demonstrated that statistical sparsity could lead to improvements over deterministic sparsity in super-resolution radar imagery. In order to carry out sparse estimation, the scaled Gaussian mixture is imposed on the scattering coefficients, which are assumed to be independently distributed. However, in practice, targets in radar images always exhibit strong spatial correlation due to the fact that a real target is physically continuous [39], [44]–[46]. To be more concrete, for example, the radar returns from a tank or an airplane will often exhibit strong spatial correlation, i.e., non-zero valued scatterers in the target region continuously residing in the range and/or cross-range dimensions. This phenomenon motivated the research in [39], [45], [46], which properly modifies the statistical sparse model. In these works, continuity in the target scene is exploited by incorporating a correlated prior in a probabilistic framework. In the sequel, we review methods that impose first-order and higher-order correlations on the sparse scattering coefficients.

A. First Order Correlation

In [39], [44], [45], a modification in spike-and-slab modeling was made so as to impose first-order spatial correlation of the coefficients. The reason for choosing the spike-and-slab prior rather than the scaled Gaussian mixture is because imposing correlation on the support of the sparse signal is more accurate and justifiable than imposing it on the amplitude of the sparse signal. As discussed in Section II, the sparsity pattern of the signal is determined by $\mathbf{W}$ in the spike-and-slab model in (7), where the parameter $\mathbf{q}$ controls the probability of $\mathbf{W}$ being non-zero.

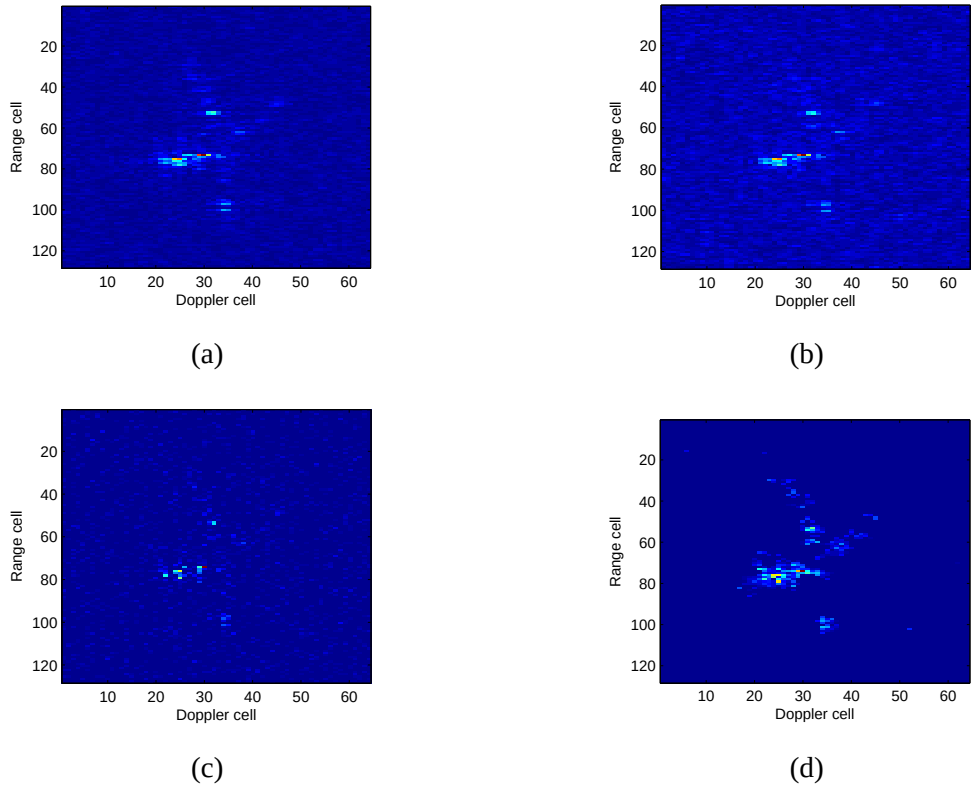

Fig. 4. Radar Images obtained by (a) range-Doppler algorithm (RDA) under noise; (b) RDA using one half of the measurements; (c) an ℓ_1 regularized method using one half of the measurements; (d) a first-order continuity method [39] using one half of the measurements

Therefore, a straightforward way to impose a continuity prior on the signal can be carried out directly on $\mathbf{W}$. However, this treatment deviates from the original intention to perform a flexible statistical modeling step. For this particular reason, it is suggested in [39], [45] to encode the first-order structural information on $\mathbf{q}$ in an intermediate way rather than straightforwardly on $\mathbf{W}$. The key modification is to replace the single Beta prior for parameter $\mathbf{q}$ by a set of Beta priors that consist of three different sets of parameters, $\{e_k, f_k\}_{k=0,1,2}$, so as to capture strongly independent, strongly continuous and non-informative priors, respectively.

More concretely, the proposed sparsity patterns that encourage continuity as well as preserve sparsity in [39], [45] are summarized as follows:

- Strong rejection: if the first-order neighborhoods of $\mathbf{X}_{mn}$ are all zero, it would be highly possible that $\mathbf{X}_{mn}$ is also zero due to the continuity of the target scene. The prior $\mathrm{Beta}(e_0, f_0)$ with $e_0 < f_0$ is utilized to make the probability q_{mn} of $W_{mn} = 0$ being large. This means

that the absence of a first-order neighborhood implies the investigated scatterer being zero with a high probability. It is noted that this rejection pattern can eliminate the undesired isolated speckles or artifacts in the radar image.

- Strong acceptance: if any of the continuity patterns for $\mathbf{X}_{mn}$ in Fig. 3 is observed, the prior that a non-zero valued a_{mn} arises with a high possibility should be imposed. This step imposes continuity of the target image. In this case, the prior $\mathrm{Beta}(e_1, f_1)$ with $e_1 > f_1$ enforces the probability q_{mn} of $W_{mn} = 1$ to be large and, thus the scatterer under test can be accepted. This implies that the occurrence of any pattern in Fig. 3 leads to one that is non-zero with a high probability. This pattern enforces first-order correlation of the scattering coefficient and, therefore continuity of the target.
- Weak rejection: apart from the scenario of strong rejection and acceptance patterns, a non-informative prior is imposed on any other neighborhood patterns for X_{mn}. The prior $\mathrm{Beta}(e_2, f_2)$ with $e_2 = f_2$ is used to impose a non-informative prior on q_{mn}. This appropriately allows the model to be effective in imposing the prior whenever necessary and to remain non-informative whenever no strong rejection or acceptance patterns appear.

By adaptively selecting from different Beta hyper-priors, the statistical model can either encourage continuity or independence, apart from mere sparsity. In this manner, the structured information can be flexibly incorporated to obtain concentrated imagery results. A key component in incorporating the prior is that it is imposed on the parameter $\mathbf{q}$ rather than directly on $\mathbf{W}$. The underlying motivation for this formulation is that it is more flexible to impose a probabilistic belief rather than a rigid support $\mathbf{W}$.

In Fig. 4, the real Yak-42 data are used to test different algorithms, where the radar image obtained with all measurements is shown in Fig. 4 (a) for reference purposes. In general, there are two issues to be considered in the evaluation of radar images. Firstly, how well the target is concentrated, i.e., more true scatterers preserved in the target region and less artifact recovered outside the target region. Secondly, how well the radar image is focused, i.e., lower side-lobed and noise. As shown in Fig. 4 (b), the radar image obtained by the range-Doppler algorithm (RDA) method is highly corrupted by noise. Although the ℓ_1 regularized approach can achieve better performance than RDA by exploiting sparsity, the obtained target image is not well concentrated and artifacts around the target are not removed, as shown in Fig. 4 (c). Notably, the method in [39] that exploits first-order continuity patterns performs best, as shown in Fig. 4 (d). More specifically, the first-order continuity method in [39] outperforms the ℓ_1 regularized approach in

terms of radar target concentration and artifact removal, and is even clearer than the reference image, obtained with all measurements as shown in Fig. 4 (a).

B. Higher-Order Correlations

It can be observed that a real radar image generally exhibits higher-order correlations than simply horizontal or vertical correlations, exploited in [39], [45]. This motivates an extension of the first-order method. To formulate a generalized framework, a more sophisticated model is developed in [46] that captures higher-order correlations based on Markov Random Fields (MRF). The MRF model is widely used in image processing for imposing structural constraints on the image. This work presents a unified framework of incorporating more complex structural information in the target scene, as compared with the simple first-order approach. In Fig. 5, the construction of the MRF model is presented, where a second-order neighborhood system is employed. This model allows continuity from four directions, i.e., horizontally, vertically, $3\pi/4$ and $1\pi/4$ diagonally. The authors argue that adopting second-order MRF enables the capture of correlations of the investigated scatterer with its nearest eight neighbors, as shown in Fig. 5.

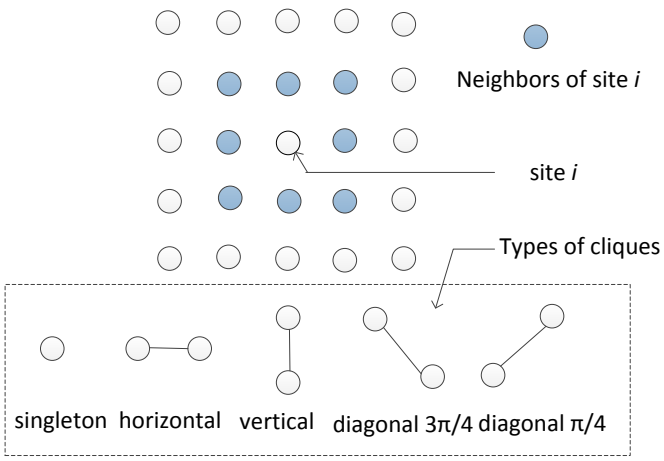

Fig. 5. An example of a second-order MRF model, where eight neighbors are considered.

To impose continuity of the target scene, the authors in [46] consider a more general structured sparse prior as compared with that in [39], [45]. A more complicated continuity prior has been proposed by modifying the spike-and-slab modeling to better preserve the weak scatterers. Moreover, the hyper-parameter selection in [39], [45] is avoided by adopting an MRF prior since all the parameters can be automatically inferred. This is a very desirable feature for

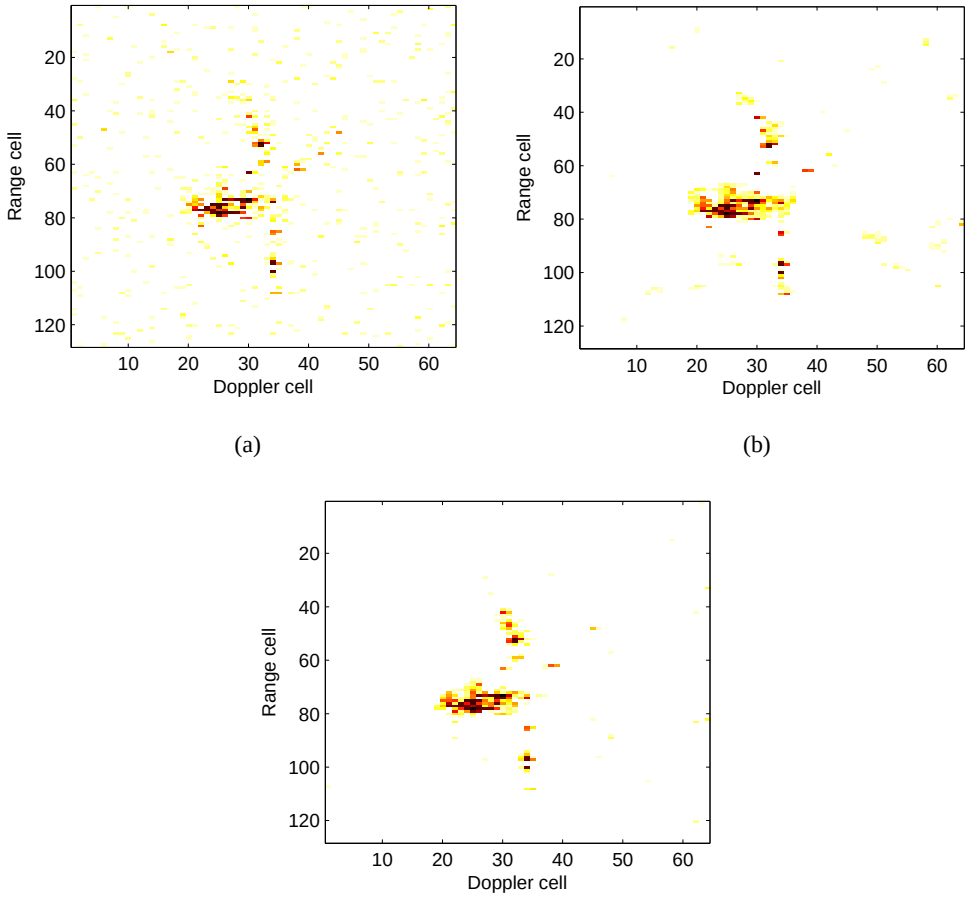

Fig. 6. Comparison of radar images with a quarter of the full measurements and SNR=5dB obtained by (a) an ℓ_1 regularized method; (b) a method in [39] using a quarter of the measurements; (c) a method in [46].

statistical inference. Based on this model, the authors employ a variational Bayesian expectation maximization (VBEM) method for inference, where an improved rate of convergence can be obtained, as compared with the method in [39], [45]. As commented earlier, the VB based method generally requires less computational complexity than the MCMC based one.

In Fig. 6, we can observe that both the first-order continuity method in [39], [45] and the second-order continuity method in [46] produce much enhanced radar images in the sense of less noise and better concentrated target region, compared with the ℓ_1 regularized method. As shown in Fig. 6 (b) and (c), the second-order continuity based method performs much better than the first-order one in terms of removing the undesirable isolated artifacts and preserving weak scatterers outside and within the target region, respectively. More importantly, the computational

time of the second-order continuity method is much less than that of the first-order one [46].

All the above demonstrates how, by incorporating structural priors in addition to sparsity, a statistical framework provides superior performance as compared to a merely sparsity based framework. The major advantage of these approaches is that they can statistically impose the structured sparsity on the signal in a rather flexible way, which allows the algorithm to adapt the structured sparse estimation in a data-driven manner. More specifically, in all the introduced models, the structural information is not directly imposed on the sparse signal itself, but on the probability distribution that determines the sparsity profile.

V. Statistical Sparsity Based Autofocus Techniques in Radar Imagery

The CS based radar imagery techniques discussed in the previous sections generally depend on the premise that pre-processing procedures, such as range cell migration correction and phase adjustment have been perfectly conducted. Unfortunately, this is not a valid assumption from a practical viewpoint, since the motion of the target cannot be precisely compensated in coarse pre-processing stages. If these errors are not properly corrected or compensated for before carrying out any CS based algorithms, the reconstructed radar image is not well concentrated.

Recently, phase error correction has been considered by utilizing a sparse recovery technique, where alternating ℓ_1 regularized approaches [14] are proposed to obtain more focused images. In these methods, the sparse scattering coefficient and the phase errors are iteratively estimated to induce sparsity, and obtain a focused radar image. Although these methods have demonstrated remarkable improvements over the conventional autofocus techniques, these regularization based methods might converge to a shallow local minimum during the iterative procedure. The alternate optimization between the sparse scattering coefficient and the phase error would inevitably result in error propagation [47]. More concretely, the alternate optimization scheme would introduce errors since the estimation accuracy of one parameter substantially influences that of another parameter. This issue is particularly severe with under-sampled data and in low SNR conditions. To appropriately overcome the above limitations, high-resolution imagery and phase error correction have been formulated in a statistical sparsity based model [47]–[49]. In this formulation, probabilistic models are imposed on the signal to encode sparsity in a statistical way. Subsequent parameter estimation is conducted within a sparse Bayesian learning framework [19], [47].

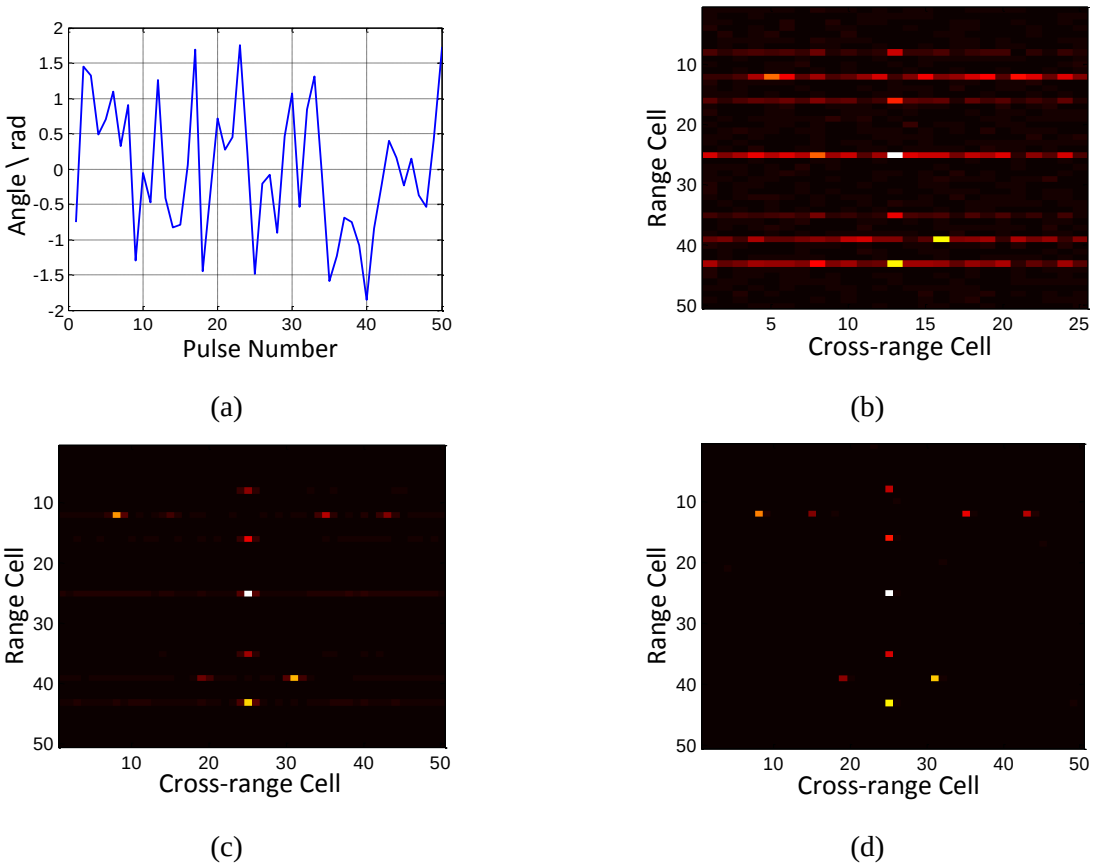

Fig. 7. Radar imaging with 25 pulses (50% of the full measurements and SNR=20dB). (a) random phase error, (b) RDA method, (c) a method in [47] (without uncertainty information) ($\mathrm{NMSE}_{\mathbf{X}} = -4.0294\mathrm{dB}$, $\mathrm{MSE}_{\varphi} = 0.4019$), (d) a method in [47] (with uncertainty information) ($\mathrm{NMSE}_{\mathbf{X}} = -12.4399\mathrm{dB}$, $\mathrm{MSE}_{\varphi} = 0.0059$).

A. Statistical Sparsity Based Autofocus

Assuming that the phase error in radar imagery exhibits range invariance [15], the mathematical model can be stated as

$$\mathbf{Y} = \mathbf{E}\mathbf{\Phi}_1\mathbf{A}\mathbf{X} + \mathbf{N}. \tag{15}$$

where $\mathbf{E} = \mathrm{diag}(e^{j\varphi_1}, ..., e^{j\varphi_P})$ denotes the phase error matrix, which is a diagonal matrix representing cross-range variant phase errors. In [47], the authors utilize the scale Gaussian mixture model to impose sparsity on $\mathbf{X}$. The estimation of $\mathbf{X}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\lambda}$ is obtained individually since they are task-specific parameters, while estimation of α_0 and $\mathbf{E}$ is performed in a global manner due to the task-invariant property.

According to the graphical model [47], the parameters, $\mathbf{X}$, $\boldsymbol{\alpha}$, $\boldsymbol{\lambda}$ and α_0, can be conveniently estimated, which is similar to that in the scaled Gaussian mixture model introduced previously.

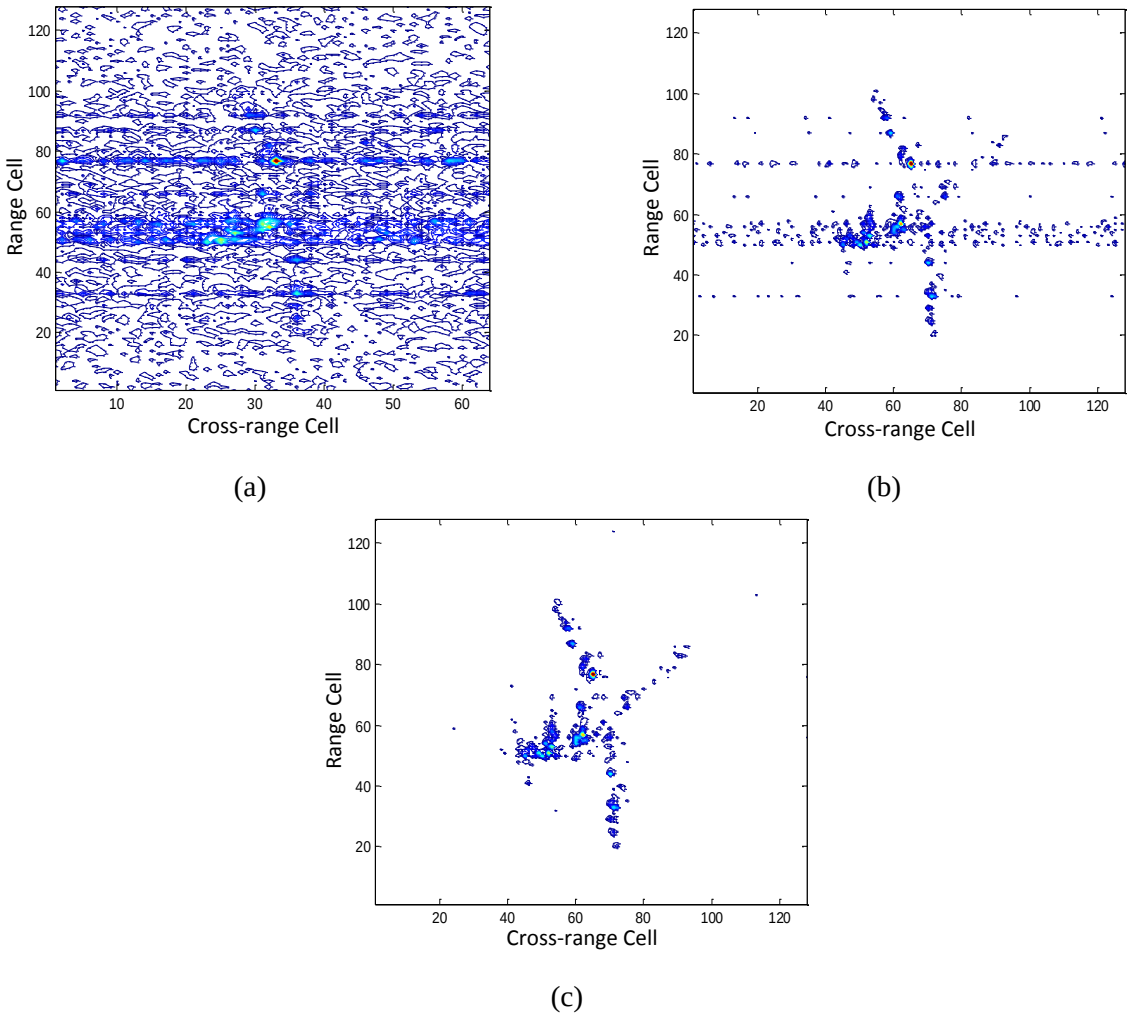

Fig. 8. Radar imagery results with one half of the measurements obtained by (a) RDA, (b) a method in [14], and (c) a method in [47].

The most straightforward way to obtain an estimate of the phase error $\mathbf{E}$ is to maximize the expected log-likelihood function as,

$$\hat{\mathbf{E}} = \arg\min_{\mathbf{E}} \langle -\ln p(\mathbf{Y}, \mathbf{X}, \boldsymbol{\alpha}, \boldsymbol{\lambda}; \mathbf{E}) \rangle_{q(\mathbf{X})q(\alpha)q(\lambda)}. \tag{16}$$

The above problem is a strictly convex optimization with a closed-form solution. By solving the optimization problem in (16), the updating formula can be obtained [49]. As a matter of fact, this updating rule for phase error is rather similar to that of the regularized approach in [14], [15], because the updating formula only uses the first-order moment of $\mathbf{X}$ to estimate $\mathbf{E}$ and the obtained covariance matrix $\boldsymbol{\Sigma}$ of $\mathbf{X}$ does not appear in this updating rule. In other words, this

formulation deviates from the original intention of utilizing higher-order statistical information in the first place.

In order to properly utilize the uncertainty information, the work in [47] proposes to incorporate the obtained covariance matrix $\mathbf{\Sigma}$ in the estimation of phase errors to obtain enhanced accuracy. Towards this end, the phase error is deliberately modeled as a complex parameter $a_i + jb_i$ rather than explicitly modeling the phase error as $e^{j\varphi_i}$. By introducing this complex parameter instead of the angle parameter φ_i, we will see that the uncertainty information can be naturally incorporated in the algorithm to achieve enhanced estimation accuracy of $\mathbf{E}$ in each iteration. In the derived updating formula in [47], it can be seen that $\mathbf{\Sigma}$, which contains uncertainty information, can be incorporated into the estimation of the phase error parameter $\mathbf{E}$. It is demonstrated in [47] that by replacing the true phase error parameters with complex-valued error parameters, the resulting scheme could utilize the estimation uncertainty information and obtain a performance gain as compared with regularized sparsity based autofocus techniques.

In Fig. 7, an illustrative example is presented to evaluate the performance of the updating rule without and with high-order uncertainty information. In this simulation, a total of 11 scatterers are present in the imaging scene. In Fig. 7(a), the random phase error is shown. In Fig. 7(c) and (d), it can be seen that both updating rules lead to a more focused image compared with the RDA method shown in Fig. 7(b). In particular, the updating rule without utilizing the uncertainty information leads to a less focused image, where undesirable side-lobe effects exist for almost all the scatterers on the imaging scene. In contrast, the image obtained by the updating rule utilizing the uncertainty information is well focused with substantially suppressed side-lobe effects. Quantitative evaluation also demonstrates that the radar image in Fig. 7 (d) provides a lower $\mathrm{NMSE}_{\mathbf{X}}$ as well as MSE_{φ} than those obtained in Fig. 7 (c) due to the inherent ability to utilize the uncertainty information of estimation of $\mathbf{X}$. This validates the motivation of utilizing the uncertainty information to achieve higher estimation accuracy, and therefore better recovered radar image. This work is also validated using the Yak-42 data. As observed from Fig. 8(a), directly applying the RDA method can barely lead to a concentrated image. In Fig. 8(b), the image obtained by ℓ_1 presents a reasonable profile of the airplane. However, it is still blurred and some of the true scatterers are not recovered correctly. In contrast, the method in [47] obtains a better concentrated image and removes most of the undesirable artifacts as seen from Fig. 8(c). With these comparisons, we conclude that the proper utilization of uncertainty information substantially enhances the performance in autofocus applications.

The statistical treatment of the sparsity based algorithm can properly utilize uncertainty information during iterations to improve the estimation accuracy. In fact, we have demonstrated how this uncertainty information can be properly used throughout this particular application. Compared with the regularized approach, the statistical sparsity based algorithms do not require the time-consuming parameter tuning for improved performance as those in the ℓ_1 regularized alternating methods.

B. Autofocus Meets Structured Sparsity

The basic idea in [47] is to iteratively estimate the sparse scatterer coefficients and the phase error to jointly induce sparsity. However, The objective of radar imaging is to obtain the most concentrated radar image rather than the sparsest one. Therefore, a merely sparsity-inducing scheme may result in undesirable image results, because it only considers sparsity as the performance measure. More concretely, the weak scatterers cannot be well-preserved and background noise cannot be desirably reduced with a simple sparsity constraint. A possible solution to obtain a more concentrated radar image rather than a mere sparse one is to exploit structured sparsity. In [49], the sparse Bayesian model is sophisticatedly modified to exploit structural sparsity. More specifically, the spatial consistency along range cells is exploited. Therefore, the framework can simultaneously cope with structured sparse signal recovery and phase error correction in an integrated manner. The focused high-resolution radar image can be obtained by iteratively estimating the sparse scatterer coefficients and phase errors to jointly obtain a structured sparse solution.

Due to the utilization of the structured sparse constraint, the proposed method preserves the target region and alleviates the over-shrinkage problem, compared to the previously presented sparsity-driven auto-focus approaches. The superior performance of the structured sparsity based technique is shown in Fig. 9. Compared with other approaches, the structured sparsity based autofocus method achieves a better concentrated image with more coefficients recovered in the target region with different under-sampling ratios.

VI. Statistical Sparsity based SAR Ground Moving Target Imaging

Imaging ground moving targets in synthetic aperture radar has become increasingly important. Conventionally, in imaging a potentially moving target, hypotheses of the target motion are constructed to match the signal by a filter bank [50]. In the scenario of closely located targets, however, their responses cannot be well distinguished from each other. Recent advances of sparsity based SAR ground moving target imaging (GMTIm) [5], [51] suggest that sparsity can be properly

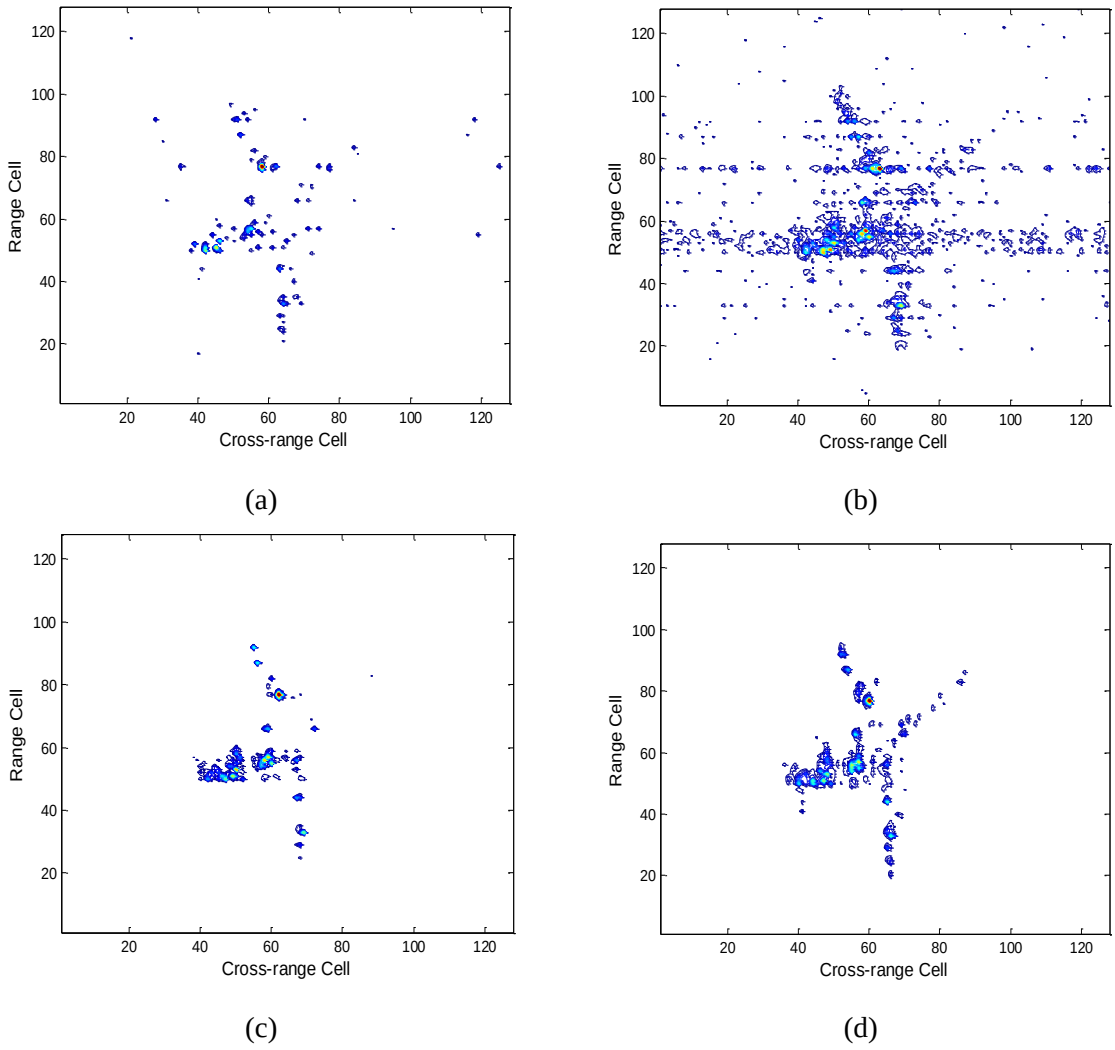

Fig. 9. Radar imagery results with SNR= 5dB using (a) ℓ_1 regularized method, (b) TV+ℓ_1 regularized method, (c) statistical sparsity based method in [47], and (d) statistical structured sparsity based method in [49].

exploited to enable multi-target processing and higher accuracy. This application is rather different from the previously introduced ones, since the received radar echoes can no longer be simply modeled as a sum of harmonics, but rather as multi-component LFM signals with unknown chirp rates. Therefore, the key challenge in SAR GMTIm is to properly formulate a mathematical model that allows a sparse representation of the images for moving targets. In [51], the signal model is constructed as a sparse linear model, where an over-complete dictionary is constructed by using a discretized velocity grid. Although empirical results demonstrate the success of the method, its performance is inhibited by the discretization errors in the dictionary. In this section,

we will briefly review two recent works based on statistical sparsity from different angles.

In [52], a statistical framework is formulated to obtain the moving target image, which could avoid the construction of a large over-complete dictionary. In particular, this work considers a K channel synthetic aperture radar system with F passes, collecting data from P azimuth angles and Q range cells. The complex-valued raw SAR image is decomposed as follows [52],

$$\mathbf{Y}_{p,f} = \mathbf{E}_{p,f} \circ (\mathbf{L}_{p,f} + \mathbf{S}_{p,f} + \mathbf{N}_{p,f}), \quad p = 1, ..., P \quad \text{and} \quad f = 1, ..., F \tag{17}$$

where $\mathbf{Y}_{p,f} \in \mathbb{C}^{Q \times K}$ denotes the raw SAR image at azimuth p and pass f, $\mathbf{E}_{p,f} \in \mathbb{C}^{Q \times K}$ is the corresponding spatial-temporal calibration error, $\mathbf{L}_{p,f} \in \mathbb{C}^{Q \times K}$ represents background clutter, $\mathbf{S}_{p,f} \in \mathbb{C}^{Q \times K}$ represents the moving target and $\mathbf{N}_{p,f} \in \mathbb{C}^{Q \times K}$ models noise. Since the number of parameters to be estimated is much larger than the number of observations $\mathbf{Y}$, proper priors must be selected for each of these parameters. The interested reader is referred to [52] for more details. Here, we only highlight the key statistical models in this formulation.

- The clutter $\mathbf{L}_{p,f}$ is decomposed into a sum of a pass-invariant background term $\mathbf{B}_p$ and a pass-specific speckle term $\mathbf{X}_{p,f}$. Assuming that the background clutter $\mathbf{B}_p$ can be represented by one of the several classes such as road or buildings, a complex Gaussian prior is used for $\mathbf{B}_p$ with a set of unknown covariance matrices that account for different classes, where each covariance matrix is modeled as an inverse Wishart distribution. Similarly, the pass-specific speckle term $\mathbf{X}_{p,f}$ is modeled probabilistically.
- Since the moving target is assumed to be sparse in the raw SAR image domain, a modified spike-and-slab model models sparsity as well as moving target signatures. More specifically, the sparsity is modeled by a Bernoulli-Beta distribution and the moving target signature is modeled as a complex Gaussian-Inverse-Wishart distribution. The rationale of this modeling is to allow a rather tractable inference.
- An additional constraint can be imposed on the hyper-parameters of the sparse moving target to encourage a smooth trajectory. It is noted that this smooth prior is constructed by modifying the Beta distribution instead of the support parameter directly, where this manipulation can be rather flexible in encoding the prior information in a probabilistic sense.

Since the work in [52] utilizes the decomposition of a raw SAR image, the construction of the dictionary as in [51] can be avoided. Since this method is formulated in a statistical framework, the algorithm could utilize the uncertainty information obtained in one parameter to enhance the

estimation of other parameters subsequently. These desirable properties of the statistical sparsity based method have led to substantial improvements over conventional methods.

Another approach for the SAR GMTIm problem is based on formulating a parametric model, where statistical sparsity is enforced [53]. In this work, the clutter is assumed to have been suppressed by off-the-shelf methods, and an initial representation of the received signal is firstly carried out by utilizing LV's distribution (LVD) [54], which is a novel TFR for representing LFM signals. However, the resolution of the LVD representation is constrained by the CPI of the target and its discretized grid [54]. It should be noted that the limited accuracy of the LVD representation may cause an unfocused target response, and thus a degraded target image. To deal with this challenge within a statistical sparsity based framework, dynamical refinement is suggested for an accurate estimation of the chirp rate in [55]. In particular, this dynamic refinement iteratively refines the initialized γ_i by LVD and the sparse target coefficient. In this way, the estimation accuracy can be improved in a statistical sparsity framework, and therefore, a concentrated moving target image can be obtained. Considering P azimuth and Q range cells, the clutter suppressed signal model can be formulated as [55],

$$\mathbf{Y} = \mathbf{E}\boldsymbol{\mathcal{A}}(\boldsymbol{\gamma}_1, ..., \boldsymbol{\gamma}_K)\mathbf{X} + \mathbf{N} \tag{18}$$

where $\mathbf{Y} \in \mathbb{C}^{P\times Q}$ is the clutter suppressed data, $\mathbf{E} \in \mathbb{C}^{P\times P}$ represents the unknown phase errors, $\mathbf{X} \in \mathbb{C}^{KN\times Q}$ models the sparse moving target to be estimated, $\gamma_i, \quad i = 1, ..., K$, is a set of parameters in the dictionary to be estimated and K is the number of moving targets. As described earlier, the dictionary $\boldsymbol{\mathcal{A}}(\boldsymbol{\gamma}_1, ..., \boldsymbol{\gamma}_K) \in \mathbb{C}^{P\times KN}$ is an over-complete one. It is constructed by concatenating K sub-dictionary, where each sub-dictionary is constructed by an LFM matrix with chirp rate γ_i. In this work [55], a scaled Gaussian mixture distribution is used to model sparsity. Similar to the work covered in Section V, statistical information is utilized to estimate the error parameter $\mathbf{E}$ and the chirp rate γ_i, where the error propagation problem during iteration is reduced [47], [55].

In Fig. 10, the canonical Gotcha data set is used for validation, and an example of the Durango target image is given to demonstrate the performance. Due to the movement of the target, the original image is substantially blurred, as observed from 10(a). After representing the received signal by LVD, the ℓ_1-norm regularization method and the conventional sparse Bayesian method are applied to obtain the moving target images as shown in Fig. 10 (b) and (c), respectively. It can be seen that the ℓ_1 regularized method and the sparse Bayesian method cannot properly focus the target image due to the representation error in LVD. In contrast, the statistical sparsity based

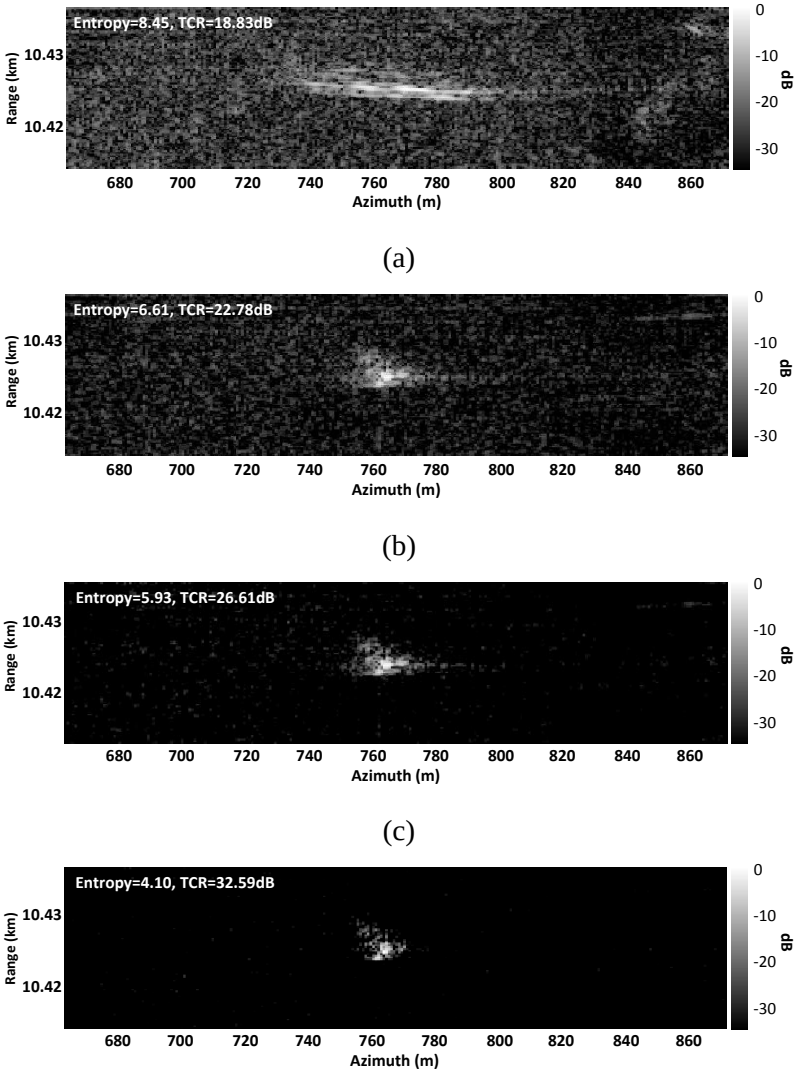

Fig. 10. Durango target image. (a) Original image; (b) image obtained by ℓ_1-norm regularization; (c) image obtained by conventional statistical sparsity based method; (d) image obtained by parametric and dynamic statistical sparsity based method in [55]

method with refinement leads to the best imaging performance in terms of better concentration and desirable noise suppression, as shown in Fig. 10(d). In particular, the target image is focused within a $5\mathrm{m} \times 5\mathrm{m}$ area that is in accordance with the Durango truth with a size of $5\mathrm{m} \times 2\mathrm{m}$.

VII. Summary and Future Directions

A. Summary

Sparsity based techniques have been reviewed from a statistical perspective, along with their recent advances in radar imagery. In various applications, it has been shown that improved performance can be obtained by adequately utilizing a statistical sparse model. These improvements obtained in the reviewed applications were largely dependent on the following core ingredients.

1) Probabilistic modeling by incorporating flexible priors in the signal, is one of the most remarkable advantages over deterministic approaches. The advantageous characteristics of the statistical framework are that it is rather flexible. In this way, the formulation could model a particular structure in a probabilistic way and also allows for fitting with the likelihood.
2) The utilization of uncertainty information during parameter estimation is important for a performance gain. Particularly, in conventional approaches, the error estimated in one stage can lead to a degraded performance in the subsequent stages. In the statistical framework, the signal estimation is conducted in a statistical manner, where the obtained statistics indicate the uncertainty in the signal estimation. Therefore, the estimation could be more accurate.

By properly manipulating the statistical sparsity models, a performance gain could be obtained.

B. Future Directions

Since the statistical sparsity based methods are very attractive, it would be very interesting to investigate the following problems in the future.

- *Computational complexity.* The statistical sparsity based methods operate in an iterative manner, where the number of iterations and the computational cost in each iteration determine the computational cost. Compared to the conventional Fourier based approach for radar imaging, the computational complexity is much higher. It is therefore imperative to develop fast algorithms, which could decrease the computational complexity or obtain fast convergence. The fast algorithms would be particularly useful for many radar applications requiring real-time processing.
- *Motion compensation errors.* In high-resolution radar imaging problem, a large coherent processing interval (CPI) is required. Then, the target movement becomes a problem as the radar line-of-sight dramatically changes. In such a scenario, even after carrying out coarse

motion compensation, range cell migration (RCM) and phase error would still be present in the radar echoes. Then, the dictionary allowing sparse representation would become more complicated, where the proposed imaging algorithm should also be able to correct RCM and phase errors. The main challenge is to properly obtain the approximated solution in the presence of a more complicated model. Towards this end, it would be particularly suitable to exploit statistical sparsity to limit error propagation. One possible way of coping with this challenge is to encode priors on the error parameters to properly regularize the solution space.

- *Temporal correlation in SAR GMTIm.* Conventionally, most SAR GMTIm algorithms focus on image formation of the moving target at one particular time instant. However, it is important to also monitor the movement of the moving target. Since the targets motion and imaging background are time-varying, simply generating one single frame image cannot provide time varying characteristics of the moving target. Therefore, it is necessary to develop temporal SAR GMTIm based on the statistical sparsity based framework, which is promising research direction in SAR-GMTIm technology. In fact, Sandia Laboratory has successfully realized Video-SAR GMTIm, where the processed results have been released on their official website. In particular, the temporal SAR GMTIm is a good candidate for applications in complicated urban scenes, where improved performance is valuable. The statistical sparsity based framework for moving target imaging in urban environments could be formulated to include the temporal smoothness constraint during the radar passes and to cope with a complicated background.
- *Improved classification performance.* An important objective of radar imagery is to classify different types of targets automatically and accurately. One should capture more structural features during target imaging by utilizing the training information obtained from the recognition stage, which will in turn greatly benefit automatic target recognition (ATR). More precisely, radar imagery should be discriminative enough for target recognition purposes. One promising future work is to incorporate appropriate priors in a statistical framework to perform discriminative radar imagery.

In summary, statistical sparsity-driven techniques have been shown to be very promising for radar imagery due to their flexibility and good statistical properties. It is expected that these applications will immensely benefit from the more recent theoretical advances in this area.

ACKNOWLEDGMENT

This work is partially supported by project fund MOE2014-T2-1-079 and RG103/14, Singapore.

REFERENCES

[1] J.C. Curlander and R.N. McDonough, *Synthetic aperture Radar*, John Wiley & Sons, 1991.

[2] E.J. Candes and M.B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[3] L.C. Potter, E. Ertin, J.T. Parker, and M. Cetin, "Sparsity and compressed sensing in radar imaging," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1006–1020, Jun. 2010.

[4] J.H. Ender, "On compressive sensing applied to radar," *Signal Processing*, vol. 90, no. 5, pp. 1402–1414, 2010.

[5] M. Cetin, I. Stojanovic, N.O. Onhon, K.R. Varshney, S. Samadi, W.C. Karl, and A.S. Willsky, "Sparsity-driven synthetic aperture radar imaging: Reconstruction, autofocusing, moving targets, and compressed sensing," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 27–40, Jul. 2014.

[6] M. Leigsnering, M. Amin, F. Ahmad, and A. M. Zoubir, "Multipath exploitation and suppression for SAR imaging of building interiors: An overview of recent advances," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 110–119, Jul. 2014.

[7] X.X. Zhu and R. Bamler, "Superresolving SAR tomography for multidimensional imaging of urban areas: Compressive sensing-based tomoSAR inversion," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 51–58, Jul. 2014.

[8] R.G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, 2007.

[9] T.A. Mariví, L. Paco, and J.J. Mallorquí, "A novel strategy for radar imaging based on compressive sensing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 12, pp. 4285–4295, 2010.

[10] G. Li, H. Zhang, X. Wang, and X. Xia, "ISAR 2-D imaging of uniformly rotating targets via matching pursuit," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 2, pp. 1838–1846, APRIL 2012.

[11] S.S. Chen, D.L. Donoho, and M.A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129–159, Jan. 2001.

[12] X. Zhu and R. Bamler, "Tomographic SAR inversion by-norm regularizationthe compressive sensing approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 10, pp. 3839–3846, 2010.

[13] V.M. Patel, G.R. Easley, D.M. Healy Jr, and R. Chellappa, "Compressed synthetic aperture radar," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 244–254, 2010.

[14] N. Onhon and M. Cetin, "A sparsity-driven approach for joint SAR imaging and phase error correction," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2075–2088, 2012.

[15] X. Du, C. Duan, and W. Hu, "Sparse representation based autofocusing technique for ISAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 3, pp. 1826–1835, 2013.

[16] I. Stojanovic and W.C. Karl, "Imaging of moving targets with multi-static SAR using an overcomplete dictionary," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 1, pp. 164–176, Feb 2010.

[17] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2346–2356, Jun. 2008.

[18] D.P. Wipf and B.D. Rao, "Sparse bayesian learning for basis selection," *IEEE Transactions on Signal Processing*, vol. 52, no. 8, pp. 2153–2164, Aug. 2004.

[19] L. Zhao, G. Bi, L. Wang, and H. Zhang, "An improved auto-calibration algorithm based on sparse bayesian learning framework," *IEEE Signal Processing Letters*, vol. 20, no. 9, pp. 889–892, 2013.

[20] L. Wang, L. Zhao, G. Bi, and C. Wan, "Hierarchical sparse signal recovery by variational bayesian inference," *IEEE Signal Processing Letters*, vol. 21, no. 1, pp. 110–113, Jan. 2014.

[21] L. Zhao, L. Wang, G. Bi, L. Zhang, and H. Zhang, "Robust frequency-hopping spectrum estimation based on sparse bayesian method," *IEEE Transactions on Wireless Communications*, vol. 14, no. 2, pp. 781–793, Feb 2015.

[22] L. Zhang, Z. Qiao, M. Xing, Y. Li, and Z. Bao, "High-resolution ISAR imaging with sparse stepped-frequency waveforms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4630–4651, Nov. 2011.

[23] H. Wang, Y. Quan, M. Xing, and S. Zhang, "ISAR imaging via sparse probing frequencies," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 3, pp. 451–455, 2011.

[24] J. Zhang, D. Zhu, and G. Zhang, "Adaptive compressed sensing radar oriented toward cognitive detection in dynamic sparse target scene," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1718–1729, 2012.

[25] Y. Huang, X. Wang, X. Li, and B. Moran, "Inverse synthetic aperture radar imaging using frame theory," *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 5191–5200, Oct. 2012.

[26] W. Rao, G. Li, X. Wang, and X. Xia, "Parametric sparse representation method for ISAR imaging of rotating targets," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 50, no. 2, pp. 910–919, Apr. 2014.

[27] C.M. Bishop, *Pattern recognition and machine learning*, vol. 1, springer New York, 2006.

[28] S.D. Babacan, R. Molina, and A.K. Katsaggelos, "Bayesian compressive sensing using laplace priors," *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 53–63, 2010.

[29] M. Figueiredo, "Adaptive sparseness for supervised learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1150–1159, 2003.

[30] N.L. Pedersen, C.N. Manchón, M. Badiu, D. Shutin, and B.H. Fleury, "Sparse estimation using bayesian hierarchical prior modeling for real and complex linear models," *Signal Processing*, vol. 115, pp. 94–109, 2015.

[31] N.G. Polson and J.G. Scott, "Shrink globally, act locally: sparse bayesian regularization and prediction," *Bayesian Statistics*, vol. 9, pp. 501–538, 2010.

[32] C.M. Carvalho, N.G. Polson, and J.G. Scott, "The horseshoe estimator for sparse signals," *Biometrika*, p. asq017, 2010.

[33] S.D. Babacan, R. Molina, and A.K. Katsaggelos, "Bayesian compressive sensing using laplace priors," *IEEE Transaction on Image Processing*, vol. 19, no. 1, pp. 53–63, Jan. 2010.

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

[35] D.G. Tzikas, A.C. Likas, and N.P. Galatsanos, "The variational approximation for bayesian inference," *IEEE Signal Processing Magazine*, vol. 25, no. 6, pp. 131–146, Nov. 2008.

[36] M. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Transactions on Signal Processing*, vol. 57, no. 6, pp. 2275–2284, 2009.

[37] H. Liu, B. Jiu, H. Liu, and Z. Bao, "Superresolution ISAR imaging based on sparse bayesian learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 8, pp. 5005–5013, Aug. 2014.

[38] L. Zhang, M. Xing, C. Qiu, J. Li, J. Sheng, Y. Li, and Z. Bao, "Resolution enhancement for inversed synthetic

aperture radar imaging under low snr via improved compressive sensing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 10, pp. 3824–3838, 2010.

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

[40] J. Xu, Y. Pi, and Z. Cao, "Bayesian compressive sensing in synthetic aperture radar imaging," *IET Radar, Sonar Navigation*, vol. 6, no. 1, pp. 2–8, Jan. 2012.

[41] B. Wang, S. Zhang, and W. Wang, "Bayesian inverse synthetic aperture radar imaging by exploiting sparse probing frequencies," *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 1698–1701, 2015.

[42] M. Luessi, S.D. Babacan, R. Molina, and A.K. Katsaggelos, "Bayesian simultaneous sparse approximation with smooth signals," *IEEE Transactions on Signal Processing*, vol. 61, no. 22, pp. 5716–5729, Nov 2013.

[43] D.P. Wipf, "Sparse estimation with structured dictionaries," in *Advances in Neural Information Processing Systems*, 2011, pp. 2016–2024.

[44] L. Wang, L. Zhao, G. Bi, and L. Zhang, "ISAR imaging by exploiting the continuity of target scene," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 6072–6076.

[45] Q. Wu, Y. Zhang, F. Ahmad, and M. Amin, "Compressive sensing based high-resolution polarimetric through-the-wall radar imaging exploiting target characteristics," *IEEE Antennas Wireless Propag. Lett.*, vol. 14, pp. 1043–1047, 2015.

[46] L. Wang, L. Zhao, G. Bi, and C. Wan, "Sparse representation-based ISAR imaging using markov random fields," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 8, pp. 3941–3953, 2015.

[47] L. Zhao, L. Wang, G. Bi, and L. Yang, "An autofocus technique for high-resolution inverse synthetic aperture radar imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 10, pp. 6392–6403, Oct 2014.

[48] Q. Wu, Y. Zhang, M. Amin, and F. Ahmad, "Autofocus bayesian compressive sensing for multipath exploitation in urban sensing," in *2015 IEEE International Conference on Digital Signal Processing (DSP)*, Jul. 2015, pp. 80–84.

[49] L. Zhao, L. Wang, G. Bi, and C. Wan, "Structured sparsity-driven autofocus algorithm for high-resolution radar imagery," *Signal Processing*, vol. 125, pp. 376 – 388, 2016.

[50] J.K. Jao and A. Yegulalp, "Multichannel synthetic aperture radar signatures and imaging of a moving target," *Inverse Problems*, vol. 29, no. 5, pp. 054009, 2013.

[51] I. Stojanovic and W.C. Karl, "Imaging of moving targets with multi-static SAR using an overcomplete dictionary," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 1, pp. 164–176, Feb. 2010.

[52] Gregory Newstadt, Edmund Zelnio, and Alfred Hero, "Moving target inference with bayesian models in SAR imagery," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 3, pp. 2004–2018, 2014.

[53] L. Yang, G. Bi, M. Xing, and L. Zhang, "Airborne SAR moving target signatures and imagery based on lvd," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 11, pp. 5958–5971, Nov 2015.

[54] X. Lv, G. Bi, C. Wan, and M. Xing, "Lv's distribution: Principle, implementation, properties, and performance," *IEEE Transactions on Signal Processing*, vol. 59, no. 8, pp. 3576–3591, 2011.

[55] L. Yang, L. Zhao, and G. Bi, "SAR ground moving target imaging algorithm based on parametric and dynamic sparse bayesian learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 4, pp. 2254–2267, Apr. 2016.

Lifan Zhao received his B.Sc degree in Electronic Engineering, Xidian University, China in 2010 and the Ph.D. degree from the School of Electrical and Electronic Engineering, Nanyang Technological University (NTU), Singapore in 2016. He is currently a research fellow in NTU. His research interests include statistical signal processing, compressed sensing, machine learning, Bayesian methods and their applications in radar and communications. He has published over 20 papers in these areas.

Lu Wang received the B.S. in electrical and electronic engineering from Xidian University in 2007 and M.Eng. degree in signal processing from National Key Lab of Radar Signal Processing (RSP), Xidian University, China, in 2010. She obtained her Ph.D. degree in 2014 from the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. She is currently an Assistant Professor in Northwestern Polytechnic University, Xi'an, China. Her major research interests include sparse signal processing, radar imaging (SAR/ISAR) and sonar signal processing.

Lei Yang received B.Sc degree in Electronic Engineering and Ph.D degree in Signal and Information Processing, both from Xidian University, in 2007 and 2012, respectively. He is currently working as a Research Fellow in the School of Electrical and Electronic Engineering, Nanyang Technological University (NTU), Singapore. His main research interests include high-resolution airborne and spaceborne synthetic aperture radar (SAR) imaging, data-driven motion error compensation (auto-focusing) algorithm for unmanned aerial vehicle (UAV) SAR, and real-time processing system for SAR.

Abdelhak M. Zoubir received his Dr.-Ing. degree from Ruhr-Universitt Bochum, Germany. He was with Queensland University of Technology, Australia, from 1992 to 1998. He then joined Curtin University of Technology, Australia, as a Professor of Telecommunications and was interim Head of the School of Electrical and Computer Engineering from 2001 to 2003. Since 2003, he has been a Professor of Signal Processing at Technische Universit´at Darmstadt, Germany. He is an IEEE Distinguished Lecturer (Class of 2010−2011), Past Chair of the IEEE Signal Processing Societys Signal Processing Theory and Methods Technical Committee, and Past Editor-in-Chief of the IEEE Signal Processing Magazine (2012−2014). His research interest lies in statistical methods for signal processing applied to telecommunications, radar, sonar, car engine monitoring, and biomedicine. He has published over 300 papers in these areas. He is a Fellow of the IEEE.

Guoan Bi received B.Sc degree in Radio communications, Dalian University of Technology, P. R. China, 1982, M.Sc degree in Telecommunication Systems and Ph.D degree in Electronics Systems, Essex University, UK, 1985 and 1988, respectively. Between 1988 and 1990, he was a research fellow at the University of Surrey, U.K.. Since 1991, he has been with the school of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. He is a Senior Member of IEEE and was the Chair of IEEE Signal Processing Society, Singapore Chapter. His research interests include DSP algorithms and hardware structures, and signal processing for various applications including sonar, radar, and communications.