Moving target inference with Bayesian models in SAR imagery

Gregory E. Newstadt, Edmund G. Zelnio, and Alfred O. Hero III, *Fellow, IEEE*
November 7, 2013

Abstract— **This work combines the physical, kinematic, and statistical properties of targets, clutter, and sensor calibration as manifested in multi-channel SAR imagery into a unified Bayesian structure that simultaneously estimates (a) clutter distributions and nuisance parameters and (b) target signatures required for detection/inference. A Monte Carlo estimate of the posterior distribution is provided that infers the model parameters directly from the data with little tuning of algorithm parameters. Performance is demonstrated on both measured/synthetic wide-area datasets.**

Index Terms—**synthetic aperture radar, moving target detection, low-rank, hierarchical Bayesian models**

I. INTRODUCTION

This work provides an algorithm for inference in multi-antenna and multi-pass synthetic aperture radar (SAR) imagery. Inference can mean many different things in this framework, including detection of moving targets, estimation of the underlying clutter distribution, estimation of the target radial velocity, and classification of pixels. To this end, the output of the proposed algorithm is an estimated posterior distribution over the variables in our model. This posterior distribution is estimated through Markov Chain Monte Carlo (MCMC) techniques. Subsequently, the inference tasks listed above are performed by appropriately using the posterior distribution. For example, detection can be done by thresholding the posterior probability that a target exists at any given location.

Recently, there has been great interest by Wright et al. [1], Lin et al. [2], Candes et al. [3] and Ding et al. [4] in the so-called robust principal component analysis (RPCA) problem that decomposes high-dimensional signals as

$$I = L + S + E, \tag{1}$$

where $I \in \mathbb{R}^{N \times M}$ is an observed high dimensional signal, $L \in \mathbb{R}^{N \times M}$ is a low-rank matrix with rank $r \ll NM$, $S \in \mathbb{R}^{N \times M}$ is a sparse component, and $E \in \mathbb{R}^{N \times M}$ is dense low-amplitude noise. This has clear connections to image processing where L can be used to model the stationary background and S represents sparse (moving) targets of interest.

Gregory Newstadt and Alfred Hero are with the Dept. of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor. Edmund Zelnio is with the Air Force Research Laboratory, Wright Patterson Air Force Base, OH 45433, USA. E-mail: ({newstage},{hero}@umich.edu and edmund.zelnio@wpafb.af.mil).

Moreover, since the background image does not change much from image to image, one would expect that L would be low-dimensional.

In [1]–[3], inference in this model is done by optimizing a cost function of the form

$$\arg \min_{L,S} \|L\|_* + \gamma \|S\|_1 + (2\mu)^{-1} \|I - L - S\|_F \tag{2}$$

where $\|\cdot\|_*$, $\|\cdot\|_1$, and $\|\cdot\|_F$ denote the matrix nuclear norm (sum of singular values), the l_1 norm, and the Frobenius norm, respectively. Sometimes, the last term is replaced by the constraint $I = L + S$ (i.e., the noiseless situation). In this optimization objective, the nuclear norm promotes a low-dimensional representation of L, the l_1 norm promotes a sparse S, and the Frobenius norm allows for small model mismatch in the presence of noise. One major drawback of these methods involves finding the algorithm parameters (e.g., tolerance levels or choices of γ, μ), which may depend on the given signal. Moreover, it has been demonstrated that the performance of these algorithms can depend strongly on these parameters.

Bayesian methods by Ding et al. [4] have been proposed that simultaneously learn the noise statistics and infer the low-rank and sparse components. Moreover, they show that their method can be generalized to richer models, e.g. Markov dependencies on the target locations. Additionally, these Bayesian inferences provide a characterization of the uncertainty of the outputs through a Markov Chain Monte Carlo (MCMC) estimate of the posterior distribution. The work by Ding et al. [4] is based on a general Bayesian framework [5] by Tipping for obtaining *sparse* solutions to regression and classification problems. Tipping's framework uses simple distributions (e.g., those belonging to the exponential class) that can be described by few parameters, known as hyperparameters. Moreover, Tipping considers a *hierarchy* where the hyperparameters themselves are assumed to have a known 'hyperprior' distribution. Often the prior and hyperprior distributions are chosen to be conjugate. Conjugate distributions have the property that the posterior and prior distributions have the same form, which makes inference/sampling from these distributions simple.. Tipping provides insight into choosing the hyperparameter distributions so as to be non-informative with respect to the prior. This latter property is important in making it possible to implement inference algorithms with few tuning parameters. Finally, Tipping provides a specialization to the 'relevance vector machine' (RVM), which can be thought of as a Bayesian version of the support vector machine. Wipf et

al. [6] provides an interpretation of the RVM as the application of a variational approximation to estimating the true posterior distribution. Wipf et al. explains the sparsity properties of the sparse Bayesian learning algorithms in a rigorous manner. Additionally, it also provides connections with other popular work in sparse problems, such as the FOCUSS and basis pursuit algorithms.

In this work, we develop a framework for inference in SAR imagery based on the algorithmic structure developed by Ding et al [4]. Whereas Ding developed an algorithm for inference in standard video imagery, this paper presents the following non-trivial extensions in order to incorporate SAR specific phenomena: (a) we consider complex-valued data rather than real-valued intensity images; (b) we model correlated noise sources based on physical knowledge of SAR phase history collection and image formation; (c) we relax the assumption of a low-rank background component by assuming that the background component lies in a low-dimensional subspace; and (d), we directly model SAR phenomenology by including terms for glints, speckle contributions, antenna gain patterns, and target kinematics. Moreover, we demonstrate the performance of the proposed algorithm on both simulated and measured datasets, showing competitive or better performance in a variety of situations.

Inference in SAR imagery is more complicated than that of standard electro-optical images. Examples of these complexities include

- SAR images have complex-valued rather than real-valued intensities, and the SAR phase information is of great importance for detection and estimation of target states. [7]–[9].
- SAR images are corrupted by spatiotemporally-varying antenna gain/phase patterns that often need to be estimated from homogeneous target-free data [10], [11].
- SAR images have spatially-varying clutter that can mask the target signature unless known a priori or properly estimated [12].
- SAR images contain motion-induced displacement and diffusion of the target response [7], [13].
- SAR images include multiple error sources due to radar collection and physical properties of the reflectors, such as angular scintillation (a.k.a. glints) [14] and speckle [15], [16].

Despite these complications, a great deal of structure exists in SAR images that can be leveraged to provide stronger SAR detection and tracking performance. This includes (a) using the coherence between multiple channels of an along-track radar in order to remove the stationary background (a.k.a, 'clutter'), (b) assuming that pixels within the image can be described by one (or a mixture) of a small number of object classes (e.g., buildings, vegetation, etc.), and (c) considering kinematic models for the target motion such as Markov smoothness priors. From this structure in SAR imagery, one might consider models that assume that the clutter lies in a low-dimensional subspace that can be estimated directly from the data. Indeed, recent work Borcea et al. [17] has shown that SAR signals can be represented as a composition of a low-rank component containing the clutter, a sparse component containing the target signatures, and additive noise.

In general, SAR images are formed by focusing the response of stationary objects to a single spatial location. Moving targets, however, will cause phase errors in the standard formation of SAR images that cause displacement and defocusing effects. Most methods designed to detect the target depend on either (a) exploiting the phase errors induced by the SAR image formation process for a single phase center system or (b) canceling the clutter background using a multiple phase center system. In this work, we provide a rich model that can combine (and exploit) both sources of information in order to improve on both methodologies.

Fienup [7] provides an analysis of SAR phase errors induced by translational motions for single-look SAR imagery. He shows that the major concerns are (a) azimuth translation errors from range-velocities, (b) azimuth smearing errors due to accelerations in range, and (c) azimuth smearing due to velocities in azimuth. Fienup also provides an algorithm for detecting targets by their induced phase errors. The algorithm is based on estimating the moving target's phase error, applying a focusing filter, and evaluating the sharpness ratio as a detection statistic. Jao [13] shows that given both the radar trajectory and the target trajectory, it is possible to geometrically determine the location of the target signature in a reconstructed SAR image. Although the radar trajectory is usually known with some accuracy, the target trajectory is unknown. On the other hand, if the target is assumed to have no accelerations, Jao provides an efficient FFT-based method for refocusing a SAR image over a selection of range velocities. Khwaja and Ma [18] provide a algorithm to exploit the sparsity of moving targets within SAR imagery; they propose a basis that is constructed from trajectories formed from all possible combinations of a set of velocities and positions. To combat the computational complexity of searching through this dictionary, the authors use compressed sensing techniques. Instead of searching over a dictionary of velocities, our work proposes to use a prior distribution on the target trajectory that can be provided a priori through road and traffic models or adaptively through observations of the scene over time.

The process of removing the stationary background in order to detect moving targets is known in the literature as 'change detection' or 'clutter suppression.' Generally, these methods require multiple views of the scene from either multiple receive channels or multiple passes of the radar. Moreover, they are based on the assumption that stationary targets will have nearly identical response when viewed at different times from the same viewpoint. In contrast, moving targets will exhibit a phase difference (namely the 'interferometric' phase) and thus can be detected as outliers. Another interpretation is that the stationary component (i.e., the clutter) lies in a low-dimensional subspace. Thus, the moving components can be detected by projecting the image into the null space of the clutter and appropriately thresholding the output.

There are several common methods for change detection with multiple looks, including displaced phase center array (DPCA) processing, space-time adaptive processing (STAP), and along-track interferometry (ATI). In DPCA, one thresholds

the difference image between two looks of the radar at slightly different times. However, the performance of DPCA (as well as STAP discussed shortly) suffers in the presence of heterogeneous clutter, such as strong returns from building edges. Whereas DPCA is a linear filter, ATI thresholds differences in the interferometric phase between the two images. Moreover, the phase responses of heterogeneous clutter tend to be insensitive to clutter amplitude and can thus be effectively removed by using ATI. Deming [8] analyzes both DPCA and ATI, showing that ATI performs well when canceling bright clutter, while DPCA performs well for canceling dim clutter. Moreover, he provides a joint algorithm that uses ATI to cancel strong discretes in the image and subsequently uses DPCA to remove small-amplitude clutter. In this paper, we compare our detection results to this joint DPCA-ATI algorithm and demonstrate competitive results, though our algorithm does not require setting thresholds on the phase/amplitude and also provides the probability of detection (as opposed to a binary output.)

STAP, like DPCA, is a linear filter for detecting moving targets from multiple looks that has been applied successfully to SAR [12]. However, STAP considers the case where $N > 2$ receive channels are available. The algorithm uses a single channel to estimate the stationary background, while the remaining $(N - 1)$ channels are used to estimate the moving component. Moreover, STAP is a matched filtering technique that adaptively chooses weights in order to project the data onto the null space of the clutter. Under ideal circumstances, STAP has the maximum signal-to-inference-and-noise-ratio (SINR) among linear filters [12]. However, STAP relies on estimating the complex-valued covariance matrix of the N-channel system, which in turn depends on the availability of homogeneous target-free secondary data. In this work, we simultaneously estimate the clutter covariance matrices as well as the target contributions. Thus, we demonstrate the capability to detect targets even in the presence of heterogeneous measurements.

Ranney and Soumekh [10], [11] develop methods for change detection from SAR images collected at two distinct times that are robust to errors in the SAR imaging process. They address error sources including inaccurate position information, varying antenna gains, and autofocus errors. They propose that the stationary components of multi-temporal SAR images can be related by a spatially-varying 2-dimensional filter. To make the change detection algorithm numerically practical, the authors propose that this filter can be well-approximated by a spatially invariant response within small subregions about any pixel in the image. This work adopts this model for the case where there are no registration errors. Under a Gaussian assumption for the measurement errors, it can be shown that the maximum likelihood estimate for the filter coefficients can be computed easily through simple least squares.

Ground Moving Target Indication (GMTI) methods involve the processing of SAR imagery to detect and estimate moving targets. Often clutter cancellation and change detection play a preprocessing role in these algorithms [19]–[22]. This work aims to combine properties of many of these algorithms into a unifying framework that simultaneously estimates the target signature and the nuisance parameters, such as clutter distributions and antenna calibrations.

The framework that is proposed in this paper contains a detailed statistical model of SAR imagery with many model variables that are jointly estimated through MCMC methods. It is natural to ask why such machinery is required for SAR inference when there are already (a) methods for MCMC inference in standard electro-optical imagery, and (b) simpler methods for SAR inference. As mentioned above, it is the authors' argument that there are sufficient complications with SAR imagery that make it difficult to use the former algorithms. Moreover, while there are indeed many simpler methods for SAR inference, these algorithms generally (a) are not robust to operating conditions and (b) do not provide estimates of their uncertainty. For example, thresholds in displaced phase center array (DPCA) processing will often need to be changed drastically depending on the radar conditions. Additionally, change detection algorithms such as DPCA provide a 0-1 output (i.e., either the target is detected or not). In contrast, the output in our framework is the probability of target existence.

Another key motivation for using a Bayesian formulation is the capability to readily impose additional structure when additional information sources are available. In particular, we consider the following two important information sources for localizing targets in SAR imagery: (a) multiple passes of the radar and (b) images formed from frame to frame (i.e., sequentially in time). In the former case, multiple passes of the radar are used to determine what the "normal" background should look like in order to detect anomalies in other passes. This is particularly useful in detecting *stationary* outliers, which cannot be detected by standard GMTI methods. In the latter case, it is desirable to account for the correlation of targets who occupy similar locations in subsequent images as well as spatially within a single frame. We impose this structure through spatial and temporal Markov properties on the sparse component.

The goal of this work is two-fold. First, we present a unifying Bayesian formulation that incorporates SAR-specific phenomena such as glints, speckle noise, and calibration errors, and is additionally able to include information from multiple passes of the radar and multiple receive channels. Second, we offer an inference algorithm through MCMC methods in order to estimate the posterior distribution given the observed SAR images. This posterior distribution can then be used for the desired inference task, such as target detection and/or estimation of the underlying clutter distribution.

The rest of the paper is organized as follows: The algorithmic structure is given in Section II. Notation is presented in Section III and the image model is provided in IV. Markov, spatial, and/or target kinematic extensions are discussed in Section V. The inference algorithm is given in Section VI. Performance is analyzed over both simulated and measured datasets in Section VII. We conclude and point to future work in Section VIII.

II. ALGORITHMIC STRUCTURE

For clarity, in this section we provide the basic structure of the algorithm described by this paper. The algorithm works as
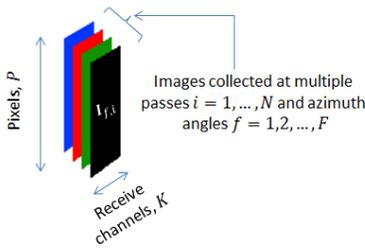
Fig. 1. SAR image sets $\boldsymbol{I}_{f,i}$ from $i = 1, 2, \ldots, N$ passes and $f = 1, 2, \ldots, F$ azimuth angles are collected in a multidimensional array. Each SAR image set consists of K images (each with P pixels) collected at different receive channels.

follows

1) SAR images sets are collected from N independent passes of the radar and from F azimuth ranges (i.e., with different azimuth angles). Each image set consists of K images (each with P pixels) collected from one of K receive channels (or antennas). Subsequently, the images ($\boldsymbol{I}_{f,i}$) are arranged in a multidimensional array indexed by their pass i and frame (azimuth range) f. This multidimensional array is shown in Figure 1.

2) A generic model is proposed that decomposes the SAR images $\boldsymbol{I}_{f,i}$ into low-rank, sparse, noise, and calibration components. Moreover, the low-rank and sparse components are decomposed into SAR-specific components that include speckle, glints, and moving targets.

3) A prior probability distribution is proposed for each type of pixel class (speckle, glints, etc.). Note that any individual pixel may be a mixture of these classes. Each probability distribution depends on various distribution parameters (e.g. the mean parameter of a Gaussian random variable). We propose a model where these parameters are also estimated from the data using so-called conjugate priors. These priors can be specified in a standard way [5] in order to be non-informative and to not require tuning.

4) Given the prior distributions in the previous step, the posterior distribution for the model parameters is derived, given the images $\boldsymbol{I}_{f,i}$. This posterior distribution is then estimated through a Markov Chain Monte Carlo (MCMC) method, namely the Gibbs Sampler.

5) Finally, the posterior distribution is provided which can be used for various tasks such as
   - Detection of moving targets and estimation of their radial velocity.
   - Determination of the clutter distribution (which can subsequently be used for STAP, even in the presence of heterogeneous noise).
   - Sensor fusion and/or efficient allocation of sensing resources.

III. NOTATION

Available is a set of SAR images of a region formed from multiple passes of an along-track radar platform with multiple antennas (i.e., phase centers.) Moreover, images are formed over distinct azimuth angle ranges that can be indexed by the frame number, f. Table I provides the indexing scheme used throughout this paper in order to distinguish between images from various antennas, frames, and/or passes. Table II provides a list of indexing conventions used to denote collections of variables.

TABLE I
INDEX VARIABLE NAMES USED IN PAPER

Index Description	Index Variable	Range
Antenna (channel)	k	$1, 2, \ldots, K$
Frame (azimuth angle)	f	$1, 2, \ldots, F$
Pass	i	$1, 2, \ldots, N$
Pixel	p	$1, 2, \ldots, P$

TABLE II
OUR DATA INDEXING CONVENTIONS

Variable	Convention	Description
$i_{k,f,i}^{(p)}$	Standard	Value at pixel p, antenna k, and frame f, pass i
$\underline{i}_{f,i}^{(p)}$	Underline	Values at pixel p, frame f, and pass i over all antennas
$\boldsymbol{i}_{f,1:N}^{(p)}$	Lower-case, Boldface	Values at pixel p and frame f over all antennas and passes
$\boldsymbol{I}_{f,i}$	Upper-case Boldface	Values over all pixels and antennas at frame f and pass i
$\boldsymbol{I}$	Upper-case, Boldface, No Indices	Values over all pixels, antennas, frames, and passes

We model the complex pixel values in SAR images with the complex-normal distribution, where we use the notation

$$\underline{w} \sim \mathcal{CN}\left(0, \boldsymbol{\Gamma}\right) \tag{3}$$

where $\mathcal{CN}(\mu, \boldsymbol{\Gamma})$ represents the complex-Normal distribution with mean μ and complex covariance matrix $\boldsymbol{\Gamma}$, and $\vec{w}$ is random vector of K complex-values (from each of K antennas.) Specifically, we directly model the correlations of pixel values among the K antennas (receive channels) through the complex covariance matrix $\boldsymbol{\Gamma}$.

IV. SAR IMAGE MODEL

We propose a decomposition of SAR images at each frame f and pass i as follows

$$\boldsymbol{I}_{f,i} = \boldsymbol{H}_{f,i} \circ \left(\boldsymbol{L}_{f,i} + \boldsymbol{S}_{f,i} + \boldsymbol{V}_{f,i}\right), \tag{4}$$

where $\boldsymbol{H}_{f,i}$ is a spatiotemporally-varying filter that accounts for antenna calibration errors, $\boldsymbol{L}_{f,i}$ is a low-dimensional representation of the background clutter, $\boldsymbol{S}_{f,i}$ is a sparse component that contains the targets of interest, $\boldsymbol{V}_{f,i}$ is zero-mean additive noise, and $\circ$ denotes the Hadamard (element-wise) product. Each of these components belongs to the space $\mathbb{C}^{P \times K}$.

As discussed earlier, this decomposition may be appropriate for SAR imagery where stationary (clutter) features in the scene don't change much from frame-to-frame, pass-to-pass, and antenna-to-antenna. The remainder of this section discusses the model in detail. Figure 2 shows a graphical representation of the model.
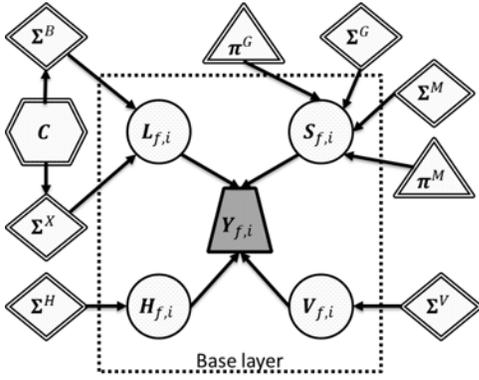
Fig. 2. This figure provides a simplified graphical model associated with proposed SAR image model. The shaded shape represents the observed random variable. The circles represent the basic parameters of the model, while the dashed lines represent hyperparameters that are also modeled as random variables.

A. Low-dimensional component, $\boldsymbol{L}_{f,i}$

We propose a decomposition of the low-rank component as

$$\boldsymbol{L}_{f,i} = \boldsymbol{B}_f + \boldsymbol{X}_{f,i}, \qquad (5)$$

where $\boldsymbol{B}_f$ is the inherent background that is identical over all passes, $\boldsymbol{X}_{f,i}$ is the speckle noise component that arises from coherent imaging in SAR. Posner [15] and Raney [16] describe speckle noise, which tends to be spatially correlated depending on the texture of the surrounding pixels.

Gierull [23] shows that the quadrature components of SAR radar channels are often modeled as zero-mean Gaussian processes. However, in the presence of heterogeneous clutter (such as in urban scenes), one must consider spatially-varying models where the clutter variance changes across radar cells. Additionally, speckle noise is usually multiplicative in nature wherein higher amplitude clutter produces higher variance noise. In this work, we approximate this multiplicative property by allowing the speckle/background distributions to vary spatially. This additive formulation lessens the computational burden, while empirical evidence suggests that the approximation is reasonable.

To account for this spatial variation, this model assumes that each background pixel can be defined by one of J classes that may be representative of roads, vegetation, or buildings within the scene. These classes are learned directly from the data so that their distributions (i.e., covariance matrices) do not need to specified a priori. Moreover, while there may be many different pixel classes, one can reasonably model the data with $J \ll P$, where P is the number of pixels in the measured images. In other words, one generally only requires a relatively small number of classes (distributions) to describe the clutter. To this end, we put a multinomial model on each object class

$$\underline{c}^{(p)} = \left\{ c_j^{(p)} \right\}_{j=1}^{J} \sim \text{Multinomial}(1; q_1, q_2, \ldots, q_J) \qquad (6)$$

where q_j is the prior probability of the j-th object class. Then the class assignment $C^{(p)}$ is the single location in $\underline{c}$ with value equal to one. We use a hidden Markov model dependency

that reflects that neighboring pixels are likely to have the same class. The class $C^{(p)}$ defines the distribution of the pixel p, where we specifically model the background and speckle components respectively as complex-normal distributed:

$$\underline{b}_f^{(p)} \sim \mathcal{CN}\left(0, \boldsymbol{\Gamma}_B^{C^{(p)}}\right), \quad \underline{x}_{f,i}^{(p)} \sim \mathcal{CN}\left(0, \boldsymbol{\Gamma}_X^{C^{(p)}}\right) \qquad (7)$$

Note that the class type specifies the distribution of the pixels and each vector of K values (e.g. background $\underline{b}_f^{(p)}$ or speckle $\underline{x}_{f,i}^{(p)}$) is drawn independently from that distribution. These prior distributions are constructed so that the pixels are conditionally independent given the pixel classes. This independence assumption allows for numerical efficiency in the inference algorithm. Moreover, pixels still maintain a correlation through (a) the determination of the pixel classification (i.e., through the hidden Markov model) and (b) the definition of the class distribution (i.e., the covariance matrices.)

B. Sparse component, $\boldsymbol{S}_{f,i}$

The sparse component contains two components: a specular noise (glint) component and a target component. We consider a shared sparsity model, wherein glint/target components are present in one antenna if and only if they are present in the other antennas. Moreover, glints are known to have a large angular dependence, in the sense that the intensity of the glint dominates in only a few azimuth angles. Thus, the indicators for glints are assumed to persist across all passes. The sparse component is modeled as

$$\boldsymbol{S}_{f,i} = \left(\boldsymbol{\Delta}_{f,1:N}^{G} \otimes \boldsymbol{1}_K^T\right) \circ \boldsymbol{G}_{f,i} + \left(\boldsymbol{\Delta}_{f,i}^{M} \otimes \boldsymbol{1}_K^T\right) \circ \boldsymbol{M}_{f,i}, \quad (8)$$

where $\boldsymbol{G}_{f,i} \in \mathbb{C}^{P \times K}$ is the specular noise (glints) component with associated indicator variables $\boldsymbol{\Delta}_{f,1:N}^{G} \in \{0,1\}^P$, $\boldsymbol{M}_{f,i} \in \mathbb{C}^{P \times K}$ is the (moving) target component with associated indicator variables $\boldsymbol{\Delta}_{f,i}^{M} \in \{0,1\}^P$, $\boldsymbol{1}_K$ is the all ones vector of size $K \times 1$, and $\otimes$ is the Kronecker product. The Kronecker product is used to denote the shared sparsity across receive channels. Once again, we assume that the glints and target components are zero-mean complex-normal distributed with covariances $\boldsymbol{\Gamma}_G$ and $\boldsymbol{\Gamma}_M$, respectively.

The indicator variable $\delta^{z,(p)}$ at pixel p, where z is representative of either g (glints) or m (moving targets), is modeled as

$$\delta^{z,(p)} \sim \text{Bernoulli}(\pi^{z,(p)}), \qquad (9)$$
$$\pi^{z,(p)} \sim \text{Beta}(a_\pi, b_\pi) \qquad (10)$$

Whereas the Gaussian distribution is a natural choice for continuous random variables, the Bernoulli distribution with parameter π is a natural choice for indicator variables, where π denotes the probability that the random variable is equal to one. Moreover, the Beta distribution is often used (as in [4]) to account for uncertainty in π. In sparse situations, we would generally expect that $\pi \ll 1$. In this model, a sparseness prior is obtained by setting $a_\pi/[a_\pi + b_\pi] \ll 1$. Alternatively, we can introduce additional structure in our model by letting a_π and b_π depend on previous frames (temporally) and/or neighboring pixels (spatially). This is particularly useful for detecting multi-pixel targets that move smoothly through a scene. Section V discusses this modification in greater detail.

C. Distribution of quadrature components

Many SAR detection algorithms rely on the ability to separate the target from the background clutter by assuming that the clutter lies in a low-dimensional subspace of the data. Consider a random vector of complex variables $\underline{w} \sim \mathcal{CN}(0, \mathbf{\Gamma})$ where w is representative of b (background), x (speckle), g (glints) or m (targets.) Under the assumptions that (a) the quadrature components of each antenna are zero-mean normal with variance σ^2 and (b) the correlation among components w_m and w_n is given by $\rho e^{-j\phi_{mn}}$, then $\mathbf{\Gamma}$ can be shown to have the form

$$\mathbf{\Gamma} = \sigma^2 \begin{bmatrix} 1 & \rho e^{-j\phi_{12}} & \cdots & \rho e^{-j\phi_{1K}} \\ \rho e^{j\phi_{12}} & 1 & \cdots & \rho e^{-j\phi_{2K}} \\ \vdots & \vdots & \ddots & \vdots \\ \rho e^{j\phi_{1K}} & \rho e^{j\phi_{2K}} & \cdots & 1 \end{bmatrix}, \quad (11)$$

where σ^2 is the channel variance, ρ is the coherence between antennas, and $\{\phi_{nm}\}_{n,m}$ are the interferometric phase differences between the antennas[1]. In an idealized model with a single point target, the interferometric phases ϕ_{mn} can be shown to be proportional to the target radial velocity [8]. In images containing only stationary targets (i.e., the background components where $\phi_{mn} = 0$), the covariance matrix has a simpler form:

$$\mathbf{\Gamma}_{background} = \sigma^2 \left((1-\rho)\mathbf{I}_{K \times K} + \rho \mathbf{1}_K \mathbf{1}_K^T \right) \quad (12)$$

where $\mathbf{I}_{K \times K}$ is the $K \times K$ identity matrix and $\mathbf{1}_K$ is the all-ones vector of length K. When there is no correlation among the antennas ($\rho = 0$), this reduces to a scaled identity matrix. In other cases, this covariance matrix provides a way to capture direct correlations among the antennas.

When the covariance matrix $\mathbf{\Gamma}$ of a multivariate Gaussian distribution is not known a priori, a common choice for a prior distribution is its conjugate prior, the Inverse-Wishart distribution. This distribution is characterized by the mean covariance and its prior weight (i.e., how strongly to weight the prior). In this paper, we use a modification to the standard model where (a) the channel variance is estimated separately from the structure of the correlation matrix, and (b) the correlation matrix mean depends on another random parameter, the correlation coefficient ρ. This additional structure aligns well with the literature that relies on strong correlations among antennas (specifically for clutter and speckle). Moreover, this model separates the learning of the channel variance σ^2, which we have no a priori knowledge about, from the learning of the correlation structure, denoted $\mathbf{\Gamma}_\rho$. The specific model is given by

$$\underline{w} \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2 \mathbf{\Gamma}_\rho\right) \quad (13)$$

$$\mathbf{\Gamma}_\rho \sim \text{InvWishart}\left(a_\Gamma((1-\rho)\mathbf{I}_{K \times K} + \rho \mathbf{1}_K \mathbf{1}_K^T), \nu_\Gamma\right) \quad (14)$$

$$\sigma^2 \sim \text{InvGamma}(a_\sigma, b_\sigma) \quad (15)$$

$$\rho \sim \text{Beta}(a_\rho, b_\rho) \quad (16)$$

[1]A more general model could account for different channel variance and coherence values, but since we use the calibration constants $\mathbf{H}_{f,i}$ to equalize the channels, the effect was seen to be relatively insignificant.

where $a_\sigma = b_\sigma = 10^{-6}$ as suggested by Tipping [5] to promote non-informative priors, (a_ρ, b_ρ) are chosen so that $\rho \approx 1$ to ensure a high coherence among the background components, ν_Γ is a parameter that controls how strongly to weight the prior covariance matrix, and a_Γ is chosen so that $E[\mathbf{\Gamma}_\rho] = (1-\rho)\mathbf{I}_{K \times K} + \rho \mathbf{1}_K \mathbf{1}_K^T$. In this work, ν_Γ is chosen to be large in order to reflect our belief that $\sigma^2 \mathbf{\Gamma}_\rho$ should be close to equation (12).

D. Calibration filter, $\mathbf{H}_{f,i}$

The calibration constants are assumed to be constant within small spatial regions $p \in Z_g$, though they may vary as a function of antenna, frame, or pass. In particular, we let

$$h_{k,f,i}^{(p)} = z_{k,f,i}(g), \forall p \in Z_g, \quad (17)$$

$$z_{k,f,i}(g) \sim \mathcal{CN}(1, (\sigma^H)^2) \quad (18)$$

where we note that if the number of pixels in class g, $|Z_g|$, is large, then maximum likelihood inference in this case yields the least-squares solution.

V. MARKOV/SPATIAL/KINEMATIC MODELS FOR THE SPARSE COMPONENT

A. Indicator probability models

This model contains multiple indicator variables with prior probabilities distributed as $\text{Beta}(a_\pi, b_\pi)$. Moreover, sparsity is obtained when $a_\pi/[a_\pi + b_\pi] \ll 1$. Alternatively, we can introduce additional structure in our model by letting a_π and b_π depend on previous frames (temporally) and/or neighboring pixels (spatially). This is especially useful for detecting multi-pixel targets that move smoothly through a scene.

Define $W^M(p, \mathbf{\Delta}_{f,i}^M)$ to be a function that maps the indicator variables $\mathbf{\Delta}_{f,i}^M$ to a real number. For example, this may be the average number of non-zero indicators in the neighborhood of pixel p, or a weighted version that puts higher value on neighboring pixels. For $f = 1$, we let

$$\begin{bmatrix} a_{1,i}^M(p) \\ b_{1,i}^M(p) \end{bmatrix} = \begin{cases} [a_H \quad b_H]^T, & W^M(p, \mathbf{\Delta}_{1,i}^M) > \varepsilon_{spatial}^M, \\ [a_L \quad b_L]^T, & \text{else,} \end{cases} \quad (19)$$

and for $f > 1$

$$\begin{bmatrix} a_{f,i}^M(p) \\ b_{f,i}^M(p) \end{bmatrix} = \begin{cases} [a_H \quad b_H]^T, & W^M(p, \mathbf{\Delta}_{f,i}^M) > \varepsilon_{spatial}^M \text{ and} \\ & \quad W^M(p, \mathbf{\Delta}_{f-1,i}^M) > \varepsilon_{temporal}^M, \\ [a_L \quad b_L]^T, & \text{else.} \end{cases} \quad (20)$$

In this paper, we choose (a_L, b_L, a_H, b_H) so that $a_L/(a_L + b + L) \ll 1$ and $a_H/(a_H + b + H) \gg 0$. A similar model can be introduced for the probabilities of the glints.

B. Target kinematic model

In some applications, such as target tracking or sequential detection, we may have access to an estimate of the kinematic state of the target(s) of interest, such as position, velocity and acceleration. This could be provided separately, or one might consider a joint estimation problem where target kinematics are being simultaneously estimated with the image model. The

Procedure 1 Gibbs Sampling Pseudocode

---

procedure $\{\boldsymbol{\Theta}\}_{i=1:N_{samples}} = \text{SARGibbs}(\boldsymbol{\Theta}_0, \boldsymbol{I})$
  $\boldsymbol{\Theta} \leftarrow \boldsymbol{\Theta}_0$
  for $iteration = 1$ to $N_{burnin} + N_{samples}$ **do**
    Sample $\sim f\left(\boldsymbol{B}, \boldsymbol{X}, \boldsymbol{G}, \boldsymbol{M}, \boldsymbol{\Delta}^G, \boldsymbol{\Delta}^M | \boldsymbol{I}, -\right)$ *//Base*
    Sample $\sim f\left(\boldsymbol{H} | \boldsymbol{I}, -\right)$ *//Calibration filter*
    Sample $\sim f\left(\boldsymbol{C} | \boldsymbol{I}, -\right)$ *//Class assignment*
    Sample $\sim f\left(\boldsymbol{\eta} | \boldsymbol{I}, -\right)$ *//Hyper-parameters*
    $\boldsymbol{\Theta}_{iteration-N_{burnin}} \leftarrow \boldsymbol{\Theta}$ if $iteration > N_{burnin}$
  end for
end procedure

---

target state estimate at any particular time could also be useful for predicting the location of the target at sequential frames. For simplicity, consider a single target at time τ whose state $\xi(\tau) = (\boldsymbol{r}(\tau), \dot{\boldsymbol{r}}(\tau))$ is known with standard errors $\boldsymbol{\Sigma}_\xi(\tau)$. Note that the uncertainty model for $(\boldsymbol{r}, \dot{\boldsymbol{r}})$ may be (a) known a prior from road maps or traffic behavior patterns, or (b) learned adaptively using some signal processing algorithm such as the Kalman or particle filters.

In standard SAR image formation, moving targets tend to appear displaced and defocused as described by Fienup [7] and Jao [13]. Moreover, Jao showed that given the radar trajectory $(\boldsymbol{q}, \dot{\boldsymbol{q}})$ and the target trajectory $(\boldsymbol{r}, \dot{\boldsymbol{r}})$, one can predict the location of the target signature within the image $\boldsymbol{p}$ by solving a system of equations that equate Doppler shifts and ranges, respectively, at each pulse:

$$\frac{d}{d\tau}\left[\|\boldsymbol{p} - \boldsymbol{q}(\tau)\|_2 - \|\boldsymbol{r}(\tau) - \boldsymbol{q}(\tau)\|_2\right]_{\boldsymbol{p}=\boldsymbol{p}^*} = 0 \qquad (21)$$

$$\|\boldsymbol{p}^* - \boldsymbol{q}(\tau)\|_2 = \|\boldsymbol{r}(\tau) - \boldsymbol{q}(\tau)\|_2, \qquad (22)$$

which can be reduced to the simpler system of equations:

$$\dot{\boldsymbol{q}}(\tau) \cdot [\boldsymbol{p}^* - \boldsymbol{q}(\tau)] = [\dot{\boldsymbol{r}}(\tau) - \dot{\boldsymbol{q}}(\tau)] \cdot [\boldsymbol{r}(\tau) - \boldsymbol{q}(\tau)] \qquad (23)$$

$$\|\boldsymbol{p}^* - \boldsymbol{q}(\tau)\|_2 = \|\boldsymbol{r}(\tau) - \boldsymbol{q}(\tau)\|_2 \qquad (24)$$

The probable locations of the target can be predicted by one of several methods, including:

- Monte Carlo estimation of the target posterior density.
- Gaussian approximation using linearization or the unscented transformation to approximate the posterior density
- Analytical approximation.

Given an estimate of the posterior density, we can modify the function W^M described in the previous section to include dependence on this kinematic information. Details of the posterior density estimation are provided in the technical report [24].

VI. INFERENCE

In this section, we provide details on estimating the posterior distribution of the model parameters given the observed SAR images. Given the estimate of the posterior distribution, one can then perform that appropriate desired task, such as detection of moving targets and/or estimation of the clutter distribution. In the former task, this can be done simply by

thresholding the probability of the target indicators. In the latter task, the posterior distribution could specify a confidence interval (or region) for the parameters of interest (such as the covariance matrices and the pixel classification probabilities).

Generally, estimating the posterior distribution on this model would be a very difficult task due to the large number of variables and the dependence among them. In particular, we use a Markov Chain Monte Carlo (MCMC) algorithm in the form of a Gibbs sampler to iteratively estimate the full joint posterior. In MCMC, this distribution is approximated by drawing samples iteratively from the conditional distribution of each (random) model variable given the most recent estimate of the rest of the variables (which we denote by $-$) [25]. Let $\boldsymbol{\Theta} = \left\{\boldsymbol{B}, \boldsymbol{X}, \boldsymbol{G}, \boldsymbol{M}, \boldsymbol{\Delta}^G, \boldsymbol{\Delta}^M, \boldsymbol{H}, \boldsymbol{C}, \boldsymbol{\eta}\right\}$ represent a current estimate of all of the model variables where $\boldsymbol{\eta}$ represents the set of all hyper-parameters. Given measurements $\boldsymbol{I}$, the inference algorithm is given in Procedure 1. Note that MCMC algorithms require a burn-in period, after which the Markov chain has become stable. The duration of the burn-in period depends on the problem and is discussed in more detail below. After the Markov chain has become stable, we collect $N_{samples}$ samples that represent the full joint distribution. Full details of the sampling procedures are given in the technical report [24].

This model requires estimation of a base layer (i.e. the direct random variables given in equations (4), (5), and (8)), the parameters of the distributions of the base layer (i.e. the covariance matrices and probabilities), and the global parameters (i.e., the clutter class assignments and the calibration filter coefficients.) To combat this numerical intractability, this model was constructed in a specific way such that (a) the hyper-parameters of the base layer were chosen to be conjugate to the base layer, and (b) the posterior distribution of the base layer is conditionally independent across pixels/frames given the other parameters. The former property allows for efficient sampling of the posterior distribution in the sense that we can sample exactly from these distributions. The latter property allows for simple parallelization of the sampling procedure over the largest dimensions of the state.

Moreover, the sampling procedures for the hyper-parameters tend to require sufficient statistics that are of significantly smaller dimension and thus more desirable from a computational viewpoint. For example, sampling of the covariance matrix $\boldsymbol{\Gamma}^M$ depends only on a $K \times K$ sample covariance matrix. It should be noted that sampling of the covariance matrices requires additional effort in order to constrain its shape to that of equation (11). In particular, we use a Metropolis-Hastings step , which can be easily done by noting that the posterior density $f(\boldsymbol{\Gamma}^W, \rho^W, (\sigma^2)^W | \boldsymbol{W})$ is proportional to an Inverse-Wishart distribution. Details specific to inference in this model are provided in the technical report [24]. For a formal presentation of Monte Carlo methods, including Gibbs samplers and Metropolis-Hastings, the authors suggest reading [25].

The computational complexity of this MCMC method is characterized both by the computational burden of a single iteration in the sampling process as well as the number of required iterations for burn-in and subsequent sampling. The

TABLE III
PARAMETERS OF SIMULATED DATASET

Parameter	Value
Pixels in image, P	$P = 100 \times 100$
Number of frames per pass, F	$F = 1$
# of antennas, K	$K = 3$
# of passes, N	$N \in \{5, 10, 20\}$
# of target pixels/image, $N_{targets}$	$N_{targets} = 20$
Clutter of background, ρ	$\rho \in \{0.9, 0.99, 0.999, 0.9999\}$
Variance of targets, σ^2_{target}	$\sigma^2_{target} = 1$
Variance of background	Either $\sigma^2_{dim} = \sigma^2_{clutter}/100$ or $\sigma^2_{bright} = \sigma^2_{clutter}$
Signal-to-noise-plus clutter (SCNR)	$SCNR \triangleq \dfrac{\sigma^2_{target}}{\sigma^2_{clutter} + \sigma^2_{noise}} \in \{0.1, 0.5, 1, 2\}$

(a) $\boldsymbol{L} + \boldsymbol{S}$

(b) $\boldsymbol{L}$

(c) $\boldsymbol{S}$
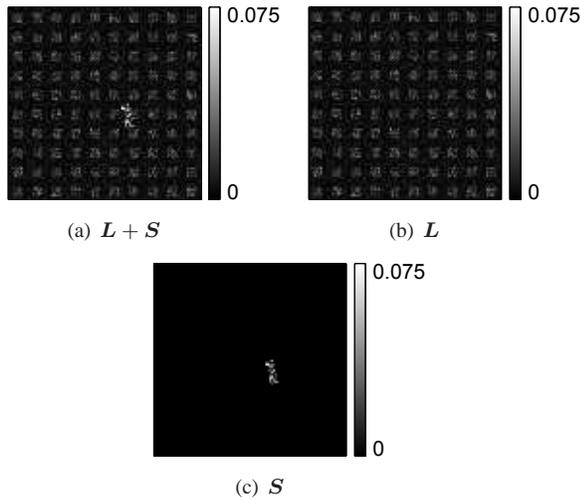
Fig. 3. This figure provides a sample image used in the simulated dataset for comparisons to RPCA methods, as well as its decomposition into low-dimensional background and sparse target components. This low SCNR image is typical of measured SAR images. Note that the target is randomly placed within the image for each of N passes. In some of these passes, the target is placed over low-amplitude clutter and can be easily detected. In other passes, the target is placed over high-amplitude clutter, which reduces the capability to detect the target.

former step is highly parallelizable and can be accomplished efficiently even for large images and multiple passes. In our experience, the required computation time was on the same order as the time required to form the images from the raw phase histories (which generally scales as $\mathcal{O}(P^2 \log P)$, where P is the number of pixels in the image.) Moreover, similar to related work [4], our experience has shown that the mean of the posterior distribution converges quickly with just a few iterations of the MCMC algorithm. Due to the parallelizability of the problem, this algorithm could potentially benefit greatly by computation on GPU's where parallelization is built-in.

VII. PERFORMANCE ANALYSIS

A. Simulation

We first demonstrate the performance of the proposed algorithm, which we refer to as the Bayes SAR algorithm, on a simulated dataset. Images were created according to the

TABLE IV
COMPARISON OF PROPOSED METHOD (BAYES SAR) TO RPCA METHODS WITH $N = 20$, $F = 1$, $K = 3$. NOTE THAT THE BAYES SAR METHOD PERFORMS ABOUT TWICE AS WELL AS EITHER OF THE RPCA METHODS FOR ALL CRITERIA. THE BAYES SAR METHOD ALSO PRODUCES A SPARSE RESULT. STANDARD ERRORS ARE PROVIDED IN PARENTHESES.

(a) Bayes SAR

SCNR	Coh.	$\frac{\|\boldsymbol{L}-\hat{\boldsymbol{L}}\|_2}{\|\boldsymbol{L}\|_2}$	$\frac{\|\boldsymbol{S}-\hat{\boldsymbol{S}}\|_2}{\|\boldsymbol{S}\|_2}$	$\frac{\|\boldsymbol{S}-\hat{\boldsymbol{S}}\|_0}{\|\boldsymbol{S}\|_0}$
10%	.900	.058 (.001)	.639 (.134)	.664 (.234)
10%	.9999	.048 (.005)	.414 (.036)	.365 (.042)
100%	.900	.056 (.001)	.155 (.015)	.152 (.009)
100%	.9999	.053 (.003)	.121 (.008)	.097 (.017)
200%	.900	.057 (.001)	.122 (.011)	.145 (.043)
200%	.9999	.053 (.005)	.117 (.016)	.094 (.009)

(b) Opt. RPCA

SCNR	Coh.	$\frac{\|\boldsymbol{L}-\hat{\boldsymbol{L}}\|_2}{\|\boldsymbol{L}\|_2}$	$\frac{\|\boldsymbol{S}-\hat{\boldsymbol{S}}\|_2}{\|\boldsymbol{S}\|_2}$	$\frac{\|\boldsymbol{S}-\hat{\boldsymbol{S}}\|_0}{\|\boldsymbol{S}\|_0}$
10%	.900	.113 (.006)	3.22 (.19)	110.9 (1.5)
10%	.9999	.113 (.006)	3.20 (.16)	108.8 (2.4)
100%	.900	.112 (.006)	1.20 (.07)	109.7 (1.9)
100%	.9999	.113 (.008)	1.20 (.08)	107.9 (2.3)
200%	.900	.116 (.010)	1.08 (.09)	110.3 (2.6)
200%	.9999	.110 (.003)	1.04 (.03)	108.7 (2.6)

(c) Bayes RPCA

SCNR	Coh.	$\frac{\|\boldsymbol{L}-\hat{\boldsymbol{L}}\|_2}{\|\boldsymbol{L}\|_2}$	$\frac{\|\boldsymbol{S}-\hat{\boldsymbol{S}}\|_2}{\|\boldsymbol{S}\|_2}$	$\frac{\|\boldsymbol{S}-\hat{\boldsymbol{S}}\|_0}{\|\boldsymbol{S}\|_0}$
10%	.900	.119 (.018)	1.04 (.08)	3.96 (.49)
10%	.9999	.116 (.022)	1.08 (.22)	3.91 (.54)
100%	.900	.126 (.029)	.768 (.082)	3.72 (.88)
100%	.9999	.125 (.023)	.754 (.061)	3.68 (.63)
200%	.900	.135 (.030)	.735 (.146)	3.93 (.95)
200%	.9999	.134 (.028)	.703 (.067)	3.86 (.82)

model given in Section IV with parameters given in Table III. The low-dimensional component was divided into one of two classes ('dim' or 'bright'). Pixels were deterministically assigned to one of these classes to resemble a natural SAR image (see Figure 3). The sparse component included a randomly placed target with multiple-pixel extent. A spatiotemporally varying antenna gain filter was uniformly drawn at random on the range $[0, 2\pi)$ for groups of pixels of size 25×25. Lastly, zero-mean IID noise was added with variance σ^2_{noise}.

The Bayes SAR model is applied to infer the low-dimensional component $\boldsymbol{L}_{f,i}$ and sparse target component $\boldsymbol{S}_{f,i}$ with estimates denoted $\hat{\boldsymbol{L}}_{f,i}$ and $\hat{\boldsymbol{S}}_{f,i}$, respectively. Hyperparameters of the model are chosen according to the Section VI. Results are given by the mean of MCMC inference with 500 burn-in iterations followed by 100 collection samples. We consider three metrics to evaluate the reconstruction errors: $\frac{\|\boldsymbol{L}-\hat{\boldsymbol{L}}\|_2}{\|\boldsymbol{L}\|_2}$, $\frac{\|\boldsymbol{S}-\hat{\boldsymbol{S}}\|_2}{\|\boldsymbol{S}\|_2}$, $\frac{\|\boldsymbol{S}-\hat{\boldsymbol{S}}\|_2}{\|\boldsymbol{S}\|_0}$, where the norm is taken over the vectorized quantities.

In comparison to the Bayes SAR model, results are given for state-of-the-art algorithms for Robust Principal Component Analysis (RCPA): an optimization-based approach proposed by Wright et al. [1] and Candes et al. [3] and a Bayesian-based approach proposed by Ding et al. [4][2]. The optimization-based
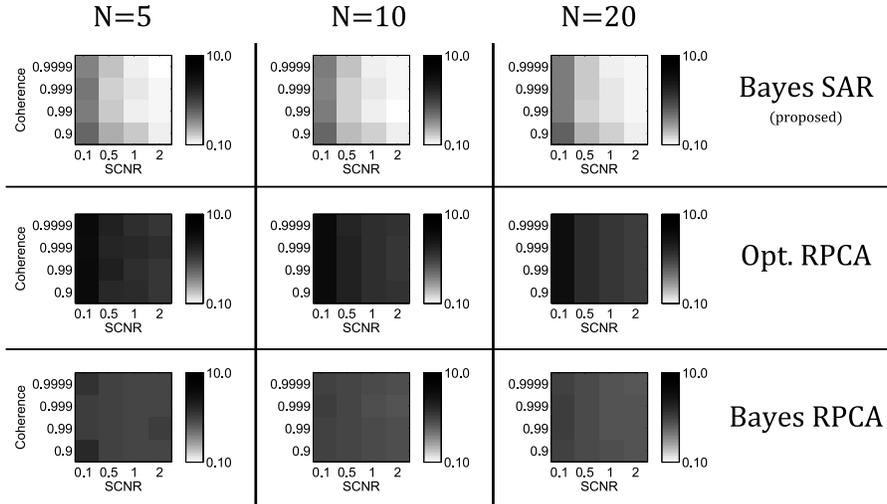
Fig. 4. This figure compares the relative reconstruction error of the target component, $\frac{\|\boldsymbol{S}-\hat{\boldsymbol{S}}\|_2}{\|\boldsymbol{S}\|_2}$, as a function of algorithm, number of passes N, coherence of antennas ρ, and signal-to-clutter-plus-noise ratio (SCNR). From top-to-bottom, the rows contains the output of the Bayes SAR algorithm (proposed), the optimization-based RPCA algorithm [1], [3], and the Bayes RPCA algorithm [4]. From left-to-right, the columns show the output for $N = 5$, $N = 10$, and $N = 20$ passes (with $F = 1$ frames per pass). The output is given by the median error over 20 trials on a simulated dataset. It is seen that in all cases, the Bayes SAR method outperforms the RPCA algorithms. Moreover, the Bayes SAR algorithm performs better if either coherence increases (i.e., better clutter cancellation) or the SCNR increases. On the other hand, the performance of the RPCA algorithms does not improve with increased coherence, since these algorithms do not directly model this relationship.

approach requires a tolerance parameter which is related to the noise level, as suggested by Ding et al. [4]. We chose this parameter in order to have the smallest reconstruction errors. The Bayesian method did not require tuning parameters, except for choosing the maximum rank of $\boldsymbol{L}_{f,i}$ which was set to 20.

Figure 4 compares the relative reconstruction error of the sparse (target) component, $\frac{\|\boldsymbol{S}-\hat{\boldsymbol{S}}\|_2}{\|\boldsymbol{S}\|_2}$, across all algorithms, number of passes N, coherence of antennas ρ, and SCNR. In all cases, the Bayes SAR method outperforms the RPCA algorithms with improving performance if either coherence or SCNR increases. Table IV provides additional numerical results for the case $N = 20$. The RCPA algorithms perform poorly in reconstructing the sparse component with relative errors near or greater than 1. This reflects the fact that (a) these algorithms miss significant sources of information, such as the correlations among antennas and among quadrature components, and (b) $N = 20$ may be too few samples to reliably estimate the principal components in these non-parametric models. In measured SAR imagery, it might be unreasonable to expect $N \gg 20$ passes of the radar, which suggests that these RPCA algorithms will likely perform poorly on such signals. In contrast, it is seen that the Bayes SAR method obtains low reconstruction errors for both low-dimensional and sparse components as either coherence or SCNR increase.

Table IV also provides standard errors on the metrics provided (i.e., the l-norms on the foreground and background components) as a measure of statistical confidence in these quantities. Note that this standard error is calculated over the 20 trials where the ground truth is known. It would also be possible to determine the predicted uncertainty of the error in the Bayesian methods by computing the standard error over the

samples in the Monte Carlo distribution. This would provide a predicted uncertainty in any one instantiation of the problem (i.e., one trial), but would not be comparable to the non-Bayesian methods (such as the optimization-based RPCA).

B. Measured data

In this section, we compare performance of the Bayes SAR approach using a set of measured data from the 2006 X-band Gotcha SAR sensor collection.[3] In particular, images were formed from phase histories collected over a scene of size 375m by 1200m for $N = 3$ passes and $K = 3$ antennas. Each image was created with a coherent processing time of 0.5 seconds with the addition of a Blackman-Harris window in the azimuth direction to reduce sidelobes. Images were created at 0.5m resolution in both the x- and y-directions. Thus each image consisted of $P = 750 \times 2400 = 1.8 \times 10^6$ pixels. Images were created at overlapping intervals spaced 0.25 seconds apart for a total of 18 seconds. Note that the ability to take advantage of correlated images (as in this case) is one of the benefits of using the proposed model/inference algorithm.

We consider three alternative approaches in comparison to the Bayes SAR approach: (1) displaced-phase center array (DPCA) processing, (2) along-track interferometry (ATI), and (3) a mixture of DPCA/ATI as described by Deming [8]. Note that all variants of ATI/DPCA depend on the chosen thresholds for phase/magnitude, respectively.

1) Comparisons to DPCA/ATI: We begin by comparing the output of the proposed algorithm across the entire 375m by 1200m scene. Figure 5 shows the output of the Bayes SAR algorithm, the DPCA output, and the ATI output. It is seen that there are significant performance gains by using

---

[3]The dataset is a superset of the data given by the Air Force Research Lab, RYA division as described in [26].
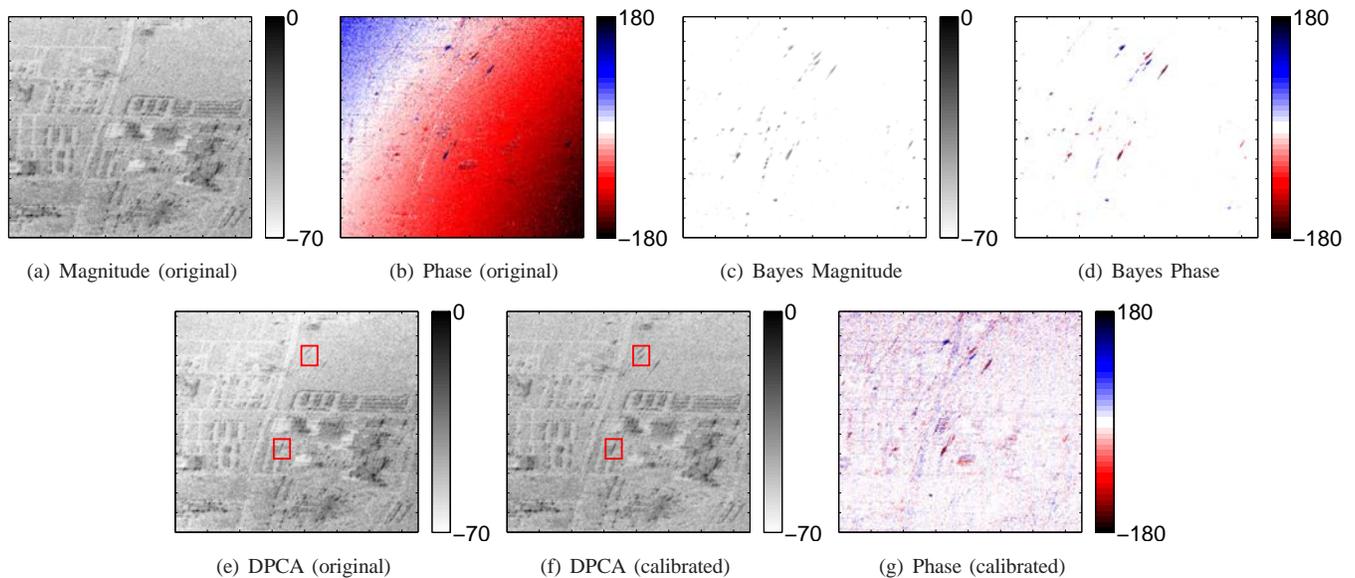
Fig. 5. This figure compares the output of the proposed algorithm as a function of magnitude and phase for a scene of size 375m by 1200m and coherent processing interval of 0.5s. The Bayes SAR algorithm takes the original SAR images in (a) and (b), estimates the nuisance parameters such as antenna miscalibrations and clutter covariances, and yields a sparse output for the target component in (c) and (d). In contrast, the DPCA and ATI algorithms are very sensitive to the nuisance parameters, which make finding detection thresholds difficult. In particular, consider the original interferometric phase image shown in (b). It can be seen that without proper calibration between antennas, there is strong spatially-varying antenna gain pattern that makes cancellation of clutter difficult. Calibration is generally not a trivial process, but to make fair comparisons to the DPCA and ATI algorithms, calibration in (f) and (g) is done by using the estimated coefficients $\boldsymbol{H}_{f,i}$ from the Bayes SAR algorithm. In (e) and (f), the outputs of the DPCA algorithm are applied to the original images (all antennas) and the calibrated images (all antennas), respectively. It should be noted that even with calibration, the DPCA outputs contain a huge number of false detections in high clutter regions. Nevertheless, proper calibration enables detection of moving targets that are not easily detected without calibration, as highlighted by the red boxes. Note that the Bayes SAR algorithm provides an output that is sparse, yet does not require tuning of thresholds as required by DPCA and/or ATI.

calibrated images as shown in (f) and (g) as compared to their original versions, (e) and (b), respectively. Furthermore, the proposed approach also provides a sparse output without choosing thresholds as required by DPCA and ATI. Note that in this figure, calibration is accomplished by using the outputs $\boldsymbol{H}_{f,i}$ from the Bayes SAR approach.

Figure 6 display the detection performance over two smaller scenes of size 125m by 125m as a function of magnitude and phase. For each scene, images are provided for sequential scenes separated by 0.5 seconds. Scene 1 contains strong clutter in the upper left region, while Scene 2 has relatively little clutter. It is seen that the proposed approach (2nd and 3rd columns) provides a sparse solution containing the targets of interest in each of the 4 images. Moreover, the 2nd column provides the estimated probability that a target occupies a given pixel, in comparison to the (0,1) output of DPCA and ATI. Although most estimated probabilities are near 1, there are a few cases where this is not the situation: in scene 2(d), a low-magnitude target is detected with low probability in the lower-right; in scene 1(b) a few target pixels from the clutter region are detected with low probability. In contrast, the performance of DPCA and ATI depend strongly on the threshold. In (a-c), the DPCA-only output provides a large number of false alarms. It is seen that the ATI/DPCA combination with 15 dB magnitude threshold over-sparsifies the solution, missing targets in (b), (c), and (d). On the other hand, the ATI/DPCA combination with 30 dB magnitude threshold detects these targets, but also includes numerous false alarms in (a) and (b). On the other hand, the proposed approach is able to detect the

targets with high fidelity regardless of the scene/image and does not require tuning of thresholds for detection.

2) Target motion models: Figure 7 shows the output of the proposed approach when prior information on the location of the targets might be available. For example, in the shown scene, targets are likely to be stopped at an intersection. The performance improvement is given for a mission scene that contains target in this high probability region. On the other hand, there are no significant performance decreases in the reference scene that does not contain targets in the intersection region. This type of processing could be extended to a tracking environment, where targets are projected to likely be in a given location within the formed SAR image as discussed in Section V.

3) Estimation of radial velocity: The dataset used in this section contained a few GPS-truthed vehicles from which we can derive (a) the 'true' location of the target within the formed SAR image, and (b) the target's radial velocity which is known to be proportional to the measured interferometric phase of the target pixels in an along-track system. To account for uncertainty in target location from the GPS sensor, we consider a 'confidence region' for pixels that have high probability of containing a target. Within these regions, each algorithm (e.g., Bayes SAR or ATI/DPCA) (a) detects pixels containing targets and (b) subsequently estimates the interferometric phase of those pixels. Note that the radial velocity is proportional to the interferometric phase up to an ambiguity factor (i.e. between 0 and 2π) which corresponds to about 7m/s. To avoid this ambiguity, each algorithm provides the radial velocity that
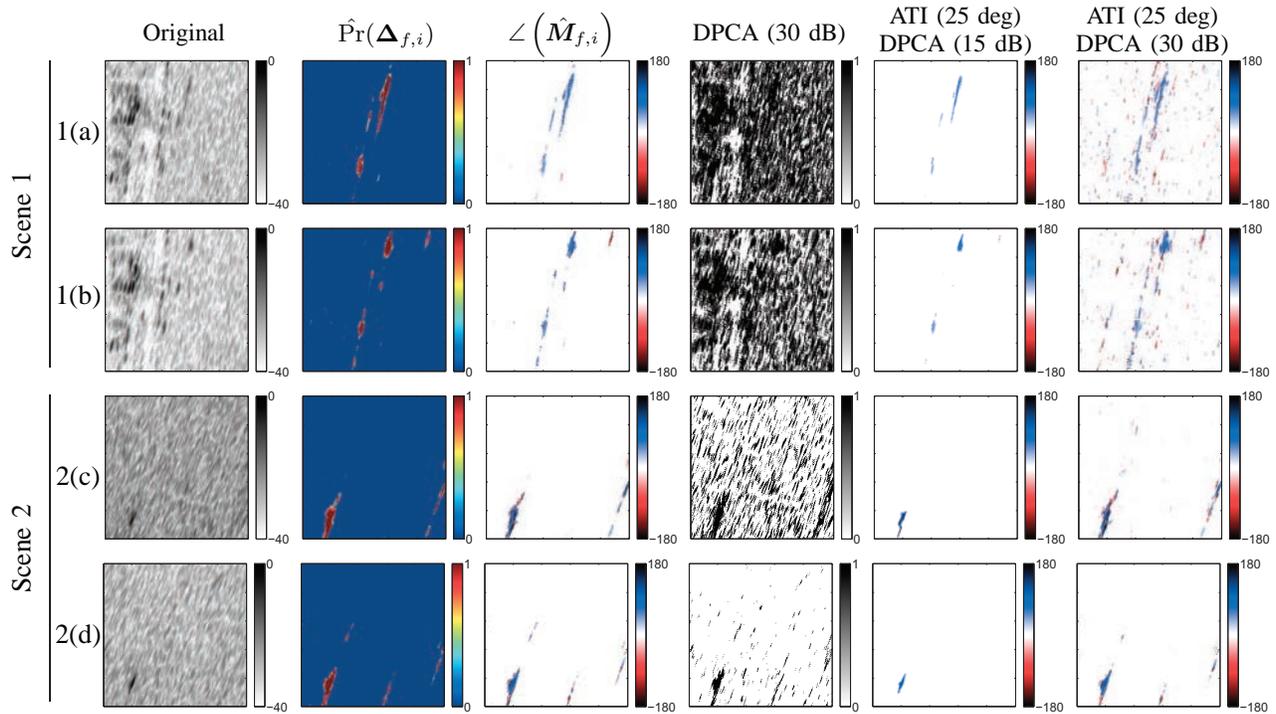
Fig. 6. This figure shows detection performance based on the magnitude/phase of the target response with comparisons between the proposed algorithm and displaced phase center array (DPCA) processing, and a mixture algorithm between DPCA and along-track interferometry (ATI). Note that DCPA and ATI declare detections if the test statistic (magnitude for DPCA and phase for ATI) are than some threshold. Results are given for two scenes of size 125m x 125m; within each scene, images were formed for two sequential 0.5 second intervals. Scene 1 contains strong clutter in the upper left region, while Scene 2 has relatively little clutter. The columns of the figure provide from left-to-right: the magnitude of the original image, the estimated probability of the target occupying a particular pixel (Bayes SAR), the estimated phase of the targets (Bayes SAR), the output of DPCA with a relative threshold of 30 dB, the output of ATI/DPCA with (25 deg, 15 dB) thresholds, and the output of ATI/DPCA with (25 deg, 30 dB) thresholds. It is seen that without phase information to cancel clutter, DPCA (30 dB) contains an overwhelming number of false alarms for scenes (a-c), although the performance is reasonable for scene (d). The ATI/DPCA algorithms provide sparser solutions by canceling the strong clutter. It is seen that the ATI/DPCA combination with 15 dB magnitude threshold over-sparsifies the solution, missing targets in (b), (c), and (d). On the other hand, the ATI/DPCA combination with 30 dB magnitude threshold detects these targets, but also includes numerous false alarms in (a) and (b). On the other hand, the proposed algorithm provides a sparse solution that detects all of these targets, while simultaneously providing a estimate of the probability of detection rather than an indicator output.
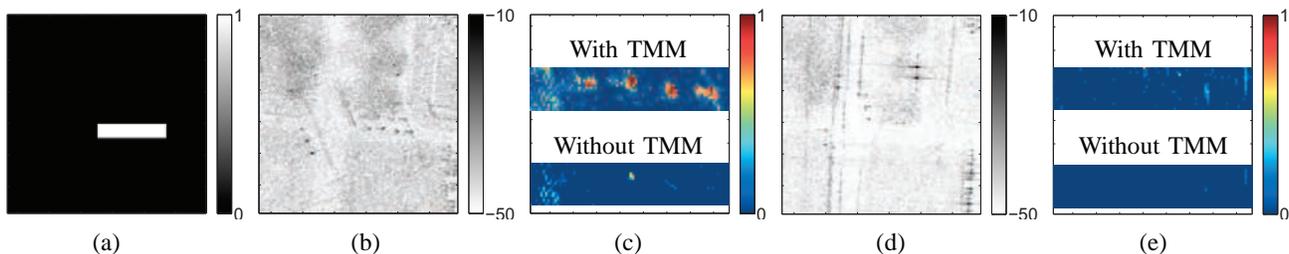
Fig. 7. This figure compares the performance of our proposed method with and without priors on target signature locations. In this scene, targets are likely to be stopped at an intersection as shown by the region in (a). A mission image containing targets is shown in (b) and a reference image without targets is shown in (d). The estimated target probabilities are shown in (c) for the mission scene where inference was done both with/without a target motion model (TMM). It can be seen that by including the prior information, we are able to detect stationary targets that cannot be detected from standard SAR moving target indication algorithms. The estimated target probabilities in the reference scene are shown in (e), showing little performance differences when prior information is included in the inference.
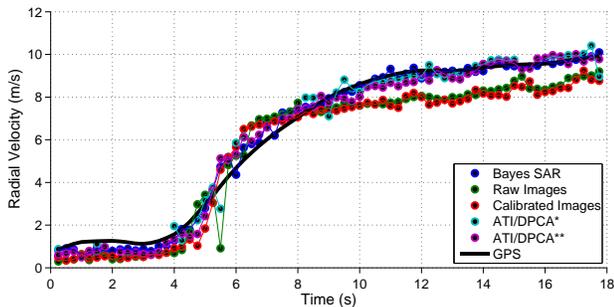
Fig. 8. This figure plots the estimated radial velocities (m/s) for a single target from measured SAR imagery over 18 seconds at 0.25 second increments. Radial velocity, which is proportional to the interferometric phase of the pixels from multiple antennas in an along-track SAR system, is estimated by computing the average phase of pixels within a region specified by the GPS-given target state (position, velocity). We compare the estimation of radial velocity from the output of the Bayes SAR algorithm, from the raw images, from the calibrated images (i.e, using the estimated calibration coefficients), and from two DPCA/ATI joint algorithms as described by Deming [8] with phase/magnitude thresholds of (25 deg, 15 dB) and (25 deg, 30 dB) respectively. For fair comparisons, the DPCA/ATI thresholds are applied to the calibrated imagery, though this is a non-trivial step in general. The black line provides the GPS-truth.

TABLE V
RADIAL VELOCITY ESTIMATION (M/S) IN MEASURED SAR DATASET. THE PROPOSED ALGORITHM (BAYES SAR) HAS LOWER BIAS AND MSE, AS WELL AS FEWER MISSED TARGETS AS COMPARED TO ALL OTHER ALTERNATIVES. MOREOVER, ALL ALGORITHMS EXCEPT 'RAW' REQUIRE ADDITIONAL CALIBRATIONS BETWEEN ANTENNAS, EXCEPT THE PROPOSED ALGORITHM WHICH ESTIMATES CALIBRATION CONSTANTS SIMULTANEOUSLY WITH THE TARGET RADIAL VELOCITY. ALSO, THE PROPOSED ALGORITHM HAS NEARLY APPROXIMATELY HALF THE ERROR OF THE ATI/DPCA ALGORITHMS WITHOUT REQUIRING TUNING OF THRESHOLDS.

Algorithm	Bias	MSE	No. Missed
Raw	0.56	0.86	7
Calibrated	0.60	0.91	0
Bayes SAR	0.11	0.16	0
ATI/DPCA*	-0.06	0.32	57
ATI/DPCA**	0.17	0.24	5

is closest to the true radial velocity (among all ambiguous choices). Note that given the region of test pixels, detection of target pixels and estimation of the interferometric phase are done independently of knowledge of the true state.

Figure 8 shows the estimated radial velocities for a single target over 18 seconds at 0.25 second increments. We compare the estimation of radial velocity from the output of the Bayes SAR algorithm, from the raw images, from the calibrated images, and from two DPCA/ATI joint algorithms as described by Deming [8] with phase/magnitude thresholds of (25 deg, 15 dB) and (25 deg, 30 dB) respectively. For fair comparisons, the DPCA/ATI thresholds are applied to the calibrated imagery, though this is a non-trivial step in general. Numerical results are summarized in Table V. It is seen that the Bayes SAR algorithm outperforms the others in terms of MSE for both targets. Moreover, the Bayes SAR algorithm never misses a target detection in this dataset, which is not the case for the DPCA/ATI algorithms. Moreover, while the calibrated images also never miss the target, there was significant bias and MSE due to the inclusion of pixels that are ignored by the Bayes

SAR and ATI/DPCA algorithms.

VIII. DISCUSSION AND FUTURE WORK

Recent work [1]–[3] has shown that it is possible to successfully decompose natural high-dimensional signals/images into low-rank and sparse components in the presence of noise, leading to the so-called robust principal component analysis algorithms. [4] introduced a Bayesian formulation of the problem that built on the success of these algorithms with the additional benefits of (a) robustness to unknown densely distributed noise with noise statistics that can be inferred from the data, (b) convergence speeds in real applications of the mean solution that are similar to those of the optimization-based procedures, and (c) characterization of the uncertainty (i.e., estimates of the posterior distribution) that could lead to improvements in subsequent inference. Moreover, the Bayesian formulation is shown to be capable of generalization to cases where additional information is available, e.g. spatial/Markov dependencies.

SAR imagery collected from a staring sensor across multiple passes, frames, and receive channels contains significant amounts of redundant information, which suggests that a low-dimensional representation for the clutter could be exploited to improve GMTI algorithms. Indeed, algorithms such as STAP already use a low-rank assumption in order to cancel clutter. On the other hand, these algorithms depend on the availability of homogeneous target-free data, thresholds for phase/magnitude-based detection which may vary across the scene, and appropriate calibration across receive channels. Moreover, SAR-specific phenomena such as complex-valued images, glints and speckle noise make it difficult to apply the previously developed RPCA methods for SAR GMTI.

This work provides a Bayesian formulation similar to [4] that (a) directly accounts for SAR-specific phenomena, (b) includes information available from staring SAR sensors (multi-pass, multi-frame, and multi-antenna), and (c) characterizes uncertainty by yielding a posterior distribution on the variables of interest given the observed SAR images. Similar to Ding et al. [4], this algorithm requires few tuning parameters since most quantities of interest are inferred directly from the data - this allows the algorithm to be robust to a large collection of operating conditions. Moreover, the performance of the proposed approach is analyzed over both simulated and measured datasets, demonstrating competing or better performance than the RPCA algorithms and ATI/DPCA.

There are several research directions which could be used in order to improve the methods described in this paper. First, the statistical model presented in this paper chose prior distributions, such as the Multivariate-Normal-Inverse-Wishart and Bernoulli-Beta distributions, for numerical efficiency. In practice, this worked reasonably well on the measured dataset. Moreover, these distributions have been applied to other Bayesian modeling problems [4], [5]. Nevertheless, one may wish to understand the sensitivity to model mismatch by analyzing performance over larger datasets. Additionally, future work could explore the tradeoff between model fidelity and computational burden. This could include generalizations

to the model, such as complex target maneuvers, multiple target classes, and explicit tracking of the target phase, as well as physical models such as multiplicative, rather than additive, speckle noise.

This method provides a rich model that can combine spatial, temporal, and kinematic information as well as infer nuisance parameters such as clutter distributions and antenna calibration errors. Nevertheless, this framework comes at the expense of significant computational burden, especially as compared to methods such as DPCA and ATI. The inference algorithm is designed explicitly to be highly parallelizable and future work should explore ways to utilize this property in order to efficiently estimate the posterior distribution.

Finally, future work will include the development of algorithms that exploit the use of a posterior distribution for improved performance in a signal processing task, e.g. detection, tracking or classification. In particular, we are interested in using algorithms for simultaneously detecting and estimating targets over a sparse scene with resource constraints, as well determining the fundamental performance limits of a SAR target tracking system.

REFERENCES

[1] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma, "Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization," *submitted to Journal of the ACM*, 2009.

[2] Z. Lin, M. Chen, and Y. Ma, "The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices," *arXiv preprint arXiv:1009.5055*, 2010.

[3] E. J. Candes, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM*, vol. 58, no. 3, 2011.

[4] X. Ding, L. He, and L. Carin, "Bayesian robust principal component analysis," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3419–3430, 2011.

[5] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *The Journal of Machine Learning Research*, vol. 1, pp. 211–244, 2001.

[6] 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, August 2004.

[7] J. R. Fienup, "Detecting moving targets in SAR imagery by focusing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 37, no. 3, pp. 794–809, 2001.

[8] R. W. Deming, "Along-track interferometry for simultaneous SAR and GMTI: application to Gotcha challenge data," in *Proceedings of SPIE*, vol. 8051, 2011, p. 80510P.

[9] R. Deming, S. MacIntosh, and M. Best, "Three-channel processing for improved geo-location performance in SAR-based GMTI interferometry," in *Proceedings of SPIE*, vol. 8394, 2012, p. 83940F.

[10] K. I. Ranney and M. Soumekh, "Signal subspace change detection in averaged multilook SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 1, pp. 201–213, 2006.

[11] M. Soumekh, "Moving target detection in foliage using along track monopulse synthetic aperture radar imaging," *IEEE Transactions on Image Processing*, vol. 6, no. 8, pp. 1148 –1163, Aug. 1997.

[12] J. H. G. Ender, "Space-time processing for multichannel synthetic aperture radar," *Electronics & Communication Engineering Journal*, vol. 11, no. 1, pp. 29–38, 1999.

[13] J. K. Jao, "Theory of synthetic aperture radar imaging of a moving target," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 9, pp. 1984–1992, 2001.

[14] B. H. Borden and M. L. Mumford, "A statistical glint/radar cross section target model," *IEEE Transactions on Aerospace and Electronic Systems*, no. 5, pp. 781–785, 1983.

[15] F. L. Posner, "Texture and speckle in high resolution synthetic aperture radar clutter," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31, no. 1, pp. 192–203, 1993.

[16] R. K. Raney and G. J. Wessels, "Spatial considerations in SAR speckle consideration," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 5, pp. 666–672, 1988.

[17] L. Borcea, T. Callaghan, and G. Papanicolaou, "Synthetic aperture radar imaging and motion estimation via robust principle component analysis," *arXiv preprint arXiv:1208.3700*, 2012.

[18] A. S. Khwaja and J. Ma, "Applications of compressed sensing for SAR moving-target velocity estimation and image compression," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 8, pp. 2848–2860, 2011.

[19] R. P. Perry, R. C. Dipietro, and R. L. Fante, "SAR imaging of moving targets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 1, pp. 188–200, 1999.

[20] S. Zhu, G. Liao, Y. Qu, Z. Zhou, and X. Liu, "Ground moving targets imaging algorithm for synthetic aperture radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 1, pp. 462–477, 2011.

[21] B. Guo, D. Vu, L. Xu, M. Xue, and J. Li, "Ground moving target indication via multichannel airborne SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 3753 –3764, Oct. 2011.

[22] J. Guo, Z. F. Li, and Z. Bao, "Adaptive clutter suppression and resolving of velocity ambiguities for an experimental three-channel airborne synthetic aperture radar-ground moving target indication system," *Radar, Sonar & Navigation, IET*, vol. 5, no. 4, pp. 426–435, 2011.

[23] C. H. Gierull, "Statistical analysis of multilook SAR interferograms for CFAR detection of ground moving targets," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 4, pp. 691–701, 2004.

[24] G. E. Newstadt, E. G. Zelnio, and A. O. Hero, III, "Moving target inference with hierarchical Bayesian models in synthetic aperture radar imagery," *Arxiv preprint arXiv:1302.4680*, Feb. 2013.

[25] C. P. Robert and G. Casella, *Monte Carlo statistical methods*. Citeseer, 2004, vol. 319.

[26] S. M. Scarborough, C. H. Casteel Jr, L. Gorham, M. J. Minardi, U. K. Majumder, M. G. Judge, E. Zelnio, M. Bryant, H. Nichols, and D. Page, "A challenge problem for SAR-based GMTI in urban environments," E. G. Zelnio and F. D. Garber, Eds., vol. 7337, no. 1. SPIE, 2009, p. 73370G. [Online]. Available: http://link.aip.org/link/?PSI/7337/73370G/1

Gregory Newstadt received the B.S. degrees (summa cum laude) from Miami University, Oxford, OH, in 2007 in Electrical Engineering and in Engineering Physics. He also received the M.S.E in Electrical Engineering: Systems (2009), M.A. in Statistics (2012) and Ph.D. in Electrical Engineering: Systems (2013) degrees from the University of Michigan, Ann Arbor, MI.

He is currently a postdoctoral researcher and lecturer at the University of Michigan, Ann Arbor, MI, in Electrical Engineering (Systems). His research interests include detection, estimation theory, target tracking, sensor fusion, and statistical signal processing.

Edmund Zelnio graduated from Bradley University, Peoria, Illinois, in 1975 and has pursued doctoral studies at The Ohio State University in electromagnetics and at Wright State University in signal processing. He has had a 37 year career with the Air Force Research Laboratory (AFRL), Wright Patterson AFB, Ohio where he has spent 35 years working in the area of automated exploitation of imaging sensors primarily addressing synthetic aperture radar. He is a former division chief and technical advisor of the Automatic Target Recognition Division of the Sensors Directorate in AFRL and serves in an advisory capacity to the Department of Defense and the intelligence community. He is currently the director of the Automatic Target Recognition Center in AFRL. He is the recipient of the 53rd DoD Distinguished Civilian Service Award and is a fellow of the Air Force Research Laboratory.

Alfred O. Hero, III received the B.S. (summa cum laude) from Boston University (1980) and the Ph.D from Princeton University (1984), both in Electrical Engineering. Since 1984 he has been with the University of Michigan, Ann Arbor, where he is the R. Jamison and Betty Professor of Engineering. His primary appointment is in the Department of Electrical Engineering and Computer Science and he also has appointments, by courtesy, in the Department of Biomedical Engineering and the Department of Statistics. In 2008 he was awarded the the Digiteo Chaire d'Excellence, sponsored by Digiteo Research Park in Paris, located at the Ecole Superieure d'Electricite, Gif-sur-Yvette, France. He has held other visiting positions at LIDS Massachussets Institute of Technology (2006), Boston University (2006), I3S University of Nice, Sophia-Antipolis, France (2001), Ecole Normale Supérieure de Lyon (1999), Ecole Nationale Supérieure des Télécommunications, Paris (1999), Lucent Bell Laboratories (1999), Scientific Research Labs of the Ford Motor Company, Dearborn, Michigan (1993), Ecole Nationale Superieure des Techniques Avancees (ENSTA), Ecole Superieure d'Electricite, Paris (1990), and M.I.T. Lincoln Laboratory (1987 - 1989).

Alfred Hero is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE). He has been plenary and keynote speaker at major workshops and conferences. He has received several best paper awards including: a IEEE Signal Processing Society Best Paper Award (1998), the Best Original Paper Award from the Journal of Flow Cytometry (2008), and the Best Magazine Paper Award from the IEEE Signal Processing Society (2010). He received a IEEE Signal Processing Society Meritorious Service Award (1998), a IEEE Third Millenium Medal (2000) and a IEEE Signal Processing Society Distinguished Lecturership (2002). He was President of the IEEE Signal Processing Society (2006-2007). He sits on the Board of Directors of IEEE (2009-2011) where he is Director Division IX (Signals and Applications).

Alfred Hero's recent research interests have been in detection, classification, pattern analysis, and adaptive sampling for spatio-temporal data. Of particular interest are applications to network security, multi-modal sensing and tracking, biomedical imaging, and genomic signal processing.