

Sub-aperture SAR Imaging with Uncertainty Quantification

Victor Churchill* Anne Gelb[†]

Abstract

In the problem of spotlight mode airborne synthetic aperture radar (SAR) image formation, it is well-known that data collected over a wide azimuthal angle violate the isotropic scattering property typically assumed. Many techniques have been proposed to account for this issue, including both full-aperture and sub-aperture methods based on filtering, regularized least squares, and Bayesian methods. A full-aperture method that uses a hierarchical Bayesian prior to incorporate appropriate speckle modeling and reduction was recently introduced to produce samples of the posterior density rather than a single image estimate. This uncertainty quantification information is more robust as it can generate a variety of statistics for the scene. As proposed, the method was not well-suited for large problems, however, as the sampling was inefficient. Moreover, the method was not explicitly designed to mitigate the effects of the faulty isotropic scattering assumption. In this work we therefore propose a new sub-aperture SAR imaging method that uses a sparse Bayesian learning-type algorithm to more efficiently produce approximate posterior densities for each sub-aperture window. These estimates may be useful in and of themselves, or when of interest, the statistics from these distributions can be combined to form a composite image. Furthermore, unlike the often-employed ℓ_p -regularized least squares methods, no user-defined parameters are required. Application-specific adjustments are made to reduce the typically burdensome runtime and storage requirements so that appropriately large images can be generated. Finally, this paper focuses on incorporating these techniques into SAR image formation process, that is, for the problem starting with SAR phase history data, so that no additional processing errors are incurred. The advantage over existing SAR image formation methods are clearly presented with numerical experiments using real-world data.

Keywords: synthetic aperture radar, sub-aperture image recovery, generalized sparse Bayesian learning, uncertainty quantification.

1 Introduction

Spotlight mode airborne synthetic aperture radar (SAR)¹ is a widely-used technology for surveillance and mapping due to its all-weather day-or-night imaging capabilities that

*Department of Mathematics, The Ohio State University, Columbus, OH, 43201 USA. E-mail: churchill.77@osu.edu

[†]Department of Mathematics, Dartmouth College, Hanover, NH, 03755 USA. Email: an-negelb@math.dartmouth.edu.

¹To be concise, we will refer to spotlight mode airborne SAR as just SAR from now on.

make it favorable over optical imaging modalities in remote sensing, [1, 21, 35]. The data are typically collected in a circular flight pattern around the scene of interest and then processed to produce images that are used for purposes such as classification and identification. It is therefore critical for practitioners to be able to rely on SAR images that clearly localize targets.

Challenges in SAR imaging

There are several well-known challenges in SAR imaging, which we discuss below, along with the existing state of the field on these issues.

- **Isotropic scattering assumption:** While it is straightforward to directly process the circularly-collected 360° -azimuth SAR phase history data into a single image estimate, i.e. so-called full-azimuth or full-aperture imaging, it forces the often used yet incorrect assumption that the imaging scene contains only isotropic scatterers. There is angular dependence of the scattered response, however. In particular, since the energy is scattered off objects in the front of the scene, then it does not mix evenly with the full scene. So-called sub-aperture techniques, where the full azimuth data are grouped into smaller overlapping sub-aperture windows and then processed separately for image recovery are often employed to alleviate this problem. Each image recovery process typically employs filtering, regularized least squares, ℓ_1 regularization, or Bayesian methods, followed by some mechanism (e.g. weighted averages) to aggregate the results from each sub-aperture recovery to form a final image. A good review of sub-aperture techniques may be found in [4], with specific examples in [15, 17, 44, 49, 50, 55, 60].
- **Speckle:** SAR is a coherent imaging system, meaning that both the collected data and the reflectivity image are complex-valued. While typically only the magnitude is viewed, the phase information should not be neglected in the image formation process, and is important for downstream tasks such as target position and amplitude estimation, [43, 42]. All coherent imaging modalities are affected by speckle, a multiplicative-noise-like phenomenon which, although is in fact signal, causes grainy-looking images and hence it is desirable to remove it. Existing methods for SAR image reconstruction from phase history data usually do not directly address speckle and post-processing operations like smoothing and filtering are typically necessary, see e.g. [3, 22, 25, 26, 39, 48]. Basic, fast methods that rely on an inverse non-uniform fast Fourier transform (NUFFT), [33], provide no speckle reduction, while sparsity-based methods that rely on ℓ_p regularization in some domain² [2, 14, 16, 17, 58, 57, 50, 53], disregard the physical meaning of speckle and instead choose to place penalties on approximate pixel magnitude values. Conflating speckle with the usual additive noise in this way makes parameter selection for the ℓ_1 regularization penalty term very difficult (and essentially without physical meaning) in practice, although some methods have tried to tackle estimating this parameter [10]. Recently, [24] incorporated the fully-developed speckle model as an image prior within a hierarchical Bayesian framework, and in doing so properly characterized the speckle as part of the signal, which could then be appropriately

²Speckle reduction is often also addressed using denoising techniques such as total variation (TV) regularization, either used directly or through a joint sparsity scheme [50, 52, 55].

reduced, for example using a sparse Bayesian learning (SBL) technique [65]. The numerical tests performed in [24] confirmed that using this model yielded improved speckle reduction over commonly used compressive sensing techniques while also providing uncertainty quantification (see below). However, due to long run-time, this method did not consider sub-aperture imaging and the correspondingly required processing of more than one aperture window.

- **Uncertainty quantification:** For the most part the existing sub-aperture and full azimuth SAR image formation methods produce a single image product, typically a maximum likelihood or maximum *a posteriori* point estimate, that approximates the unknown ground truth. Predictions from these methods are not probabilistic, and therefore provide no information about the statistical confidence with which we can trust the features in the resulting images, e.g., which are more likely actual objects in the scene and which are more likely attributed to speckle or noise. This makes forming reliable images difficult, particularly in SAR where even many synthetically-created examples have unknown true reflectivity.³ In [24], an MCMC-based method was used to sample the posterior density of a hierarchical Bayesian model for full-azimuth SAR image reconstruction from phase history data, giving much more robust information about the distribution of the image to affirm the confidence with which each estimate can be trusted. Sampling-based methods using this same prior structure have been developed to quantify uncertainty in basic real-valued linear inverse problems such as image reconstruction, see e.g. [9], and have also been applied to SAR imaging tasks such as moving target inference, [46], passive SAR image reconstruction, [69], and speckle noise model selection, [38]. In addition, several SAR image reconstruction methods [29, 68, 71, 72] have leveraged SBL estimation procedures, [65], or Bayesian compressed sensing [36] to analyze the same posterior.

The purpose of this paper therefore is to address these challenges by introducing an efficient sub-aperture SAR imaging method that models and reduces speckle and returns an approximate posterior density, allowing for uncertainty quantification. This will be accomplished by adapting the speckle characterization and reduction techniques, along with the uncertainty quantification, of the method in [24]. Significantly, we are able to address these challenges *within* the process of reconstructing SAR images directly from phase history data, as opposed to relying on altering images that have already been reconstructed or otherwise processed.

Contributions

This investigation provides several significant contributions to the SAR image formation literature, specifically to sub-aperture imaging. First, to the best of our knowledge, sub-aperture approaches for wide-angle SAR imaging have previously been limited to point estimates. To this end, our proposed method provides approximate posterior densities for each sub-aperture image based on a deterministic estimation procedure [32] to analyze the constructed hierarchical Bayesian prior. Second, while many sub-aperture methods have been shown to yield highly resolved SAR images, [4, 15, 17, 44,

³We point out that there are some instances for which benchmarks for despeckling have been established, [27]. However, these tests operate on simulated magnitude-only images that have already been formed, while the focus of this paper is on reconstructing complex-valued images from real-world phase history data.

49, 50, 55, 60], whether through a direct prior on each sub-aperture image or a joint sparsity approach, these techniques all require some user selected thresholds and/or parameters, which makes them less robust to modifications within the data. By contrast, the proposed method is automated, meaning that it does not require any parameter tuning for different data sets. While some methods have been proposed to aid in selection of the regularization parameters, e.g. [10, 59], the task remains challenging in practice and at the very least requires some extra processing. The method we propose is also able to incorporate TV regularization (as well as higher order TV, and most other regularization operators, e.g. wavelets) into its hierarchical Bayesian prior. The ability to regularize in other domains illustrates the general flexibility of the Bayesian model, whereby additional priors can be added to account for a variety of constraints, such as phase error correction [56, 67, 58, 62, 47, 70]. To the best of our knowledge, this is the first time an SBL-type prior has been used with TV for SAR image formation. Third, we demonstrate that this novel approach is also significantly faster than other sparsity (ℓ_1 -regularized least squares) methods. Furthermore, among other subjective advantages in the actual appearance of the image, our technique increases image contrast and reduces speckle. Finally, a challenging real-world data example is considered to demonstrate the new methodology, and support its use. Real-world SAR image formation is a very large problem, and we make application-specific efficient storage and runtime adjustments due to the prohibition of traditional matrix-based methods for linear inverse problems, as even storing dense matrices of the necessary size is problematic. Parallel processing of the sub-aperture images is also included in the implementation.

Organization

The rest of this paper is organized as follows. Section 2 gives necessary background on SAR imaging and existing estimation techniques. Section 3 derives the hierarchical Bayesian prior of [65] from scratch, emphasizing the incorporation of coherent imaging, speckle, and sparsity using conjugate priors, as in [24], and then explains the new sub-aperture imaging algorithm and its properties. Section 4 shows a real-world example using the Air Force Research Laboratory’s GOTCHA Volumetric Data Set 1.0, [13]. Some concluding remarks and ideas for future directions are provided in Section 5.

2 Background

As noted in Section 1, although it is straightforward to directly process the circularly-collected 360° -azimuth SAR phase history data into a single full-azimuth image estimate, it forces the often used yet incorrect assumption that the imaging scene contains only isotropic scatterers. By using opposite halves of the full azimuth to recover the same underlying scene, Figure 1 demonstrates that this assumption is clearly false. It is reasonable to assume, however, that the information received is the same within a small aperture window. Therefore the *sub-aperture imaging approach*, [16, 44, 55], which forms a SAR image by combining a finite number of individual reconstructions coming from small overlapping aperture windows of data, helps to mitigate the effects of conflicting information resulting from the anisotropic nature of the scatterers. Observe in Figure 2(right) that simply applying a non-uniform Fourier transform (NUFFT) to each sub-aperture window and then combining the results by taking the maximum modulus over

the sub-apertures to form a single image clearly yields more localized features.

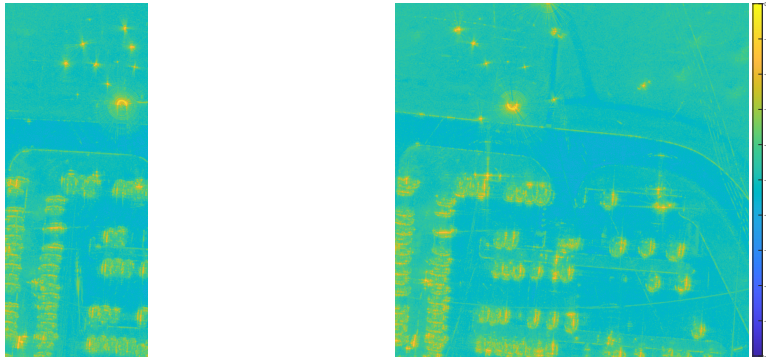


Figure 1: Parking lot images formed using opposite halves of full azimuth coverage (2048×2048 pixels, NUFFT image formation).

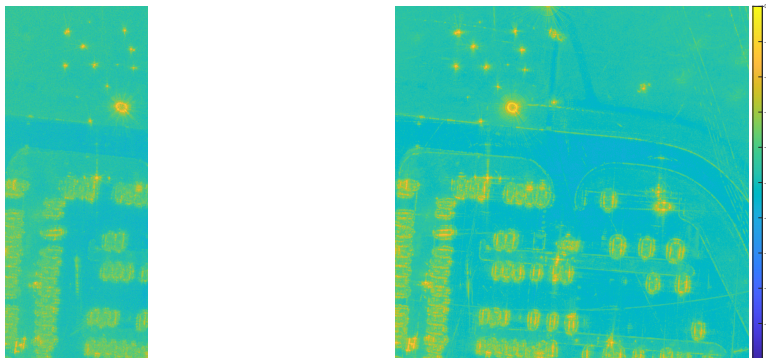


Figure 2: Parking lot images formed via (left) full-aperture and (right) sub-aperture methods (2048×2048 pixels, NUFFT image formation with 60 sub-apertures each spanning 8° with 2° of overlap).

2.1 SAR modeling

The ensuing description of the SAR image formation inverse problem below generally follows the development in [35, 58]. Let $f : \Omega \rightarrow \mathbb{C}$ denote the two-dimensional reflective scene of scattering objects that we want to recover, where f is defined over

$$\Omega := \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \leq U^2\}.$$

At a particular position in the sensing process, indicated by an azimuth angle $\theta \in [0, 2\pi]$, the transmitted linear FM chirp mixes with the scene in a way that depends upon θ , the angle from which the chirp is emitted.⁴ In the far field case, once the transmitted signal reaches the scene it has essentially a planar wave front, and thus the points in the scene along each line perpendicular to the direction of the chirp all mix with the same values. Hence the two-dimensional setup is often simplified to a one-dimensional

⁴In practice there is also a relevant angle of elevation which is not critical to the development of our method.

process by compressing the scatterers along each of these lines to a single point. This compression is commonly referred to as the projection or Radon transform of f at the angle θ , and is denoted $p : \Omega \rightarrow \mathbb{R}^2$. It can be expressed mathematically as

$$p(\theta, u) = \iint_{x^2+y^2 \leq U^2} f(x, y) \delta(u - x \cos \theta - y \sin \theta) dx dy, \quad (1)$$

where δ is the Dirac delta function and $u \in \mathbb{R}$ is the slant range position.

The linear FM chirp that is transmitted and mixed with the scene is the real part of

$$s(t) = \begin{cases} e^{i(\omega t + \alpha t^2)}, & |t| \leq \frac{T}{2} \\ 0, & \text{otherwise} \end{cases}, \quad (2)$$

where ω is the carrier frequency, 2α is the chirp rate, and $T > 0$ is the pulse duration. This chirp signal mixes with the scene to yield reflected signals of the form

$$r(\theta, t) = \int_{-U}^U \text{Re} \{p(\theta, u) s(t - \tau_0 - \tau(u))\} du, \quad (3)$$

where τ_0 is the round trip time required for the chirp to travel to the scene center and $\tau(u)$ is the additional travel time for any particular position in the scene u . If R is the distance from the transmitter/receiver to the scene center and c is the speed of light in a vacuum, we have $\tau_0 = 2R/c$ and $\tau(u) = 2u/c$.

A deramping process is implemented to extract approximate instantaneous frequency information (i.e. the classical Fourier transform of f) from the chirp response, ultimately yielding the approximation⁵

$$\hat{f}_\theta(t) := \int_{-U}^U p(\theta, u) e^{-iku} du \approx r(\theta, t), \quad (4)$$

where the spatial frequencies k , measured in cycles/meter, are given by

$$k = k(t) := \frac{2}{c}(\omega + 2\alpha(t - \tau_0)). \quad (5)$$

This approximation makes several assumptions about the transmitted signal, the scene, and how accurately the data are measured. For example, as already noted, in far field spotlight SAR, the distance R from the transmitter/receiver to the scene center is far enough so that the transmitted signal has essentially a planar wave front (i.e. any curvature of the scene geometry can be safely ignored). Moreover, the chirp rate α is small enough so that higher order terms can be ignored in the approximation of (4). A third assumption often made is that the round trip propagation time is exactly known. This is generally not the case, and this lack of precision causes a phase error in the signal recovery. Autofocusing is commonly employed to reduce the shearing effect caused by the phase error. We do not discuss the ramifications of phase error in the current investigation, but point readers to [5, 58, 66, 56, 67, 58, 62, 47, 70] for general information on autofocusing techniques. Finally, (5) assumes that the scene scatterers in spotlight SAR are isotropic, meaning that the reflection is the same regardless of the azimuth angle.

⁵More details may be found in [35, 58].

Figures 1 and 2 respectively demonstrate the inaccuracy in the image recovery due to this faulty assumption and how using a sub-aperture approach helps to alleviate this issue.

The projection slice theorem [35] is used to conveniently rewrite (4) as

$$\hat{f}_\theta(t) = \int_{-U}^U p(\theta, u) e^{-iku} du = \iint_{x^2+y^2 \leq U^2} f(x, y) e^{-ik(x \cos \theta + y \sin \theta)} dx dy. \quad (6)$$

Hence forming an image from SAR phase history data can be modeled as the inverse of a continuous non-uniform Fourier transform.

To discretize the problem, let temporal frequency values be given by t_m for $m = 1, \dots, M$, and a set of azimuth angles by θ_l for each sub-aperture $l = 1, \dots, L$. From (5) we can define

$$k_m = \frac{2}{c}(\omega + 2\alpha(t_m - \tau_0)), \quad m = 1, \dots, M, \quad (7)$$

as the discretized spatial frequencies, leading to the linear system model for each sub-aperture image formation

$$\hat{\mathbf{f}}^{(l)} = \mathbf{F}^{(l)} \mathbf{f}^{(l)} + \mathbf{n}^{(l)}, \quad l = 1, \dots, L. \quad (8)$$

The discrete forward operator $\mathbf{F}^{(l)} : \mathbb{C}^N \rightarrow \mathbb{C}^M$, $l = 1, \dots, L$, is the two-dimensional discrete non-uniform Fourier transform matrix for the sub-aperture data collection modeled by (6) that maps the concatenated l -th sub-aperture reflectivity image $\mathbf{f}^{(l)} \in \mathbb{C}^N$ to the corresponding vertically-concatenated phase history data $\hat{\mathbf{f}}^{(l)} \in \mathbb{C}^M$, where N is the number of pixels in the each image, M is the length of the data in each window, and L is the number of sub-aperture windows.⁶ Note that by using the discrete Fourier transform in (8) we introduce both aliasing error and the Gibbs phenomenon. As these errors are typically smaller in scale than the data acquisition error, they are simply folded into the additive noise in (8) and not separately considered. We also observe that (8) is equivalent to the full-aperture case when $L = 1$, i.e. a single aperture window with full azimuth information.

Since it is reasonable to assume that within a small aperture the scatterers in the scene are isotropic, the method proposed in this investigation uses the sub-aperture imaging approach of dividing the full azimuth data into sub-apertures which may or may not overlap. Finally, we assume in (8) that the model and measurement error $\mathbf{n}^{(l)}$ is complex noise that is circularly-symmetric white Gaussian distributed. That is, $\mathbf{n}_i^{(l)} \sim \mathcal{CN}(0, 1/\beta^{(l)})$ i.i.d. for all pixels i , where $(\beta^{(l)})^{-1} > 0$ is the noise variance. This yields the Gaussian likelihood function

$$p(\hat{\mathbf{f}}^{(l)} | \mathbf{f}^{(l)}, \beta^{(l)}) \propto (\beta^{(l)})^M \exp \left(-\frac{\beta^{(l)}}{2} \|\hat{\mathbf{f}}^{(l)} - \mathbf{F}^{(l)} \mathbf{f}^{(l)}\|_2^2 \right), \quad l = 1, \dots, L, \quad (9)$$

which measures the goodness of fit of the model (8), where $\|\mathbf{g}\|^2 = \mathbf{g}^* \mathbf{g}$ with $\mathbf{g}^*$ the conjugate transpose of $\mathbf{g}$. The initial objective is to infer $\mathbf{f}^{(l)}$ from $\hat{\mathbf{f}}^{(l)}$ in each sub-aperture, and then combine each of the L recovered windowed images into a single image (if desired).

⁶In fact it is possible to have $M = M(l)$, $l = 1, \dots, L$, but for ease of presentation we choose the number of frequencies in each window to remain constant.

2.2 Estimation techniques

One straightforward way to estimate each $\mathbf{f}^{(l)}$ from $\hat{\mathbf{f}}^{(l)}$ is to maximize the likelihood function. From the Gaussian likelihood in (9), this estimate is

$$\begin{aligned}\mathbf{f}_{ML}^{(l)} &= \arg \max_{\mathbf{f}} \left\{ p(\hat{\mathbf{f}}^{(l)} | \mathbf{f}, \beta) \right\} \\ &= \arg \min_{\mathbf{f}} \left\{ \|\hat{\mathbf{f}}^{(l)} - \mathbf{F}^{(l)} \mathbf{f}\|_2^2 \right\}, \quad l = 1, \dots, L.\end{aligned}\tag{10}$$

Assuming approximate orthogonality of $\mathbf{F}^{(l)}$, we have $\mathbf{f}_{ML}^{(l)} \approx (\mathbf{F}^{(l)})^* \hat{\mathbf{f}}^{(l)}$, which only requires an inverse non-uniform fast Fourier transform (NUFFT) application to invert the data, [31, 40, 34]. Specifically, an individual reflectivity image can be found by interpolating the typically polar grid of measured samples in frequency space to an equally spaced rectangular grid, then computing an inverse *uniform* fast Fourier transform, [1]. While the NUFFT has the advantage of being computationally efficient, the noisy data and model error can still degrade image quality. Hence an improvement is generally needed. Each image in Figures 1 and 2 was formed using the above NUFFT method. We highlight the differences between Figure 2(left), which uses $L = 1$ full-azimuth window, and Figure 2(right), which shows the maximum modulus composite image of $L = 60$ sub-aperture windows of 8° with 2° of overlap.

An assumption about sparsity is often used to improve image quality. Indeed, there are many sparsity-based SAR image formation methods, see e.g. [17] for a good overview and [2, 14, 16, 17, 50, 51, 53, 57, 58] for specific examples. Because SAR images are frequently sparse in the image domain (more precisely in the magnitude of the image), an ℓ_1 (or more generally ℓ_p) norm penalty term on the presumed sparse domain of $\mathbf{f}^{(l)}$ is often added to improve on the results obtained using (10). In this regard it is important to recall that each $\mathbf{f}^{(l)}$ is complex, and only its magnitude, $|\mathbf{f}^{(l)}|$, has a corresponding sparse domain, as the phase is not modeled as sparse, [35]. Since $|\mathbf{f}^{(l)}|$ is not differentiable, it is convenient to instead employ a unitary diagonal matrix $\Theta^{(l)}$ such that $(\Theta^{(l)})^* \mathbf{f}^{(l)} \approx |\mathbf{f}^{(l)}|$, where $\Theta_{jj}^{(l)} = \tilde{\mathbf{f}}_j^{(l)} / |\tilde{\mathbf{f}}_j^{(l)}|$ is extracted from some cheaply computed approximation $\tilde{\mathbf{f}}^{(l)}$, such as the NUFFT, [55]. Thus we arrive at

$$\mathbf{f}_{\ell_1}^{(l)} = \arg \min_{\mathbf{f}} \left\{ \frac{\beta^{(l)}}{2} \|\hat{\mathbf{f}}^{(l)} - \mathbf{F}^{(l)} \mathbf{f}\|_2^2 + \lambda \|\mathbf{T}(\Theta^{(l)})^* \mathbf{f}\|_1 \right\}, \quad l = 1, \dots, L, \tag{11}$$

where $\mathbf{T}$ is an appropriate sparsifying operator⁷. The solution is typically found via the alternating direction method of multipliers (ADMM), [11], although other methods are also available. Besides being an inexact representation, approximating $|\mathbf{f}^{(l)}|$ in this way has an important consequence. In particular, since (11) regularizes the sparsity of an approximation to the magnitude (instead of the magnitude itself), the regularization term no longer corresponds to any prior distribution since data are considered in forming the initial estimate. An alternate formulation is to solve for the phase explicitly, e.g. [17]. Nevertheless, for practical purposes such empirically based priors are regularly used [73]. A more commonly discussed difficulty in using (11) is that it often requires fine tuning of the sparsity prior parameter λ and noise variance $1/\beta^{(l)}$ (often combined into a single

⁷In general we can have $\mathbf{T} = \mathbf{T}^{(l)}$, $l = 1, \dots, L$, but choosing sparsifying operators for different sub-apertures may be difficult in practice without prior knowledge of the scene.

parameter). Assuming there is enough information to choose the regularization parameters, see e.g. [10, 59], the ℓ_1 regularization approach, often referred to as compressive sensing, [12], can be a highly effective way to compute a point estimate image for SAR. Figure 3 shows four realizations of (11) using $\mathbf{T} = \mathbf{I}$ (reflecting presumed sparsity in the image magnitude) with $\beta = 1$ and $\lambda = 1/20, 1/40, 1/60, 1/80$. Compared with the NUFFT images, the ℓ_1 method removes a lot of the “background” scattering which may be advantageous. However, it is also apparent that the results are quite sensitive to the choice of the regularization parameter. The images are 2048×2048 pixels and use $L = 12$ windows each spanning 40° with an overlap of 10° . The composite maximum method, as explained in Section 2.3 below, is then used to combine the sub-apertures. Code and parameters (i.e. λ and the number of iterations) for this method are taken from [54, 55].

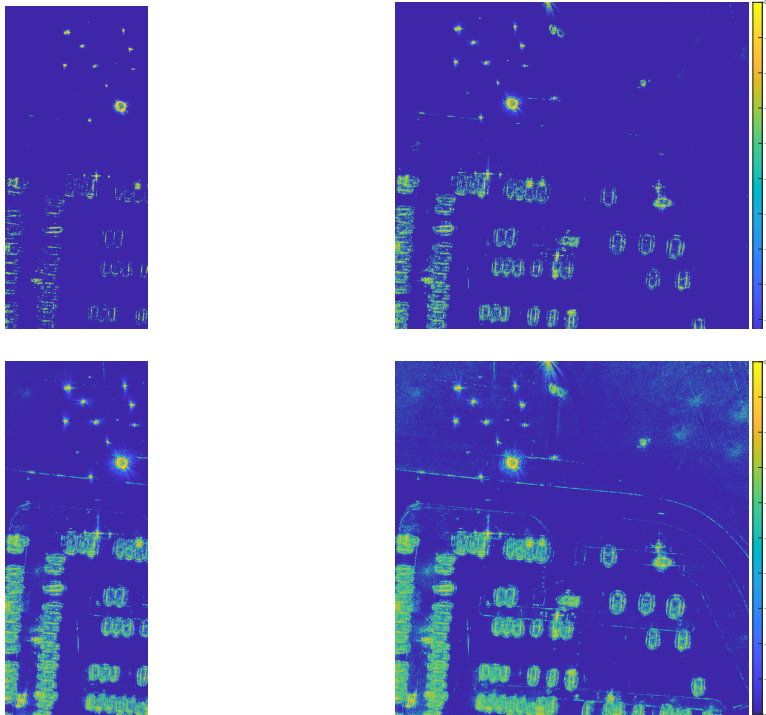


Figure 3: Composite maximum sub-aperture MAP estimate images from the GOTCHA data set, [13], formed with the ℓ_1 method (11) using $\mathbf{T} = \mathbf{I}$ with $\beta = 1$ and $\lambda = 1/20, 1/40, 1/60, 1/80$ (left to right, top to bottom). Each image is obtained using the composite maximum estimate in (14).

Another observation can be made from the construction of (11) from (10). Specifically, had the prior probability distribution

$$p(\mathbf{f}^{(l)}|\lambda) \propto \exp\left(-\lambda\|\mathbf{T}|\mathbf{f}^{(l)}|\|_1\right) \quad (12)$$

been invoked, the resulting posterior would be

$$\begin{aligned} p(\mathbf{f}^{(l)}|\hat{\mathbf{f}}^{(l)}, \beta^{(l)}, \lambda) &\propto p(\hat{\mathbf{f}}^{(l)}|\mathbf{f}^{(l)}, \beta^{(l)})p(\mathbf{f}|\lambda) \\ &\propto \exp\left(-\frac{\beta^{(l)}}{2}\|\hat{\mathbf{f}}^{(l)} - \mathbf{F}^{(l)}\mathbf{f}^{(l)}\|_2^2 - \lambda\|\mathbf{T}|\mathbf{f}^{(l)}|\|_1\right), \quad l = 1, \dots, L. \end{aligned} \quad (13)$$

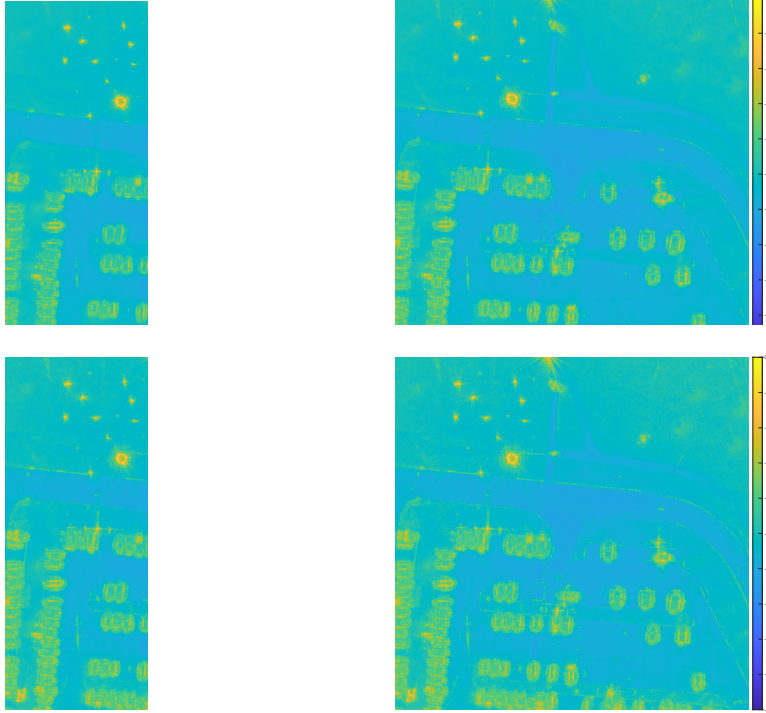


Figure 4: Composite maximum sub-aperture MAP estimate images from the GOTCHA data set, [13], formed with the TV- ℓ_1 method (11) with $\beta = 1$ and $\lambda = 1/40, 1/80, 1/120, 1/160$ (left to right, top to bottom). Each image is obtained using the composite maximum estimate in (14).

Observe that maximizing (13) yields (11), and indeed (11) is known as a maximum a posteriori (MAP) estimate. Of course this is not the only prior distribution that can be used and others would invoke other *a priori* beliefs. This discussion also makes clear that the regularization penalty term within the cost function imposes the *a priori* belief specified in the prior probability distribution, [61]. Regardless of which prior distribution is chosen, without additional information, the parameters λ or β would be difficult to choose. Moreover, finding the maximum is generally not the best way to interrogate a posterior. Finally, the single image reconstruction provides no quantification of the certainty for which the estimate should be trusted. Hence in what follows we take the position that densities should be sought rather than point estimates, and that the parameters of the prior should also be estimated.

Many choices can be made for $\mathbf{T}$, but for comparison purposes, in our experiments we show images formed using two popular sparsity-encouraging image formation methods: (i) ℓ_1 regularization [63], where sparsity is promoted in the image itself; and (ii) total variation (TV) regularization [52], where sparsity is promoted in the gradients in the image. Both can be codified as (11) with sparsifying operator $\mathbf{T}$ chosen appropriately, and both have been extensively applied in SAR (see e.g. [58, 28, 55, 2, 14]). Alternatively, the sub-apertures can also be modeled jointly, e.g. [50, 55], to take advantage of the presumably small differences between neighboring sub-aperture images. Figure 4 shows four realizations of (11) using the two-dimensional anisotropic TV operator for $\mathbf{T}$ with $\beta = 1$ and $\lambda = 1/40, 1/80, 1/120, 1/160$. Compared with the NUFFT and ℓ_1 images, we

notice the piecewise constant smoothing effect of TV regularization which, depending on λ , appears to blur some objects in the scene. The images are 1024×1024 pixels and use $L = 12$ windows each spanning 40° with an overlap of 10° . The composite maximum method provided in (14) discussed in Section 2.3 below is once again used to combine the sub-apertures. We note that the smaller image size is due to long runtime when using a non-identity regularization operator. As before, code and parameters (i.e. λ and the number of iterations) for this method are taken from [54, 55].

2.3 Combining sub-aperture images

As mentioned briefly in the descriptions of the figures, if desired, sub-aperture images can be combined to form a single composite image. One way to achieve this is by taking the argument of maximum modulus over the sub-apertures for each image pixel, [44]. Specifically, we compute elementwise

$$\mathbf{f}_{max}^{composite} = \arg \max_{l=1,\dots,L} |\mathbf{f}^{(l)}|, \quad (14)$$

which can be interpreted as choosing the most significant scattering response at a particular location in the scene over all of the sub-aperture windows. Although non-coherent and subject to a reduction of information, this combination of sub-apertures provides a single image to examine and display, and furthermore is able to somewhat mollify the problem of the faulty isotropic scattering assumption.

3 Sub-aperture SAR Image Formation with UQ

We now derive the proposed method for sub-aperture SAR imaging with uncertainty quantification. We begin by specifying the hierarchical Bayesian model followed by the SBL-type estimation procedure.

3.1 Hierarchical prior for speckle and sparsity

With the likelihood given by (9), a prior for each latent variable, $\mathbf{f}^{(l)}$, $l = 1, \dots, L$, is required to compute a posterior. The prior expresses a belief about a quantity before observation. For example, as already noted above an ℓ_1 prior can be used to enforce a sparsity assumption on the magnitude of $\mathbf{f}^{(l)}$ in some domain. Here we encode into the prior the fact that SAR images are affected by the speckle phenomenon. As discussed in Section 1, speckle, which is manifested as a granular pattern of bright and dark spots through an image, occurs in all coherent imaging and is often mischaracterized as noise, [35]. Although it is in fact signal, speckle still decreases image contrast and so it is desirable to remove it. Speckle reduction is often addressed using denoising techniques such as total variation (TV) regularization, which reduces to the form given by (11). But as shown in [24] and will be further demonstrated in our numerical experiments, treating speckle as noise results in an unnatural smoothing of the speckle. Therefore, here we instead follow [24, 28, 35] by directly incorporating the speckle into the prior, so that it is properly characterized as part of the signal.

In fully-developed speckle, we assume $\text{Re}(\mathbf{f}_i^{(l)})$ and $\text{Im}(\mathbf{f}_i^{(l)})$, i.e. the real and imaginary parts of the i th pixel of the image $\mathbf{f}^{(l)}$, are i.i.d. Gaussian with variance $1/\alpha_i^{(l)}$. Hence $\mathbf{f}^{(l)}$

is circularly-symmetric complex Gaussian, i.e. $\mathbf{f}^{(l)} \sim \mathcal{CN}(\mathbf{0}, \text{diag}(1/\boldsymbol{\alpha}^{(l)}))$, with density

$$p(\mathbf{f}^{(l)}|\boldsymbol{\alpha}^{(l)}) \propto \prod_{i=1}^N \alpha_i^{(l)} \exp\left(-\frac{1}{2}\|\sqrt{\boldsymbol{\alpha}^{(l)}} \odot \mathbf{f}^{(l)}\|_2^2\right), \quad l = 1, \dots, L, \quad (15)$$

where $\odot$ is elementwise multiplication. Thus the prior on the magnitude of the i th pixel $|\mathbf{f}_i^{(l)}| = \sqrt{\text{Re}(\mathbf{f}_i^{(l)})^2 + \text{Im}(\mathbf{f}_i^{(l)})^2}$ is Rayleigh with mean proportional to $1/\alpha_i^{(l)}$.

Remark 3.1 (Multiplicative noise model). *Because any change in the magnitude of each pixel $|\mathbf{f}_i^{(l)}|$ is proportional to $1/\alpha_i^{(l)}$, the speckle phenomenon has also been modeled as a multiplicative noise, [28]. By contrast, here we address the speckle directly by including it in our model with the prior (15), and later simultaneously estimating the associated speckle parameter $1/\alpha^{(l)}$ rather than using post-image-formation techniques. Based on this inversely proportional relationship between $|\mathbf{f}_i^{(l)}|$ and $\alpha_i^{(l)}$, we can deduce that our method successfully reduces speckle when the corresponding speckle parameter estimate is large.*

While the prior (15) and likelihood (9) are indeed enough to derive each posterior $p(\mathbf{f}^{(l)}|\hat{\mathbf{f}}^{(l)}, \boldsymbol{\alpha}^{(l)}, \beta^{(l)})$ and compute a MAP estimate for each sub-aperture, this estimate would not necessarily be representative of such a posterior. Nor does it provide information about the statistical confidence of the estimate for each recovered pixel value, or for any other recovered features of the image, [45]. Moreover, the regularization parameters for both the cost function and prior in the MAP estimate approach (analogous to $\beta^{(l)}$ and $\boldsymbol{\alpha}^{(l)}$ here) are *user-specified*. However they are truly unknown and therefore should be inferred. For these reasons we seek the joint posterior $p(\mathbf{f}^{(l)}, \boldsymbol{\alpha}^{(l)}, \beta^{(l)}|\hat{\mathbf{f}}^{(l)})$, and hence will be simultaneously estimating the speckle parameter $\boldsymbol{\alpha}^{(l)}$ and noise parameter $\beta^{(l)}$, which will lend clarity when determining whether or not the speckle reduction techniques are actually working.⁸

To calculate $p(\mathbf{f}^{(l)}, \boldsymbol{\alpha}^{(l)}, \beta^{(l)}|\hat{\mathbf{f}}^{(l)})$, we need to define priors on $\boldsymbol{\alpha}^{(l)}$ and $\beta^{(l)}$. We first invoke a conjugate Gamma prior for $\beta^{(l)}$. That is, $\beta^{(l)} \sim \Gamma(c, d)$ with density

$$p(\beta^{(l)}|c, d) \propto (\beta^{(l)})^{c-1} \exp(-d\beta^{(l)}), \quad l = 1, \dots, L. \quad (16)$$

Similarly, a conjugate Gamma prior is invoked on each element of $\boldsymbol{\alpha}^{(l)}$, i.e. $\alpha_i^{(l)} \sim \Gamma(a, b)$. By independence, $\boldsymbol{\alpha}^{(l)} \sim \Gamma(a, b)$ with

$$p(\boldsymbol{\alpha}^{(l)}|a, b) \propto \prod_{i=1}^N (\alpha_i^{(l)})^{a-1} \exp\left(-b \sum_{i=1}^N \alpha_i^{(l)}\right), \quad l = 1, \dots, L. \quad (17)$$

Note the dependence of (16) and (17) on parameters a , b , c , and d , which as in [9, 65] are chosen rather than inferred. Following [65], we choose $c = d = 0$ so that the resulting improper prior on the noise variance (16) is uniform over a logarithmic scale, making all scales equally likely. Regarding a and b , analogous parameters for a real-valued model in [9] are chosen to reflect the uncertainty in the latent variable, making the prior uninformative. On the other hand in [24, 65], $a = b = 0$, resulting in an improper

⁸Without a reference ground truth image, speckle statistics are typically only estimated from small regions of already formed images, [3].

prior $p(\mathbf{f}_i^{(l)}) \sim 1/|\mathbf{f}_i^{(l)}|$, which is peaked at zero and hence encourages sparsity.⁹ As already noted, the reflectivity in SAR images is presumably sparse, so that using (17) with parameters $a, b \approx 0$ seems reasonable. Importantly, fixing a, b, c and d in this way removes any need for user-defined parameters in this model.

The form of the posterior is achieved through a hierarchical Bayesian model where the likelihood parameters $\mathbf{f}^{(l)}$ and $\beta^{(l)}$ are given priors with prior parameters (hyperparameters) $\boldsymbol{\alpha}^{(l)}$, c , and d . Moving up to the final level of hierarchy in this model, the hyperparameter $\boldsymbol{\alpha}^{(l)}$ is given a prior (called a hyperprior) with hyperhyperparameters a and b . By Bayes' theorem, the joint posterior for each $\mathbf{f}^{(l)}$, $\boldsymbol{\alpha}^{(l)}$, and $\beta^{(l)}$, $l = 1, \dots, L$, is

$$\begin{aligned} p(\mathbf{f}^{(l)}, \boldsymbol{\alpha}^{(l)}, \beta^{(l)} | \hat{\mathbf{f}}^{(l)}) &\propto p(\hat{\mathbf{f}}^{(l)} | \mathbf{f}^{(l)}, \beta^{(l)}) p(\beta^{(l)} | 0, 0) p(\mathbf{f}^{(l)} | \boldsymbol{\alpha}^{(l)}) p(\boldsymbol{\alpha}^{(l)} | 0, 0) \\ &\propto (\beta^{(l)})^{M-1} \exp \left(-\frac{\beta^{(l)}}{2} \|\hat{\mathbf{f}}^{(l)} - \mathbf{F}^{(l)} \mathbf{f}^{(l)}\|_2^2 - \frac{1}{2} \|\sqrt{\boldsymbol{\alpha}^{(l)}} \odot \mathbf{f}^{(l)}\|_2^2 \right). \end{aligned} \quad (18)$$

It is important to make the distinction between the speckle model introduced in [28] and the subsequent methodology developed in [24]. While the implementation in [28] provides an appropriate characterization of speckle, it then uses TV regularization to obtain a MAP estimate. By contrast, the characterization in [24] is followed by a full posterior recovery using a sampling method which enables quantification of the uncertainty. Note that the methods in both [28] and [24] are for full azimuth imaging ($L = 1$ in (18)).

3.1.1 Regularization in other domains

Applying the above sparsifying prior in domains other than the imaging domain is challenging. Some methods have been proposed that accommodate specific domains, e.g. TV [8, 18, 19, 20, 23], however a generalized regularization model, for which the only restriction was that the regularization operator satisfy the so-called common kernel condition (which will be discussed in more detail in Section 3.2), [37], was not considered until [32].¹⁰

Hence, while (18) properly models speckle, as mentioned earlier one may wish to enforce sparsity in some other domain, e.g. some transformation of the magnitude $\mathbf{T}|\mathbf{f}^{(l)}|$. In order to include this, one can alter (15) to

$$p(\mathbf{f}^{(l)} | \boldsymbol{\alpha}^{(l)}) \propto \prod_{i=1}^N \alpha_i^{(l)} \exp \left(-\frac{1}{2} \|\sqrt{\boldsymbol{\alpha}^{(l)}} \odot (\mathbf{T}|\mathbf{f}^{(l)}|)\|_2^2 \right), \quad l = 1, \dots, L. \quad (19)$$

Letting $\Theta^{(l)}$ be a unitary diagonal matrix as in Section 2.2 such that $(\Theta^{(l)})^* \mathbf{f}^{(l)} \approx |\mathbf{f}^{(l)}|$, we then have

$$p(\mathbf{f}^{(l)} | \boldsymbol{\alpha}^{(l)}, \Theta^{(l)}) \propto \prod_{i=1}^N \alpha_i^{(l)} \exp \left(-\frac{1}{2} \|\sqrt{\boldsymbol{\alpha}^{(l)}} \odot (\mathbf{T}(\Theta^{(l)})^* \mathbf{f}^{(l)})\|_2^2 \right). \quad (20)$$

⁹To ensure numerical robustness, we choose all parameters, a, b, c, d to be machine precision rather than 0 in our implementation.

¹⁰Incidentally, the method introduced in [32] restricts the noise to be independent but not necessarily identically distributed, as is required in other approaches. This desirable property may be useful for fusing multiple time-dependent SAR data sets, but is not part of the current investigation.

We can then write the final posterior for each $l = 1, \dots, L$ as

$$p(\mathbf{f}^{(l)}, \boldsymbol{\alpha}^{(l)}, \beta^{(l)} | \hat{\mathbf{f}}^{(l)}, \Theta^{(l)}) \propto p(\hat{\mathbf{f}}^{(l)} | \mathbf{f}^{(l)}, \beta^{(l)}) p(\beta^{(l)} | 0, 0) p(\mathbf{f}^{(l)} | \boldsymbol{\alpha}^{(l)}, \Theta^{(l)}) p(\boldsymbol{\alpha}^{(l)} | 0, 0) \quad (21)$$

$$\propto (\beta^{(l)})^{M-1} \exp \left(-\frac{\beta^{(l)}}{2} \|\hat{\mathbf{f}}^{(l)} - \mathbf{F}^{(l)} \mathbf{f}^{(l)}\|_2^2 - \frac{1}{2} \|\sqrt{\boldsymbol{\alpha}^{(l)}} \odot (\mathbf{T}(\Theta^{(l)})^* \mathbf{f}^{(l)})\|_2^2 \right).$$

3.2 Estimation Techniques Revisited

At this stage the proposed method deviates from [24], where $\mathbf{f}^{(l)}$, $\boldsymbol{\alpha}^{(l)}$, and $\beta^{(l)}$ were sampled from the joint posterior using a single 360° azimuth window ($L = 1$). Requirements for storage and memory of samples for $\mathbf{f}^{(l)}$, $\boldsymbol{\alpha}^{(l)}$, and $\beta^{(l)}$ were already challenging in the single aperture window case, limiting image reconstruction to images of $N = 512^2$ and smaller. In particular, there were $\sim 5 \times 10^5$ sampled parameters to consider, each with ~ 1300 samples. In the current framework, as a consequence of appropriately modeling the anisotropic nature of scatterers in the scene, the composite image reconstruction model in (8) is now roughly L times larger than problem size considered in the corresponding models in [24]. Specifically in the context of (18), we would essentially be reconstructing L sub-aperture images of size N . For example, the corresponding composite imaging problem with $L = 60$ would require us to store samples of $\sim 3 \times 10^7$ parameters.

To limit such extreme storage and memory requirements, we now instead follow a Bayesian coordinate descent (BCD) algorithm similar to that of generalized sparse Bayesian learning method [32] to deterministically choose a density estimate. SBL-type methods have been used extensively for reconstructing SAR images from phase history data, see e.g. [41, 71, 72]. In addition to these direct applications of SBL in SAR image reconstruction from phase history data, the posterior density arrived at in SBL has also been used in sampling schemes for a variety of SAR tasks such as moving target recognition [46], model selection for speckle [38], as well as both direct image reconstruction [24] and composite image reconstruction [69]. To form an image, the individual (conditional) posterior for $\mathbf{f}^{(l)}$ and to the parameters $\boldsymbol{\alpha}^{(l)}$ and $\beta^{(l)}$, $l = 1, \dots, L$, are updated in sequence. The individual posterior for each $\mathbf{f}^{(l)}$ is

$$p(\mathbf{f}^{(l)} | \hat{\mathbf{f}}^{(l)}, \boldsymbol{\alpha}^{(l)}, \beta^{(l)}, \Theta^{(l)}) \propto \exp \left(-\frac{\beta^{(l)}}{2} \|\hat{\mathbf{f}}^{(l)} - \mathbf{F}^{(l)} \mathbf{f}^{(l)}\|_2^2 - \frac{1}{2} \|\sqrt{\boldsymbol{\alpha}^{(l)}} \odot (\mathbf{T}(\Theta^{(l)})^* \mathbf{f}^{(l)})\|_2^2 \right), \quad (22)$$

hence

$$\mathbf{f}^{(l)} | \hat{\mathbf{f}}^{(l)}, \boldsymbol{\alpha}^{(l)}, \beta^{(l)}, \Theta^{(l)} \sim \mathcal{CN}(\boldsymbol{\mu}^{(l)}, \boldsymbol{\Sigma}^{(l)}) \quad (23)$$

with

$$\boldsymbol{\mu}^{(l)} = \beta^{(l)} \boldsymbol{\Sigma}^{(l)} (\mathbf{F}^{(l)})^* \hat{\mathbf{f}}^{(l)} \quad (24)$$

and

$$\boldsymbol{\Sigma}^{(l)} = \left[\beta^{(l)} (\mathbf{F}^{(l)})^* \mathbf{F}^{(l)} + (\mathbf{T} \Theta^{(l)})^* \text{diag}(\boldsymbol{\alpha}^{(l)}) \mathbf{T} \Theta^{(l)} \right]^{-1}. \quad (25)$$

We note that for this to hold, $\mathbf{F}^{(l)}$ and $\mathbf{T}\Theta^{(l)}$ must satisfy the common kernel condition [32, 37, 64]:

$$\text{kernel}(\mathbf{F}^{(l)}) \cap \text{kernel}(\mathbf{T}\Theta^{(l)}) = \{\mathbf{0}\}, \quad l = 1, \dots, L,$$

which of course depends on the choice of $\mathbf{T}$, which can be non-invertible and even non-square. Since $\mathbf{F}^{(l)}$ is a Fourier transform operator with a full-rank matrix representation then it has trivial kernel. Therefore, for the purposes of our examples, we confirmed via direct computation that both $\mathbf{T}$ being the identity and the anisotropic TV operator when multiplied by $\Theta^{(l)}$ yield matrices with trivial kernels.

Per Bayesian coordinate descent [32], $\alpha^{(l)}$ and $\beta^{(l)}$, $l = 1, \dots, L$, are approximated by the means of their conditional posteriors. Due to the conjugate prior relationship, both conditional posteriors are Gamma-distributed with means given by

$$\mu_{\beta^{(l)}} = \frac{M}{\|\hat{\mathbf{f}}^{(l)} - \mathbf{F}^{(l)}\mathbf{f}^{(l)}\|_2^2} \quad (26)$$

and

$$\mu_{\alpha^{(l)}} = \frac{1}{\|\mathbf{T}(\Theta^{(l)})^*\mathbf{f}^{(l)}\|_2^2}. \quad (27)$$

Remark 3.2 Note that while $\Theta^{(l)}$ is not included in the estimation problem as an inferred variable, it is not fixed by an initial estimate as in Section 2.2, e.g. as $\Theta_{jj}^{(l)} = (\mathbf{f}_{ML})_j^{(l)} / |(\mathbf{f}_{ML})_j^{(l)}|$. In the iterative estimation procedure that follows we update each $\Theta^{(l)}$ using the same formula with the current iterate for $\mathbf{f}^{(l)}$:

$$\Theta_{jj}^{(l)} = \mathbf{f}_j^{(l)} / |\mathbf{f}_j^{(l)}|, \quad j = 1, \dots, N, \quad l = 1, \dots, L. \quad (28)$$

This technique was used in [58] to keep the quantity $(\Theta^{(l)})^*\mathbf{f}^{(l)}$ real-valued as the iterates change. Observe that due to construction of the prior this has no impact for the case when $\mathbf{T} = \mathbf{I}$.

The algorithm therefore consists of alternating computation of $\boldsymbol{\mu}^{(l)}$ using $\boldsymbol{\Sigma}^{(l)}$ from (24) and (25), $\mu_{\beta^{(l)}}$ and $\mu_{\alpha^{(l)}}$ from (26) and (27), and $\Theta^{(l)}$ from (28), until a convergence criterion has been achieved. The computational challenge is in using the covariance $\boldsymbol{\Sigma}$ to compute $\boldsymbol{\mu}$ at each step. However, as mentioned earlier, since $\mathbf{F}$ is applied via a NUFFT, [31], if $\mathbf{T} = \mathbf{I}$ (the assumption generally used for the SAR speckle model) then $\boldsymbol{\Sigma}^{-1} = \beta\mathbf{F}^H\mathbf{F} + (\mathbf{T}\Theta^{(l)})^*\text{diag}(\boldsymbol{\alpha})\mathbf{T}\Theta^{(l)}$ is diagonal and hence efficiently inverted with a cheap elementwise division by $\beta + \alpha$, [24]. If $\mathbf{T}$ is more general, then solving the linear system is more involved. Nevertheless, in many cases $\boldsymbol{\Sigma}^{-1}$ is typically sparse, e.g. if $\mathbf{T}$ is a TV operator. Algorithm 1 shows the exact steps used in our examples. We emphasize that, in practice, $\boldsymbol{\Sigma}$ itself is not computed until after Algorithm 1 has been executed and uncertainty quantification information is desired, and instead $\boldsymbol{\Sigma}^{-1}$ is initially treated only as the left-hand side of a linear system which is solved to find $\boldsymbol{\mu}$.

Many typical convergence criteria can be used. In our implementation, we use a standard convergence criterion based on the relative change of subsequent iterates (see details in Section 4), although convergence could also be based on speckle reduction in a particular region based on α , etc. Moreover, due to the loop in Algorithm 1, the sub-aperture images can be formed roughly L -times faster by operating in parallel.

Algorithm 1 Sub-aperture SAR image formation

Initiate $\boldsymbol{\mu}_0^{(l)} = (\mathbf{F}^{(l)})^* \hat{\mathbf{f}}^{(l)}$, $\Theta_0^{(l)} = \boldsymbol{\mu}_0^{(l)} / |\boldsymbol{\mu}_0^{(l)}|$, and convergence tolerance ϵ .
for $l = 1, \dots, L$ **do**
 $k = 0$;
 while $\frac{\|\boldsymbol{\mu}_{k+1}^{(l)} - \boldsymbol{\mu}_k^{(l)}\|_2}{\|\boldsymbol{\mu}_k^{(l)}\|_2} > \epsilon$ **do**
 $[\mu_{\alpha^{(l)}}]_{k+1} = \frac{1+2a}{|\mathbf{T}(\Theta^{(l)})^* \boldsymbol{\mu}_k^{(l)}|^2 + 2b}$;
 $[\mu_{\beta^{(l)}}]_{k+1} = \frac{M+2c}{\|\hat{\mathbf{f}}^{(l)} - \mathbf{F}^{(l)} \boldsymbol{\mu}_k^{(l)}\|_2^2 + 2d}$;
 $\boldsymbol{\mu}_{k+1}^{(l)} = [[\mu_{\beta^{(l)}}]_{k+1} (\mathbf{F}^{(l)})^* \mathbf{F}^{(l)} + (\mathbf{T} \Theta^{(l)})^* \text{diag}([\mu_{\alpha^{(l)}}]_{k+1}) \mathbf{T} \Theta^{(l)}]^{-1} [\mu_{\beta^{(l)}}]_{k+1} \boldsymbol{\mu}_0^{(l)}$;
 $\Theta_{k+1}^{(l)} = \boldsymbol{\mu}_{k+1}^{(l)} / |\boldsymbol{\mu}_{k+1}^{(l)}|$;
 $k = k + 1$;
 end while
 $\mu_{\alpha^{(l)}} = [\mu_{\alpha^{(l)}}]_{k+1}$;
 $\mu_{\beta^{(l)}} = [\mu_{\beta^{(l)}}]_{k+1}$;
 $\boldsymbol{\mu}^{(l)} = \boldsymbol{\mu}_{k+1}^{(l)}$;
 $\Theta^{(l)} = \Theta_{k+1}^{(l)}$;
end for

The results of Algorithm 1 are the means of the conditional posteriors defined by the final iterates $\{\mu_{\alpha^{(l)}}, \mu_{\beta^{(l)}}, \boldsymbol{\mu}^{(l)}, \Theta^{(l)}\}_{l=1}^L$. At this point, final covariances $\boldsymbol{\Sigma}^{(l)}$ can be computed as in (25), and being a probability distribution, many estimates can be derived from these distributions on $\mathbf{f}^{(l)}$, or they can be sampled to form other estimates, etc. This provides more information than is common in sub-aperture wide angle SAR image reconstruction. In addition, note that $\mu_{\alpha^{(l)}}$ and $\mu_{\beta^{(l)}}$ are also themselves point estimates of the speckle and noise parameters, respectively. This is significant as typical estimates for such parameters (if possible at all) would require additional processing as they are not included in the image formation process itself. Moreover, as discussed in more detail in [24] with respect to the $L = 1$ (full azimuth) case, in general we have no intuition for $\alpha^{(l)}$ and $\beta^{(l)}$. Our method allows us to encode this uncertainty by choosing uninformative priors. As a result, the estimates we obtain help to lend clarity when determining whether or not the speckle (and noise) reduction techniques are actually working. Finally, without a ground truth reference image, any additional processing to obtain this information would have to rely on pre-formed images.

3.3 Combining sub-aperture densities

As mentioned above in (14) and originating in [44], taking the argument of the maximum modulus over the sub-apertures of each image pixel provides a way to form a single composite image from multiple sub-aperture image estimates. This typically brightens the appearance of the image compared with full-azimuth imaging (see Figure 2), but is non-coherent and a reduction of the information in each sub-aperture image estimate. Furthermore, it is applied to image estimates, which are already a reduction of information compared to the full posterior density.

We therefore propose a method for combining the sub-aperture *densities* themselves.

This option assumes that each sub-aperture image is independent of the others and takes advantage of the properties of summing independent Gaussian random variables. We have that each sub-aperture image is modeled by a Gaussian posterior with mean $\boldsymbol{\mu}^{(l)}$ and covariance $\boldsymbol{\Sigma}^{(l)}$, which are computed using the final estimates for the noise variance and speckle parameter. If we assume each of these random variables are independent, then their average is also a Gaussian random variable.¹¹ That is, the coherent composite posterior is modeled as

$$\mathbf{f}_{mean}^{composite} \sim \mathcal{CN}(\boldsymbol{\mu}^{composite}, \boldsymbol{\Sigma}^{composite}), \quad (29)$$

where

$$\boldsymbol{\mu}^{composite} = \frac{1}{L} \sum_{l=1}^L \boldsymbol{\mu}^{(l)} \quad (30)$$

and

$$\boldsymbol{\Sigma}^{composite} = \frac{1}{L^2} \sum_{l=1}^L \boldsymbol{\Sigma}^{(l)}. \quad (31)$$

There are many ways to interrogate this composite distribution. For example, taking the elementwise square root of the diagonal of $\boldsymbol{\Sigma}^{composite}$ gives the standard deviation image for the composite posterior, which gives insight beyond a single point estimate into the spread of the density. For instance, around 68% of the mass lies within a standard deviation of the mean.

We furthermore note the significance of the availability to estimate the composite speckle parameter

$$\mu_{\alpha}^{composite} = \frac{1}{L} \sum_{l=1}^L \mu_{\alpha}^{(l)}. \quad (32)$$

Similar to the variance, (32) can provide confirmation that speckle has been reduced. Finally, it is of course also possible to use the maximum option to recover a non-coherent point estimate while relying on the mean option to obtain other information.

4 Computational Results

4.1 GOTCHA SAR Data

The GOTCHA Volumetric SAR Data Set 1.0 consists of SAR phase history data of a parking lot scene collected at X-band with a 640 MHz bandwidth with full azimuth coverage at 8 different elevation angles with full polarization, [13]. The GOTCHA SAR phase history data were collected when a plane carrying a sensor flew a roughly circular measurement flight around a parking lot near the Sensors Directorate Building at Wright-Patterson Air Force Base in Dayton, Ohio. The parking lot contains various targets

¹¹We recognize that given the typical processing that uses overlapping data sub-apertures, azimuthal angular independence may not typically be a good assumption. However, we suggest that non-overlapping sub-apertures can be taken. Furthermore, any angular independence argument would also depend on the scene itself.



Figure 5: Optical images of parking lot being imaged in GOTCHA dataset, [13]. Note scene contains a variety of calibration targets, such as primitive reflectors like the tophat shown, a Toyota Camry, forklift, and tractor.

including civilian vehicles, construction vehicles, calibration targets, primitive reflectors, and military vehicles. Figure 5 shows optical images of the targets. Note that because this is real-world data, the elevation angle is not perfectly constant, and the path is not perfectly circular. The center frequency is 9.6GHz and bandwidth is 640MHz. This public release data has been used extensively for testing new SAR image formation methods, [6, 7, 30, 55], and is available from <https://www.sdms.afrl.af.mil/index.php?collection=gotcha>.

4.2 Experiments

We now demonstrate the accuracy, efficiency, and robustness of the proposed method for sub-aperture SAR image reconstruction from phase history data with uncertainty quantification. Note that the ground truth reflectivity image is unknown, preventing the computation of standard error statistics such as the relative error. This is the case even in synthetically-created SAR examples, where the true reflectivity is still unknown. Therefore, the uncertainty quantification information the proposed method provides is all the more valuable, as it is able to quantify how much we should trust pixel values and structures in the image even in the absence of ground truth. Throughout, all reflectivity images $\mathbf{f}$ are displayed in decibels (dB):

$$20 \log_{10} \left(\frac{|\mathbf{f}|}{\max |\mathbf{f}|} \right), \quad (33)$$

with a minimum of -100 dB and maximum of 0 dB. Lesser or greater values are assigned the minimum or maximum.

In what follows, we generally hesitate to make subjective claims about the appearance or “quality” of these reconstructions, as all certainly have advantages and disadvantages. Thus depending on what the image is being used for, different image reconstruction methods or image features (e.g. smoothness, sparsity, etc.) may be more useful. We instead focus on the new methodology, its capabilities, as well as objective comparisons.

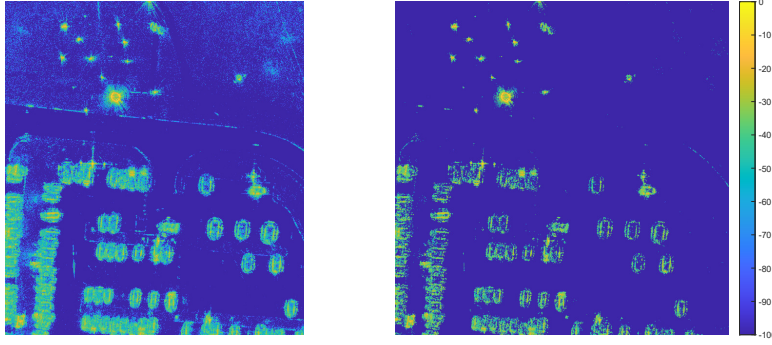


Figure 6: Sub-aperture composite **mean** images formed with Algorithm 1 using $\mathbf{T} = \mathbf{I}$ and convergence tolerance $\epsilon = 0.1$ (left) and $\epsilon = 0.01$ (right) from the GOTCHA data set, [13].

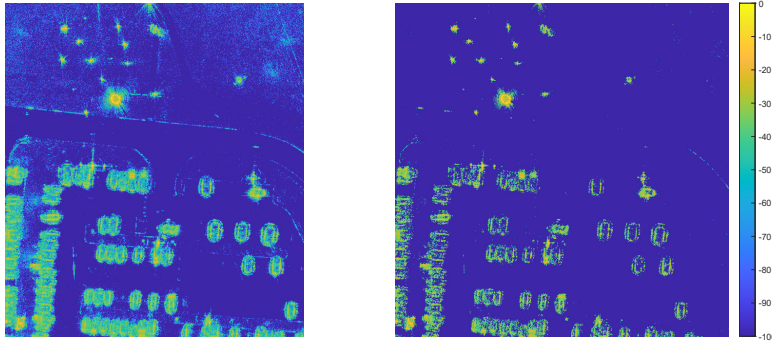


Figure 7: Sub-aperture composite **max** images formed with Algorithm 1 using $\mathbf{T} = \mathbf{I}$ and convergence tolerance $\epsilon = 0.1$ (left) and $\epsilon = 0.01$ (right) from the GOTCHA data set, [13].

4.2.1 Benchmarking sparsity testing

Figures 6 and 7 show the composite mean and max estimates for the GOTCHA parking lot scene using $L = 12$ windows each spanning 40° with 10° of overlap (see Section 4.2.2 below for more on sub-aperture size). Compared with one another, the mean and max composite estimates have a very similar appearance, which for a Gaussian aligns with our expectation that the mean and mode would be similar. As demonstrated in Figures 6 and 7, when compared with other sub-aperture methods in Figures 2 (right), 3, and 4, Algorithm 1 is able to reduce background speckle appearance while localizing bright targets.

We note that different convergence tolerances yield different results. For example, if $\epsilon = 0.1$ (corresponding to a 10% relative error iterate threshold), the sub-aperture regions each run for ~ 3 iterations and still have visible speckle in them, evidenced by the speckle parameter image. The image contains all features of the original image with increased contrast. However, if $\epsilon = 0.01$ (corresponding to a 1% relative error iterate threshold), the sub-aperture regions each run for ~ 12 iterations and have very little visible speckle in them. Strong targets are localized more precisely, but the image is less interpretable to human eyes. Regardless of the composite technique, our tests found that lowering ϵ further does not remarkably change the image appearance. This is essentially mandated by the criterion itself which limits relative changes. We therefore

consider $\epsilon = 0.01$ to correspond to the converged state.¹² Given this relatively simple specification of ϵ , we emphasize the difference in choosing this non-physical parameter compared with the challenge of selecting the physically-relevant regularization parameter used in other sparsity-based schemes. Furthermore, the main intent of imposing ϵ at all is to increase the speed of the algorithm, as discussed below.

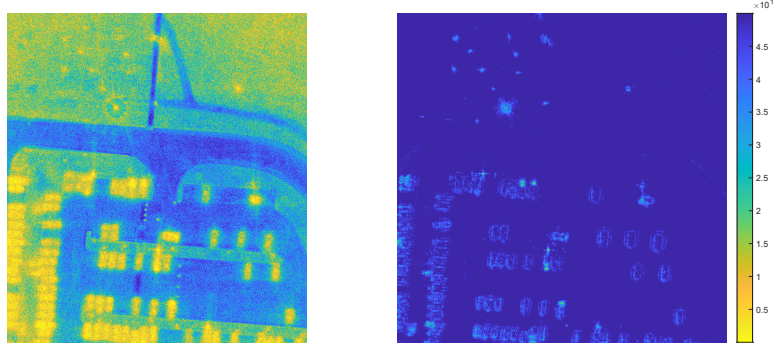


Figure 8: Sub-aperture composite **speckle parameter** images as in (32) formed with Algorithm 1 using $\mathbf{T} = \mathbf{I}$ and convergence tolerance $\epsilon = 0.1$ (left) and $\epsilon = 0.01$ (right) from the GOTCHA data set, [13]. Note the color scale here is inverted to conform with the other figures.

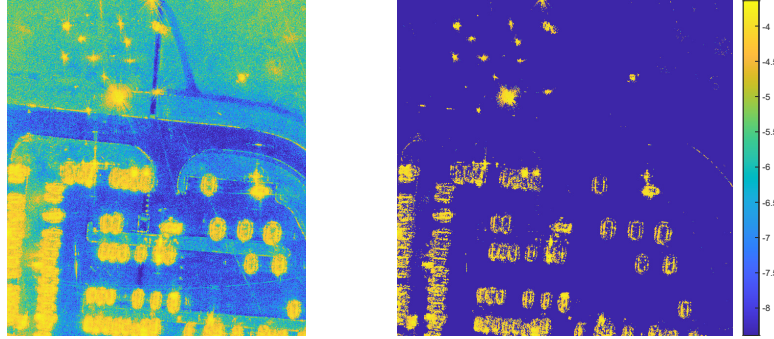


Figure 9: Sub-aperture composite **standard deviation** images as in (31) formed with Algorithm 1 using $\mathbf{T} = \mathbf{I}$ and convergence tolerance $\epsilon = 0.1$ (left) and $\epsilon = 0.01$ (right) from the GOTCHA data set, [13].

Furthermore, the proposed method also supplies extra information in Figures 8 and 9, which show the speckle parameter and standard deviation images as in (32) and (31), respectively. These images confirm the speckle and noise reduction resulting from the sparsity-inducing hierarchical Bayesian prior, indicating tight confidence intervals in the background with the only significant variance occurring where the mean estimate predicted signal returns.

¹²In this regard, we note that tests were conducted to confirm this is the case where each sub-aperture was run for 1000 iterations.

Remark 4.1 We note that for brevity we limit the exposition on extra information to these few examples, and emphasize that they are certainly not the only statistics that can be utilized. As was done in [24], median estimates and confidence intervals can also be computed.

Method	Variance
NUFFT	$5.5275e - 08$
$\ell_1, \lambda = 1/20$	$2.0548e - 16$
Alg. 1, T=I, $\epsilon = .1$, mean	$2.4433e - 18$
Alg. 1, T=I, $\epsilon = .1$, max	$3.5544e - 16$
Alg. 1, T=I, $\epsilon = .01$, mean	$2.1432e - 28$
Alg. 1, T=I, $\epsilon = .01$, max	$1.0437e - 26$

Table 1: Variance for a small 50×50 -pixel homogeneous subregion with each algorithm to show speckle reduction in 2048×2048 images.

To quantify the improvement and speckle reduction, Table 1 shows the variance of each image in a small (50×50 pixel) homogeneous region containing no targets in the center of the image in the intersection entering the parking lot. We choose to compare the most sparsifying ℓ_1 example tested ($\lambda = 1/20$), as this will serve as a lower variance estimate compared with other regularization parameter choices. This type of measurement has been used to evaluate speckle reduction, [3]. For the result, we see that all methods perform quite well, with all but one Algorithm 1 composite resulting in smaller variance compared with the most sparsifying ℓ_1 method over the homogeneous subregion, indicating superior small scale speckle reduction.

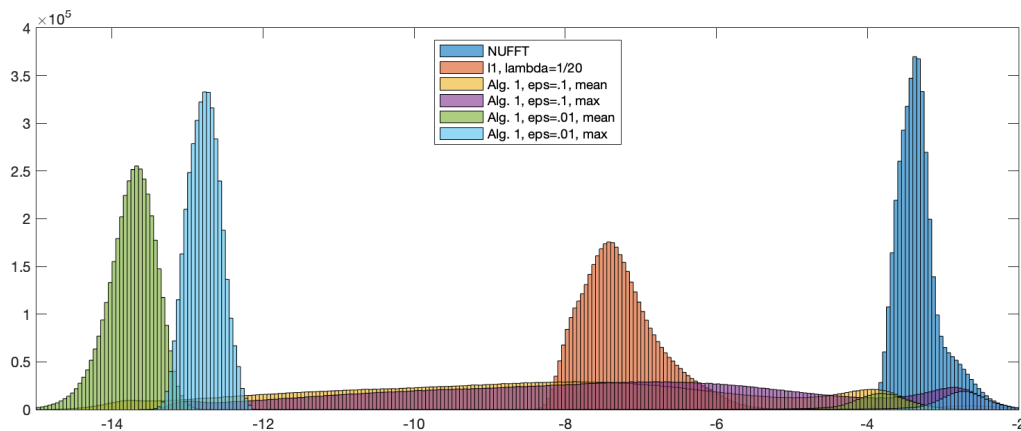


Figure 10: Histograms of $\log_{10} |\mu|$, where μ is the composite mean and max resulting from Algorithm 1 with different convergence tolerances, and the point estimates of the NUFFT and ℓ_1 regularization with $\lambda = 1/20$ comparing their sparsification. Observe that Algorithm 1 sparsifies the image more and has visibly increased contrast.

image size	2048 ²
NUFFT	8.4s
ℓ_1 (20 iters.)	2534.7s
Algorithm 1 (T=I , $\epsilon = 0.1$)	21.3s
Algorithm 1 (T=I , $\epsilon = 0.01$)	94.2s

Table 2: Sub-aperture imaging runtimes for $L = 12$ windows each spanning 40° with 10° overlap.

Figure 10 shows a histogram of the $\log_{10}$ absolute value of various composite image estimates. The separation between the left and right “humps” of the histogram of the modulus values in each reconstruction displayed in Figure 10 further confirms that Algorithm 1 significantly increases contrast in the image while dampening background speckle. That is, Algorithm 1 better separates pixels into categories: signal (the small hump on the right) and no signal (the large hump on the left). This separation of pixel values makes it easier to distinguish image features. In other words, it increases the contrast. In addition, the figure also shows that neither the NUFFT nor ℓ_1 reconstructions sparsifies the image to the extent done by Algorithm 1. In both cases, the left hump, which inherently represents pixels with no target reflectivity, is at least many orders of magnitude closer to zero. Since the goal of TV regularization is to sparsify its TV transform, it is not included in Figure 10.

Finally, Table 2 shows another advantage – the speed of reconstruction versus other sparsity encouraging methods, especially when the image size grows. Computations were done on a Macbook Pro with a 2.6 GHz 6-Core Intel Core i7 processor and 16GB memory. Implementation is achieved by using the NUFFT from [31], as well as comparison code from [54, 55].

4.2.2 Varying the number of sub-apertures

An interesting phenomenon occurs when varying the number of sub-apertures by widening or narrowing the sub-aperture size. Based on the relationship between the isotropic scattering assumption and full-azimuth imaging, the reader may expect a (nearly) linear relationship: using more sub-apertures will give stronger returns (i.e. more non-zero pixels and a “brighter” image) and more features of the scene will be shown in the composite estimate. However, our testing indicates that there is a wide range of sub-aperture angles around 60° that demonstrate good performance. In particular, Figure 11 shows composite mean estimates using Algorithm 1 with $\mathbf{T} = \mathbf{I}$ and $\epsilon = 0.01$ for 10° , 60° , and 360° (full-azimuth) sub-aperture sizes with no overlap. Respectively, these correspond to $L = 36, 6, 1$ sub-aperture windows. We see that both the 10° and 360° show less signal in these final estimates than the 60° case. While this observation deserves more investigation, we can interpret these results as follows. If the sub-aperture is too small, as in the 10° case, then scattered returns are also small and perhaps confused with noise. Therefore the SBL approach does not return much significant signal in each sub-aperture. By contrast, if the 360° full-azimuth aperture is used, the image suffers from overall weak returns due to the isotropic scattering assumption over such a wide angle.

Having sub-apertures of 60° seems to mitigate both issues – the angle is wide enough to pick out plenty of significant scatterers, but not so big that full-azimuth problems arise. This phenomenon is also observed with composite max estimates and varying ϵ , as well as for sub-aperture sizes that fall between those shown in the figure. Finally, we emphasize that the desired appearance of these composite estimates is subjective (more returns is not necessarily superior), and that furthermore because the selection of sub-aperture size happens during data processing, data collection itself is unaffected, and so multiple configurations can be easily and quickly tested depending on the scene.

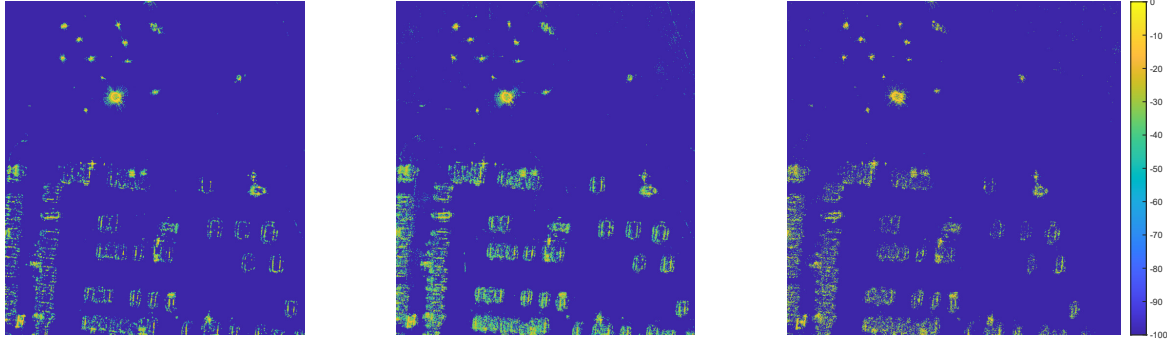


Figure 11: Composite mean estimates using Algorithm 1 ($\mathbf{T} = \mathbf{I}$, $\epsilon = 0.01$) for 10° , 60° , and 360° sub-aperture sizes with no overlap.

4.2.3 Total variation regularization testing

Finally, as mentioned earlier, it is possible to incorporate a non-identity regularizer into Algorithm 1 as long as the kernels of the regularization matrix and the forward Fourier operator have trivial intersection. Due to the now sparse (as opposed to diagonal) covariance inversion problem that must take place at every step, this is a much more computationally expensive problem. Further efforts will be devoted to accelerating this task if these early tests show promising results. Therefore, as a proof of concept, we limit the problem size to 512×512 pixels, where we focus on a subregion of the parking lot from the same GOTCHA data set, [13]. Specifically, the subregion is in the center-right of the full parking lot image and contains the forklift and a few other vehicles. To account for the smaller image size, we use $L = 60$ sub-aperture windows each spanning 8° with 2° of overlap.

Figures 12 and 13 show mean and max composite estimates of the subregion using two-dimensional anisotropic TV regularization. That is, the regularization operator is

$$\mathbf{T} = \begin{bmatrix} \mathbf{D} \otimes \mathbf{I} \\ \mathbf{I} \otimes \mathbf{D} \end{bmatrix}, \quad (34)$$

where $\otimes$ denotes the Kronecker product, $\mathbf{I}$ is the $N \times N$ identity matrix, and $\mathbf{D} \in \mathbb{R}^{N \times N}$ defined by

$$\mathbf{D}(i, j) = \begin{cases} 1 & i = j \\ -1 & i = j + 1 \end{cases}. \quad (35)$$

When applied to a concatenated image vector, using such an operator $\mathbf{T}$ encourages sparsity in the horizontal and vertical differences in the image. Some of the typical smoothing behavior associated with TV is apparent, with oversmoothing of some important features. Further testing is also required to understand the visible differences in the mean and max composite images, particularly for $\epsilon = 0.5$ (left).

Uncertainty quantification information for the TV case also requires further investigation. In particular, further testing is also needed in assigning physical meaning to the “speckle parameter” α here, as it does not represent speckle in this context. In addition, similar to the computation of Algorithm 1 itself with a non-identity regularizer, the computation of the composite covariance is still expensive even at this limited size so a standard deviation image is not available.

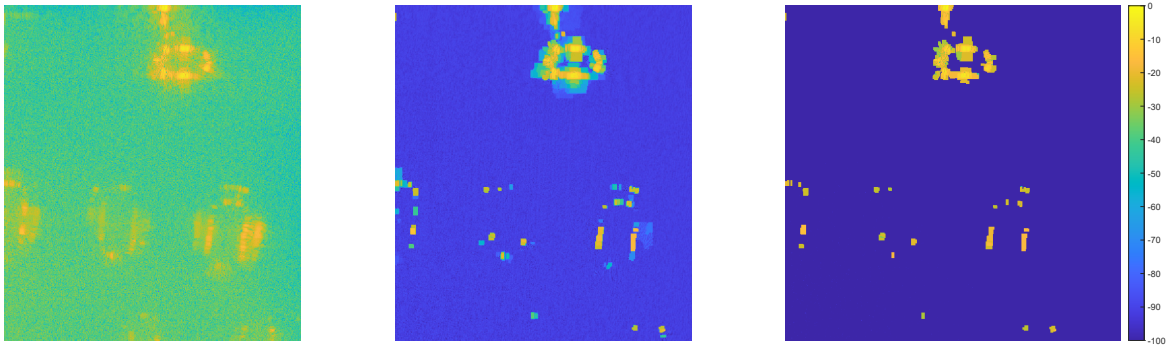


Figure 12: Sub-aperture composite **mean** images formed with Algorithm 1 using the two-dimensional anisotropic TV operator for $\mathbf{T}$ and (left) $\epsilon = 0.5$, (center) $\epsilon = 0.1$, and (right) $\epsilon = 0.01$ from the GOTCHA data set, [13].

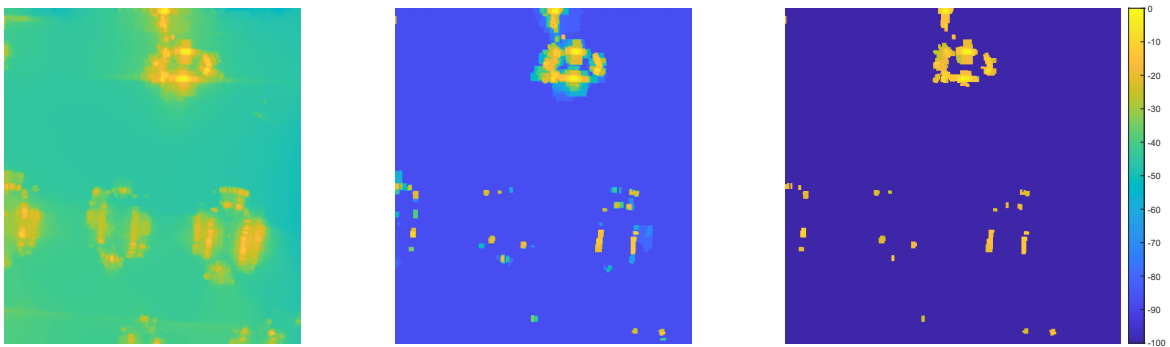


Figure 13: Sub-aperture composite **max** images formed with Algorithm 1 using the two-dimensional anisotropic TV operator for $\mathbf{T}$ and (left) $\epsilon = 0.5$, (center) $\epsilon = 0.1$, and (right) $\epsilon = 0.01$ from the GOTCHA data set, [13].

5 Conclusion

The sub-aperture imaging approach separately analyzes small aperture windows of data to form image estimates that can combine into a single composite reconstruction. This

can help to mitigate the effects of conflicting information resulting from the anisotropic nature of the scatterers in the scene. In this paper, a hierarchical Bayesian prior and corresponding SBL-type estimation procedure is used to develop a method that models sub-aperture posterior densities as opposed to producing point estimates. This extra uncertainty quantification information provides a more robust result for practitioners to be confident in, with access to estimates for the parameters governing speckle and noise, as well as regarding the standard deviation of image pixels. A new coherent composite image combination method based on summing independent Gaussians is also presented. Furthermore, the method allows for a variety of admissible non-identity regularization operators such as TV. In terms of the estimates generated, we demonstrate with a real-world numerical example that compared to existing methods, the proposed algorithm reduces speckle, improves contrast, and is orders of magnitude closer in speed to the fast NUFFT method which allows large images to be processed. Future work will focus on increasing efficiency as well as further exploring using non-identity regularizers in this framework. We also plan to incorporate autofocusing techniques into our method.

Acknowledgements

This work is partially supported by AFOSR #F9550-22-1-0411, NSF-DMS #1912685, ONR #N00014-20-1-2595.

References

- [1] ANDERSSON, F., MOSES, R., AND NATTERER, F. Fast Fourier methods for synthetic aperture radar imaging. *IEEE Transactions on Aerospace and Electronic Systems* 48, 1 (2012), 215–229.
- [2] ARCHIBALD, R., GELB, A., AND PLATTE, R. B. Image reconstruction from undersampled Fourier data using the polynomial annihilation transform. *Journal of Scientific Computing* 67, 2 (2016), 432–452.
- [3] ARGENTI, F., LAPINI, A., BIANCHI, T., AND ALPARONE, L. A tutorial on speckle reduction in synthetic aperture radar images. *IEEE Geoscience and Remote Sensing Magazine* 1, 3 (2013), 6–35.
- [4] ASH, J., ERTIN, E., POTTER, L. C., AND ZELNIO, E. Wide-angle synthetic aperture radar imaging: Models and algorithms for anisotropic scattering. *IEEE Signal Processing Magazine* 31, 4 (2014), 16–26.
- [5] ASH, J. N. An autofocus method for backprojection imagery in synthetic aperture radar. *IEEE Geoscience and Remote Sensing Letters* 9, 1 (2012), 104–108.
- [6] AUSTIN, C. D. *Sparse methods for model estimation with applications to radar imaging*. PhD thesis, The Ohio State University, 2012.
- [7] AUSTIN, C. D., ERTIN, E., AND MOSES, R. L. Sparse multipass 3D SAR imaging: Applications to the GOTCHA data set. In *Algorithms for Synthetic Aperture Radar Imagery XVI* (2009), vol. 7337, International Society for Optics and Photonics, p. 733703.

- [8] BABACAN, S. D., MOLINA, R., AND KATSAGGELOS, A. K. Sparse bayesian image restoration. In *2010 IEEE International Conference on Image Processing* (2010), IEEE, pp. 3577–3580.
- [9] BARDSLEY, J. M. MCMC-based image reconstruction with uncertainty quantification. *SIAM Journal on Scientific Computing* 34, 3 (2012), A1316–A1332.
- [10] BATU, O., AND CETIN, M. Parameter selection in sparsity-driven sar imaging. *IEEE Transactions on Aerospace and Electronic Systems* 47, 4 (2011), 3040–3050.
- [11] BOYD, S., PARIKH, N., CHU, E., PELEATO, B., ECKSTEIN, J., ET AL. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning* 3, 1 (2011), 1–122.
- [12] CANDÈS, E. J., ROMBERG, J., AND TAO, T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory* 52, 2 (2006), 489–509.
- [13] CASTEEL JR, C. H., GORHAM, L. A., MINARDI, M. J., SCARBOROUGH, S. M., NAIDU, K. D., AND MAJUMDER, U. K. A challenge problem for 2D/3D imaging of targets from a volumetric data set in an urban environment. In *Algorithms for Synthetic Aperture Radar Imagery XIV* (2007), vol. 6568, International Society for Optics and Photonics, p. 65680D.
- [14] ÇETIN, M., AND KARL, W. C. Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization. *IEEE Transactions on Image Processing* 10, 4 (2001), 623–631.
- [15] CETIN, M., AND MOSES, R. L. SAR imaging from partial-aperture data with frequency-band omissions. In *Defense and Security* (2005), International Society for Optics and Photonics, pp. 32–43.
- [16] CETIN, M., AND MOSES, R. L. Sar imaging from partial-aperture data with frequency-band omissions. In *Algorithms for Synthetic Aperture Radar Imagery XII* (2005), vol. 5808, International Society for Optics and Photonics, pp. 32–43.
- [17] ÇETIN, M., STOJANOVIĆ, I., ÖNHON, N. Ö., VARSHNEY, K., SAMADI, S., KARL, W. C., AND WILLSKY, A. S. Sparsity-driven synthetic aperture radar imaging: Reconstruction, autofocusing, moving targets, and compressed sensing. *IEEE Signal Processing Magazine* 31, 4 (2014), 27–40.
- [18] CHANTAS, G., GALATSANOS, N., LIKAS, A., AND SAUNDERS, M. Variational bayesian image restoration based on a product of t -distributions image prior. *IEEE transactions on image processing* 17, 10 (2008), 1795–1805.
- [19] CHANTAS, G., GALATSANOS, N. P., MOLINA, R., AND KATSAGGELOS, A. K. Variational bayesian image restoration with a product of spatially weighted total variation image priors. *IEEE transactions on image processing* 19, 2 (2009), 351–362.
- [20] CHANTAS, G. K., GALATSANOS, N. P., AND LIKAS, A. C. Bayesian restoration using a new nonstationary edge-preserving image prior. *IEEE Transactions on Image Processing* 15, 10 (2006), 2987–2997.
- [21] CHENEY, M., AND BORDEN, B. *Fundamentals of Radar Imaging*. SIAM, 2009.

- [22] CHIERCHIA, G., EL GHECHE, M., SCARPA, G., AND VERDOLIVA, L. Multitemporal sar image despeckling based on block-matching and collaborative filtering. *IEEE Transactions on Geoscience and Remote Sensing* 55, 10 (2017), 5467–5480.
- [23] CHURCHILL, V., AND GELB, A. Estimation and uncertainty quantification for piecewise smooth signal recovery. *Journal of Computational Mathematics* (2022 (to appear)).
- [24] CHURCHILL, V., AND GELB, A. Sampling-based spotlight SAR image reconstruction from phase history data for speckle reduction and uncertainty quantification. *SIAM Journal on Uncertainty Quantification* (2022 (to appear)).
- [25] COZZOLINO, D., PARRILLI, S., SCARPA, G., POGGI, G., AND VERDOLIVA, L. Fast adaptive nonlocal sar despeckling. *IEEE Geoscience and Remote Sensing Letters* 11, 2 (2013), 524–528.
- [26] DAOUI, A., YAMNI, M., KARMOUNI, H., SAYYOURI, M., QJIDAA, H., ET AL. Stable computation of higher order charlier moments for signal and image reconstruction. *Information Sciences* 521 (2020), 251–276.
- [27] DI MARTINO, G., PODERICO, M., POGGI, G., RICCIO, D., AND VERDOLIVA, L. Benchmarking framework for sar despeckling. *IEEE Transactions on geoscience and remote sensing* 52, 3 (2013), 1596–1615.
- [28] DONG, X., AND ZHANG, Y. SAR image reconstruction from undersampled raw data using maximum a posteriori estimation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8, 4 (2014), 1651–1664.
- [29] DUAN, H., ZHANG, L., FANG, J., HUANG, L., AND LI, H. Pattern-coupled sparse bayesian learning for inverse synthetic aperture radar imaging. *IEEE Signal Processing Letters* 22, 11 (2015), 1995–1999.
- [30] ERTIN, E., AUSTIN, C. D., SHARMA, S., MOSES, R. L., AND POTTER, L. C. GOTCHA experience report: Three-dimensional SAR imaging with complete circular apertures. In *Algorithms for Synthetic Aperture Radar Imagery XIV* (2007), vol. 6568, International Society for Optics and Photonics, p. 656802.
- [31] FESSLER, J. A., AND SUTTON, B. P. Nonuniform fast Fourier transforms using min-max interpolation. *IEEE Transactions on Signal Processing* 51, 2 (2003), 560–574.
- [32] GLAUBITZ, J., GELB, A., AND SONG, G. Generalized sparse bayesian learning and application to image reconstruction. *SIAM Journal on Uncertainty Quantification* (2022 (to appear)).
- [33] GORHAM, L. A., AND MOORE, L. J. SAR image formation toolbox for MATLAB. In *Algorithms for Synthetic Aperture Radar Imagery XVII* (2010), vol. 7699, International Society for Optics and Photonics, p. 769906.
- [34] GREENGARD, L., AND LEE, J.-Y. Accelerating the nonuniform fast Fourier transform. *SIAM Review* 46, 3 (2004), 443–454.
- [35] JAKOWATZ, C. V., WAHL, D. E., EICHEL, P. H., GHIGLIA, D. C., AND THOMPSON, P. A. *Spotlight-mode synthetic aperture radar: A signal processing approach*. Springer Science & Business Media, 2012.
- [36] JI, S., XUE, Y., AND CARIN, L. Bayesian compressive sensing. *IEEE Transactions on Signal Processing* 56, 6 (2008), 2346–2356.

- [37] KAIPIO, J., AND SOMERSALO, E. *Statistical and computational inverse problems*, vol. 160. Springer Science & Business Media, 2006.
- [38] KARAKUŞ, O., KURUOĞLU, E. E., AND ALTINKAYA, M. A. Generalized bayesian model selection for speckle on remote sensing images. *IEEE Transactions on Image Processing* 28, 4 (2018), 1748–1758.
- [39] LEE, J.-S., JURKEVICH, L., DEWAELE, P., WAMBACQ, P., AND OOSTERLINCK, A. Speckle filtering of synthetic aperture radar images: A review. *Remote sensing reviews* 8, 4 (1994), 313–340.
- [40] LEE, J.-Y., AND GREENGARD, L. The type 3 nonuniform FFT and its applications. *Journal of Computational Physics* 206, 1 (2005), 1–5.
- [41] LIU, H., JIU, B., LIU, H., AND BAO, Z. Superresolution isar imaging based on sparse bayesian learning. *IEEE Transactions on Geoscience and Remote Sensing* 52, 8 (2013), 5005–5013.
- [42] MOORE, L. J., RIGLING, B. D., AND PENNO, R. P. Characterization of phase information of synthetic aperture radar imagery. *IEEE Transactions on Aerospace and Electronic Systems* 55, 2 (2018), 676–688.
- [43] MOORE, L. J., RIGLING, B. D., PENNO, R. P., AND ZELNIO, E. G. Using phase for radar scatterer classification. In *Algorithms for Synthetic Aperture Radar Imagery XXIV* (2017), vol. 10201, International Society for Optics and Photonics, p. 102010J.
- [44] MOSES, R. L., POTTER, L. C., AND CETIN, M. Wide-angle sar imaging. In *Algorithms for Synthetic Aperture Radar Imagery XI* (2004), vol. 5427, SPIE, pp. 164–175.
- [45] NAGY, J. G., AND O’LEARY, D. P. Image restoration through subimages and confidence images. *Electronic Transactions on Numerical Analysis* 13 (2002), 22–37.
- [46] NEWSTADT, G., ZELNIO, E., AND HERO, A. Moving target inference with bayesian models in sar imagery. *IEEE Transactions on Aerospace and Electronic Systems* 50, 3 (2014), 2004–2018.
- [47] ONHON, N. Ö., AND CETIN, M. A sparsity-driven approach for joint sar imaging and phase error correction. *IEEE Transactions on Image Processing* 21, 4 (2011), 2075–2088.
- [48] PARRILLI, S., PODERICO, M., ANGELINO, C. V., AND VERDOLIVA, L. A nonlocal sar image denoising algorithm based on llmmse wavelet shrinkage. *IEEE Transactions on Geoscience and Remote Sensing* 50, 2 (2011), 606–616.
- [49] PAULSON, C. *Utilizing Glint Phenomenology to Perform Classification of Civilian Vehicles Using Synthetic Aperture Radar*. PhD thesis, University of Florida, 2013.
- [50] POTTER, L. C., ERTIN, E., PARKER, J. T., AND CETIN, M. Sparsity and compressed sensing in radar imaging. *Proceedings of the IEEE* 98, 6 (2010), 1006–1020.
- [51] POTTER, L. C., SCHNITER, P., AND ZINIEL, J. Sparse reconstruction for radar. In *Algorithms for Synthetic Aperture Radar Imagery XV* (2008), vol. 6970, SPIE, pp. 9–23.

- [52] RUDIN, L. I., OSHER, S., AND FATEMI, E. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena* 60, 1-4 (1992), 259–268.
- [53] SAMADI, S., ÇETIN, M., AND MASNADI-SHIRAZI, M. A. Sparse representation-based synthetic aperture radar imaging. *IET radar, sonar & navigation* 5, 2 (2011), 182–193.
- [54] SANDERS, T. MATLAB imaging algorithms: Image reconstruction, restoration, and alignment, with a focus in tomography. <http://www.toby-sanders.com/software>, <https://doi.org/10.13140/RG.2.2.33492.60801>.
- [55] SANDERS, T., GELB, A., AND PLATTE, R. B. Composite SAR imaging using sequential joint sparsity. *Journal of Computational Physics* 338 (2017), 357–370.
- [56] SANDERS, T., AND SCARNATI, T. Combination of correlated phase error correction and sparsity models for sar. In *Computational Imaging II* (2017), vol. 10222, SPIE, pp. 63–73.
- [57] SCARNATI, T. *Recent Techniques for Regularization in Partial Differential Equations and Imaging*. Arizona State University, 2018.
- [58] SCARNATI, T., AND GELB, A. Joint image formation and two-dimensional autofocus for synthetic aperture radar data. *Journal of Computational Physics* 374 (2018), 803–821.
- [59] SCARNATI, T., AND GELB, A. Variance based joint sparsity reconstruction of synthetic aperture radar data for speckle reduction. In *Algorithms for Synthetic Aperture Radar Imagery XXV* (2018), vol. 10647, International Society for Optics and Photonics, p. 106470R.
- [60] STOJANOVIC, I., CETIN, M., AND KARL, W. C. Joint space aspect reconstruction of wide-angle sar exploiting sparsity. In *Algorithms for Synthetic Aperture Radar Imagery XV* (2008), vol. 6970, SPIE, pp. 37–48.
- [61] STUART, A. M. Inverse problems: a Bayesian perspective. *Acta numerica* 19 (2010), 451–559.
- [62] SU, W., QIN, Y., WANG, H., AND YANG, Q. Joint isar imaging and phase error correction based on sparse bayesian learning. *Int. J. Sig. Proc. Sys* 4, 6 (2016), 487–493.
- [63] TIBSHIRANI, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)* (1996), 267–288.
- [64] TIKHONOV, A. N., GONCHARSKY, A., STEPANOV, V., AND YAGOLA, A. G. *Numerical methods for the solution of ill-posed problems*, vol. 328. Springer Science & Business Media, 1995.
- [65] TIPPING, M. E. Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research* 1, Jun (2001), 211–244.
- [66] UGUR, S., AND ARIKAN, O. SAR image reconstruction and autofocus by compressed sensing. *Digital Signal Processing* 22, 6 (2012), 923–932.
- [67] WU, C., DENG, B., WANG, H., QIN, Y., AND SU, W. A sparse bayesian approach for joint sar imaging and phase error correction. In *2015 4th International Conference on Computer Science and Network Technology (ICCSNT)* (2015), vol. 1, IEEE, pp. 1383–1386.

- [68] WU, J., LIU, F., JIAO, L., AND WANG, X. Compressive sensing sar image reconstruction based on bayesian framework and evolutionary computation. *IEEE Transactions on Image Processing* 20, 7 (2011), 1904–1911.
- [69] WU, Q., ZHANG, Y. D., AMIN, M. G., AND HIMED, B. High-resolution passive sar imaging exploiting structured bayesian compressive sensing. *IEEE Journal of Selected Topics in Signal Processing* 9, 8 (2015), 1484–1497.
- [70] XU, G., XING, M., ZHANG, L., LIU, Y., AND LI, Y. Bayesian inverse synthetic aperture radar imaging. *IEEE Geoscience and Remote Sensing Letters* 8, 6 (2011), 1150–1154.
- [71] XU, J.-P., PI, Y.-N., AND CAO, Z.-J. Bayesian compressive sensing in synthetic aperture radar imaging. *IET Radar, Sonar & Navigation* 6, 1 (2012), 2–8.
- [72] XUE, M., SANTIAGO, E., SEDEHI, M., TAN, X., AND LI, J. Sar imaging via iterative adaptive approach and sparse bayesian learning. In *Algorithms for Synthetic Aperture Radar Imagery XVI* (2009), vol. 7337, International Society for Optics and Photonics, p. 733706.
- [73] ZHANG, J., GELB, A., AND SCARNATI, T. Empirical bayesian inference using a support informed prior. *SIAM/ASA Journal on Uncertainty Quantification* 10, 2 (2022), 745–774.