Reconstruction From Aperture-Filtered Samples With Application to Scatterometer Image Reconstruction

Brent A. Williams, *Member, IEEE*, and David G. Long, *Fellow, IEEE*

***Abstract*—This paper approaches scatterometer image reconstruction as the inversion of a discrete noisy aperture-filtered sampling operation. Aperture-filtered sampling is presented and contrasted with conventional and irregular sampling. Discrete reconstruction from noise-free aperture-filtered samples is investigated and contrasted with conventional continuous reconstruction approaches. The discrete approach enables analytical treatment of the reconstruction grid resolution and the effective resolution imposed by the sampling and reconstruction operations. The noisy case is also explored. A reconstruction estimator based on maximum *a posteriori* (MAP) estimation is proposed to recover the conventional samples from noisy scatterometer measurements. This approach enables the scatterometer noise distribution to be appropriately accounted for in the reconstruction operation. The MAP and conventional reconstruction approaches are applied to the SeaWinds scatterometer and the Advanced Wind Scatterometer, and the effective resolution of the different methods is analyzed. The MAP approach produces results consistent with the well-established scatterometer image reconstruction (SIR) algorithm. The MAP approach significantly enhances the resolution at the expense of increased noise. Although a detailed noise-versus-resolution tradeoff analysis is beyond the scope of this paper, the new framework allows for a more general treatment than the ad hoc tuning parameters of the SIR algorithm.**

***Index Terms*—Inverse problems, irregular sampling, maximum *a posteriori* (MAP) estimation, scatterometer image reconstruction.**

I. Introduction

A SCATTEROMETER is an active microwave device (radar) that measures the normalized radar cross section (σ^0) averaged spatially over an aperture function. Typically, each radar pulse is partitioned into several measurements using range-Doppler processing so that the spatial aperture function of each measurement is a combination of the antenna footprint and range-Doppler processing. The sampling geometry results in irregularly spaced aperture functions with different shapes. A scatterometer makes several measurements with different azimuth angles, incidence angles, and, possibly, polarizations of the same spatial region.

Since scatterometers make multiple overlapping measurements of the Earth's surface, these measurements may be combined to produce σ^0 images. These images provide valuable geophysical information for land and ice studies [1], [2]. Several imaging methods have been proposed for scatterometer image reconstruction. Perhaps, the most simple is to create a gridded product by averaging all measurements whose centers fall into a particular grid element. Gridding produces relatively low-resolution images. Another imaging technique employs a weighted average on a higher resolution grid. This is the basis of the averaging (AVE) algorithm [3], which sets the value of each pixel to the average of all the σ^0 measurements, weighted by the value of the respective aperture functions at each pixel. Some common methods that further enhance the resolution are based on the additive algebraic reconstruction technique (AART) or the multiplicative algebraic reconstruction technique (MART) [3], [4]. For the noisy scatterometer σ^0 imaging problem, the MART algorithm tends to produce a less noisy estimate of the conventional samples than the AART algorithm. Practical considerations lead to the scatterometer image reconstruction (SIR) algorithm, which is a normalized version of MART that tends to converge faster and with less noise [3], [4].

Scatterometer σ^0 imaging algorithms proposed in the literature [3], [4] are based on continuous noise-free reconstruction operators. Continuous reconstruction operations allow for a very general but somewhat restricted result. More practical results can be obtained by considering the discrete inverse problem. For example, continuous methods do not use knowledge of the noise distribution and do not provide a simple method of determining the resolution, while a discrete approach can handle both these issues directly. Note also that the commonly used SIR algorithm is tuned using ad hoc methods in order to reduce the effects of noise and the filtering artifacts imposed by the aperture functions. These ad hoc methods make it difficult to analytically evaluate the quality of the estimates. An estimator that uses the noise distribution can be expected to perform better, is theoretically more appropriate, and allows the quality of the estimates to be analyzed using standard estimation-theory tools.

Although for some scenarios, the scatterometer image reconstruction problem is well posed (meaning that there are enough linearly independent measurements to estimate each pixel of the image), in most practical applications, the problem is ill posed. Thus, some form of regularization is required to estimate the σ^0 image from the scatterometer measurements. The AVE, AART, MART, and SIR algorithms all have a built-in regularization, but, as noted earlier, they do not fully account for noise.

Reconstruction techniques based on ill-posed inverse-problem theory and convex regularization have been developed for remotely sensed data [5], [6]. However, these methods are developed primarily for synthetic aperture radar and impose some implicit assumptions, such as coherent measurement scenarios, that are not appropriate for scatterometry. A slightly more generic approach is required for scatterometry.

Manuscript received December 23, 2009; revised August 10, 2010; accepted September 19, 2010. Date of publication December 10, 2010; date of current version April 22, 2011.

B. A. Williams was with the Electrical and Computer Engineering Department, Brigham Young University, Provo, UT 84602 USA. He is now with the Jet Propulsion Laboratory, Pasadena, CA 91109 USA.

D. G. Long is with the Electrical and Computer Engineering Department, Brigham Young University, Provo, UT 84602 USA (e-mail: long@ee.byu.edu).

Digital Object Identifier 10.1109/TGRS.2010.2086063

This paper treats scatterometer image reconstruction as the inversion of a noisy aperture-filtered sampling operation, where the noise can be expressed as additive but need not be independent of the signal (i.e., such as the case with fading). The sampling model is presented and made discrete, and noise-free reconstruction is explored and contrasted with conventional approaches. Reconstruction from noisy aperture-filtered samples is then examined. A reconstruction estimator based on maximum *a posteriori* (MAP) estimation is proposed for σ^0 imaging from scatterometers. Examples from the Advanced Wind Scatterometer (ASCAT) and the SeaWinds scatterometer are illustrated, and the resolution enhancement is compared with conventional approaches.

The focus of this paper is to develop a theoretical framework that allows the scatterometer noise model to be appropriately accounted for in the image-reconstruction problem, to present analytical results concerning resolution that were previously unobtainable, and to develop the first step toward an optimal reconstruction estimator. Although a new reconstruction algorithm is developed, in practice, the results are not expected to be a significant improvement over the SIR algorithm, which has been empirically tuned for decades to tradeoff noise and resolution and perform well for scatterometer applications. Rather, consistency between the results simultaneously validates both the rigorous and empirical methods. The MAP approach produces a higher resolution estimate than the SIR algorithm at the expense of also enhancing the noise. The new framework allows for a more general approach to optimizing the noise-versus-resolution tradeoff, although a detailed noise-versus-resolution analysis is beyond the scope of this paper.

This paper is organized as follows. Section II examines and contrasts conventional, irregular, and aperture-filtered sampling theories. Section III explores inverse operators for the discrete noise-free case and develops theoretical tools to analyze resolution. Section IV discusses estimation from noisy samples and develops a scatterometer image reconstruction estimator that directly accounts for scatterometer noise and compares the results with conventional approaches. Finally, Section V concludes this paper.

II. Sampling

Sampling is the process of converting a continuous-index signal, such as a scalar- or vector-valued function, into a discrete-index signal, termed as a sequence. In a Hilbert space (i.e., a complete vector space with an inner product defined [7]), sampling can be represented as a series of inner products with sampling functions [8]. This section contrasts conventional regular sampling, irregular sampling, and sampling with irregularly spaced aperture functions (i.e., aperture-filtered sampling) in a Hilbert space. A discrete-sampling formulation is considered, and the scatterometer-sampling model is presented.

A. Sampling Formulations

In conventional sampling theory, the sampling of a continuous-index signal $s(x)$ in a Hilbert space can be represented by performing an inner product with a Dirac delta function centered at the sample location x_n for regularly spaced samples (i.e., for the 1-D case, $x_n = nt$, where t is the sample spacing and n is an integer) [8]. This sampling produces a sequence in which each sample represents the value of the original signal at the corresponding sample location. This formulation also applies to irregular sampling with irregularly placed x_n's. The sampling operation can be written as a vector of inner products

$$\vec{s} = \begin{bmatrix} s[x_1] \\ \vdots \\ s[x_N] \end{bmatrix} = \begin{bmatrix} \int \delta(x - x_1) s(x) dx \\ \vdots \\ \int \delta(x - x_n) s(x) dx \end{bmatrix} \tag{1}$$

where $\vec{s}$ is the vector of samples of the continuous-index signal and N is the total number of samples. While in theory N may be infinite, for all practical applications, N must be finite. In this formulation, $s[x_n]$ is the value of $s(x)$ evaluated at $x = x_n$.

Note that if $s(x)$ is bandlimited to ω_0, the delta functions may be replaced with sinc functions $\mathrm{sinc}(\omega_0(x - x_n)) = \sin(\pi\omega_0(x - x_n))/\pi\omega_0(x - x_n)$, and the same result is obtained. Also, if the signal is bandlimited and periodic, the delta functions may be replaced with Dirichlet kernels $D(\omega_0(x - x_n)) = \sin((2\omega_0 + 1)(x - x_n)/2)/(2\omega_0 + 1)\sin((x - x_n)/2)$, with integration over only one signal period. For regular sampling of bandlimited signals, the original continuous-index signal can be reconstructed with sinc-function interpolation, while the bandlimited periodic signals can be recovered with Dirichlet-kernel interpolation. These interpolation operations are equivalent to low-pass filtering [8]. For irregular sampling, bandlimited signals can be reconstructed from irregularly spaced samples as long as the samples are sufficiently dense [4], [9]–[11].

For a more general sampling operation, the delta functions in (1) can be replaced with arbitrary aperture functions $A_n(x)$ that may have a different functional form (shape) for each sample. Sampling with arbitrary aperture functions produces the sampling operation

$$\vec{g} = \begin{bmatrix} \int A_1(x) s(x) dx \\ \vdots \\ \int A_N(x) s(x) dx \end{bmatrix} = \mathcal{A}s(x) \tag{2}$$

where $\mathcal{A}$ is the sampling operator and $\vec{g}$ represents the aperture-filtered samples. We use the notation $\vec{g}$ instead of $\vec{s}$ to represent the aperture-filtered samples since the samples do not necessarily represent the value of the original signal at a particular location due to the aperture function. Bandlimited signals can be reconstructed from samples made with distributed sampling functions [11], [12]. However, these results apply to general bandlimited signals, and more practical results may be found for the finite discrete case (i.e., bandlimited periodic signals).

If the aperture functions are bandlimited, the sampling operation can be expressed as (see Appendix I)

$$\vec{g} = \begin{bmatrix} \vec{A}_1^T \vec{s} \\ \vdots \\ \vec{A}_N^T \vec{s} \end{bmatrix} = \mathbf{A}\vec{s} = \mathcal{A}s(x) \tag{3}$$

where $\mathbf{A}$ is a matrix operator that operates on the conventional regularly spaced samples $\vec{s}$ of the bandlimited version

of the continuous-index signal $s(x)$, $\vec{A}_n$'s represent conventional sampling of the bandlimited aperture functions, and T represents the transpose (or Hermitian transpose for complex signals). The same result is obtained if $s(x)$ is bandlimited even if each $A_n(x)$ is not bandlimited, where the rows of $\mathbf{A}$ are conventional samples of bandlimited versions of the aperture functions. Moreover, if $s(x)$ or each $A_n(x)$ is bandlimited and periodic, $\mathbf{A}$ is a finite-dimensional matrix and can be analyzed with standard linear algebra (see Appendix I).

Having a discrete mapping from the conventional samples to the aperture-filtered samples allows a unique approach to reconstruction. Reconstructing the original signal can be thought of as first reconstructing $\vec{s}$ from $\vec{g}$ by inverting $\mathbf{A}$ and then performing Dirichlet-kernel interpolation on the reconstructed $\vec{s}$ to produce the continuous-index signal $s(x)$. Reconstruction from aperture-filtered samples is discussed in more detail in Section III.

B. Discrete Processing of Continuous-Index Signals

In practice, sampling is limited to a finite number of samples. With a finite number of samples, the signal may be *treated* as bandlimited and periodic. Constraining the signal to be bandlimited and periodic may seem restrictive but, as noted in Appendix I-B, is appropriate for most practical applications. That is, most practical signals can be approximated arbitrarily close by a bandlimited periodic signal. This approximation may introduce aliasing, but the aliasing is negligible if the bandlimit and the period are chosen appropriately.

For the rest of this paper, we restrict our attention to periodic bandlimited signals, although the bandlimit may be arbitrarily high. This restriction greatly simplifies the mathematics and allows the reconstruction analysis to be performed using linear algebra instead of real analysis, i.e., using matrix theory instead of linear operator theory. Furthermore, focusing on discrete problems allows analysis of the noisy inverse problem using standard estimation theory.

The choice of the signal bandlimit and period is problem dependent. A general signal period may be chosen as anything larger than the extent of the data or region that is to be processed. Once a signal period is selected, an appropriate bandlimit can be imposed. If the signal bandlimit or signal spectrum is known, a bandlimit may be chosen to minimize aliasing, but for many applications, no *a priori* knowledge about the signal is available. Nevertheless, there may be a fundamental bandlimit imposed by the aperture functions. If every aperture function is bandlimited by some ω_0, the sampling operation on the continuous-index signal is equivalent to the same sampling operation on the bandlimited version of the signal. This equivalence implies that no portion of the signal with frequency content higher than ω_0 can be recovered. Thus, assuming a bandlimit that is the highest bandlimit of the aperture functions is sufficient to recover all the information about the signal that is possible from the aperture-filtered samples alone. If the aperture functions are not bandlimited, a high bandlimit may be chosen so that the resulting aliasing is sufficiently small. The bandlimit and period determine the number of conventional samples required to represent the continuous-index signal, as well as the sample spacing (i.e., the pixel or grid resolution).

C. Scatterometer-Sampling Model

In scatterometry, σ^0 measurements represent noisy aperture-filtered samples of the 2-D σ^0 field. The σ^0 field may be reconstructed using various slice measurements of a similar flavor (i.e., that have the same geometry, frequency, and polarization). Measurements of a given flavor sample the same σ^0 field and can be combined. Furthermore, for land and ice imaging purposes, all slices of a given polarization and frequency may be combined by assuming negligible azimuthal variation and by adjusting the σ^0 values to a common incidence angle [3]. In this paper, we deal primarily with the SeaWinds scatterometer and ASCAT. For SeaWinds, the incidence-angle adjustment is not necessary since the slices of a given polarization have a similar incidence angle. For ASCAT, incidence-angle normalization to 40° is used.

SeaWinds is a Ku-band scatterometer that orbits the Earth in a Sun-synchronous near-polar orbit. The instrument has a scanning pencil-beam antenna with two beams at different incidence angles and polarizations. The v-pol beam is at a nominal incidence angle of 54°, while the h-pol beam is at an incidence angle of 46°. The two scanning beams produce a swath with four "flavors" (v-pol fore and aft looking and h-pol fore and aft looking) in the inner portion of the swath and two flavors in the outer portion of the swath where there is only one beam. The backscatter return from each pulse from each beam is partitioned into several "slices" using range-Doppler processing. Each slice is considered to be statistically independent, and each has its own aperture function or slice spatial response function [13].

The ASCAT scatterometer is a C-band v-pol instrument in near-polar orbit that has two sets of three stationary fan-beam antennas pointed at different azimuth angles. The system applies a type of pulse compression to obtain range resolution, producing slice σ^0 measurements with a relatively wide range of incidence angles. This sampling results in a swath in which each point is sampled by multiple beams with differing azimuth angles [14].

Although SeaWinds and ASCAT cover a large percentage of the Earth's surface daily, the σ^0 fields are generally reconstructed only over particular regions of interest. That is, we are only interested in a region of finite extent, which means that the signal and the slice response functions can be assumed to repeat periodically outside the region of interest.

The SeaWinds and ASCAT scatterometers are designed for large-scale ocean-wind vector measurements rather than σ^0 imaging. Although it is possible to use several days' worth of data to obtain a determined or overdetermined sampling operator (or sampling matrix), for many practical scatterometer imaging applications, scatterometer-sampling operators are underdetermined. That is, the scatterometer image reconstruction problem is generally ill-posed, meaning that we are attempting to estimate more parameters than we have independent measurements. This suggests that there is no unique reconstruction and that additional constraints on the signal must be applied in order to reconstruct σ^0 images. In the next two sections, we consider both the well-posed (fully determined or overdetermined sampling matrix) and ill-posed (underdetermined sampling matrix) cases for completeness but focus much of the discussion on the ill-posed case.

III. Noise-Free Reconstruction

This section considers discrete reconstruction from noise-free aperture-filtered samples. The ability to reconstruct the original signal depends on whether the mapping $\mathbf{A}$ is reversible. The system $\vec{g} = \mathbf{A}\vec{s}$ represents a linear system of equations. Depending on the structure of the forward operator $\mathbf{A}$, the system may be fully determined, overdetermined, or underdetermined. Each of the three cases is considered in the following, and the conditions required for exact reconstruction from the aperture-filtered samples are explored. Optimum inverse mappings that enable reconstruction of $\vec{s}$ from $\vec{g}$ are also presented for the overdetermined, fully determined, and underdetermined cases. Resolution limits imposed by the sampling are also explored, and the optimum pixel resolution for the SeaWinds and ASCAT scatterometers are derived.

A. Sampling Matrix: Overdetermined, Fully Determined, or Underdetermined

Whether the sampling matrix $\mathbf{A}$ is overdetermined, fully determined, or underdetermined depends on the density of the samples, the signal bandlimit, the signal period, and the linear independence of the aperture functions. As noted in the following and in Appendix I, the reconstructed signal bandlimit is limited by the highest bandlimit of the aperture functions, i.e., the number of conventional samples in the region of interest, and the condition of the sampling matrix are determined completely by the sensor. A fully determined system results if each aperture function is linearly independent and if the number of aperture-filtered samples is equal to the number of conventional samples required to represent the bandlimited periodic signal. An overdetermined system results if there are more aperture-filtered samples than the required number of conventional samples and if the number of linearly independent aperture functions is equal to the number of required conventional samples. An underdetermined system results if there are fewer aperture-filtered samples than the required number of conventional samples or if the number of linearly independent aperture functions is less than the required number of conventional samples.

Reconstructing the conventional samples $\vec{s}$ from the aperture-filtered samples $\vec{g}$ requires a slightly different reconstruction operation for each of the three cases: fully determined, overdetermined, and underdetermined. First, the conventional samples are reconstructed from the aperture-filtered samples; then, for each case, the reconstruction of the continuous-index signal is performed using Dirichlet-kernel interpolation with the reconstructed conventional samples

$$\hat{s}(x) = \sum_{x_n} \hat{s}[x_n] D\left(\omega_0(x - x_n)\right) \tag{4}$$

where $\hat{s}[x_n]$ is the reconstructed conventional sample corresponding to location x_n and $D(\omega_0(x - x_n))$ is the Dirichlet kernel with the same period and bandlimit imposed on the signal.

1) Fully Determined Case: For the case when the system is fully determined, matrix $\mathbf{A}$ is square and full rank. In this case $\mathbf{A}$ is invertible, and there exists a unique mapping from the aperture-filtered samples $\vec{g}$ back to the conventional samples $\vec{s}$. Therefore, if $\mathbf{A}$ is an $N \times N$ full-rank matrix, every signal that can be represented by N conventional samples can be reconstructed from the aperture-filtered samples.

For this case, the discrete reconstruction operation is

$$\hat{\vec{s}} = \mathbf{A}^{-1}\vec{g} \tag{5}$$

where $\hat{\vec{s}}$ represents the reconstructed conventional samples and $\mathbf{A}^{-1}$ is the standard matrix inverse of $\mathbf{A}$. The continuous-index signal is reconstructed using (4), where $\hat{\vec{s}}$ is the vector form of the reconstructed discrete signal $\hat{s}[x_n]$.

2) Overdetermined Case: The overdetermined case occurs when $\mathbf{A}$ has more rows than columns and is a full-row rank. By discarding redundant samples, a consistent overdetermined system may be transformed into a determined system without loss of information. That is, the aperture-filtered samples that correspond to the linearly dependent rows of $\mathbf{A}$ may be discarded, producing a mapping from $\vec{s}$ to the remaining aperture-filtered samples that is square and full rank and that represents a fully determined system. Thus, if $\mathbf{A}$ is an $M \times N$ matrix with N linearly independent rows, every signal that can be represented by N conventional samples can be reconstructed from the aperture-filtered samples.

When $\mathbf{A}$ is overdetermined, instead of discarding data, a least square inverse may be used for the reconstruction operation. A least square approach produces the same result in theory but is numerically more stable. Furthermore, with additive white Gaussian noise, the least square inverse produces the minimum-variance unbiased estimate of the conventional samples [7], [15]. Applying the least square inverse results in the discrete reconstruction operation

$$\hat{\vec{s}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\vec{g}. \tag{6}$$

The continuous-index signal is reconstructed using (4).

3) Underdetermined Case: The underdetermined case occurs if the row rank of $\mathbf{A}$ is less than the number of conventional samples required to represent the bandlimited signal. For this case, there is no one-to-one mapping that maps the range space of $\mathbf{A}$ back to the entire domain of $\mathbf{A}$. When $\mathbf{A}$ is underdetermined, there is no unique inverse; nevertheless, additional constraints may be imposed in order to obtain a pseudoinverse. Here, some practical results are summarized and contrasted with the conventional SIR approaches, while Appendix II considers the underdetermined case in detail.

Since there are several different constraints that can be imposed to obtain a pseudoinverse, there are also several different pseudoinverses that can produce a different reconstruction from the same data. One of the most common constraints is the minimum-norm constraint which results in the Moore–Penrose pseudoinverse [7]. Note that the AART algorithm also employs the minimum-norm constraint and so the AART algorithm converges to the same result as the Moore–Penrose pseudoinverse followed by Dirichlet interpolation. The Moore–Penrose pseudoinverse is particularly useful because it is also a linear inverse. This condition is not generally the case for other constraints. For example, the maximum-entropy constraint used in the MART and SIR algorithms results in a nonlinear reconstruction operation. Although the MART and SIR methods tend to produce a less noisy reconstruction with actual data, for the noise-free treatment here, the Moore–Penrose pseudoinverse

is suggested, and the scatterometer noise is handled explicitly later.

The reconstruction of the conventional samples from the aperture-filtered samples using the Moore–Penrose pseudoinverse is expressed as

$$\hat{\vec{s}} = \mathbf{A}^{\dagger}\vec{g} \tag{7}$$

where $\hat{\vec{s}}$ represents the reconstructed conventional samples and $\mathbf{A}^{\dagger}$ is the Moore–Penrose pseudoinverse of $\mathbf{A}$ defined by

$$\mathbf{A}^{\dagger} = \mathbf{V}_A(\mathbf{\Sigma}_A)^{-T}\mathbf{U}_A^T \tag{8}$$

where $\mathbf{U}_A\mathbf{\Sigma}_A\mathbf{V}_A^T$ is the singular value decomposition of $\mathbf{A}$ and $(\mathbf{\Sigma}_A)^{-T}$ is defined as the transpose of $\mathbf{\Sigma}_A$ with the nonzero elements replaced by their reciprocals [7]. The reconstruction of the continuous-index signal is performed using Dirichlet-kernel interpolation with the reconstructed conventional samples as described in (4).

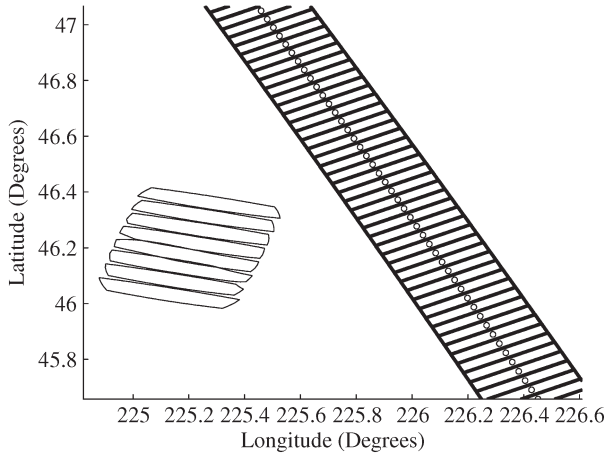

Fig. 1. Typical slice spatial response functions from SeaWinds and ASCAT for one pulse. The 6-dB contours are shown. The boxes with the circles in them are ASCAT SZF (slice) measurements; the contours on the left are SeaWinds slices.

B. Pixel and Effective Resolution

The optimum uniform-sample spacing (i.e., pixel resolution) and optimum bandlimit to use when reconstructing the signal have not been extensively explored in literature. A bound on the frequency recoverability from irregular sampling theory is given in [4] (i.e., there is a bandlimit such that every signal with that bandlimit can be recovered). The bound is determined by the sampling density, suggesting that the reconstruction grid resolution is a function of the irregular sampling density. However, for aperture-filtered sampling, assuming that the signal has a lower bandlimit than the highest bandlimit of the aperture functions may introduce aliasing. With aperture-filtered samples, the sample spacing must be determined by the bandlimit or approximate bandlimit of the aperture functions rather than by just the density of the aperture-filtered samples. The density of the samples is related to the condition of the sampling matrix (whether it is overdetermined, fully determined, or underdetermined) and does not directly impose a bandlimit on the signal.

1) SeaWinds and ASCAT Resolution: Here, the optimal regular pixel gridding is considered for the SeaWinds and ASCAT scatterometers. The pixel resolution for each scatterometer is derived from the narrow width of the measurement response functions.

The image grid pixel size is determined by the bandlimit of the slice spatial response functions. Fig. 1 shows typical 6-dB contours of the slice spatial response functions for a given pulse from SeaWinds and from ASCAT [16], [17]. For SeaWinds, the 6-dB slice contours are approximately 6 km in the narrow direction and 25 km wide in the long direction. For ASCAT, the contours are about 4.2 km in the narrow direction and 20–35 km in the long direction. To obtain a nominal pixel resolution, each of the slice response functions is approximated by a sinc-squared function that has the same 6-dB width as the slices in the narrow and long directions. The sinc-squared functions can be represented by regular samples with spacing corresponding to about half the narrow 6-dB beamwidth. This sample spacing is scaled by a factor of $1/\sqrt{2}$ to allow for the worst case slice orientation with respect to the gridding axes (i.e., 45°). For SeaWinds, the sinc-function approximation suggests a nominal conventional-sample spacing of about $6\ \text{km}/2\sqrt{2} \approx 2.12$ km [18]. This sample spacing is consistent with the 2.225-km pixel spacing found empirically to be approximately the resolution-enhancement limit for SeaWinds. For ASCAT, the pixel resolution is about $4.2\ \text{km}/2\sqrt{2} \approx 1.5$ km. Note that the range filtering of the σ^0 values performed onboard the ASCAT spacecraft degrades the effective reconstruction resolution. Thus, for ASCAT, a coarser conventional-sample spacing may be appropriate. From empirical observations, it seems that the resolution-enhancement limit is about 4 to 6 km for ASCAT. To be consistent between the data sets, we process both the SeaWinds and ASCAT data on the standard 2.225-km grid.

2) Effective Resolution: Since scatterometer-sampling operators are typically underdetermined, the effective resolution may be different from the pixel resolution defined by the bandlimit of the aperture functions. Furthermore, the effective resolution may vary throughout the swath. We define the effective resolution in terms of the width of the point-target response to the sampling and reconstruction operation.

For the linear reconstruction operations outlined previously, the resolution of the reconstructed images is only a function of the sampling matrix $\mathbf{A}$ and the reconstruction operation. If $\mathbf{A}$ is fully determined or overdetermined, then the resolution of the reconstructed images is the Nyquist sampling corresponding to the highest bandlimit of the aperture functions and is shift invariant. However, if $\mathbf{A}$ is underdetermined, then the resolution limit is as defined earlier, but the effective resolution may be lower and may vary throughout the image. The effective resolution at location i of the image can be defined by the width of the impulse response (at location i) of the sampling and reconstruction operation and can be calculated analytically, independent of the signal. The impulse response for the Moore–Penrose pseudoinverse $\vec{h}_i$ at location i is defined as

$$\vec{h}_i = \mathbf{A}^{\dagger}\mathbf{A}\vec{\delta}_i \tag{9}$$

where $\vec{\delta}_i$ is a discrete delta function centered at location i. Since the AART algorithm employs the same constraint as the Moore–Penrose pseudoinverse, this effective-resolution result also applies to the AART algorithm (in the limit of infinite

iterations). Note that if a nonlinear reconstruction operation is applied, the effective resolution may be a function of the signal as well as the sampling.

Recall that the MART and the SIR algorithms employ a maximum-entropy constraint. When the sampling matrix is underdetermined, applying a maximum-entropy constraint does not result in a linear inverse operation. That is, the discrete versions of the MART and SIR reconstruction operations cannot be expressed as a matrix applied to the data. Therefore, the effective resolution of these methods is more difficult to analyze theoretically.

IV. Noise

This section considers how noise added to the aperture-filtered samples affects the signal recovery. A noise model is introduced and optimality criteria are presented. Reconstruction from noisy aperture-filtered samples is explored. A scatterometer image reconstruction estimator is presented based on MAP estimation, and an example is illustrated. The MAP resolution enhancement is compared with conventional approaches. Methods of reducing the noise or variability of the estimates are discussed.

A. White Gaussian Noise

As previously noted, prior scatterometer reconstruction methods did not incorporate the measurement noise model. Neglecting noise entirely is equivalent to assuming additive white Gaussian noise, and it can be shown that for white Gaussian noise, the reconstruction operators shown previously are the minimum-variance unbiased estimators of the conventional samples [7], [15]. For many applications, the noise can be modeled as additive white Gaussian noise; however, for scatterometry, this is only a good approximation in some very restrictive situations, such as when the image is relatively constant. Scatterometry has a more complicated noise model, and this is one reason why the MART and SIR algorithms tend to produce a less noisy reconstruction than the AART method. Using knowledge of the scatterometer noise model, it may be possible to obtain a more ideal reconstruction estimator while appropriately accounting for the noise.

B. Noise Model

For the noise model, it is assumed that a zero-mean random variable $\vec{\nu}$ is added to the aperture-filtered samples $\vec{g}$. The noisy sampling operation may be written as

$$\vec{g}_\nu = \vec{g} + \vec{\nu} = \mathbf{A}\vec{s} + \vec{\nu} \tag{10}$$

where $\vec{\nu}$ is the noise random vector and $\vec{g}_\nu$ represents the noisy aperture-filtered samples.

Noisy scatterometer σ^0 measurements can be represented as Gaussian random variables, where the variances are quadratic functions of the means [19]. This noise distribution embodies the receiver noise as well as fading. Measurements are assumed to be statistically independent. The noise distribution has the form

$$f(\vec{g}_\nu|\vec{g}) = \frac{\exp\left\{-\frac{1}{2}(\vec{g_v} - \mathbf{A}\vec{s})^T\mathbf{R}^{-1}(\vec{g_v} - \mathbf{A}\vec{s})\right\}}{(2\pi)^{\frac{M}{2}}|\mathbf{R}|^{\frac{1}{2}}} \tag{11}$$

where M is the number of noisy aperture-filtered samples (i.e., the length of $\vec{g_v}$), which may be greater than N (i.e., the length of $\vec{s}$) for an overdetermined system. The covariance $\mathbf{R}$ of the vector of σ^0 measurements is a diagonal matrix, and each diagonal element $R_{i,i}$ can be expressed as

$$\begin{aligned} R_{i,i} &= \alpha_i(\vec{g}_i)^2 + \beta_i\vec{g}_i + \gamma_i \\ &= \alpha_i\left(\vec{A}_i^T\vec{s}\right)^2 + \beta_i\vec{A}_i^T\vec{s} + \gamma_i \end{aligned} \tag{12}$$

where $g_i = \vec{A}_i^T\vec{s}$ is the ith noise-free σ^0 measurement and α_i, β_i, and γ_i are parameters that are a function of the scatterometer design and the measured receiver noise power [19]. With noise, a reconstruction estimator of $\vec{s}$ from the noisy $\vec{g}_\nu$ is defined to reconstruct the signal.

C. Optimality Criteria

We seek an optimal estimator in the sense that the estimates are unbiased and have the minimum covariance of all possible estimators. Depending on the structure of the noise process, the estimator may be linear or nonlinear. A lower bound on the covariance of any unbiased estimator is the Cramer–Rao bound [7], [20]. The unbiased estimator that achieves the Cramer–Rao bound is the optimal estimator.

The Cramer–Rao lower bound on the covariance $\mathbf{R}(\hat{\vec{s}})$ of any unbiased estimator is the inverse of the Fisher information matrix $\mathbf{J}(\vec{s})$

$$\mathbf{R}(\hat{\vec{s}}) \geq \mathbf{J}(\vec{s})^{-1} \tag{13}$$

in the sense that $\mathbf{R}(\hat{\vec{s}}) - \mathbf{J}(\vec{s})^{-1}$ is positive semidefinite, where the components of $\mathbf{J}(\vec{s})$ are defined as [7]

$$J_{i,j}(\vec{s}) = E\left(\frac{\partial}{\partial s_i}\log\{f(\vec{g}_\nu|\vec{g})\}\right)\left(\frac{\partial}{\partial s_j}\log\{f(\vec{g}_\nu|\vec{g})\}\right) \tag{14}$$

where $f(\vec{g}_\nu|\vec{g})$ is the joint probability density function (pdf) of the noisy samples given the noise-free samples and s_i and s_j represent the ith and jth components of the noise-free discrete signal $\vec{s}$. The covariance in (13) represents the minimum covariance that can be obtained with any unbiased estimator, although a biased estimator may produce a lower covariance.

D. Reconstruction Estimators

In this section, reconstruction approaches from noisy data are examined for the case in which the Fisher information matrix is invertible and the case in which it is singular. For each case, methods of obtaining estimates with low bias and low expected squared error are explored.

1) Invertible Fisher Information Matrix: Note that $\mathbf{J}(\vec{s})$ is an $N \times N$ matrix, where N is the length of the vector $\vec{s}$. If $\mathbf{A}$ is fully determined or overdetermined, then generally, $\mathbf{J}(\vec{s})$ is invertible. If $\mathbf{J}(\vec{s})$ is invertible, every bandlimited periodic signal that can be represented by N conventional samples is observable in the sense that the conventional samples may be estimated with finite precision or variance of the estimates from the aperture-filtered samples. If $\mathbf{J}(\vec{s})$ is invertible, each conventional sample may be estimated directly; however, unless

$\mathbf{J}(\vec{s})$ is diagonal for every $\vec{s}$, the conventional samples must be estimated simultaneously.

A minimum-variance unbiased estimator is desirable; however, in most applications, there is no general method to find such an estimator. Nevertheless, maximum-likelihood (ML) estimators are asymptotically unbiased and asymptotically efficient (i.e., minimum variance) [7]. For these reasons, we propose that, lacking further information, an ML approach be used to estimate the conventional samples from the noisy aperture-filtered samples when the Fisher information is invertible.

An ML estimate of the conventional samples $\hat{\vec{s}}$ can be written as

$$\hat{\vec{s}} = \arg\max_{\vec{s}} \{f(\vec{g}_\nu|\vec{g})\} \tag{15}$$

which is the argument that maximizes the likelihood (i.e., $f(\vec{g}_\nu|\vec{g})$) of obtaining the noisy aperture-filtered samples given the noise-free aperture-filtered samples, where the noise-free aperture-filtered samples are a function of the conventional samples $\vec{g} = \mathbf{A}\vec{s}$. Reconstruction of the continuous-index signal is accomplished via Dirichlet-kernel interpolation from the ML estimates of the conventional samples.

2) Singular Fisher Information Matrix: If $\mathbf{A}$ represents an underdetermined system, $\mathbf{J}(\vec{s})$ is not invertible. A singular Fisher information suggests that the parameters (i.e., the conventional samples) are not observable from the data alone, i.e., the conventional samples cannot be estimated with any degree of precision using only the noisy aperture-filtered samples. However, if additional constraints are imposed, such as a signal model or a prior distribution on the signal, estimates of the conventional samples may be obtained.

There are three philosophically different approaches for dealing with unobservable parameters: model-based estimation, variational analysis, and Bayes estimation theory. Each of these methods effectively injects prior information about the signal so that the parameters that are unobservable from the aperture-filtered samples alone become observable using the aperture-filtered samples and the prior information.

For model-based estimation, a signal model is imposed such that the parameters of the model are observable, and the corresponding Fisher information matrix for estimating the model parameters is invertible. The signal is then reconstructed from the model-parameter estimates. This method requires imposing a signal model. If some knowledge of the structure of the signal is available, a model that describes the signal structure may be chosen. If prior knowledge of the signal is limited, it is generally desirable to impose a model that preserves all the information in the aperture-filtered samples and that minimizes aliasing.

Variational analysis imposes additional constraints on the ML objective function and simultaneously optimizes some linear combination (usually a convex combination) of the ML objective function with the constraints. This approach is similar to how the information-preserving pseudoinverse is obtained in Appendix II. For this method, the constraints can be chosen somewhat arbitrarily. Furthermore, the relative weights assigned to each constraint can also be chosen arbitrarily. Although variational analysis is a powerful tool, unless there is good reason to choose particular constraints and relative weights, variational analysis is *ad hoc*.

Bayes estimation imposes prior information about the signal via a prior distribution of the signal. The general form of a Bayesian estimator modified for our application is [7]

$$\hat{\vec{s}} = \arg\min_{\vec{s}} \left\{ \int L(\vec{s}, \vec{g}_\nu) f(\vec{s}|\vec{g}_\nu) d\vec{s} \right\} \tag{16}$$

where $L(\vec{s}, \vec{g}_\nu)$ is a loss function and $f(\vec{s}|\vec{g}_\nu)$ is the conditional posterior distribution. The most common loss functions are the squared-error loss function and the uniform loss function. Using a uniform loss function produces a MAP estimator which has as its objective function, the ML pdf, scaled by the prior distribution.

Bayes estimation is extremely powerful, but it implicitly assumes that the signal is a random process, and it requires knowledge of the signal distribution. Prior distributions can be obtained for particular applications either empirically from a collection of data, theoretically from knowledge of the physical process that produces the signal, or by assuming a maximum-entropy distribution given knowledge of a few qualities of the signal (such as the signal mean, power, spectrum, or region of support).

We take a Bayesian MAP estimation approach to solve the noisy inverse problem when the sampling operator is underdetermined. The MAP estimator can be expressed as

$$\hat{\vec{s}} = \arg\max_{\vec{s}} \{f(\vec{g}_\nu|\vec{g}) f(\vec{s})\} \tag{17}$$

where $f(\vec{s})$ is the prior distribution of $\vec{s}$. Using a prior is mathematically equivalent to (although philosophically different from) assuming additional statistically independent data representing noisy conventional samples. That is, performing ML estimation with the aperture-filtered samples and the additional data, whose noise distribution is the prior, results in the same exact expression as the MAP estimator.

With the scatterometer noise distribution, it is difficult to verify if the MAP estimator is a minimum-variance unbiased estimator. Nevertheless, since the MAP estimator can be expressed as an ML estimator with additional data (as described previously), it can exhibit similar asymptotic qualities. The MAP estimator is a good candidate for many noise distributions, often resulting in low-bias and low-variance estimates.

E. Scatterometer Reconstruction Estimator

Here, we apply the MAP estimation approach to the SIR problem. The MAP estimator can handle both the well-posed and ill-posed problems, and, under certain conditions, MAP estimation degenerates to ML estimation for well-posed problems. For completeness, the scatterometer ML estimator is discussed in Appendix III.

1) MAP Reconstruction: Reconstruction is accomplished by estimating the conventional (uniformly spaced) samples $\vec{s}$ of the σ^0 field using a MAP estimator. The MAP estimator searches for the conventional samples $\vec{s}$ that maximize the ML function scaled by the prior. This process is equivalent to maximizing the linear combination of the log-likelihood function and the log of the prior. The ML function is the pdf of the noisy σ^0 measurements, and the prior is a pdf of the σ^0

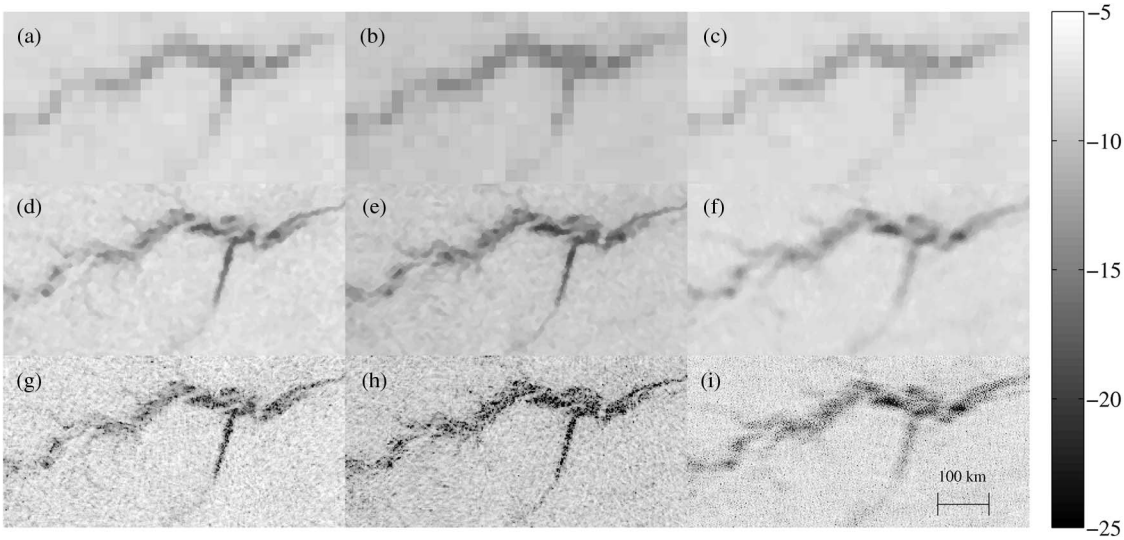

Fig. 2. Reconstructed σ^0 images (in decibels) from SeaWinds and ASCAT over the Amazon using four days' worth of data (JD 201-204) in 2008. Images (a)–(c) are gridded images of SeaWinds h-pol, SeaWinds v-pol, and ASCAT, respectively. Images (d)–(f) are SIR images of SeaWinds h-pol, SeaWinds v-pol, and ASCAT, respectively. Images (g)–(i) are MAP images of SeaWinds h-pol, SeaWinds v-pol, and ASCAT, respectively. The diagonal streaks in the river are actual features.

image. The scatterometer noise model results in the MAP log-likelihood objective function

$$L_{\mathrm{MAP}} = -\sum_i \left[\frac{\left(g_{\nu,i} - \vec{A}_i^T \vec{s}\right)^2}{2R_{i,i}} + 1/2 \log\{2\pi R_{i,i}\} \right] + \log f(\vec{s}) \tag{18}$$

where $f(\vec{s})$ is the prior pdf.

The local maxima of the MAP objective function can be found by setting the gradient equal to zero and solving the corresponding system of equations. However, the resulting system of equations is somewhat complicated so we use a gradient-search method to find a local maximum of (18) near an initial guess. The gradient-search method begins with an initial value computed using the AVE algorithm and moves incrementally in the direction of the gradient until convergence to the maxima.

For practical purposes, the search method can be adjusted to follow the gradient with respect to the parameters expressed in decibels (i.e., we estimate the decibel image). This forces the estimates to always be positive and makes it easier to handle the update when the σ^0 values have a low signal-to-noise ratio. This approach also tends to converge faster than searching the linear image.

2) Priors: The form of the prior depends on the application and what assumptions about the signal are reasonable. The standard approach to obtain a prior is using an empirical distribution from a large collection of data. However, because different surfaces (i.e., land, ice, ocean, or vegetation) have such different responses, empirical priors for scatterometer imaging applications may be multimodal and difficult to express as a functional form that can be differentiated.

Another approach to obtain a prior is to employ a maximum-entropy distribution. Using a maximum-entropy prior ensures that the least amount of structure is imposed on the signal under certain constraints (e.g., for a given mean and variance). For scatterometer imaging, a one-sided distribution is appropriate since the noise-free σ^0 measurements represent a magnitude. The maximum-entropy one-sided distribution with one parameter is the exponential distribution. The larger the mean of the exponential distribution is, the larger is the entropy. Thus, we may use an exponential distribution with an arbitrarily large mean to regularize the problem. However, because of the structure of the scatterometer noise, using an exponential prior requires a large number of iterations for the gradient-search algorithm to converge, which may not be practical.

For convenience, we use a log-normal prior with a mean as the AVE image and a tunable variance. The smaller the variance is, the closer is the estimate to the smooth AVE image. The larger the variance is, the less the result is influenced by the prior. A tunable variance allows a tradeoff between resolution enhancement and noise amplification. This prior distribution for each s_i can be expressed as

$$f(s_i) = \frac{10}{\log(10) s_i p \sqrt{2\pi}} \exp \frac{\left(10 \log_{10}(s_i) - 10 \log_{10}(s_{i,\mathrm{AVE}})\right)^2}{-2p^2} \tag{19}$$

where $s_{i,\mathrm{AVE}}$ is the AVE image value at index i and p is the variance. For this paper, we set p to be very large in order to obtain the highest resolution reconstruction, leaving the noise-versus-resolution tradeoff for future investigation.

3) Example: Now, consider a practical example. Fig. 2 shows multiorbit gridded, SIR, and MAP images of the Amazon made from SeaWinds and ASCAT data. For ASCAT, the incidence-angle-normalized images are plotted. As expected, the ASCAT effective resolution is lower due to onboard spatial filtering. The results of the reconstruction algorithms, shown by the bottom two rows of images in Fig. 2, enhance the resolution compared with the gridded product (top row). The MAP images (bottom row) contain more details than the SIR images (middle row), although the MAP images seem to be noisier. The noise in the MAP images can be attenuated by filtering the images or by tuning the variance of the log-normal prior, which can produce images of comparable quality with the SIR images.

Note the different responses of the surface to the different polarizations and frequencies. This combined C- and Ku-band

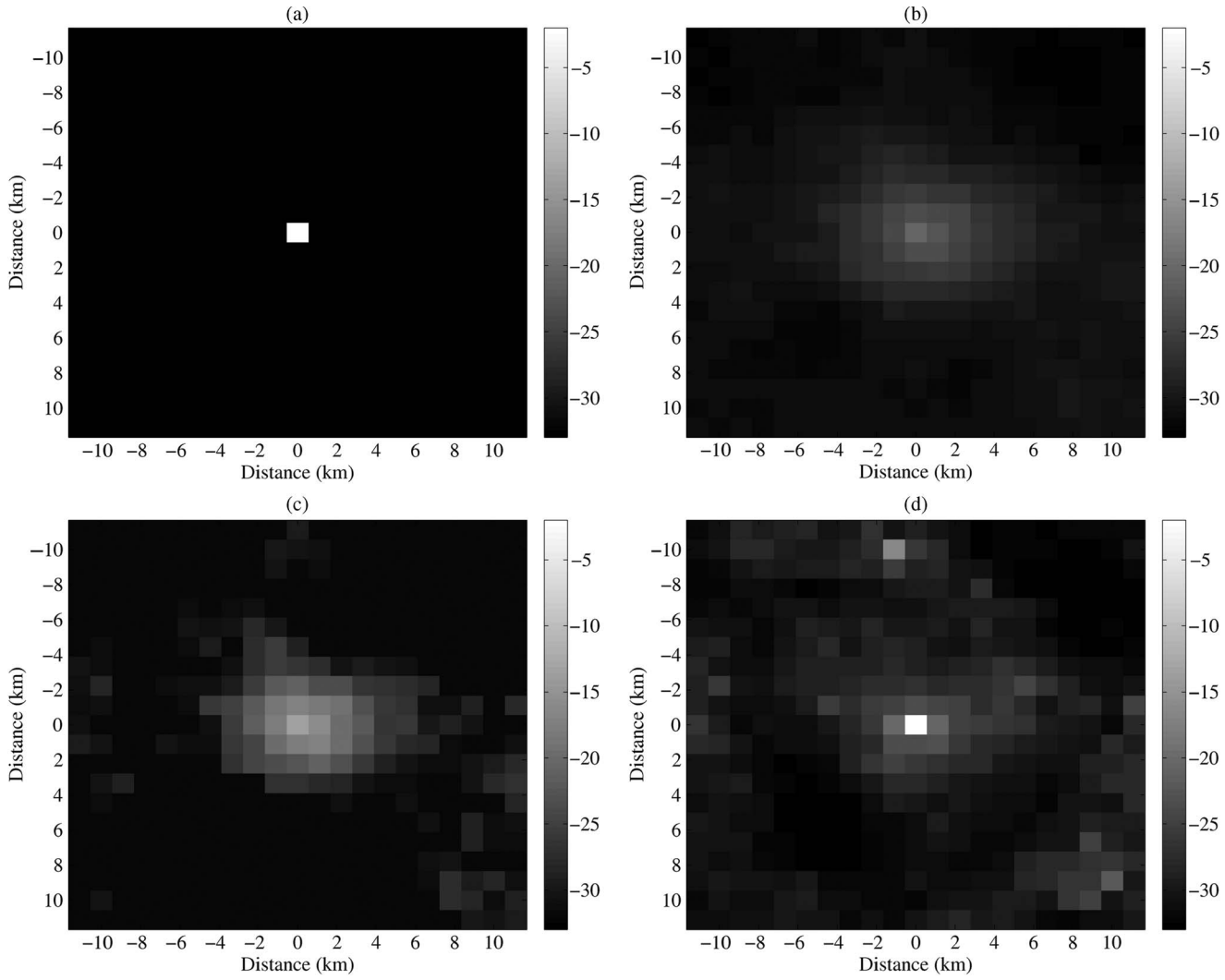

Fig. 3. (a) Monte Carlo simulation synthetic point target and spatial responses for the (b) AVE, (c) SIR, and (d) MAP reconstruction methods.

information may be valuable for land, forest, and ice studies. Obtaining collocated measurements of both frequencies may provide additional geophysical information about the state of the surface, which is an additional incentive for the development of a dual-frequency system [21].

F. Resolution Revisited

The effective resolution of the MAP and SIR reconstruction operations are difficult to evaluate analytically. Therefore, Monte Carlo simulation is employed to evaluate the resolution enhancement and the effective resolution of the different reconstruction methods.

First, a synthetic point target is measured using actual SeaWinds sampling geometry for ten consecutive days. Using more days' worth of samples allows the noise to be more suppressed in the reconstruction. Next, Monte Carlo noise is added to these synthetic measurements, and the AVE, SIR, and MAP reconstruction methods are applied. The resulting images represent noisy realizations of the spatial response functions of the different reconstruction methods. Fig. 3 shows a synthetic point target and the noisy spatial response functions of the AVE, SIR, and MAP reconstruction methods for a particular noise realization. Moreover, Fig. 4 shows the profiles of the horizontal and vertical cuts of the point-target-response images. As expected, the AVE response is the smoothest and the most spread out, and the peak value is the most attenuated of the reconstruction methods. The SIR response is sharper than the AVE, but not as sharp as the MAP response. The MAP result has the sharpest response, and the height is recovered most accurately, suggesting that the effective resolution of the MAP result is higher than the AVE and SIR methods. Nevertheless, the MAP image has the noisiest response.

Another general method for testing the effective resolution in the presence of noise is spectral analysis. In conventional signal processing, this is a powerful tool that can indicate the frequencies where the noise begins to dominate the signal (i.e., where the spectrum levels out, indicating the highest frequency where the signal is distinguishable from the noise). Although spectral analysis is typically useful for resolution analysis, because the noise of the reconstructed estimates is not isotropic, the spectrum of the noise process is not defined, and spectral analysis may be misleading. For this application, a more sophisticated method must be applied to develop an optimal

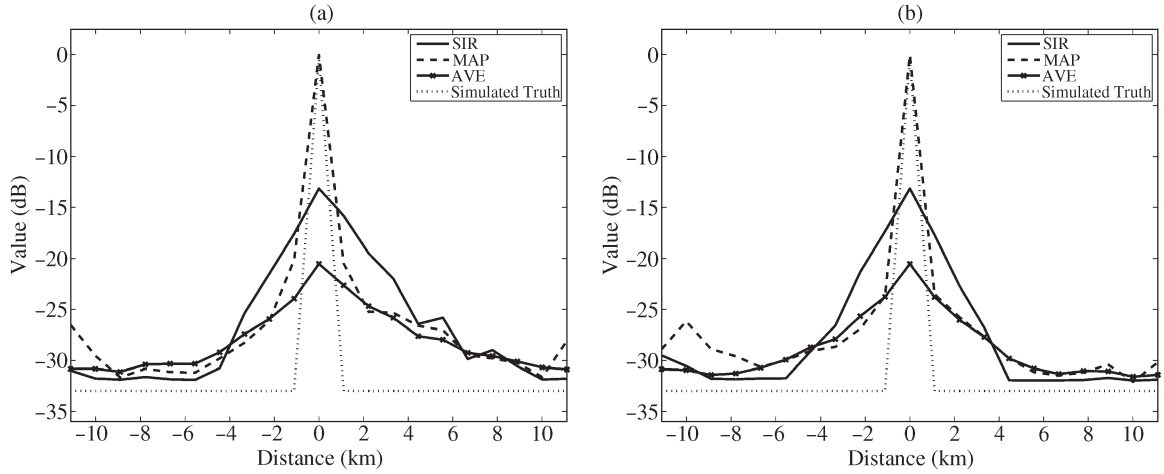

Fig. 4. (a) Horizontal and (b) vertical cuts of the point target and the AVE, SIR, and MAP spatial responses.

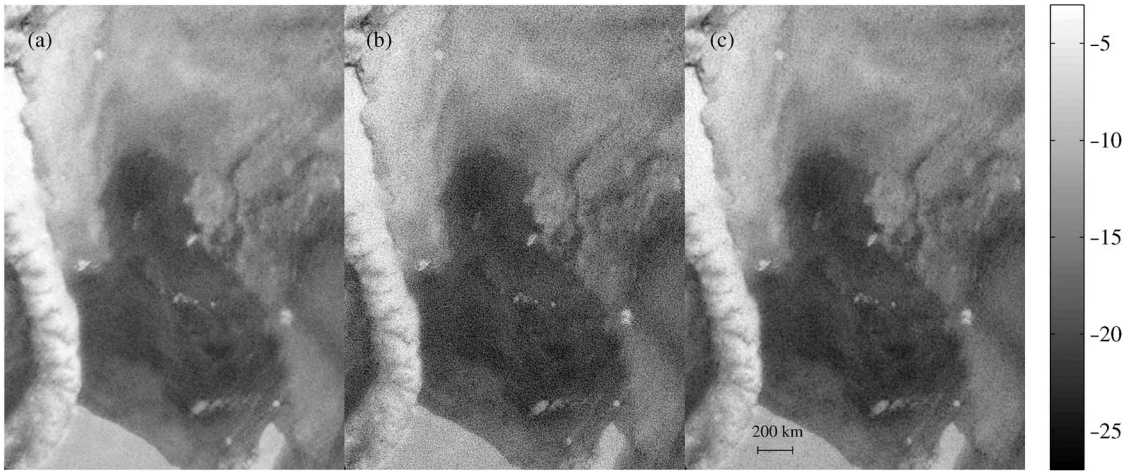

Fig. 5. (a) SIR image, (b) MAP image, and (c) filtered MAP image (in decibels) of the Weddell Sea using SeaWinds v-pol measurements on days 200–204 of 2008.

noise-versus-resolution tradeoff and to evaluate the resolution of the signal that is "highly observable" (i.e., distinguishable from the noise). Such a treatment is beyond the scope of this paper and will be dealt with in a future paper.

G. Noise Reduction

The covariance of the estimates may be too large to be useful for some applications. Potential noise-reduction operations include filtering or averaging of the aperture-filtered samples (prefiltering), low-order model-based estimation, Bayes estimation with a more informative prior (i.e., a prior with lower entropy), and filtering of the reconstructed signal (postfiltering). Prefiltering the aperture-filtered samples reduces both the signal power and the noise power, but it is difficult to track which components of the signal and noise are being attenuated. Low-order model-based methods may introduce aliasing. Bayes estimation reduces noise but relies on knowledge of the signal. Postfiltering of the estimates, however, reduces the total noise power and the total signal power in a way that can be tracked and that does not introduce aliasing. Thus, postfiltering may be suitable for reducing the variability of the estimates if the application requires.

We note that, although there may be more information in the MAP images because they have a higher resolution, there may not be much more useful information because of the higher noise level. Both the methods of low-pass filtering the MAP image and of applying an informative prior [i.e., reducing p in (19)] can be tuned to produce images of comparable quality to the SIR results. Nevertheless, the new approach allows for a more general and more optimal noise-versus-resolution tradeoff. A rigorous treatment of this tradeoff is beyond the scope of this paper and is left for future investigation. For the purpose of this paper, it is sufficient to verify that the new approach can produce images of comparable quality to the previous empirically tested SIR result—thus simultaneously validating the empirical and rigorous approaches.

Fig. 5 shows a SIR image, a MAP image, and a filtered MAP image of the Weddell Sea using four days worth of SeaWinds v-pol data. The bright regions on the left and lower left corner is the Antarctic Peninsula and the Ronne Ice Shelf, respectively. The dark and gray regions are sea ice, while the small bright targets are large tabular icebergs embedded in the sea ice. The filtered MAP image is filtered with a simple 5×5 pixel (i.e., 11 km) sliding averaging window. Although the MAP image is more noisy than the SIR, the filtered MAP

image has a similar noise level and resolves the same structures as the SIR image. This consistency validates the new MAP approach and suggests that the raw MAP estimates contain at least as much information as the SIR results. Further, a more optimal noise-versus-resolution tradeoff may be possible with a filter that accounts for the spatially varying variability of the estimates.

V. Conclusion

This paper has approached scatterometer image reconstruction as the inversion of a noisy aperture-filtered sampling operation, focusing on bandlimited, periodic signals. Reconstruction from noise-free and noisy samples are explored. A more theoretically appropriate reconstruction algorithm is proposed based on MAP estimation, which can reconstruct more details than the SIR algorithm but with enhanced noise. Examples from SeaWinds and ASCAT are presented.

This paper has presented several important observations and results. They include the following: 1) deriving reconstruction operations for the noisy and noise-free cases; 2) demonstrating that in the noise-free case, exact reconstruction is guaranteed only when the sampling operation is fully determined or overdetermined; 3) illustrating that prefiltering and model-based methods are generally not equivalent to postfiltering and may introduce undesirable artifacts, such as aliasing; 4) showing that the conventional-sample spacing (i.e., pixel resolution) for a given application is determined by the bandlimit of the aperture-functions; and 5) presenting a theoretically more appropriate scatterometer image reconstruction algorithm based on MAP estimation.

Future work will consider noise-versus-resolution tradeoffs. Analysis of the quality of the MAP estimates will also be investigated in more detail. The theory will be modified to include aperture-filtered sampling of fields that are related by a pointwise nonlinearity, such as scatterometer ocean-surface-wind field reconstruction. Bayes estimation with a quadratic loss function will also be considered. Connections with synthetic aperture array processing will also be explored.

Appendix I
Discrete Equivalence of Bandlimited Signals

Here, the conditions are considered under which the continuous-index sampling operation $\mathcal{A}$ is equivalent to a discrete linear operation on the conventional samples. First, it is shown that if each $A_n(x)$ is bandlimited, the sampling can be represented by (3). The same result is obtained if $s(x)$ is bandlimited, even if each $A_n(x)$ is not bandlimited. Next, it is shown that if $s(x)$ or each $A_n(x)$ is bandlimited and periodic, $\mathbf{A}$ is a finite-dimensional matrix and can be analyzed with standard linear algebra. An argument for approximating practical signals as bandlimited and periodic is also presented.

A. Discrete Equivalence

First, consider the case in which each $A_n(x)$ is bandlimited to ω_0 and both $s(x)$ and each $A_n(x)$ are in the Hilbert space of square integrable functions (L_2). From conventional reconstruction theory, recall that any function $f(x)$ that is bandlimited to ω_0 can be represented by sinc-function interpolation from the conventional (uniformly spaced) samples $f[x_n]$, and $f(x)$ can thus be expressed as

$$f(x) = \sum_n f[x_n] \text{sinc}\left(\omega_0(x - x_n)\right). \tag{20}$$

If each of the aperture functions $A_n(x)$ are bandlimited to ω_0, each row of the sampling operation in (2) can be written as

$$\begin{aligned}\int A_n(x)s(x)dx &= \int \sum_i A_n[x_i] \text{sinc}\left(\omega_0(x - x_i)\right) s(x)dx \\ &= \sum_i A_n[x_i] \int \text{sinc}\left(\omega_0(x - x_i)\right) s(x)dx \\ &= \sum_i A_n[x_i] s_{\text{BL}}[x_i] \end{aligned} \tag{21}$$

where $s_{\text{BL}}[x_i]$ represents the conventional samples of a bandlimited version of $s(x)$ and $A_n[x_i]$ represents the conventional samples of the aperture functions. The sampling operation can thus be written as a discrete linear operation on the samples of a bandlimited version of the signal.

By a similar argument, if $s(x)$ is bandlimited but the $A_n(x)$'s are not necessarily bandlimited, a similar result is obtained

$$\begin{aligned}\int A_n(x)s(x)dx &= \int \sum_i \text{sinc}\left(\omega_0(x - x_i)\right) A_n(x) s[x_i] \\ &= \sum_i A_{n,\text{BL}}[x_i] s[x_i] \end{aligned} \tag{22}$$

where $A_{n,\text{BL}}[x_i]$ represents the bandlimited version of the nth aperture function.

Therefore, if either $s(x)$ or each $A_n(x)$ is bandlimited, the formulation in (3) is obtained. Here, $\vec{s}$ represents the sinc-function samples of $s(x)$ [i.e., conventional samples of a low-pass filtered version of $s(x)$], and the rows of $\mathbf{A}$ represent sinc-function samples of the aperture functions. In general, $\vec{s}$ and $\mathbf{A}$ are infinite dimensional.

It can be shown that $\mathbf{A}$ and $\vec{s}$ are finite dimensional if either the aperture functions or the signal are bandlimited and periodic. Suppose that each $A_n(x)$ is bandlimited and periodic. Dirichlet-kernel interpolation then reconstructs the aperture functions from conventional samples. Each row of the sampling operation becomes

$$\begin{aligned}\int_{\mathcal{P}} A_n(x)s(x)dx &= \int_{\mathcal{P}} \sum_{i=1}^{\mathcal{P}} A_n[x_i] D\left(\omega_0(x - x_i)\right) s(x)dx \\ &= \sum_{i=1}^{\mathcal{P}} A_n[x_i] \int_{\mathcal{P}} D\left(\omega_0(x - x_i)\right) s(x)dx \\ &= \sum_{i=1}^{\mathcal{P}} A_n[x_i] s_{\text{BL},P}[x_i] \end{aligned} \tag{23}$$

where $D(\omega_0(x - x_i))$ represents the Dirichlet kernel, $s_{\text{BL},P}[x_i]$ represents conventional samples of the bandlimited periodic version of $s(x)$, and $\mathcal{P}$ represents the fundamental period of the aperture functions. Also, if the $A_n(x)$'s can be represented as periodic but are not necessarily bandlimited but $s(x)$ is bandlimited and periodic, by symmetry, a similar result is obtained with the periodic and bandlimited versions of the aperture functions.

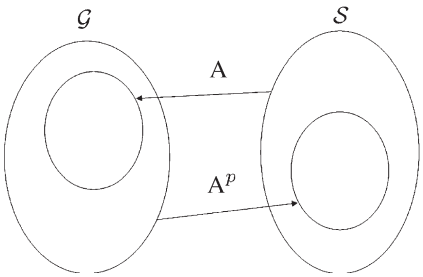

Fig. 6. Spaces associated with the sampling matrix $\mathbf{A}$ and the mapping $\mathbf{A}^p$.

B. Bandlimited Periodic Approximation of Practical Signals

Note the few fundamental qualities of signals in practical applications: they have finite extent, have finite energy, are bounded, and are generally continuous (or at least piecewise continuous). Signals with finite extent can be extended to be periodic, and if they satisfy the Dirichlet conditions, they can be exactly represented by the Fourier series. The Dirichlet conditions for the Fourier series require that a periodic signal be absolutely integrable over a signal period, be of bounded variation, and have finitely many discontinuities in order to be represented exactly by the Fourier series [22]. Most practical signals satisfy these conditions. Furthermore, above some frequency, the energy in the Fourier series of the signal must decrease toward zero since practical signals have finite energy. For a practical signal, a bandlimit and period may be chosen such that a bandlimited periodic version of the signal exists where the approximation error is sufficiently small. Thus, for practical applications, most signals can be approximated arbitrarily close by bandlimited periodic signals.

APPENDIX II
NOISE-FREE RECONSTRUCTION: UNDERDETERMINED CASE

This appendix discusses reconstruction for the noise-free underdetermined case. Different pseudoinverses of the sampling matrix are considered, different methodologies of regularizing the problem are discussed, and a particular reconstruction estimator is proposed.

The underdetermined case occurs if the row rank of $\mathbf{A}$ is less than the number of conventional samples required to represent the bandlimited signal. For this case, there is no one-to-one mapping that maps the range space of $\mathbf{A}$ back to the entire domain of $\mathbf{A}$. If $\mathbf{A}$ is an underdetermined $M \times N$ matrix, then only some signals that can be represented by N conventional samples may be reconstructed. That is, there is a subspace of the domain of $\mathbf{A}$ over which an inverse mapping (i.e., a pseudoinverse) may be defined. Here, restrictions on the class of signals that allow every signal of the class to be recovered for a given sampling are investigated.

The class of recoverable signals is a subspace of the signal space. Fig. 6 shows the spaces associated with the sampling matrix $\mathbf{A}$ and a pseudoinverse $\mathbf{A}^p$. Here, $\mathbf{A} : \mathcal{S} \rightarrow \mathcal{G}$, where $\mathcal{S}$ is the domain and represents the Hilbert space of conventional discrete signals $\vec{s}$ and $\mathcal{G}$ is the codomain and represents the Hilbert space of aperture-filtered signals $\vec{g}$. Although $\mathbf{A}$ is not invertible over the entire domain and codomain, the domain and codomain may be restricted so that $\mathbf{A}$ is bijective over these subspaces. In this case, there is a unique inverse over the subspaces, which is a pseudoinverse of $\mathbf{A}$. A pseudoinverse of $\mathbf{A}$ is any mapping $\mathbf{A}^p : \mathcal{G} \rightarrow \mathcal{S}$ (see Fig. 6) such that $\mathbf{A}^p\mathbf{A}\vec{s} = \vec{s}$ for every $\vec{s}$ in the range space of $\mathbf{A}^p$, where $\mathbf{A}^p$ need not be a linear operator.

There are generally several different subspaces over which an inverse of $\mathbf{A}$ can be defined. Each such subspace is associated with a different pseudoinverse. Restricting the codomain to any subspace of the range space of $\mathbf{A}$ is sufficient to constrain $\mathbf{A}$ to be a surjection (i.e., onto). Moreover, deciding which subspace of $\mathcal{S}$ to use is equivalent to imposing a signal model. That is, a signal model may be chosen whose range space spans a subspace of $\mathcal{S}$. Estimating the parameters of a low-order model and then constructing an estimate of the signal $\vec{s}$ using the model and the estimated parameters constitute an inverse of $\mathbf{A}$ over a subspace of $\mathcal{S}$ (i.e., a pseudoinverse of $\mathbf{A}$).

Note that some pseudoinverses may not preserve all the information in the aperture-filtered samples. In order to preserve all the information in the samples, the pseudoinverse must be reversible by $\mathbf{A}$. This information-preserving constraint can be thought of as requiring each point in the range space of $\mathbf{A}$ to map back to itself through the pseudoinverse followed by the forward projection (i.e., $\mathbf{A}\mathbf{A}^p\vec{g} = \vec{g}$ for all $\vec{g}$ in the range space of $\mathbf{A}$). This property does not necessarily hold for every pseudoinverse mapping $\mathbf{A}^p$, but there always exists a pseudoinverse mapping that is information preserving.

An information-preserving pseudoinverse can be defined using constrained optimization, i.e., by choosing for each point $\vec{g}$ in the range space of $\mathbf{A}$ the point $\vec{s}$ in $\mathcal{S}$ that maps to the point $\vec{g}$ that minimizes some metric $d(\vec{s}, \vec{z})$ defined in $\mathcal{S}$. In other words, we minimize $d(\vec{s}, \vec{z})$ subject to $\vec{g} = \mathbf{A}\vec{s}$, where $d(\vec{s}, \vec{z})$ is a metric that represents the distance between the vector $\vec{s}$ and some predetermined vector $\vec{z}$. $\vec{z}$ may be a particular signal (e.g., an expected signal) for which we want to find the closest signal $\vec{s}$ to $\vec{z}$ that produces the obtained aperture-filtered samples. For many applications, $\vec{z}$ is taken to be $\vec{0}$ so that $d(\vec{s}, \vec{0})$ is a vector norm.

Using the L_2-norm $d(\vec{s}, \vec{0}) = \|x\|_2^2$ produces the Moore–Penrose pseudoinverse [7]. This constrained optimization approach using the L_2-norm is similar to what is done in the AART and Grochenig's irregular sampling theorem [4], [11]. Another common constraint used for nonnegative signals is to maximize the signal entropy $-\sum_i s_i \log(s_i)$ subject to $\vec{g} = \mathbf{A}\vec{s}$. This is the basis for the MART and the SIR algorithm [4]. Although the L_2-norm constraint results in a linear pseudoinverse, a different constraint may produce a nonlinear pseudoinverse for a linear sampling operator. Formulating the pseudoinverse as a constrained optimization problem is powerful because it can be extended to general linear and nonlinear operators.

It is interesting to consider what happens to signals that are not in the range space of the pseudoinverse when the sampling and reconstruction processes are applied. The reconstructed signal is guaranteed to be in the range space of the pseudoinverse; however, if the original $\vec{s}$ is not in the range space of the pseudoinverse, the reconstructed signal may not be the closest signal to $\vec{s}$ in the range space of the pseudoinverse. This artifact is a generalized form of aliasing. For a linear pseudoinverse, if some portions of the original signal that are orthogonal to the range space of the pseudoinverse are not mapped to zero in the sampling operation, the energy in those components is preserved in the aperture-filtered samples. The pseudoinverse

then erroneously maps this energy into its range space. Unless everything orthogonal to the range space of the pseudoinverse is in the null space of the sampling operator, the pseudoinverse introduces aliasing. The Moore–Penrose pseudoinverse forces signal components orthogonal to its range space to be in the null space of the sampling operator and is therefore an antialiasing pseudoinverse, whereas reconstruction using the MART or SIR algorithms may introduce aliasing and may even result in multiple solutions.

If nothing is known about the signal, we suggest that the structure to impose on the signal be a function of $\mathbf{A}$ so as to force the pseudoinverse to be linear, information preserving, and antialiasing. Where no *a priori* knowledge of the signal structure is available and noise is negligible, we propose that the reconstruction be done using the Moore–Penrose pseudoinverse.

Using the Moore–Penrose pseudoinverse implies that for each point in the range space of $\mathbf{A}$, the simplest or lowest energy signal that could have produced the aperture-filtered samples is chosen as the inverse because the Moore–Penrose pseudoinverse can be obtained by minimizing the L_2-norm, as described earlier. This implication is consistent with the notion that signals with low energy are easier to produce and more likely to occur in nature than signals with high energy. The Moore–Penrose pseudoinverse can also be calculated conveniently.

Appendix III
Scatterometer ML Estimator

Here, the scatterometer ML reconstruction estimator is obtained assuming that the sampling matrix is not underdetermined. It is also shown that an analytic form for the estimator may be obtained, although the expression is rather complicated, and multiple solutions (local maxima) are possible.

The ML estimator for the scatterometer noise model searches for the conventional samples $\vec{s}$ that maximize the log-likelihood function

$$L_{\mathrm{ML}} = -\sum_i \left[\frac{\left(g_{\nu,i} - \vec{A}_i^T \vec{s}\right)^2}{2R_{i,i}} + \frac{1}{2}\log\{2\pi R_{i,i}\} \right]. \quad (24)$$

The local maxima are obtained by setting the gradient to zero and solving the resulting system of equations. The partial derivative of L_{ML} with respect to the jth component of $\vec{s}$ is

$$\frac{\partial L_{\mathrm{ML}}}{\partial s_j} = \sum_i \frac{-A_{i,j}}{R_{i,i}} \left[-g_{\nu,i} + \vec{A}_i^T \vec{s} + \alpha_i \vec{A}_i^T \vec{s} + \beta_i/2 - \frac{\left(g_{\nu,i} - \vec{A}_i^T \vec{s}\right)^2 \left(\alpha_i \vec{A}_i^T \vec{s} + \beta_i/2\right)}{R_{i,i}} \right] = \vec{A}_j^T \vec{K}(\vec{s}). \quad (25)$$

Taking the partial derivatives with respect to each component and setting them equal to zero produces the system

$$\mathbf{A}^T \vec{K}(\vec{s}) = \vec{0} \quad (26)$$

which implies that $\vec{K}(\vec{s}) = \vec{0}$ or that $\vec{K}(\vec{s})$ is in the null space of $\mathbf{A}^T$. If $\mathbf{A}$ is fully determined or overdetermined, there is no null space, and the only solutions are when $\vec{K}(\vec{s}) = \vec{0}$. If each element of $\vec{K}(\vec{s})$ is set to zero, cubic equations in $\vec{A}_i^T \vec{s}$ are obtained for each i that have at least one and up to three real roots. Each of these roots can be solved analytically via the cubic equation. The solutions to the entire system of equations result in a linear system in $\vec{s}$. That is, if $\vec{z}$ is a solution to the cubic system of equations $\vec{K}(\vec{s}) = \vec{0}$, then $\vec{A}_i^T \vec{s} = z_i$ for each i, and the linear system $\mathbf{A}\vec{s} = \vec{z}$ is obtained. If $\mathbf{A}$ is fully determined, then $\hat{\vec{s}} = \mathbf{A}^{-1}\vec{z}$, and if it is overdetermined, then $\hat{\vec{s}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\vec{z}$.

There may be more than one maximum of L_{ML}. To find the ML estimate, all of the local maxima must be found and checked to see which one has the highest L_{ML} value. This process is too cumbersome for a practical implementation and so a simple gradient search is used to find a local maximum near an initial guess.

References

[1] D. Long, M. Drinkwater, B. Holt, S. Saatchi, and C. Bertoia, "Global ice and land climate studies using scatterometer image data," *EOS, Trans. Amer. Geophys. Union*, vol. 82, no. 43, p. 503, 2001.

[2] P. Yu, D. Clausi, and S. Howell, "Fusing AMSR-E and QuikSCAT imagery for improved sea ice recognition," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 1980–1989, Jul. 2009.

[3] D. G. Long, P. J. Hardin, and P. T. Whiting, "Resolution enhancement of spaceborne scatterometer data," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 3, pp. 75–700, May 1993.

[4] D. S. Early and D. G. Long, "Image reconstruction and enhanced resolution imaging from irregular samples," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 291–302, Feb. 2001.

[5] Y. V. Shkvarko, "Unifying experiment design and convex regularization techniques for enhanced imaging with uncertain remote sensing data—Part I: Theory," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 82–95, Jan. 2010.

[6] Y. V. Shkvarko, "Unifying experiment design and convex regularization techniques for enhanced imaging with uncertain remote sensing data—Part II: Adaptive implementation and performance issues," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 96–111, Jan. 2010.

[7] T. D. Moon and W. C. Stirling, *Mathemetical Methods and Algorithms for Signal Processing*, T. Robbins, Ed. Upper Saddle River, NJ: Prentice-Hall, 2000.

[8] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, A. V. Oppenheim, Ed., 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1999.

[9] H. G. Feichtinger and K. Grochenig, "Irregular sampling theorems and series expansions of band-limited functions," *J. Math. Anal. Appl.*, vol. 167, no. 2, pp. 530–556, Jul. 1992.

[10] H. G. Feichtinger and K. Grochenig, "Iterative reconstruction of multivariate band-limited functions from irregular sampling values," *SIAM J. Math. Anal.*, vol. 23, no. 1, pp. 244–261, Jan. 1992.

[11] K. Grochenig, "Iterative reconstruction of multivariate band-limited functions from irregular sampling values," *Math. Comput.*, vol. 59, no. 199, pp. 181–194, Jul. 1992.

[12] R. Franz and D. G. Long, "Multidimensional reconstruction from irregular samples," in *Proc. Opt. Soc. Amer. Tech. Dig.—Signal Recovery and Synthesis*, Albuquerque, NM, Nov. 5–8, 2001, pp. 135–137.

[13] T. Lungu, *QuikSCAT Science Data Product Users Manual Overview and Geophysical Data Products*. Pasadena, CA: Jet Propulsion Lab., Sep. 2006.

[14] *ASCAT Wind Product User Manual*, KNMI, The Netherlands, Aug. 2009. [Online]. Available: www.osi-saf.org/biblio/docs/ss3_pm_ascat_1_7.pdf

[15] B. A. Williams, "Signal processing methods for ultra high resolution scatterometry," Ph.D. dissertation, Brigham Young Univ., Provo, UT, 2010.

[16] I. S. Ashcraft and D. G. Long, "The spatial response function of SeaWinds backscatter measurements," in *Proc. SPIE Earth Observing Syst.*, W. L. Barnes, Ed., Bellingham, WA, Aug. 2003, vol. 5151, pp. 609–618.

[17] R. Lindsley and D. G. Long, "Adapting the SIR algorithm to ASCAT," Microw. Remote Sens. Lab., Brigham Young Univ., Provo, UT, 2009, Tech. Rep.
[18] B. A. Williams, M. P. Owen, and D. G. Long, "The ultra high resolution QuikSCAT product," in *Proc. IEEE Radar Conf.*, Pasadena, CA, May 2009, pp. 1–6.
[19] R. E. Fischer, "Standard deviation of scatterometer measurements from space," *IEEE Trans. Geosci. Remote Sens.*, vol. GRS-10, no. 2, pp. 106–113, Apr. 1972.
[20] S. D. Gupta, S. K. Mitra, P. S. S. N. V. P. Rao, and J. K. Grosh, *Selected Papers of C.R. Rao*. New York: Wiley, 1994.
[21] B. W. Stiles, S. M. Hristova-Veleva, R. S. Dunbar, S. Chan, S. L. Durden, D. Esteban-Fernandez, E. Rodriguez, W. L. Poulsen, R. W. Gaston, and P. S. Callahan, "Obtaining accurate ocean surface winds in hurricane conditions: A dual-frequency scatterometry approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 8, pp. 3101–3113, Aug. 2010.
[22] A. V. Oppenheim and A. S. Willsky, *Signals and Systems*, A. V. Oppenheim, Ed., 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

Brent A. Williams (S'04–M'08) received the B.S. degree in electrical engineering and the Ph.D. degree in electrical engineering from Brigham Young University, Provo, UT, in 2005 and 2010, respectively, where he has recently finished his dissertation titled "Signal Processing Methods for Ultra High Resolution Scatterometry."

From 2005 to 2010, he was with the Microwave Earth Remote Sensing Laboratory exploring ultra-high resolution ocean wind scatterometry and σ^0 image reconstruction. He is currently with the Jet Propulsion Laboratory, Pasadena, CA, continuing research in scatterometry.

Dr. Williams is a member of Eta Kappa Nu. He was the recipient of the IEEE Geoscience and Remote Sensing Society Interactive Session Prize Paper Award in 2006 and the IEEE Geoscience and Remote Sensing Society Mikio Takagi Student Prize Award in 2007 and 2010, respectively.

David G. Long (S'80–SM'98–F'08) received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1989.

From 1983 to 1990, he was with the Jet Propulsion Laboratory (JPL), National Aeronautics and Space Administration (NASA), Pasadena, CA, where he developed advanced radar remote-sensing systems. While at JPL, he was the Project Engineer on the NASA Scatterometer Project which flew from 1996 to 1997. He also managed the SCANSCAT project, the precursor to SeaWinds, which was launched in 1999 and 2002. He is currently a Professor with the Electrical and Computer Engineering Department, Brigham Young University (BYU), Provo, UT, where he teaches upper division and graduate courses in communications, microwave remote sensing, radar, and signal processing and is the Director of the BYU Center for Remote Sensing. He is the Principal Investigator on several NASA-sponsored research projects in remote sensing. He has over 390 publications in signal processing and radar scatterometry. His research interests include microwave remote sensing, radar theory, space-based sensing, estimation theory, signal processing, and mesoscale atmospheric dynamics.

Dr. Long is currently an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS. He was the recipient of the NASA Certificate of Recognition for several times.