Feature-preserving regularization method for complex-valued inverse problems with application to coherent imaging

Müjdat Çetin
Sabanci University
Faculty of Engineering and Natural Sciences
34956, Istanbul, Turkey
E-mail: mcetin@sabanciuniv.edu

William C. Karl
Boston University
Multi-Dimensional Signal Processing Laboratory
Boston, Massachusetts 02215

Alan S. Willsky, MEMBER SPIE
Massachusetts Institute of Technology
Laboratory for Information and Decision Systems
Cambridge, Massachusetts 02139

Abstract. We propose a method for feature-preserving regularized reconstruction in coherent imaging systems. In our framework, image formation from measured data is achieved through the minimization of a cost functional, designed to suppress noise artifacts while preserving features such as object boundaries in the reconstruction. The cost functional includes nonquadratic regularizing constraints. Our formulation effectively deals with the complex-valued and potentially random-phase nature of the scattered field, which is inherent in many coherent systems. We solve the challenging optimization problems posed in our framework by developing and using an extension of half-quadratic regularization methods. We present experimental results from three coherent imaging applications: digital holography, synthetic aperture radar, and ultrasound imaging. The proposed technique produces images where coherent speckle artifacts are effectively suppressed, and important features of the underlying scenes are preserved. *© 2006 Society of Photo-Optical Instrumentation Engineers.* [DOI: 10.1117/1.2150368]

Subject terms: coherent imaging; image reconstruction; regularization; computed imaging; half-quadratic regularization; nonquadratic optimization; inverse problems.

Paper 040843R received Nov. 10, 2004; revised manuscript received Apr. 22, 2005; accepted for publication May 24, 2005; published online Jan. 10, 2006.

1 Introduction

This paper addresses image reconstruction problems in coherent imaging. Coherent imaging is based on recording spatial and/or temporal variations in both the intensity of a scattered field and its phase.[1] Many microwave, optical, and acoustic sensing applications use coherent imaging, and particular modalities include synthetic-aperture radar (SAR), holography, sonar, ultrasound, and laser imaging, among others. In both coherent and incoherent imaging tasks, reconstruction of an image from observed data is often an ill-posed inverse problem. Solution of such inverse problems can be achieved through regularization methods, which turn the problem into a well-posed one and prevent the amplification of measurement noise during the reconstruction process. However, one limitation of straightforward regularization methods, such as Tikhonov regularization,[2] is the suppression of important features in the resulting imagery, such as edges. Recently this issue has been successfully addressed by feature-preserving regularization techniques in incoherent imaging applications, such as restoration of blurred and noisy optical images[3] and reconstruction in x-ray tomography.[4]

Coherent image reconstruction poses additional challenges that do not appear in incoherent imaging. First, the signals involved are in general complex-valued. Furthermore, in many problems, including SAR and holography of diffuse objects, the phase of the scattered field is a highly random quantity.* This leads to two complications. First, due to constructive and destructive interference of scatterers within a resolution cell, conventional coherent images suffer from speckle artifacts. (Speckle appears when the surface being imaged has roughness at the scale of the illuminating wavelength.) Second, due to the complex-valued and possibly random-phase nature of the fields, straightforward application of image reconstruction methods originally designed for incoherent imaging may not produce accurate reconstructions, as we experimentally demonstrate in Sec. 3.

To address these challenges, we propose a feature-preserving regularization method specifically for coherent imaging tasks. The approach we present involves the minimization of a cost functional that contains *nonquadratic regularization* constraints. Such nonquadratic constraints have been shown to lead to feature preservation by preferring reconstructions that are *sparse* in terms of the features of interest.[6] Our framework is general enough to handle various features (as we demonstrate later), but for the sake of concreteness at this point, let us assume that the features of interest are the boundaries between distinct physically meaningful regions in the scene. The goal then is to reconstruct images where various imaging artifacts and noise are suppressed, while object and region boundaries (edges) are

*This property is known to enable high-quality reconstructions from limited Fourier-offset data in coherent imaging.[5] For this reason, Fourier transform holograms are often constructed using a diffuser to impart essentially random phase to each point in the original scene before recording.

0091-3286/2006/$22.00 © 2006 SPIE

preserved. The regularization constraints in our framework achieve artifact suppression by imposing smoothness on the *magnitudes* of the reconstructed complex-valued field reflectivities (or transmission coefficients). The nonquadratic aspect of these regularizing functionals leads to edge preservation, similar to the case in incoherent imaging problems.[3,4] To solve the resulting optimization problems, we provide a formal extension of half-quadratic regularization techniques[7] to complex-valued, random-phase fields. This constitutes the major technical contribution of our work.

There are a number of publications that are related to some of the coherent-imaging issues that we address. The implications of the random-phase nature of coherent images in terms of the quality of the reconstructions has been analyzed in Refs. 5 and 8. The work in Ref. 9 presents a maximum likelihood technique for reconstructing complex-valued, random-phase images from Fourier-offset data using the expectation-maximization algorithm. Bayesian techniques have been used for filtering complex-valued, speckled images in Ref. 10, and for ultrasound Doppler spectral analysis based on autoregressive models in Ref. 11. A technique for image reconstruction from noisy digital holograms based on the method of projection onto convex sets (POCS) has been developed in Ref. 12. These last three papers are somewhat related to our approach in that they use regularizing constraints. A number of more recent publications have a closer relation to our perspective for coherent imaging, in particular in their emphasis on preservation of edges or other features. A Bayesian approach for the nonlinear inverse scattering problem of tomographic imaging using microwave or ultrasound probing has been proposed in Ref. 13. In Refs. 14 and 15, maximum-entropy regularization has been used for image reconstruction from sparsely sampled coherent field data. The work in Ref. 16 proposes a regularized autoregressive model for spectral estimation, with application to medical ultrasonic radio-frequency images. Another method for spectral estimation involves regularization through a circular Gibbs-Markov model.[17] A statistical deconvolution technique for diffuse ultrasound scattering has been proposed in Ref. 18, where sampling techniques are used for inference. In Ref. 19, anisotropic diffusion[20] has been used for ultrasound speckle reduction and coherence enhancement. The total variation-based regularization method proposed in Ref. 21 has been applied to coherent imaging, in particular to near-field acoustic holography. Finally, in Ref. 22, a penalized-likelihood image reconstruction technique has been proposed for image-plane holography, which uses incoherent illumination.

Our approach is significantly different from this body of previous work in a number of ways. First, we consider the random-phase aspect (and deal with the effects of speckle) much more explicitly than any of the previous papers on inverse problems in coherent imaging. Second, the structures of the energy functionals used in our framework are quite different from what has been used in previous work, and this structure allows the use of a variety of regularizing constraints within a single framework. Third, the algorithm we use for optimization, namely an extension of half-quadratic regularization, is new. We demonstrate the performance of the proposed method on examples from a number of coherent imaging applications. With enhanced speckle and artifact suppression, as well as feature preservation, the images produced by our method appear to yield more accurate reconstructions than conventional coherent imaging techniques.

In Sec. 2, we present our nonquadratic regularization-based approach. We first develop the method with ℓ_p-norm-based potential functionals, and then extend it to other nonquadratic potentials. Section 3 contains the experimental results, and we conclude in Sec. 4.

2 Nonquadratic Regularization for Complex-Valued Problems

This section contains the description of the nonquadratic technique we propose in this paper. We start by describing the general form of the observation models we consider. We then formulate an optimization problem for coherent imaging, which involves a cost functional based on ℓ_p norms. To minimize this cost functional, we propose an algorithm based on half-quadratic regularization, and provide a statistical interpretation of this strategy. Finally we generalize our method to incorporate nonquadratic cost functionals other than ℓ_p norms.

2.1 *Observation Model*

In this paper, we are concerned with inverse problems in which the sensor measurements y are related to the underlying, unknown field f, through a Fredholm integral equation of the first kind:

$$y(\mathbf{x}) = \int_{\Omega} T(\mathbf{x},\mathbf{x}')f(\mathbf{x}')\,\mathrm{d}\mathbf{x}' + w(\mathbf{x}), \tag{1}$$

where Ω is the spatial region of interest for the reconstruction, and w is additive measurement noise. The argument of f corresponds to 2-D or 3-D spatial coordinates, and the arguments of y and w depend on the domain of the measurements in specific applications.

We assume that the integral kernel T, which models the relationship between the underlying field and the measured data, is known. For example, T may be a band-limited, possibly frequency-offset Fourier transform operator, where the physics of the problem, the sensor parameters, and the observation geometry determine the exact structure. Another example for T, used in tomographic imaging modalities, is projection-type operators, related to the Radon transform.[23] Yet another form arising in many applications is convolutional operators. For some particular observation models that are of interest in our work (and that we use in our experimental analysis), see Refs. 24 for digital holography, Refs. 25 and 26 for SAR, and Refs. 14, 27, and 28 for ultrasound.

In many coherent imaging applications, which involve, e.g., multiple scattering and other second-order phenomena, the exact equations governing the observation process are actually nonlinear. In such scenarios, approximate linear observation models as in Eq. (1) can be obtained through first-order solutions, which exclude all but primary scattering. Such linear models include the well-known Born approximation and the physical optics approximation. These linear approximations yield acceptable results in many

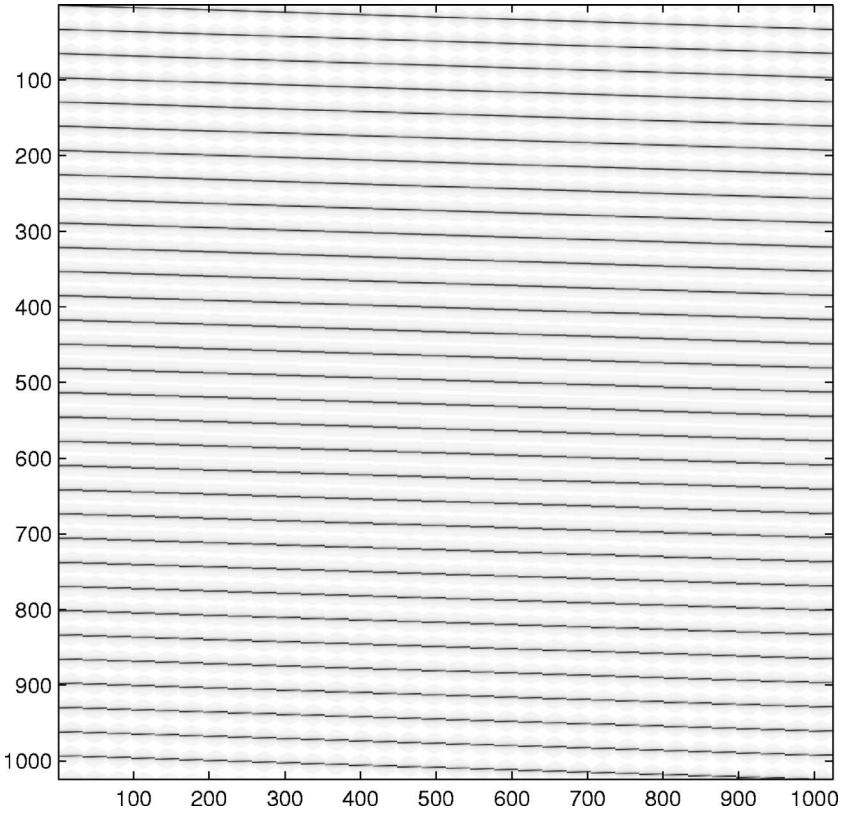

Fig. 1 Grayscale plot (black corresponds to the maximum value, and white to the minimum) of the magnitude of the elements in a SAR projection matrix for a 32×32 field. The radar in this example operates in the X band with a center frequency of 10 GHz, and the underlying scene is viewed through an angular span of 2.3 deg.

practical situations, where our techniques are directly applicable. On the other hand, it is certainly of interest to develop inversion techniques based on more accurate nonlinear models. While we address only linear problems in this paper, the key ideas we present are potentially useful in nonlinear problems as well, and our method could be generalized to such cases.

In practice, we discretize the relationship in Eq. (1) and use the following model for the coherent observation process:

$$\mathbf{y} = \mathbf{T}\mathbf{f} + \mathbf{w}, \tag{2}$$

where $\mathbf{y}, \mathbf{f}$, and $\mathbf{w}$ are the sampled data, the unknown image, and noise, respectively, all column-stacked as vectors. Similarly, $\mathbf{T}$ is a matrix representing the discrete observation kernel. To provide some flavor of such discrete operators, in Fig. 1 we illustrate a tomographic projection operator that arises in one of the applications of interest in this paper, namely, SAR. The operator is complex-valued, and we show only the magnitudes of the elements of the matrix as a grayscale plot. Each column of the matrix corresponds to one spatial location in the underlying image, and describes how the reflectivity at that location contributes to the projectional radar observations. Each row of the matrix corresponds to one particular data point (one sample in the discretized radar return at a particular observation angle), and describes the effect of various spatial locations in the scene on that data point.

Given the observation model in Eq. (2), the objective is to obtain a reconstruction of $\mathbf{f}$, based on the data $\mathbf{y}$. Conventional image formation techniques vary depending on the particular modality and sensor model, and include algorithms based on beamforming, filtered backprojection, and inverse Fourier transformation, among others.

2.2 *Cost Functional Based on ℓ_p Norms*

We propose to find the reconstructed image $\hat{\mathbf{f}}$ as the minimizer of the following cost functional:

$$J_0(\mathbf{f}) = \|\mathbf{y} - \mathbf{T}\mathbf{f}\|_2^2 + \lambda \|\mathbf{D}|\mathbf{f}|\|_p^p, \tag{3}$$

where $\|\cdot\|_p$ denotes the ℓ_p norm, $\mathbf{D}$ is a matrix to be described below, $|\mathbf{f}|$ denotes the vector of magnitudes of the complex-valued vector $\mathbf{f}$, and $\lambda, p < 2$ are scalar parameters.[†] Note that the formulation of Eq. (3) takes into account the forward model $\mathbf{T}$ and starts from the observed sensor data $\mathbf{y}$, and hence is not simply a postprocessing of a formed image.

The first term of $J_0(\mathbf{f})$ in Eq. (3) is a data fidelity term, while the second term incorporates prior information regarding both the behavior of the field $\mathbf{f}$ and the nature of the features of interest in the resulting reconstructions. In par-

[†]When $p < 1$, the triangle inequality is not satisfied and it would be more precise to use the term "quasi-norm" rather than "norm." However, we ignore this subtlety and use the term "ℓ_p norm" for any value of p.

ticular, the nonquadratic structure of the second term provides feature preservation,[3,4] where the matrix $\mathbf{D}$ determines the kind of features to be preserved. For example, if we are interested in reconstructing images consisting of spatially extended objects and regions, with slowly varying physical properties (such as reflectivities) within the regions, then a good choice for $\mathbf{D}$ is a discrete approximation to the 2-D spatial derivative operator (gradient). With this choice, the second term in Eq. (3) becomes a piecewise smoothness constraint, imposing smoothness within regions and allowing sharp transitions across the region boundaries, leading to edge preservation. In Sec. 3, we show examples demonstrating the use of this choice of $\mathbf{D}$ on digital holography and SAR imaging. For a discussion of the structure of 2-D discrete derivative operators, see the Appendix (Sec. 5.1).

While edge-preserving reconstruction is of interest in many coherent-imaging tasks, one might also be interested in other features. For example, rather than spatially extended objects, an application might involve imaging spatially localized scatterers. In that case, we would be interested in preserving the scattering amplitudes of the strong scatterers in the scene, while suppressing noise and artifacts. In our framework, this could be achieved by choosing $\mathbf{D}$ to be an identity operator in Eq. (3). Such constraints have been shown to lead to superresolution.[29] In Sec. 3, we show examples demonstrating the use of this choice of $\mathbf{D}$ on ultrasound imaging.

In order to avoid problems due to nondifferentiability of the ℓ_p norm around the origin when $p \leqslant 1$, we use a smooth approximation to the ℓ_p-norm in Eq. (3).[3] This leads to the following slightly modified cost functional to be used in practice for numerical purposes:

$$J(\mathbf{f}) = \|\mathbf{y} - \mathbf{T}\mathbf{f}\|_2^2 + \lambda \sum_{i=1}^{M} [|(\mathbf{D}|\mathbf{f}|)_i|^2 + \epsilon]^{p/2}, \tag{4}$$

where $\epsilon \geqslant 0$ is a small constant, $(\cdot)_i$ denotes the i'th element of a vector, and M is the length of the vector $\mathbf{D}|\mathbf{f}|$. Note that $J(\mathbf{f}) \to J_0(\mathbf{f})$ as $\epsilon \to 0$.[‡]

Nonquadratic regularizing constraints such as ℓ_p norms have previously been shown to produce feature-preserving solutions in problems such as image restoration[3] and x-ray tomography,[4] where the signals involved are real-valued. In contrast, we are interested in coherent systems such as SAR and holography, where the processed signals are complex-valued. In many cases of interest, the phase of the unknown complex-valued field $\mathbf{f}$ is highly random, and uncorrelated with the phase at neighboring pixels. Based on this observation, in such coherent imaging problems, regularizing constraints such as smoothness should be applied explicitly to the *magnitudes* $|\mathbf{f}|$ of the complex-valued reflectivities $\mathbf{f}$. In our framework, this is achieved through the expression $\mathbf{D}|\mathbf{f}|$ in Eq. (4). This nonlinearity in $\mathbf{f}$ makes the optimization problem more challenging than those arising in incoherent imaging applications. In the next subsection, we propose an extension of half-quadratic regularization methods[7] to complex-valued, random-phase fields for achieving efficient and robust numerical solution of the optimization problems of the form (4), posed in our framework.

2.3 *Half-Quadratic Regularization for Coherent Imaging*

The main idea in half-quadratic regularization is to introduce and optimize a new cost functional, which has the same minimum as the original nonquadratic cost functional [in our case, $J(\mathbf{f})$], but one which can be manipulated with linear algebraic methods. In incoherent imaging applications, such a new cost functional is obtained by augmenting the original cost functional with an *auxiliary* vector.

Currently available half-quadratic regularization methods designed for incoherent imaging cannot handle the more complicated structure of the optimization problems involved in coherent imaging. In order to deal with such complications, we propose using two auxiliary vectors, $\mathbf{b}$ and $\mathbf{s}$, matched to the structure of the problem, to form an augmented cost functional $K(\mathbf{f},\mathbf{b},\mathbf{s})$ which satisfies

$$\inf_{\mathbf{b},\mathbf{s}} K(\mathbf{f},\mathbf{b},\mathbf{s}) = J(\mathbf{f}). \tag{5}$$

In particular, we construct $K(\mathbf{f},\mathbf{b},\mathbf{s})$ in such a way that it is quadratic in $\mathbf{f}$ (hence the name half-quadratic) and easy to minimize in $\mathbf{b}$ and $\mathbf{s}$. Then the minimization of $K(\mathbf{f},\mathbf{b},\mathbf{s})$ can be performed through a block coordinate descent approach.

Now, let us consider our particular cost functional $J(\mathbf{f})$ of Eq. (4). We can show that the following augmented cost functional $K(\mathbf{f},\mathbf{b},\mathbf{s})$ satisfies the relationship (5) for the particular $J(\mathbf{f})$ of Eq. (4) (see Appendix, Sec. 5.2):

$$K(\mathbf{f},\mathbf{b},\mathbf{s}) = \|\mathbf{y} - \mathbf{T}\mathbf{f}\|_2^2 + \lambda \sum_{i=1}^{M} \left\{ \mathbf{b}_i [|(\mathbf{D}\mathbf{S}\mathbf{f})_i|^2 + \epsilon] + \left(\frac{p}{2\mathbf{b}_i}\right)^{\frac{p}{2-p}} \left(1 - \frac{p}{2}\right) \right\}, \tag{6}$$

where

$$\mathbf{S} = \mathrm{diag}\{\exp(-j\mathbf{s}_l)\}, \tag{7}$$

with $\mathbf{s}_l$ the l'th element of the vector $\mathbf{s}$, and $\mathrm{diag}\{\cdot\}$ a diagonal matrix whose l'th diagonal element is given by the expression inside the braces. Due to Eq. (5), $J(\mathbf{f})$ and $K(\mathbf{f},\mathbf{b},\mathbf{s})$ share the same minima in $\mathbf{f}$. Note that $K(\mathbf{f},\mathbf{b},\mathbf{s})$ is a quadratic function with respect to $\mathbf{f}$.[§] We benefit from the half-quadratic structure through the use of an iterative block coordinate descent method on $K(\mathbf{f},\mathbf{b},\mathbf{s})$, in order to find the field $\hat{\mathbf{f}}$ that also minimizes $J(\mathbf{f})$:

$$\hat{\mathbf{s}}^{(n+1)} = \arg\min_{\mathbf{s}} K(\hat{\mathbf{f}}^{(n)}, \hat{\mathbf{b}}^{(n)}, \mathbf{s}), \tag{8}$$

[‡]Note that there is still some nondifferentiability left in $J(\mathbf{f})$, due to $|\mathbf{f}|$. One could in principle apply a similar smooth approximation for this term. However, to keep the notation simple, we ignore this subtlety in our development. One could avoid any practical difficulties this might cause simply by defining the phase at the origin of the complex plane to be zero.

[§]We have obviously omitted the recipe for finding a valid $K(\mathbf{f},\mathbf{b},\mathbf{s})$ from $J(\mathbf{f})$ here. We just want to point out that, given any feature-preserving cost functional $J(\mathbf{f})$, the augmented cost functional can be found by using convex duality relationships, and we refer the interested reader to Ref. 7.

$$\hat{\mathbf{b}}^{(n+1)} = \arg\min_{\mathbf{b}} K(\hat{\mathbf{f}}^{(n)}, \mathbf{b}, \hat{\mathbf{s}}^{(n+1)}), \tag{9}$$

$$\hat{\mathbf{f}}^{(n+1)} = \arg\min_{\mathbf{f}} K(\mathbf{f}, \hat{\mathbf{b}}^{(n+1)}, \hat{\mathbf{s}}^{(n+1)}), \tag{10}$$

where n denotes the iteration number. Using results from Sec. 5.2, we obtain

$$\hat{\mathbf{s}}_i^{(n+1)} = \phi[(\hat{\mathbf{f}}^{(n)})_i], \tag{11}$$

$$\hat{\mathbf{b}}_i^{(n+1)} = \frac{p}{2[(\mathbf{D}\hat{\mathbf{S}}^{(n+1)}\hat{\mathbf{f}}^{(n)})_i^2 + \epsilon]^{1-p/2}}, \tag{12}$$

$$[\mathbf{T}^H\mathbf{T} + \lambda(\hat{\mathbf{S}}^{(n+1)})^H\mathbf{D}^T\mathrm{diag}\{\hat{\mathbf{b}}_i^{(n+1)}\}\mathbf{D}\hat{\mathbf{S}}^{(n+1)}]\hat{\mathbf{f}}^{(n+1)} = \mathbf{T}^H\mathbf{y}, \tag{13}$$

where $\phi[z]$ denotes the phase of the complex number z. We can substitute Eqs. (11) and (12) into (13) to obtain a single iterative expression for $\hat{\mathbf{f}}^{(n+1)}$, which would then constitute the overall iterative algorithm.

Note that each iteration in Eq. (13) requires the solution of a set of linear equations for the unknown $\hat{\mathbf{f}}^{(n+1)}$. The coefficient matrix of this set of equations is Hermitian, positive semidefinite, and usually sparse. Hence these equations may themselves be efficiently solved using iterative approaches. We use the conjugate gradient (CG) algorithm for this solution, and terminate it when the ℓ_2 norm of the relative residual becomes smaller than a threshold $\delta_{\mathrm{CG}} > 0$.[30] We run the iteration (13) until $\|\hat{\mathbf{f}}^{(n+1)} - \hat{\mathbf{f}}^{(n)}\|_2^2 / \|\hat{\mathbf{f}}^{(n)}\|_2^2 < \delta$, where $\delta > 0$ is a small constant. In the Appendix (Sec. 5.3), we show that this algorithm is convergent in terms of the cost functional. For algorithms of this type, stronger results on the convergence of the iterates exist,[4,31] requiring certain assumptions on the nature of the cost functionals[4] or on the nature of the local minima.[31] For the specific algorithm we present here, we have not yet carried out such a more detailed analysis. In our algorithm, we use a stopping criterion based on the relative change in the iterates $\hat{\mathbf{f}}^{(n)}$ as stated, and we have not run into any convergence problems in practice. In general, the algorithm appears to be reaching a local minimum from any initialization.

2.4 *Statistical Interpretation of Half-Quadratic Regularization*

It is well known that optimization problems of the form in Eq. (3) can also be interpreted as statistical estimation problems (see, e.g., Ref. 32). In particular, the same optimization problem is reached when we try to find the maximum *a posteriori* (MAP) estimate of the field $\mathbf{f}$ based on the data $\mathbf{y}$ using a Gaussian, independent identically distributed noise model, together with a generalized Gaussian prior model for the field reflectivity magnitudes, where the spatial dependence structure is governed by the matrix $\mathbf{D}$. The phase distribution is assumed to be uniform and spatially independent. As an example, when $p=1$, we have a Laplacian prior model for the field magnitudes. This heavy-tailed nature of the prior distribution is what leads to preservation of features such as edges. Note that the prior distribution here is *non-Gaussian*, and spatially *stationary*.

Now, let us interpret our half-quadratic regularization-based algorithm statistically. First note that the cost functional in Eq. (6) is a quadratic function of the field $\mathbf{f}$. Consequently, the coordinate-descent-based minimization in Eqs. (11)–(13) essentially solves a sequence of quadratic minimization problems for the field [although this is not explicitly shown, it might be observed from the linear structure of the iteration in Eq. (13)]. However, the quadratic problems contain field-dependent weights involving the auxiliary vectors $\mathbf{b}$ and $\mathbf{s}$. From an estimation standpoint, we essentially have a Gaussian prior for the field, but the distribution is nonstationary due to the field-dependent weighting, which is adaptively determined. Hence, the half-quadratic regularization-based algorithm might be viewed as replacing the original stationary, non-Gaussian problem with a series of nonstationary but Gaussian problems.

Table 1 Families of potential functionals used. Here p is a parameter determining the shape of the functionals, and ϵ is a small smoothing constant.

$\psi_1(x)$	$(x^2+\epsilon)^{p/2}$
$\psi_2(x)$	$\dfrac{(x^2+\epsilon)^{p/2}}{1+(x^2+\epsilon)^{p/2}}$
$\psi_3(x)$	$\log[1+(x^2+\epsilon)^{p/2}]$

2.5 *Extension to Other Nonquadratic Functionals*

In Sec. 2.2, we have formulated the image reconstruction problem using a particular family of regularizing functionals, namely ℓ_p norms. We now generalize our framework and iterative algorithm to incorporate a wider range of potentially useful choices, which have previously found use in incoherent image restoration and reconstruction problems.[4,7,33,34] To this end, let us consider the following general form for the cost functional:

$$J(\mathbf{f}) = \|\mathbf{y} - \mathbf{T}\mathbf{f}\|_2^2 + \lambda \sum_i \psi((\mathbf{D}|\mathbf{f}|)_i), \tag{14}$$

where ψ denotes the regularizing functional.

Three particular classes of functionals ψ we consider in this paper are shown in Table 1.[‖] Note that the use of ψ_1 leads to constraints in terms of approximate ℓ_p norms, which is precisely what we have discussed in Sec. 2.2. The potential functional ψ_2 is based on previous work in Ref. 33. Special cases of ψ_2 for $p=1$ and $p=2$ yield the potential functionals used in Refs. 7 and 4, respectively. Finally, ψ_3 is a generalized version of the potential functional proposed in Ref. 34. Note that these potential functionals can more generally be expressed in terms of x/Δ, where Δ is a scaling parameter. We use a fixed Δ, and omit it in our analysis for notational simplicity.

[‖]One might subtract an appropriate constant from each potential functional to set $\psi_k(0)=0$ $(k=1, 2, 3)$; however, we have chosen not to do so in Table 1, to keep the notation simpler.

Table 2 The updates for the auxiliary variable **b** for each of the three potential functionals

Potential functional	Associated $\hat{\mathbf{b}}_i^{(n+1)}$
ψ_1	$\dfrac{p/2}{[(\mathbf{D}\hat{\mathbf{s}}^{(n+1)}\hat{\mathbf{f}}^{(n)})_i^2+\epsilon]^{1-p/2}}$
ψ_2	$\dfrac{p/2}{[(\mathbf{D}\hat{\mathbf{s}}^{(n+1)}\hat{\mathbf{f}}^{(n)})_i^2+\epsilon]^{1-p/2}\{[(\mathbf{D}\hat{\mathbf{s}}^{(n+1)}\hat{\mathbf{f}}^{(n)})_i^2+\epsilon]^{p/2}+1\}^2}$
ψ_3	$\dfrac{p/2}{[(\mathbf{D}\hat{\mathbf{s}}^{(n+1)}\hat{\mathbf{f}}^{(n)})_i^2+\epsilon]^{1-p/2}\{[(\mathbf{D}\hat{\mathbf{s}}^{(n+1)}\hat{\mathbf{f}}^{(n)})_i^2+\epsilon]^{p/2}+1\}}$

We minimize $J(\mathbf{f})$ in Eq. (14) by using the half-quadratic regularization-based coordinate descent strategy in Eqs. (8)–(10). This requires finding and using the augmented cost functional $K(\mathbf{f},\mathbf{b},\mathbf{s})$ that satisfies the condition in Eq. (5) for the particular potential functional ψ used in $J(\mathbf{f})$ of Eq. (14). For the sake of brevity, we do not give the expressions for $K(\mathbf{f},\mathbf{b},\mathbf{s})$ for each of the potential functionals in Table 1, but rather mention how the iterations for the ℓ_p-norm case, given by Eqs. (11)–(13), would be affected by the use of a different functional. In fact, the only modification needed in the iterative algorithm of Eqs. (11)–(13) is the update for $\hat{\mathbf{b}}_i^{(n+1)}$ in Eq. (12). Table 2 shows the form of these updates for the three potential functionals of Table 1. Note that the framework we have presented is not limited to the three specific potential functionals we have used as examples, and other functionals might be used as well.

3 Experimental Results

We demonstrate the performance of our techniques on three imaging applications: digital holography, SAR, and ultrasound. For particular sensor models in these applications, see Refs. 14, 24, and 25. In the cost functional of Eq. (4), we find that values of p around 1 appear to yield good results for the applications we consider here. As a result, we use $p=1$ in all of our experimental results in this paper. We choose the hyperparameter λ, which appears in the cost functional $J(\mathbf{f})$ of Eq. (4), based on subjective qualitative assessment of the formed imagery. We set the approximation parameter in the nonquadratic potentials in Table 1 to be $\epsilon=10^{-5}$, which is small enough not to affect the behavior of the solution. For the termination condition of our iterative algorithm, we use $\delta=10^{-6}$ and a CG tolerance of $\delta_{\mathrm{CG}}=10^{-3}$.

Figure 2 contains the results of currently available methods for a holography experiment. The magnitude of the underlying complex-valued scene is shown in Fig. 2(a). The phase of the scene at each pixel is uniformly distributed, and uncorrelated with the phase at other pixels. We consider the case of Fraunhofer diffraction[1,24] and compute a band-limited Fourier hologram, which constitutes the measured data. The amount of data we have after band-limitation is equal to 76% of the hologram data that would be needed to form a full-resolution reconstruction of the original image. The image in Fig. 2(b) is the magnitude of the conventional reconstruction from the hologram. This result is dominated by coherent speckle artifacts. We now show how incoherent image-processing techniques can fail in this problem. In Fig. 2(c), we show the result of an incoherent edge-preserving reconstruction method. In particular we use nonquadratic regularization with ℓ_p-norm-based constraints.[3,4] Since such techniques have been designed for real-valued signals, they are not able to treat the magnitude and phase components properly. This leads to some smoothing in the real and imaginary components of the field; however, a speckle-dominated magnitude image is produced, which shows only minor improvement over the conventional image of Fig. 2(b). In Fig. 2(d), we present the result of applying a variant of anisotropic diffusion[20] to the magnitude of the conventionally reconstructed image. Some speckle suppression seems to have been achieved; however, a significant amount of detail in the scene has been lost.

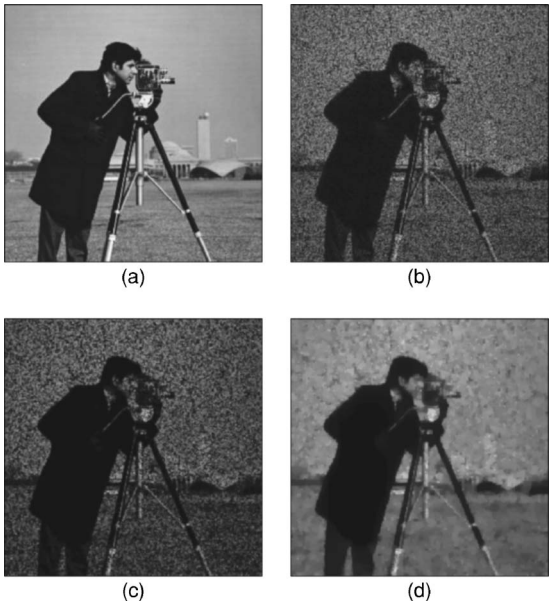

Fig. 2 Reconstruction of an image from its band-limited Fourier hologram using currently available techniques. (a) Original scene. (b) Conventional reconstruction. (c) Reconstruction by an edge-preserving regularization method designed for incoherent imaging. (d) Postprocessing of the conventionally reconstructed image by an anisotropic diffusion-based method.

In Fig. 3, we present the results of the technique we have proposed in Sec. 2, with each of the three regularizing potentials from Table 1, and $p=1$. In this experiment, we choose **D** to be a discrete approximation to the 2-D spatial derivative operator. With suppressed speckle and preserved edges, our method provides what appears to be an accurate reconstruction of the original scene in Fig. 2(a). These results demonstrate the power of our model-based coherent image reconstruction approach as compared to standard coherent image formation [Fig. 2(b)], incoherent edge-preserving regularization [Fig. 2(c)], and anisotropic diffusion-based postprocessing for image enhancement [Fig. 2(d)].

For the remaining examples, we only present images produced by conventional imaging and our nonquadratic regularization-based method. An additional analysis similar

Fig. 3 Reconstruction of an image from its band-limited Fourier hologram by the technique proposed in Sec. 2, with the following choices of regularizing functionals from Table 1, and $p=1$: (a) ψ_1, (b) ψ_2, (c) ψ_3.

to that carried out for the digital holography example of Fig. 2 and 3 yields qualitatively very similar results.

Our next example is from X-band SAR imaging, where we use a tomographic observation model.[25] Figure 4(a) contains a conventional SAR image of three vehicles in a field containing some trees. Speckle artifacts, clearly visible in this reconstruction, make, e.g., automatic segmentation of SAR images very challenging. In contrast, the images produced by our method (with $p=1$ and $\mathbf{D}$ being a derivative operator), shown in Fig. 4(b)–4(d) for different regularizing potentials ψ, produce regions (vehicle, tree, shadow, background) that appear to be more easily separable.

Our final results are from ultrasound imaging motivated by the application of nondestructive evaluation (NDE). One of the goals in nondestructive evaluation is to image the internal structure of homogeneous materials to detect material defects such as cracks. We present experimental results based on data collected at the Large Ultrasound Test Facility[35] at Boston University. The goal in this experimental setup is to image the cross section of an aluminum object (modeling the crack) immersed in a tank full of water (modeling the homogeneous material). Data are collected in a monostatic data acquisition configuration by mechanically scanning a single transducer through a set of aperture coordinates above the tank. At each data collection point, we record a broadband echo signal. For the experiments reported here, we only use frequency-domain data at a temporal frequency of 730 kHz, although our approach could also use data at multiple frequencies. For the mathematical model relating the underlying image to the observed data, we use the physical optics approximation, as in Ref. 14. This leads to a Green's function, or a complex-valued point spread function (PSF), which we use to construct the matrix $\mathbf{T}$ in Eq. (2). This theoretical observation model appears to be in good agreement with the experimental PSF we have obtained using a spherical point target in our experimental setup. Further details of this experimental setup are beyond the scope of the current paper, and will be described elsewhere. Let us now start presenting our image reconstruction results. The synthetic image in Fig. 5(a) shows the U-shaped cross section of the aluminum object, based on the true dimensions of the object, and its actual relative location within the viewing geometry. This synthetic image is just to help visualize the "underlying true field" in this experiment, and the results we present next are based on measured data and not on synthetically generated data. In Fig. 5(b), we show a conventional image, reconstructed using a regularized pseudoinverse technique.[36] Such techniques are widely used in a variety of inverse problems. This image exhibits some artifacts, making it difficult to determine the shape of the imaged object (hence the shape and structure of the crack in NDE). In this application, the goal is to image narrow cracks rather than spatially distributed objects; hence in our methods we use $\mathbf{D}=\mathbf{I}$ in Eqs. (4) and (14). Our technique ($p=1$) produces the images in Fig. 5(c)–5(e) where artifacts are reduced, and the shape of the aluminum object is preserved.

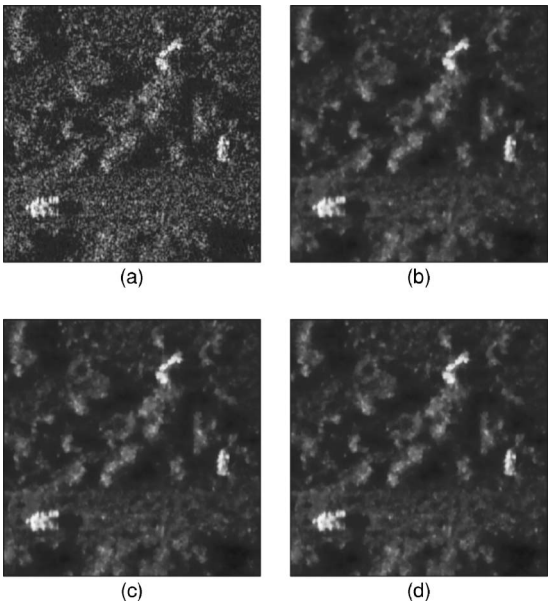

Fig. 4 (a) Conventional SAR image of a scene. (b),(c),(d) Reconstructions produced by the technique proposed in Sec. 2, with the following choices of regularizing functionals from Table 1, and $p=1$: (b) ψ_1, (c) ψ_2, (d) ψ_3.

4 Conclusions

We have presented an optimization-based method for image formation in coherent systems. Our approach is based on cost functionals that are extensions of nonquadratic regularization techniques. The cost functionals are constructed

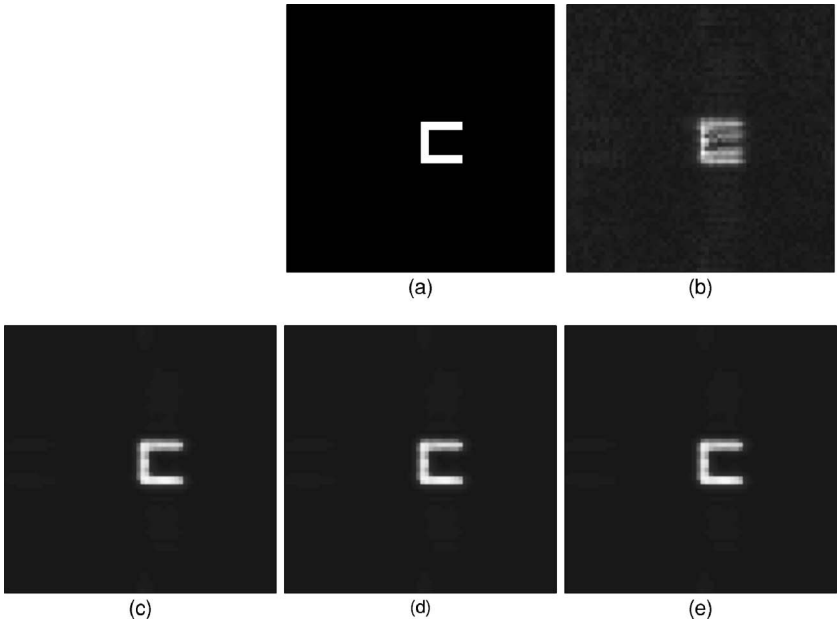

Fig. 5 Ultrasound imaging experiments based on measured 730-kHz data. (a) Synthetic image describing the underlying scene to be imaged. (b) Conventional reconstruction. (c),(d),(e) Reconstructions produced by the technique proposed in Sec. 2, with the following choices of regularizing functionals from Table 1, and $p=1$: (c) ψ_1, (d) ψ_2, (e) ψ_3.

in such a way to achieve noise and artifact suppression together with feature preservation in the resulting images, while taking into account the nature of the signals involved in coherent imaging. In order to efficiently solve the optimization problems formulated for coherent imaging, we extend and use half-quadratic regularization methods. Our experimental study has shown the effectiveness of this strategy in obtaining reconstructions that are superior in a number of ways to conventional coherent images. The improvements provided by these reconstructions appear to be promising for visual and automatic interpretation of the underlying scenes. One interesting direction for future work is the extension of the techniques presented in this paper to coherent imaging problems involving nonlinear observation models.

5 Appendix

5.1 *Discrete 2-D Derivative Operators*

In our method, we use smoothness constraints on a field, which require the spatial derivatives of the field. We use the horizontal and vertical first-order difference operators in approximating such derivatives. Derivatives of the field in other directions, such as the diagonals, may be used as well; however, we have found the use of horizontal and vertical derivatives sufficient. Consider a real-valued, sampled field $\mathbf{z}$, column-stacked as a vector of length $N=N_xN_y$, where N_x and N_y denote the numbers of rows and columns, respectively, in the 2-D field. We can compute first differences of this field, $\mathbf{D}_x\mathbf{z}$ and $\mathbf{D}_y\mathbf{z}$, in the horizontal and vertical directions, respectively, where the discrete derivative operators are given by

$$\mathbf{D}_x = \begin{bmatrix} -\mathbf{I} & \mathbf{I} & & \\ & \ddots & \ddots & \\ & & -\mathbf{I} & \mathbf{I} \end{bmatrix} \tag{15}$$

and

$$\mathbf{D}_y = \begin{bmatrix} \mathbf{D}_1 & & & \\ & \mathbf{D}_1 & & \\ & & \ddots & \\ & & & \mathbf{D}_1 \end{bmatrix}, \tag{16}$$

with

$$\mathbf{D}_1 = \begin{bmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix}. \tag{17}$$

Note that, since we take first differences between neighboring pixels, it is appropriate to have the discrete derivatives defined on the locations between the adjacent pixels. With these definitions, $\mathbf{D}_x$ then has a size of $N_y(N_x-1)\times N_xN_y$, and $\mathbf{D}_y$ has a size of $N_x(N_y-1)\times N_xN_y$. Hence, these are nonsquare operators. However, if the use of square derivative operators is desired, the preceding definitions can be augmented by derivatives defined at the boundary of the field. This may be preferred, for example, when one wants to associate each derivative to a pixel location.

We now describe two ways to compute the smoothness constraint terms, of the form $\|\mathbf{Dz}\|_p^p$, that appear in objective functionals such as that in Eq. (3). The discussion can eas-

ily be generalized to smoothness constraints with other potential functionals, such as those considered in Sec. 2.5.

The first approach is based on treating the horizontal and vertical derivatives separately when imposing a smoothness constraint. This is achieved by defining the 2-D discrete derivative operator $\mathbf{D}$ as follows:

$$\mathbf{D} \triangleq \begin{bmatrix} \mathbf{D}_x \\ \mathbf{D}_y \end{bmatrix}. \tag{18}$$

With this definition, we can write $\|\mathbf{Dz}\|_p^p$ as

$$\|\mathbf{Dz}\|_p^p = \sum_{i=1}^{M} |(\mathbf{Dz})_i|^p = \sum_{i=1}^{M_x} |(\mathbf{D}_x\mathbf{z})_i|^p + \sum_{i=1}^{M_y} |(\mathbf{D}_y\mathbf{z})_i|^p$$
$$= \|\mathbf{D}_x\mathbf{z}\|_p^p + \|\mathbf{D}_y\mathbf{z}\|_p^p, \tag{19}$$

where $M_x \triangleq N_y(N_x-1)$, $M_y \triangleq N_x(N_y-1)$, and $M=M_x+M_y$.

The second approach is based on treating the gradient at each pixel location as a two-element vector $[(\mathbf{D}_x\mathbf{z})_i\ (\mathbf{D}_y\mathbf{z})_i]^T$, composed of the horizontal and vertical gradients, and using the ℓ_2 norms of such gradients at all locations in the field for the computation of the overall ℓ_p norm:

$$\|\mathbf{Dz}\|_p^p \triangleq \sum_{i=1}^{N} [|(\mathbf{D}_x\mathbf{z})_i|^2 + |(\mathbf{D}_y\mathbf{z})_i|^2]^{p/2}. \tag{20}$$

Two things must be noted here. First, the use of a *linear* operator $\mathbf{D}$ is only conceptual in this case, because no such explicit matrix exists. Second, this approach requires a one-to-one correspondence between horizontal and vertical derivatives at each location in the scene; hence in this case we use square $(N\times N)$ derivative operators $\mathbf{D}_x, \mathbf{D}_y$.

In our method, we make use of both approaches described; however, all the mathematical expressions in the body of this paper are based on the first approach. Note that when $p=2$, the two approaches are identical, with the use of square derivative operators. To make the association between the two approaches clear, let us consider square derivative operators, and examine the first approach in this case:

$$\|\mathbf{Dz}\|_p^p = \sum_{i=1}^{N} |(\mathbf{D}_x\mathbf{z})_i|^p + \sum_{i=1}^{N} |(\mathbf{D}_y\mathbf{z})_i|^p \tag{21}$$

$$= \sum_{i=1}^{N} [|(\mathbf{D}_x\mathbf{z})_i|^p + |(\mathbf{D}_y\mathbf{z})_i|^p]. \tag{22}$$

Let us compare this expression with the second approach, given in Eq. (20). There the ℓ_2 norm of the gradient vector at each location is used in the computation of the overall ℓ_p norm. In contrast, the first approach, as shown in Eq. (22), corresponds to using an ℓ_p norm for the gradient vector $[(\mathbf{D}_x\mathbf{z})_i\ (\mathbf{D}_y\mathbf{z})_i]^T$ at each location. This association lets us compare the consequences of using the two approaches. For example, when $p<2$, the first approach used in a smoothness constraint would favor horizontal and vertical edges over diagonal edges, more than in the second approach.

5.2 *Half-Quadratic Functional for ℓ_p-Norm-Based Regularization*

The objective of this subsection is to prove the relationship (5), which we repeat below, between the particular functionals $J(\mathbf{f})$ of Eq. (4) and $K(\mathbf{f},\mathbf{b},\mathbf{s})$ of Eq. (6):

$$\inf_{\mathbf{b},\mathbf{s}} K(\mathbf{f},\mathbf{b},\mathbf{s}) = J(\mathbf{f}). \tag{23}$$

This relationship shows that $K(\mathbf{f},\mathbf{b},\mathbf{s})$ of Eq. (6) is a valid augmented cost functional to be used in half-quadratic regularization for the functional $J(\mathbf{f})$ of Eq. (4).

To keep the derivation simple, we consider a 1-D signal $\mathbf{f}$, rather than a 2-D field, in this subsection. The results however can easily be extended to the 2-D case. We assume the following structure for the discrete 1-D derivative operator $\mathbf{D}$:

$$\mathbf{D} = \begin{bmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix}, \tag{24}$$

which simply consists of two-element differences.

Let us now find $\mathbf{s}$ and $\mathbf{b}$ that minimize $K(\mathbf{f},\mathbf{b},\mathbf{s})$ of Eq. (6). First consider $\mathbf{s}$. The portion of $K(\mathbf{f},\mathbf{b},\mathbf{s})$ that depends on $\mathbf{s}$ is the following:

$$\sum_{i=1}^{M} \mathbf{b}_i |(\mathbf{DSf})_i|^2. \tag{25}$$

Based on the structures of $\mathbf{D}$ in Eq. (24) and $\mathbf{S}$ in Eq. (7), we have

$$(\mathbf{DSf})_i = -\exp(-j\mathbf{s}_i)\,\mathbf{f}_i + \exp(-j\mathbf{s}_{i+1})\,\mathbf{f}_{i+1}, \tag{26}$$

and consequently

$$|(\mathbf{DSf})_i|^2 = |\mathbf{f}_i|^2 + |\mathbf{f}_{i+1}|^2 - 2\Re[|\mathbf{f}_i|\,|\mathbf{f}_{i+1}|\exp(j\{\phi[(\mathbf{f})_i] - \phi[(\mathbf{f})_{i+1}]\})\exp[j(\mathbf{s}_{i+1}-\mathbf{s}_i)]]. \tag{27}$$

Here $\phi[(\mathbf{f})_i]$ denotes the phase of the complex number $\mathbf{f}_i$. The sum in Eq. (25) takes its minimum value when the product inside the outermost brackets in Eq. (27) has a zero imaginary part for all i. Hence the minimizing $\mathbf{s}$ satisfies

$$\mathbf{s}_{i+1} - \mathbf{s}_i + \phi[(\mathbf{f})_i] - \phi[(\mathbf{f})_{i+1}] = 0. \tag{28}$$

We could have obtained this result by the following qualitative argument as well. We want to minimize Eq. (25), which is a weighted sum of squared norms of the differences between complex-number pairs of the form $z_i = \exp(-j\mathbf{s}_i)\,\mathbf{f}_i$. The variables we have for optimization are $\mathbf{s}_i$ for all i, hence we can essentially choose the phase of each complex number. Naturally, the minimum is obtained when the complex numbers z_i have identical phase, since this makes the norm of the difference between two complex numbers as small as possible. This is exactly what the condition in Eq. (28) implies: the optimum $\mathbf{s}_i$ should "rotate" $\mathbf{f}_i$ in such a way that the resulting z_i have the same phase for

all i. Note that we still have a freedom in choosing what that identical phase is. If we simply choose it to be 0, then we have the following optimal $\mathbf{s}$:

$$\mathbf{s}_i = \phi[(\mathbf{f})_i] \quad \forall\, i. \tag{29}$$

Note that with this $\mathbf{s}$, we have $\mathbf{Sf}=|\mathbf{f}|$. Hence,

$$\inf_{s} K(\mathbf{f},\mathbf{b},\mathbf{s}) = \|\mathbf{y}-\mathbf{T}\mathbf{f}\|_2^2 + \lambda \sum_{i=1}^{M} \left[\mathbf{b}_i [|(\mathbf{D}|\mathbf{f}|)_i|^2 + \epsilon] + \left(\frac{p}{2\mathbf{b}_i}\right)^{\frac{p}{2-p}} \left(1-\frac{p}{2}\right) \right]. \tag{30}$$

Next, let us consider $\mathbf{b}$. Differentiating the summand in Eq. (30) and setting it equal to zero, we obtain the following condition for the minimizing $\mathbf{b}$:

$$\mathbf{b}_i = \frac{p}{2[|(\mathbf{D}|\mathbf{f}|)_i|^2 + \epsilon]^{1-p/2}}. \tag{31}$$

Substituting Eq. (31) in $K(\mathbf{f},\mathbf{b},\mathbf{s})$, we obtain the result we desire:

$$\inf_{\mathbf{b},\mathbf{s}} K(\mathbf{f},\mathbf{b},\mathbf{s}) = \|\mathbf{y}-\mathbf{T}\mathbf{f}\|_2^2 + \lambda \sum_{i=1}^{M} [|(\mathbf{D}|\mathbf{f}|)_i|^2 + \epsilon]^{p/2} = J(\mathbf{f}), \tag{32}$$

which shows that Eq. (5) holds for $J(\mathbf{f})$ of Eq. (4) and $K(\mathbf{f},\mathbf{b},\mathbf{s})$ of Eq. (6).

5.3 *Convergence of the Algorithm in Sec. 2.3*

Let us consider the sequence $K_n = K(\hat{\mathbf{f}}^{(n)}, \hat{\mathbf{b}}^{(n+1)}, \hat{\mathbf{s}}^{(n+1)})$, and show that it is convergent. From Eqs. (8) and (9), we have

$$K(\hat{\mathbf{f}}^{(n)}, \hat{\mathbf{b}}^{(n)}, \hat{\mathbf{s}}^{(n+1)}) \leqslant K(\hat{\mathbf{f}}^{(n)}, \hat{\mathbf{b}}^{(n)}, \hat{\mathbf{s}}^{(n)}) \quad \forall\, n, \tag{33}$$

$$K(\hat{\mathbf{f}}^{(n)}, \hat{\mathbf{b}}^{(n+1)}, \hat{\mathbf{s}}^{(n+1)}) \leqslant K(\hat{\mathbf{f}}^{(n)}, \hat{\mathbf{b}}^{(n)}, \hat{\mathbf{s}}^{(n+1)}) \quad \forall\, n, \tag{34}$$

which implies

$$K(\hat{\mathbf{f}}^{(n)}, \hat{\mathbf{b}}^{(n+1)}, \hat{\mathbf{s}}^{(n+1)}) \leqslant K(\hat{\mathbf{f}}^{(n)}, \hat{\mathbf{b}}^{(n)}, \hat{\mathbf{s}}^{(n)}). \tag{35}$$

Similarly, from Eq. (10), we have

$$K(\hat{\mathbf{f}}^{(n+1)}, \hat{\mathbf{b}}^{(n+1)}, \hat{\mathbf{s}}^{(n+1)}) \leqslant K(\hat{\mathbf{f}}^{(n)}, \hat{\mathbf{b}}^{(n+1)}, \hat{\mathbf{s}}^{(n+1)}) \quad \forall\, n. \tag{36}$$

Now, let us consider the difference:

$$K_n - K_{n-1} = \left[K(\hat{\mathbf{f}}^{(n)}, \hat{\mathbf{b}}^{(n+1)}, \hat{\mathbf{s}}^{(n+1)}) - K(\hat{\mathbf{f}}^{(n)}, \hat{\mathbf{b}}^{(n)}, \hat{\mathbf{s}}^{(n)}) \right] + \left[K(\hat{\mathbf{f}}^{(n)}, \hat{\mathbf{b}}^{(n)}, \hat{\mathbf{s}}^{(n)}) - K(\hat{\mathbf{f}}^{(n-1)}, \hat{\mathbf{b}}^{(n)}, \hat{\mathbf{s}}^{(n)}) \right]. \tag{37}$$

Using Eqs. (35) and (36), we obtain

$$K_n - K_{n-1} \leqslant 0 \quad \forall\, n, \tag{38}$$

which means that the sequence K_n is decreasing. Since it is bounded below and decreasing, the sequence K_n converges. Hence the algorithm in Sec. 2.3 is convergent in terms of the cost functional. A similar convergence result can be applied to the variants of this algorithm described in later sections.

Acknowledgments

We would like to thank in Emmanuel Bossy and Robin Cleveland, from the Acoustics Group at Boston University, for collecting and providing the ultrasound data. This work was supported in part by the Army Research Office under grants DAAD19-00-1-0466 and DAAG55-97-1-0013, the Air Force Office of Scientific Research under grants F49620-00-0362 and F49620-96-1-0028, and the National Institutes of Health under grant NINDS 1 R01 NS34189.

References

1. J. M. Blackledge, *Quantitative Coherent Imaging*, Academic Press, San Diego, CA (1989).
2. A. N. Tikhonov, "Solution of incorrectly formulated problems and the regularization method," *Sov. Math. Dokl.* **4**, 1035–1038 (1963).
3. C. R. Vogel and M. E. Oman, "Fast, robust total variation-based reconstruction of noisy, blurred images," *IEEE Trans. Image Process.* **7**(6), 813–824 (1998).
4. P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Trans. Image Process.* **6**(2), 298–310 (1997).
5. D. C. Munson, Jr., and J. L. C. Sanz, "Image reconstruction from frequency-offset Fourier data," *Proc. IEEE* **72**(6), 661–669 (1984).
6. D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization," *Proc. Natl. Acad. Sci. U.S.A.* **100**(5), 2197–2202 (2003).
7. D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. Pattern Anal. Mach. Intell.* **14**(3), 367–383 (1992).
8. J. He and S. X. Pan, "Magnitude reconstruction of complex images from incomplete Fourier phase data," in *16th Annual Conf. of the IEEE Industrial Electronics Society*, Vol. **1**, pp. 357–362 (1990).
9. B. K. Jennison and J. P. Allebach, "Maximum likelihood image reconstruction from Fourier-offset data using the expectation-maximization algorithm," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Vol. **4**, pp. 2597–2600 (1991).
10. M. Rabbani and B. E. A. Saleh, "Bayesian filtering of speckled images," *Opt. Commun.* **53**(3), 147–152 (1985).
11. J. F. Giovannelli, A. Herment, and G. Demoment, "A Bayesian approach to ultrasound Doppler spectral analysis," in *IEEE Ultrasonics Symp.*, pp. 1055–1058 (1993).
12. C. P. Mariadassou and B. Yegnanarayana, "Image reconstruction from noisy digital holograms," *IEE Proc. F, Radar Signal Process.* **137**(5), 351–356 (1990).
13. H. Carfantan and A. Mohammad-Djafari, "A Bayesian approach for nonlinear inverse scattering tomographic imaging," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Vol. **4**, pp. 2311–2314 (1995).
14. D. J. Battle, R. P. Harrison, and M. Hedley, "Maximum entropy image reconstruction from sparsely sampled coherent field data," *IEEE Trans. Image Process.* **6**(8), 1139–1147 (1997).
15. D. J. Battle, "Maximum entropy regularisation applied to ultrasonic image reconstruction," PhD Thesis, Univ. of Sydney (1999).
16. J. M. Gorce, D. Friboulet, J. D'hooge, B. Bijnens, and I. E. Magnin, "Regularized autoregressive models for a spectral estimation scheme dedicated to medical ultrasonic radio-frequency images," in *IEEE Ultrasonics Symp.*, pp. 1461–1464 (1997).
17. P. Ciuciu, J. Idier, and J. F. Giovannelli, "Regularized estimation of mixed spectra using a circular Gibbs-Markov model," *IEEE Trans. Image Process.* **49**(10), 2202–2213 (2001).
18. O. Husby, T. Lie, T. Langø, J. Hokland, and H. Rue, "Bayesian 2-D deconvolution: a model for diffuse ultrasound scattering," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **48**(1), 121–130 (2001).
19. K. Z. Abd-Elmoniem, A.-B. M. Yousef, and Y. M. Kadah, "Real-time speckle reduction and coherence enhancement in ultrasound imaging via nonlinear anisotropic diffusion," *IEEE Trans. Biomed. Eng.* **49**(9), 997–1014 (2002).
20. P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.* **12**(7), 629–639 (1990).
21. A. Abubakar and P. M. van den Berg, "A multiplicative weighted l^2-norm total variation regularization for deblurring algorithms," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Vol. **4**, pp. 3545–3548 (2002).
22. S. Sotthivirat and J. A. Fessler, "Penalized-likelihood image reconstruction for digital holography," *J. Opt. Soc. Am. A* **21**(5), 737–750

(2004).
23. A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, IEEE Press, New York (1988).
24. L. P. Yaroslavskii and N. S. Merzlyakov, *Methods of Digital Holography*, Consultants Bureau, New York (1980).
25. D. C. Munson, Jr., J. D. O'Brien, and W. K. Jenkins, "A tomographic formulation of spotlight-mode synthetic aperture radar," *Proc. IEEE* **71**, 917–925 (Aug. 1983).
26. C. V. Jakowatz, Jr., D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and P. A. Thompson, *Spotlight-mode Synthetic Aperture Radar: A Signal Processing Approach*, Kluwer Academic Publishers, Norwell, MA (1996).
27. J. A. Jensen, *Estimation of Blood Velocities Using Ultrasound: A Signal Processing Approach*, Cambridge University Press, Cambridge, UK (1996).
28. J. T. Ylitalo and H. Ermert, "Ultrasound synthetic aperture imaging: monostatic approach," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **41**(3), 333–339 (1994).
29. D. L. Donoho, I. M. Johnstone, J. C. Koch, and A. S. Stern, "Maximum entropy and the nearly black object," *J. R. Stat. Soc. Ser. B. Methodol.* **54**(1), 41–81 (1992).
30. G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins Univ. Press, Baltimore (1996).
31. A. H. Delaney and Y. Bresler, "Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography," *IEEE Trans. Image Process.* **7**(2), 204–221 (1998).
32. W. C. Karl, "Regularization in image restoration and reconstruction," in *Handbook of Image and Video Processing*, A. Bovik, Ed., pp. 141–160, Academic Press (2000).
33. S. Geman and D. E. McClure, "Statistical methods for tomographic image reconstruction," in *Proc. 46th Sess. Int. Statistical Institute, Bull. ISI* **52** (1987).
34. T. Hebert and R. Leahy, "A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors," *IEEE Trans. Med. Imaging* **8**, 194–202 (1989).
35. Boston University Acoustics Group, Medical Ultrasound Testbed (MedBED), Web page, http://www.bu.edu/paclab/censsis/medbed/
36. R. L. Lagendijk and J. Biemond, *Iterative Identification and Restoration of Images*, Kluwer, Boston, (1991).

Müjdat Çetin received his BS degree from Boğaziçi University, Istanbul, Turkey, in 1993, his MS degree from the University of Salford, Manchester, UK, in 1995, and his PhD degree from Boston University, Boston, MA, USA, in 2001, all in electrical engineering. Since 2001, he has been a member of the Laboratory for Information and Decision Systems, M.I.T., Cambridge, MA, USA, where he now holds the title of research scientist. Since September 2005 he has also been an assistant professor at Sabanci University, Istanbul, Turkey. He has served in various organizational capacities, including special session organizer, session chair, and technical program committee member, for a number of conferences, including the SPIE Conference on Algorithms for Synthetic Aperture Radar Imagery, the IEEE International Conference on Image Processing, and the IEEE International Conference on Acoustics, Speech, and Signal Processing. His research interests are in the broad areas of statistical signal and image processing, inverse problems, computer vision, and data fusion.

William C. Karl received his PhD degree in electrical engineering and computer science in 1991 from the Massachusetts Institute of Technology, Cambridge, where he also received his SM, EE, and SB degrees. He held the position of staff research scientist with the Brown-Harvard-M.I.T. Center for Intelligent Control Systems and the M.I.T. Laboratory for Information and Decision Systems from 1992 to 1994. He joined the faculty of Boston University in 1995, where he is currently professor of electrical and computer engineering. Since January 1996 he has also held a joint appointment in the Department of Biomedical Engineering at Boston University. He has been an associate editor of the *IEEE Transactions on Image Processing*. He has also served in various organizational capacities, including session organizer and chair for the 36th Asilomar Conference on Signals, Systems and Computers, special session on inverse problems in imaging; session organizer and chair for the 2000 Conference in Information Sciences and Systems, special session on medical imaging; and member of the organizing committee for the First SIAM Conference on the Life Sciences. Dr. Karl's research interests are in the areas of multidimensional signal and image processing, geometric estimation, detection, and medical signal and image processing. He is a member of the Institute of Electrical and Electronics Engineers and Sigma Xi.

Alan S. Willsky joined the M.I.T. faculty in 1973 and is currently the Edwin Sibley Webster Professor of Electrical Engineering. He is a founder, member of the board of directors, and chief scientific consultant of Alphatech, Inc. From 1998–2002 he served as a member of the U.S. Air Force Scientific Advisory Board. He has received several awards, including the 1975 American Automatic Control Council Donald P. Eckman Award, the 1979 ASCE Alfred Noble Prize, and the 1980 IEEE Browder J. Thompson Memorial Award. Dr. Willsky has held visiting positions in England and France and various leadership positions in the IEEE Control Systems Society (which made him a Distinguished Member in 1988). He has delivered numerous keynote addresses and is coauthor of the undergraduate text *Signals and Systems*. His research interests are in the development and application of advanced methods of estimation and statistical signal and image processing. Methods he has developed have been successfully applied in a variety of applications, including failure detection, surveillance systems, biomedical signal and image processing, and remote sensing.