SIAM J. APPL. MATH.
Vol. 64, No. 5, pp. 1749–1774
© 2004 Society for Industrial and Applied Mathematics

SINGULAR INTEGRALS, IMAGE SMOOTHNESS, AND THE RECOVERY OF TEXTURE IN IMAGE DEBLURRING*

ALFRED S. CARASSO†

Abstract. Total variation (TV) image deblurring is a PDE-based technique that preserves edges, but often eliminates vital small-scale information, or *texture*. This phenomenon reflects the fact that most natural images are not of bounded variation. The present paper reconsiders the image deblurring problem in Lipschitz spaces $\Lambda(\alpha, p, q)$, wherein a wide class of nonsmooth images can be accommodated. A new and fast FFT-based deblurring method is developed that can recover texture in cases where TV deblurring fails completely. Singular integrals, such as the Poisson kernel, are used to create an effective new image analysis tool that can calibrate the lack of smoothness in an image. It is found that a rich class of images $\in \Lambda(\alpha, 1, \infty) \cap \Lambda(\beta, 2, \infty)$, with $0.2 < \alpha,\ \beta < 0.7$. The Poisson kernel is then used to regularize the deblurring problem by appropriately constraining its solutions in $\Lambda(\alpha, 2, \infty)$ spaces, leading to new L^2 error bounds that substantially improve on the Tikhonov–Miller method. This so-called *Poisson Singular Integral* or *PSI* method is only one of an infinite variety of singular integral deblurring methods that can be constructed. The method is found to be well-behaved in both the L^1 and L^2 norms, producing results closely matching those obtained in the theoretically optimal, but practically unrealizable, case of true Wiener filtering. Deblurring experiments on synthetically defocused images illustrate the PSI method's very significant improvements over both the total variation and Tikhonov–Miller methods. In addition, successful reconstructions with *inexact* prior Lipschitz space information, highlight the robustness and practicality of the PSI method.

Key words. image deblurring, total variation, nonsmooth images, loss of texture, Lipschitz spaces, Besov spaces, semigroup approximations, singular integrals, Poisson kernel, Gaussian kernel, recovery of texture, Tikhonov–Miller method, true Wiener filtering, Poisson Singular Integral method, PSI method

AMS subject classifications. 94A08, 65R30, 65T60, 47D06

DOI. 10.1137/S0036139903428306

1. Introduction.

The space $BV(R^2)$ of functions of bounded variation, normed by the "total variation" seminorm $\int_{R^2} |\nabla f| dxdy$, plays an important role in much recent work in image analysis. See, e.g., [11], [13], [14], [15], [16], [20], [21], [26], [33], and [39]. In particular, highly successful applications of the total variation approach to image denoising have been well-documented. In contrast, total variation image deblurring is generally not well-behaved, and often results in unacceptable loss of fine scale information. This phenomenon is now believed traceable to an improper choice of function space [24]. The present paper reconsiders the image deblurring problem in Lipschitz spaces $\Lambda(\alpha, p, q)$, wherein a wide class of nonsmooth images can be accommodated. A new and fast FFT-based deblurring technique is developed that can demonstrably recover texture in cases where total variation deblurring fails completely. The approximation properties of certain singular integral operators are intimately linked to such Lipschitz spaces [3], [4], [36]. Here, these properties are

*Received by the editors May 20, 2003; accepted for publication (in revised form) December 16, 2003; published electronically July 2, 2004.
http://www.siam.org/journals/siap/64-5/42830.html
†Mathematical and Computational Sciences Division, National Institute of Standards and Technology, Gaithersburg, MD 20899 (alfred.carasso@nist.gov). This work was performed by an employee of the U.S. Government or under U.S. Government contract. The U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes. Copyright is owned by SIAM to the extent not limited by these rights. A United States patent application is pending on part of the work described herein.

exploited in two distinct ways. In the first part of the paper, singular kernels are used to create an effective new FFT-based image analysis tool that can calibrate the lack of smoothness in an image. This tool can be used in contexts unrelated to deblurring, e.g., as a sharpness analysis tool in performance evaluation of imaging systems or image reconstruction software [38], or as a tool for detecting and quantifying "fine-structure" content in images. In the second part of the paper, singular integrals are used as regularization tools in the deblurring problem. Specifically, we show how to stabilize ill-posedness by using the Poisson kernel to impose a priori constraints, in appropriate $\Lambda(\alpha, p, \infty)$ spaces, on the desired nonsmooth deblurred image. This so-called *Poisson Singular Integral* or *PSI* method, is only one of an infinite variety of singular integral deblurring methods that can be constructed. In particular, Gaussian kernels may also be used, leading to the *Gaussian Singular Integral* or *GSI* method. Restricting attention to the case of defocus blurs, we derive L^2 error bounds for the PSI method for images in $\Lambda(\alpha, 2, \infty)$, and demonstrate robust recovery of fine structure in synthetically blurred images. Similar results hold for other types of blurs.

Extensive numerical experiments with known exact solutions indicate that the PSI method is remarkably well-behaved. In both the L^1 and L^2 norms, relative errors in the PSI method are found to closely approximate those obtained in the theoretically optimal, but practically unrealizable, case of *true* Wiener filtering. The latter method requires prior knowledge of the exact power spectra of both the noise and unknown desired sharp image, i.e., a total of $8N^2$ prior data values for a $2N \times 2N$ image. Since the PSI method requires only 4 prior data values, its ability to closely track Wiener filtering is especially noteworthy. The availability of reliable fast deblurring methods is of major significance in the case of nonsmooth images. The true value of the Lipschitz exponent α in the desired sharp image is usually not known in advance, although a plausible range of values for α can often be deduced. In section 10, we document the practicality and robustness of the PSI method, by showing that good quality reconstructions can often be obtained with *inexact*, but plausible, Lipschitz space information. Such initial restorations can then be refined interactively. Here, fast algorithms enable *simultaneous* computation and display of large numbers of trial deblurred images, resulting from multiple choices for α and/or some of the regularization parameters. We stress that the PSI method is exclusively intended for deblurring and is not intended for denoising.

2. Lack of smoothness of images. In [27], a new analytical framework for image processing is introduced, whereby a given image $f(x, y)$ is conceptualized as being the sum of three components, $f(x, y) = u(x, y) + v(x, y) + w(x, y)$. Loosely speaking, $u(x, y)$ contains the edges and the other high-priority information that is sufficient for object recognition, $v(x, y)$ contains the fine-scale details and other low-priority information that is often not necessary for recognition, and $w(x, y)$ represents noise. The $v(x, y)$ component is called *texture*. One example of $v(x, y)$ might be the hair in a photograph of a person's face. Another example of $v(x, y)$ might be the heat-shield tiles in an image of the Columbia space shuttle. The ability to resolve individual hairs is generally not necessary for identification. In several image processing tasks, such as compression, segmentation, or face recognition, this texture component can often be neglected. However, there are numerous other situations, such as in the ill-fated Columbia episode, where $v(x, y)$ may be of paramount interest. It is shown in [27] that only the $u(x, y)$ component can generally be expected to lie in $BV(R^2)$. In [24], it is proved that most natural images are *not* of bounded variation, because the texture component $v(x, y)$ generally has infinite total variation.

Denoising and deblurring are two basic image processing tasks where total variation restoration has been extensively applied. Such restoration can be accomplished most effectively by solving an initial value problem for an appropriate nonlinear anisotropic diffusion equation, using the stepwise marching scheme described in [26]. In deblurring, one typically starts with a degraded image $g(x, y)$ which differs from the desired true image $f(x, y)$ in that the $u(x, y)$ component is blurred but recognizable, the $v(x, y)$ component is seriously attenuated and often not recognizable, and the $w(x, y)$ component is usually small. Reconstructing $v(x, y)$ while keeping $w(x, y)$ small, is the prime objective in numerous medical, astronomical, industrial, and scientific contexts [9], [10]. However, while total variation deblurring sharpens $u(x, y)$ and keeps $w(x, y)$ small, the texture component $v(x, y)$ is often eliminated due to the "staircase effect" [13], [20], [29], [30], [39]. This is in accordance with the analyses in [24], [27].

Let $x = (x_1, x_2) \in R^2$. Postulating $f(x) \in BV(R^2)$ means that $f(x)$ is constrained to satisfy,

$$\int_{\mathbf{R^2}} |f(x+h) - f(x)| dx \leq Const\ |h|. \tag{1}$$

However, from the standpoint of modeling texture, it is advantageous to consider functions $f(x)$ satisfying weaker constraints, such as

$$\left\{ \int_{\mathbf{R^2}} |f(x+h) - f(x)|^p dx \right\}^{1/p} \leq Const\ |h|^\alpha, \qquad 0 < \alpha < 1. \tag{2}$$

Such an f lies in $\Lambda(\alpha, p, \infty)$. With $0 < \alpha < 1,\ 1 \leq p < \infty$, the Lipschitz (Besov) spaces $\Lambda(\alpha, p, q)$ [36], [37], consist of the class of functions $f(x) \in L^p(R^2)$ with finite seminorm $\| f \|_{\alpha pq}$, where

$$\| f \|_{\alpha pq} = \left\{ \int_{R^2} \left(|h|^{-\alpha} \| f(x+h) - f(x) \|_p \right)^q dh/|h|^2 \right\}^{1/q}, \qquad 1 \leq q < \infty, \tag{3}$$

$$\| f \|_{\alpha p \infty} = \sup_{h \in R^2} \left\{ |h|^{-\alpha} \| f(x+h) - f(x) \|_p \right\}, \qquad q = \infty. \tag{4}$$

For given p and q, functions with larger values of α are better behaved, or "smoother", than functions with smaller values of α, and functions in $\Lambda(\alpha, p, q_1)$ are smoother than functions in $\Lambda(\alpha, p, q_2)$ if $q_1 < q_2$. In fact, the following continuous embedding results are proved in [37, Theorem 9]

$$\Lambda(\alpha_2, p, q_1) \subset \Lambda(\alpha_1, p, q_2), \qquad 0 < \alpha_1 \leq \alpha_2 < 1;\ 1 \leq q_1 \leq q_2 \leq \infty. \tag{5}$$

Also, in R^2,

$$\Lambda(\alpha, p, q) \subset \Lambda(\beta, r, q), \qquad \alpha - 2/p = \beta - 2/r, \quad p \leq r. \tag{6}$$

Let $r = 2$, let the pair (α, p) satisfy $2/(1+\alpha) < p \leq 2$, and let $\beta = 1 + \alpha - 2/p$. Then, $0 < \beta \leq \alpha$, and it follows from (5) and (6) that

$$\Lambda(\alpha, p, q) \subset \Lambda(\beta, 2, q) \subset \Lambda(\beta, 2, \infty) \subset L^2(R^2). \tag{7}$$

This result will be important in what follows.

For given fixed p with $1 < p < \infty$ and $q = 2p/(2+\alpha p)$, a class of Lipschitz spaces $\Lambda(\alpha, q, q) \subset L^p(R^2)$ is considered in [18], [19], and shown to contain common types of images. A method for empirically estimating image smoothness is developed in [18], [19], based on analyzing the behavior of *lossy wavelet compression* of the image $f(x, y)$. In [12], the spaces $\Lambda(\alpha, q, q) \subset L^2(R^2),\ q = 2/(1+\alpha)$, are advocated as being particularly appropriate for accommodating a rich variety of real images in an L^2 setting. Lossy wavelet compression is again used to estimate image smoothness, and values of α in the range $0.4 < \alpha < 0.75$ are reported in [12] for a class of 24 test images $\in \Lambda(\alpha, \frac{2}{1+\alpha}, \frac{2}{1+\alpha})$. Such α values are an indication of true image smoothness only when the image is largely *noise free.* If the noise component $w(x, y)$ is not sufficiently small, artificially low values of α must be expected. A basic limitation of the above wavelet compression approach is the restriction on q, which precludes consideration of the larger spaces $\Lambda(\alpha, p, \infty) \supset \Lambda(\alpha, p, q)$.

The present independent method of estimating image smoothness rests on an entirely different analytical basis, and requires neither wavelet expansions nor image compression. Instead, the method uses fast FFT algorithms to convolve the image with a specific type of kernel, and then analyzes how well this convolved image approximates the original image as the kernel approaches the Dirac δ-function. This simple direct approach permits consideration of the spaces $\Lambda(\alpha, p, \infty),\ 1 \le p < \infty$. The results obtained here are compatible with those obtained in [18], [19], [12], and [24]. We indeed find that most natural images are not of bounded variation, and that a rich variety of images $\in \Lambda(\alpha, 1, \infty)$ with $0.2 < \alpha < 0.7$.

Remark 1. We deal with high resolution images $f(x, y)$ of size 512×512 or 1024×1024 pixels. Such an $f(x, y)$ may be viewed as a piecewise constant or trigonometric polynomial approximation to the original intensity field $f^\infty(x, y)$, or as some other kind of finite dimensional representation of the infinite dimensional object f^∞. All norms are equivalent on a finite dimensional space. Hence, even if $f^\infty(x, y)$ is not of bounded variation, the discrete total variation norm for $f(x, y)$ is always finite, though it may be very large. To estimate smoothness properties of $f^\infty(x, y)$ by examination of the finite dimensional representation $f(x, y)$ requires some sagacity. In [18, section 4B, section 5B], the authors stress that in their method of estimating the value of α by monitoring the rate of convergence as a function of the number $\mathcal{N}$ of nonzero wavelet coefficients, one must restrict attention to *low values* of $\mathcal{N}$. At high values of $\mathcal{N}$, the fact that $f(x, y)$ is actually piecewise constant causes the error to decrease much too rapidly, resulting in an *artificially high* reading for α that diverges from true behavior in $f^\infty(x, y)$. This same finite dimensionality pitfall occurs in the present approach, but wears a different guise. See Remark 2 and the discussion surrounding Figures 1 and 2 below.

We shall use the spaces $\Lambda(\alpha, 1, \infty)$ and $\Lambda(\alpha, 2, \infty)$ for examining and classifying image smoothness. However, deblurring applications will be limited to the spaces $\Lambda(\beta, 2, \infty) \subset L^2(R^2)$, wherein all spaces $\Lambda(\alpha, p, q),\ 2/(1+\alpha) < p \le 2,\ 1 \le q \le \infty$, are continuously embedded. The spaces $\Lambda(\alpha, 2, \infty)$ will be shown to contain a rich and significant class of images.

3. The spaces $\Lambda(\alpha, p, q)$ and the Poisson singular integral. Define the Fourier transform $\hat{h}(\xi, \eta)$ of $h(x, y) \in L^1(R^2)$ by

$$\mathcal{F}\{h\} = \hat{h}(\xi, \eta) \equiv \int_{R^2} h(x, y) e^{-2\pi i(\xi x + \eta y)} dx dy. \tag{8}$$

For each fixed $t > 0$, consider the Poisson kernel in R^2

$$\psi(x, y, t) = \frac{t}{2\pi(x^2 + y^2 + t^2)^{3/2}}, \qquad (x, y) \in R^2. \tag{9}$$

We have

$$\hat{\psi}(\xi, \eta, t) = e^{-t\rho}, \qquad \rho = \sqrt{\xi^2 + \eta^2}. \tag{10}$$

For each $t > 0$, define the linear operator U^t on $L^p(R^2)$, $1 \le p < \infty$, by

$$U^t f = \int_{R^2} \psi(u, v, t) f(x - u, y - v) du dv. \tag{11}$$

It can be shown that $\lim_{t \downarrow 0} \| U^t f - f \|_p = 0$. Moreover, defining U^0 to be the identity operator, we have that for $s, t \ge 0$, $U^t U^s = U^{t+s}$. In fact, $\{U^t\}_{t \ge 0}$ is a holomorphic contraction semigroup on $L^p(R^2)$. See [3]. We may write $U^t = e^{-tA}$, where $-A$ is the infinitesimal generator of U^t. Here, A corresponds to the fractional differential operator $(-\Delta)^{1/2}$. Note that for $t > 0$, U^t maps $L^p(R^2)$ into $D(A)$, so that $AU^t f$ is well-defined for arbitrary $f \in L^p$. In general, this is not the case for nonholomorphic semigroups. The Gauss singular integral, where the two-dimensional Gaussian kernel is used in lieu of ψ in (9), defines an analogous holomorphic semigroup W^t, with $A = -\Delta$. Many such singular integral semigroups S^t exist. A very rich variety can be constructed by *subordination* [7], [40]. For small $t > 0$, S^t behaves as an approximate identity on L^p. There is a large literature on how well $S^t f$ approximates f as $t \downarrow 0$. See [3], [4], [5], [35], [36], [37], and the references therein. As $t \downarrow 0$, we have $\| S^t f - f \|_p = o(1)$ for arbitrary $f \in L^p$, $\| S^t f - f \|_p = O(t)$ if and only if $f \in D(A)$, and $\| S^t f - f \|_p = o(t)$ if and only if $S^t f = f$ for all $t \ge 0$. Thus, the optimal rate is always $O(t)$. Of particular interest in this paper is the case of *nonoptimal* approximation, where $f \notin D(A)$ yet retains sufficient smoothness that $\| S^t f - f \|_p = O(t^\alpha)$, $0 < \alpha < 1$, as $t \downarrow 0$. While complete theories exist for a wide class of singular kernels, the simplest such theory revolves around the Poisson semigroup U^t in (11). We have from [37, Theorem 4], the following result.

THEOREM 1. *Let U^t, $t > 0$, be the Poisson integral operator in* (11), *and let $0 < \alpha < 1$, $1 \le p, q < \infty$. Then, $f \in \Lambda(\alpha, p, q)$ if and only if*

$$\int_0^\infty \left(t^{-\alpha} \| U^t f - f \|_p\right)^q dt/t \ < \ \infty. \tag{12}$$

For $q = \infty$, we have $f \in \Lambda(\alpha, p, \infty)$ if and only if

$$\sup_{t>0} \ t^{-\alpha} \| U^t f - f \|_p \ < \ \infty. \tag{13}$$

Using the embedding results in (7) together with (13) leads to the following corollary.

THEOREM 2 (corollary). *Let $f \in \Lambda(\alpha, p, q)$, with $2/(1 + \alpha) < p \le 2$, and let $\beta = 1 + \alpha - 2/p$. Then, in the L^2 norm*

$$\sup_{t>0} \ t^{-\beta} \| U^t f - f \|_2 \ < \ \infty. \tag{14}$$

4. Periodized problems, the Poisson summation formula, and FFT algorithms. The above results can be used to fashion a practical image analysis tool. Theoretically, given any image $f(x,y)$ in $L^1(R^2)$, one could use the Fourier transform (8) to form

$$\mathcal{F}\left\{U^t f\right\} = e^{-t\rho}\hat{f}(\xi,\eta), \qquad \rho = \sqrt{\xi^2+\eta^2}, \tag{15}$$

for sequences of positive t-values tending to zero. Inverse transformation is always possible on account of the factor $e^{-t\rho}$, and this can be used to produce an infinite sequence of positive numbers $\mu_n = \{\| U^{t_n} f - f \|_1 \,/\, \| f \|_1\}$ with $t_n \downarrow 0$. If every such sequence (t_n, μ_n), ultimately lies below the curve $\mu(t) = C\ t^\alpha,\ 0 < t \le \bar{t}$, for suitably chosen constants $C > 0$ and $0 < \alpha < 1$, then $\| U^t f - f \|_1 \le C \| f \|_1\, t^\alpha$, as $t \downarrow 0$, and $f(x,y) \in \Lambda(\alpha, 1, \infty)$ by Theorem 1. However, this does not lead to a practical procedure.

On the other hand, Theorems 1 and 2 remain valid in the periodic case [36], [37]. Here, the image $f(x,y)$ and the kernel $\psi(x,y,t)$ in (9) are now periodized [5], [6]. Let Ω denote the unit square $-1/2 < x, y \le 1/2$ in R^2. The image $f(x,y)$ is now viewed as originally defined on Ω from which it is extended by periodicity to all of R^2. Let

$$\hat{f}(\xi,\eta) = \int_\Omega f(x,y) e^{-2\pi i(\xi x + \eta y)} dx dy. \tag{16}$$

Define the periodized Poisson kernel $\psi^*(x,y,t)$ by

$$\psi^*(x,y,t) = \sum_{k,m=-\infty}^{\infty} \psi(x+k, y+m, t), \qquad t > 0, \quad (x,y) \in R^2, \tag{17}$$

and let

$$U^t f = \int_\Omega \psi^*(u,v,t) f(x-u, y-v) du dv, \qquad t > 0. \tag{18}$$

The Poisson summation formula, [1], [5], [6], [22], [37], can be used to show that the periodized Poisson kernel has a complex Fourier series with Fourier coefficients again given by (10), but where ξ, η are now *integers* running from $-\infty$ to $+\infty$. Moreover,

$$U^t f = \sum_{\xi,\eta=-\infty}^{\infty} e^{-t\rho} \hat{f}(\xi,\eta) e^{2\pi i(x\xi + y\eta)}, \qquad t > 0, \quad \rho = \sqrt{\xi^2+\eta^2}. \tag{19}$$

Again the factor $e^{-t\rho}$ assures uniform convergence of the Fourier series in (19). Let

$$f_N(x,y) = \sum_{\xi,\eta=-N}^{N} e^{-t\rho} \hat{f}(\xi,\eta) e^{2\pi i(x\xi + y\eta)}, \qquad t > 0, \quad \rho = \sqrt{\xi^2+\eta^2}. \tag{20}$$

Since $L^p(\Omega) \subset L^1(\Omega)$, $p > 1$, we may apply this approach to any $f \in L^p$, and $\| U^t f - f_N \|_p$ can be made arbitrarily small by choosing N large enough in (20). Next, given the $2J \times 2J$ digitized image $f(x,y)$ with $J > N$, the discrete Fourier transform [2] is now the appropriate numerical tool for analyzing this periodized problem. One can use FFT algorithms to form the Fourier coefficients $\hat{f}(\xi,\eta)$, $-J \le \xi, \eta \le J$, and then apply the filter $(e^{-t\rho} - 1)$ as in (15). An inverse FFT then yields an accurate approximation to $U^t f - f$ at each of the $2J \times 2J$ pixels, for each small $t > 0$. We

may then examine the discrete L^p relative error in Poisson approximation as $t \downarrow 0$, and locate constants C and α such that $\| U^t f - f \|_p \leq C \| f \|_p t^\alpha,\ 0 < t \leq \bar{t}$. In summary, we have constructed an accurate numerical procedure, based on correct mathematical analysis, for assessing membership in any $\Lambda(\alpha, p, \infty)$ space. Equally important, the values of C and α constitute a priori information that will be useful in stabilizing the ill-posed deblurring problem.

Remark 2. Analogously to the case of lossy wavelet compression discussed in Remark 1, there is a finite dimensionality pitfall in the above singular integral methodology that *necessitates the exclusion of very small values of* $t > 0$. Let $f^\infty(x, y)$ be the original image intensity field as in Remark 1, and assume that $f^\infty(x, y) \in \Lambda(0.5, p, \infty)$, so that $\| U^t f^\infty - f^\infty \|_p = O(\sqrt{t})$ as $t \downarrow 0$, by Theorem 1. Let $f(x, y)$ be the $2J \times 2J$ digitized image corresponding to $f^\infty(x, y)$. We shall show that at very small values of $t > 0$, the behavior of $\| U^t f - f \|_p$ *diverges* from true behavior in $f^\infty(x, y)$, resulting in a *false reading* for α. Let $S^t = e^{-tA}$ be any contraction semigroup on $L^p(R^2)$. As already pointed out, if $f \in D(A)$, $\| S^t f - f \|_p = O(t)$ as $t \downarrow 0$. This follows from

$$S^t f - f = \int_0^t \frac{d}{du}(S^u f) du = -\int_0^t S^u A f du, \tag{21}$$

so that $\| S^t f - f \|_p \leq t \| Af \|_p$, for all $t > 0$. In addition, $\| S^t f - f \|_p \approx t \| Af \|_p$, for all sufficiently small $t > 0$, because $S^u A f \approx Af$ for all sufficiently small u. In the above Poisson semigroup U^t, the unbounded operator A is defined as follows in Fourier space

$$\mathcal{F}\{Af\} = \rho \hat{f}(\xi, \eta), \qquad \rho = \sqrt{\xi^2 + \eta^2}. \tag{22}$$

Since the digitized $2J \times 2J$ image $f(x, y)$ is a trigonometric polynomial, it is always $\in D(A)$ and $\| Af \|_p$ is always finite, although it may be very large. Consequently, with a possibly large positive constant K, we always have $\| U^t f - f \|_p \leq Kt$ for all $t > 0$, as well as actual linear behavior $\| U^t f - f \|_p \approx Kt$ for all sufficiently small t, irrespective of the behavior of $\| U^t f^\infty - f^\infty \|_p$ at these same values of t. This phenomenon is well-illustrated in Figures 1 and 2 below.

5. Application to real images. The following examples illustrate the use of the Poisson singular integral approach. Our first example, in Figure 1, is the 512×512 Mandrill image highlighted in [24] as an example of an image $\notin BV(R^2)$. The above FFT procedure was used to obtain the L^1 and L^2 relative errors in Poisson approximation

$$\mu(t) = \| U^t f - f \|_p \,/\, \| f \|_p, \qquad p = 1, 2, \tag{23}$$

at 300 values of t given by $t_n = 0.5(0.95)^n,\ n = 1, 300$. For the L^1 norm, a plot of $\mu(t)$ versus t on a log-log scale produced the solid curve A in Figure 1. Least squares fitting was used to find the two distinct majorizing dashed straight lines Γ and Σ. For each dashed line, the y-axis intercept value obtained by least squares was slightly increased so as to make each line lie visibly above the solid curve A; however, the slope of each line remains the same as that obtained from least squares. The line Γ, defined by $\log \mu(t) = 3.2 + 0.994 \log t$, accurately captures the misleading linear trend in (23) for very small values of t, while being grossly inaccurate at larger values of t. It was obtained by excluding data corresponding to $\log t > -7$ from the least squares fit. The line Γ implies that $\| U^t f - f \|_1 < 24.53 \| f \|_1 t^{0.994}$

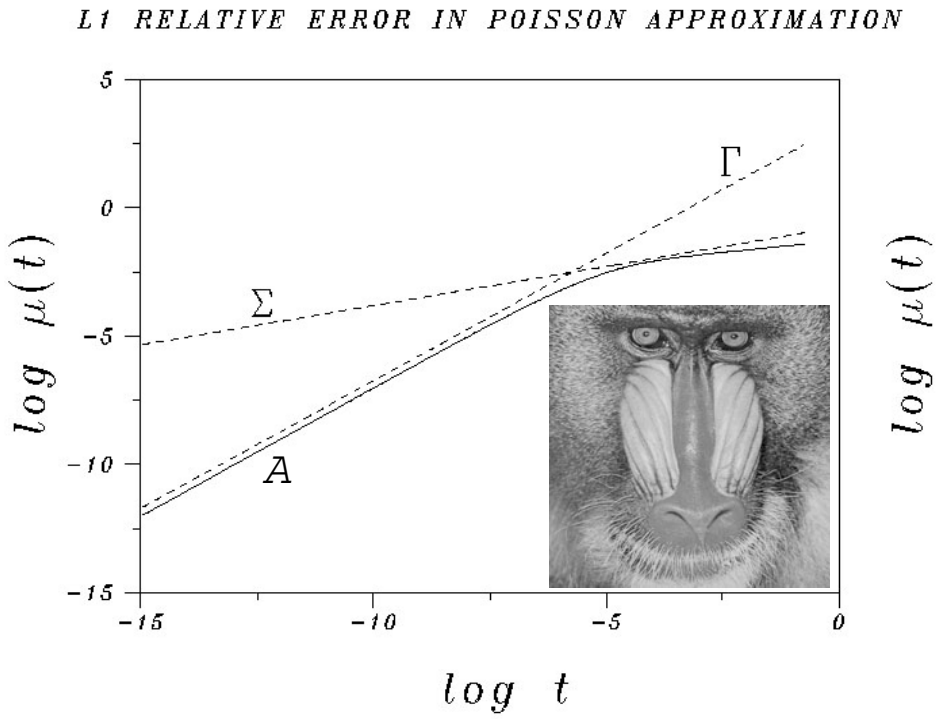

FIG. 1. 512×512 *Mandrill image was identified in* [24] *as not in* $BV(R^2)$. *This is confirmed in above graphical use of Theorem* 1, *using FFT techniques discussed in section* 4. *Solid curve* A *is a plot of* $\mu(t) = \| U^t f - f \|_1 \,/\, \| f \|_1$ *versus* t, *on a log-log scale. Majorizing dashed straight line* Γ, *defined by* $\log \mu(t) = 3.2 + 0.994 \log t$, *accurately captures linear behavior in* (23) *for very small values of* t, *but is grossly inaccurate at larger values of* t. *Linear behavior at very small* t *is misleading, and falsely implies that image is of bounded variation. (See Remarks* 1 *and* 2*). Majorizing dashed straight line* Σ, *defined by* $\log \mu(t) = -0.75 + 0.306 \log t$, *accurately reflects behavior for* $-6 \le \log t \le -1$, *while being grossly inaccurate at very small* t. *Behavior along* Σ *is taken to be true behavior in Mandrill image, implying image* $\in \Lambda(0.306, 1, \infty)$ *with* $\| U^t f - f \|_1 \le 0.472 \, \| f \|_1 \, t^{0.306}$, $0 < t \le 0.1$.

for all $t > 0$. As emphasized in Remark 2, this correct statement primarily reflects the fact that the 512×512 Mandrill image lies in a finite dimensional space; it *does not* describe the smoothness properties of the intensity field $f^\infty(x, y)$ that gave rise to the digitized Mandrill image. The majorizing dashed straight line Σ, defined by $\log \mu(t) = -0.75 + 0.306 \log t$, accurately reflects behavior of (23) for $-6 \le \log t \le -1$, while being grossly inaccurate at very small values of t. The line Σ was obtained by excluding all data corresponding to $\log t < -6$ from the least squares fit. Note that this still leaves over 100 data points remaining. The behavior along Σ indicates that $\| U^t f - f \|_1 \le 0.472 \, \| f \|_1 \, t^{0.306}$, $0 < t \le 0.1$, and this is taken to be the true behavior in the Mandrill image. From (13), this implies that the Mandrill image $\in \Lambda(0.306, 1, \infty)$, and hence, is not of bounded variation. The behavior in the L^2 norm is strikingly similar, and indicates the image $\in \Lambda(0.271, 2, \infty)$. Estimates of α in any other discrete L^p norm can be obtained similarly. All α estimates shown in this paper were obtained using the above procedure of constructing the line Σ in log-log plots of $\mu(t)$, after excluding all data corresponding to $\log t < -6$. As in [18,

section 5B], we have occasionally found contradictory examples where the value of α in the L^2 norm was greater than that in the L^1 norm. When that happened, a new Σ line was constructed for the L^2 trace, based on excluding data corresponding to $\log t < -5$. It is recommended that data for very small values of t always be included in log-log plots of $\mu(t)$, so as to enable clear identification of the spurious linear trend, prior to rejecting that part of the data.

Our second example, in Figure 2(A), is a 1024×1024 Whirlpool galaxy image, taken at the National Optical Astronomy Observatory, (NOAO/AURA/NSF), by T. Rector and M. Ramirez. As in the case of Figure 1, Poisson integral approximation in L^1 was used to obtain the solid curve A, and the line Σ_A was constructed using least squares. This procedure was repeated for the L^2 norm. The results indicate that Figure 2(A) satisfies $\| U^t f - f \|_1 \leq 0.6 \| f \|_1 t^{0.530}$, $0 < t \leq 0.1$, and that Figure 2(A) $\in \Lambda(0.530, 1, \infty) \cap \Lambda(0.462, 2, \infty)$. Interestingly, if we sharpen Figure 2(A) using the *APEX method* [9], we obtain the image in Figure 2(B). This enhanced image displays significant fine scale detail not readily visible in the original image, and strongly resembles a Whirlpool galaxy plate taken by Milton Humason in 1950 using the 200 inch Mt. Palomar telescope. See [34, plate 26]. Here, L^1 Poisson analysis produced the solid curve B and the majorizing line Σ_B. We find that Figure 2(B) $\in \Lambda(0.239, 1, \infty) \cap \Lambda(0.230, 2, \infty)$, and thus has substantially *lower* values of α than does Figure 2(A). This result is highly plausible. Presumably, any low-pass blurring process that may have affected Figure 2(A) would have attenuated fine scale features, and thereby *increased* the values of α. The result also indicates that APEX processing of image (A) produced relatively more sharpening in the L^1 norm than in the L^2 norm.

The nine images in Figure 3 and Table 1 form an interesting collection that includes natural as well as man made objects, exhibiting a wide range of sizes. The last row contains a nanoscale electron microscopy micrograph, a galactic scale object, and a cosmological scale structure. Along with the three images in Figures 1 and 2, this paper has applied the Poisson integral method to 12 high resolution images, and we have found that, in either $\Lambda(\alpha, 1, \infty)$ or $\Lambda(\alpha, 2, \infty)$, the values of α lie in the range $0.2 < \alpha < 0.7$. This range of values is compatible with that found in [12], [18], [19], using an entirely different method. Moreover, while $\Lambda(\alpha, 2, \infty)$ are smaller spaces than are $\Lambda(\alpha, 1, \infty)$, they are evidently wide enough to contain each of these 12 images, albeit with smaller values of α. Notice also that the values of the constant C in Table 1 are confined to a very narrow range in both L^1 and L^2.

Remark 3. Following [18], [19], the values of C and α reported in Table 1 and elsewhere in this paper, are given to two and three decimal places. Such precision must be viewed with skepticism. These values are based on the use of the Poisson operator U^t, together with the particular sequence $t_n = 0.5(0.95)^n$, $n = 1, 300$. Also, the specific interval $-6 \leq \log t \leq 0$ was chosen for the least squares fit to $\log \mu(t)$. However, different sequences t_n tending to zero might be used, as well as slightly larger or slightly smaller t-intervals for the least squares fit. In addition, the Gaussian operator G^t may be used in place of U^t, in which case the image Lipschitz exponent $\alpha = 2\delta$, where δ is the slope of the corresponding Σ line. It is found that such variations in the basic methodology result in slightly different values for C and α. For this reason, all reported values of C and α should probably be *rounded to one decimal place*. It may not be feasible to determine true image Lipschitz space parameters to higher place accuracy.

The PSI deblurring method to be described in section 7 below requires prior knowledge of the values of C and α in the desired unknown deblurred image. In

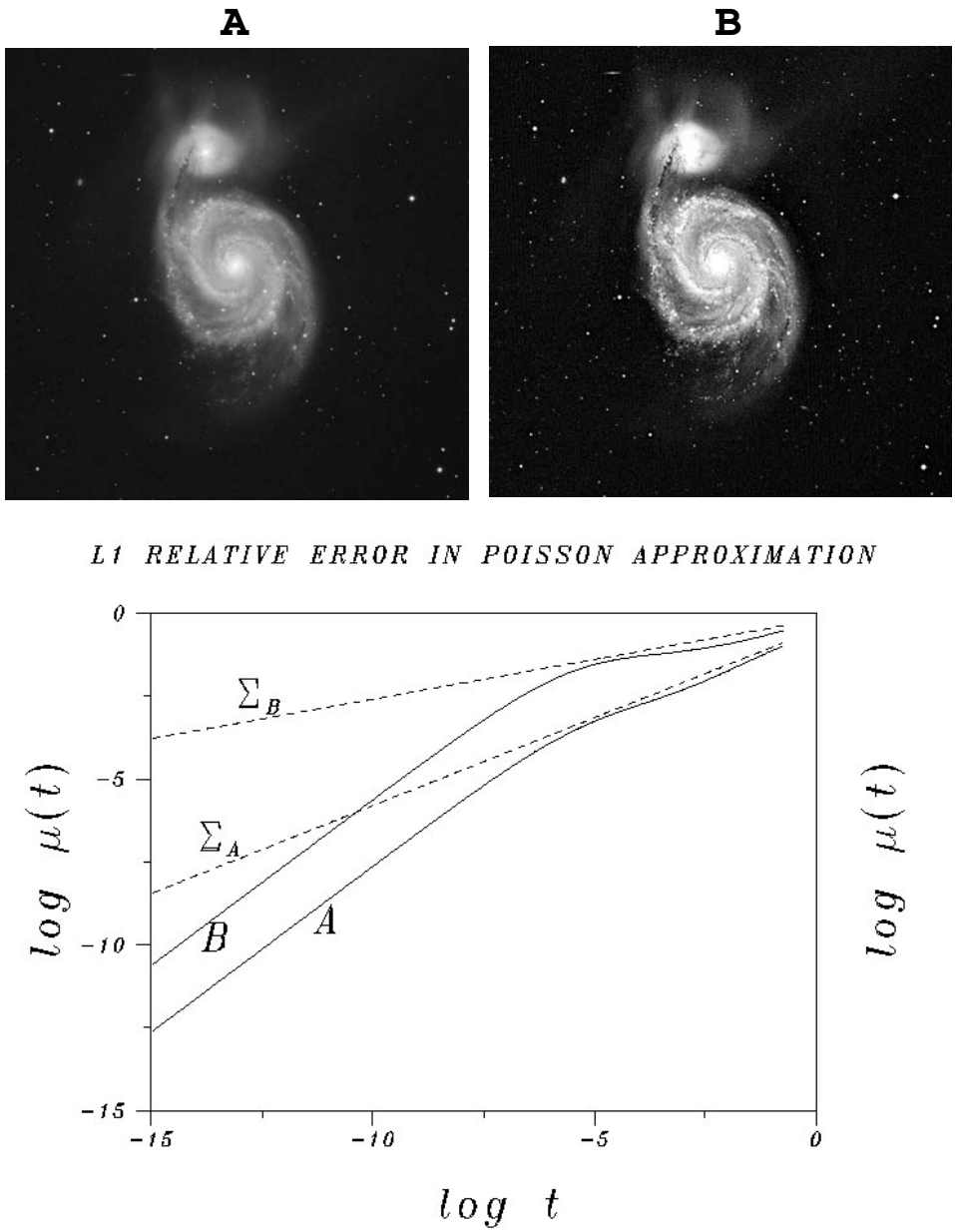

FIG. 2. *Whirlpool galaxy M*51. *Original and enhanced images have noticeably different* L^1 *Poisson traces* $\mu(t) = \| U^t f - f \|_1 / \| f \|_1$, *reflecting sharply distinct Lipschitz exponents.* (A) *Original* 1024×1024 *image taken by T. Rector and M. Ramirez, National Optical Astronomy Observatory, (NOAO/AURA/NSF).* L^1 *Poisson relative error* $\mu(t)$, *shown in solid trace A, is majorized by dashed straight line* Σ_A *defined by* $\log \mu(t) = -0.5 + 0.530 \log t$. *This implies that image* (A) $\in \Lambda(0.530, 1, \infty)$. (B) *Blind deconvolution of* (A) *using APEX method* [9], *brings out significant fine scale detail, and results in solid trace* B, *majorized by dashed straight line* Σ_B *defined by* $\log \mu(t) = -0.2 + 0.239 \log t$. *This indicates that deblurred image* (B) $\in \Lambda(0.239, 1, \infty)$. *Image* (B) *strongly resembles* [34, plate 26] *taken by Milton Humason using* 200 *inch Mt. Palomar telescope. For* $\log\ t < -7$, *solid traces* A *and* B *have identical slopes of* 0.994. *This confirms the observation in Remark* 2 *that behavior at very small* t *is artificial and disconnected from true image smoothness.*

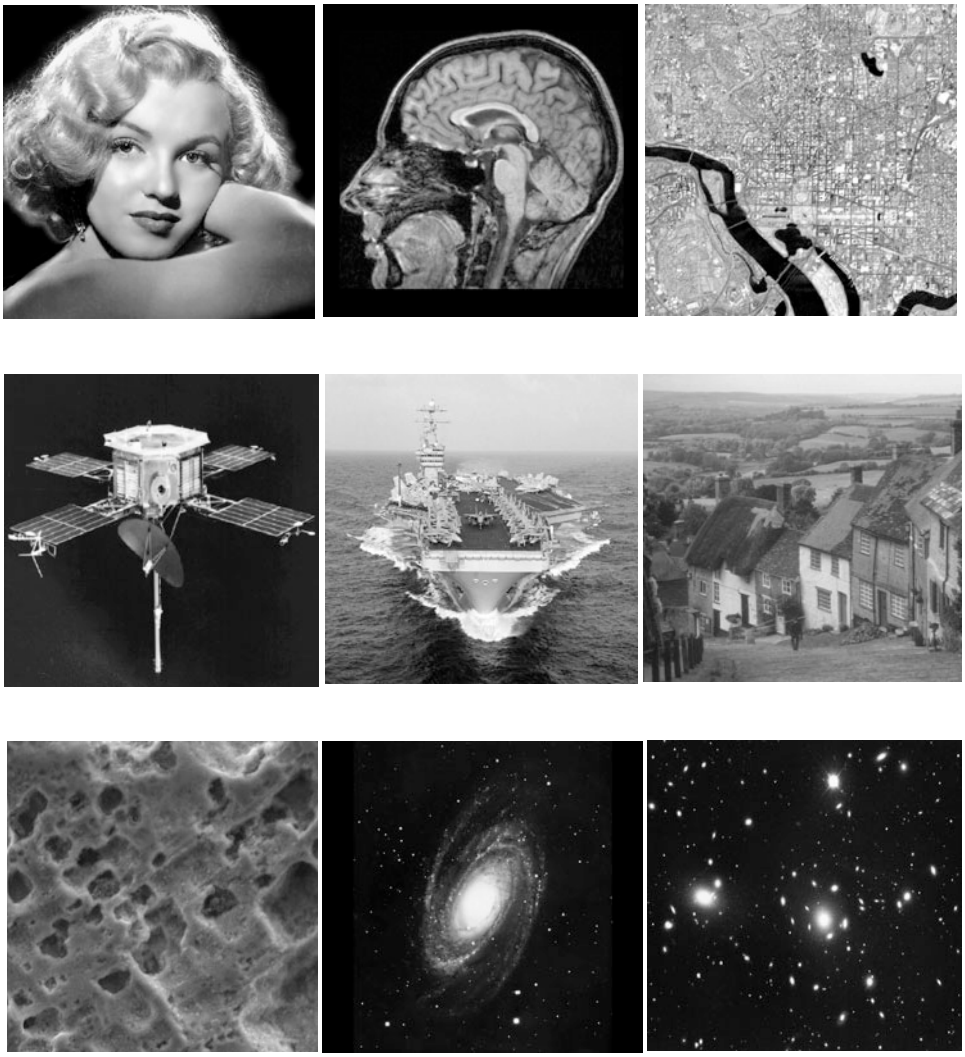

FIG. 3. *A significant class of high-resolution* 8*-bit images have Lipschitz exponents* α *in the range* $0.2 < \alpha < 0.7$, *in either* L^1 *or* L^2, *and are not of bounded variation.*

general, these values will not be known. However, as shown in Figure 2, it is reasonable to assume that α in the deblurred image will be *lower* than in the given blurred image, provided that image is relatively noise free. Inspection of Table 1, or of other more extensive tables pertaining to the types of images under consideration, may indicate plausible initial estimates for (C, α). As will be shown in section 10, the PSI method is sufficiently robust as to produce good quality reconstructions, even with *inexact* Lipschitz data. Moreover, given a fast algorithm, because of the narrow range of values involved in both C and α, it is feasible to refine such reconstructions by simultaneous computation and display of multiple trial restorations, based on neighboring values of (C, α). Efficient exploration in parameter space is usually the key to the successful solution of inverse problems, when such problems can be solved.

TABLE 1

Values of $(C,\ \alpha)$ *in* $\| U^t f - f \|_p \le C \| f \|_p t^\alpha,\ 0 < t \le 0.1,\ p = 1, 2,$ *for each image* $f(x, y)$ *in Figure* 3*, when* U^t *is Poisson operator in* (19)*. The* (C, α) *values shown below may be more meaningful if rounded to one decimal place. See Remark* 3*.*

Image	*Size*	$(C,\ \alpha)\ \in \Lambda(\alpha, 1, \infty)$	$(C,\ \alpha)\ \in \Lambda(\alpha, 2, \infty)$
Marilyn Monroe	512^2	$C = 0.77,\ \ \alpha = 0.565$	$C = 0.68,\ \ \alpha = 0.474$
Sagittal brain MRI	512^2	$C = 1.28,\ \ \alpha = 0.590$	$C = 1.02,\ \ \alpha = 0.520$
Washington DC Landsat	512^2	$C = 0.45,\ \ \alpha = 0.341$	$C = 0.55,\ \ \alpha = 0.340$
Mariner 5 spacecraft	512^2	$C = 0.90,\ \ \alpha = 0.448$	$C = 0.99,\ \ \alpha = 0.417$
USS Eisenhower	512^2	$C = 0.47,\ \ \alpha = 0.420$	$C = 0.50,\ \ \alpha = 0.362$
English village	512^2	$C = 0.49,\ \ \alpha = 0.472$	$C = 0.55,\ \ \alpha = 0.439$
Nanoscale micrograph	1024^2	$C = 0.45,\ \ \alpha = 0.415$	$C = 0.55,\ \ \alpha = 0.415$
Spiral galaxy M81	1024^2	$C = 0.68,\ \ \alpha = 0.365$	$C = 0.78,\ \ \alpha = 0.327$
Cluster of galaxies	1024^2	$C = 0.65,\ \ \alpha = 0.222$	$C = 0.97,\ \ \alpha = 0.216$

6. Image deblurring in $L^2(R^2)$. We now consider the image deconvolution problem $Pf = g$ with a known *shift-invariant* point spread function (psf) $p(x, y)$,

$$Pf \equiv p(x, y) \otimes f(x, y) = g(x, y), \qquad g(x, y) = g_e(x, y) + n(x, y). \tag{24}$$

Here, $\otimes$ denotes convolution, $g(x, y)$ is the given recorded noisy blurred image, $g_e(x, y)$ is the hypothetical exact blurred image that would have been recorded in the absence of any noise, and $n(x, y)$, presumed small, represents the cumulative effects of all noise processes and other errors affecting final acquisition of the digitized array $g(x, y)$. The noise may be *multiplicative*. Neither $g_e(x, y)$ nor $n(x, y)$ are known, only their sum $g(x, y)$. Denoting the unknown exact sharp image by $f_e(x, y)$, we have

$$Pf_e = p(x, y) \otimes f_e(x, y) = g_e(x, y). \tag{25}$$

Given only (24), we seek a solution $f(x, y)$ in (24) such that $Pf \approx g$, *and* such that $\| f - f_e \|_2$ is small. To achieve this goal, some a priori information about f_e and n is always necessary. Most real images $f_e(x, y)$ contain fine scale features, sharp edges, and other kinds of nondifferentiable singularities. Deblurring techniques that impose stabilizing constraints in the form of prescribed bounds on partial derivatives of $f(x, y)$ in (24), are generally inapplicable, although they are often used. Penalties for such use include smoothing out of sharp features, and possible loss of vital diagnostic information. Indeed, the desire to accurately reconstruct edges and other sharp singularities was the principal reason for developing total variation methods. In fact, several deblurring methods actually exist that do not require prescribed bounds on derivatives [8].

A wide variety of blurs can be used as illustrative examples in (24). Here, we consider the case of uniform defocus blur, where the psf is proportional to the characteristic function of a disc of radius R. This is the so-called "pillbox" model [17], [25], [16], [31]. If $R > 0$ is the radius of the "circle of confusion", the psf for defocus blur is given by

$$p(x, y) = \begin{cases} (\pi R^2)^{-1}, & x^2 + y^2 \le R^2, \\ 0, & x^2 + y^2 > R^2. \end{cases} \tag{26}$$

This has a Fourier transform given by the "sombrero function" [23, p. 72]

$$\hat{p}(\xi, \eta) = 2J_1(R\rho)/(R\rho), \qquad \rho = \sqrt{\xi^2 + \eta^2}, \tag{27}$$

where $J_1(x)$ is the Bessel function of the first kind of order 1. In our numerical experiments below on $2N \times 2N$ images, the expression (27) is used to blur images by Fourier domain multiplication with a preselected $R > 0$, and (ξ, η) are integers with $-N \leq \xi, \eta \leq N$. Rather than interpret R as a radius, we simply observe that the severity of such a blur is determined by the number of zeroes[1] in $|\hat{p}(\rho)|$ on $0 < \rho \leq N$.

6.1. True Wiener filtering and the Tikhonov–Miller method. Wiener filtering [32, p. 356], is an important example of a method that does not impose differentiability constraints. It assumes instead that the power spectra $|\hat{n}(\xi, \eta)|$ and $|\hat{f}_e(\xi, \eta)|$ of each of $n(x, y)$ and $f_e(x, y)$ are known. When this is the case, Wiener filtering produces a solution $f^w(x, y)$ in (24) defined as follows in Fourier space

$$\hat{f}^w(\xi, \eta) = \frac{\overline{\hat{p}}(\xi, \eta)\hat{g}(\xi, \eta)}{|\hat{p}(\xi, \eta)|^2 + |\hat{n}(\xi, \eta)|^2/|\hat{f}_e(\xi, \eta)|^2}, \tag{28}$$

where $\overline{z}$ denotes the complex conjugate of z. Under some additional conditions, it can be shown that $f^w(x, y)$ is an approximate solution of $Pf = g$ that *minimizes* the error $\| f - f_e \|_2$ over all $f \in L^2$. In practice, the power spectra $|\hat{n}(\xi, \eta)|$ and $|\hat{f}_e(\xi, \eta)|$ are very seldom known in advance, and true Wiener filtering is almost never realizable. However, the solution (28) is of considerable theoretical interest because of its optimality property. Note that numerous ad hoc versions of (28) exist, in which more readily available quantities are substituted in place of the required, but unavailable, true power spectra. Such versions are sometimes called Wiener filtering by an abuse of terminology. However, these substitute versions *do not* satisfy the Wiener optimality criterion, nor do they elicit the same degree of theoretical interest.

One of the best-known rigorously analyzable and feasible versions of Wiener filtering is the Tikhonov–Miller method [28], now considered canonical in image deblurring [25]. Significantly, this method makes no a priori assumptions regarding the statistical character of the data noise. For nondifferentiable images, Tikhonov–Miller restoration requires the following a priori information: an upper bound $\epsilon > 0$ for the L^2 norm of the noise $n(x, y)$ in the blurred image $g(x, y)$, and an upper bound M for the L^2 norm of the unblurred image f_e

$$\| n \|_2 = \| Pf_e - g \|_2 \leq \epsilon, \qquad \| f_e \|_2 \leq M, \qquad \epsilon/M \ll 1. \tag{29}$$

It is assumed that ϵ and M are compatible with the existence of an $f_e(x, y) \in L^2$ satisfying (29). Tikhonov–Miller restoration is defined as the unique function $f^\tau(x, y)$ such that

$$f^\tau(x, y) = Arg \min_{f \in L^2(R^2)} \left\{ \| Pf - g \|_2^2 + (\epsilon/M)^2 \| f \|_2^2 \right\}. \tag{30}$$

As will be seen from Theorem 3, where the Tikhonov–Miller method corresponds to the special case $\Gamma_{\overline{t}} = 0$, this minimum problem has a unique solution satisfying

$$Q_\tau f^\tau = P^* g, \qquad Q_\tau = P^* P + (\epsilon/M)^2 I. \tag{31}$$

Moreover, there holds the following best-possible error bound for Tikhonov–Miller reconstruction

$$\| f^\tau - f_e \|_2 \leq \epsilon\sqrt{2} \| Q_\tau^{-1/2} \|_2, \tag{32}$$

[1] The first five positive zeroes of $J_1(x)$ are 3.83171, 7.01559, 10.17347, 13.32369, and 16.47063.

where

$$\| Q_\tau^{-1/2} \|_2 = \sup_{\xi,\eta} \left\{ |\hat{p}(\xi,\eta)|^2 + (\epsilon/M)^2 \right\}^{-1/2}. \tag{33}$$

Given the psf $p(x,y)$, together with the a priori information ϵ, M, one can always find the maximum value in the $2N \times 2N$ array on the right of (33). As in (28) we may implement (31) in Fourier space. We have

$$\hat{f}^\tau(\xi,\eta) = \frac{\overline{\hat{p}}(\xi,\eta)\hat{g}(\xi,\eta)}{|\hat{p}(\xi,\eta)|^2 + (\epsilon/M)^2}. \tag{34}$$

Moreover, from (29) and Parseval's relation

$$\int_{R^2} |\hat{n}(\xi,\eta)|^2 d\xi d\eta \le \epsilon^2, \qquad \int_{R^2} |\hat{f}_e(\xi,\eta)|^2 d\xi d\eta \le M^2. \tag{35}$$

Therefore, the Tikhonov–Miller method can be viewed as an approximate version of true Wiener filtering where the unavailable pointwise values of the spectra in (28) are replaced by more readily available integrals of these quantities. However, it may be anticipated that since true Wiener filtering requires prior knowledge in the form of $8N^2$ numbers for a $2N \times 2N$ image, whereas the Tikhonov–Miller method requires only 2, less accurate results must generally be expected from the latter method.

7. The Poisson Singular Integral (PSI) method for images $\in \Lambda(\alpha, 2, \infty)$. The preceding discussion was necessary to set the stage for the PSI method. Here, in addition to the a priori constraints (29), the behavior of $\| U^t f_e - f_e \|_2$ on $0 \le t \le \bar{t}$ is assumed known, as in the case of Table 1. The constants $C_{\bar{t}}$ and α are now used to place a further constraint on $f_e(x,y)$. For any $f \in L^2(R^2)$, we have on Fourier transforming $f - U^t f$ and using (10),

$$\mathcal{F}\left\{ f - U^t f \right\} = \left(1 - e^{-t\rho}\right) \hat{f}(\xi,\eta), \qquad \rho = \sqrt{\xi^2 + \eta^2}. \tag{36}$$

Therefore, from Parseval's theorem,

$$\int_0^t \| U^s f - f \|_2^2 \, ds = \int_0^t ds \int_{R^2} \left(1 - e^{-s\rho}\right)^2 |\hat{f}(\xi,\eta)|^2 d\xi d\eta. \tag{37}$$

For fixed $t > 0$, define $\hat{z}(\xi,\eta,t) \ge 0$ by

$$\hat{z}(\xi,\eta,t) = \left\{ \int_0^t \left(1 - e^{-s\rho}\right)^2 ds \right\}^{1/2} = \left\{ t + \frac{4e^{-t\rho} - e^{-2t\rho} - 3}{2\rho} \right\}^{1/2}. \tag{38}$$

It follows directly from the integral definition in (38) that for any fixed $t > 0$, $\hat{z}(\rho,t)$ is a strictly increasing function of ρ, and that $\hat{z}(0,0,t) = 0$. For fixed $t > 0$, define the linear operator $Z(t)$ in $L^2(R^2)$ by

$$Z(t)f = \int_{R^2} \hat{z}(\xi,\eta,t)\hat{f}(\xi,\eta)e^{2\pi i(\xi x + \eta y)} d\xi d\eta. \tag{39}$$

Then, from (37),

$$\int_0^t \| U^s f - f \|_2^2 \, ds = \| Z(t)f \|_2^2. \tag{40}$$

For any $f_e \in \Lambda(\alpha, 2, \infty)$, $0 < \alpha < 1$, we have $\| U^s f_e - f_e \|_2 \leq C_{\bar{t}} \| f_e \|_2 s^\alpha$, $0 \leq s \leq \bar{t}$, where $C_{\bar{t}}$ is a positive constant depending on $\bar{t}$, f_e and α. Therefore, with $\| f_e \|_2 \leq M$,

$$\| Z(\bar{t}) f_e \|_2^2 \leq \frac{C_{\bar{t}}^2 M^2 \bar{t}^{1+2\alpha}}{1+2\alpha}. \tag{41}$$

Define

$$\Gamma_{\bar{t}} = \left\{ \frac{1+2\alpha}{C_{\bar{t}}^2 \, \bar{t}^{1+2\alpha}} \right\}^{1/2}. \tag{42}$$

The exact image $f_e(x, y)$ satisfies the following a priori constraints:

$$\| P f_e - g \|_2 \leq \epsilon, \qquad (\epsilon/M) \| f_e \|_2 \leq \epsilon, \qquad (\epsilon/M) \, \Gamma_{\bar{t}} \| Z(\bar{t}) f_e \|_2 \leq \epsilon. \tag{43}$$

Fix $\bar{t} > 0$, and consider the minimization problem

$$f^\psi(x, y) = Arg \min_{f \in L^2(R^2)} \left\{ \| Pf - g \|_2^2 + (\epsilon/M)^2 \left(\| f \|_2^2 + \Gamma_{\bar{t}}^2 \| Z(\bar{t}) f \|_2^2 \right) \right\}. \tag{44}$$

As will be seen in Theorem 3 below, this minimum problem has a unique solution satisfying

$$Q_\psi f^\psi = P^* g, \qquad Q_\psi = P^* P + (\epsilon/M)^2 \left\{ I + \Gamma_{\bar{t}}^2 Z(\bar{t})^* Z(\bar{t}) \right\}. \tag{45}$$

The function $f^\psi(x, y)$ in (44) is defined to be the PSI deblurred image. Moreover, there holds the following error bound for PSI deblurring

$$\| f^\psi - f_e \|_2 \leq \epsilon \sqrt{3} \| Q_\psi^{-1/2} \|_2, \tag{46}$$

where

$$\| Q_\psi^{-1/2} \|_2 = \sup_{\xi, \eta} \left\{ |\hat{p}(\xi, \eta)|^2 + (\epsilon/M)^2 \left(1 + \Gamma_{\bar{t}}^2 |\hat{z}(\xi, \eta, \bar{t})|^2 \right) \right\}^{-1/2}. \tag{47}$$

Given the psf $p(x, y)$, together with the a priori information ϵ, M, and $\Gamma_{\bar{t}}$, one can always find the maximum value in the $2N \times 2N$ array on the right of (47). Again, as in (28) and (34), f^ψ can be found explicitly in Fourier space. We have

$$\hat{f}^\psi(\xi, \eta) = \frac{\bar{\hat{p}}(\xi, \eta) \hat{g}(\xi, \eta)}{|\hat{p}(\xi, \eta)|^2 + (\epsilon/M)^2 \{ 1 + \Gamma_{\bar{t}}^2 |\hat{z}(\xi, \eta, \bar{t})|^2 \}}. \tag{48}$$

Equations (45)–(48) should be compared with equations (31)–(34). Tikhonov–Miller deblurring can then be seen as an extreme case of PSI deblurring, the case where $f_e(x, y)$ is presumed no smoother than the most general L^2 function, so that $\| U^t f_e - f_e \|_2 = o(1)$ as $t \downarrow 0$. This corresponds to $C_{\bar{t}} = \infty$ in (41), and hence, $\Gamma_{\bar{t}} = 0$ in (48).

THEOREM 3. *Fix $\bar{t} > 0$ and let the exact image $f_e(x, y)$ satisfy the a priori constraints* (43). *Let $f^\psi(x, y)$ minimize* (44), *and let Q_ψ be the positive self-adjoint operator on $L^2(R^2)$ given by*

$$Q_\psi = P^* P + (\epsilon/M)^2 \left\{ I + \Gamma_{\bar{t}}^2 Z(\bar{t})^* Z(\bar{t}) \right\}. \tag{49}$$

Then f^ψ is the unique solution of $Q_\psi f^\psi = P^* g$, and f^ψ satisfies

$$\begin{aligned} &\| Pf^\psi - g \|_2^2 + (\epsilon/M)^2 \left\{ \| f^\psi \|_2^2 + \Gamma_{\bar{t}}^2 \| Z(\bar{t}) f^\psi \|_2^2 \right\} \le 3\epsilon^2, \\ &\| P(f^\psi - f_e) \|_2^2 + (\epsilon/M)^2 \left\{ \| f^\psi - f_e \|_2^2 + \Gamma_{\bar{t}}^2 \| Z(\bar{t})(f^\psi - f_e) \|_2^2 \right\} \le 3\epsilon^2. \end{aligned} \tag{50}$$

This implies the L^2 error bound

$$\| f^\psi - f_e \|_2 \le \epsilon \sqrt{3} \| Q_\psi^{-1/2} \|_2, \tag{51}$$

where

$$\| Q_\psi^{-1/2} \|_2 = \sup_{\xi,\eta} \left\{ |\hat{p}(\xi,\eta)|^2 + (\epsilon/M)^2 \left(1 + \Gamma_{\bar{t}}^2 |\hat{z}(\xi,\eta,\bar{t})|^2\right) \right\}^{-1/2}. \tag{52}$$

Proof. Let $\mathcal{H}$ denote the Hilbert space direct sum $L^2(R^2) \bigoplus L^2(R^2) \bigoplus L^2(R^2)$ with elements $[u, v, w]$, scalar product $([u_1, v_1, w_1], [u_2, v_2, w_2]) \equiv \langle u_1, u_2 \rangle + \langle v_1, v_2 \rangle + \langle w_1, w_2 \rangle$, and norm $|||\ |||$. Let $\tilde{P} : L^2(R^2) \mapsto \mathcal{H}$ be defined by $\tilde{P} f = [Pf,\ \omega f,\ \omega \Gamma_{\bar{t}} Z(\bar{t}) f]$, where $\omega = (\epsilon/M)$, and let $\tilde{g} = [g, 0, 0]$. We seek to minimize $|||\tilde{P} f - \tilde{g}|||$ over all $f \in L^2(R^2)$. The normal equation $\tilde{P}^* \tilde{P} f^\psi = \tilde{P}^* \tilde{g}$ gives $Q_\psi f^\psi = P^* g$ with Q as in (49). By hypothesis $|||\tilde{P} f_e - \tilde{g}|||^2 \le 3\epsilon^2$. The minimizing element f^ψ is such that $\tilde{P} f^\psi$ is the orthogonal projection in $\mathcal{H}$ of $\tilde{g}$ on the range of $\tilde{P}$. By the Pythagorean theorem

$$|||\tilde{P} f^\psi - \tilde{g}|||^2 + |||\tilde{P}(f_e - f^\psi)|||^2 = |||\tilde{P} f_e - \tilde{g}|||^2 \le 3\epsilon^2. \tag{53}$$

This proves (50). We now establish (51). From (49), (50),

$$\| Q_\psi^{1/2}(f_e - f^\psi) \|_2^2 = \langle Q_\psi(f_e - f^\psi), (f_e - f^\psi) \rangle = |||\tilde{P}(f_e - f^\psi)|||^2 \le 3\epsilon^2. \tag{54}$$

Hence,

$$\| f_e - f^\psi \|_2 = \| Q_\psi^{-1/2} Q_\psi^{1/2} (f_e - f^\psi) \|_2 \le \epsilon\sqrt{3} \| Q_\psi^{-1/2} \|_2 . \qquad \square \tag{55}$$

8. A preliminary deblurring experiment. In the following controlled experiment, knowledge of the exact solution $f_e(x, y)$ is used to derive *exact* values for all parameters that constitute a priori information in each of the above three methods. Such exact knowledge is not available in practice. The experiment is primarily of theoretical interest. It is designed to illustrate major differences in behavior, and to properly locate the PSI method in relation to Wiener filtering and the Tikhonov–Miller method. The PSI method with *inexact* information is discussed in section 10.

The 8-bit 512×512 Marilyn Monroe image $f_e(x, y)$ in Figure 3 was synthetically defocused by Fourier domain multiplication with the expression in (27) using $R = 0.06$. This produced the exact blurred image $g_e(x, y)$. Multiplicative noise $n(x, y)$ was then added to g_e as follows. Each pixel value $g_e(x, y)$ was perturbed by adding to it the quantity $n(x, y) = 0.03\sigma(x, y) g_e(x, y)$, where $\sigma(x, y)$ is an array of uniformly distributed random numbers in the range $[-1, 1]$. We term this process "adding 3% noise". With varying percentages, we shall use the same process in all our experiments. Note that no noise is thereby added at points where $g_e(x, y) = 0$. The resulting $g(x, y) = g_e(x, y) + n(x, y)$ is shown in Figure 4(A). We find $\| n \|_2 = \epsilon = 2.247$, and $\| f_e \|_1 = 107.59$, $\| f_e \|_2 = M = 131.13$. Therefore $\epsilon/M = 0.01713$. From Table

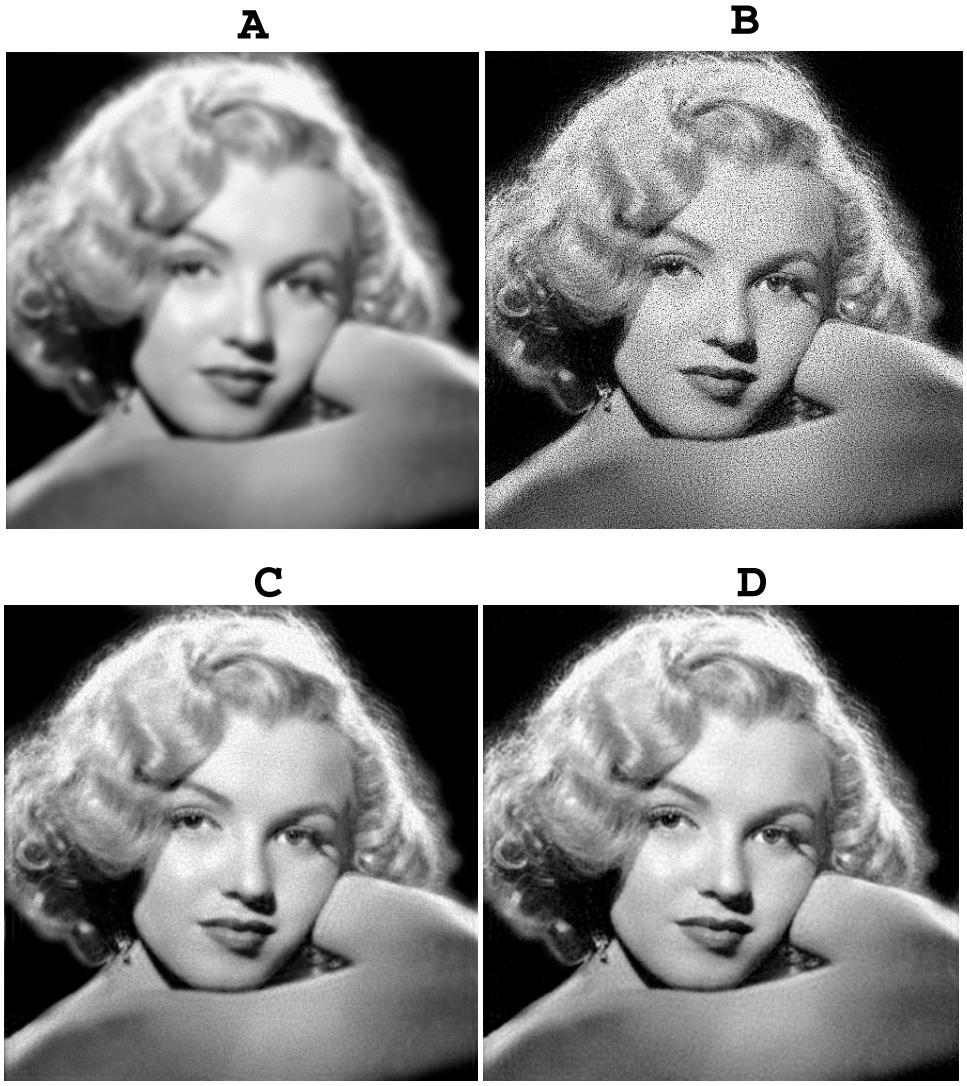

FIG. 4. *Instructive deblurring experiment with exact a priori information highlights significant differences in behavior in above three FFT-based methods.* (A) *Defocused Marilyn Monroe image with* $R = 0.06$ *and* 3% *multiplicative noise.* (B) *Tikhonov–Miller method with exact parameters* ϵ *and* M*, brings out significant noise.* (C) *PSI method with exact parameters* $\epsilon,\ M,\ \alpha = 0.474,\ C_{\bar{t}} = 0.68$. (D) *True Wiener filtering with exact power spectra* $|\hat{n}(\xi,\eta)|,\ |\hat{f}_e(\xi,\eta)|$. *Realizable PSI deblurring closely matches unrealizable true Wiener filtering.*

TABLE 2
Behavior in defocused Marilyn Monroe image in Figure 4.

Deblurring method	L^1 *relative error*	L^2 *relative error*
Tikhonov–Miller (B)	29.82%	34.17%
Poisson Singular Integral (C)	6.89%	9.04%
True Wiener filtering (D)	6.03%	7.88%

1, we have $\| U^t f_e - f_e \|_2 \leq 0.68 \| f \|_2 t^{0.474}$, $0 < t \leq 0.1$. With $\bar{t} = 0.1$, (42) gives $\Gamma_{\bar{t}} = 19.33$. Next, using FFT algorithms, we obtain the exact power spectra $|\hat{n}(\xi,\eta)|$, $|\hat{f}_e(\xi,\eta)|$. We are now ready to compare these three FFT-based procedures under optimal conditions for each method.

The Tikhonov–Miller reconstruction is shown in Figure 4(B). Significantly, this reconstruction is quite noisy, despite the use of *exact* values for ϵ and M. While the regularizing information in (29) prevents *explosive* noise amplification, it is obviously insufficient to prevent serious noise contamination. This is generally the case in the Tikhonov–Miller method. The PSI restoration is shown in Figure 4(C). Here, the additional information that $f_e \in \Lambda(\alpha, 2, \infty)$, together with the values of the constants $C_{\bar{t}}$ and α, were evidently decisive in eliminating noise. The Wiener filtered solution, shown in Figure 4(D), appears only slightly better than the PSI solution. However, the very major difference between true Wiener filtering and the approximate version known as the Tikhonov–Miller method, is another significant result brought out by this deblurring experiment.

It is instructive to study the L^1 and L^2 relative error pattern shown in Table 2. It is widely assumed in practice that the L^2 minimum error property of true Wiener filtering remains valid for the more feasible, approximate versions of such filtering. This is emphatically not the case. The Tikhonov–Miller relative errors are more than *four times larger* than the true Wiener errors. On the other hand, relative errors in the PSI method are only slightly larger than those for true Wiener filtering. Put another way, the PSI method appears to be a feasible procedure that can very substantially improve upon the Tikhonov–Miller method.

Insight into how this improvement comes about can be gained by an analysis of the respective error bounds for each method. Notice that each of the denominators on the right-hand sides of (34) and (48) are radially symmetric functions of (ξ,η), while this is not the case in (28). These denominators play a dual role. They define the actual regularization procedures in (34) and (48), and they define the resulting error bounds in (33) and (47). Because of the radial symmetry, a one-dimensional picture tells the whole story. Define the respective Tikhonov–Miller and PSI *error bound functions* $\theta_\tau(\xi)$, $\theta_\psi(\xi)$ as follows

$$\begin{aligned} \theta_\tau(\xi) &= \left\{ |\hat{p}(\xi,0)|^2 + (\epsilon/M)^2 \right\}^{-1/2}, \\ \theta_\psi(\xi) &= \left\{ |\hat{p}(\xi,0)|^2 + (\epsilon/M)^2 (1 + \Gamma_{\bar{t}}^2 |\hat{z}(\xi,0,\bar{t})|^2) \right\}^{-1/2}. \end{aligned} \tag{56}$$

In Figure 5, we plot $\theta_\tau(\xi)$ and $\theta_\psi(\xi)$ as determined by the actual parameter values that entered the deblurring experiment in Figure 4. The significant differences in these two curves translate into fundamental differences in the Fourier space regularization that defines the corresponding procedures. From (27), we see that $\theta_\tau(\xi)$ has a maximum of $M/\epsilon = 58.36$, at every point $\xi > 0$ where $J_1(0.06\ \xi) = 0$. There are 4 such points on $0 < \xi \leq 256$. The curve $\theta_\psi(\xi)$ also develops maxima at these same points, but these maxima are about *five times smaller* than those in $\theta_\tau(\xi)$, owing to the additional term involving $\hat{z}(\xi,0,\bar{t})$. Since the error estimate in each method is proportional to the maximum along the corresponding curve, it is natural to find substantially smaller errors in Figure 4(C) than in Figure 4(B).

9. Comparing total variation deblurring with PSI deblurring. The use of initial value PDE methods in image processing and computer vision has mushroomed into an important new branch of applied mathematics. The basic idea originates in

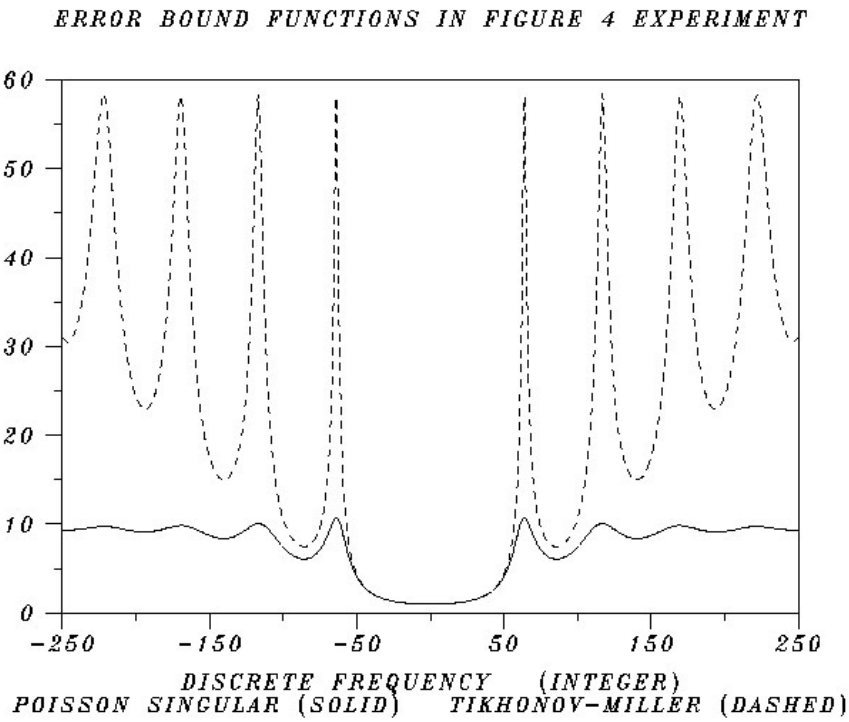

FIG. 5. *Plot of error bound functions $\theta_\tau(\xi)$ (dashed curve), and $\theta_\psi(\xi)$ (solid curve), as defined in* (56)*, for the deblurring experiment in Figure* 4*. Maximum value in θ_τ is more than five times larger than in θ_ψ. Qualitative difference in behavior in these two curves implies significant difference in Fourier domain regularization in the PSI and Tikhonov–Miller methods. Difference in maximum values explains large difference in L^2 relative errors in Figures* 4(B) *and* 4(C).

gradient descent methods for minimizing appropriate energy functionals. Instructive surveys of this general set of ideas may be found in [11] and [39].

The total variation approach introduced in [33] is one of the most popular PDE methods, and it is primarily designed to recover edges in the original image. Given the deconvolution problem $Pf = g$ as in (24), TV deblurring presupposes the exact sharp image $f_e(x, y) \in BV(R^2)$, and it produces an image $f^{tv}(x, y)$ defined by

$$f^{tv}(x,y) = Arg \min_{f \in BV(R^2)} \left\{ (\lambda/2) \parallel Pf - g \parallel_2^2 + \int_{R^2} |\nabla f| dxdy \right\}. \tag{57}$$

This means that $f^{tv}(x, y)$ is the solution of

$$P^* P f^{tv} - \lambda^{-1} \nabla . \left(\frac{\nabla f^{tv}}{|\nabla f^{tv}|} \right) = P^* g. \tag{58}$$

Here, $\lambda > 0$ is a regularization parameter that can be tuned. Provided the noise level in $g(x, y)$ is small, larger values of λ produce sharper images. Too large a value of λ leads to computational instability. Unlike the cases in (31) and (45), (58) is a nonlinear deconvolution problem that cannot be solved explicitly in Fourier space. In fact, considerable effort is generally required to obtain f^{tv} for large size imagery. In pure denoising applications, where $P = I$, this effort is usually warranted by the

quality of the resulting restoration. Recently, a new time dependent evolutionary approach to (58) has been developed [26], whereby $f^{tv}(x,y)$ is obtained as the steady state solution to the following nonlinear anisotropic diffusion problem

$$\left\{ \begin{array}{l} u_t = -\lambda|\nabla u|\ P^*(Pu-g) + |\nabla u|\ \nabla.\left(\nabla u/\{\sqrt{|\nabla u|^2+\beta}\}\right), \\ u(x,y,0) = g(x,y), \end{array} \right. \tag{59}$$

where the given blurred image $g(x,y)$ is used as the initial value. In addition, $u(x,y,t)$ satisfies homogeneous Neumann conditions at the boundary of the unit square Ω. In (59), $\beta > 0$ is a small constant designed to prevent division by zero. In [26, section 5], an efficient new explicit finite difference scheme for (59) is proposed. This scheme has improved stability and edge-enhancing properties, and converges rapidly to the desired steady state solution. Accordingly, we shall use that method in our total variation deblurring experiments.

This paper has drawn attention to the fact that most images are not smooth. The PSI method is predicated on locating $f_e(x,y)$ in the correct Lipschitz space, while TV deblurring assumes $f_e(x,y) \in BV(R^2)$. It may be argued that such refined smoothness measures are primarily applicable to $f^\infty(x,y)$, the original intensity field that gave rise to the digitized finite dimensional object $f_e(x,y)$, but may not be meaningful for $f_e(x,y)$ itself. Indeed, since all norms are equivalent in finite dimensional space, it remains to be seen whether such abstruse function space notions are ultimately of any computational significance.

Our first experiment involves a slightly defocused image with very little noise. The original sharp USS Eisenhower image is shown in Figure 6(A). Fourier space multiplication with (27) using $R = 0.03$, followed by the addition of 0.1% multiplicative noise, produced the blurred Figure 6(B). Because of the low noise level, we chose β small and λ large in (59), as recommended in [26]. With $\beta = 0.0001,\ \Delta t = 0.1(\Delta x)^2$ and $\lambda = 300$, we obtained Figure 6(C) at $T = 100\Delta t$. Higher values of λ were computationally unstable. Moreover, the resulting TV image did not improve if more time steps were taken. Figure 6(D) is the PSI deblurred image using exact values for $\epsilon,\ M$, and using $\alpha = 0.362,\ C_{\bar{t}} = 0.50$, from Table 1. Zooming on selected parts of the image in Figures 6(E) and 6(F), clearly shows significant loss of structural detail in the TV image, as compared with PSI deblurring. For completeness, the L^1 relative errors in this experiment were as follows: true Wiener filtering (not shown) 1.67%, PSI method 2.18%, TV deblurring 6.83%, and Tikhonov–Miller (not shown) 4.71%.

In our second experiment, the sharp English village image in Figure 7(A) was moderately defocused using $R = 0.06$, and 0.1% multiplicative noise was again added to form Figure 7(B). With β and Δt as in Figure 6(C), it was possible to choose $\lambda = 500$, and obtain Figure 7(C) at $T = 100\Delta t$. Again, no improvement was noted with more time steps. Figure 7(D) is the PSI deblurred image using the exact values for $\epsilon,\ M$, together with $\alpha = 0.439,\ C_{\bar{t}} = 0.55$, from Table 1. Because of the stronger blur, more information is now lost in TV deblurring. Zooming in on the first three houses in Figures 7(E) and 7(F), we see that the windows and roof shingles have been virtually eliminated in the TV image. The L^1 relative errors in this experiment were as follows: true Wiener filtering (not shown) 1.98%, PSI method 3.05%, TV deblurring 6.70%, and Tikhonov–Miller (not shown), 7.42%.

10. The PSI method with inexact Lipschitz data. The controlled experiments in sections 8 and 9 were designed to illustrate important theoretical points, and involved use of the PSI method with exact prior Lipschitz space data. In fact,

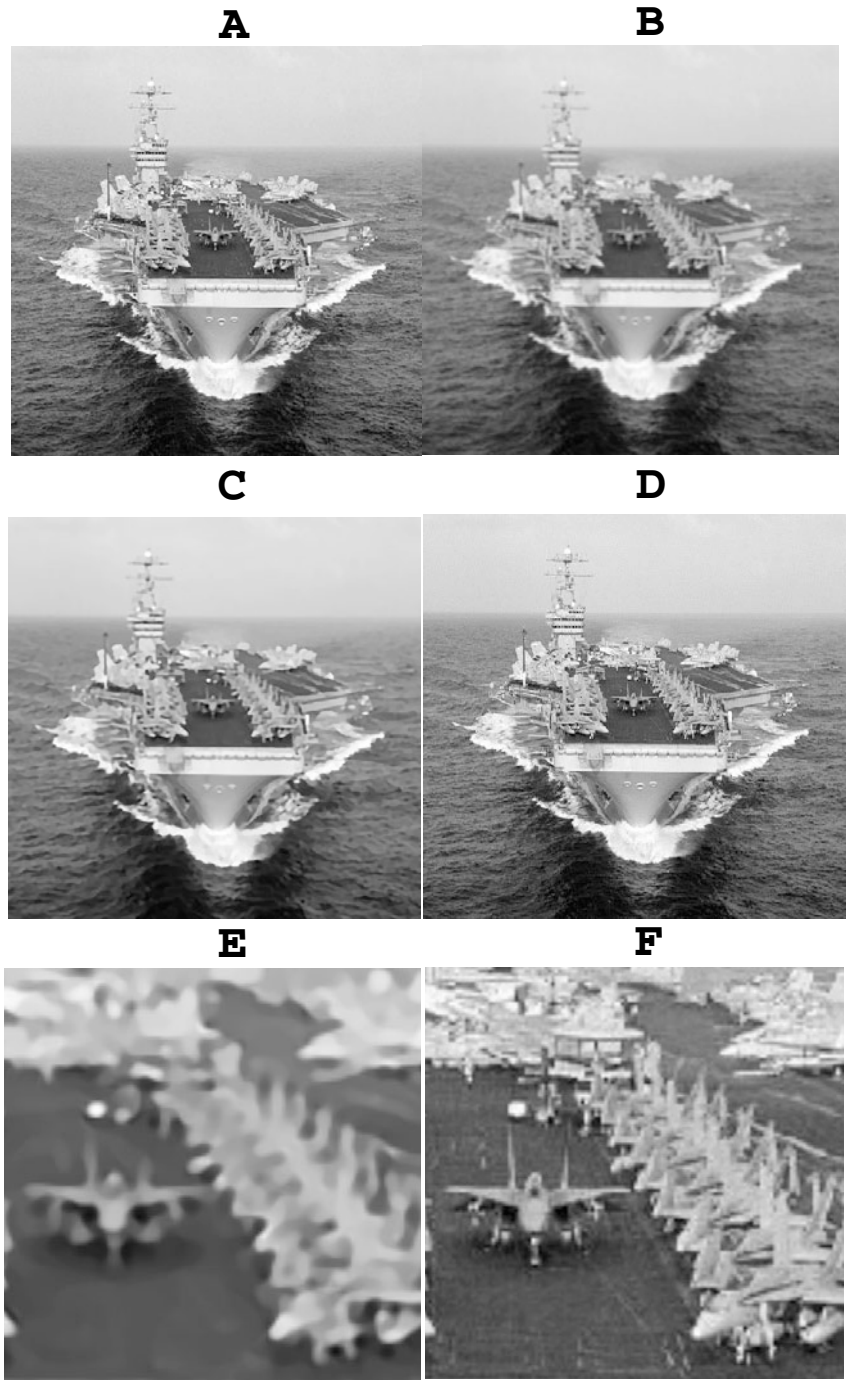

FIG. 6. *Comparison of total variation and PSI deblurring on mildly blurred image. Zooming on selected parts of the image enables meaningful comparisons between the two methods.* (A) *Original sharp USS Eisenhower image.* (B) *Mildly defocused image with* $R = 0.03$ *and* 0.1% *multiplicative noise.* (C) *Total variation deblurring by applying finite difference scheme in* [26, section 5] *to* (59), *with* $\beta = 0.0001,\ \lambda = 300,\ \Delta t = 0.1(\Delta x)^2,\ T = 100\Delta t$. (D) *PSI deblurring using exact a priori parameters,* $\epsilon,\ M,\ \alpha = 0.362,\ C_{\overline{t}} = 0.50$. (E) *Zooming in TV deblurred image reveals significant loss of structural detail.* (F) *Zooming on same region in PSI image.*

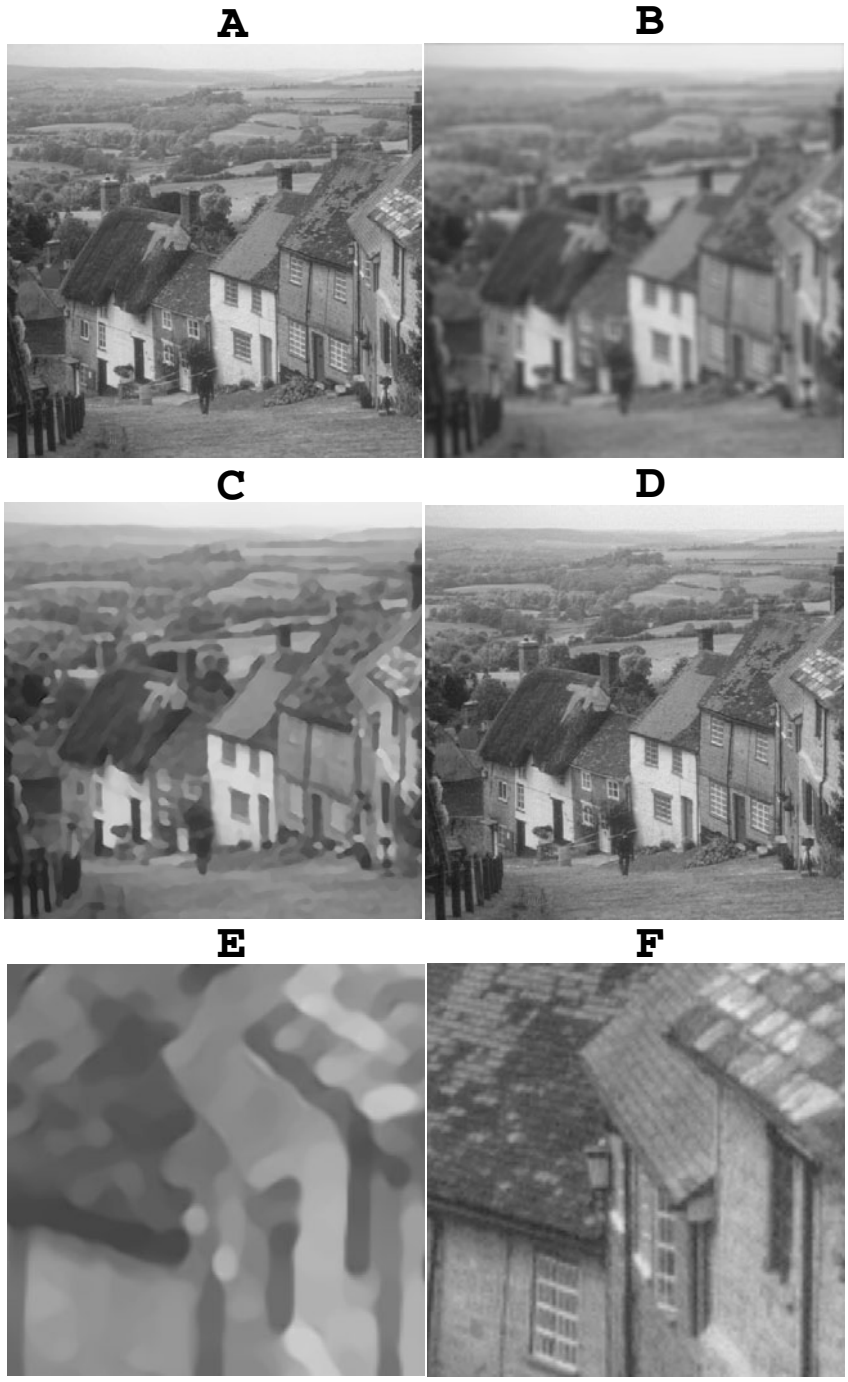

FIG. 7. *Comparison of total variation and PSI methods on moderately blurred image. Zooming now reveals unacceptable loss of content in TV deblurring.* (A) *Original sharp English village image.* (B) *Moderately defocused image with* $R = 0.06$ *and* 0.1% *multiplicative noise.* (C) *Total variation deblurring using scheme in* [26, section 5] *with* $\beta = 0.0001,\ \lambda = 500,\ \Delta t = 0.1(\Delta x)^2,\ T = 100\Delta t$. (D) *PSI method with exact a priori parameters,* $\epsilon,\ M,\ \alpha = 0.439,\ C_{\overline{t}} = 0.55$. (E) *Zooming in image* (C) *reveals loss of windows and roof shingles.* (F) *Zooming on same region in PSI image.*

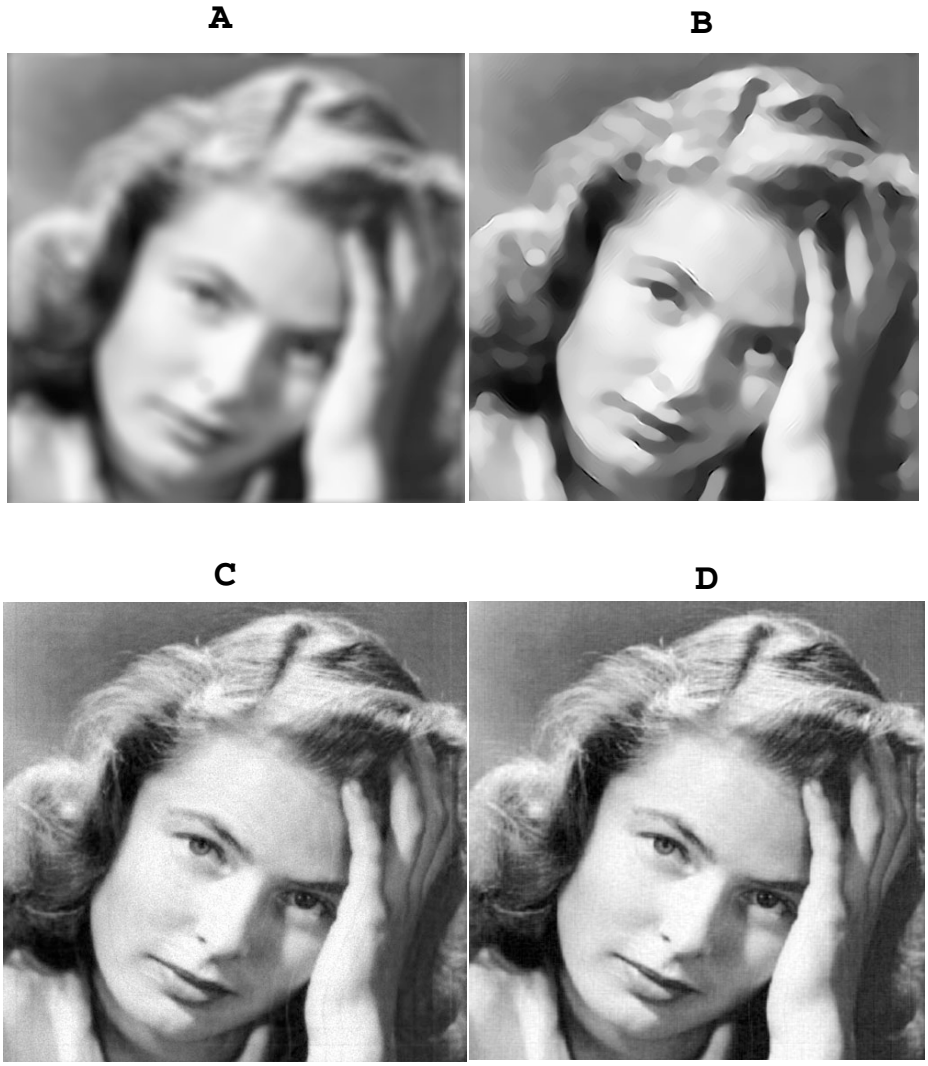

FIG. 8. *Use of plausible guess. Robust PSI method can produce useful reconstructions, even with inexact Lipschitz data.* (A) *Moderately defocused Ingrid Bergman image with* $R = 0.12$ *and* 0.5% *multiplicative noise.* (B) *Total variation deblurring using scheme in* [26, section 5], *with* $\lambda = 400,\ \beta = 0.0001,\ \Delta t = 0.1(\Delta x)^2$, *and* $T = 150\Delta t$, *produces lifeless, mannequin-like appearance.* (C) *PSI method using plausible guess* $\alpha = 0.5,\ C_{\bar{t}} = 0.5$. (D) *True Wiener filtering with exact power spectra* $|\hat{n}(\xi,\eta)|,\ |\hat{f}_e(\xi,\eta)|$. *Fast PSI deblurring, with fictitious Lipschitz data, produces good first approximation to unrealizable optimal Wiener image.*

TABLE 3
Behavior in moderately defocused Ingrid Bergman image in Figure 8.

Deblurring method	L^1 *relative error*	L^2 *relative error*
Tikhonov–Miller (not shown)	24.15%	27.33%
Total Variation (B)	5.79%	8.45%
PSI with plausible guess (C)	4.75%	5.76%
True Wiener filtering (D)	3.53%	4.40%

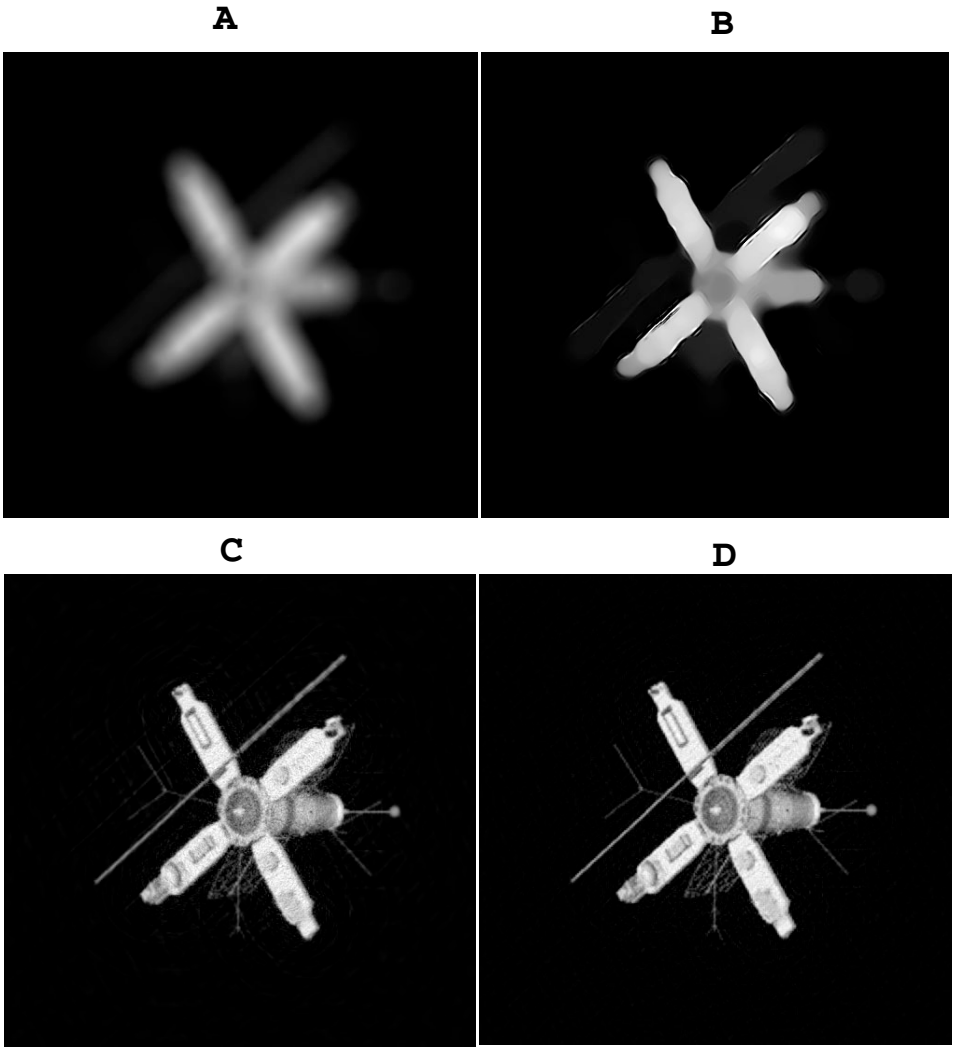

FIG. 9. *Use of substitute information. Robust PSI method produces remarkably good reconstruction using Lipschitz space data corresponding to image of "similar" object.* (A) *Strongly defocused USAF satellite image with* $R = 0.25$ *and* 0.5% *multiplicative noise.* (B) *Total variation deblurring using scheme in* [26, section 5], *with* $\lambda = 400,\ \beta = 0.0001,\ \Delta t = 0.1(\Delta x)^2$, *and* $T = 150\Delta t$, *results in severe loss of texture.* (C) *PSI method using substitute Lipschitz data* $\alpha = 0.417,\ C_{\bar{t}} = 0.99$, *obtained from Mariner* 5 *image in Figure* 3. (D) *True Wiener filtering with exact power spectra* $|\hat{n}(\xi,\eta)|,\ |\hat{f}_e(\xi,\eta)|$. *Fast PSI deblurring, using substitute data, closely matches unrealizable optimal Wiener image.*

TABLE 4
Behavior in strongly defocused USAF satellite image in Figure 9.

Deblurring method	L^1 *relative error*	L^2 *relative error*
Tikhonov–Miller (Not shown)	37.95%	33.56%
Total Variation (B)	32.91%	31.82%
PSI based on similar object (C)	20.69%	16.83%
True Wiener filtering (D)	17.50%	13.04%

the PSI method is a robust method of great practical significance that can produce useful reconstructions even with *inexact* Lipschitz data. Inspection of Table 1 shows that the values of (C, α) are typically confined to a narrow range. Indeed, a plausible guess for (C, α) might be $(0.5, 0.5)$ in many cases. In other situations, a sharp image of a similar object might provide useful values for (C, α). Such initial reconstructions can then be interactively refined through *fast* simultaneous computation and display of multiple trial PSI images, corresponding to neighboring (C, α) values. We now give two examples that show how good such *initial* reconstructions can be.

Figure 8(A) is an 8-bit 512×512 synthetically defocused Ingrid Bergman image, obtained using (27) with $R = 0.12$, followed by adding 0.5% multiplicative noise. Total variation deblurring using the scheme in [26, section 5], with $\lambda = 400$, $\beta = 0.0001$, and $T = 150\Delta t$, produces the lifeless, mannequin-like appearance shown in Figure 8(B). However, the PSI method with the plausible guess $\alpha = 0.5$, $C_{\bar{t}} = 0.5$, produces Figure 8(C). This initial reconstruction is already in good qualitative agreement with the true Wiener image in Figure 8(D). The L^1 and L^2 relative errors in Table 3 indicate that the PSI method, even with such inexact data, significantly improves upon the Tikhonov–Miller and total variation methods.

Our final experiment involves the strongly defocused USAF satellite image in Figure 9(A), where $R = 0.25$ and 0.5% multiplicative noise was added. As may be expected in such a severely blurred image, total variation deblurring, shown in Figure 9(B), results in severe loss of structural detail. Here, the Mariner 5 image shown in Figure 3 may be considered a "similar" object, and the corresponding $\Lambda(\alpha, 2, \infty)$ information in Table 1, $C = 0.99$, $\alpha = 0.417$, may be used in the PSI method. Remarkably, this produces the reconstruction shown in Figure 9(C). While faint honeycomb artifacts are visible against the dark background in Figure 9(C), this initial PSI image is an excellent approximation to the optimal true Wiener image shown in Figure 9(D). Relative errors in this experiment are shown in Table 4.

In Figure 4, the PSI method's improvement over the Tikhonov–Miller method can be traced to the fact that the constraints in (29) allowed the solution to be *too rough*. In Figures 6 through 9, PSI's improvement over the total variation method stems from the fact that the minimum principle (57) forces the solution to be *too smooth*. Apparently, the use of singular integrals to calibrate image smoothness, together with the direct use of that information in constraining the solutions of the deblurring problem, constitutes an important new idea in image deconvolution.

REFERENCES

[1] S. Bochner, *Harmonic Analysis and the Theory of Probability*, University of California Press, Berkeley, CA, 1955.

[2] E. O. Brigham, *The Fast Fourier Transform*, Prentice-Hall, Englewood Cliffs, NJ, 1974.

[3] P. L. Butzer and H. Berens, *Semi-Groups of Operators and Approximation*, Springer-Verlag, New York, 1967.

[4] P. L. Butzer and R. J. Nessel, *Favard classes for n-dimensional singular integrals*, Bull. Amer. Math. Soc., 72 (1966), pp. 493–498.

[5] P. L. Butzer and R. J. Nessel, *Fourier Analysis and Approximation*, Academic Press, New York, 1971.

[6] A. P. Calderón and A. Zygmund, *Singular integrals and periodic functions*, Studia Math., 14 (1954), pp. 249–271.

[7] A. S. Carasso and T. Kato, *On subordinated holomorphic semigroups*, Trans. Amer. Math. Soc., 327 (1991), pp. 867–877.

[8] A. S. Carasso, *Linear and nonlinear image deblurring: A documented study*, SIAM J. Numer. Anal., 36 (1999), pp. 1659–1689.

[9] A. S. Carasso, *The APEX method in image sharpening and the use of low exponent Lévy stable laws*, SIAM J. Appl. Math., 63 (2002), pp. 593–618.

[10] A. S. Carasso, D. S. Bright, and A. E. Vladár, *The APEX method and real-time blind deconvolution of scanning electron microscope imagery*, Optical Engineering, 41 (2002), pp. 2499–2514.

[11] V. Caselles, J. M. Morel, G. Sapiro, and A. Tannenbaum, eds., *Special Issue on Partial Differential Equations and Geometry-Driven Diffusion in Image Processing*, IEEE Trans. Image Process., 7, 1998.

[12] A. Chambolle, R. A. DeVore, N. Lee, and B. J. Lucier, *Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage*, IEEE Trans. Image Process., 7 (1998), pp. 319–335.

[13] A. Chambolle and P. L. Lions, *Image recovery via total variation minimization and related problems*, Numer. Math., 76 (1997), pp. 167–188.

[14] T. Chan, A. Marquina, and P. Mulet, *High-order total variation-based image restoration*, SIAM J. Sci. Comput., 22 (2000), pp. 503–516.

[15] T. F. Chan and J. Shen, *Mathematical models for local nontexture inpaintings*, SIAM J. Appl. Math., 62 (2002), pp. 1019–1043.

[16] T. F. Chan and C. K. Wong, *Total variation blind deconvolution*, IEEE Trans. Image Process., 7 (1998), pp. 370–375.

[17] S. Chaudhuri and A. N. Rajagopalan, *Depth from Defocus: A Real Aperture Imaging Approach*, Springer-Verlag, New York, 1999.

[18] R. A. DeVore, B. Jawerth, and B. J. Lucier, *Image compression through wavelet transform coding*, IEEE Trans. Inform. Theory, 38 (1992), pp. 719–746.

[19] R. A. DeVore and B. J. Lucier, *Classifying the smoothness of images: Theory and application to wavelet image processing*, in Proceedings of the 1994 IEEE International Conference on Image Processing, Austin TX, Vol. 2, IEEE Press, Los Alamitos, CA, 1994, pp. 6–10.

[20] D. C. Dobson and F. Santosa, *Recovery of blocky images from noisy and blurred data*, SIAM J. Appl. Math., 56 (1996), pp. 1181–1198.

[21] S. Durand and J. Froment, *Reconstruction of wavelet coefficients using total variation minimization*, SIAM J. Sci. Comput., 24 (2003), pp. 1754–1767.

[22] W. Feller, *An Introduction to Probability Theory and Its Applications*, 2nd ed., Vol. 2, Wiley, New York, 1971.

[23] J. D. Gaskill, *Linear Systems, Fourier Transforms, and Optics*, Wiley, New York, 1978.

[24] Y. Gousseau and J.-M. Morel, *Are natural images of bounded variation?*, SIAM J. Math. Anal., 33 (2001), pp. 634–648.

[25] R. L. Lagendijk and J. Biemond, *Iterative Identification and Restoration of Images*, Kluwer Academic Publishers, Norwell, MA, 1991.

[26] A. Marquina and S. Osher, *Explicit algorithms for a new time dependent model based on level set motion for nonlinear deblurring and noise removal*, SIAM J. Sci. Comput., 22 (2000), pp. 387–405.

[27] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, Univ. Lecture Ser. 22, AMS, Providence RI, 2001.

[28] K. Miller, *Least squares methods for ill-posed problems with a prescribed bound*, SIAM J. Math. Anal., 1 (1970), pp. 52–74.

[29] M. Nikolova, *Local strong homogeneity of a regularized estimator*, SIAM J. Appl. Math., 61 (2000), pp. 633–658.

[30] M. Nikolova, *Minimizers of cost-functions involving nonsmooth data-fidelity terms. Application to the processing of outliers*, SIAM J. Numer. Anal., 40 (2002), pp. 965–994.

[31] E. L. O'Neill, *Introduction to Statistical Optics*, Addison-Wesley, Reading, MA, 1963.

[32] W. K. Pratt, *Digital Image Processing*, 2nd ed., Wiley, New York, 1991.

[33] L. Rudin and S. Osher, *Total variation based image restoration with free local constraints*, in Proceedings of the 1994 IEEE International Conference on Image Processing, Austin, TX, Vol. 1, 1994, IEEE Press, Los Alamitos, CA, pp. 31–35.

[34] A. Sandage, *The Hubble Atlas of Galaxies*, Publication 618, Carnegie Institution of Washington, Washington, DC, 1961.

[35] A. M. Stokolos and W. Trebels, *On the rate of almost everywhere convergence of Abel-Cartwright means on $L^p(R^n)$*, Result. Math., 34 (1998), pp. 373–380.

[36] M. H. Taibleson, *Lipschitz classes of functions and distributions in E_n*, Bull. Amer. Math. Soc., 69 (1963), pp. 487–493.

[37] M. H. Taibleson, *On the theory of Lipschitz spaces of distributions on Euclidean n-space.* I. *Principal properties*, J. Math. Mechanics, 13 (1964), pp. 407–478.

[38] A. E. Vladár, M. T. Postek, and M. P. Davidson, *Image sharpness measurement in scanning electron microscopy*, Scanning, 20 (1998), pp. 24–34.

[39] J. Weickert, *Anisotropic Diffusion in Image Processing*, Teubner, Stuttgart, Germany, 1998.

[40] K. Yosida, *Fractional powers of infinitesimal generators and the analyticity of the semigroups generated by them*, Proc. Japan Acad., 36 (1960), pp. 86–89.