

NISTIR 7727

**Fractional Diffusion, Low Exponent
Levy Stable Laws, and Slow Motion
Denoising of Helium Ion Microscope
Nanoscale Imagery**

Alfred S. Carasso
András E. Vladár

NISTIR 7727

**Fractional Diffusion, Low Exponent
Levy Stable Laws, and Slow Motion
Denoising of Helium Ion Microscope
Nanoscale Imagery**

Alfred S. Carasso

*Applied and Computational Mathematics Division
Information Technology Laboratory*

András E. Vladár

*Precision Engineering Division
Manufacturing Engineering Laboratory*

September 2010



U.S. Department of Commerce
Gary Locke, Secretary

National Institute of Standards and Technology
Patrick D. Gallagher, Director

FRACTIONAL DIFFUSION, LOW EXPONENT LÉVY STABLE LAWS, AND ‘SLOW MOTION’ DENOISING OF HELIUM ION MICROSCOPE NANOSCALE IMAGERY

ALFRED S. CARASSO* AND ANDRÁS E. VLADÁR†

Abstract. Helium ion microscopes (HIM) are capable of acquiring images with better than 1 nm resolution, and HIM images are particularly rich in morphological surface details. However, such images are generally quite noisy. A major challenge is to denoise these images while preserving delicate surface information. This paper presents a powerful *slow motion* denoising technique, based on solving linear fractional diffusion equations forward in time. The method is easily implemented computationally, using fast Fourier transform (FFT) algorithms. When applied to actual HIM images, the method is found to reproduce the essential surface morphology of the sample with high fidelity. In contrast, such highly sophisticated methodologies as Curvelet Transform denoising, and Total Variation denoising using split Bregman iterations, are found to eliminate vital fine scale information, along with the noise. Image Lipschitz exponents are a useful image metrology tool for quantifying the fine structure content in an image. In this paper, this tool is applied to rank order the above three distinct denoising approaches, in terms of their texture preserving properties. In several denoising experiments on actual HIM images, it was found that fractional diffusion smoothing performed noticeably better than split Bregman TV, which in turn, performed slightly better than Curvelet denoising.

Key words. HIM images; image denoising; image texture; image metrology; total variation; curvelet transform; low exponent Lévy stable laws; image Lipschitz exponents; surface morphology.

1. Introduction. This paper presents an easily implemented denoising methodology, based on solving linear fractional diffusion equations using fast Fourier transform (FFT) algorithms. When applied to state of the art nanoscale imagery, this method can outperform computationally more sophisticated denoising techniques based on *Curvelet Transform* thresholding, or on minimizing image *Total Variation*.

Both scanning He ion microscopes (HIM), and scanning electron microscopes (SEM), are capable of acquiring images with better than 1 nm resolution. Such imagery is very much needed in nano-scale research and development and production. HIM images provide better surface-related information than SEM images, but are generally noisier. SEM images at the highest magnifications are also prone to noise. The important nano-structures are usually very small, and only the sharpest focused beams can resolve them. Such beams, formed with pA currents, involve a few hundred to a few thousand atoms, and do not generate a lot of signal. This necessarily produces very noisy images. Any technique that can improve this situation is very much needed, and would result in otherwise unavailable acquisition speed and/or information. However, preserving the fidelity of the essential sample-related fine details is of paramount interest.

This paper develops a useful approach to this difficult problem, based on solving linear fractional diffusion equations. The use of fractional diffusion equations and related isotropic Lévy stable laws, were introduced in image analysis in [4], where they were applied to solve an important class of image deblurring problems. Subsequently, it was discovered that *low exponent* Lévy stable laws could be successfully applied in *blind deconvolution* of a large variety of real blurred imagery of considerable scientific

*Applied and Computational Mathematics Division, National Institute of Standards and Technology, Gaithersburg, MD 20899. (alfred.carasso@nist.gov).

†Precision Engineering Division, National Institute of Standards and Technology, Gaithersburg, MD 20899. (andras.vladar@nist.gov)

interest, including SEM images, and Hubble space telescope and other astronomical images, [5], [7], [8]. In such deblurring applications, one solves *ill-posed* fractional diffusion equations backwards in time, with the blurred image as data at time $t = 1$. In contrast, the present denoising application involves solving *well-posed* linear equations forward in time, with the noisy image as data at time $t = 0$. In that context, the significance of low Lévy exponents is expressed by the sharp inequality developed in Eq. (15) below. That inequality indicates that for small $t > 0$, low Lévy exponent fractional diffusion smoothing retains considerably more of the fine structure in the initial data, than does Gaussian smoothing, which corresponds to smoothing with the classical heat conduction equation.

In an entirely different direction, there has been considerable interest in recent years in the use of nonlinear anisotropic diffusion equations in image denoising, [1], [15], [16], [17]. Here, the denoised image is defined to be the *steady-state* solution to this nonlinear problem. The *Total Variation (TV)* approach is the best-known example of this class of methods. Up to date surveys of *TV* denoising methodologies together with useful software, may be found in [10] and [12]. The *TV* approach is especially useful when the ideal sharp image is piecewise smooth, and consists of isolated smooth objects with well-defined edges. Such images belong to the class $BV(R^2)$ of functions of *bounded variation*, which plays an essential role in this theory. Important examples of effective *TV* denoising of such images, in the field of medical computed tomography, are discussed in [23].

Another important class of methods centers on wavelet transforms, and filtering the image by appropriately thresholding the wavelet coefficients. Recently, a more effective approach has been developed, based on the use of *Curvelets* that can better represent curved edges in the image, [3], [20]. In [18], curvelet transform denoising is again successfully applied to computed tomography brain slices.

While both *TV* and *Curvelet* denoising work quite well on $BV(R^2)$ images, many important classes of images $f(x, y)$ display significant fine scale details or *texture*, together with amorphous features, and do not belong to $BV(R^2)$, [6], [13]. Use of *TV* or *Curvelet* processing of such imagery may eliminate texture along with the noise, while preserving edges. The L^1 Lipschitz exponent α , where $0 < \alpha \leq 1$, is a mathematical index that can capture the fine-structure content and degree of unsmoothness in an image, provided that image is relatively *noise free*. Images that are of bounded variation (including smoothly differentiable images) have $\alpha = 1$. The value of α decreases systematically with increasing roughness. Images with significant non differentiable small scale structures typically have $\alpha \ll 1$. A method of estimating image Lipschitz exponents is developed in [6] and applied to image restoration in [9].

HIM images are examples of images where texture and detailed surface morphology are of prime interest, and for which *TV* and *Curvelet* denoising may not be appropriate. In this paper, we exhibit several examples of real HIM data, where fractional diffusion denoising is superior to both *Curvelet* and *split Bregman TV* denoising, in retaining essential sample-related features. These improvements may not be visually apparent in the reduced size images in the printed issue of this journal. However, significant enhancement becomes evident when the *on-line* version of this paper is viewed at full size on a modern high resolution device, such as a wide screen, active matrix, liquid crystal display (LCD) monitor. By estimating image Lipschitz exponents α before and after denoising, we can use α as a useful metric to rank order these three distinct denoising methods in their ability to retain texture, and confirm the visual results.

2. Fourier space characterization of textured imagery. For purposes of theoretical analysis it is helpful to think of images $f(x, y)$ as functions in $L^p(\mathbb{R}^2)$, $p = 1, 2$, i.e., functions such that

$$(1) \quad \|f\|_p = \left\{ \int_{\mathbb{R}^2} |f(x, y)|^p dx dy \right\}^{1/p} < \infty, \quad p = 1, 2.$$

Define the 2D Fourier transform of $f(x, y)$ by

$$(2) \quad \mathcal{F}\{f\} = \hat{f}(\xi, \eta) \equiv \int_{\mathbb{R}^2} f(x, y) \exp\{-2\pi i(\xi x + \eta y)\} dx dy.$$

The L^1 norm of $f(x, y)$ is proportional to the total image radiant flux, and conservation of $\|f\|_1$ is a desirable attribute in any image processing method. Since $f(x, y) \geq 0$, it follows from Eq. (2) that $\hat{f}(0, 0) = \|f\|_1$.

Let $|\nabla f| = (f_x^2 + f_y^2)^{1/2}$. The class $BV(\mathbb{R}^2)$ is the class of functions $f(x, y)$ for which $|\nabla f| \in L^1(\mathbb{R}^2)$. Thus, $\|\nabla f\|_1$ is finite. Assume $|\nabla f| \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ so that both $\|\nabla f\|_1$ and $\|\nabla f\|_2$ are finite. Then, from Parseval's theorem

$$(3) \quad \int_{\mathbb{R}^2} |\nabla f(x, y)|^2 dx dy = \int_{\mathbb{R}^2} (f_x^2 + f_y^2) dx dy = \int_{\mathbb{R}^2} (\xi^2 + \eta^2) |\hat{f}(\xi, \eta)|^2 d\xi d\eta < \infty.$$

This implies that $|\hat{f}(\xi, \eta)|$ must decay sufficiently fast at infinity to make the last integral converge. Images $f(x, y)$ with significant fine structure need not satisfy the assumption $|\nabla f| \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$, and both $\|\nabla f\|_1$ and $\|\nabla f\|_2$, may be infinite. In that case, $|\hat{f}(\xi, \eta)|$ does not decay fast enough at infinity.

3. Total variation (TV) denoising. Given a noisy image $f(x, y)$, and the regularization parameter $\omega > 0$, total variation denoising seeks a function $f^T(x, y)$ with finite $\|\nabla f^T\|_1$, such that

$$(4) \quad f^T(x, y) = \text{Arg} \min_{u \in BV(\mathbb{R}^2)} \{ \|\nabla u\|_1 + \omega/2 \|u - f\|_2^2 \}.$$

There are several methods that can be used to solve this minimization problem, [10], [12], [15]. The method in [15] obtains $f^T(x, y)$ as the unique steady-state solution to a nonlinear anisotropic diffusion equation with $f(x, y)$ as initial data. A more effective method is the split Bregman iteration discussed in [12]. In this paper, we apply a MATLAB implementation of the split Bregman approach, as developed by the authors in [12], to denoise HIM images. Split Bregman TV denoising of $f(x, y)$ typically reduces $\|\nabla f\|_1$ considerably, and need not preserve $\|f\|_1$.

4. Curvelet denoising. The curvelet transform is designed to represent edges and other singularities along curves, more efficiently than traditional wavelet methods, [3], [20]. In this paper we use a MATLAB implementation of curvelet denoising developed by the authors in [20], and made available in their CurveLab package. The basic idea is discussed in [20, section 5]. The noisy image $f(x, y)$ is assumed corrupted by white noise with a noise level σ_n . Let W_λ be the noisy curvelet coefficients corresponding to $f(x, y)$. The code estimates the variances σ_λ of the coefficients W_λ , using knowledge of σ_n together with a Monte Carlo simulation that estimates the L^2 norms of individual curvelets. The denoised image curvelet coefficients \overline{W}_λ are then obtained by thresholding the noisy coefficients as follows

$$(5) \quad \overline{W}_\lambda = W_\lambda, \quad \text{if } |W_\lambda|/\sigma_n \geq k\sigma_\lambda, \quad \overline{W}_\lambda = 0, \quad \text{if } |W_\lambda|/\sigma_n < k\sigma_\lambda.$$

Here, $k = 3$, except for the finest scale, where $k = 4$. As in the *TV* case, curvelet denoising significantly reduces $\|\nabla f\|_1$, and need not preserve $\|f\|_1$.

The curvelet denoising experiments discussed in [3], [20], and [18], involve synthetically noised images, where both the noise level σ_n , as well as the type of noise, are known. In our experiments below on real HIM images, such knowledge is unavailable, and we must use educated guesses for the input σ_n .

5. Linear fractional diffusion equations and Lévy stable denoising. For fixed β with $0 < \beta \leq 1$, consider the linear fractional diffusion initial value problem in $L^2(\mathbb{R}^2)$,

$$(6) \quad w_t = -(-\Delta)^\beta w, \quad t > 0, \quad w(x, y, 0) = f(x, y),$$

where Δ denotes the 2D Laplacian. This reduces to the classical heat equation when $\beta = 1$. Eq.(6) has the unique Fourier space solution

$$(7) \quad \hat{w}(\xi, \eta, t) = \exp\{-t(\xi^2 + \eta^2)^\beta\} \hat{f}(\xi, \eta), \quad t > 0,$$

from which $w(x, y, t)$ can be found by inverse Fourier transformation

$$(8) \quad w(x, y, t) = \int_{\mathbb{R}^2} \exp\{2\pi i(\xi x + \eta y)\} \exp\{-t(\xi^2 + \eta^2)^\beta\} \hat{f}(\xi, \eta) d\xi d\eta.$$

In Eq. (7), the function

$$(9) \quad \hat{h}(\xi, \eta, t) = \exp\{-t(\xi^2 + \eta^2)^\beta\}, \quad t > 0,$$

is the Fourier transform of the Green's function for the linear fractional diffusion equation in Eq. (6). For each fixed $t > 0$, Eq. (9) is also the Fourier transform of an *isotropic Lévy stable probability density function* with exponent 2β , [11], [14], [19]. In physical (x, y) space, such probability densities are bell-shaped functions with increasingly heavy tails as β decreases from $\beta = 1$. The choice $\beta = 1$ corresponds to the Gaussian density, while $\beta = 1/2$ corresponds to the Lorentzian density. For other values of β , the corresponding density is not known in closed form in physical (x, y) space. As t decreases, these functions become steeper and narrower, approaching the 2D Dirac δ -function as $t \downarrow 0$. In the image deblurring applications discussed in [5], [7], [8], such Lévy stable laws play a vital role as candidate *point spread functions*. In the present application, denoising is accomplished by effectively blurring the noisy image with such narrow point spread functions.

We note that from Eq.(7), $\hat{w}(0, 0, t) = \hat{f}(0, 0)$, $t > 0$. Hence,

$$(10) \quad \|w(\cdot, t)\|_1 = \|f\|_1, \quad t > 0,$$

and the linear diffusion smoothing process in Eq.(6) conserves the image L^1 norm.

6. Monotonicity and the significance of low β values. In the Hilbert space $L^2(\mathbb{R}^2)$, the unique solution $w(x, y, t)$ in Eq. (8) satisfies

$$(11) \quad \|w(\cdot, t) - f\|_2^2 = \int_{\mathbb{R}^2} \{1 - \exp(-t(\xi^2 + \eta^2)^\beta)\}^2 |\hat{f}(\xi, \eta)|^2 d\xi d\eta.$$

Hence,

$$(12) \quad \|w(\cdot, t_2) - f\|_2 \geq \|w(\cdot, t_1) - f\|_2, \quad t_2 \geq t_1 \geq 0.$$

At the same time, from Eq. (3), $|\nabla w(x, y, t)|$ satisfies

$$(13) \quad \|\nabla w(\cdot, t)\|_2^2 = \int_{R^2} (\xi^2 + \eta^2) \exp(-2t(\xi^2 + \eta^2)^\beta) |\hat{f}(\xi, \eta)|^2 d\xi d\eta.$$

Hence,

$$(14) \quad \|\nabla w(\cdot, t_2)\|_2 \leq \|\nabla w(\cdot, t_1)\|_2, \quad t_2 \geq t_1 \geq 0.$$

If $|\nabla f| \in L^2(R^2)$, then $\|\nabla w(\cdot, t)\|_2 \leq \|\nabla f\|_2$, $t \geq 0$. If $|\nabla f| \notin L^2(R^2)$, then $\|\nabla w(\cdot, t)\|_2$ is finite for $t > 0$, but becomes infinite as $t \downarrow 0$. The rate at which this happens depends on β , as $\|\nabla w(\cdot, t)\|_2 = O(t^{-1/2\beta})$, $t \downarrow 0$. In fact, with $e = 2.71828\dots$, the following sharp inequality is a consequence of Eq. (13),

$$(15) \quad \|\nabla w(\cdot, t)\|_2 \leq \sup_{\rho \geq 0} \{\rho \exp(-t\rho^{2\beta})\} \|f\|_2 = \{2\beta t e\}^{-1/2\beta} \|f\|_2, \quad t > 0.$$

In the case of Gaussian smoothing, corresponding to $\beta = 1$, this inequality implies $\|\nabla w(\cdot, t)\|_2 = O(t^{-1/2})$, as $t \downarrow 0$. In contrast, $\|\nabla w(\cdot, t)\|_2 = O(t^{-5})$, as $t \downarrow 0$, when $\beta = 0.1$. This indicates that for small $t > 0$, the solution $w_\beta(x, y, t)$, with $\beta \ll 1$, retains considerably more of the small scale features in the initial data $f(x, y)$, than is the case with Gaussian smoothing. As a practical consequence, low exponent Lévy stable smoothing results in more moderate reductions in $\|\nabla f\|_1$ than is the case with TV or curvelet denoising, while conserving $\|f\|_1$.

7. Computational considerations. We deal exclusively with square images $g(x, y)$ of size $2N \times 2N$ pixels, and the fast Fourier transform (FFT) is the primary computational tool used in this paper. In order to render mathematical formulae more transparent, we use the same notation, $\hat{g}(\xi, \eta)$, for both discrete and continuous Fourier transforms. In the discrete FFT case, the frequencies ξ and η are understood to be integer-valued and to range from $-N$ to N . Likewise, $g(x, y)$ denotes both discrete and continuous images. In the discrete case, the variables x, y are measured in pixels and range from 1 to $2N$.

We may compute the solution $w(x, y, t)$ in Eq. (6) at any given $t > 0$, by using the forward and inverse FFT to implement the operations in Eq. (7) and Eq. (8) respectively. However, a more efficient recursive procedure can be used. With $\hat{h}(\xi, \eta, t)$ as in Eq. (9), let $\varpi = 1.0/K$ for a fixed positive integer K sufficiently large, and define $Q(\xi, \eta) = \hat{h}(\xi, \eta, \varpi)$. With $\hat{w}(\xi, \eta, 0) = \hat{f}(\xi, \eta)$, consider the following recursion

$$(16) \quad \hat{w}(\xi, \eta, k\varpi) = Q(\xi, \eta) \hat{w}(\xi, \eta, (k-1)\varpi), \quad k = 1, 2, 3, \dots$$

An inverse FFT in Eq.(16) at each integer k , generates the solution $w(x, y, t)$ at as many discrete times $t_k = k\varpi$, as desired. Diagnostic statistical information about $w(x, y, t)$ can also be calculated for selected values of t_k as t increases. Of particular interest are the discrete L^p norms, $p = 1, 2$, defined as follows

$$(17) \quad \|w_d(\cdot, t)\|_p = \left\{ (2N)^{-2} \sum_{x,y=1}^{2N} |w(x, y, t)|^p \right\}^{1/p}.$$

As in Eq.(10), we have $\|w_d(\cdot, t)\|_1 = \|w_d(\cdot, 0)\|_1$, $t > 0$. We also define the discrete analogs of $\|\nabla w(\cdot, t)\|_p$, $p = 1, 2$, as follows

$$(18) \quad \|\nabla_d w(\cdot, t)\|_p = \left\{ (2N)^{-2} \sum_{x,y=1}^{2N-1} (\{w_x(x, y, t)\}^2 + \{w_y(x, y, t)\}^2)^{p/2} \right\}^{1/p}.$$

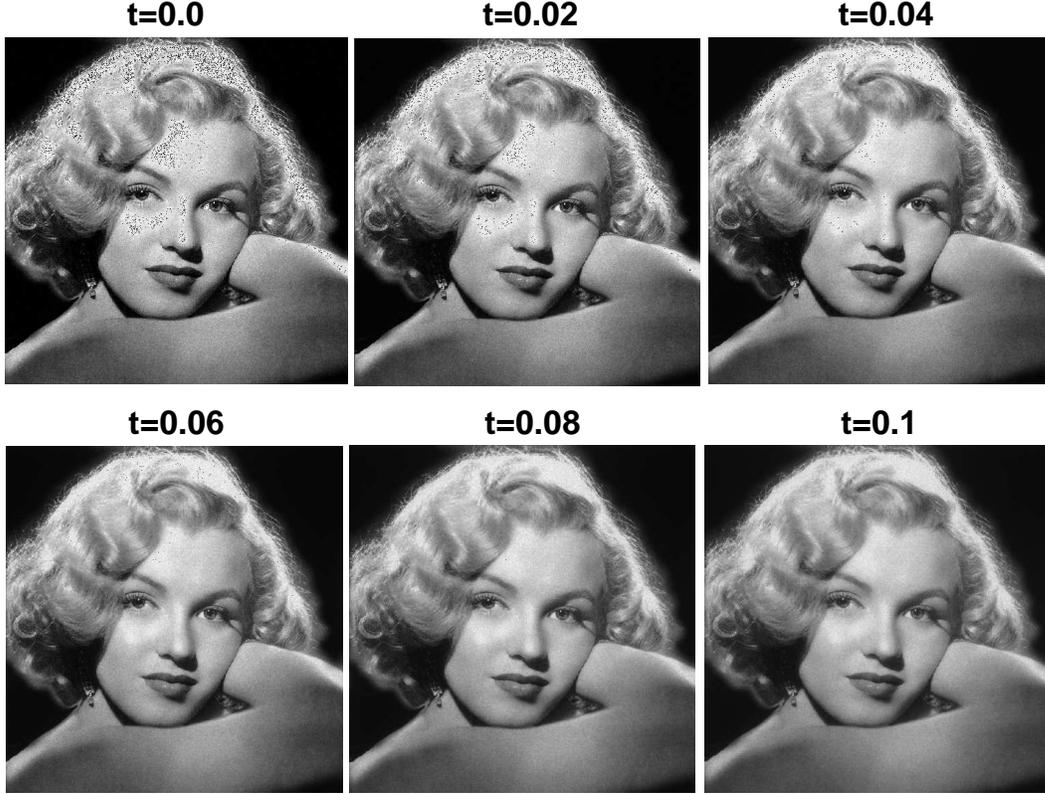


FIG. 1. Poisson noised Marilyn Monroe image $f(x, y)$ at $t = 0$, is progressively denoised using the recursion in Eq.(16) to solve Eq.(20) from $t = 0$ to $t = 0.1$, with $\beta = 0.2$. In that time interval, for $p = 1, 2$, $\|\nabla_d w(\cdot, t)\|_p$, decreases monotonically by a factor $\lambda \approx 0.46$, while $\|w_d(\cdot, t)\|_1$ is conserved.

where

$$(19) \quad \begin{aligned} w_x(x, y, t) &= (2N)^{-1} (w(x+1, y, t) - w(x, y, t)), \\ w_y(x, y, t) &= (2N)^{-1} (u(x, y+1, t) - u(x, y, t)). \end{aligned}$$

Note that for a mathematical image $f(x, y) \in L^p(R^2)$ as considered in section 2, we may have $\|\nabla f\|_p = \infty$. However, the corresponding $2N \times 2N$ pixels image $f(x, y)$ will have a finite value for $\|\nabla_d f\|_p$, although that value will be relatively large.

8. Minimum principle for ‘slow motion’ Lévy stable denoising. Given a noisy image $f(x, y)$, calculate $\|\nabla_d f\|_p$, $p = 1, 2$. Fix β , with $0 < \beta < 1/2$, and consider the evolution problem

$$(20) \quad w_t = -(-\Delta)^\beta w, \quad t > 0, \quad w(x, y, 0) = f(x, y).$$

As indicated in the previous section, using the recursion in Eq.(16) and FFT algorithms, we may readily solve Eq. (20) and calculate $\|\nabla_d w(\cdot, t)\|_p$, $p = 1, 2$, for selected t values as t increases. Analogously to the minimum principle in Eq. (4), we pose the following minimization problem. For given fixed λ with $0 < \lambda < 1$,

$$(21) \quad f^L(x, y) = \text{Arg} \min_{t>0} \{ \|w(\cdot, t) - f\|_2 \ni \|\nabla_d w(\cdot, t)\|_2 \leq \lambda \|\nabla_d f\|_2 \}.$$

Estimating the Lipschitz exponent in Sydney image

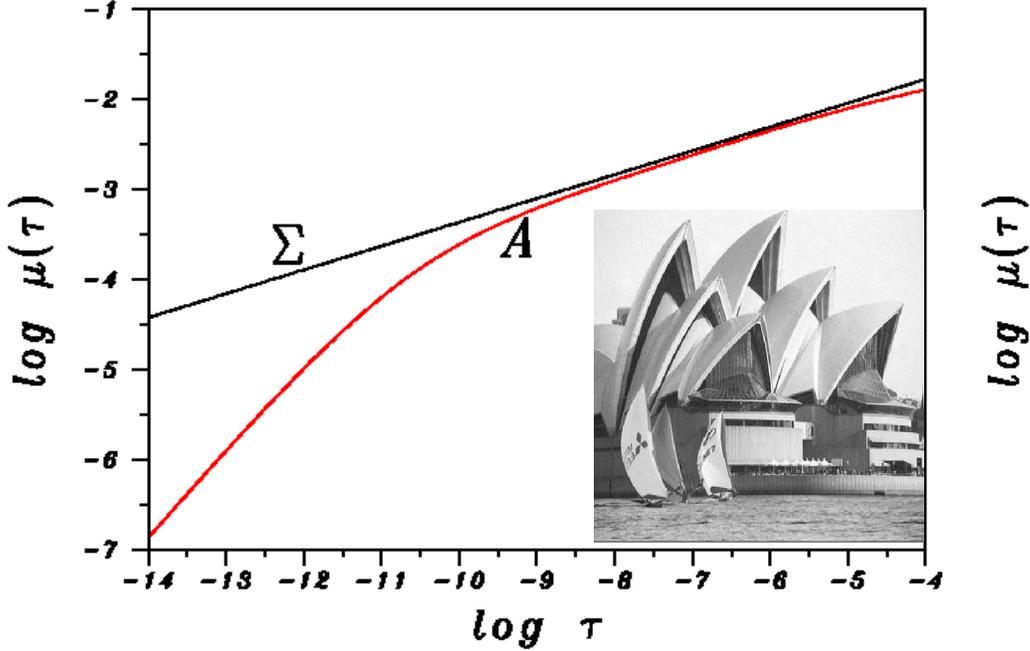
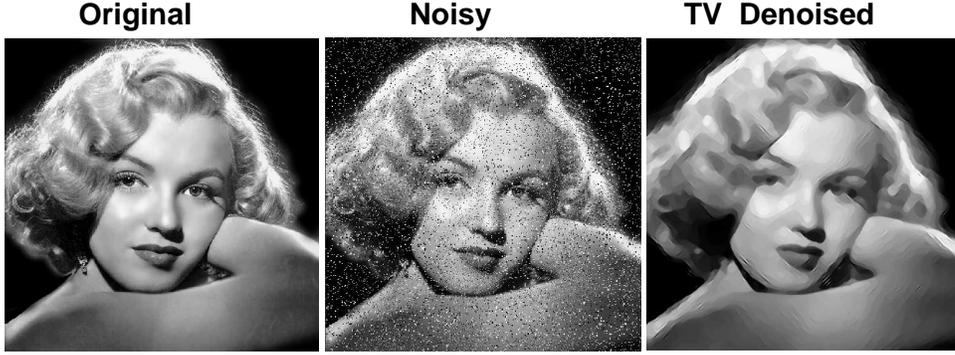


FIG. 2. With $\mu(\tau)$ as in Eq.(24), red curve A is a plot of $\mu(\tau)$ versus τ , on a log-log scale. Image Lipschitz exponent equals twice the slope of majorizing line Σ . That slope is 0.265, indicating that the Sydney image has L^1 Lipschitz exponent $\alpha = 0.530$.

In view of the monotonicity results in Eq.(12) and Eq. (13), this minimum principle has the unique solution $f^L(x, y) = w(x, y, t^\dagger)$, where $t^\dagger > 0$ is the earliest time at which $\|\nabla_d w(\cdot, t)\|_2 \leq \lambda \|\nabla_d f\|_2$. As shown in Figure 1, one can monitor this denoising process as t increases from $t = 0$ to $t = t^\dagger$, by displaying the image evolution, and evaluating the accompanying diagnostic information, $\|w^d(\cdot, t)\|_p$ and $\|\nabla_d w(\cdot, t)\|_p$, $p = 1, 2$. In Figure 1, a Poisson noised 512×512 pixel Marilyn Monroe image is used as initial data in Eq. (20) with $\beta = 0.2$. Initially, $\|\nabla_d f\|_1 = 11,000$, $\|\nabla_d f\|_2 = 15,000$. These gradients decay monotonically to the values 5000 and 6900, respectively, at $t = 0.1$, while $\|w_d(\cdot, t)\|_1$ is conserved. Thus, $t = 0.1$ would correspond to t^\dagger , had λ been chosen to be ≈ 0.46 . Such displays enable the user to decide whether important small-scale information has been smoothed out, along with the noise, prior to reaching t^\dagger . The process can then be restarted with different values of β and λ . These are valuable exploratory options in practice. Such options are unavailable in the Total Variation method which, given the regularization parameter ω , produces a single final denoised image, defined as the limit of the convergent Split Bregman iteration in [12], or as the steady-state solution in the nonlinear anisotropic diffusion problem in [15]. Likewise, given the input noise level σ_n , curvelet denoising results in a single final denoised image.



Lipschitz exponents after noising and denoising

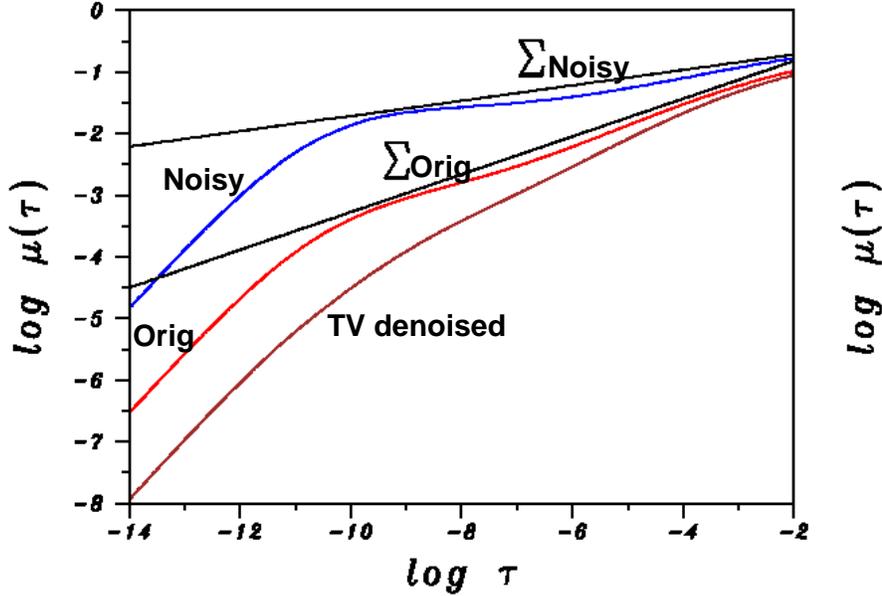


FIG. 3. Noising an image decreases the Lipschitz exponent, while some forms of denoising can eliminate texture along with the noise, and increase the Lipschitz exponent. Slopes of lines Σ_{Noisy} and Σ_{Orig} are, respectively, 0.130 and 0.297, indicating that the salt and pepper noisy image has a substantially smaller Lipschitz exponent, 0.260, than does the original noiseless image, 0.594. The corresponding Σ line for the TV denoised trace, (brown curve), was not drawn to avoid clutter. That line has slope 0.406, indicating a Lipschitz exponent of 0.812, substantially higher than the original noiseless image. TV denoising often removes important texture, and leads to oversmoothed images.

9. Image fine structure and Lipschitz exponents. Most natural images $f(x, y)$ are not smoothly differentiable functions of x and y , but display edges, localized sharp features, and other significant fine scale details or *texture*. The image Lipschitz exponent measures the fine structure content of an image, provided that image is relatively noise free. An image $f(x, y)$ has L^1 Lipschitz exponent α , if and only if

$$(22) \quad \int_{\mathbb{R}^2} |f(x + h_1, y + h_2) - f(x, y)| dx dy \leq Const |h|^\alpha, \quad |h| \rightarrow 0,$$

where $|h| = (h_1^2 + h_2^2)^{1/2}$, and α is fixed with $0 < \alpha \leq 1$. An image that is of bounded variation, or smoother, has $\alpha = 1$. The value of α decreases with increasing fine structure. Most natural images have $\alpha < 0.6$, and are not of bounded variation.

In [6], [9], an effective method for estimating image Lipschitz exponents is developed, based on a major theorem in [22]. See also [2], [21]. For fixed $\tau > 0$, define the linear operator G^τ by means of the Fourier series

$$(23) \quad \{G^\tau f\}(x, y) = \sum_{m, n=-\infty}^{\infty} \exp\{-\tau(m^2 + n^2)\} \hat{f}_{mn} \exp\{2\pi i(xm + yn)\},$$

where \hat{f}_{mn} are the Fourier coefficients of the image $f(x, y)$, the latter assumed defined on the unit square. Let $\mu(\tau)$ be the L^1 relative error in approximating $f(x, y)$ with the Fourier series $\{G^\tau f\}(x, y)$,

$$(24) \quad \mu(\tau) = \|G^\tau f - f\|_1 / \|f\|_1, \quad \tau > 0.$$

As shown in [22], an image $f(x, y)$ has Lipschitz exponent α if and only if $\mu(\tau) = O(\tau^{\alpha/2})$ as $\tau \downarrow 0$. Because of the exponential decay, the infinite series in Eq. (23) can be well-approximated by a finite sum for each fixed τ . Such a sum can be formed using FFT algorithms, and $G^\tau f$ can be evaluated for each fixed $\tau_n > 0$ in a sequence $\{\tau_n\}$ tending to zero, together with $\mu(\tau_n)$. By plotting $\mu(\tau_n)$ vs τ_n on a log-log scale, positive constants C and α can be located such that $\mu(\tau) \leq C \tau^{\alpha/2}$ as $\tau \downarrow 0$.

Figure 2 describes this Lipschitz estimation procedure as applied to a 512×512 pixels Sydney image $f(x, y)$. The above FFT procedure was used to obtain $\mu(\tau_n)$ in Eq. (24) at 400 values $\tau_n = 0.5(0.95)^n$, $n = 1, 400$. A plot of $\mu(\tau)$ versus τ on a log-log scale produced the red curve A in Figure 2. The curve A exhibits a characteristic *elbow shape*. It consists of a straight line segment with slope ≈ 1 , beginning near $\log \tau = -15$ and continuing to near $\log \tau = -10$. There is then a transition to a different regime, one that is more slowly increasing and that continues to near $\log \tau = 0$. As explained more fully in [6], [9], the rapidly varying portion for $\log \tau < -10$ is a fallacious finite-dimensionality artifact, unrelated to the true image Lipschitz exponent. Only the slowly varying part of A is relevant to estimating Lipschitz exponents. Least squares fitting on $-9 \leq \log \tau \leq -4$ was used to find the majorizing straight line Σ for the slowly varying part of the red curve A . The y -axis intercept value obtained by least squares was slightly increased so as to make the line Σ lie visibly above the red curve A . However, the slope of Σ remains the same as that obtained from least squares. The line Σ is defined by $\log \mu(\tau) = -0.902 + 0.265 \log \tau$, implying that $\mu(\tau) \leq 0.493 \tau^{0.265}$ as $\tau \downarrow 0$. According to the theorem in [22], the Sydney image has Lipschitz exponent $\alpha = 0.530$.

10. Noise contamination, smoothing, and image Lipschitz exponents.

The behavior of image Lipschitz exponents α before and after processing, is of major interest. As shown in Figure 3, an original sharp, noiseless, 512×512 pixel Marilyn Monroe image $f(x, y)$, has $\alpha = 0.594$ and $\|\nabla_d f\|_1 = 5600$. This is synthetically contaminated by adding salt and pepper noise with density 0.1. Such noise artificially decreases α to the value $\alpha = 0.260$, while simultaneously increasing $\|\nabla_d f\|_1$ to the value 22000. As previously noted, the Lipschitz exponent is a true measure of image fine structure, only if that image is relatively noise free. A variety of algorithms may now be applied to denoise the noisy Marilyn Monroe image, with varying degrees of success. Some algorithms may smooth out genuine fine details along with the noise, decreasing $\|\nabla_d f\|_1$ and simultaneously increasing α , often well beyond their true

values in the original noiseless image. This is the case with the *TV* denoised image in Figure 3, which has $\|\nabla_{df}\|_1 = 2010$, and $\alpha = 0.812$. Here, a procedure different from the split Bregman method was used for *TV* denoising. This is the previously mentioned approach in [15], where the denoised image corresponds to the steady-state solution in a nonlinear anisotropic diffusion equation. The finite difference scheme in [15] was used with $\Delta t = 0.1(\Delta x)^2$, regularization parameter $\Lambda = 2.0$, and forward integration carried for 300 time steps Δt . In contrast, denoising the Poisson noised Marilyn Monroe image in Figure 1, using Lévy stable fractional diffusion, results in $\|\nabla_{df}\|_1 = 5000$, which is close to the true value in the sharp noiseless image. These observations will be helpful in the next section.

11. Denoising state of the art HIM imagery. The Helium Ion Microscope (HIM) is a new type of microscope that works by scanning a well-focused He ion beam over the surface of the sample in a raster pattern. The technique is very similar to the method used in scanning electron microscopes (SEM). As in the SEM case, the most significant signal is produced by the secondary electrons (SE), especially the SE1 electrons that are generated right at the point where the charged beam hits the sample. These electrons carry information about the finest morphological details of the sample. In the HIM case, more of these electrons are produced than in the SEM case, with a higher proportion of SE1 and other SE signals. Consequently, the resulting secondary electron images are richer in surface details. The price for this is a generally somewhat worse signal-to-noise ratio. The He beam current is typically smaller, and even the use of sub-pA current is feasible. With certain samples it is recommended to use low beam currents to avoid the erosion of the samples.

The resulting higher noise levels hamper the extraction and interpretation of needed morphological information from HIM images. The resolution of HIMs and state-of-the-art SEMs is better than 1 nm, and in some cases, it is under 0.5 nm. While efforts to further improve resolution are currently focused on new hardware designs, it is important to explore and develop software-based solutions as well. One solution lies in denoising HIM and SEM images, which when done well, can substantially improve image quality. High-frequency components of the signal carry the information about the fine details, sharp edges, and fast transition in grey levels. Imaging instruments typically have a transfer function that shows worsening signal-to-noise ratio for fine details, because the noise is commonly worse in the high-frequency range. Simple denoising methods that merely delete the high-frequency portion of the signal are not helpful. Methods that introduce unacceptably large distortions, or lead to significant blurring, are not very valuable either, especially for measurement purposes. Many of the currently available denoising methods fall into this category. However, the split Bregman total variation minimization method, and the Curvelet Thresholding method, are two highly sophisticated denoising methods that have undergone intensive development over the last ten years, and have found useful application in many areas, [12], [3], [20]. Accordingly, it is appropriate to evaluate the fractional diffusion method on HIM imagery by comparing it with these two well-esteemed approaches.

The three HIM sample images discussed below involve signal-dependent noise of unknown characteristics and intensity. Educated guesses must be used for the input regularization parameters. After some preliminary experimentation, a value of $\omega = 0.025$ in Eq.(4) was selected for the split Bregman *TV* method, and a value $\sigma_n = 30$ in Eq.(5) was adopted for the Curvelet Thresholding method. These values were used for all three samples, and they are well within the ranges recommended by the authors in [12], [3], [20], in the MATLAB implementations of their methods. In

the fractional diffusion method, we used $\beta = 0.2$ in Eq.(20), and selected the solution at $t = 0.1$ as the denoised image in all three samples. This turned out to be equivalent to having chosen $\lambda \approx 0.33$ in Eq.(21), so that $t^\dagger = 0.1$.

The following discussion presupposes access to the on-line version of this journal, together with a high-resolution computer screen. Figure 4 deals with an actual HIM image of an Au-decorated gold on carbon sample, with a field of view of 600 nm. In Figure 4, the Lévy denoised image has maintained fidelity to the surface texture in the foreground, as is evident from the jagged edges, as well as to the small structures in the background. Such background structures are not well-recovered in the *TV* denoised image, while the foreground surface texture has been smoothed. The background structures are better recovered in the Curvelet denoised image; however, the jagged edges in the foreground surface texture have now been eliminated. This becomes more evident when examining a portion of the sample, as shown in Figure 5. In Table 1, we see that Lévy stable denoising conserves $\|f_d\|_1$, while reducing $\|\nabla_d f\|_1$ by a factor of three. However, *TV* and Curvelet denoising do not conserve $\|f_d\|_1$, and reduce $\|\nabla_d f\|_1$, by a factor of seven or more.

The behavior of the L^1 Lipschitz exponent α in Figure 6 provides a useful metric in the denoising experiment in Figures 4 and 5. The original noisy HIM image has $\alpha = 0.236$, while the Lévy stable image has $\alpha = 0.462$. The *TV* and Curvelet images have substantially higher values of α , namely 0.778 and 0.845, respectively. This is consistent with the behavior of $\|\nabla_d f\|_1$ in Table 1.

A very different sample image is considered in Figure 7 but the denoising results show a similar pattern. Here, a salt crystal on radiolaria HIM image, with a field of view of 90 μm is denoised. By examining a portion of the image in Figure 8, it is evident that surface morphology is better reproduced in the Lévy stable image than it is in the *TV* and Curvelet images. The Lipschitz traces in this experiment (not shown) are very similar to those in Figure 6. The original noisy HIM image has $\alpha = 0.288$, while the Lévy stable image has $\alpha = 0.492$. The *TV* and Curvelet images have higher values of α , namely 0.666 and 0.674, respectively. Again, Table 2 shows quite significant reductions in $\|\nabla_d f\|_1$ in the *TV* and Curvelet images, as compared with the Lévy stable image. In this example, the 30% drop in $\|f_d\|_1$ after split Bregman *TV* processing, is striking.

The last example, in Figures 9 and 10, is instructive. This is a gold on carbon sample, with a field of view of 300 nm. The original HIM image appears very noisy, and $\|\nabla_d f\|_1 = 47000$. Lévy stable denoising removes a considerable amount of noise, with $\|\nabla_d f\|_1$ reduced to 15000, but the resulting surfaces are still fuzzy. However, such surface fuzziness may be characteristic of the sample, much like the surface of a peach. The *TV* and Curvelet images exhibit aggressive denoising, resulting in $\|\nabla_d f\|_1 \leq 3500$, and surfaces as smooth as the surface of an apple. The Lipschitz traces in this experiment (not shown), are again very similar to those in Figure 6. The original noisy HIM image has $\alpha = 0.085$, while the Lévy stable image has $\alpha = 0.211$. The *TV* and Curvelet images have significantly higher values of α , namely 0.696 and 0.704, respectively. In this example, Lévy stable denoising provides the microscopist with a more prudent reconstruction, to be considered alongside the *TV* and Curvelet images.

12. Concluding remarks. Lévy stable denoising exploits the fact that in the evolution equation $w_t = -(-\Delta)^\beta w$, $t > 0$, $w(x, y, 0) = f(x, y)$, the solution $w(x, y, t)$ satisfies $\|\nabla w(\cdot, t)\|_2 = O(t^{-1/2\beta})$, $t \downarrow 0$. Thus, if β is chosen small, and $f(x, y)$ is not smooth, $\|\nabla w(\cdot, t)\|_2$ blows up much more rapidly as $t \downarrow 0$, than would be the case

with Gaussian smoothing ($\beta = 1$). This suggests that for such small β , the solution $w_\beta(x, y, t)$ retains a great deal of the fine structure in the initial data $f(x, y)$, at small values of $t > 0$. This expectation is borne out in practice. When this method is applied to noisy HIM imagery, it is found that a considerable portion of the noise is removed, without significantly altering the original information content of the image. In Figures 4, 5, 7, 8, 9 and 10, jagged edges in the surface texture, and several other aspects in the original surface morphology, are preserved in the Lévy stable image. This implies fidelity to high frequency information. An important diagnostic is provided by the value of $\|\nabla_{df}\|_1$, which was typically reduced by a factor of 3 in the Lévy images.

Surprisingly, such fidelity to the original surface morphology was not feasible with the split Bregman *TV* method, or with the Curvelet Thresholding method. The *TV* method succeeds in its aim at minimizing $\|\nabla_{df}\|_1$ in the denoised image. In Figures 4 and 9, $\|\nabla_{df}\|_1$ is reduced by factors of 7 and 13, respectively, after split Bregman *TV* denoising. Unexpectedly, Curvelet denoising produces even more severe reductions in $\|\nabla_{df}\|_1$. Clearly, significant fine structure information has been removed, along with the noise, in the *TV* and Curvelet images. This is confirmed by the sizeable increases in image Lipschitz exponents.

A very significant advantage of the fractional diffusion method, is the *slow motion* option which enables monitoring of the denoising process, and the possibility of backtracking to a more optimal image. With unknown noise characteristics, it is difficult to guess optimal choices for the regularization parameters in the split Bregman *TV* and Curvelet Thresholding methods, and equally difficult to decide whether a better denoised image may even be feasible. However, preliminary denoising experiments using the fast Lévy stable method, may provide useful guidance in the choice of regularization parameters in the *TV* or Curvelet method.

REFERENCES

- [1] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing*, Springer, New York, 2002.
- [2] P. L. Butzer and H. Berens, *Semi-Groups of Operators and Approximation*, Springer-Verlag, New York, 1967.
- [3] E. J. Candès, L. Demanet, D.L. Donoho and L. Ying, "Fast discrete curvelet transforms", *SIAM Multiscale Model. Simul.* 5 (2006), pp. 861–899. See also <http://www.curvelet.org>
- [4] A. S. Carasso, "Image restoration and diffusion processes," *Proc. SPIE 2035*, 255–266, *Mathematical Methods in Medical Imaging II*, Joseph N. Wilson; David C. Wilson; Eds. (1993).
- [5] A. S. Carasso, "The APEX method in image sharpening and the use of low exponent Lévy stable laws", *SIAM J. Appl. Math.* 63, 593–618 (2002).
- [6] A. S. Carasso, "Singular integrals, image smoothness, and the recovery of texture in image deblurring," *SIAM J. Appl. Math.* 64, 1749–1774 (2004).
- [7] 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, 2499-2514 (2002).
- [8] A. S. Carasso, "APEX blind deconvolution of color Hubble space telescope imagery and other astronomical data", *Optical Engineering* 45, 107004 (2006).
- [9] A. S. Carasso and A. E. Vladár, "Calibrating image roughness by estimating Lipschitz exponents, with applications to image restoration," *Optical Engineering* 47, 037012 (2008).
- [10] J. Dahl, P. C. Hansen, S. H. Jensen, and T. L. Jensen, "Algorithms and software for total variation image reconstruction via first-order methods," *Numer Algor* 53, 67–92 (2010).
- [11] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. 2, Second ed., Wiley, New York, 1971.
- [12] T. Goldstein and S. Osher, "The split Bregman method for L^1 regularized problems," *SIAM J. Imaging Sci.* 2, 323–343 (2009). See also <http://www.math.ucla.edu/tagoldst/code.html>
- [13] Y. Gousseau and J. M. Morel, "Are natural images of bounded variation ?" *SIAM J. Math. Anal.* 33, 634–648 (2001).

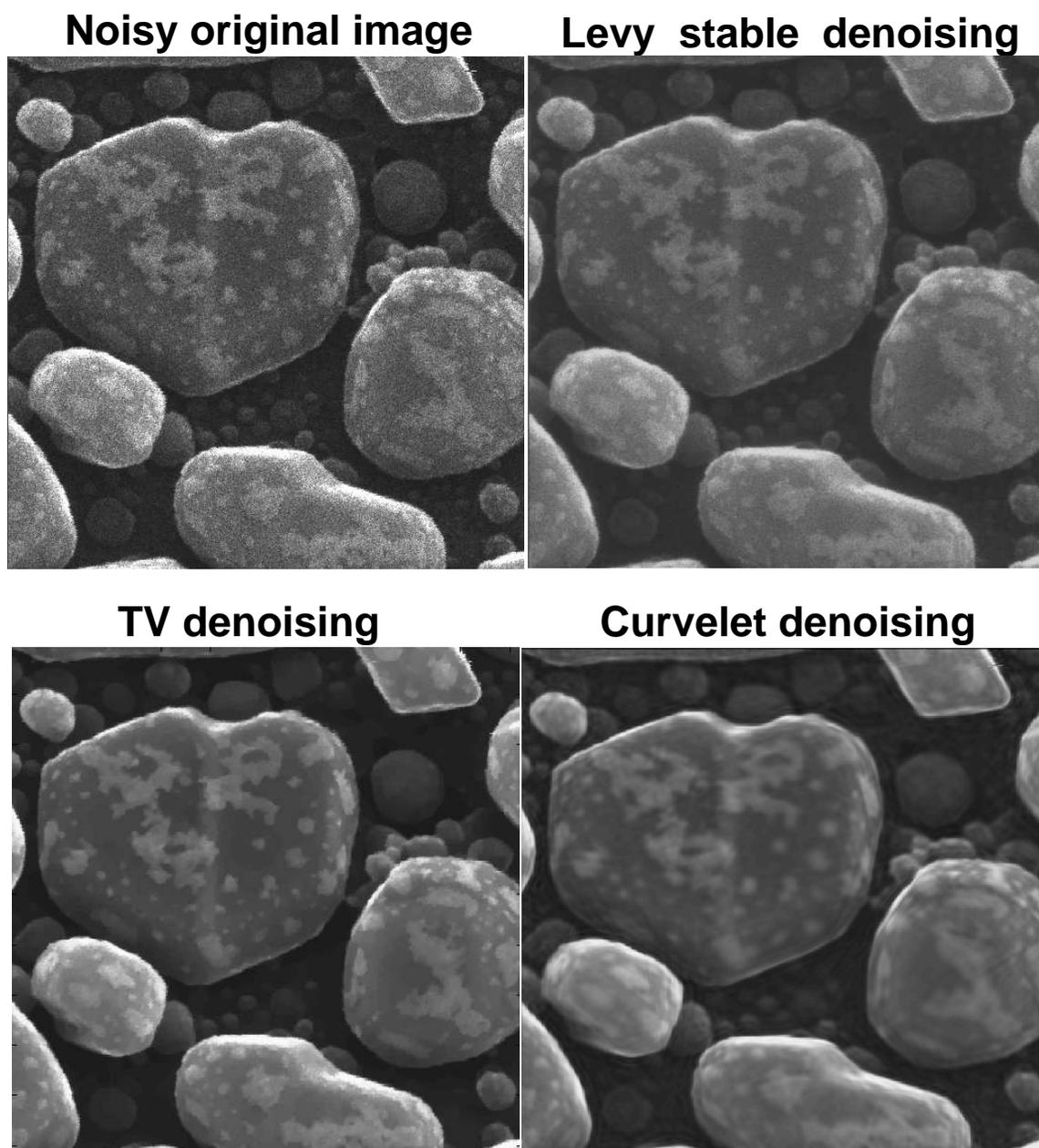


FIG. 4. Denoising of actual HIM Au-decorated gold on carbon sample, using Lévy fractional diffusion, split Bregman TV, and Curvelet thresholding. Here, field of view is 600 nm. Important loss of structural detail is evident in TV and Curvelet images. This is confirmed by examining a portion of the sample as shown in Figure 5, as well as the behavior of $\|\nabla_d f\|_1$ in Table 1, and the Lipschitz traces in Figure 6. The original noisy image has Lipschitz exponent $\alpha = 0.236$; the Lévy stable image has $\alpha = 0.462$; the split Bregman TV image has $\alpha = 0.778$; and the Curvelet image has $\alpha = 0.845$.

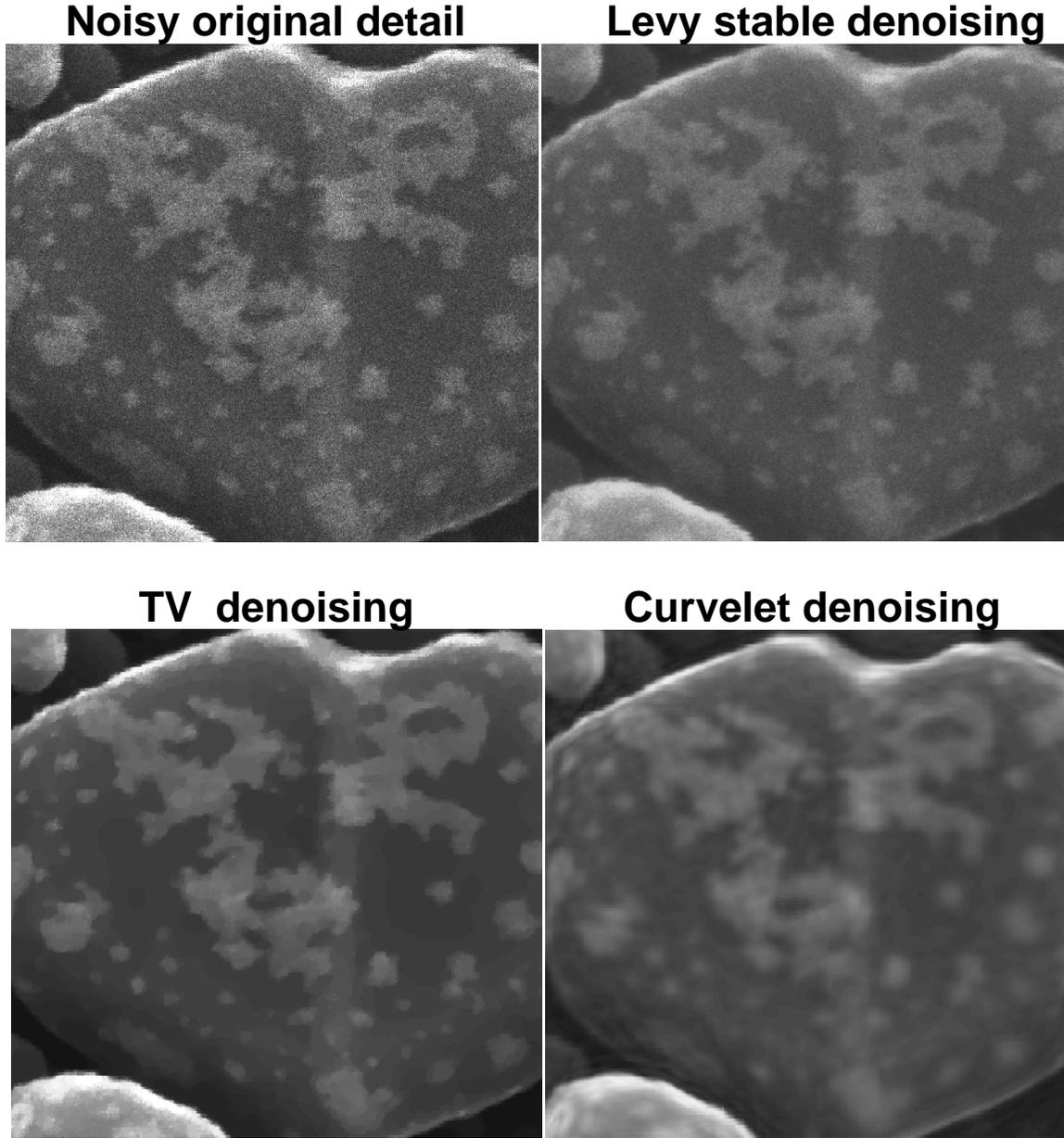


FIG. 5. Poor recovery of surface morphology in TV and Curvelet denoising experiments on Au-decorated gold on carbon sample in Figure 4. Here, field of view is 300 nm. This behavior is compatible with Table 1 below, and the Lipschitz traces in Figure 6.

TABLE 1.

Behavior of $\|f_d\|_1$, $\|f_d\|_2$, and $\|\nabla_d f\|_1$, in Figure 4 denoising.

Note severe $\|\nabla_d f\|_1$ reduction in Curvelet and TV denoising.

Image $f(x, y)$	$\ f_d\ _1$	$\ f_d\ _2$	$\ \nabla_d f\ _1$
Noisy original	88	99	25000
Lévy stable ($\beta = 0.2$, $t^\dagger = 0.1$)	88	94	8500
Split Bregman TV ($\omega = 0.025$)	74	89	3400
Curvelet thresholding ($\sigma_n = 30$)	81	91	2700

Behavior of Lipschitz exponents in Figure 4

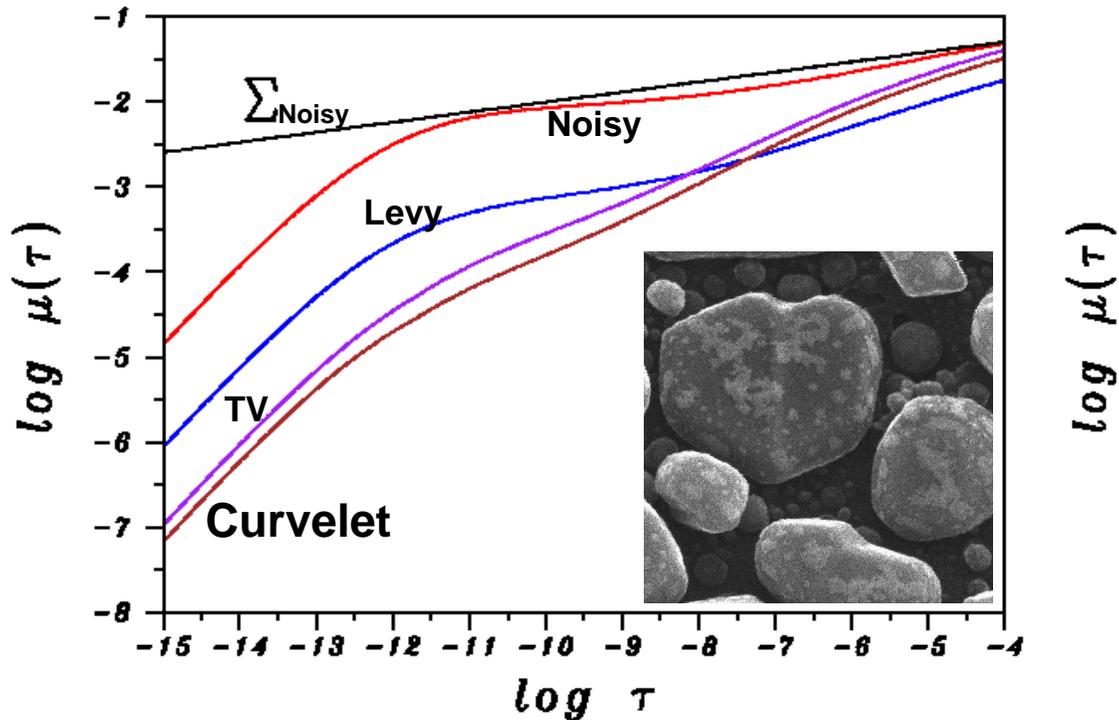


FIG. 6. Behavior of L^1 image Lipschitz exponents in denoising experiment on Au-decorated gold on carbon sample in Figures 4 and 5. Here, only majorizing Σ line for original trace is shown. Remaining Σ lines are not drawn to avoid clutter. Red curve is original noisy trace with Lipschitz exponent $\alpha = 0.236$. Blue curve is Lévy denoised trace with exponent $\alpha = 0.462$. Purple curve corresponds to split Bregman TV trace, with substantially larger exponent $\alpha = 0.778$. Brown curve is Curvelet trace with exponent $\alpha = 0.845$. These values quantify the loss of fine-structure evident in Figure 4 and 5.

- [14] E. Lukacs, *Characteristic Functions*, Second ed., Griffin, London, 1970.
- [15] 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, 387–405 (2000).
- [16] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12, 629–639 (1990).
- [17] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D.* 60, 259–268 (1992).
- [18] R. Sivakumar, "Denoising of computer tomography images using curvelet transform", *ARPN Journal of Engineering and Applied Sciences* 2, 21–26 (2007).
- [19] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Chapman and Hall, New York, 1994.
- [20] J. L. Starck, E. J. Candès and D. L. Donoho. "The curvelet transform for image denoising", *IEEE Transactions on Image Processing*, 11 670–684 (2002).
- [21] A. M. Stokolos and W. Trebels, "On the rate of almost everywhere convergence of Abel-Cartwright means on $L^p(\mathbb{R}^n)$ ", *Result. Math.* 34, 373–380 (1998).
- [22] M. H. Taibleson, "On the theory of Lipschitz spaces of distributions on Euclidean n-space. I.

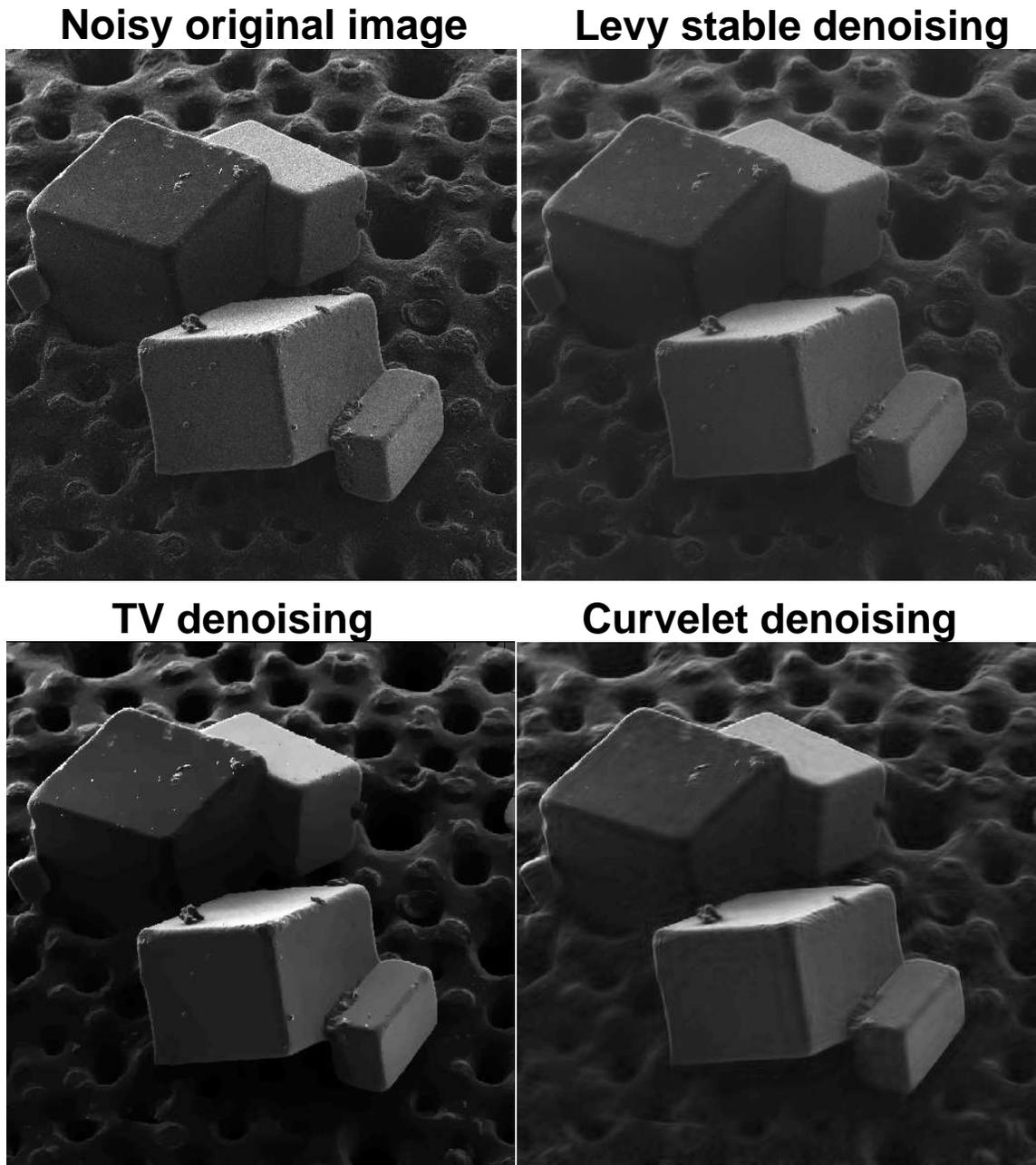


FIG. 7. Denoising of actual HIM salt crystals on radiolaria sample, using Lévy fractional diffusion, split Bregman TV, and Curvelet thresholding. Here field of view is $90 \mu\text{m}$. Loss of structural detail is evident in TV and Curvelet images. This is confirmed by examining a portion of the sample as shown in Figure 8, as well as the behavior of $\|\nabla_d f\|_1$ in Table 2. Note 30% drop in $\|f_d\|_1$, after split Bregman TV processing, in Table 2. Lipschitz traces for this experiment (not shown), are very similar those in Figure 6. Original noisy image has Lipschitz exponent $\alpha = 0.288$. Lévy denoised image has $\alpha = 0.492$. Split Bregman TV denoised image has $\alpha = 0.666$. Curvelet denoised image has $\alpha = 0.674$.

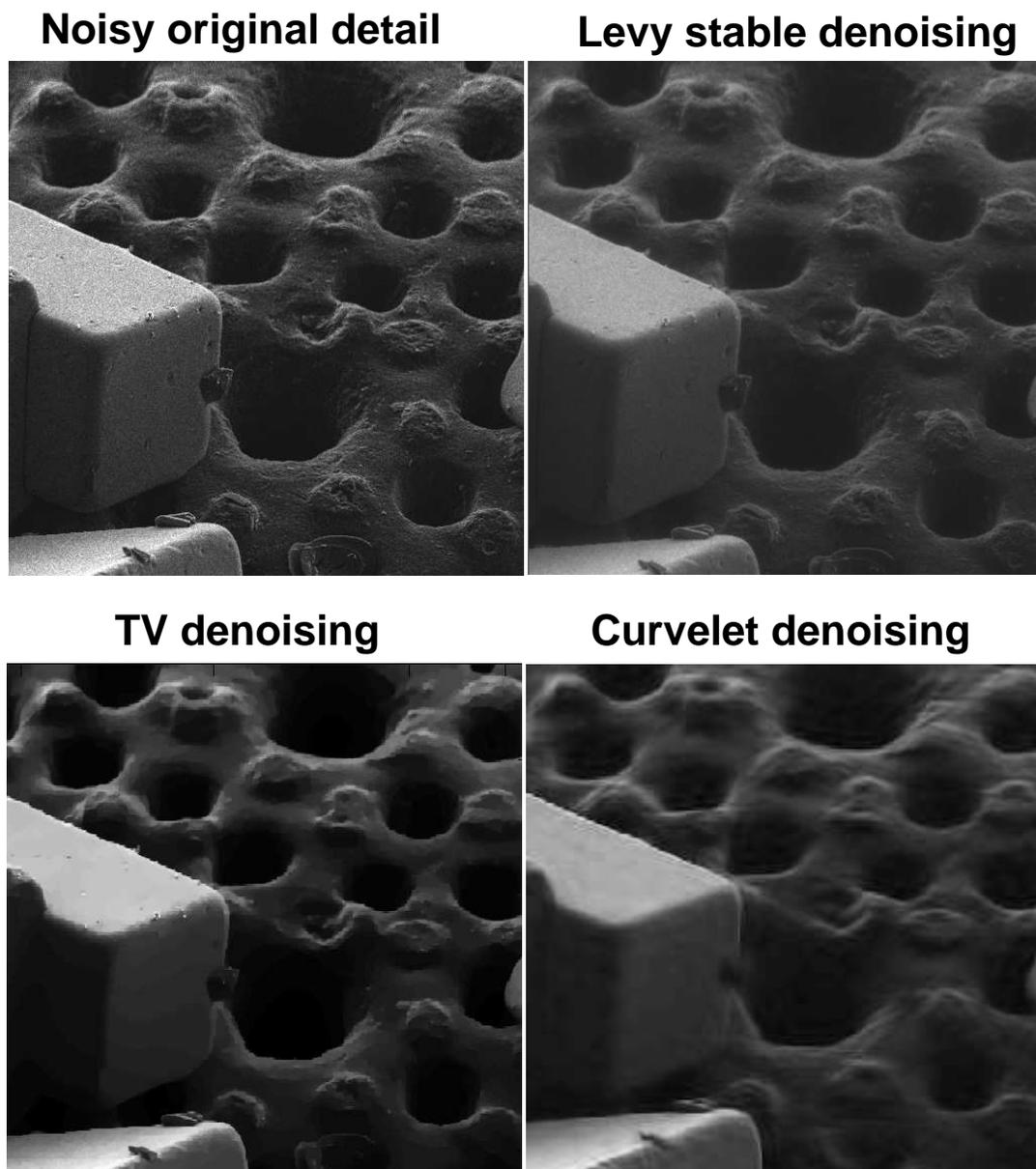


FIG. 8. Examination of parts of the salt crystals on radiolaria sample in Figure 7, under magnification, reveals significant erosion of surface morphology in TV and Curvelet images, as compared with Lévy stable image. Here, field of view is $45 \mu\text{m}$. Image Lipschitz exponents can quantify the loss of fine structure.

TABLE 2.

Behavior of $\|f_d\|_1$, $\|f_d\|_2$, and $\|\nabla_d f\|_1$, in Figure 7 denoising.
 Note severe $\|\nabla_d f\|_1$ reduction in Curvelet and TV denoising.

Image $f(x, y)$	$\ f_d\ _1$	$\ f_d\ _2$	$\ \nabla_d f\ _1$
Noisy original	54	65	14000
Lévy stable ($\beta = 0.2$, $t^\dagger = 0.1$)	54	61	4900
Split Bregman TV ($\omega = 0.025$)	37	57	3000
Curvelet thresholding ($\sigma_n = 30$)	46	56	2600

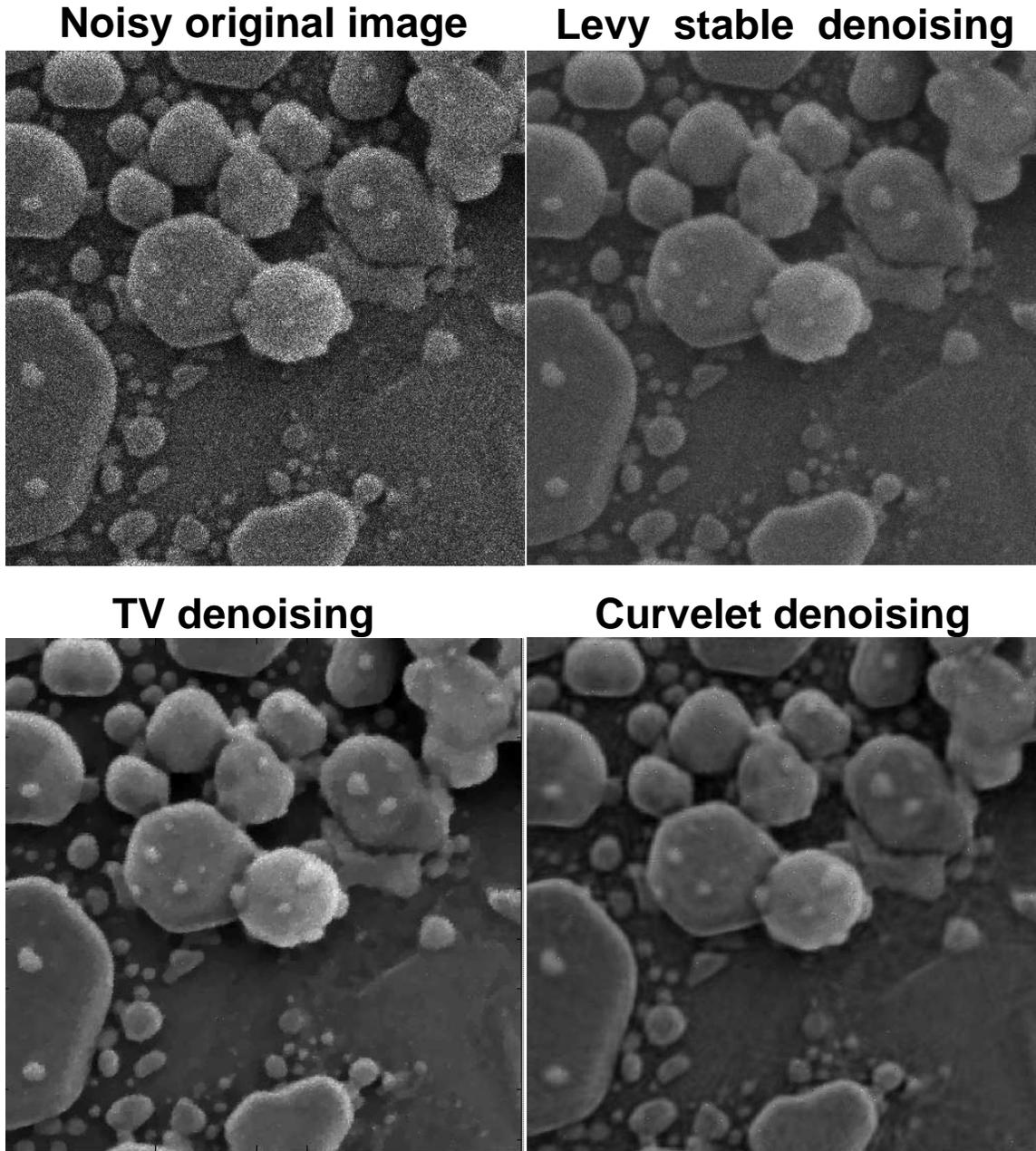


FIG. 9. Denoising of actual HIM gold on carbon sample, using Lévy fractional diffusion, split Bregman TV, and Curvelet thresholding. Here, field of view is 300 nm. TV and Curvelet denoising produce smoother surfaces than Lévy stable denoising. However, true surfaces may not be smooth. Lipschitz traces for this experiment (not shown), are again very similar to those in Figure 6. Here, the original noisy image has $\alpha = 0.085$; the Lévy stable image has $\alpha = 0.211$; the split Bregman TV image has $\alpha = 0.696$; and the Curvelet image has $\alpha = 0.704$.

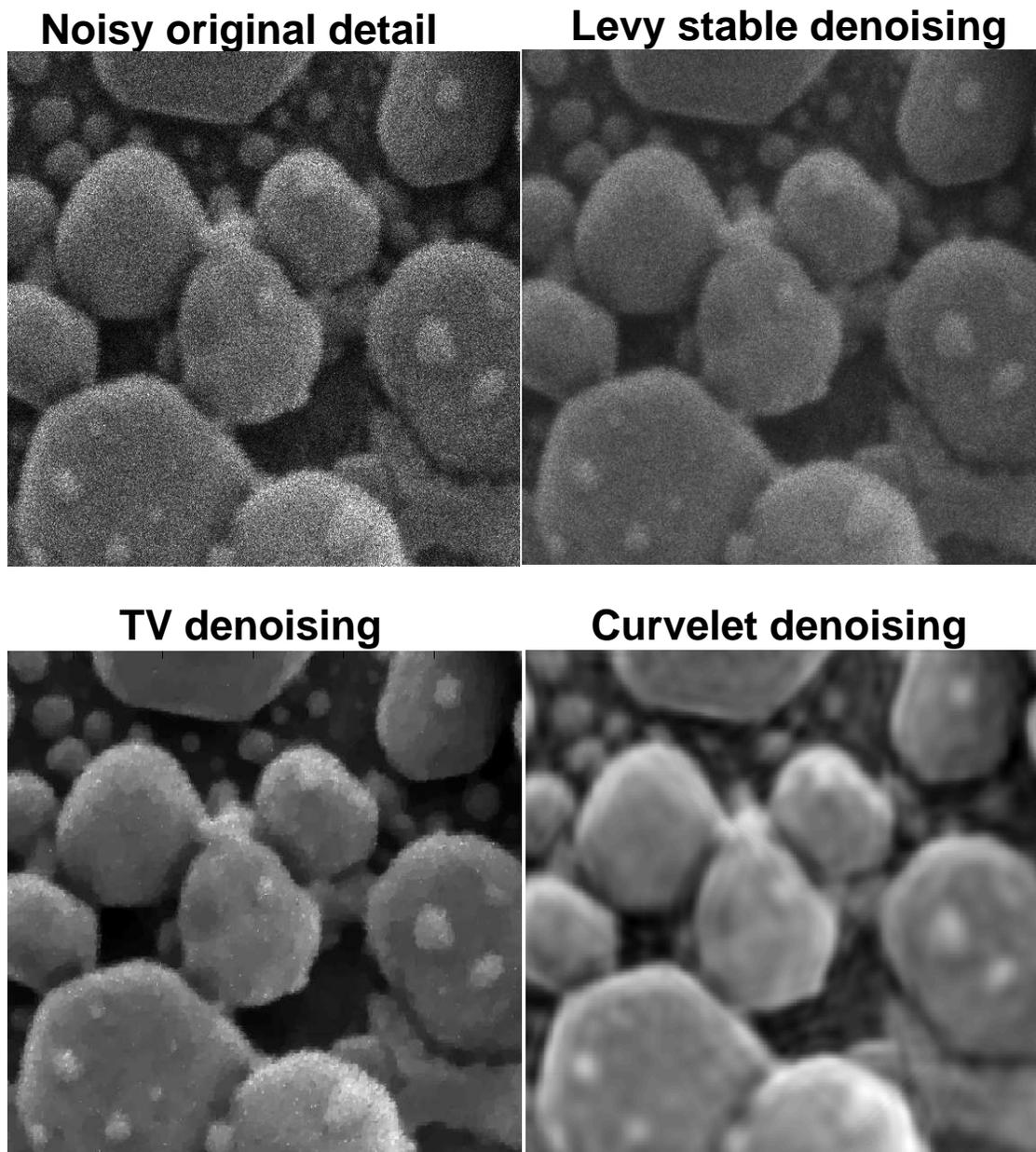


FIG. 10. Examination of parts of the gold on carbon sample in Figure 9, under magnification, reveals extent of TV and Curvelet smoothing of surface texture, as compared with Lévy stable image. Here, field of view is 150 nm.

TABLE 3.

Behavior of $\|f_d\|_1$, $\|f_d\|_2$, and $\|\nabla_d f\|_1$, in Figure 9 denoising.
 Note severe $\|\nabla_d f\|_1$ reduction in Curvelet and TV denoising.

Image $f(x, y)$	$\ f_d\ _1$	$\ f_d\ _2$	$\ \nabla_d f\ _1$
Noisy original	74	84	47000
Lévy stable ($\beta = 0.2$, $t^\dagger = 0.1$)	74	78	15000
Split Bregman TV ($\omega = 0.025$)	73	82	3500
Curvelet thresholding ($\sigma_n = 30$)	64	70	3000

- Principal properties", *Journal of Mathematics and Mechanics* 13, 407–478 (1964).
- [23] X. Q. Zhang and J. Froment, "Total variation based Fourier reconstruction and regularization for computerized tomography," *IEEE 2005 Nuclear Science Symposium Conference Record*, Volume 4, 2332–2336 (2005).