# TWO NEW NONLINEAR NONLOCAL DIFFUSIONS FOR NOISE REDUCTION

PATRICK GUIDOTTI, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA AT IRVINE, IRVINE, CA 92697-3875 (GPATRICK@MATH.UCI.EDU) AND JAMES V. LAMBERS, DEPARTMENT OF ENERGY RESOURCES ENGINEERING, STANFORD UNIVERSITY, STANFORD, CA 94305-2220 (LAMBERS@STANFORD.EDU).

Abstract. Two new nonlocal nonlinear diffusion models for noise reduction are proposed, analyzed and implemented. They are both a close relative of the celebrated Perona-Malik equation. In a way, they can be viewed as a new regularization paradigm for Perona-Malik. They do preserve and enhance the most cherished features of Perona-Malik while delivering well-posed equations which admit a stable natural discretization. Unlike other regularizations, however, certain piecewise smooth functions are (meta)stable equilibria and, as a consequence, their dynamical behavior and that of their discrete implementations can be fully understood and do not lead to any "paradox". The presence of nontrivial equilibria also explains why blurring is kept in check. One of the models has been proved to be well-posed. Numerical experiments are presented that illustrate the main features of the new models and that provide insight into their interesting dynamical behavior as well as demonstrate their effectiveness as a denoising tool.

## 1. Introduction

Denoising is one of the fundamental problems in image processing. Together with deblurring, it is an unavoidable component of image acquisition. The literature about this topic is vast and many methods have been devised, which perform the task to various degrees of satisfaction. In a recent paper [8], Buades, Coll, and Morel provide a nice overview of the major categories of available methods. They

This research was supported by the National Science Foundation under award number DMS-0712875.

all have their advantages and disadvantages. The common starting point is an image

$$u(i) = v(i) + n(i)\,,\ i \in 1, \dots, N$$

which is a collection of information about $N$ pixels. In the experiments performed in this paper it will always be a gray scale image. In this case $u$ is scalar-valued and takes on quantized integer values between 0 and 255. The image is typically the superposition of "real" informational content $v$ and some added noise $n$ originating, e.g., in the acquisition phase. Any denoising method aims at recovering the $v$-component by ridding the image of the superposed noise. This is clearly an impossible task as small scale real features of the image can become indistinguishable from artificial features created by noise. It follows that all known denoising algorithms tend to produce artifacts such as blurring, staircasing, and others well-known undesirable phenomena.

In this paper two new PDE-based models for noise reduction are proposed in the spirit of the original Perona-Malik model but with a fundamentally new twist. Even though the new equations can be thought of as new regularizations of the Perona-Malik equation, they distinguish themselves from other known regularizations in at least one crucial respect. Their nonlinearity applied to piecewise smooth functions is still singular (non-smooth) and, as a consequence, piecewise smooth functions are equilibria for the corresponding evolution. Thus the proposed equations are not only a new well-posed regularization of Perona-Malik but they show that (ill-posed) forward-backward diffusion is not a necessary ingredient in a recipe for an effective PDE-based denoising tool capable of preserving edges. They can in fact achieve this goal as a consequence of the nontrivial dynamical behavior engendered by the presence of non-trivial steady-states.

A brief overview will be given of the Perona-Malik equation and of a series of fixes proposed over the years in order to cure its mathematical and/or practical shortcomings. Non PDE-based methods also offer a valid alternative for denoising but will not take center stage in this paper. The interested reader is, however,

referred to [8] and the references cited therein for a more exhaustive catalog of methods available in the literature. The proposed equations read

$$\begin{cases} u_t = \frac{1}{1+c^2[(-\triangle)^{1-\varepsilon}u]^2}\triangle u & \text{in } \Omega \text{ for } t>0\,, \\ \partial_\nu u = 0 & \text{on } \partial\Omega \text{ for } t>0\,, \\ u(0)=u_0 & \text{in } \Omega \text{ for } t=0\,, \end{cases} \tag{1.1}$$

and

$$\begin{cases} u_t = \operatorname{div}\big(\frac{1}{1+c^2|\nabla^{1-\varepsilon}u|^2}\nabla u\big) & \text{in } \Omega \text{ for } t>0\,, \\ u \text{ periodic} & \text{for } t>0\,, \\ u(0)=u_0 & \text{in } \Omega \text{ for } t=0\,. \end{cases} \tag{1.2}$$

where $\Omega$ is the unit square in $\mathbb{R}^2$ and represents the image canvas, and $c>0$ is a threshold parameter which determines the size of the gradients which will be preserved. Thus, in these models, pixel values are thought of as a discrete realization of a continuous image function $u$. The original image is $u_0$ and the nonlinear evolution is supposed to reduce its noise component. The proposed equations are characterized by the nonlocal dependence introduced in the nonlinear diffusivity. The nonlocal terms will be properly defined in the next section. Equation (1.2) is directly related to the original Perona-Malik equation which is obtained in the limiting case where $\varepsilon = 0$. In one space dimension, equation (1.1) can also be related to the Perona-Malik equation by a similar argument and a change of dependent variable. The details are postponed to the next section. In both cases, comparison with Perona-Malik will reveal that the new equations can be viewed either as a nonlinear diffusion with slightly reduced intensity in the nonlinearity, or as a new

type of regularization. The Perona-Malik equation

$$\begin{cases} u_t = \operatorname{div}\big(\frac{1}{1+|\nabla u|^2}\nabla u\big) & \text{in } \Omega \text{ for } t>0\,, \\ \partial_\nu u = 0 & \text{on } \partial\Omega \text{ for } t>0\,, \\ u(0)=u_0 & \text{in } \Omega \text{ for } t=0\,. \end{cases} \tag{1.3}$$

was proposed in [22] and is probably the most well-known nonlinear diffusion model for denoising in a class of models commonly referred to as *anisotropic diffusions*. A good reference for such models is the book by Weickert [27], which is an outgrowth of his thesis [26], and the references cited later in this introduction.

While discretizations of (1.3) yield quite an effective denoising tool, the equation is itself mathematically ill-posed, as has been pointed out in [17]. By computing the divergence in (1.3) one in fact obtains

$$\operatorname{div}\big(\frac{1}{1+|\nabla u|^2}\nabla u\big) = \frac{1}{1+|\nabla u|^2}\big[\triangle u - 2\frac{\nabla u^T D^2 u \nabla u}{1+|\nabla u|^2}\big]\,.$$

It follows that if $u$ has a sharp gradient across a hyper-surface, diffusion in the normal direction can reverse its sign. Thus (1.3) is a so-called forward-backward diffusion equation and this is precisely the structure which has been exploited in [17] to prove ill-posedness. This is not very satisfactory because the lack of well-posedness makes it impossible to derive theoretical results predicting and explaining the behavior of an algorithm based on (1.3). Some researchers [13, 15, 7, 6] have therefore focused their attention on understanding the properties of semi-discrete versions of (1.3), where the spatial variable is discretized. The results are confined to the one dimensional case.

Denoising methods obtained from discretizations of (1.3) also have practical limitations in the form of poor performance in flat regions (see [8]) and in the form of artifacts like staircasing, as will be pointed out below (see Fig. 1 and also [17, 26]). It is therefore only natural that significant effort was profused in order to overcome the theoretical and practical limitations of the Perona-Malik model. Within the framework of second-order models, the main avenue followed in this

endeavor was that of regularizations and relaxations. The latter fall into two main categories: Spatial and spatio-temporal regularizations and/or relaxations. Spatial regularizations are obtained by smoothing the argument of the nonlinearity by a $C^\infty$-kernel $G_\sigma$, thus leading to the equation

$$\begin{cases} u_t = \operatorname{div}\left(\frac{1}{1+|\nabla G_\sigma * u|^2}\nabla u\right), \\ \partial_\nu = 0\,, \\ u(0) = u_0\,. \end{cases}$$

Typically $G_\sigma$ is a Gaussian and $\sigma$ determines the scale beyond which regularization occurs. This approach goes back to [9] but also see [1, 20]. More recently, a number of researchers have begun considering spatio-temporal regularizations and/or time relaxations of (1.3). In this case the argument of the nonlinearity is regularized by a space-time convolution or substituted with relaxed versions of it as in [21, 11, 10]. Relaxation models with no spatial regularizations have been considered in [4, 5]. The latter are limiting cases of the model considered in [21]. In [2], the author tries to mimic a typical discretization of (1.3) whereby the equation is implicitly linearized by evaluating the nonlinearity at a previous time step. He therefore proposes a purely temporal regularization with a non-smooth kernel. This leads to a delayed Perona-Malik equation. The new equation is locally well-posed and seems to introduce significantly less blurring than the more conventional regularization techniques.

Limitations of (1.3) and, more generally, second-order models were already pointed out early on in [14]. Observing that they tend to produce "patchy" results, the authors argue that higher order discrete methods might be a good alternative. Many a research effort went since then into the use of nonlinear fourth order diffusions in image processing [28, 12, 18]. Since this paper focuses mainly on the second-order case, a detailed discussion of these models and their literature is omitted. The ideas of this paper can, however, be extended to a variety of fourth-order models and will be the topic of another forthcoming paper.

One of the possible reasons why the Perona-Malik equation works well in practice might be that characteristic functions are, at least formally in some sense, stationary solutions of the equation. In a one dimensional setting, this was formulated via a generalized concept of stationary solution in [17]. Having characteristic functions as stationary solutions is a desirable property for any denoising equation but is not enjoyed by any of the known regularizations of the Perona-Malik equation. From the variational point of view and in the one-dimensional setting, piecewise constant functions can be viewed as stationary points of

$$E(u) = \frac{1}{2}\int_\Omega \log(1+|\nabla u|^2)\,dx\,,$$

for which (1.3) is a gradient flow, provided one interprets $u_x$ as the absolutely continuous part of the first derivative ([7]).

The equations proposed in this paper are new and possess a number of welcome and desirable properties from both the theoretical and the practical point of view. First of all (1.1) can be shown to be well-posed for $\varepsilon > 0$ (globally in time in a more restrictive range of $\varepsilon$) for smooth initial data and to admit continuous piecewise affine functions as stationary solutions in its one-dimensional formulation. The local well-posedness of (1.2) is studied in a forthcoming paper, where it will be shown to admit characteristic functions of smooth sets as stationary solutions. This is due to the reduction in nonlinearity intensity obtained for $\varepsilon > 0$. In fact the gradient of a characteristic function of a domain $D \subset \Omega$ with smooth boundary $\partial D$ is a distribution (of line integral type) supported on the boundary. On the other hand the fractional gradient used in (1.2) produces a function with an integrable singularity along the boundary $\partial D$ and which is otherwise smooth. Thus the nonlinearity applied to it gives an at least continuous function which vanishes precisely on $\partial D$. The product

$$\frac{1}{1+|\nabla^{1-\varepsilon}u|^2}\nabla u$$

can therefore be justified and, for $\varepsilon > 0$ small enough, it vanishes. For $\varepsilon = 0$, the product can not be made sense of as it involves multiplication of distributions.

Numerical experiments will show that denoising based on (1.1) and (1.2) lead to visually distinct results which can be directly traced back to the equation having two distinct classes of nontrivial stationary solutions as described above.

Well-posedness of (1.1) and (1.2) make working with discretizations of them very safe. Secondly, they allow for natural, FFT-based, efficient and stable pseudo-spectral discretizations. Experiments performed with them demonstrate their effectiveness (especially in the case of (1.2)) at noise-reduction with remarkable preservation of sharp edges. They also appear to avoid known shortcomings of discretizations of the Perona-Malik equation itself such as staircasing and poor performance in flat regions. Choosing $\varepsilon = \frac{1}{2}$ for the nonlinearity in the diffusion coefficient in (1.1) and $\varepsilon = 0$ in (1.2) lead the same intensity as in the leading order term of (1.3). Numerical experiments with realistic images will show that $\varepsilon < \frac{1}{2}$ and close to $\frac{1}{2}$ for (1.1) and $\varepsilon > 0$ and small for (1.2) are a good choice. Thus (1.1) and (1.2) effectively correspond to a nonlinearity intensity reduction or "regularization" of (1.3) through the use of fractional order differential operators.

Fractional derivatives have been used in [19] for edge detection. The authors report that fractional derivatives can successfully be employed in order to enhance selectivity and robustness in the presence of noise.

Finally, it would be quite natural to consider equations (1.1) and (1.2) with an additive fidelity term such as $\lambda(u_0 - u)$ for some $\lambda > 0$. Numerical experiments, however, exhibit little to no benefit. This might be due to the fact that the fidelity term destroys the dynamical properties of the equations as simple piecewise smooth functions are no longer equilibria unless the initial image is itself a simple piecewise smooth image.

It should be observed that methods other than PDE-based nonlinear diffusions have been studied and implemented in the context of noise reduction. Here the focus is more on the former and the interested reader is therefore referred to the already cited overview paper [8] and the references found therein for a wider perspective on the subject. It is, however, worthwhile pointing out that the recent developments

in local filtering methods, such as the nonlocal means method proposed in [8], also have a distinct nonlocal feature.

The paper is structured as follows. In the next section, well-posedness results are quoted which have been obtained for (1.1), whereas the well-posedness of (1.2) will be the topic of a future paper. The following sections introduce discretizations of (1.1) and (1.2), which are simple and prove efficient and stable. Finally, in the last section, results will be presented of experiments on real images. Particularly insightful one dimensional experiments will also be presented.

## 2. The Equations

In this section equations (1.1) and (1.2) are properly formulated. Starting with (1.1), consider the evolution equation in the function space $\mathrm{L}_2(\Omega)$. Then the $\mathrm{L}_2(\Omega)$-realization $A$ of $-\triangle$ with Neumann boundary condition is defined through

$$Au = -\triangle u\,,\ u \in \mathrm{dom}(A)\,, \mathrm{dom}(A) = \left\{u \in \mathrm{H}^2(\Omega) \,\middle|\, \partial_\nu u = 0\right\}. \tag{2.1}$$

This operator is induced by the form

$$a(u,v) = \int_\Omega \nabla u \cdot \nabla v \, dx\,,\ u, v \in \mathrm{H}^1(\Omega)\,,$$

and is therefore positive semi-definite. This can also be seen by explicitly computing its eigenvalues and eigenfunctions in the simple geometry considered here. This will be exploited in the choice of numerical method later. It is therefore possible to define fractional powers $A^\rho$ of $A$ for $\rho \in (0,1)$ using its symbol. In fact

$$A = \mathcal{C}^{-1}\,\mathrm{diag}[(\lambda_n)_{n\in\mathbb{N}}]\mathcal{C}$$

where $\mathcal{C}$ is the two-dimensional cosine transform and the sequence $(\lambda_n)_{n\in\mathbb{N}}$ contains the ordered eigenvalues of $A$. Then one simply defines

$$A^\rho = \mathcal{C}^{-1}\,\mathrm{diag}[(\lambda_n^\rho)_{n\in\mathbb{N}}]\mathcal{C} \tag{2.2}$$

as the eigenvalues are all non-negative. It follows that equation (1.1) is a well-defined quasi-linear nonlocal diffusion. As for equation (1.2) the starting point is that

$$\partial_z = \mathcal{F}_z^{-1}\,\mathrm{diag}[(2\pi i n)_{n\in\mathbb{Z}}]\mathcal{F}_z$$

where $\mathcal{F}_z$ is the partial Fourier transform with respect to $z = x, y$. The fractional differentiation operator appearing in (1.2) is defined through

$$\nabla^\rho = \begin{bmatrix} \partial_x^\rho \\ \partial_y^\rho \end{bmatrix} \tag{2.3}$$

where

$$\partial_z^\rho = \mathcal{F}_z^{-1}\,\mathrm{diag}[\big((2\pi n)^\rho\, e^{i\rho\frac{\pi}{2}\,\mathrm{sign}(n)}\big)_{n\in\mathbb{Z}}]\mathcal{F}_z$$

In either case, the choice of $\rho = 0$ leads to a fully nonlinear ill-posed diffusion equation, which for (1.2) coincides with the classical Perona-Malik equation.

## 3. Analytical Results

It has been proved in [16] that (1.1) admits local-in-time smooth solutions for smooth initial data which can be shown to be global for appropriate choice of $\varepsilon$.

**Theorem 3.1.** *Let $\varepsilon \in (0,1]$ and assume that $u_0 \in \mathrm{W}^{2-2/p}_{p,N}(\Omega)$ and that $p > \frac{2+n}{2\varepsilon}$. Then equation (1.1) has a unique maximal solution $u : [0, T_{max}] \to \mathrm{L}_p(\Omega)$ with*

$$u \in \mathrm{L}_p\big([0,T_{max}], \mathrm{W}^2_{p,N}(\Omega)\big) \cap \mathrm{W}^1_p\big([0,T_{max}), \mathrm{L}_p(\Omega)\big)\,.$$

*Furthermore, if $n = 1$ and $\varepsilon > 3/4$, then the solution is global in time and converges to a trivial steady-state. The subscript $N$ is used to indicate the closed subspace of functions satisfying homogeneous Neumann condition.*

Furthermore, a Lyapunov functional exists for smooth solutions of (1.1), which is given by

$$E(u) = \int_\Omega |\nabla u|^2\, dx\,. \tag{3.1}$$

This follows from the validity of the following conservation relation

$$\frac{1}{2}\frac{d}{dt}\int_\Omega |\nabla u|^2\,dx = -\int_\Omega \frac{|Au|^2}{1+|A^{1-\varepsilon}u|^2}\,dx\,, \tag{3.2}$$

which is easily verified. Smooth solutions do eventually evolve to trivial steady-states, but they do so by flirting for a long time with nearby non-smooth metastable "edgy" states. This offers a possible explanation for the effectiveness of the proposed denoising method which is most strikingly illustrated by one- and two-dimensional experiments in a later section.

In a one-dimensional setting, equation (1.1) turns out to be very closely related to (1.2), and can therefore be viewed as a new kind of regularization of (1.3). This is best seen by first observing that $v(x) = \int_0^x u(y)\,dy$ satisfies

$$\begin{cases} v_t - \frac{1}{1+v_{xx}^2}v_{xx} = 0 & \text{in } (0,1) \text{ for } t>0\,,\\ v(t,0)=0\,,\ v(t,1)=\int_0^1 u_0(y)\,dy & \text{for } t>0\,,\\ v(0,\cdot)=\int_0^\cdot u_0(y)\,dy & \text{in } [0,1]\,, \end{cases} \tag{3.3}$$

whenever $u$ is a solution of (1.3) in its one dimensional formulation. To see this, take into account that the average of $u_0$ is preserved during time evolution as follows from (1.3). Choosing $w(x) = v(x) - x\int_0^1 u_0(y)\,dy\,,\ x\in[0,1]\,$, the boundary condition can be "homogenized" to give

$$\begin{cases} w_t - \frac{1}{1+w_{xx}^2}w_{xx} = 0 & \text{in } (0,1) \text{ for } t>0\,,\\ w(t,0)=0=w(t,1) & \text{for } t>0\,,\\ w(0,\cdot)=\int_0^\cdot u_0(y)\,dy - (\cdot)\int_0^1 u_0(y)\,dy & \text{in } [0,1]\,. \end{cases} \tag{3.4}$$

Conversely, any solution of (3.3) leads to a solution of (1.3) by setting

$$u(x) = v_x(x)\,,\ x\in[0,1]\,.$$

Notice that if $u$ is a piecewise constant function, then $v$ is a continuous piecewise affine function. Thus continuous piecewise affine functions play the same role for (3.4) as piecewise constant functions do for (1.3).

Equation (3.4) is fully nonlinear and clearly presents the same analytical difficulties as the original Perona-Malik equation. It is, however, more closely related to (1.1). Equation (1.1) is in fact obtained from it through nonlinearity intensity reduction by using the fractional powers of the Laplacian.

Constants are stationary solutions of (1.1) in its one- and two-dimensional formulations, and they eventually determine the asymptotic behavior of smooth solutions. In its one-dimensional formulation, however, (1.1) admits continuous piecewise affine functions as stationary solutions as well. This fact is proved in [16] and is a welcome feature of the modified equations, as it is at the origin of the interesting and practically useful medium-term behavior of the solutions. Observe that these functions are only "morally" stationary solutions of the original (1.3), as a rigorous proof would require the availability of a well-defined multiplication for distributions (generalized functions). Equation (1.2), on the other hand, has the advantage of admitting characteristic functions of domains with smooth boundaries as stationary solutions. This will make the denoising method based on its discretization very effective. A more thorough analysis of its analytical properties will be performed elsewhere, but numerical experiments performed in this paper suggest that (1.2) gives rise to a dynamical behavior similar to that of (1.1) with the role of piecewise affine functions taken on by characteristic functions. It should be observed that (1.2) conserves the $\mathrm{L}_1(\Omega)$-norm of positive solutions and admits the $\mathrm{L}_2(\Omega)$-norm as a Lyapunov functional, since

$$\frac{1}{2}\frac{d}{dt}\int_\Omega u^2\,dx = -\int_\Omega \frac{|\nabla u|^2}{1+|\nabla^{1-\varepsilon}u|^2}\,dx\,.$$

## 4. Numerical Implementation

Next, numerical discretizations of (1.1) and (1.2) are derived and used to perform numerical experiments intended to illustrate and demonstrate the claimed improvements on the classical Perona-Malik equation.

4.1. **The One-Dimensional Case.** The periodic differentiation operator $D$ is discretized spectrally by means of the fast Fourier transform $\mathcal{F}_n$, which yields the discretization

$$D_n = \mathcal{F}_n^{-1}\Lambda_n\mathcal{F}_n \tag{4.1}$$

where $n = 2^m$ denotes the number of grid points used and

$$\Lambda_n = 2\pi i \,\mathrm{diag}\left(-\frac{n}{2}+1, -\frac{n}{2}+2, \ldots, 0, 1, \ldots, \frac{n}{2}\right).$$

The time variable is discretized by an implicit Euler scheme so that

$$\mathbf{u}^{k+1} = A_n^{-1}\mathbf{u}^k \tag{4.2}$$

where the matrix $A_n$ is defined by

$$A_n = \left[\mathrm{id}_n + D_n \frac{h_t}{1 + c^2(D_n^{1-\varepsilon}\mathbf{u}^k)^2} D_n\right], \tag{4.3}$$

$\mathbf{u}^k$ is the spatial $n$-vector at time $t_k$, and $h_t > 0$ is the time step. Observe that by setting $\varepsilon = 0$, a discretization of the classical Perona-Malik equation is recovered. The exponentiation of $\Lambda_n$ is carried out as described in Section 2; that is,

$$(ik)^{1-\varepsilon} = |k|^{1-\varepsilon} e^{i\pi/2(1-\varepsilon)\mathrm{sign}(k)}, \; k = -n/2+1, \ldots, n/2.$$

For each time step, the iterative method GMRES (see [24]) is used to solve the system (4.2). It uses Arnoldi iteration (see [3]) by building a set of vectors $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ that form an orthogonal basis for the Krylov subspace

$$K_m = \mathrm{span}\{\mathbf{u}^k, A_n\mathbf{u}^k, A_n^2\mathbf{u}^k, \ldots, A^{m-1}\mathbf{u}^k\}. \tag{4.4}$$

where $m \ll n$. Then

$$AQ_m = Q_{m+1}H_m \tag{4.5}$$

where

$$Q_m = \begin{bmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_m \end{bmatrix} \tag{4.6}$$

and $H_m$ is an upper Hessenberg matrix.

For $j = 1, 2, \ldots$, an approximate solution $\mathbf{u}_j^{k+1}$ is obtained by computing

$$\mathbf{u}_j^{k+1} = Q_j \mathbf{v}_j \,,$$

where $\mathbf{v}_j$ is the solution of the much smaller system

$$H_j \mathbf{v}_j = \|\mathbf{u}^k\|_2 \mathbf{e}_1, \quad \mathbf{e}_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T . \tag{4.7}$$

This can be solved efficiently using Gaussian elimination, since only one element per column needs to be eliminated. GMRES iteratively computes approximations $\mathbf{u}_j^{k+1}$, for $j = 1, 2, \ldots$, until the residual

$$\|A_n \mathbf{u}_j^{k+1} - \mathbf{u}^k\|_2 = \|H_j \mathbf{v}_j - \|\mathbf{u}^k\|_2 \mathbf{e}_1\|_2 \tag{4.8}$$

is sufficiently small. For one-dimensional problems, this method performs very effectively, converging to the desired tolerance after only a few iterations.

In view of the spectral discretization of the Laplacian used in (4.2), and the variable diffusion coefficient, the resulting matrix is dense in both physical and Fourier space. However, a method such as GMRES, though designed for sparse systems, is still practical in this case. This is because matrix-vector multiplication can be implemented in $O(n \log n)$ time by employing the FFT to multiply vectors by $A_n$, and performing multiplication by the diffusion coefficient in physical space.

4.2. **Two-Dimensional Problems.** The discretization of (1.1) and (1.2) in two dimensions is a direct generalization of the periodic, one-dimensional case discussed

above. A spectral discretization is again chosen for the periodic differentiation operator, implemented using the fast Fourier transform. For simplicity, it is assumed that the grid consists of $n = 2^m$ grid points for each dimension, in which case the discretization of the differentiation operator $D_z$, where $z = x, y$, has the form

$$A_{z;n,n} = \mathcal{F}_{n,n}^{-1} \Lambda_{z;n,n} \mathcal{F}_{n,n}$$

where the diagonal entry of $\Lambda_{x;n,n}$ corresponding to the Fourier coefficient with wave number $(j, k)$ is $2\pi ij$, and the diagonal entry of $\Lambda_{y;n,n}$ corresponding to the same Fourier coefficient is $2\pi ik$. The resulting scheme is

$$\mathbf{u}^{k+1} = \left[ \mathrm{id}_{n,n} + \sum_{z=x,y} D_{z;n,n} \frac{h_t}{1 + c^2 (\sum_{z=x,y} [D_{z;n,n}^{1-\varepsilon} \mathbf{u}^k]^2)} D_{z;n,n} \right]^{-1} \mathbf{u}^k \,. \tag{4.9}$$

For the Neumann case, the Laplacian $A$ is discretized in a similar manner, except that the fast cosine transform $C_{n,n}$ is used. The resulting discretization has the form

$$A_{n,n} = \mathcal{C}_{n,n}^{-1} \Lambda_{n,n} \mathcal{C}_{n,n}$$

where the diagonal entry of $\Lambda_{n,n}$ corresponding to the Fourier cosine coefficient with wave number $(j, k)$ is $-4\pi^2(j^2 + k^2)$. The resulting scheme is

$$\mathbf{u}^{k+1} = \left[ \mathrm{id}_{n,n} + \frac{h_t}{1 + c^2 [(-A_{n,n})^{1-\varepsilon} \mathbf{u}^k]^2} A_{n,n} \right]^{-1} \mathbf{u}^k. \tag{4.10}$$

As in the one-dimensional case, GMRES is used to solve the systems (4.9) and (4.10), but because of the greater storage requirements in this case, and the fact that GMRES requires $O(nj)$ storage where $n$ is the length of the right-hand side and $j$ is the number of iterations, the Arnoldi iteration is restarted after every 25 iterations, using the most recently computed approximation $\mathbf{u}_j^{k+1}$ as the initial vector for the Krylov subspace (4.4). The time step $h_t$ is adaptively reduced if restarting becomes necessary, or increased if GMRES converges to the desired tolerance in very few iterations.

Because of the added computational expense required for larger images, in terms of storage and iterations of GMRES, such images are handled by decomposing them into blocks of a manageable size, such as $128 \times 128$, and denoised independently of one another. When using periodic boundary conditions, a "padding" border of 16 pixels is added around each block, reflecting the image across its boundary as needed, in order to prevent artifacts from appearing at the interfaces between blocks or on the boundary of the entire image. This strategy also allows for parallelization of the method. It is not used for (1.1) as homogeneous Neumann conditions do not cause boundary effects.

## 5. Numerical Experiments

5.1. **One-Dimensional Case.** The one-dimensional situation is interesting in itself, but only a few experiments will be shown here which support the claims made so far about the features of the equations under consideration. They effectively illustrate the main properties of the proposed diffusions. A variety of subsequent two-dimensional experiments will corroborate the one-dimensional findings, in particular for (1.2).

5.2. **Non-Divergence Equation.** Consider the smooth initial datum

$$u_0(x) := 100x^2(1-x^2)\,,\ x \in [0,1]\,. \tag{5.1}$$

Figure 1 depicts the first derivative of the function $u^k$ after 100 time steps of size $h_t = 0.06$ for $m = 8$. The blue curve corresponds to the Perona-Malik solutions and staircasing is apparent. The red, magenta and cyan curves correspond to solution of the modified equation for $\varepsilon = 0.3, 0.2, 0.1$, respectively, and $c = 1$. It is an illustration of the effect of intensity reduction in the nonlinear diffusion coefficient. The intensity $\varepsilon$ clearly acts as a regularization parameter on the original Perona-Malik equation. It appears that the effects of the Perona-Malik equation are preserved with the notable exception of staircasing, which does not occur in the regularized equation unless $\varepsilon$ is chosen very small. Notice that this does, however,

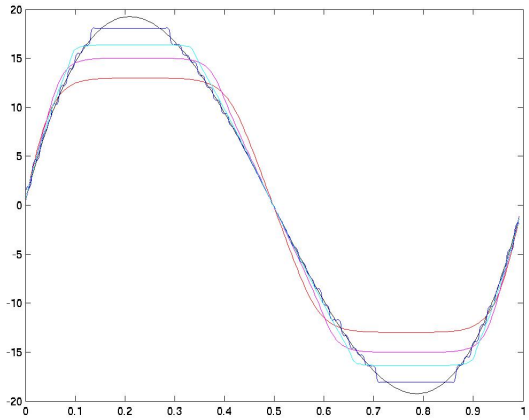


FIGURE 1. The gradient of the solution of (5.1) for different values of $\varepsilon$.

not come at the expense of introducing undesired blurring. This is due to the fact that the regularizing effect is mitigated by the presence of non-trivial steady-states which temporarily attract the solution. This feature of equations (1.1) and (1.2) plays a crucial role, as will become apparent in all experiments shown in this section. Some decrease in contrast is the only price paid. Notice that the black curve depicts the first derivative of the initial value. Comparison with the other curves also reveals some deblurring. Continuous piecewise affine functions are steady-states of (1.1). It is, however, not clear if a smooth solution of (1.1) can develop a singularity in finite time. In any event, solutions feel the presence of these nontrivial steady-states, which are in fact metastable equilibria. A solution emanating from a smooth initial datum never seems to leave the smooth regime but still tends to evolve along a path which is at all times close to continuous piecewise affine functions. Diffusion, however, slowly but surely drives it to a trivial steady-state. This behavior is exemplified in Figure 2, where various stages of the evolution of a smooth solution with initial condition

$$u_0(x) = 10\sin(6\pi x) + 2\sin(4\pi x)\,,\ x \in [0,1]\,. \tag{5.2}$$

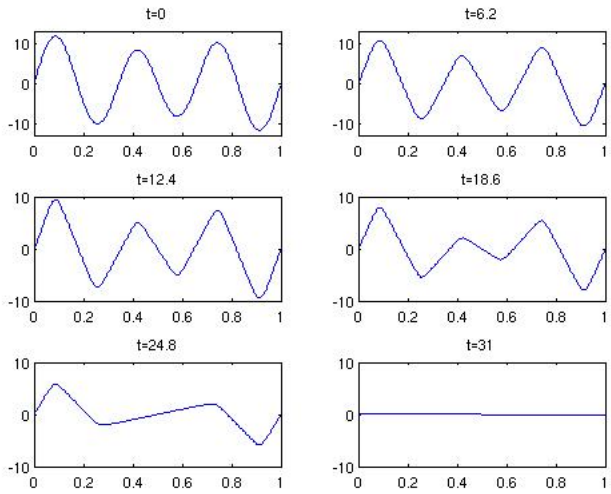


FIGURE 2. Typical evolution of a smooth initial datum along a "path of continuous almost piecewise affine functions" according to (1.1).

are depicted and where

$$\varepsilon = 0.3\,,\ c = 0.05\,,\ m = 8\,,\ h_t = 0.05\,.$$

5.3. **Divergence Equation.** A similar behavior is exhibited by solutions of (1.2) where, however, the role of continuous piecewise affine functions is played by piecewise constant functions. Figure 3 depicts the evolution of a smooth initial datum typified by

$$u_0(x) = 10\cos(6\pi x) + 4\sin(4\pi x)\,,\ x \in [0,1]\,,$$

with

$$\varepsilon = 0.3\,,\ c = 10\,,\ m = 8\,,\ h_t = 0.05\,.$$

In this case, the highly oscillatory perturbation of a simple step function is also considered to show how equation (1.2) and, similarly (1.1), have an ability to differentiate between local features, which are annihilated, and global features, which are preserved. This remarkable quality of the equations is a combination of the

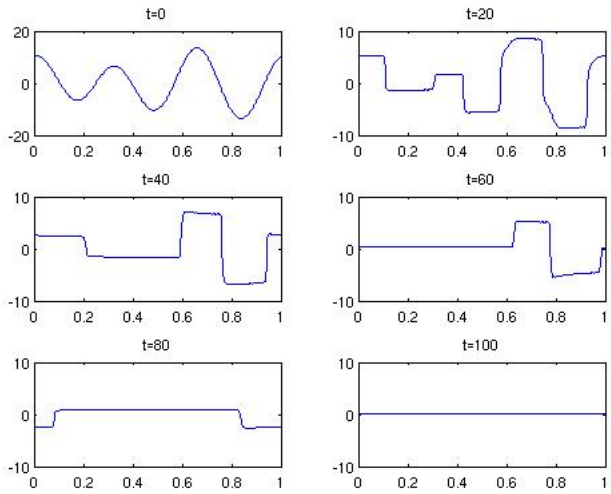


FIGURE 3. Typical evolution of a solution of (1.2) always close to a step function.

specifics of the nonlocality introduced in the nonlinearity, and the fact that step (piecewise affine) functions are steady states. Figure 4 depicts the evolution of

$$u_0(x) = \chi_{[0.25,0.75]} + 0.3\sin(64\pi\, x)\,,\ x \in [0,1]\,,$$

for

$$\varepsilon = 0.15\,,\ c = 10\,,\ m = 8\,,\ h_t = 0.05\,,$$

and clearly demonstrates the claim.

5.4. **Two-dimensional Case.** Real images are considered in this section. Additive white noise is added to the original image before it is fed to equations (1.1) and (1.2). More precisely, every pixel is independently corrupted by a uniformly distributed random variable with values in $[-20,20]\cap\mathbb{N}$ if not otherwise stated. The chosen experiments show that the qualities predicted and already observed in the one-dimensional setting are also observed in two dimensions and with real images for (1.2), whereas equation (1.1) does not seem to perform as well in two dimensions as in the one-dimensional case. This is probably due to the fact that continuous

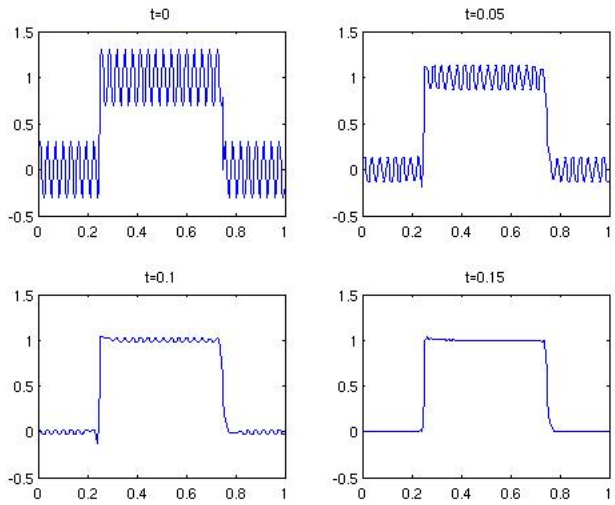


FIGURE 4. The fate of high frequency oscillations superposed on the simple step function $\chi_{[0.25,0.75]}$.

piecewise linear functions of two variables do not provide a flexible enough set of equilibria to allow for good approximation of the original image.

5.5. **Non-Divergence Equation.** Figures 5 and 6 show two experiments that illustrate the denoising effect of (1.1) on real images. Depicted are, from left to right and top down, the original image corrupted with noise, the output of the time integration, the method noise, and the uncorrupted original image. The method noise is the difference between the input and the output of the algorithm. The degree to which it resembles noise is an indication of the method efficacy (see [8]).

5.6. **Divergence Equation.** The same images as for the testing of (1.1) and additional images are used in this section. They are depicted in Figures 8–10 and they clearly demonstrate the efficacy of equation (1.2) as a denoising tool. In all Figures the same parameters are used. The image size is $512 \times 512$, $\varepsilon = 0.1$, $c = 0.5$, the time step is variable as explained in Section 4, and the integration time is $T = 9$. The corresponding SNR and the change in PSNR are summarized in Table 1 below. It should not come as a surprise that the method performs slightly better when the

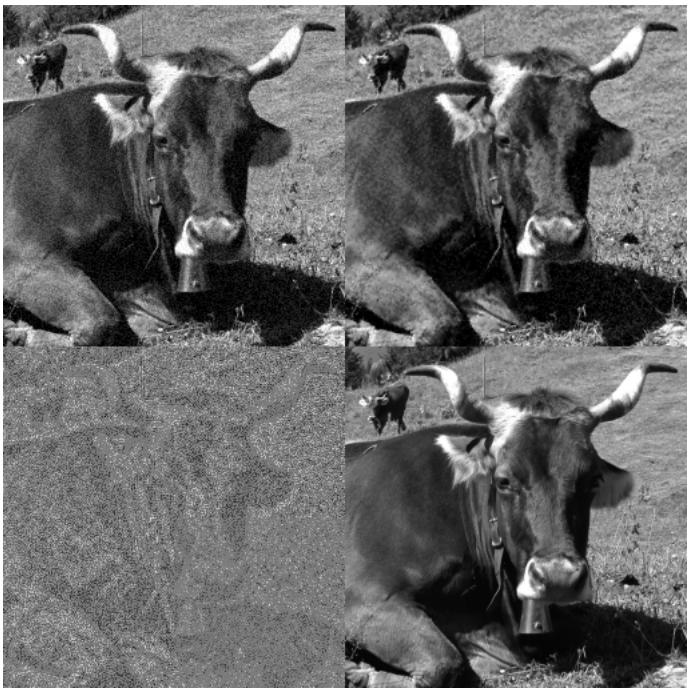

FIGURE 5. The denoising effect obtained with (1.1) with $\varepsilon = 0.6$ and $c = 1$ on a noisy cow. Depicted are the initial noisy condition, the final output of the denoising method at $T = 18$, the method noise and the original image. For the noisy image SNR= 5.01. PSNR goes from 27.00 to 28.30.

| Image | boats | cow | naive painting | downtown L.A. |
|---|---|---|---|---|
| SNR | 4.05 | 5.09 | 4.21 | 1.66 |
| PSNR in | 26.89 | 27.01 | 26.92 | 26.87 |
| PSNR out | 31.38 | 30.67 | 33.73 | 33.62 |

TABLE 1. SNR and PSNR for the input and output images, respectively.

underlying image has a more "cartoonish" structure.

Figures 11 and 12 depict the evolution of two images over a somewhat larger time range than needed for denoising. It shows how the evolution, just as in the one dimensional examples, tends to happen along a path that is always close to an appropriate piecewise constant function. The contrast is steadily reduced between adjacent regions which orderly merge into larger areas under the influence of diffusion. The evolution thus naturally delivers an interesting segmentation of the

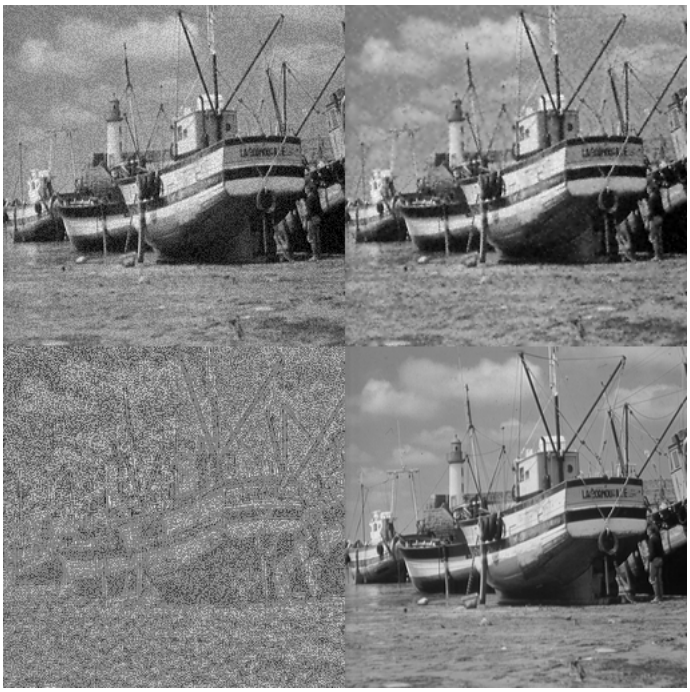

FIGURE 6. The denoising effect obtained with (1.1) with $\varepsilon = 0.6$ and $c = 1$ on the boats image. Depicted are the same combination of images as in Figure 5, with $T = 36$. SNR is 3.93 and PSNR goes from 26.88 to 29.10.

original image. Observe that no artificial terraces are ever created in the process, not in the long time examples nor in the previous examples, but edges always correspond to a feature present in the original image. As a consequence, the typical "patchiness" of nonlinear diffusions and visible in Figures 5–6 is avoided.

Finally, it should be observed that the methods deliver robust results even in the presence of noise of larger intensity, as demonstrated in Figure 13, where the boats image is now corrupted by Gaussian white noise with standard deviation 30. The solution is depicted at equidistant time intervals up to time $T = 2.34$ with parameters $c = 0.1$, $\varepsilon = 0.1$ and variable time step.

Figure 14 depicts the denoising result obtained by (1.2) on the $512 \times 512$ boats image which has been corrupted with Gaussian white noise with standard deviation 20. An output PSNR of 28.77 is obtained with parameters $c = 0.1$, $\varepsilon = 0.1$ and

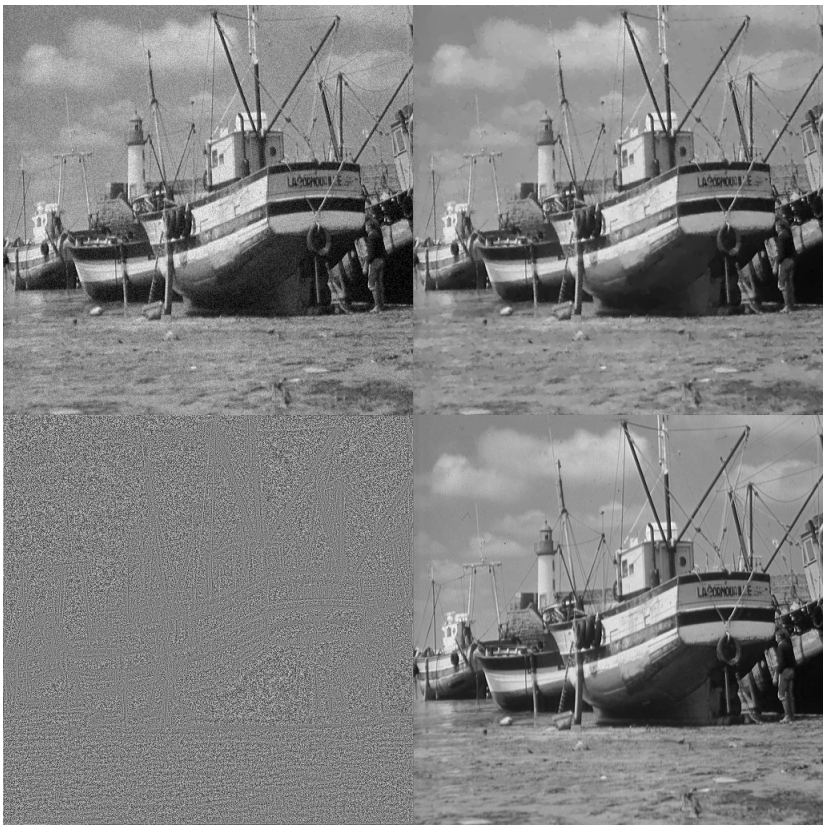

FIGURE 7. Denoising effect of (1.2) on a standard test figure. SNR is 4.05 and PSNR goes from 26.89 to 31.38.

$T = 2.34$. It compares well with published denoising experiments [25], the result of which the authors characterize as state-of-the-art. They are performed by means of the Bayesian least squares estimator based on the GSM model [23] where the output PSNR varies between 26.86 and 29.62 depending on the choice of neighborhood size.

## 6. Conclusions

In this paper a novel PDE-based denoising technique has been proposed and analyzed. Its main feature is the nonlinear nonlocal diffusion coefficient obtained by the use of fractional derivatives. Two equations, in particular, have been proposed and analyzed, one in divergence form, (1.2), and the other not, (1.1). The new

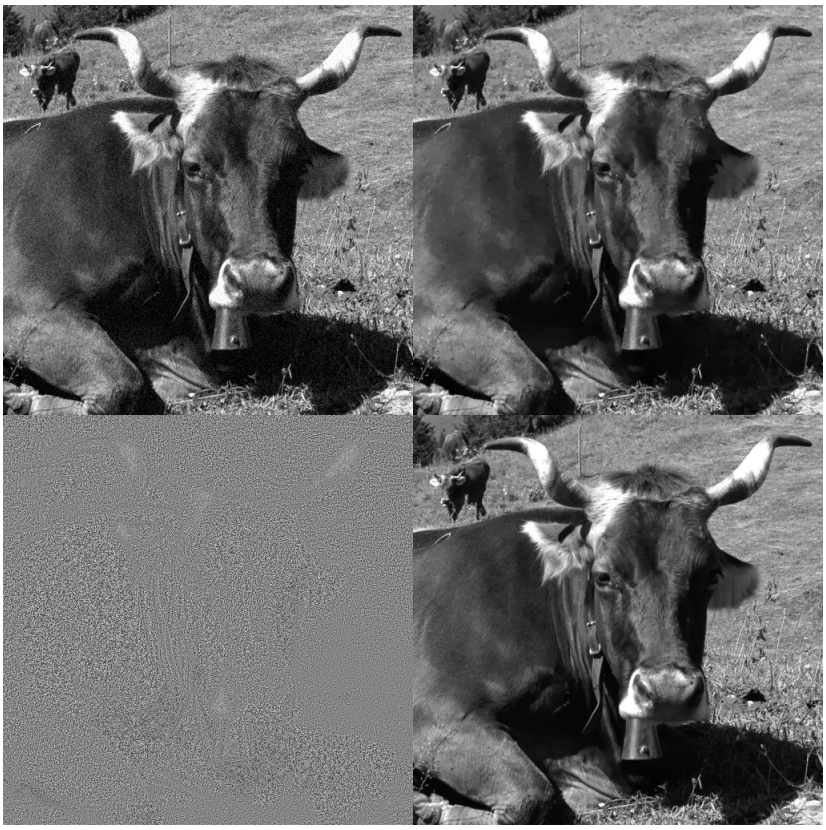

FIGURE 8. Denoising effect of (1.2) on a noisy cow. SNR is 5.09 and PSNR goes from 27.01 to 30.67.

equations can be interpreted as a novel regularization paradigm for the celebrated Perona-Malik equation. Equation (1.2), in particular, performs better than traditional regularizations. The reason for this is found in the combination of the degree of nonlocality and the (meta)stability of certain piecewise smooth functions. The first allows (1.1) and (1.2) to differentiate between local and global feature while the second prevents this from happening by indiscriminately annihilating information from the small scales. Equation (1.1) has already been proved well-posed whereas (1.2) will be analyzed in a future paper but is believed to be locally well-posed. Thus the new equations are not ill-posed, preserve many of the cherished features of Perona-Malik, and even improve on it (no staircasing nor poor performance on "flat" regions) as the many experiments performed in the previous section clearly

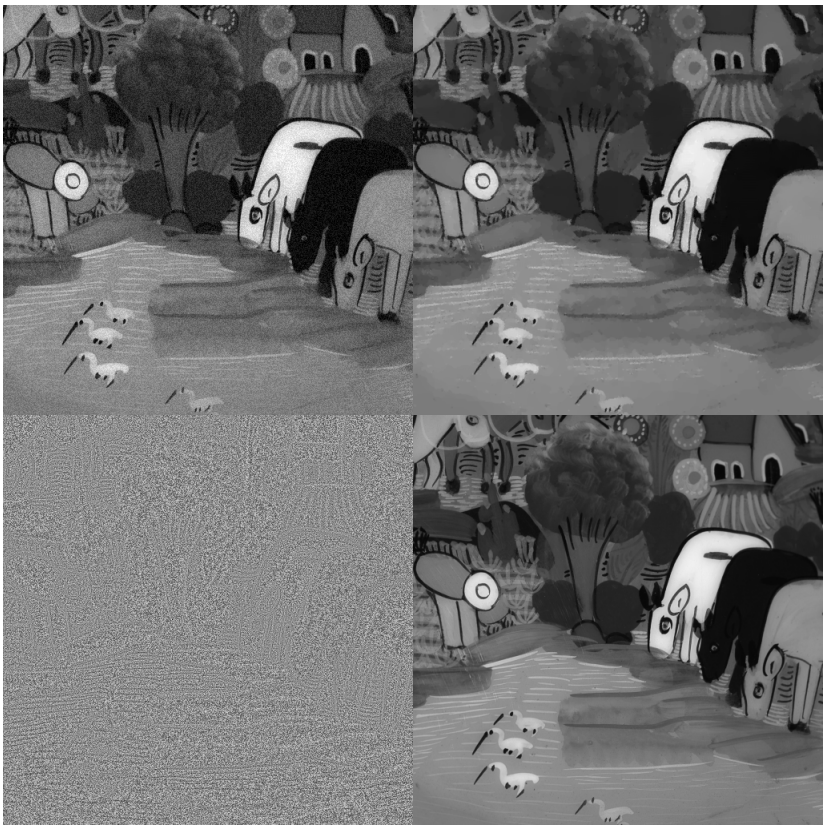

FIGURE 9. Denoising effect of (1.2) on a noisy naive painting. SNR is 4.21 and PSNR goes from 26.92 to 33.73.

demonstrate. They also allow for stable and efficient pseudo-spectral discretizations which are well-suited to capture the nonlocal derivatives needed.

The fact that fractional derivatives are used instead of a more standard regularization with a smooth kernel is crucial. In fact applying the fractional derivative to non-smooth data still produces singularities albeit integrable ones at the edges. It is, in this sense, a much milder regularization. In particular, it allows for non trivial equilibria for the system which engender an interesting dynamical behavior and explain the salient features of the corresponding implementations.

## REFERENCES

[1] L. ALVAREZ, P.-L. LIONS, AND J.-M. MOREL, *Image selective smoothing and edge-detection by non-linear diffusion. II.*, SIAM J. Numer. Anal., 29 (1992), pp. 845–866.

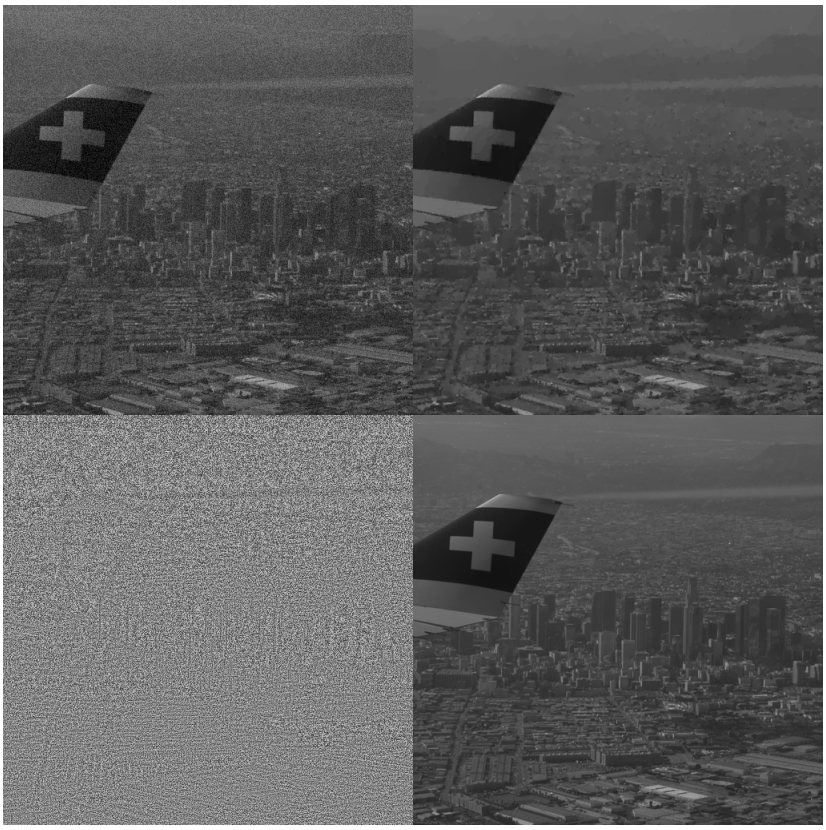

FIGURE 10. Denoising effect of (1.2) on noisy downtown L.A. SNR is 1.66 and PSNR goes from 26.87 to 33.62.

[2] H. AMANN, *Time-Delayed Perona-Malik Problems*, Acta Math. Univ. Comenianae, LXXVI (2007), pp. 15–38.

[3] W. E. ARNOLDI, *The principle of minimized iteration in the solution of the matrix eigenvalue problem*, Quarterly of Applied Mathematics, 9 (1951), pp. 17–25.

[4] A. BELAHMIDI, *Equations aux dérivées partielles appliquées à la restoration et à l'agrandissement des images.*, Ph.D. Thesis, Université Paris-Dauphine, Paris, 2003.

[5] A. BELAHMIDI AND A. CHAMBOLLE, *Time-delay regularization of anisotropic diffusion and image processing*, M2AN Math. Model. Numer. Anal., 39 (2005), pp. 231–251.

[6] G. BELLETTINI, M. NOVAGA, AND M. PAOLINI, *Convergence for long-times of a semidiscrete Perona-Malik Equation in one dimension*, Preprint, (2008).

[7] G. BELLETTINI, M. NOVAGA, M. PAOLINI, AND C. TORNESE, *Classification of the Equilibria for the semi-discrete Perona-Malik Equation*, Preprint, (2008).

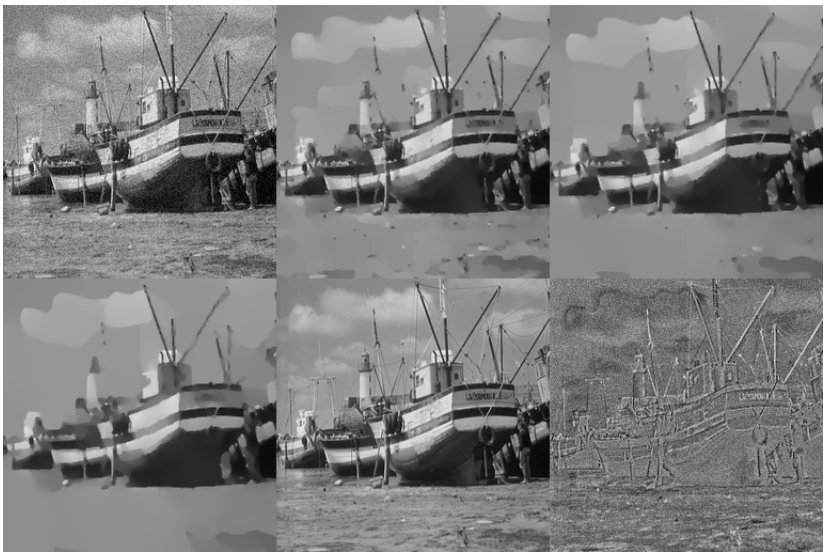

FIGURE 11. The segmentation effect of (1.2) on the boat image obtained by long time integration.

FIGURE 12. The segmentation effect of (1.2) on the cow image obtained by long time integration.

[8] A. BUADES, B. COLL, AND J. MOREL, *A Review of Image Denoising Algorithms, with a New One*, Multiscale Modeling and Simulations, 4 (2005), pp. 490–530.

[9] F. CATTÉ, P.-L. LIONS, J.-M. MOREL, AND T. COLL, *Image selective smoothing and edge-detection by non-linear diffusion*, SIAM J. Numer. Anal., 29 (1992), pp. 182–193.

FIGURE 13. An experiment with intense noise. SNR is 1.59 and PSNR goes from 18.84 to 26.70.

[10] Y. CHEN AND P. BOSE, *On the incorporation of time-delay regularization into curvature-based diffusion*, J. Math. Imaging Vision, 14 (2001), pp. 149–164.

[11] G. COTTET AND M. E. AYYADI, *A volterra type model for image processing*, IEEE Trans. Image Processing, 7 (1998), pp. 292–303.

[12] S. DIDAS, J. WEICKERT, AND B. BURGETH, *Stability and local Feature Enhancement of Higher Order Nonlinear Diffusion Filtering*, Pattern Recognition, 3663 (2005), pp. 451–458.

[13] S. ESEDOGLU, *An analysis of the perona-malik scheme*, Comm. Pure Appl. Math., 54 (2001), pp. 1442–1487.

[14] S. GEMAN, D. MCCLURE, AND D. GEMAN, *A Nonlinear Filter for Film Restoration and Other Problems in Image Processing*, Graphical Models and Image Processing, 54 (1992), pp. 281–289.

[15] M. GHISI AND M. GOBBINO, *Gradient Estimates for the Perona-Malik Equation*, Math. Ann., 3 (2007), pp. 557–590.

[16] P. GUIDOTTI, *A New Well-posed Nonlinear Nonlocal Diffusion*, Submitted.

[17] S. KICHENASSAMY, *The Perona-Malik paradox*, SIAM J. Appl. Math., 57 (1997), pp. 1328–1342.

[18] M. LYSAKER, A. LUNDERVOLD, AND X. TAI, *Noise Removal Using Fourth Order Differential Equations with Applications to Medical Magnetic Resonance Images in Space-Time*, IEEE Transaction on Image Processing, 12 (2003), pp. 1579–1590.

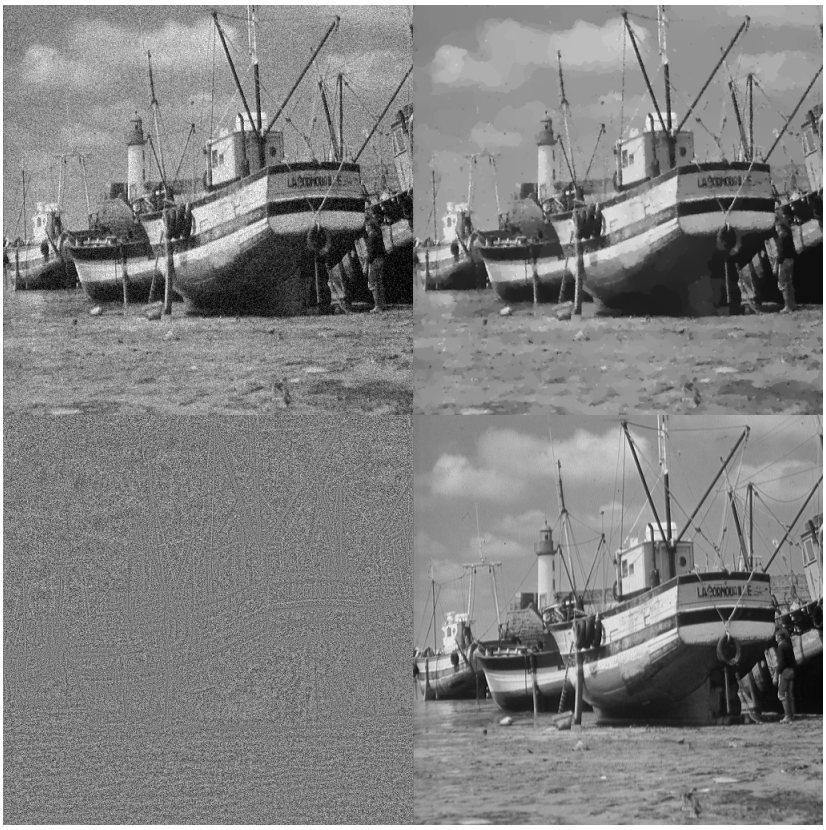

FIGURE 14. Denoising effect of (1.2) on the boats image where the original is corrupted by Gaussian white noise with $\sigma = 20$.

[19] B. Matthieu, P. Melchior, A. Outstaloup, and C. Ceyral, *Fractional Differentiation for Edge Detection*, Signal Processing, 83 (2003), pp. 2421–2432.

[20] M. Nitzberg and T. Shiota, *Nonlinear image smoothing with edge and corner enhancement*, Tech. Report 90-2, Harvard University, Cambridge, MA, 1990.

[21] ———, *Nonlinear image filtering with edge and corner enhancement*, IEEE Trans. Pattern Anal. and Machine Intelligence, 14 (1992), pp. 826–833.

[22] P. Perona and J. Malik, *Scale-space and edge detection using anistotropic diffusion*, IEEE Transactions Pattern Anal. Machine Intelligence, 12 (1990), pp. 161–192.

[23] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, *Scale-space and edge detection using anistotropic diffusion*, IEEE Transactions Image Process., 12 (2003), pp. 1338–1351.

[24] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.

[25] S. Tan and L. Jiao, *Multishrinkage: analytical form for a bayesian wavelet estimator based on the multivariate Laplacian model*, Optics Letters, 32 (2007), pp. 2583–2585.

[26] J. Weickert, *Anisotropic Diffusion in Image Processing*, Ph.D. Thesis, Universität Kaiserslautern, Kaiserlautern, 1996.

[27] ———, *Anisotropic Diffusion in Image Processing*, ECMI Series, Teubner Verlag, Stuttgart, 1998.

[28] Y. You and M. Kaveh, *Fourth order partial differential equations for noise removal*, IEEE Transaction on Image Processing, 9 (2000), pp. 1723–1730.