SIAM J. IMAGING SCIENCES
Vol. 8, No. 4, pp. 2487–2518

A Total Fractional-Order Variation Model for Image Restoration with Nonhomogeneous Boundary Conditions and Its Numerical Solution*

Jianping Zhang† and Ke Chen‡

Abstract. To overcome the weakness of a total variation based model for image restoration, various high order (typically second order) regularization models have been proposed and studied recently. In this paper we analyze and test a fractional-order derivative based total α-order variation model which can outperform the currently popular high order regularization models. There exist several previous works using total α-order variations for image restoration; however, first, no analysis has been done yet, and second, all tested formulations, differing from each other, utilize the zero Dirichlet boundary conditions which are not realistic (while nonzero boundary conditions violate definitions of fractional-order derivatives). This paper first reviews some results of fractional-order derivatives and then analyzes the theoretical properties of the proposed total α-order variational model rigorously. It then develops four algorithms for solving the variational problem—one based on the variational Split-Bregman idea and three based on direct solution of the discretize-optimization problem. Numerical experiments show that, in terms of restoration quality and solution efficiency, the proposed model can produce highly competitive results, for smooth images, to two established high order models: the mean curvature and the total generalized variation.

Key words. fractional-order derivatives, total α-order variation, PDE, image denoising, image inverse problems, optimization methods

AMS subject classifications. 62H35, 65N22, 65N55, 74G65, 74G75

DOI. 10.1137/14097121X

1. Introduction.

This paper presents a fractional-order derivative based regularizer for variational image restoration. It may be used for other imaging models such as image registration. Denote an observed image by $z = z(x)$, $x \in \Omega \subset \mathbb{R}^d$, where Ω is the bounded domain of the image with d space dimension and has a Lipschitz boundary. Here we consider $d = 2$ and mainly the image denoising problem with an additive noise; i.e., we assume $z = u + \eta_0$ with η_0 representing some unknown Gaussian noise of mean zero and deviation σ, but most results are applicable to $d > 2$ and other noise models.

1.1. Image inverse problem.

Restoring the unknown u (without any restrictions) from z is an inverse problem. According to the maximum likelihood principle [40], most image

*Received by the editors June 2, 2014; accepted for publication (in revised form) September 2, 2015; published electronically November 5, 2015. The research of the authors was supported by the UK EPSRC grant (EP/K036939/1) and the National Natural Science Foundation of China (NSFC project 11301447).
http://www.siam.org/journals/siims/8-4/97121.html

†Department of Mathematical Sciences, The University of Liverpool, Liverpool, Merseyside L6972L, United Kingdom, and School of Mathematics and Computational Science, Xiangtan University, Xiangtan, Hunan 411105, People's Republic of China (cmit@liv.ac.uk, jpzhang@xtu.edu.cn).
‡Centre for Mathematical Imaging Techniques and Department of Mathematical Sciences, The University of Liverpool, Liverpool, Merseyside L6972L, United Kingdom (k.chen@liv.ac.uk, www.liv.ac.uk/cmit).

processing problems involve solving the least-squares problem

$$\min_u \int_\Omega |P(u) - z|^2 dx, \tag{1.1}$$

measuring the fidelity to z. For example, $P(u) = u$ for image denoising, $P(u)$ takes the template image $T(x + u(x))$ (and $z = R(x)$ for a reference image) for image registration, and $P(u) = P_{\Omega_1}(u(x))$ for image inpainting with $\Omega_1 \subset \Omega$ the subdomain with missing data.

The problem (1.1) is in general ill-posed due to nonuniqueness; therefore, how to effectively solve it becomes a fundamental task in image sciences. The most popular idea is to regularize it so that the resulting well-posed problem admits a unique solution. The classical regularization technique by Tikhonov and Arsenin [76] is to add a smoothing regularization term into the energy functional to derive the minimization problem

$$\min_u \int_\Omega |P(u) - z|^2 dx + \lambda \int_\Omega |\nabla u|^2 dx, \tag{1.2}$$

where λ is a positive constant. This model cannot preserve image edges, though it is simple to use. The total variation (TV) model by Rudin, Osher, and Fatemi [68], or the ROF model,

$$\min_u \int_\Omega |\nabla u| dx, \qquad \int_\Omega |P(u) - z|^2 dx = \sigma^2, \quad P(u) = u, \tag{1.3}$$

is widely used, where σ is an estimate of the error η_0 between the noisy image z and the true data u. The ROF model preserves the image edges by seeking solutions of piecewise constant functions in the space of bounded variation (BV) functions. A variety of methods based on the TV regularization have been developed to deal with the imaging problems such as image restoration [1, 2, 10, 82], image registration [49, 38, 63], image decomposition [62, 39, 33], image inpainting [47, 41, 42, 23], and image segmentation [16, 77]. Restoring smooth images in some applications where edges are not the main features presents difficulties for the ROF model as it can yield the so-called blocky (staircase) effects. Another disadvantage of the model is the loss of image contrasts [53]. It should be remarked that the recently popular method by the iterative regularization technique [61] can reduce the staircasing effect and improve on the image contrast to some extent; additionally it provides a fast implementation.

1.2. High order regularization. To remedy the above-mentioned two drawbacks (staircasing and contrast), two types of alternative regularizer to TV regularization have been proposed in the literature. The first type introduces higher order regularization into image variational models [22, 24, 7, 53, 74, 32, 15, 84]. The mean curvature based variation denoising model was studied in [53, 54, 17, 84], where the regularized solution u is obtained by solving the fourth order Euler–Lagrange equation. Bredies, Kunisch, and Pock [15] proposed the total generalized variation (TGV) regularizer involving a linear combination of higher order derivatives and the TV of u to model the image denoising, while Chang, Tai, and Xing [26] considered a nonlinear combination of regularizer based on first and second order derivatives. For image inpainting, a high order regularization based on Euler's elastica of u is used in [24]. Similarly, Euler's elastica energy [57] and mean curvature [37, 30] are also proposed to transform the template image $T(x+u)$ to map the reference image $R(x)$ in image registration;

see also [36, 52]. The above-mentioned high order regularization methods are effective, but due to high nonlinearity efficient numerical solution is a major issue.

The second type introduces fractional-order derivatives, which are widely studied in other research subjects beyond image processing [3, 5, 6, 8, 85], into regularization of images. For example, Bai and Feng [11] first introduced fractional-order derivative into anisotropic diffusion equations for noise removal,

$$\frac{\partial u}{\partial t} = -D_x^{\alpha *}(c(|D^\alpha u|)D^\alpha u) - D_y^{\alpha *}(c(|D^\alpha u|)D^\alpha u), \tag{1.4}$$

where $c(\cdot)$ denotes the divergence parameter and $D_x^{\alpha *}$ denotes the adjoint operator of D_x^α, which may be viewed as a generalization of the Perona–Malik model. Although the above equation can be related to the Euler–Lagrange equations of an energy functional with the fractional derivative of the image intensity, generalizing commonly used PDE models, the energy minimization models are not studied as such. The discrete Fourier transform is used to implement the numerical algorithm assuming a periodic input image at its borders [11]. See also [46, 45, 50, 67] for more motivations and studies based on the above diffusion equation. Chen et al. [29, 28, 27] considered the fractional-order TV-L^2 image denoising model

$$\min_u \left\{ E(u) := \int_\Omega \sqrt{(D_x^\alpha u)^2 + (D_y^\alpha u)^2} d\Omega + \frac{\lambda}{2}\|u - f\|_2^2 \right\} \tag{1.5}$$

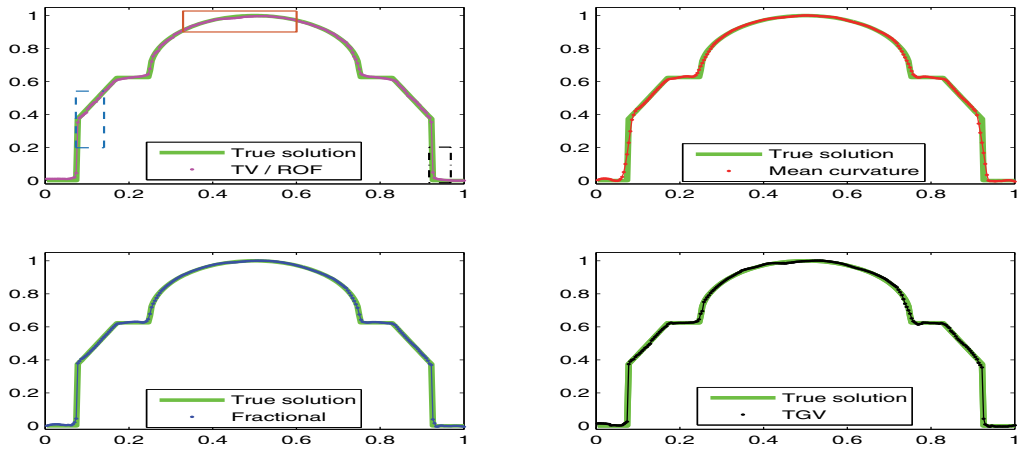
and numerically obtained improved denoising results over the Perona–Malik and ROF models; however, no analysis was given. There, they converted this primal formulation into a dual problem for the new dual variable $\mathbf{p} = (p_1, p_2)$ by $u = f - \mathrm{div}^\alpha \mathbf{p}/\lambda$ and used a dual algorithm using the gradient descent idea similar to the Chambolle method [18] for the ROF. In [81], the authors proposed a *discrete* optimization framework for the image denoising problem where the fractional order derivative is used to model the regularization term,

$$\min_u \left\{ \sum_{i,j=1}^{N} |(\nabla^\alpha u)_{i,j}| + 1/2 \sum_{j=0}^{L} 2^{-2js_j} |[\lambda(f-u)_j]|^2,\ 1 \le \alpha \le 2, 0 \le s_j \le 1 \right\}, \tag{1.6}$$

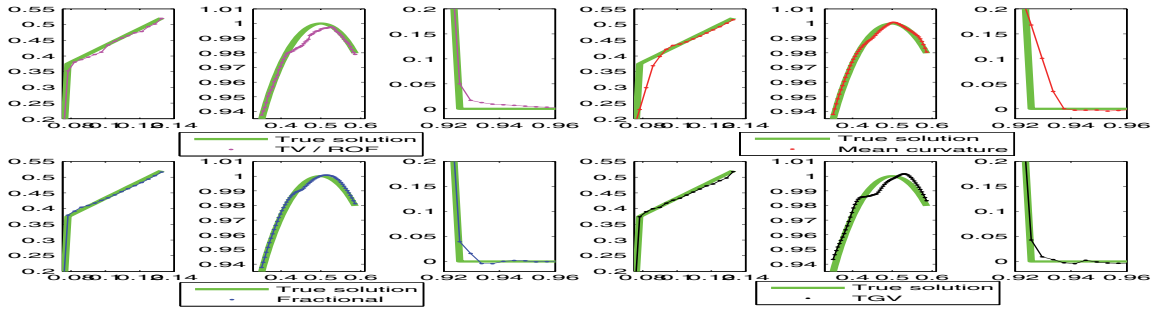
which is solved by an alternating projection algorithm. See also [21].

These works have reflected good performance of the fractional-order derivative in achieving a satisfactory compromise such as no staircasing and in preserving important fine-scale features such as edges and textures. Figure 1 presents the denoising results for testing the total fractional-order variation model in comparison to the previously mentioned three methods (namely, the TV, mean curvature, and the TGV). The true image was chosen to possess partly smooth and partly nonsmooth regions, while the test data is after adding zero mean Gaussian noise with variance $\sigma = 5/255$. As can be seen, there are the regions where the total variation type models (TV, TGV, and the total fractional-order variation) show better edge enhancement results in comparison to the mean curvature model (see Figure 1(b) for local representations of the left and right boxes in Figure 1(a)), while in smooth regions (see Figure 1(b) for local representations of the middle box in Figure 1(a)) the total fractional-order variation model restores a purely parabolic segment perfectly. The encouraging results motivated us to investigate this new model more closely.

There have been several other works involving discrete forms of an α-order derivative proposed to tackle image registration problems [78, 55] and image inpainting problems [83]. Comparing with the first type of high order models [15, 37, 30], a fractional-order model (type two) is less nonlinear and hence is more amenable to developing fast iterative solvers.



(a) Four models.



(b) Local representations.

Figure 1. *Comparisons of four models in denoising an artificial example.*

Clearly there is strong evidence to suggest that fractional-order derivatives may be effective regularizers for imaging applications. There is an urgent need to establish a rigorous theory for the total α-order variation based variational model so that further applications to image inverse problems can be considered in a systematic way.

1.3. Our contributions. This work is substantially different from previous studies. We mainly focus on the continuous total α variation-based model (instead of discrete formulation), its analysis, and associated numerical algorithms.

Our contributions are fourfold:

- We analyze properties of the total α-order variation, laying foundations for applications to imaging inverse problems as a regularizer.
- We give a new method for treating nonzero Dirichlet boundary conditions which rep-

resents a generalization of similar results that existed only in one to two dimensions.
- We establish the convexity, the solvability, and a solution theory for the total α-order variation model to make it more advantageous to work with than high order and nonconvex counterparts (such as a mean curvature based model) which are not gradient based and whose solutions are not as theoretically well known.
- We propose and test four solution algorithms (respectively, the Split-Bregman based algorithm, the forward-backward algorithm, the Nesterov accelerated method, and the fast iterative shrinkage-thresholding algorithm (FISTA)) to solve the underlying total α-order variation model. We also compare with related models.

We hope our work will motivate further studies and facilitate future applications of the α-order variation based regularizer to other imaging problems in the community.

The rest of the paper is organized as follows. Section 2 reviews the definitions and basic properties of the fractional-order derivative. Section 3 first defines the total α-order variation and the space of functions of fractional-order bounded variations. In this space, it then analyzes the the existence and the uniqueness of the solution of the total α-order variation based model for denoising. In section 4, a boundary condition regularization method for treating nonzero Dirichlet boundary conditions is proposed to effectively employ and compute the fractional-order derivatives of an image. Section 5 first discusses the discretization of the fractional-order derivatives by a finite difference method and presents a Split-Bregman scheme for effective solution. Section 6 takes the alternative discretize-optimize solution approach and develops three optimization-based algorithms (the forward-backward algorithm, the Nesterov accelerated method, and FISTA) to solve the image denoising model. Experimental results are shown in section 7, and the paper is concluded with a summary in section 8.

2. Review of fractional-order derivatives.

This section reviews definitions and simple properties of a fractional-order derivative which has a long history and may be considered as a generalization of the integer order derivatives. Three popular definitions to be reviewed are the Riemann–Liouville (RL), the Grünwald–Letnikov (GL), and the Caputo definitions [56, 60, 64].

In this paper, a fraction $\alpha \in \mathbb{R}^+$ is assumed to lie in between two integers $n-1, n$; i.e., $0 \leq \ell = n-1 < \alpha < n$ and a fractional α-order differentiation at point $x \in \mathbb{R}$ is denoted by the differential operator $D^{\alpha}_{[a,x]}$, where a and x are the bounds of the integral over a one-dimensional (1D) computational domain. Undoubtedly, the gamma function is very important for the study of the fractional derivative, which is defined by the integral [64]

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1}\, dt.$$

One of the basic properties is that $\Gamma(z+1) = z\Gamma(z)$ and hence $\Gamma(n) = n!$. Before introducing formal definitions, we review the following informative but classical example.

Example 1. Abel's integral equation, with

$$\frac{1}{\Gamma(\alpha)} \int_0^x \frac{\psi(\tau)}{(x-\tau)^{1-\alpha}} d\tau = f(x), \quad x > 0, \tag{2.1}$$

has the solution given by the well-known formula

$$\psi(x) = \frac{1}{\Gamma(1-\alpha)}\frac{d}{dx}\int_0^x \frac{f(\tau)}{(x-\tau)^\alpha}d\tau, \quad x > 0. \tag{2.2}$$

This example helps us to understand the formal definitions of fractional derivatives. In fact, for $0 < \alpha < 1$, (2.1) taking on the form $I^\alpha_{[0,x]}\psi(x) := D^{-\alpha}_{[0,x]}\psi(x) = f(x)$ is called the fractional α-order left RL *integral* of $\psi(x)$, and (2.2) taking on the form $D^\alpha_{[0,x]}f(x) = \psi(x)$ is defined as the fractional α-order left RL *derivative* of $f(x)$. As operators, under suitable conditions [69], we have $D^{-\alpha}_{[0,x]}D^\alpha_{[0,x]} = I$, where I denotes the identity operator.

The first definition of a general order α derivative is the left-sided RL derivative

$$D^\alpha_{[a,x]}f(x) = \frac{1}{\Gamma(n-\alpha)}\left(\frac{d}{dx}\right)^n \int_a^x \frac{f(\tau)d\tau}{(x-\tau)^{\alpha-n+1}}. \tag{2.3}$$

Subsequently the right-sided RL and the Riesz–RL (central) fractional derivative are, respectively, given by

$$D^\alpha_{[x,b]}f(x) = \frac{(-1)^n}{\Gamma(n-\alpha)}\left(\frac{d}{dx}\right)^n \int_x^b \frac{f(\tau)d\tau}{(\tau-x)^{\alpha-n+1}}$$

and

$$D^\alpha_{[a,b]}f(x) = \frac{1}{2}\left(D^\alpha_{[a,x]}f(x) + (-1)^n D^\alpha_{[x,b]}f(x)\right).$$

The second definition is the GL left-sided derivative denoted by

$${}^GD^\alpha_{[a,x]}f(x) = \lim_{h\to 0}\frac{1}{h^\alpha}\sum_{j=0}^{[\frac{x-a}{h}]}(-1)^j\begin{pmatrix}\alpha\\ j\end{pmatrix}f(x-jh), \quad \begin{pmatrix}\alpha\\ j\end{pmatrix} = \frac{\alpha(\alpha-1)\dots(\alpha-j+1)}{j!}, \tag{2.4}$$

which resembles the definition for an integer order derivative, where $[\vartheta]$ is the integer such that $\vartheta - 1 < [\vartheta] \le \vartheta$. The third definition is the Caputo order α derivative defined by

$${}^CD^\alpha_{[a,x]}f(x) = \frac{1}{\Gamma(n-\alpha)}\int_a^x \frac{f^{(n)}(\tau)d\tau}{(x-\tau)^{\alpha-n+1}}, \tag{2.5}$$

where $f^{(n)}$ denotes the nth-order derivative of function $f(x)$. The right-sided derivative and the Riesz–Caputo fractional derivative are similarly defined by

$${}^CD^\alpha_{[x,b]}f(x) = \frac{(-1)^n}{\Gamma(n-\alpha)}\int_x^b \frac{f^{(n)}(\tau)d\tau}{(\tau-x)^{\alpha-n+1}}, \quad {}^CD^\alpha_{[a,b]}f(x) = \frac{1}{2}\left({}^CD^\alpha_{[a,x]}f(x) + (-1)^n{}^CD^\alpha_{[x,b]}f(x)\right).$$

When $\alpha = n-1$ is an integer, the above left-sided RL definition reduces to the usual definition for a derivative. One notes that when a function is $n-1$ times continuously differentiable and its nth derivative is integrable, the fractional derivatives by the above definitions are equivalent subject to homogeneous boundary conditions [64]. However, we do not require such equivalence for our image function u; refer to Remark 2.

Fractional derivatives have many interesting properties; below we review a few that are relevant to this work.

Linearity. For a fractional derivative $D^{\alpha}_{[a,x]}$, by any of the above three definitions, one has

$$D^{\alpha}_{[a,x]}(p\, f(x) + q\, g(x)) = p\, D^{\alpha}_{[a,x]} f(x) + q\, D^{\alpha}_{[a,x]} g(x)$$

for any fractional differentiable functions $f(x), g(x)$ and $p, q \in \mathbb{R}$. This property will be shortly used to prove convexity and to derive the first order optimal conditions.

Zero fractional derivatives. An integer derivative of an image u at pixels of flat regions may be close to zero, but the left RL derivative of a constant intensity function is not zero. One advantage of minimizing an RL derivative instead of the TV (image gradients) could be a nonconstant solution. It would be interesting to know the kinds of functions that have zero α-order derivatives.

Lemma 2.1 (singularity). *Assume that $D^{\alpha}_{[a,x]}$ is one of the above three fractional-order derivative operators. For any noninteger $\alpha > 0$ and $x > a$, there exists a nonconstant value function $f(\tau)$ in $(a, x]$ such that $D^{\alpha}_{[a,x]} f(x) = 0$.*

Proof. We give explicit constructions. Here we consider only the RL and Caputo derivatives; for the GL derivative, we can derive a similar conclusion through their equivalency.

1. Assume that $0 < \alpha < 1$; for some $x > 0$, if $f(\tau)$ is taken as

   $$f(\tau) = (x - 2\tau)(x - \tau)^{\alpha}$$

   for any $\tau \in (0, x]$ in Abel's inverse transform (2.2), then $D^{\alpha}_{[0,x]} f(x) = \psi(x) = 0$.
2. Assume that $\alpha > 1$; if $f(\tau)$ is taken as

   $$f(\tau) = (x - \tau)^{\alpha - 1}$$

   for any $\tau \in (a, x]$ in the α-order RL derivative, then $D^{\alpha}_{[a,x]} f(x) = 0$.
3. Assume that $\alpha > 0$ in the Caputo derivative definition; if $f(\tau)$ is taken as

   $$f(\tau) = (x - \tau)^{n-1}$$

   for any $\tau \in (a, x]$ in (2.5), then ${}^{C}D^{\alpha}_{[a,x]} f(x) = 0$.

Actually for any $\alpha > 0$, the left RL $D^{\alpha}_{[0,x]} f(x) = 0$ if $f(x) = x^{\alpha - k}$ for all $k = 1, 2, \ldots, 1 + \ell$ (note that $\ell = [\alpha] = n - 1$); refer to [48]. ∎

Remark 1. For our later applications in section 7, we take $1 < \alpha < 2$. Hence we have the left RL $D^{\alpha}_{[0,x]} f(x) = 0$ if $f(x) = x^{\alpha - 1}$ or $x^{\alpha - 2}$; i.e., $f(x) = x^{0.6}$ or $x^{-0.4}$ when $\alpha = 1.6$. For the Caputo derivative, ${}^{C}D^{\alpha}_{[0,x]} f(x) = 0$ if $f(x) = 1$ or x.

Boundary conditions. For the left RL derivative $D^{\alpha}_{[a,x]} f(x)$ of $f(x)$, one assumes that $f(a) = 0$ or $f(b) = 0$ for the right RL derivative; otherwise, there is a singularity at the endpoint. So the Riesz–RL derivative would require $f(a) = f(b) = 0$. One solution for nonzero Dirichlet boundary conditions for f would be to extract a linear approximation $g(x)$ (that coincides with f at $x = a, b$) and consider $D^{\alpha}_{[a,x]}(f(x) - g(x))$; however, there is no such method for the two-dimensional (2D) case. In Jumarie's work [51], a simple alternative is to modify the RL derivative to

$$D^{\alpha}_{[a,x]} f(x) = \frac{1}{\Gamma(n - \alpha)} \left(\frac{d}{dx}\right)^{n} \int_{a}^{x} \frac{f(\tau) - f(a)}{(x - \tau)^{\alpha - n + 1}} d\tau,$$

also ensuring that the new fractional derivative of a constant is equal to zero and removing the singularity at $x = a$ [8]. In section 4, we present one method for treating nonzero Dirichlet boundary conditions in two dimensions.

3. The total α-order variation and its related model. This section first studies the properties of the total α-order variation, second analyzes a total α-order variation based denoising model, and finally presents a numerical algorithm. For the classical TV based model, its solution lies in a suitable space called the function space $\mathrm{BV}(\Omega)$ of bounded variation [25, 70, 15]. As seen in Figure 1, the total fractional-order variation model can preserve both edges and smoothness of an image; we anticipate from the former that its solution should lie in a space similar to the BV space and from the latter that the smoothness is due to the nonlocal nature of the new regularizer.

It turns out that for total α-order variation using α-order derivatives, a suitable space is the space $\mathrm{BV}^\alpha(\Omega)$ of functions of α-bounded variation on Ω which will be defined and studied next. The work of this section is motivated by analysis of the TV [1, 9, 18] and of the total generalized variation (TGV) [15].

In variational regularization methods, integration by parts involves the space of test functions in addition to the main solution space. Before discussing the total α-order variation, we give the following definition.

Definition 3.1 (a space of test functions). *Let $\mathcal{C}^\ell(\Omega, \mathbb{R}^d)$ denote the space of ℓ-order continuously differentiable functions. Furthermore, for any $\mathcal{C}^\ell(\Omega, \mathbb{R}^d) \ni v : \Omega \mapsto \mathbb{R}^d$, if the $(\ell+1)$th order derivative $v^{(\ell+1)}$ is integrable and $\frac{\partial^i v(x)}{\partial n^i}|_{\partial\Omega} = 0$ for all $i = 0, 1, \ldots, \ell$, then v is a compactly supported continuous-integrable function in Ω. Therefore, the ℓ-compactly supported continuous-integrable function space is denoted by $\mathscr{C}_0^\ell(\Omega, \mathbb{R}^d)$.*

Definition 3.2 (total α-order variation). *Let K denote the space of special test functions*

$$K := \Big\{ \boldsymbol{\phi} \in \mathscr{C}_0^\ell(\Omega, \mathbb{R}^d) \;\Big|\; |\boldsymbol{\phi}(x)| \le 1 \text{ for all } x \in \Omega \Big\},$$

where $|\boldsymbol{\phi}| = \sqrt{\sum_{i=1}^d \phi_i^2}$. Then the total α-order variation of u is defined by

$$TV^\alpha(u) := \sup_{\boldsymbol{\phi} \in K} \int_\Omega \Big(- u \; div^\alpha \, \boldsymbol{\phi} \Big) dx,$$

where $div^\alpha \boldsymbol{\phi} = \sum_{i=1}^d \frac{\partial^\alpha \phi_i}{\partial x_i^\alpha}$ and $\frac{\partial^\alpha \phi_i}{\partial x_i^\alpha}$ denotes a fractional α-order derivative $D^\alpha_{[a,b]}\phi_i$ of ϕ_i along the x_i direction.

We note that $\mathrm{TV}^\alpha(u)$ is the same for any definition of $\frac{\partial^\alpha \phi_i}{\partial x_i^\alpha}$ because $\boldsymbol{\phi}$ satisfies the equivalence conditions. However, for our applications in this paper, $\frac{\partial^\alpha u}{\partial x_i^\alpha}$ is generally not the same for different fractional derivatives (not even in the distributional sense).

Based on the α-BV seminorm, the α-BV norm is defined by

$$\|u\|_{\mathrm{BV}^\alpha} = \|u\|_{L^1} + \mathrm{TV}^\alpha(u), \tag{3.1}$$

and further the space of functions of α-bounded variation on Ω can be defined by

$$\mathrm{BV}^\alpha(\Omega) := \Big\{ u \in L^1(\Omega) \;\Big|\; \mathrm{TV}^\alpha(u) < +\infty \Big\}. \tag{3.2}$$

Lemma 3.3 (lower semicontinuity). *Let $\{u^k(x)\}$ be a sequence from $BV^\alpha(\Omega)$ which converges in $L^1(\Omega)$ to a function $u(x)$. Then $TV^\alpha(u) \le \liminf_{k\to\infty} TV^\alpha(u^k)$.*

Proof. Since $u_k \in BV^\alpha(\Omega)$, for any $\boldsymbol{\phi}(x) \in \mathscr{C}_0^\ell(\Omega, \mathbb{R}^d)$ such that $|\boldsymbol{\phi}(x)| \le 1$ on Ω, then $\mathrm{div}^\alpha\, \boldsymbol{\phi}$ is bounded; hence

$$\int_\Omega \Big(- u\ \mathrm{div}^\alpha\ \boldsymbol{\phi}\Big) dx = \liminf_{k\to+\infty} \int_\Omega \Big(- u_k\ \mathrm{div}^\alpha\ \boldsymbol{\phi}\Big) dx \le \liminf_{k\to+\infty} \mathrm{TV}^\alpha(u_k)$$

from $u_k \to u$ in $L^1(\Omega)$. Taking $\sup_{\boldsymbol{\phi}(x)}$ in the above inequality, we have lower semicontinuity from $\mathrm{TV}^\alpha(u) \le \liminf_{k\to+\infty} \mathrm{TV}^\alpha(u_k)$ (*see* [1, 34] for the TV case). ∎

Lemma 3.4. *The space $BV^\alpha(\Omega)$ is a Banach space.*

Proof. First we can see that $\mathrm{BV}^\alpha(\Omega)$ is a normed space following immediately from the definitions of $\|u\|_{L^1(\Omega)}$ and total α-order variation $\mathrm{TV}^\alpha(u)$, so it remains only to prove completeness. Suppose $\{u^k\}$ is a Cauchy sequence in $\mathrm{BV}^\alpha(\Omega)$; then, by the definition of the norm, it must also be a Cauchy sequence in $L^1(\Omega)$. According to the completeness of $L^1(\Omega)$, there exists a function u in $L^1(\Omega)$ such that $u^k \to u$ in $L^1(\Omega)$.

Since $\{u^k\}$ is a Cauchy sequence in $\mathrm{BV}^\alpha(\Omega)$, $\|u^k\|_{\mathrm{BV}^\alpha}$ is bounded. Thus $\mathrm{TV}^\alpha(u^k)$ is bounded as $k \to \infty$, by the lower semicontinuity of $\mathrm{TV}^\alpha(u)$ in $\mathrm{BV}^\alpha(\Omega)$ space (see Lemma 3.3); one shows that $u \in \mathrm{BV}^\alpha(\Omega)$.

We shall show that $u^k \to u$ in $\mathrm{BV}^\alpha(\Omega)$. We know that for any $\epsilon > 0$ there exists a positive integer N such that $\|u^k - u^j\|_{BV^\alpha(\Omega)} < \epsilon$ for any $j, k > N$; hence one has $\mathrm{TV}^\alpha(u^k - u^j) < \epsilon$. Since $u^k \to u$ in $L^1(\Omega)$, thus $u^j - u^k \to u^j - u$ in $L^1(\Omega)$. Hence by Lemma 3.3,

$$\mathrm{TV}^\alpha(u^j - u) \le \liminf_{k\to\infty} \mathrm{TV}^\alpha(u^j - u^k) \le \epsilon,$$

which shows that $u^k \to u$ in $\mathrm{BV}^\alpha(\Omega)$; therefore, $\mathrm{BV}^\alpha(\Omega)$ is a Banach space. ∎

Remark 2. In the literature [64], the equivalence of different fractional derivatives requires stringent continuity conditions; e.g., one has ${}^C D^\alpha_{[a,b]}\eta(x) = D^\alpha_{[a,b]}\eta(x)$ in the test space $\mathscr{C}_0^{n-1}([a,b], \mathbb{R})$. However, for imaging applications (the objective function u), we do not require such equivalence.

To distinguish the two definitions, we shall continue using the superscript C for C derivative based quantities such as ${}^C\mathrm{div}^\alpha$ and ${}^C\nabla^\alpha$, while no superscript means that a quantity is based on the RL derivative.

For any positive integer $p \in \mathbf{N}^+$, let $W_p^\alpha(\Omega) = \left\{u \in L^p(\Omega) \;\middle|\; \|u\|_{W_p^\alpha(\Omega)} < +\infty\right\}$ be a function space embedding with the norm

$$\|u\|_{W_p^\alpha(\Omega)} = \left(\int_\Omega |u|^p dx + \int_\Omega |\nabla^\alpha u|^p dx\right)^{1/p}, \text{ where } \nabla^\alpha u = \left(\frac{\partial^\alpha u}{\partial x_1}, \ldots, \frac{\partial^\alpha u}{\partial x_d}\right)^T.$$

For any $\xi(x) \in W_1^\alpha([a,b])$ and $\eta(x) \in \mathscr{C}_0^{n-1}([a,b],\mathbb{R})$,

$$\begin{aligned}
&\int_a^b \xi(x) \cdot {}^C D_{[a,b]}^\alpha \eta(x) dx \\
&= (-1)^n \int_a^b \eta(x) \cdot D_{[a,b]}^\alpha \xi(x) dx \ + \sum_{j=0}^{n-1} (-1)^j D_{[a,b]}^{\alpha-n+j} \xi(x) \frac{\partial^{n-j-1}\eta(x)}{\partial x^{n-j-1}}\Big|_{x=a}^{x=b} \\
&= (-1)^n \int_a^b \eta(x) \cdot D_{[a,b]}^\alpha \xi(x) dx
\end{aligned} \tag{3.3}$$

gives the α-order integration by parts formula (see [4]).

Furthermore, applying (3.3) twice, we have shown the relationship

$$\int_\Omega u(x) \Big((-1)^{nC} \mathrm{div}^\alpha \Big) \boldsymbol{\phi}(x) dx = \int_\Omega \boldsymbol{\phi}(x) \cdot \nabla^\alpha u(x) dx, \tag{3.4}$$

where $u(x) \in W_1^\alpha(\Omega)$ and $\boldsymbol{\phi}(x) \in \mathscr{C}_0^\ell(\Omega, \mathbb{R}^d)$; clearly the operator $(-1)^{nC}\mathrm{div}^\alpha$ is the adjoint of operator ∇^α. Note that, for $\boldsymbol{\phi}(x) \in \mathscr{C}_0^\ell(\Omega, \mathbb{R}^d)$, we have ${}^C\mathrm{div}^\alpha \boldsymbol{\phi} = \mathrm{div}^\alpha \boldsymbol{\phi}$, which may not be true if $\boldsymbol{\phi}(x)$ is in a different space.

Proposition 3.5. *Assume that $u \in W_1^\alpha(\Omega)$; then $TV^\alpha(u) = \int_\Omega |\nabla^\alpha u| dx$.*

Proof. For any $\alpha > 0$, using the dual relationship (3.4), one can obtain that

$$\int_\Omega u(x) \mathrm{div}^\alpha \boldsymbol{\phi}(x) dx \ = (-1)^n \int_\Omega \boldsymbol{\phi}(x) \cdot \nabla^\alpha u(x) dx,$$

and in addition $|\boldsymbol{\phi}| \le 1$ in K implies that

$$\boldsymbol{\phi}_0(x) = \begin{cases} (-1)^n \nabla^\alpha u / |\nabla^\alpha u|, & |\nabla^\alpha u(x)| \neq 0, \\ 0 & \text{otherwise} \end{cases}$$

can maximize the functional $\int_\Omega \boldsymbol{\phi}(x) \cdot \nabla^\alpha u(x) dx = \int_\Omega |\nabla^\alpha u| dx$. By multiplying $\boldsymbol{\phi}_0$ by a suitable characteristic ℓ-compactly supported continuous function η_ϵ in Ω (e.g., $\eta_\epsilon \in \mathscr{C}_0^\ell(\Omega, \mathbb{R}^d)$) and then mollifying (see [1] for TV and [9, 43]), the new $\int_\Omega \boldsymbol{\phi}_\epsilon(x) \cdot \nabla^\alpha u(x) dx$ with $\boldsymbol{\phi}_\epsilon \in K$ is arbitrarily close to $\int_\Omega |\nabla^\alpha u| dx$ as $\epsilon \to 0$ [34]; hence one shows that $\mathrm{TV}^\alpha(u) = \int_\Omega |\nabla^\alpha u| dx$ by taking $\sup_{\boldsymbol{\phi}_\epsilon \in K} \int_\Omega u(x) \cdot \mathrm{div}^\alpha \boldsymbol{\phi}_\epsilon(x) dx \ = \sup_{\boldsymbol{\phi}_\epsilon \in K} (-1)^n \int_\Omega \boldsymbol{\phi}_\epsilon(x) \cdot \nabla^\alpha u(x) dx$. ∎

Remark 3. Since $u \in W_1^\alpha(\Omega)$ leads to $\mathrm{TV}^\alpha(u) = \int_\Omega |\nabla^\alpha u| dx$, in fact, it is easy to show that the lower semicontinuity $\int_\Omega |\nabla^\alpha u| dx \le \liminf_{k\to\infty} \int_\Omega |\nabla^\alpha u^k| dx$ holds in the space $W_1^\alpha(\Omega)$ similarly to the TV case [34].

Lemma 3.6. *The space $W_p^\alpha(\Omega)$ is a Banach space.*

Proof. The $p = 1$ case is clear. Now, for $p \neq 1$, let q satisfy $1/p + 1/q = 1$. To obtain the lower semicontinuity, taking $u \in W_p^\alpha(\Omega)$ and $\psi(x) \in \{\boldsymbol{\phi} \in \mathscr{C}_0^\ell(\Omega, \mathbf{R}^d) \mid \|\boldsymbol{\phi}(x)\|_{L^q(\Omega)} \le 1 \text{ for all } x \in \Omega\}$, the inequality

$$\begin{aligned}
\int_\Omega (-1)^n \nabla^\alpha u \psi dx = \int_\Omega u \mathrm{div}^\alpha \psi dx \ &= \liminf_{k\to+\infty} \int_\Omega u^k \mathrm{div}^\alpha \psi dx \ = \liminf_{k\to+\infty} \int_\Omega (-1)^n \nabla^\alpha u^k \psi dx \\
&\le \liminf_{k\to+\infty} \left(\int_\Omega |\nabla^\alpha u^k|^p dx \right)^{1/p} \left(\int_\Omega |\psi|^q dx \right)^{1/q} \le \liminf_{k\to+\infty} \left(\int_\Omega |\nabla^\alpha u^k|^p dx \right)^{1/p}
\end{aligned}$$

holds; further one has $\left(\int_\Omega |\nabla^\alpha u|^p dx\right)^{1/p} \le \liminf_{k\to+\infty} \left(\int_\Omega |\nabla^\alpha u^k|^p dx\right)^{1/p}$. Then we can deduce the result, following reasoning similar to that in the proof of Lemma 3.4. ∎

Lemma 3.7. *The following embedding results hold:* $W_2^\alpha(\Omega) \subseteq W_1^\alpha(\Omega) \subseteq BV^\alpha(\Omega) \subseteq L^1(\Omega)$.

Proof. First, from the definitions of $BV^\alpha(\Omega)$ and $W_p^\alpha(\Omega)$, we can see that $BV^\alpha(\Omega) \subset L^1(\Omega)$ and $W_1^\alpha(\Omega) \subset L^1(\Omega)$. Second, for any $f \in W_1^\alpha(\Omega)$ and $\boldsymbol{\phi} \in K$, we have

$$\int_\Omega f \mathrm{div}^\alpha \boldsymbol{\phi}\, dx = (-1)^n \int_\Omega \boldsymbol{\phi}(x) \cdot \nabla^\alpha f(x) dx \le \|\nabla^\alpha f\|_{L^1(\Omega)} < +\infty,$$

i.e., $f \in BV^\alpha(\Omega)$ or $W_1^\alpha(\Omega) \subseteq BV^\alpha(\Omega)$. Finally, $W_2^\alpha(\Omega) \subseteq W_1^\alpha(\Omega)$ follows from $L^2(\Omega) \subseteq L^1(\Omega)$. ∎

Lemma 3.8. *The functional* $TV^\alpha(u)$ *is convex.*

Proof. The proof follows the linearity of fractional-order derivatives, and the positively homogeneous and subadditive properties of $\mathrm{TV}^\alpha(u)$. ∎

Theory for a total α-order variation model. We are now ready to analyze model (1.5) or the total α-order variation model in a more precise form:

$$\min_{u\in \mathrm{BV}^\alpha(\Omega)} \left\{ E(u) := \mathrm{TV}^\alpha(u) + \frac{\lambda}{2} F(u) \right\}, \qquad F(u) = \int_\Omega |u - z|^2 dx. \tag{3.5}$$

To focus on the total α-order variation model in $\Omega = (0,1)\times(0,1) \subset \mathbb{R}^2$, we assume $1 < \alpha < 2$; the following theorem establishes convexity of the minimization problem (3.5).

Theorem 3.9 (convexity). *The functional $E(u)$ in $BV^\alpha(\Omega)$ is convex for $\lambda \ge 0$ and strictly convex if $\lambda > 0$.*

Proof. Since $F(u)$ is a strictly convex functional, the proof follows from Lemma 3.8. ∎

If a Banach space X is reflexive (separable), then every bounded sequence in X (in X^*) has a weakly ($weak^*$) convergent subsequence [80, Prop. 38.2]. Although $\mathrm{BV}^\alpha(\Omega)$ is not reflexive, it is the dual of a separable space. Therefore, we can give the following definition.

Definition 3.10 (a *weak** topology). *In $BV^\alpha(\Omega)$, a weak $BV^\alpha - w^*$ topology is defined as*

$$u_j \xrightarrow[BV^\alpha - w^*]{*} u \iff u_j \xrightarrow[L^1(\Omega)]{} u \quad \text{and} \quad \int_\Omega \boldsymbol{\phi} \cdot \nabla^\alpha u_j\, dx \longrightarrow \int_\Omega \boldsymbol{\phi} \cdot \nabla^\alpha u\, dx$$

for all $\boldsymbol{\phi}$ in $\mathscr{C}_0^0(\Omega, \mathbb{R}^d)$.

From Definition 3.10, we may derive the weak compactness of $\mathrm{BV}^\alpha(\Omega)$ on the $weak^*$ topology. This, combined with the weak lower semicontinuity of $E(u)$ and boundedness of Banach space $\mathrm{BV}^\alpha(\Omega)$ (i.e., u is bounded in Banach space $\mathrm{BV}^\alpha(\Omega)$), yields the following result.

Theorem 3.11 (existence). *The functional $E(u) : BV^\alpha(\Omega) \to \mathbb{R}$ has a minimum.*

Proof. The proof follows reasoning similar to the proof of [80, Prop. 38.12(d)]). ∎

Theorem 3.12 (uniqueness). *The functional $E(u)$ has a unique minimizer in $BV^\alpha(\Omega)$ when $\lambda > 0$.*

Proof. The convexity result of Theorem 3.9 leads to uniqueness of solutions. Refer to [80, Theorem 47C]. ∎

We remark that similar existence and uniqueness theories of the TV problem can be found in [1, 19, 9].

4. Nonzero Dirichlet boundary conditions and regularization. The standard definitions for fractional derivatives require a function to have zero Dirichlet boundary conditions due to end singularity, but for imaging applications such conditions are unrealistic and too restrictive. To obtain the system for finding the unknown intensities u at inner nodes of a discretization grid, we have to use boundary conditions, but the difficulties caused by them in fractional derivative computations would be hard to overemphasize; inaccurate boundary conditions can easily lead to the oscillations near boundaries, so proper treatment of the boundary conditions for problems involving fractional derivatives is crucial.

In this section, we shall reduce nonzero Dirichlet boundary conditions to zero ones so that standard definitions and our introduced algorithms become applicable. The basic idea of boundary regularization is to introduce an auxiliary unknown function which satisfies the zero boundary conditions. In this way, the nonzero boundary conditions move to the right-hand side of the equation as a new known quantity.

We recall that, in the 1D case, if the boundary conditions are nonzero,

$$u(0) = a,\ u(1) = b,$$

we can reduce them to zero boundary conditions by introducing an auxiliary function $e(x) = a(1-x) + bx$. Precisely taking $\bar{u}(x) = u(x) - e(x)$ [65], we then have

$$\bar{u}(0) = 0,\ \bar{u}(1) = 0, \quad \bar{u}'(0) = \bar{u}'(1) = 0,$$

and a Neumann boundary condition is imposed by artificially extending the boundary values; i.e., $e'(0) = e'(1) = 0$ on $\partial\Omega$.

Below we generalize the above 1D idea to the 2D case, assuming that the four corners of the solution are given or accurately estimated:

$$u(0,0) = a,\ u(0,1) = b,\ u(1,0) = c,\ u(1,1) = d.$$

With a, b, c, d known, at any image point $(x, y) \in \Omega$, a bilinear auxiliary function satisfying the above four conditions, $e_1(x,y) = a + (c-a)x + (b-a)y + (d+a-c-b)xy$, can be constructed to lead to

$$\bar{u}(x,y) = u(x,y) - e_1(x,y), \tag{4.1}$$

which takes zero values at all four corners.

If boundary conditions $u(0,y) = a_1(y),\ u(1,y) = a_2(y),\ u(x,0) = b_1(x),\ u(x,1) = b_2(x)$ at $\partial\Omega$ are known a priori, then we can easily verify that

$$\begin{aligned} \bar{a}_1(y) &:= \bar{u}(0,y) = a_1(y) - e_1(0,y), \quad \bar{a}_2(y) := \bar{u}(1,y) = a_2(y) - e_1(1,y), \\ \bar{b}_1(x) &:= \bar{u}(x,0) = b_1(x) - e_1(x,0), \quad \bar{b}_2(x) := \bar{u}(x,1) = b_2(x) - e_1(x,1) \end{aligned}$$

define the new Dirichlet conditions for $\bar{u}(x,y)$.

We can achieve zero conditions at the edges using the auxiliary function $e_2(x,y) = ((1-x)\bar{a}_1(y) + x\bar{a}_2(y)) + ((1-y)\bar{b}_1(x) + y\bar{b}_2(x))$. It is clear to see that the new image $\tilde{u}(x,y) = u(x,y) - e_1(x,y) - e_2(x,y)$ satisfies

$$\tilde{u}(x,y)|_{\partial\Omega} = 0. \tag{4.2}$$

Remark 4. It remains to address the question of how to obtain estimates of $u(x,y)$ at corners and edges:

1. The true intensities $a := u(0,0)$, $b := u(1,0)$, $c := u(1,0)$, and $d := u(1,1)$ in four corner points are unknown a priori. To build the auxiliary function $e_1(x,y)$, the solutions approximating them should be solved from the observed image $z(x,y)$ by the local smoothing or other simple techniques.
2. Similarly, the true edge intensities $a_1(y)$, $a_2(y)$, $b_1(x)$, and $b_2(x)$ are also not given. A reconstruction step on boundary $\partial\Omega$ is necessary in order to capture a robust solution. To proceed, we can apply a 1D model.

According to Remark 4, we can propose a complete procedure for regularizing boundary conditions for 2D variational image inverse problems in edges and corners.

- First, we restore image intensities in four corner points from an observed image z. A natural technique would be a local smoothing operator for the region of corner points. The oscillations could also be reduced by many variational methods to local regions.
- Second, in order to reconstruct four edges from the restored intensities $z(0,y)$, $z(1,y)$, $z(x,0)$, and $z(x,1)$, the total α-order variation regularization would be used to solve four 1D inverse problems, i.e., solve an equation like (3.5):

$$\min_u \left\{ E^{1D}(u) = \int_a^b \left| \frac{d^\alpha u}{dx^\alpha} \right| dx + \frac{\lambda^{1D}}{2} \int_a^b (u-z)^2 dx \right\}. \tag{4.3}$$

Thus, through $e_1(x,y), e_2(x,y)$, we see that (4.1) reduces to finding the new image $\bar{u}(x,y)$ with zero Dirichlet conditions, and hence the standard definitions of fractional derivatives for $\bar{u}(x,y)$ apply.

5. Discretization and Split-Bregman algorithm.

Since solution uniqueness of our variational model (3.5) is resolved, we now consider how to seek a numerical solution of the total α-order variation model. We first reformulate it in preparation for employment of an efficient solver and then discuss some discretization details (by finite differences) before presenting our Algorithm 1.

5.1. A Split-Bregman formulation.

Inspired by Goldstein and Osher's Split-Bregman work [44], we introduce a special and new variable $\boldsymbol{d}(x) = (d_1(x), d_2(x))^T$ to the total α-order variation based model (3.5) to derive the following constrained optimization problem:

$$\min_{u,\boldsymbol{d}} \int_\Omega |\boldsymbol{d}| dx + \frac{\lambda}{2} F(u), \quad \text{s.t. } \boldsymbol{d} = \nabla^\alpha u. \tag{5.1}$$

To enforce the constraint condition, we transfer it into the Bregman formulation

$$\begin{aligned}
(u^{k+1}, \boldsymbol{d}^{k+1}) =& \min_{u,\boldsymbol{d}} \int_\Omega |\boldsymbol{d}| dx + \frac{\lambda}{2} F(u) - \int_\Omega \langle \boldsymbol{p}_d^k, \boldsymbol{d} - \boldsymbol{d}^k \rangle dx \\
& - \int_\Omega \langle \boldsymbol{p}_u^k, u - u^k \rangle dx + \frac{\mu}{2} \int_\Omega |\boldsymbol{d} - \nabla^\alpha u|^2 dx, \\
\boldsymbol{p}_u^{k+1} =& \boldsymbol{p}_u^k - \mu (\nabla^\alpha)^T (\nabla^\alpha u^{k+1} - \boldsymbol{d}^{k+1}), \\
\boldsymbol{p}_d^{k+1} =& \boldsymbol{p}_d^k - \mu (\boldsymbol{d}^{k+1} - \nabla^\alpha u^{k+1}).
\end{aligned}$$

The above iterative scheme can be simplified to the two-step algorithm [44, 71, 72]:

$$\min_{u,\boldsymbol{d}} \int_\Omega |\boldsymbol{d}|dx + \frac{\mu}{2}\int_\Omega \left|\boldsymbol{d} - \nabla^\alpha u + \frac{\boldsymbol{p}}{\mu}\right|^2 dx + \int_\Omega |\boldsymbol{p}|^2 dx + \frac{\lambda}{2}F(u) \tag{5.2}$$

with the multiplier updated by iteration $\boldsymbol{p}^{k+1} = \boldsymbol{p}^k - \gamma(\boldsymbol{d} - \nabla^\alpha u)$, where $\boldsymbol{p}(x) = (p_1(x), p_2(x))^T$. Here the two main subproblems of (5.2) are

$$\begin{aligned} &\text{Subproblem } \boldsymbol{d} && \min_{\boldsymbol{d}} \int_\Omega |\boldsymbol{d}|dx + \frac{\mu}{2}\int_\Omega \left|\boldsymbol{d} - \nabla^\alpha u + \frac{\boldsymbol{p}}{\mu}\right|^2 dx; \\ &\text{Subproblem } u && \min_u J(u) := \frac{\mu}{2}\int_\Omega \left|\boldsymbol{d} - \nabla^\alpha u + \frac{\boldsymbol{p}}{\mu}\right|^2 dx + \frac{\lambda}{2}F(u). \end{aligned} \tag{5.3}$$

Further note that subproblem $\boldsymbol{d}$ has a closed-form solution [44], while subproblem u is determined by the associated Euler–Lagrange equation as shown below.

Theorem 5.1. *Let $u(x)$ be a minimizer of functional $J(u)$ from (5.3). Then $u(x)$ satisfies the first order optimal condition*

$$(-1)^n \mu\, {}^C div^\alpha \left(\nabla^\alpha u - \boldsymbol{d} - \frac{\boldsymbol{p}}{\mu}\right) + \lambda(u - z) = 0 \tag{5.4}$$

with one of the following sets of boundary conditions:

(i) fixed: $u(x)\big|_{\partial\Omega} = b_1(x)$, *and* $\frac{\partial u(x)}{\partial n}|_{\partial\Omega} = b_2(x)$,

(ii) homogeneous: $D^{\alpha-2}(\nabla^\alpha u - \boldsymbol{d} - \frac{\boldsymbol{p}}{\mu}) \cdot \mathbf{n}|_{\partial\Omega} = 0$, $D^{\alpha-1}(\nabla^\alpha u - \boldsymbol{d} - \frac{\boldsymbol{p}}{\mu}) \cdot \mathbf{n}|_{\partial\Omega} = 0$,

where $\mathbf{n}$ denotes the unit outward normal and ${}^C div^\alpha$ denotes the divergence operator based on the C derivative.

Proof. Refer to the appendix. ∎

5.2. Discretization of the fractional derivative. Before introducing the finite difference discretization of the fractional derivative, we define a spatial partition (x_k, y_l) (for all $k = 0, 1, \ldots, N+1; l = 0, 1, \ldots, M+1$) of image domain Ω. Assume u has a zero Dirichlet boundary condition (practically, we apply the regularization method section 4 first before discretization). Here we mainly consider the discretization of the α-order fractional derivative at the inner point (x_k, y_l) (for all $k =, 1, \ldots, N; l = 0, 1, \ldots, M$) on Ω along the x-direction by using the approach

$$\begin{aligned} D^\alpha_{[a,b]} f(x_k, y_l) &= \frac{\delta_0^\alpha f(x_k, y_l)}{h^\alpha} + O(h) \;= \frac{1}{2}\left(\frac{\delta_-^\alpha f(x_k, y_l)}{h^\alpha} + \frac{\delta_+^\alpha f(x_k, y_l)}{h^\alpha}\right) + O(h) \\ &= \frac{1}{2}\left(h^{-\alpha}\sum_{j=0}^{k+1} \omega_j^\alpha f^l_{k-j+1} + h^{-\alpha}\sum_{j=0}^{N-k+2} \omega_j^\alpha f^l_{k+j-1}\right) + O(h), \end{aligned} \tag{5.5}$$

which is applicable to both the RL and C derivatives [66, 79], where $f^l_s = f_{s,l}$, $\omega_j^{(\alpha)} = (-1)^j \binom{\alpha}{j}$, $j = 0, 1, \ldots, N+1$, and

$$\omega_0^{(\alpha)} = 1; \quad \omega_j^{(\alpha)} = \left(1 - \frac{1+\alpha}{j}\right)\omega_{j-1}^{(\alpha)} \quad \text{for } j > 0.$$

Alternative discretization for fractional derivatives in the Fourier space can be found in [11, 45].

Observe from (5.5) that the first order estimate of the α-order fractional $D^{\alpha}_{[a,b]}f(x_k, y_l)$ along the x-direction at the point (x_k, y_l) with a fixed y_l is a linear combination of $N+2$ values $\{f_0^l, f_1^l, \ldots, f_N^l, f_{N+1}^l\}$.

After incorporating the zero boundary condition in the matrix approximation of fractional derivative, all N equations of fractional derivatives along the x direction in (5.5) can be written simultaneously in the matrix form (denote $w = \omega_0^{\alpha} + \omega_2^{\alpha}$)

$$\begin{pmatrix} \delta_0^{\alpha} f(x_1, y_l) \\ \delta_0^{\alpha} f(x_2, y_l) \\ \vdots \\ \vdots \\ \delta_0^{\alpha} f(x_N, y_l) \end{pmatrix} = \underbrace{\frac{1}{2h^{\alpha}} \begin{pmatrix} 2\omega_1^{\alpha} & w & \omega_3^{\alpha} & \cdots & \omega_N^{\alpha} \\ w & 2\omega_1^{\alpha} & \ddots & \ddots & \vdots \\ \omega_3^{\alpha} & \ddots & \ddots & \ddots & \omega_3^{\alpha} \\ \vdots & \ddots & \ddots & 2\omega_1^{\alpha} & w \\ \omega_N^{\alpha} & \cdots & \omega_3^{\alpha} & w & 2\omega_1^{\alpha} \end{pmatrix}}_{B_N^{\alpha}} \underbrace{\begin{pmatrix} f_1^l \\ f_2^l \\ \vdots \\ \vdots \\ f_N^l \end{pmatrix}}_{f}. \tag{5.6}$$

From the definition of fractional-order derivative (5.5), for any $1 < \alpha < 2$, the coefficients $\omega_k^{(\alpha)}$ have the following properties [64, 79]:

(1) $\omega_0^{(\alpha)} = 1,\ \omega_1^{(\alpha)} = -\alpha < 0,$ (2) $1 \geq \omega_2^{(\alpha)} \geq \omega_3^{(\alpha)} \geq \cdots \geq 0,$
(3) $\sum_{k=0}^{\infty} \omega_k^{(\alpha)} = 0,$ (4) $\sum_{k=0}^{m} \omega_k^{(\alpha)} \leq 0\ (m \geq 1).$

Hence by the Gershgorin circle theorem, one can derive that matrix B_N^{α} in (5.6) is a symmetric and negative definite Toeplitz matrix (i.e., $-B_N^{\alpha}$ is a positive definite Toeplitz matrix).

We recall that the Kronecker product $A \otimes B$ of the $p \times q$ matrix $A = [a_{ij}]$ and the $n \times m$ matrix $B = [b_{rt}]$ is the $np \times mq$ matrix having the block structure $A \otimes B := [a_{ij}B]$. Further, vector $(A \otimes B)x$ can be computed by matrix scheme BXA^T (i.e., $[(A \otimes B)x]_s = [BXA^T]_{j,i}$ with $s = (i-1)m + j$), where the $m \times q$ matrix X is the reshape of the vector x along its column.

Let $U \in \mathbb{R}^{N \times M}$ denote the solution matrix at all nodes $(kh_x; lh_y)$, $k = 1, \ldots, N$, $l = 1, \ldots, M$, corresponding to x-direction and y-direction spatial discretization nodes. Denote by $\vec{u} \in \mathbb{R}^{\mathbb{N}M \times 1}$ the ordered solution vector of U. The direct and discrete analogue of differentiation of arbitrary α order derivative is

$$u_x^{(\alpha)} = (I_M \otimes B_N^{\alpha})\vec{u} = B_x^{(\alpha)}\vec{u},$$

where $u_x^{(\alpha)} = (u_{11}^{(\alpha)}, \ldots, u_{N1}^{(\alpha)}, u_{12}^{(\alpha)}, \ldots, u_{NM}^{(\alpha)})^T$, $\vec{u} = (u_{11}, \ldots, u_{N1}, u_{12}, \ldots, u_{NM})^T$. Similarly, the αth order y-direction derivative of $u(x; y)$ is approximated by

$$u_y^{(\alpha)} = B_y^{(\alpha)}\vec{u} = (B_M^{\alpha} \otimes I_N)\vec{u}, \quad \text{where } u_y^{(\alpha)} = (u_{11}^{(\alpha)}, \ldots, u_{1M}^{(\alpha)}, u_{21}^{(\alpha)}, \ldots, u_{NM}^{(\alpha)})^T.$$

5.3. The Split-Bregman algorithm. In discrete form, we are ready to state the discretized equations in structured matrix form. The discrete scheme of (5.4) is given by

$$(-1)^n\mu\Bigg(\Big((B_x^{(\alpha)})^T(B_x^{(\alpha)}\vec{u})+(B_y^{(\alpha)})^T(B_y^{(\alpha)}\vec{u})\Big)-\Big((B_x^{(\alpha)})^T\vec{d}_1+(B_y^{(\alpha)})^T\vec{d}_2\Big)$$
$$-\frac{1}{\mu}\Big((B_x^{(\alpha)})^T\vec{p}_1+(B_y^{(\alpha)})^T\vec{p}_2\Big)\Bigg)+\lambda(\vec{u}-\vec{z})=0$$

with discretizations $\vec{d}_i=\big(d_{11}^i,\ldots,d_{N1}^i,d_{12}^i,\ldots,d_{NM}^i\big)^T$ and $\vec{p}_i=\big(p_{11}^i,\ldots,p_{N1}^i,p_{12}^i,\ldots,p_{NM}^i\big)^T$ of vectors $\boldsymbol{d}$ and $\boldsymbol{p}$ $(i=1,2)$. A matrix approximation equation is given as

(5.7)
$$\underbrace{\Big((B_N^\alpha)^T(B_N^\alpha U)+U(B_M^\alpha)^TB_M^\alpha\Big)+\bar{\lambda}U}_{WU}=\underbrace{\bar{\lambda}Z+\Big((B_N^\alpha)^TD_1+D_2B_M^\alpha\Big)+\frac{1}{\mu}\Big((B_N^\alpha)^TP_1+P_2B_M^\alpha\Big)}_{F},$$

where D_i and P_i are $N\times M$-size reshape matrices of vectors $\vec{d}_i$ and $\vec{p}_i$ for $i=1,2$, $\bar{\lambda}=(-1)^n\lambda/\mu$. The following justifies the use of a conjugate gradient method for $WU=F$.

Theorem 5.2. *The weighted matrices inner product $\langle WU,U\rangle=\sum_{ij}(\sum_k W_{ik}U_{kj})U_{ij}$ is positive for any matrix $U\neq 0$, where W is a known positive definite operator.*

Proof. For any matrix $U\neq 0$, it is easy to show that

$$\begin{aligned}\langle WU,U\rangle&=\Big\langle\Big((B_N^\alpha)^T(B_N^\alpha U)+U(B_M^\alpha)^TB_M^\alpha\Big)+\bar{\lambda}U,\ U\Big\rangle\\&=\langle B_N^\alpha U,B_N^\alpha U\rangle+\langle U(B_M^\alpha)^T,U(B_M^\alpha)^T\rangle+\bar{\lambda}\langle U,U\rangle\\&=\|B_N^\alpha U\|_F^2+\|U(B_M^\alpha)^T\|_F^2+\bar{\lambda}\|U\|_F^2>0,\end{aligned}$$

which completes the proof. ∎

An implementation of this method is summarized below.

Algorithm 1 (Split-Bregman iterations (PDE-SB)).

Step 1. *Boundary regularization for an observed image z;*

Step 2. *Given initial matrices $P_1^{k=0}$, $P_2^{k=0}$, and $U^{k=0}$;*

Step 3. *Solve subproblem* **d***: Compute the auxiliary matrix $\binom{D_1}{D_2}$ from the closed-form solution*

$$\begin{pmatrix}D_1\\D_2\end{pmatrix}^{k+1}=shrink\left(\begin{pmatrix}B_N^\alpha U^{k+1}\\U^{k+1}(B_M^\alpha)^T\end{pmatrix}+\begin{pmatrix}P_1\\P_2\end{pmatrix}^k,\frac{1}{\mu}\right)$$

by solving the Moreau–Yosida problem with the l^1 regularization;

Step 4. *Solve subproblem u: Find the solution U^{k+1} of (5.7) with an effective parameter λ μ by the CG method;*

Step 5. *Update $\binom{P_1}{P_2}^{k+1}=\binom{P_1}{P_2}^k+\gamma\left(\binom{B_N^\alpha U^{k+1}}{U^{k+1}(B_M^\alpha)^T}-\binom{D_1}{D_2}^{k+1}\right)$ with $\gamma\in(0,1]$;*

Step 6. *Check the stopping condition;*

– *If $|U^k-U^{k+1}|<\epsilon$,*
stop and return $U^*:=U^{k+1}$;

– *else*
  $k := k + 1$, *go back to Step* 1*;*
– *end*

Step 7. *Accept the correct solution* U *from boundary regularization.*

6. Optimization based numerical methods. As many variational models are increasingly solved by the discretize-optimize approach, we now present three related algorithms for model (5.1) after applying a finite difference discretization. In this section, we assume that we have the zero Dirichlet boundary conditions for u mainly to simplify the notation.

As in section 5, the αth order derivative $u_x^{(\alpha)}$ of $u(x;y)$ along all x-direction nodes in Ω can be given by matrix $B_N^\alpha U$, and similarly for $U(B_M^\alpha)^T$ for the y-direction (as U is the solution matrix).

Define $\langle U, V\rangle = \sum_{ij} U_{ij}V_{ij}$, and let $V_1 = \{p \mid 0 \le p \le 1\}$, $V_2 = \{p \mid |p| \le 1\}$. Then using the discrete setting introduced above, the discretized problem of model (3.5) is

$$\min_{U\in V_1} \max_{\Phi\in V_2} G(U, D^*\Phi) + \frac{\lambda}{2} H(U), \tag{6.1}$$

where $H(U) = \sum_{ij}(U_{ij} - Z_{ij})^2$ and $G(U, D^*\Phi) = \langle U, D^*\Phi\rangle = \sum_{ij} U_{ij}\left(B_N^\alpha\Phi_1 + \Phi_2(B_M^\alpha)^T\right)_{ij}$, due to $D^*\Phi = B_N^\alpha\Phi_1 + \Phi_2(B_M^\alpha)^T$. We also have the adjoint relationship $\langle U, D^*\Phi\rangle = \langle DU, \Phi\rangle$ with $DU = (B_N^\alpha U, U(B_M^\alpha)^T)$ and $\Phi = (\Phi_1, \Phi_2)$. In line with the literature, this model can be denoted by the convex optimization problem in a generic notation by

$$\min_x \{f_1(x) + f_2(x)\}, \quad \text{i.e., } \min_{x,y} \{f_1(x) + f_2(y)\} \text{ s.t. } x = y, \tag{6.2}$$

where one views $x = U$, $f_1(x) = \max_{\Phi\in V_2} G(U, D^*\Phi)$, $f_2(x) = H(U)$. We also need the notation

$$\operatorname{prox}_{f_1}^{\lambda}(x^k) := \arg\min_{x\in V_1} \left\{ f_1(x) + \frac{1}{2\lambda}\|x - x^k\|^2 \right\},$$

where f_1 can be any other convex function and $\lambda > 0$.

To solve (6.2) by the methods to be presented, computation of the proximal point $\operatorname{prox}_{f_1}^{\lambda}(x)$ is a major and nontrivial step. We consider how to compute it when $D = \nabla^\alpha$, borrowing ideas from solving a similar problem of TV regularization. In a dual setting, Chambolle [18, 20] first proposed a discrete dual method by optimizing a cost function consisting of two variants [18, 29]. Recently, one variant of this scheme was employed in [29] to effectively solve a fractional image model by a dual transform. The other variant is used in [27].

Define two projections as

$$\operatorname{Proj}_{V_1}(p) = \begin{cases} 0, & p < 0, \\ p, & 0 \le p \le 1, \\ 1, & 1 \le p, \end{cases} \qquad \operatorname{Proj}_{V_2}(p) = \frac{p}{\max(1, \|p\|)}.$$

Note that $\frac{\partial G(x, D^*\Phi)}{\partial x} = D^*\Phi$ and that the optimal solution is

$$x = \operatorname{prox}_{f_1}^{\gamma}(x^k) = \operatorname{Proj}_{V_1}(\bar{x}), \tag{6.3}$$

where $\bar{x} = x^k - \gamma D^*\Phi$ and Φ is unknown. Based on methods of Chambolle [18] and Beck and Teboulle [13], we see that (6.3) can be used to reduce the min-max problem

$$\min_{x\in V_1}\left\{\max_{\Phi\in V_2}\langle x, D^*\Phi\rangle + \frac{1}{2\gamma}\|x-x^k\|^2\right\}$$

to the dual problem $\max_{\Phi\in V_2}\langle \mathrm{Proj}_{V_1}(\bar{x}), D^*\Phi\rangle + \frac{1}{2\gamma}\|\mathrm{Proj}_{V_1}(\bar{x}) - x^k\|^2$ and further to

$$\begin{aligned}
&\langle \mathrm{Proj}_{V_1}(\bar{x}), D^*\Phi\rangle + \frac{1}{2\gamma}\|\mathrm{Proj}_{V_1}(\bar{x}) - x^k\|^2 \quad = \langle \mathrm{Proj}_{V_1}(\bar{x}), D^*\Phi\rangle \\
&\quad + \frac{1}{2\gamma}\|\mathrm{Proj}_{V_1}(\bar{x}) - x^k + \gamma D^*\Phi\|^2 - \frac{1}{2\gamma}\|\gamma D^*\Phi\|^2 - \frac{1}{2\gamma}2\langle \mathrm{Proj}_{V_1}(\bar{x}) - x^k, \gamma D^*\Phi\rangle \\
&= \frac{1}{2\gamma}\|\mathrm{Proj}_{V_1}(\bar{x}) - (x^k - \gamma D^*\Phi)\|^2 - \frac{1}{2\gamma}\|\gamma D^*\Phi\|^2 + \frac{1}{2\gamma}2\langle x^k, \gamma D^*\Phi\rangle \\
&= \frac{1}{2\gamma}\|\mathrm{Proj}_{V_1}(\bar{x}) - (x^k - \gamma D^*\Phi)\|^2 - \frac{1}{2\gamma}\left(\|\gamma D^*\Phi\|^2 - 2\langle x^k, \gamma D^*\Phi\rangle + \|x^k\|^2\right) + \frac{1}{2\gamma}\|x^k\|^2 \\
&= \frac{1}{2\gamma}\|\mathrm{Proj}_{V_1}(\bar{x}) - (x^k - \gamma D^*\Phi)\|^2 - \frac{1}{2\gamma}\|x^k - \gamma D^*\Phi\|^2 + \frac{1}{2\gamma}\|x^k\|^2 \\
&= \frac{1}{2\gamma}\left(\|\bar{x} - \mathrm{Proj}_{V_1}(\bar{x})\|^2 - \|\bar{x}\|^2 + \|x^k\|^2\right),
\end{aligned} \tag{6.4}$$

i.e., $\max_{\Phi\in V_2}\langle \mathrm{Proj}_{V_1}(\bar{x}), D^*\Phi\rangle + \frac{1}{2\gamma}\|\mathrm{Proj}_{V_1}(\bar{x}) - x^k\|^2 = -\frac{1}{2\gamma}\min_{\Phi\in V_2} h(\Phi)$, where $h(\Phi) = \|x^k - \gamma D^*\Phi\|^2 - \|(x^k - \gamma D^*\Phi) - \mathrm{Proj}_{V_1}(x^k - \gamma D^*\Phi)\|^2 - \|x^k\|^2 = \|\bar{x}\|^2 - \|\bar{x} - \mathrm{Proj}_{V_1}(\bar{x})\|^2 - \|x^k\|^2$.

Below we consider the operator $S(\bar{x}) = \|\bar{x} - \mathrm{Proj}_{V_1}(\bar{x})\|^2 = \inf_y\{\delta_{V_1}(y) + \frac{1}{2\gamma}\|y - \bar{x}\|^2\}$. Since its gradient is $\nabla_{\bar{x}} S(\bar{x}) = 2(\bar{x} - \mathrm{Proj}_{V_1}(\bar{x}))$, we get

$$\nabla_\Phi h(\Phi) = -2\gamma D(\mathrm{Proj}_{V_1}(x^k - \gamma D^*\Phi)).$$

The minimization problem $\min_{\Phi\in V_2} h(\Phi)$ can be solved to obtain the Φ-update as follows:

- $\bar{\Phi} = \Phi^n - L(h)\nabla_\Phi h(\Phi^n)$,
- $\Phi^{n+1} = \mathrm{Proj}_{V_2}(\bar{\Phi}) = \mathrm{Proj}_{V_2}(\Phi^n + 2L(h)\gamma D(\mathrm{Proj}_{V_1}(x^k - \gamma D^*\Phi^n)))$,

using the gradient projection scheme of $h(\Phi)$ [13]. Here $L(h) \le 16\gamma^2$ is the Lipschitz constant. Finally, the proximal point $\mathrm{prox}^\gamma_{f_1}(x^k)$ is given by (6.3) once Φ is obtained; see also [13].

6.1. Forward-backward algorithm. Various applications in sparse optimizations stimulated the search for simple and efficient first order methods. The forward-backward scheme for (6.2) is based (as the name suggests) on recursive application of an explicit forward step with respect to f_2, i.e.,

$$\min_x\left\{\underbrace{f_2(x^k) + \langle \nabla f_2(x^k), x - x^k\rangle}_{l(x)} + \frac{1}{2\gamma}\|x - x^k\|^2\right\},$$

and followed by an implicit backward step with respect to f_1, i.e.,

$$\min_x\left\{f_1(x) + \frac{1}{2\gamma}\|x - x^k\|^2\right\}. \tag{6.5}$$

The scheme decouples the contributions of the functions f_1 and f_2 in a gradient descent step [12]. The scheme is also known under the name of proximal gradient methods [73, 31, 71, 12], since the implicit step relies on the computation of the so-called proximity operator.

The forward-backward algorithm is summarized as follows.

Algorithm 2 (forward-backward algorithm (FB) [12]).

- *Fix initial* x_0, *set* $\epsilon \in [0, \min\{1, 1/\beta\}]$, β *(a Lipschitz parameter);*
- *For* $k \geq 0$
  - *Step* 1. $\gamma_k \in [\epsilon, 2/\beta - \epsilon]$, $\lambda_k \in [\epsilon, 1]$;
  - *Step* 2. $y_k = prox_l^{\gamma_k}(x_k)$;
  - *Step* 3. $x_{k+1} = prox_{f_1}^{\gamma_k}(y_k)$;
  - *Step* 4. $x_{k+1} = x_k + \lambda_k(x_{k+1} - x_k)$;
  - *Step* 5. *Stop when* $\|x_{k+1} - x_k\|$ *is small enough; otherwise continue.*

6.2. Nesterov's method. As a gradient based method, though simple, the above method can exhibit a slow speed of convergence. For this reason, Nesterov [58] proposed an improved gradient method aiming to accelerate and modify the classical forward-backward splitting algorithm, while achieving an almost optimal convergence rate. As a consequence of this breakthrough, a few recent works have followed up on the idea and improved techniques for some specific problems in signal or image processing [13, 10].

Recently Nesterov [59] presented an accelerated multistep version, which converges as $O(\frac{1}{r^2})$ (r is the iteration number). For a problem of type (6.2), this new method introduced a composite gradient mapping. We now show the algorithm as follows.

Algorithm 3 (Nesterov accelerated method (Nesterov [59])).

- *Fix initial* x_0, b_0, *set* $y_0 = x_0$, *and* β *(a Lipschitz parameter);*
- *For* $k \geq 0$
  - *Step* 1. *Find* $a = a_k$ *from the quadratic equation* $\frac{a^2}{2(b_k+a)} = \frac{1+b_k}{\beta}$;
  - *Step* 2. $v_{=} prox_{f_1}^{b_k}(x_k - y_k)$;
  - *Step* 3. $z_{k+1} = \frac{b_k x_k + a_k v_k}{b_k + a_k}$;
  - *Step* 4. $x_{k+1} = prox_{f_1}^{\beta^{-1}}(z_{k+1} - \beta^{-1}\nabla f_2(z_{k+1}))$;
  - *Step* 5. $y_{k+1} = y_k + a_k \nabla f_2(x_{k+1})$;
  - *Step* 6. $b_{k+1} = b_k + a_k$;
  - *Step* 7. *Stop when* $\|x_{k+1} - x_k\|$ *is small enough; otherwise continue.*

6.3. FISTA method. Beck and Teboulle [13, 14] proposed a fast iterative shrinkage thresholding algorithm (FISTA) to solve the image denoising and deblurring model. The method applies the idea of Nesterov to the forward-backward splitting framework, resulting in the same optimal convergence rate as Nesterov's method but wider applicability. It can be applied to a variety of practical problems arising from sparse signal recovery, image processing, and machine learning and hence has become a standard algorithm.

Applying it to (6.2), we obtain Algorithm 4 below.

Algorithm 4 (FISTA (Beck–Teboulle [12, 13, 14])).

- *Fix initial* x_0, *set* $z_0 = x_0$ *and* $t_0 = 1$, β *(a Lipschitz parameter);*
- *For* $k \geq 0$
  - *Step* 1. $y_k = z_k - \beta^{-1}\nabla f_2(z_k)$;

Step 2. $x_{k+1} = prox_{f_1}^{\beta^{-1}}(y_k)$;
Step 3. $t_{k+1} = \frac{1+\sqrt{4t_k^2+1}}{2}$;
Step 4. $z_{k+1} = x_k + (1 + \frac{t_k - 1}{t_k})(x_{k+1} - x_k)$;
Step 5. *Stop when* $\|x_{k+1} - x_k\|$ *is small enough; otherwise continue.*

7. Numerical results. Finally, we present some numerical results from using the four presented algorithms denoted by

PDE-SB: PDE based Split-Bregman (Algorithm 1);
Opti-FB: Optimization based forward-backward (Algorithm 2);
Opti-Nesterov: Optimization based Nesterov accelerated method (Algorithm 3);
Opti-FISTA: Optimization based FISTA (Algorithm 4),

and their comparisons with related methods. In all tests, an initial solution is the noisy image $z(x, y)$; the algorithms solving the diffusion equation or optimization problem are stopped after achieving a relative residual of 10^{-4} or a relative error of 10^{-8} within 1000 outer and 15 inner iterations. Here we mainly compare the solution's visual quality, the *signal-to-noise ratio* (snr), and the *peak signal-to-noise ratio* (psnr) values which are given by

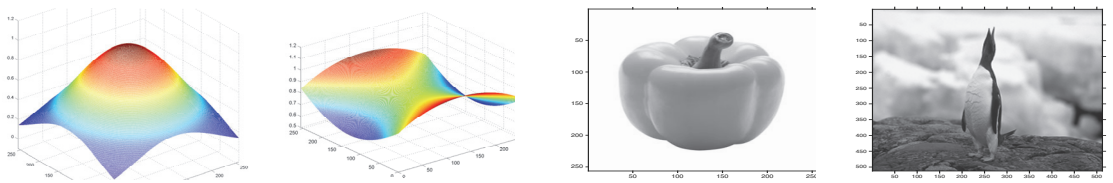
$$snr(u, u^*) = 10 \log_{10} \frac{\|u^* - mean(u^*)\|_F^2}{\|u - u^*\|_F^2}, \; psnr(u, u^*) = 10 \log_{10} \frac{n_x n_y \left(\max_{i,j} u_{i,j}^* \right)^2}{\|u - u^*\|_F^2},$$

where $mean(u^*)$ is an average value of the true image u^*, and n_x and n_y denote the size of the test image z. It should be noted, however, that these valuations do not always correlate with human perception. In real life situations, the two measures are also not possible because the true image is not known.

In general, an optimization problem may be solved many times to select a suitable regularization parameter λ or to optimize the solution for the underlying inverse problem; a solution is accepted when some stopping criterion is satisfied. It remains to carry out a systematic study on our new model as in [82] for the TV model. However, we shall use the best (numerical) λ for all models in the following tests.

For denoising, $F(u) = (u - z)^2$ is the L^2 measure between the solution u and the observed image z. To intuitively describe the denoising ability, four sets of data will be used in this part (also see Figure 2):

P1: *Problem* 1 - Parabolic surfaces;
P2 : *Problem* 2 - Saddle surface;
P3: *Problem* 3 - Pepper;
P4 : *Problem* 4 - Penguin.



(a) **P1**. (b) **P2**. (c) **P3**. (d) **P4**.

Figure 2. *Test datasets.*

Though our framework is readily applicable to image deblurring and image registration, here we only present denoising results.

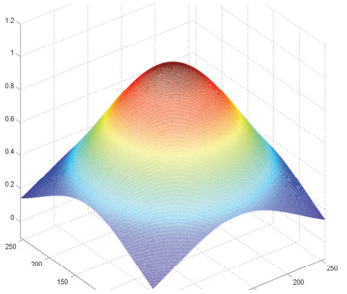
7.1. Performance comparisons of boundary regularization. We first test the idea from section 4. On the one hand, the variational framework seeks the boundary conditions of a nonzero Dirichlet or a Neumann type on $\partial\Omega$, and also real images do have nonzero boundary conditions. On the other hand, fractional-order derivatives require homogeneous boundary conditions (as used in the work of many authors) due to end singularity. In order to aid accurate computation of the discretized fractional-order derivative, in our work, a boundary processing technique (section 4) has been proposed to transform nonzero boundary conditions of observed data z into zero boundary conditions; hence a consequent matrix approximation to the fractional derivative operator $D^{\alpha}_{[0,\,1]}$ can use a zero Dirichlet boundary condition.

Here we test the performance and effectiveness of our boundary regularization against no regularization. The experiment is carried out on **P1**. Parabolic surfaces are as shown in Figure 3; i.e., Figure 3(a) shows a synthetic image of size 256×256 and range [0, 1], and the noisy image in Figure 3(b) shows an added zero mean value Gaussian random noise with a mean variance $\delta = \frac{15}{256}$. For the boundary regularization case, the approximation $u|_{\partial\Omega}$ from the observed data $z|_{\partial\Omega}$ is from applying the 1D fractional-order variation model as described in section 4. The treated case is named "treated," and the results are depicted in Figures 3(d) and 3(f)), where $psnr$= 47.53 and snr=35.43. The solution obtained from assuming zero boundary conditions for u is named "nontreated" and is depicted in Figures 3(c) and 3(e), where $psnr = 23.69$ and $snr = 10.38$. Clearly our boundary regularization treatment is effective.

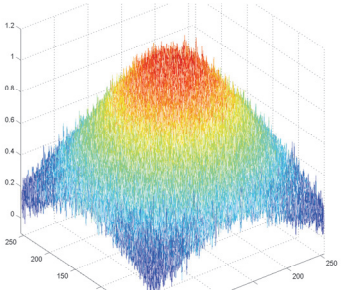
7.2. Comparisons of Algorithms 1–4. In Table 1, we compare the restoration quality (via $psnr$ and snr) of the four algorithms. There, all four test datasets are used. In the cases of synthetic images **P1** and **P2** with noise variation $\delta = \frac{10}{255}$, λ is taken as 12000 and 3800, respectively, and $\alpha = 1.6$. In the cases of natural images **P3** and **P4** with noise variation $\delta = \frac{5}{255}$, λ is taken as 18000 and 20000, respectively, and $\alpha = 1.4$. One can see from Table 1 that Opti-Nesterov and PDE-SB perform similarly in terms of the best restoration quality (via psnr and snr). However, in efficiency (computation time *cpu(s)*), Opti-FISTA and PDE-SB are the best, while Opti-Nesterov takes more computational time than the other three algorithms. Evidently, overall, PDE-SB (Algorithm 1) shows the most consistently good performance in the considered tested cases.

7.3. Sensitivity tests for α and λ. Since our model (3.5) contains two main parameters, α for the order of differentiation and λ as the coupling parameter for a regularized inverse problem, it is of interest to test their sensitivity on the restoration quality. Here we shall test all algorithms' sensitivity using the image **P2** (saddle surface of size 256×256, after adding zero mean value Gaussian random noise image of range [0, 1] and $\delta = \frac{10}{256}$).

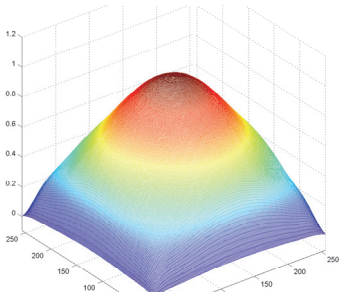
Varying λ in a large range from 400 to 60000, all four algorithms are tested on this synthetic image. The results are shown in Figures 4(a) and 4(c) for different stopping criteria. (**GSC:** The general stopping criteria with the relative residual 10^{-4}, relative error 10^{-8}, and inner iterations 10. **SSC:** The strong stopping criteria with the relative residual 10^{-7}, relative error 10^{-10}, and inner iterations 25.) While we see results different from the TV denoising



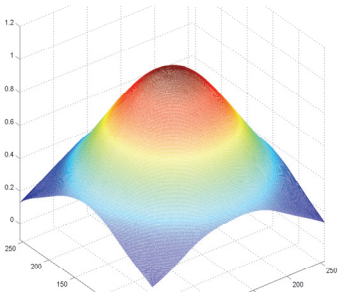
(a) True surface.



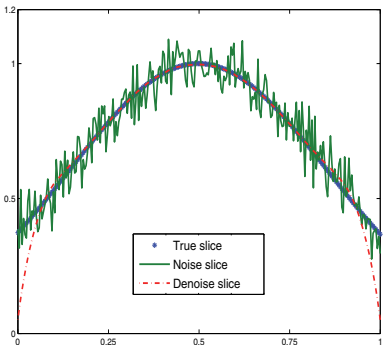
(b) Noise surface.



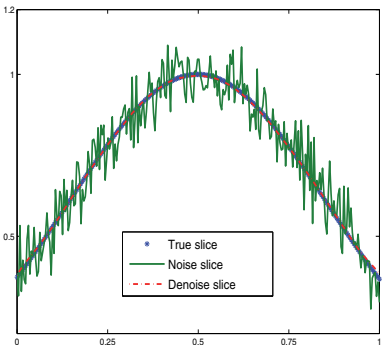
(c) Nontreated $r = 10.93\%$.



(d) Treated $r = 0.66\%$.



(e) Slice for nontreated: Bad.



(f) Slice for treated: Good.

Figure 3. *Test for* **P1***—Comparisons between treated and nontreated cases for nonzero boundary conditions* ($\delta = \frac{15}{256}$) *using PDE-SB. The treated case has* psnr $= 47.53$ *and* snr $= 35.43$*, while the nontreated case has* psnr $= 23.69$ *and* snr $= 10.38$*. Clearly our boundary regularization (section 4) is effective, while direct application of a fractional model leads to incorrect boundary restoration. Here the error* $r = \|u - u^*\|_F / \|u^*\|_F$.

case, where the regularization parameter λ is crucial for restoration quality [82], Figures 4(a) and 4(c) show that our total α-order variation regularization model still obtains a satisfactory solution for a large range of λ; this is a pleasing observation. Of course, there exists the issue of an optimal choice.

Next, we vary $\alpha \in (1, 2)$ from 1.1 to 1.9. Figures 4(b) and 4(d) show the four algorithms' restored results responding to two stopping conditions **GSC** and **SSC**. As represented, the smaller α leads to the blocky (staircase) effects in u, and the larger α will make solution u

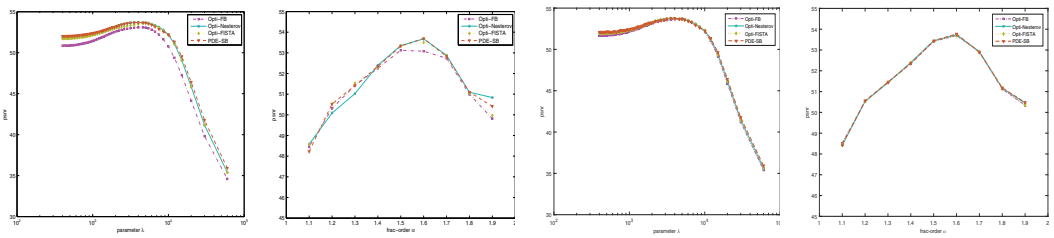
Table 1

Comparisons of optimizing algorithms, where $\delta = \frac{10}{255}$ for saddle and parabolic surfaces and $\delta = \frac{5}{255}$ for pepper and penguin images.

	Opti-FB			Opti-Nesterov			Opti-FISTA			PDE-SB		
	snr	*psnr*	*cpu(s)*	*snr*	*psnr*	*cpu(s)*	*snr*	*psnr*	*cpu(s)*	*snr*	*psnr*	*cpu(s)*
P1	36.78	50.09	16.83	36.91	50.22	27.23	36.94	50.24	16.71	36.96	50.27	14.53
P2	31.04	53.08	17.28	31.61	53.67	28.43	31.46	53.50	18.14	31.63	53.69	15.09
P3	29.21	43.29	15.96	29.40	43.49	16.09	29.40	43.49	9.75	29.48	43.56	8.16
P4	25.19	38.01	14.68	25.35	38.15	16.27	25.34	38.15	8.62	25.34	38.14	8.45

too smooth along the x_1- and x_2-directions, respectively. For denoising, our test suggests that $\alpha = 1.6$ is suitable for smooth problems because the diffusion coefficients are almost isotropic in all regions, leading to smooth deformation fields, and $\alpha = 1.4$ is appropriate for nonsmooth problems because the diffusion coefficients are close to zero in regions representing large gradients of the fields, allowing discontinuities at those regions.

We should emphasize that the stopping criteria have impacted the actual numerical implementation. In other words, if we drop the limit on the maximal number of inner iterations and relative residuals (and relative errors), some methods take too long but obtain more satisfactory results.



(a) **GSC**: psnr vs. λ (b) **GSC**: psnr vs. α (c) **SSC**: psnr vs. λ (d) **SSC**: psnr vs. α

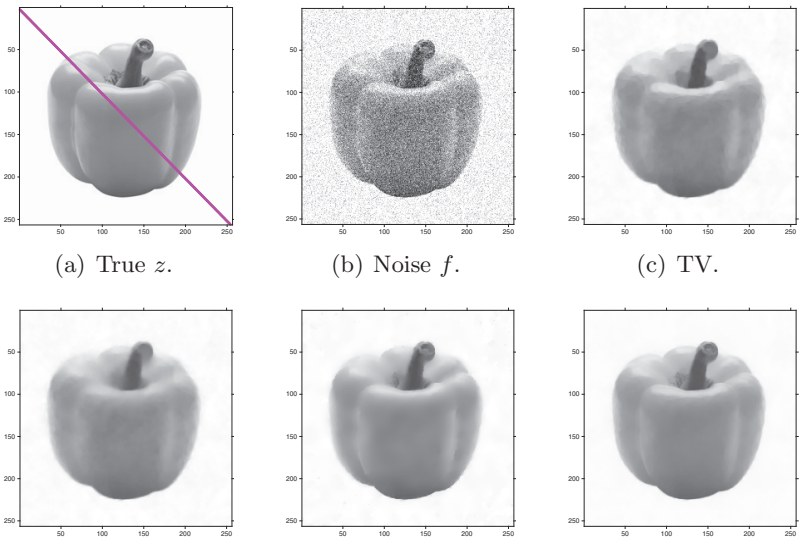
Figure 4. *Sensitivity test of Algorithms 1–4 to parameters λ (with fixed $\alpha = 1.6$) and α (with fixed $\lambda = 3800$) in the cases of the **GSC** and **SSC** stopping conditions.*

7.4. Comparisons with other nonfractional variational models. In this test, we compare our total α-variation model (PDE-SB) with three popular methods for variational image denoising. The first compared approach is naturally the TV model proposed by Rudin, Osher, and Fatemi [68] because the total α-order variation model in this work is inspired by it. The second compared work is the mean curvature model [75], which also addresses the problem of restoring a good result for a smooth image; their approach is different from ours since it is focused on higher order regularization and a multigrid method. See also [54, 17, 84]. The third compared approach is the TGV model [15] involving a combination of first order and higher order derivatives to reduce the staircasing effect of the bounded variation functional.

In Table 2, we first compare the restoration quality (via psnr and snr) and efficiency (computation times *cpu(s)*) of four approaches by testing the artificial images (**P1** (Parabolic

surface), **P2** (Saddle surface)) and the natural images (**P3** (Pepper), **P4** (Penguin)); in each approach relevant parameters are shown in Table 2. We see that, with the empirically optimal parameters λ^*, the differences between the four models are very small, though our new and convex model is slightly better. In other tests where such optimal parameters are not used, our new model performs more robustly and better.

In order to present more visual differences, some stronger regularization parameters ($\lambda^*/2$) and higher noise variations (with the noise level $\delta = \frac{30}{255}$) are tested. Such visual differences in solutions can be seen in Figures 5 and 7, respectively, for restoring noisy and natural images **P3** (Pepper, Figure 5(b)) and **P4** (Penguin, Figure 7(b)). While ROF denoising leads to blocky results, the mean curvature model performs better in the smooth regions but exhibits more smooth results near discontinuities. The TGV model leads to further improvements over the aforementioned models. The total fractional-order variation model leads to significantly better results. The reason is that the new model tries to approximate the image based on affine functions or nonlocal high order smooth functions, which is clearly better in this case; in other words, our approach is more effective in eliminating the noise for smooth images and is competitive with high order methods; in efficiency the new approach (PDE-SB) is much faster than the TGV and the mean curvature. We also plot four error results between the restored and true images along a diagonal (magenta) line in Figure 5(a) for comparison with Figure 6 and present zoomed-in surfaces inside the rectangle (magenta in Figure 7(a)). In Figure 8, we see that PDE-SB produces the best restored surface, which shows a major advantage (or better performance) of our total α-order variation model (3.5) when the test image is smooth, and even when the contrast between meaningful objects and the background is low.



(a) True z. (b) Noise f. (c) TV.

(d) Mean curvature. (e) TGV. (f) Our approach.

Figure 5. *Comparison* I*: Comparisons of our PDE-SB with TV, mean curvature, and TGV models.*

8. Conclusions. The total α-order variation regularization model with fractional-order derivative is potentially useful in modeling all imaging problems. In this paper we rigorously

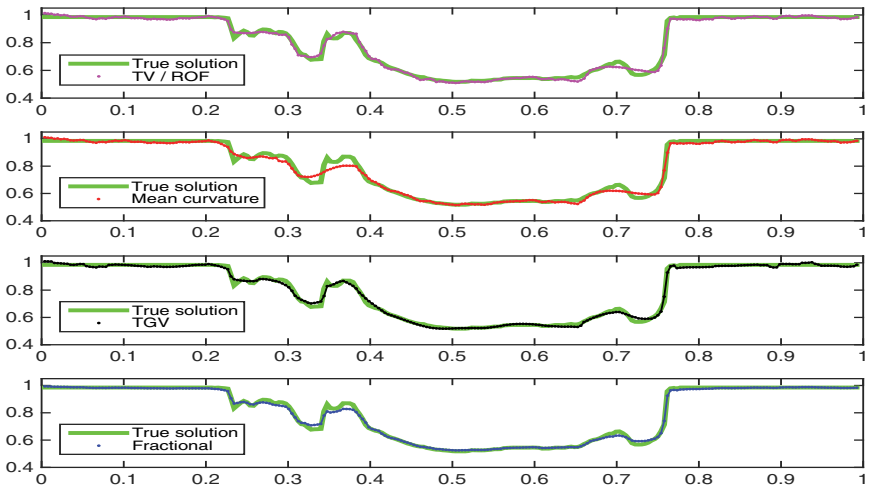
Table 2

*Comparisons of four models in restoration quality: The total α variation model (3.5), mean curvature [54, 17, 75], TV [68, 35, 25], and TGV [15] models for synthetic images (**P1** in Figure 2(a) and **P2** in Figure 2(b)) and natural images (**P3** in Figure 5(a) and **P4** in Figure 7(a)) with different noise variances $\delta_j = \frac{\hat{\delta}_j}{255}$ (correspondingly we use λ_j). We first fix $\mu = 1.1$, $\gamma = 1$, $\alpha = 1.6$ for **P1**–**P2**, $\alpha = 1.15$ and $\alpha = 1.1$ for **P3**–**P4**, respectively, in the total α-order variation model, and $\gamma = 19$, $\beta = 10^{-5}$ in the mean curvature model, and two weight parameters of the first and second order terms in the TGV model ($\nu_0 = 1$, $\nu_1 = 2$ for **P1**–**P2**; $\nu_0 = 1$, $\nu_1 = 0.5$ for **P3**–**P4**); other parameters are as shown on the* "para" *rows. λ^{1D} from (4.3) is required by the new model only.*

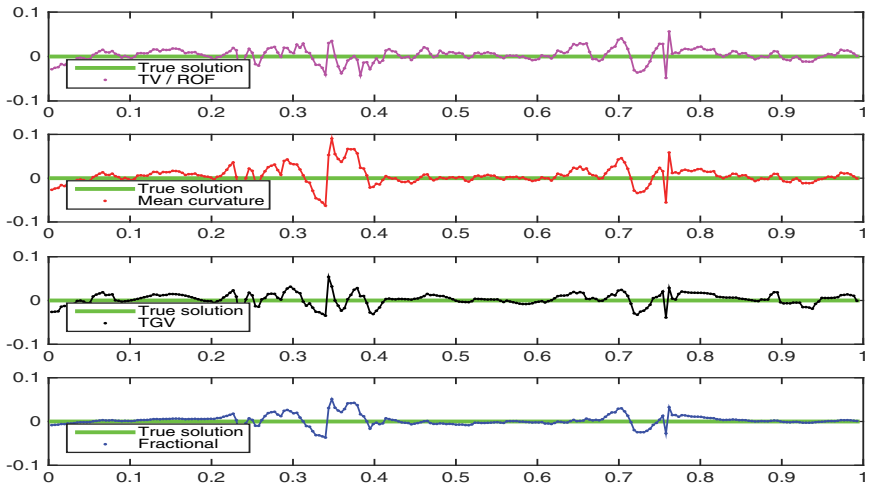
		Mean curvature [75]		TV [68]		TGV [15]		Total α-order model (3.5)	
	$\hat{\delta}$	*snr*	*psnr*	*snr*	*psnr*	*snr*	*psnr*	*snr*	*psnr*
	10	33.44	46.74	32.17	45.52	36.41	49.72	37.55	50.86
P1	20	30.19	43.50	29.55	42.83	33.03	46.33	33.52	46.83
		$\lambda_1 = 1/0.4 \times 256^2$		$\lambda_1 = 1026$		$\lambda_1 = 1/1.2 \times 256^2$		$\lambda_1^{1D} = 0.1$, $\lambda_1 = 21900$	
	Para	$\lambda_2 = 1/0.03 \times 256^2$		$\lambda_2 = 535$		$\lambda_2 = 1/0.6 \times 256^2$		$\lambda_2^{1D} = 0.1$, $\lambda_2 = 14400$	
	10	27.27	49.31	23.09	45.13	30.75	52.68	32.02	54.18
P2	20	22.88	44.92	19.45	41.49	25.62	47.51	26.48	48.54
		$\lambda_1 = 1/0.9 \times 256^2$		$\lambda_1 = 883$		$\lambda_1 = 1/0.9 \times 256^2$		$\lambda_1^{1D} = 1$, $\lambda_1 = 1800$	
	Para	$\lambda_2 = 1/0.01 \times 256^2$		$\lambda_2 = 488$		$\lambda_2 = 1/0.5 \times 256^2$		$\lambda_2^{1D} = 0.2$, $\lambda_2 = 1800$	
	10	20.43	38.80	20.08	38.35	20.40	38.78	20.48	38.86
	15	18.76	37.11	18.01	36.69	18.68	37.12	18.84	37.20
P3	20	17.48	35.82	17.17	35.33	17.55	35.87	17.57	35.90
		$\lambda_1 = 1/16 \times 256^2$		$\lambda_1 = 2216$		$\lambda_1 = 1/55 \times 256^2$		$\lambda_1^{1D} = 1$, $\lambda_1 = 16500$	
	Para	$\lambda_2 = 1/14 \times 256^2$		$\lambda_2 = 1373$		$\lambda_2 = 1/26 \times 256^2$		$\lambda_2^{1D} = 0.1$, $\lambda_2 = 9300$	
		$\lambda_3 = 1/6 \times 256^2$		$\lambda_3 = 893$		$\lambda_3 = 1/12 \times 256^2$		$\lambda_3^{1D} = 0.01$, $\lambda_3 = 6200$	
	5	25.16	37.95	24.85	37.58	25.39	38.20	25.34	38.14
	10	21.72	34.60	21.33	34.07	21.82	34.71	21.75	34.62
P4	15	19.26	32.05	18.66	31.29	19.44	32.21	19.42	32.20
		$\lambda_1 = 1/9 \times 256^2$		$\lambda_1 = 3341$		$\lambda_1 = 1/49 \times 256^2$		$\lambda_1^{1D} = 0.1$, $\lambda_1 = 24000$	
	Para	$\lambda_2 = 1/5 \times 256^2$		$\lambda_2 = 1856$		$\lambda_2 = 1/20 \times 256^2$		$\lambda_2^{1D} = 0.1$, $\lambda_2 = 8000$	
		$\lambda_3 = 1/6 \times 256^2$		$\lambda_3 = 1095$		$\lambda_3 = 1/11 \times 256^2$		$\lambda_3^{1D} = 0.1$, $\lambda_3 = 18500$	

analyzed a simple variational model using total α-order variation for image denoising. One Split-Bregman based algorithm and three optimization-based algorithms were developed to solve the resulting image inverse problem. Instead of using the usual fixed and zero boundary conditions, we proposed a boundary regularization method to treat the fractional-order derivatives. Numerical results show that the PDE-based Split-Bregman algorithm (PDE-SB) performs similarly to (though more stably than) optimization-based approaches, while our boundary regularization method is essential for getting good results for imaging denoising. Moreover, PDE-SB outperforms currently competitive variational models in terms of restoration quality. There are still outstanding issues with our proposed model and algorithms; among these, optimal selection of λ is to be addressed. Future work will also consider generalization of this work to other image inverse problems.

Appendix. Proof of Theorem 5.1. To shorten the proof, let ω be a function in $W_1^\alpha(\Omega)$ to be specified shortly. For $u \in W_1^\alpha(\Omega) \subset \mathrm{BV}^\alpha(\Omega)$, we compute the first order G-derivative



(a) u



(b) error $u - u^*$

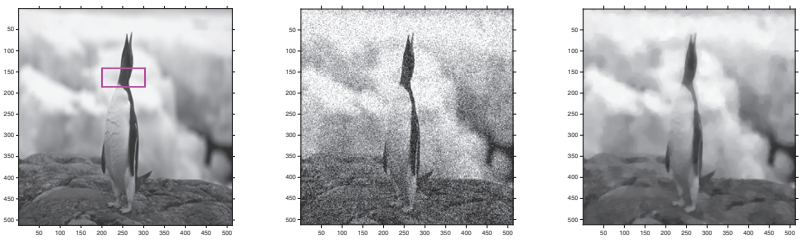
Figure 6. *Comparison* I*: The slice presentations of four restorations along a diagonal line in Figure* 5(a).

(Gateaux) of the functional $J(u)$ in the direction ω by

$$J'(u)\omega = \lim_{t\to 0} \frac{J(u+t\omega) - J(u)}{t} = \lim_{t\to 0} \frac{Q(u+t\omega) - Q(u)}{t} + \frac{\lambda}{2} \frac{F(u+t\omega) - F(u)}{t}, \tag{A.1}$$

where $Q(u) = \frac{\mu}{2} \int_\Omega |\boldsymbol{d} - \nabla^\alpha u + \frac{\boldsymbol{p}}{\mu}|^2 dx$ (see (5.3)). Using the Taylor series with respect to t yields

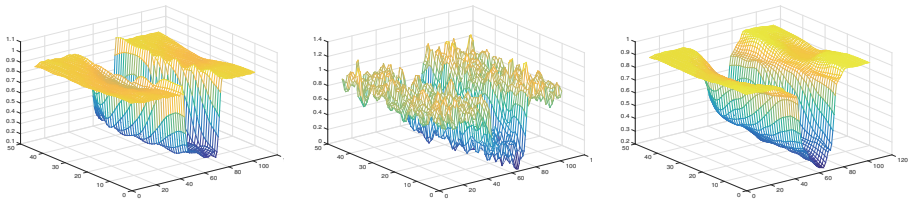
$$J'(u)\omega = \int_\Omega \boldsymbol{W} \cdot \nabla^\alpha \omega dx + \lambda \int_\Omega (u - z)\, \omega dx \tag{A.2}$$



(a) True z. (b) Noise f. (c) TV.

(d) Mean curvature. (e) TGV. (f) Our approach.

Figure 7. *Comparison* II*: Comparisons of PDE-SB with TV, mean curvature, and TGV models.*



(a) True z. (b) Noise f. (c) TV.

(d) Mean curvature. (e) TGV. (f) Our approach.

Figure 8. *Comparison* II*: Display of zoom-in surfaces inside of the red rectangle in Figure* 7(a)*. Visibly the left image (MC) is too smooth, and the middle (TGV) has some noise, while the right (PDE-SB) appears best.*

with $\boldsymbol{W} = -\mu(\boldsymbol{d} - \nabla^\alpha u + \frac{\boldsymbol{p}}{\mu})$. Recall that

(A.3)

$$\int_\Omega \boldsymbol{W} \cdot \nabla^\alpha \omega dx = (-1)^n \int_\Omega \omega^C \mathrm{div}^\alpha \boldsymbol{W} dx - \sum_{j=0}^{n-1} (-1)^j \int_0^1 D_{[a,b]}^{\alpha-n+j} W_1 \frac{\partial^{n-j-1}\omega(x)}{\partial x_1^{n-j-1}}\Big|_{x_1=0}^{x_1=1} dx_2$$
$$- \sum_{j=0}^{n-1} (-1)^j \int_0^1 D_{[c,d]}^{\alpha-n+j} W_2 \frac{\partial^{n-j-1}\omega(x)}{\partial x_2^{n-j-1}}\Big|_{x_2=0}^{x_2=1} dx_1,$$

where we note that $n = 2$ for $1 < \alpha < 2$. Next consider two case studies.

(i). Given $u(x)|_{\partial\Omega} = b_1(x)$ and $\frac{\partial u(x)}{\partial n}|_{\partial\Omega} = b_2(x)$, since $(u(x) + t\omega(x))|_{\partial\Omega} = (u(x))|_{\partial\Omega} = b_1(x)$ and $\frac{\partial(u(x)+t\omega(x))}{\partial n}|_{\partial\Omega} = \frac{\partial u(x)}{\partial n}|_{\partial\Omega} = b_2(x)$, it suffices to take $\omega \in \mathscr{C}_0^1(\Omega, \mathbb{R})$. Such a choice ensures $\frac{\partial^i \omega(x)}{\partial n^i}|_{\partial\Omega} = 0, i = 0, 1 \;\Rightarrow\; \frac{\partial^{n-j-1}\omega(x)}{\partial x_1^{n-j-1}}|_{x_1=0 \text{ or } 1} = \frac{\partial^{n-j-1}\omega(x)}{\partial x_1^{n-j-1}}|_{x_2=0 \text{ or } 1} = 0,\ n-j-1 = 0, 1$. Hence (A.1) with (A.2) reduces to (5.4).

(ii). Keep $\omega \in W_1^\alpha(\Omega)$. Since $\frac{\partial^{n-j-1}\omega(x)}{\partial x_1^{n-j-1}}|_{x_1=0 \text{ or } 1} \neq 0$, $\frac{\partial^{n-j-1}\omega(x)}{\partial x_1^{n-j-1}}|_{x_2=0 \text{ or } 1} \neq 0$, the boundary terms in (A.3) can only diminish if

$$D_{[a,b]}^{\alpha-n+j} W_1\Big|_{x_1=0 \text{ or } 1} = 0 \;\text{ and }\; D_{[c,d]}^{\alpha-n+j} W_2\Big|_{x_2=0 \text{ or } 1} = 0 \;\Rightarrow\; D^{\alpha-n+j}\boldsymbol{W}\cdot n = 0, j = 0, 1.$$

The proof is complete.

Remark 5. In imaging applications, the above first set (i) of boundary conditions seems unreasonable because one hardly knows a priori what b_1, b_2 should be. The second set (ii) of boundary conditions appears complicated and might be simplified as follows.

From [64, section 2.3.6, page 75], if $W_1(x)$ has a sufficient number of continuous derivatives, then $D_{[0,\,1]}^{\alpha-n+j} W_1|_{x_1=0 \text{ or } 1} = 0$ for any $\alpha \in (1, 2)$ is equivalent to $\frac{\partial^j W_1}{\partial x_1^j}|_{x_1=0 \text{ or } 1} = 0$ $(j = 0, 1)$; i.e.,

$$W_1\Big|_{x_1=0 \text{ or } 1} = 0 \quad\text{and}\quad \frac{\partial W_1}{\partial x_1}\Big|_{x_1=0 \text{ or } 1} = 0.$$

Indeed, if the nth derivative of $u(x)$ is integrable in $[0, 1]$, then $W_1|_{x_1=0 \text{ or } 1} = 0$ is equivalent to

$$u(x)\Big|_{x_1=0 \text{ or } 1} = 0 \;\text{ and }\; \frac{\partial u(x)}{\partial x_1}\Big|_{x_1=0 \text{ or } 1} = 0;$$

on the other hand, $\frac{\partial^k u(x)}{\partial x_1^k}|_{x_1=0 \text{ or } 1} = 0$ (for all $k = 0, 1, 2$) are equivalent to $\frac{\partial^\alpha u(x)}{\partial x_1^\alpha}|_{x_1=0 \text{ or } 1} = 0$ and $\frac{\partial^{1+\alpha} u(x)}{\partial x_1^{1+\alpha}}|_{x_1=0 \text{ or } 1} = 0$; hence one has $\frac{\partial W_1}{\partial x_1}|_{x_1=0 \text{ or } 1} = 0$. The derivations of W_2 are similar to those of W_1.

REFERENCES

[1] R. Acar and C. R. Vogel, *Analysis of bounded variation penalty methods for ill-posed problems*, Inverse Problems, 10 (1994), pp. 1217–1229.

[2] V. Agarwal, A. V. Gribok, and M. A. Abidi, *Image restoration using l-1 norm penalty function*, Inverse Probl. Sci. Eng., 15 (2007), pp. 785–809.

[3] O. P. Agrawal, *Formulation of Euler-Lagrange equations for fractional variational problems*, J. Math. Anal. Appl., 272 (2002), pp. 368–379.

[4] O. P. Agrawal, *Fractional variational calculus in terms of Riesz fractional derivatives*, J. Phys. A, 40 (2007), pp. 62–87.

[5] R. Almeida and D. F. M. Torres, *Calculus of variations with fractional derivatives and fractional integrals*, Appl. Math. Lett., 22 (2009), pp. 1816–1820.

[6] R. Almeida and D. F. M. Torres, *Necessary and sufficient conditions for the fractional calculus of variations with Caputo derivatives*, Commun. Nonlinear Sci. Numer. Simul., 16 (2011), pp. 1490–1500.

[7] L. Ambrosio and S. Masnou, *A direct variational approach to a problem arising in image reconstruction*, Interfaces Free Bound., 5 (2003), pp. 63–82.

[8] A. Atangana and A. Secer, *A note on fractional order derivatives and table of fractional derivatives of some special functions*, Abstr. Appl. Anal., 2013 (2013), 279681.
[9] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, 2nd ed., Appl. Math. Sci. 147, Springer-Verlag, New York, 2006.
[10] J. F. Aujol, *Some first-order algorithms for total variation based image restoration*, J. Math. Imaging Vision, 34 (2009), pp. 307–327.
[11] J. Bai and X.-C. Feng, *Fractional-order anisotropic diffusion for image denoising*, IEEE Trans. Image Process., 16 (2007), pp. 2492–2502.
[12] H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, Springer Optim. Appl. 49, Springer, New York, 2011.
[13] A. Beck and M. Teboulle, *Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems*, IEEE Trans. Image Process., 18 (2009), pp. 2419–2434.
[14] A. Beck and M. Teboulle, *A fast iterative shrinkage-thresholding algorithm for linear inverse problems*, SIAM J. Imaging Sci., 2 (2009), pp. 183–202.
[15] K. Bredies, K. Kunisch, and T. Pock, *Total generalized variation*, SIAM J. Imaging Sci., 3 (2010), pp. 492–526.
[16] X. Bresson, S. Esedoglu, P. Vandergheynst, J.-P. Thiran, and S. Osher, *Fast global minimization of the active contour/snake model*, J. Math. Imaging Vision, 28 (2007), pp. 151–167.
[17] C. Brito-Loeza and K. Chen, *Multigrid algorithm for high order denoising*, SIAM J. Imaging Sci., 3 (2010), pp. 363–389.
[18] A. Chambolle, *An algorithm for total variation minimization and applications*, J. Math. Imaging Vision, 20 (2004), pp. 89–97.
[19] A. Chambolle and P. L. Lions, *Image recovery via total variation minimization and related problems*, Numer. Math., 76 (1997), pp. 167–188.
[20] A. Chambolle and T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, J. Math. Imaging Vision, 40 (2011), pp. 120–145.
[21] R. H. Chan, A. Lanza, S. Morigi, and F. Sgallari, *An adaptive strategy for the restoration of textured images using fractional order regularization*, Numer. Math. Theory Methods Appl., 6 (2013), pp. 276–296.
[22] T. Chan, A. Marquina, and P. Mulet, *High-order total variation-based image restoration*, SIAM J. Sci. Comput., 22 (2000), pp. 503–516.
[23] T. Chan, A. M. Yip, and F. E. Park, *Simultaneous total variation image inpainting and blind deconvolution*, Internat. J. Imaging Syst. Tech., 15 (2005), pp. 92–102.
[24] T. F. Chan, S. H. Kang, and J. Shen, *Euler's elastica and curvature-based inpainting*, SIAM J. Appl. Math., 63 (2002), pp. 564–592.
[25] T. F. Chan and J. Shen, *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*, SIAM, Philadelphia, 2005.
[26] Q. S. Chang, X. C. Tai, and L. Xing, *A compound algorithm of denoising using second-order and fourth-order partial differential equations*, Numer. Math. Theory Methods Appl., 2 (2009), pp. 353–376.
[27] D. Chen, Y. Chen, and D. Xue, *Fractional-order total variation image restoration based on primal-dual algorithm*, Abstr. Appl. Anal., 2013 (2013), 585310.
[28] D. Chen, Y. Chen, and D. Xue, *Three fractional-order TV-l^2 models for image denoising*, J. Comput. Inform. Syst., 9 (2013), pp. 4773– 4780.
[29] D. Chen, S. S. Sun, C. R. Zhang, Y. Q. Chen, and D. Y. Xue, *Fractional-order TV-L2 model for image denoising*, Cent. Eur. J. Phys., 11 (2013), pp. 1414–1422.
[30] N. Chumchob, K. Chen, and C. Brito-Loeza, *A fourth-order variational image registration model and its fast multigrid algorithm*, Multiscale Model. Simul., 9 (2011), pp. 89–128.
[31] P. L. Combettes and V. R. Wajs, *Signal recovery by proximal forward-backward splitting*, Multiscale Model. Simul., 4 (2005), pp. 1168–1200.
[32] G. Dal Maso, I. Fonseca, G. Leoni, and M. Morini, *A higher order model for image restoration: The one-dimensional case*, SIAM J. Math. Anal., 40 (2009), pp. 2351–2391.

[33] V. Duval, J. F. Aujol, and L. Vese, *Projected gradient based color image decomposition*, in Scale Space and Variational Methods in Computer Vision, Lecture Notes in Comput. Sci. 5567, Springer, Berlin, 2009, pp. 295–306.
[34] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, Stud. Adv. Math., CRC Press, Boca Raton, FL, 1991.
[35] R. P. Fedkiw, G. Sapiro, and C. W. Shu, *Shock capturing, level sets, and PDE based methods in computer vision and image processing: A review of Osher's contributions*, J. Comput. Phys., 185 (2003), pp. 309–341.
[36] B. Fischer and J. Modersitzki, *Fast diffusion registration*, Contemp. Math., 313 (2002), pp. 117–128.
[37] B. Fischer and J. Modersitzki, *Curvature based image registration*, J. Math. Imaging Vision, 18 (2003), pp. 81–85.
[38] C. Frohn-Schauf, S. Henn, and K. Witsch, *Multigrid based total variation image registration*, Comput. Vis. Sci., 11 (2008), pp. 101–113.
[39] J. B. Garnett, T. M. Le, Y. Meyer, and L. A. Vese, *Image decompositions using bounded variation and generalized homogeneous Besov spaces*, Appl. Comput. Harmon. Anal., 23 (2007), pp. 25–56.
[40] S. Geman and D. Geman, *Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images*, IEEE Trans. Pattern Anal. Machine Intell., 6 (1984), pp. 721–741.
[41] P. Getreuer, *Total variation inpainting using split Bregman*, Image Process. On Line, 2 (2012), 147157.
[42] S. N. Ghate, S. Achaliya, and S. Raveendran, *An algorithm of total variation for image inpainting*, Internat. J. Comput. Electron. Res., 1 (2012), pp. 124–130.
[43] E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, Springer, New York, 1984.
[44] T. Goldstein and S. Osher, *The split Bregman method for L1-regularized problems*, SIAM J. Imaging Sci., 2 (2009), pp. 323–343.
[45] P. Guidotti, *A new nonlocal nonlinear diffusion of image processing*, J. Differential Equations, 246 (2009), pp. 4731–4742.
[46] P. Guidotti and J. V. Lambers, *Two new nonlinear nonlocal diffusions for noise reduction*, J. Math. Imaging Vision, 33 (2009), pp. 25–37.
[47] W. Guo and L. H. Qiao, *Inpainting based on total variation*, in Proceedings of the International Conference on Wavelet Analysis and Pattern Recognition (Beijing, 2007), Vol. 2, IEEE, Washington, DC, 2007, pp. 939–943.
[48] R. Hilfer, *Applications of Fractional Calculus in Physics*, World Scientific, Singapore, 2000.
[49] L. Hömke, C. Frohn-Schauf, S. Henn, and K. Witsch, *Total variation based image registration*, in Image Processing Based on Partial Differential Equations, Springer, New York, 2007, pp. 343–361.
[50] M. Janev, S. Pilipović, T. Atanacković, R. Obradović, and N. Ralević, *Fully fractional anisotropic diffusion for image denoising*, Math. Comput. Model., 54 (2011), pp. 729–741.
[51] G. Jumarie, *Modified Riemann-Liouville derivative and fractional Taylor series of nondifferentiable functions: Further results*, Comput. Math. Appl., 51 (2006), pp. 1367–1376.
[52] H. Köstler, K. Ruhnau, and R. Wienands, *Multigrid solution of the optical flow system using a combined diffusion-and curvature-based regularizer*, Numer. Linear Algebra Appl., 15 (2008), pp. 201–218.
[53] M. Lysaker, A. Lundervold, and X. C. Tai, *Noise removal using fourth-order partial differential equation with application to medical magnetic resonance images in space and time*, IEEE Trans. Image Process., 12 (2003), pp. 1579–1590.
[54] M. Lysaker, S. Osher, and X. C. Tai, *Noise removal using smoothed normals and surface fitting*, IEEE Trans. Image Process., 13 (2004), pp. 1345–1357.
[55] A. Melbourne, N. Cahill, C. Tanner, M. Modat, D. J. Hawkes, and S. Ourselin, *Using fractional gradient information in non-rigid image registration: Application to breast MRI*, in Medical Imaging 2012: Image Processing, Proc. SPIE 8314, 2012, 83141Z.
[56] K. S. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*, Wiley Interscience, New York, 1993.
[57] J. Modersitzki, *Numerical Methods for Image Registration*, Oxford University Press, Oxford, UK, 2004.
[58] Y. Nesterov, *A method for unconstrained convex minimization problem with the rate of convergence* $o(1/k^2)$, Soviet Math. Dokl., 269 (1983), pp. 543–547.

[59] Y. Nesterov, *Gradient methods for minimizing composite functions*, Math. Program., 140 (2013), pp. 125–161.
[60] K. B. A. Oldham and J. A. Spanier, *The Fractional Calculus: Theory and Applications of Differentiation and Integration to Arbitrary Order*, Dover, New York, 2006.
[61] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, *An iterative regularization method for total variation-based image restoration*, Multiscale Model. Simul., 4 (2005), pp. 460–489.
[62] S. Osher, A. Solé, and L. Vese, *Image decomposition and restoration using total variation minimization and the H^{-1} norm*, Multiscale Model. Simul., 1 (2003), pp. 349–370.
[63] T. Pock, M. Urschler, C. Zach, R. Beichel, and H. Bischof, *A duality based algorithm for TV-L1-optical-flow image registration*, in Medical Image Computing and Computer-Assisted Intervention, Springer, New York, 2007, pp. 511–518.
[64] I. Podlubny, *Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and Some of Their Applications*, Math. Sci. Engrg. 198, Academic Press, San Diego, CA, 1999.
[65] I. Podlubny, *Matrix approach to discrete fractional calculus*, Fract. Calc. Appl. Anal., 3 (2000), pp. 359–386.
[66] I. Podlubny, A. Chechkin, T. Skovranek, Y. Chen, and B. M. Vinagre-Jara, *Matrix approach to discrete fractional calculus* II: *Partial fractional differential equations*, J. Comput. Phys., 228 (2009), pp. 3137–3153.
[67] P. D. Romero and V. F. Candela, *Blind deconvolution models regularized by fractional powers of the Laplacian*, J. Math. Imaging Vision, 32 (2008), pp. 181–191.
[68] L. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259–268.
[69] S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Fractional integrals and derivatives: Theory and Applications*, CRC Press, Boca Raton, FL, 1993.
[70] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen, *Variational Methods in Imaging*, Appl. Math. Sci. 167, Springer, New York, 2009.
[71] S. Setzer, *Split Bregman algorithm, Douglas-Rachford splitting and frame shrinkage*, in Scale Space and Variational Methods in Computer Vision, Lecture Notes in Comput. Sci. 5567, Springer, Berlin, 2009, pp. 464–476.
[72] S. Setzer, *Operator splittings, Bregman methods and frame shrinkage in image processing*, Internat. J. Comput. Vis., 92 (2011), pp. 265–280.
[73] Z. Shen, K.-C. Toh, and S. Yun, *An accelerated proximal gradient algorithm for frame-based image restoration via the balanced approach*, SIAM J. Imaging Sci., 4 (2011), pp. 573–596.
[74] G. Steidl, S. Didas, and J. Neumann, *Relations between higher order TV regularization and support vector regression*, in Scale Space and PDE Methods in Computer Vision, Lecture Notes in Comput. Sci. 3459, Springer, Berlin, 2005, pp. 515–527.
[75] L. Sun and K. Chen, *A new iterative algorithm for mean curvature-based variational image denoising*, BIT, 54 (2014), pp. 523–553.
[76] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, John Wiley & Sons, New York, 1977.
[77] M. Unger, T. Pock, W. Trobin, D. Cremers, and H. Bischof, *TVSeg-interactive total variation based image segmentation*, in BMVC: 2008 British Machine Vision Conference, Leeds, UK, 2008, pp. 1–10.
[78] R. Verdú-Monedero, J. Larrey-Ruiz, J. Morales-Sánchez, and J. L. Sancho-Gómez, *Fractional regularization term for variational image registration*, Math. Probl. Eng., 2009 (2009), 707026.
[79] H. Wang and N. Du, *Fast solution methods for space-fractional diffusion equations*, J. Comput. Appl. Math., 255 (2014), pp. 376–383.
[80] E. Zeidler, *Nonlinear Functional Analysis and Its Applications.* III. *Variational Methods and Optimization*, Springer-Verlag, New York, 1985.
[81] J. Zhang, Z. Wei, and L. Xiao, *Adaptive fractional-order multi-scale method for image denoising*, J. Math. Imaging Vision, 43 (2012), pp. 39–49.
[82] J. Zhang, K. Chen, and B. Yu, *An iterative Lagrange multiplier method for constrained total-variation-based image denoising*, SIAM J. Numer. Anal., 50 (2012), pp. 983–1003.

[83] Y. Zhang, Y. F. Pu, J. R. Hu, and J. L. Zhou, *A class of fractional-order variational image inpainting models*, Appl. Math. Inform. Sci, 6 (2012), pp. 299–306.
[84] W. Zhu and T. F. Chan, *Image denoising using mean curvature of image surface*, SIAM J. Imaging Sci., 5 (2012), pp. 1–32.
[85] Z. Y. Zhu, G. G. Li, and C. J. Cheng, *A numerical method for fractional integral with applications*, Appl. Math. Mech., 24 (2003), pp. 373–384.