

Galerkin method with splines for total variation minimization

Qianying Hong¹, Ming-Jun Lai², Leopold Matamba Messi³ and Jingyue Wang¹ 

Journal of Algorithms & Computational Technology
Volume 13: 1–16
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1748301819833046
journals.sagepub.com/home/act



Abstract

Total variation smoothing methods have been proven to be very efficient at discriminating between structures (edges and textures) and noise in images. Recently, it was shown that such methods do not create new discontinuities and preserve the modulus of continuity of functions. In this paper, we propose a Galerkin–Ritz method to solve the Rudin–Osher–Fatemi image denoising model where smooth bivariate spline functions on triangulations are used as approximating spaces. Using the extension property of functions of bounded variation on Lipschitz domains, we construct a minimizing sequence of continuous bivariate spline functions of arbitrary degree, d , for the TV- L^2 energy functional and prove the convergence of the finite element solutions to the solution of the Rudin, Osher, and Fatemi model. Moreover, an iterative algorithm for computing spline minimizers is developed and the convergence of the algorithm is proved.

Keywords

Finite element, Galerkin method, image denoising, spline, total variation

Received 1 June 2018; Revised received 2 September 2018; accepted 26 October 2018

Introduction

A few decades ago, Rudin et al.¹ proposed a constrained total variation minimization method for image enhancement. Suppose you have an image f filled with artifacts and your goal is to reduce these artifacts and enhance the quality of your image while preserving the details of the image as much as possible. Assuming that the image f is a function defined on regular domain $\Omega \subset \mathbb{R}^2$, Rudin, Osher and Fatemi's (ROF) approach is to solve the following penalized total variation minimization problem

$$\arg \min_{u \in L^2(\Omega)} \lambda \int_{\Omega} |Du| + \frac{1}{2} \int_{\Omega} |u-f|^2 dx \quad (1)$$

where $\int_{\Omega} |Du|$ is the total variation of u , and λ is a positive parameter controlling the fidelity of the recovered image to the initial image f . From now on we will denote the total variation of a function u by

$$J(u) := \int_{\Omega} |Du| \quad (2)$$

We shall refer to the minimization problem (equation (1)) as the ROF model, and denote its objective functional by

$$E_{\lambda}^f(u) := \lambda J(u) + \frac{1}{2} \int_{\Omega} |u-f|^2 dx \quad (3)$$

Notice that a minimal condition for the ROF model to be defined is that the image f be a square integrable function over Ω in which case the domain of $E_{\lambda}^f(u)$ is $BV(\Omega) \cap L^2(\Omega)$, where $BV(\Omega)$ stands for the space of functions of bounded variation over Ω .

¹Department of Mathematics, Fuzhou University, Fuzhou, Fujian, China

²Department of Mathematics, The University of Georgia, Athens, GA, USA

³Mathematical Biosciences Institute, The Ohio State University, Columbus, OH, USA

Corresponding author:

Jingyue Wang, Department of Mathematics, Fuzhou University, Fuzhou, Fujian 350108, China.

Email: jingyuewang@fzu.edu.cn



The last decades witnessed excellent progress and exciting development regarding the analysis of the ROF model. The interested reader is referred to Chambolle et al.² for a comprehensive introduction on this topic. Since the introduction of the model, much understanding is gained on the analytic properties of its solution (c.f. Caselles et al.,³ Allard,^{4,5} and Caselles et al.⁶). For instance, it is known that, the solution of the model does not introduce new discontinuity compared to the input data. Specifically, it was proved in Caselles et al.⁶ that the jump set of the solution is always contained in the jump set of the input image; and if Caselles et al.³ f has modulus of continuity ω , then so does the minimizer of $E_\lambda^f(u)$ provided that Ω is convex. A concise outline of recent progress on the study of the properties of the continuous model can be found at Chambolle et al.⁷

On the computational side of the model, early algorithms were based on smooth approximations of the primal form of the variational problem (equation (1)) (c.f. Acar and Vogel⁸ and Chambolle and Lions⁹). Chambolle¹⁰ first proposed a gradient descent algorithm that operated on the dual form of the original minimization problem. In addition, He proved that this algorithm converged to the exact solution of the discretized model (c.f. Chambolle^{10,11}). Much progress had since been made in the 2000s to develop highly accelerated first-order algorithms (c.f. Aujol and Dossal,¹² Beck and Teboulle,¹³ Beck,¹⁴ Combettes and Wajs,¹⁵ Esser et al.,¹⁶ Goldstein and Osher,¹⁷ Zhu and Chan,¹⁸ and Zhu et al.¹⁹) to solve the dual problem. See Chambolle and Pock²⁰ for a recent survey on such numerical methods. Independent of the aforementioned approaches, there is another class of combinatorial optimization methods that are very efficient and have been extensively studied by the computer vision community (c.f. Boykov et al.,²¹ Darbon and Sigelle,^{22,23} and Goldfarb and Yin²⁴). We note that in all aforementioned works, the discretization of the ROF model is always carried out on a fixed rectangular grid where finite difference schemes are convenient to compute the total variation. Error bound of such discretization was derived in Lai et al.²⁵ and Wang and Lucier.²⁶

To the best of our knowledge, Dobson and Vogel²⁷ (Theorem 2.2, p. 1782) are among the first to investigate solving the total variation minimization problem with Galerkin schemes. They gave a sufficient condition for the convergence of a Galerkin scheme for the ROF model. However, they also observed that the said condition is easily achieved if the solution of the ROF model is sufficiently smooth and suggested that more research be done under less stringent regularity assumptions. Along this line of study, many well-developed methods in the theory of finite element

method are applied to this problem (c.f. Lai and Messi,²⁸ Matamba Messi,²⁹ Bartels,^{30,31} Bartels and Milicevic,³² Chen and Tai,³³ Feng and Prohl,³⁴ Feng et al.,³⁵ Litvinov et al.,³⁶ Stamm and Wihler,³⁶ Tian and Yuan,³⁷ and Xu et al.³⁸). It is worth noting that all these works employ either piecewise-constant or piecewise-linear finite elements. They also presume certain regularity conditions on the input data with the exception of Bartels' work.

This paper addresses Dobson and Vogel's question by constructing a convergent Galerkin scheme regardless of the regularity of the solution of the ROF model. The novelty of our method is we extend the conventional piecewise-linear finite-elements-based approach to bivariate spline element with arbitrary degree, and prove the convergence of the numerical scheme. We also propose a relaxation algorithm for solving the variation problem using bivariate splines. We note in numerical experiments that unlike other algorithms that use finite difference schemes, the spline algorithm that we propose does not show the notorious "staircase" effect. We also gain flexibility with the capability to select any polygonal region in an image for local treatment; this was not easy with algorithms that required rectangular patches when doing local denoising. Our method is also anticipating the day when we will be able to capture images with non-rectangular CCD devices.

The paper is structured as follows. The next section is devoted to the necessary mathematical preliminaries on functions of bounded variation and bivariate spline functions. In the *Galerkin approximation with smooth splines* section, we review relevant properties of the ROF model and prove the main result (Theorem 11) of this paper. We introduce an algorithm for approximating the terms of the spline minimizing sequence and prove its convergence in the *Fixed point relaxation algorithm* section. In the last section, we show the numerical evidence of its effectiveness at smoothing images.

Preliminary results

In this section and throughout the paper, the planar domain Ω is assumed polygonal, unless otherwise noted. We also remind the reader that by domain of \mathbb{R}^2 , we mean a connected open subset.

Functions of bounded variation

A function $u : \Omega \rightarrow \mathbb{R}$ is said to be of bounded variation if $u \in L^1(\Omega)$ and its total variation

$$J(u) := \sup \left\{ \int_{\Omega} u \operatorname{div}(\varphi) dx : \varphi \in C_c^1(\Omega, \mathbb{R}^2) \right\} \quad (4)$$

$$|\varphi(x)| \leq 1, \forall x \in \Omega \} \quad (5)$$

is finite. For example any function $u \in W^{1,1}(\Omega)$ is of bounded variation with total variation

$$\int_{\Omega} |Du| = \int_{\Omega} |\nabla u| dx \quad (6)$$

The set of functions of bounded variation, denoted $BV(\Omega)$, is a Banach space for the norm

$$\|u\|_{BV} := \|u\|_{L^1} + \int_{\Omega} |Du| \quad (7)$$

Furthermore, if $u \in BV(\Omega)$, then its distributional derivative,⁴⁰ Du , is a finite vector-valued Radon measure on Ω , and the total variation of Du induces a Borel measure on Ω known as the total variation measure of u over Ω .

The following result asserts that a function u defined on a domain Ω of \mathbb{R}^2 with zero total variation must be constant. In particular, if $\Omega = \mathbb{R}^2$, then the total variation is a norm on $BV(\mathbb{R}^2)$ equivalent to the norm $\|\cdot\|_{BV}$ defined in equation (7).

Theorem 1 (Poincaré Inequality^{41,42}). Suppose that Ω is a bounded Lipschitz domain of \mathbb{R}^2 . Then there exists a constant C depending only on Ω such that

$$\|u - u_{\Omega}\|_{L^2(\Omega)} \leq C \int_{\Omega} |Du|, \quad \forall u \in BV(\Omega) \quad (8)$$

where $u_{\Omega} = \frac{1}{|\Omega|} \int_{\Omega} u(x) dx$ is the average value of u over Ω . If $\Omega = \mathbb{R}^2$, then there exists $C > 0$ such that for any compactly supported function $u \in BV(\mathbb{R}^2)$

$$\|u\|_{L^2(\mathbb{R}^2)} \leq C \int_{\mathbb{R}^2} |Du| \quad (9)$$

Another property of functions of bounded variation that is central to our contribution in this work is the existence of an extension operator from $BV(\Omega)$ into $BV(\mathbb{R}^2)$ that does not turn the boundary of Ω into a singular set for the total variation measure.

Theorem 2⁴¹ (Proposition 3.21, p. 131). Suppose that Ω is a bounded Lipschitz domain. Then for any bounded open set $A \subset \mathbb{R}^2$ such that Ω is relatively compact in A , there exists a bounded linear extension operator $T_A : BV(\Omega) \rightarrow BV(\mathbb{R}^2)$ such that the following hold:

a. For any $u \in BV(\Omega)$, $|DT_{Au}|(\Gamma) = 0$ and the support of T_{Au} is contained in A .

b. The restriction of T_A to $W^{1,1}(\Omega)$ is a bounded linear operator into $W^{1,1}(\mathbb{R}^2)$.

Proof. The proof is constructive and parallels the construction of an extension operator³⁹ from $W^{1,1}(\Omega)$ to $W^{1,1}(\mathbb{R}^2)$ using a partition of unity argument. The complete proof is found in Ambrosio et al.⁴⁰ \square

We now give the properties of the total variation functional $J : L^1(\Omega) \rightarrow [0, \infty]$ that play a primordial role in proving the existence and uniqueness of the solution for the ROF model.

Proposition 3. The total variation functional $J : L^1(\Omega) \rightarrow [0, +\infty]$ satisfies the properties:

- (a) J is positively 1-homogeneous, i.e., $J(tu) = tJ(u)$, $\forall t \geq 0$ and $\forall u \in BV(\Omega)$;
- (b) J is convex, i.e., $J(tu + (1-t)v) \leq tJ(u) + (1-t)J(v)$, $\forall t \in [0, 1]$, $\forall u, v \in L^1(\Omega)$;
- (c) J is lower semicontinuous, i.e., if (u_n) is a sequence which converges in $L^1(\Omega)$ to u , then

$$J(u) \leq \liminf_{n \rightarrow \infty} J(u_n) \quad (10)$$

Proof. The proof of the proposition is straightforward with (a) and (b) arising from the definition of the total variation, while (c) is a consequence of Lebesgue Dominated Convergence Theorem. \square

To establish the main result of this paper, we need to construct a sequence of smooth functions that converges in $L^1(\Omega)$ for which the equality holds in equation (10). By exploiting the extension property of functions of bounded variation (see Theorem 2 above), we will use the following lemma to achieve this goal.

Lemma 4⁴² (Proposition 1.15). Suppose $u \in BV(\Omega)$. If $A \subset \subset \Omega$ is a relatively compact open subset of Ω such that

$$\int_{\partial A} |Du| = 0 \quad (11)$$

then

$$\int_A |Du| = \lim_{\epsilon \rightarrow 0} \int_A |D(u * \eta_{\epsilon})| \quad (12)$$

where $\eta_{\epsilon}(x) = \epsilon^{-2} \eta(x/\epsilon)$ and η is radially symmetric mollifier.

Bivariate spline functions

Let Δ be a triangulation of Ω . A spline function on the triangulation Δ is a function s defined on Ω such that

for any triangle $T \in \Delta$, the restriction $s|_T$ of s to T is a polynomial. The degree of a spline function is the maximum degree of its restrictions to elements of the triangulation Δ . The space of spline functions of degree d on Δ is denoted by

$$\mathcal{S}_d^{-1}(\Delta) := \{s : \Omega \rightarrow \mathbb{R} : s|_T \in \mathbb{P}_d \forall T \in \Delta\}$$

where \mathbb{P}_d is the vector space of bivariate polynomials of degree less than or equal to d . The space of smooth spline functions of degree d and order $0 \leq r \leq d$, $\mathcal{S}_d^r(\Delta)$, is defined by

$$\begin{aligned} \mathcal{S}_d^r(\Delta) &= C^r(\Omega) \cap \mathcal{S}_d^{-1}(\Delta) \\ &= \{s \in C^r(\Omega) : s|_t \in \mathbb{P}_d, \forall T \in \Delta\} \end{aligned}$$

Given a basis of the polynomial space \mathbb{P}_d , it is easy to see that $\mathcal{S}_d^{-1}(\Delta)$ is isomorphic to \mathbb{R}^N where $N = \#(\Delta) \binom{d+2}{2}$ and $\#(\Delta)$ is the number of triangles in Δ , while the space of smooth splines $\mathcal{S}_d^r(\Delta)$ is a subspace of \mathbb{R}^N of the form⁴²

$$\mathcal{S}_d^r(\Delta) \equiv \{c \in \mathbb{R}^N : A(r)c = \mathbf{0}\} \quad (13)$$

where $A(r)$ is an $(r+1)(d+1)E \times N$ matrix encoding the smoothness condition across the interior edges of the triangulation Δ , and E is the number of interior edges of Δ . Notice that we can use a different basis of \mathbb{P}_d for each triangle $T \in \Delta$ and in such instance we shall write

$$\mathcal{S}_d^{-1}(\Delta_h) = \prod_{T \in \Delta} \mathbb{P}_d^T$$

For our purposes in this paper, we shall use the Bernstein-Bézier basis of \mathbb{P}_d^T for each triangle $T \in \Delta$.

Spline functions have been used with much success in the numerical computation of partial differential equations using variational methods⁴⁴⁻⁴⁸ and more recently for the numerical simulation of the Darcy-Stokes equation.⁴⁹ In general, spline functions may be utilized as approximation spaces to study some classes of variational equations using the Galerkin method. Their appeal to us in this work is twofold. Firstly, bivariate spline functions possess good approximation power in the Sobolev spaces $W^{m,p}(\Omega)$ as illustrated by the following theorem.

Theorem 5⁴³ (Theorem 10.2, p. 277). Suppose that Δ is a regular triangulation of Ω of mesh size $h > 0$. Let $p \in [1, \infty]$ and $d \in \mathbb{N}$ be given. Then for every

$u \in W^{d+1,p}(\Omega)$, there exists a spline function $s_u \in \mathcal{S}_d^0(\Delta)$ such that

$$\begin{aligned} \|D_1^\alpha D_2^\beta (u - s_u)\|_{L^p(\Omega)} &\leq Kh^{d+1-\alpha-\beta} |u|_{d+1,p} \\ \forall 0 \leq \alpha + \beta &\leq d \end{aligned} \quad (14)$$

where K depends only on d and the smallest angle of Δ , and

$$|u|_{d+1,p} = \sum_{\alpha+\beta=d+1} \|D_1^\alpha D_2^\beta u\|_{L^p(\Omega)}$$

Secondly, the differential operators $D_1^\alpha D_2^\beta$ are bounded linear operators between the spaces $\mathcal{S}_d^{-1}(\Delta)$ and $\mathcal{S}_{d-\alpha-\beta}^{-1}(\Delta)$. This property is known in the literature as the Markov inequality.

Theorem 6 (Markov inequality,⁴² Theorem 2.32). Let Δ be a triangulation of Ω . Let $p \in [1, \infty)$ and $d \in \mathbb{N}$ be fixed. There exists a constant K depending only on d such that for all nonnegative integers α and β with $0 \leq \alpha + \beta \leq d$, we have

$$\|D_1^\alpha D_2^\beta s\|_{L^p(\Omega)} \leq \frac{K}{\rho^{\alpha+\beta}} \|s\|_{L^p(\Omega)}, \forall s \in \mathcal{S}_d^{-1}(\Delta) \quad (15)$$

where $\rho = \min\{\rho_t : t \in \Delta\}$ with ρ_t the inradius of the triangle t .

Remark 7. Notice that one can define a spline functions space on any simplicial tiling of the domain Ω and all polynomials based finite element spaces are preferential subspaces of the spline space $\mathcal{S}_d^{-1}(\Delta_h)$ for some choice of d .

Natural images contain structural information in the form of edges and textures. Resolving these entities with continuous spline functions will require fine triangulations which in turn demand high resolution images. An alternative is to use high order spline functions on moderate size triangulations as these are more capable of capturing the variations corresponding to edges than continuous spline functions on such triangulations. The following result gives the relationship between the degree and the order of the spline function to guarantee suitable approximation power in Sobolev spaces.

Theorem 8⁴³ (Theorem 10.10). Let $d \geq 3r + 2$ and suppose Δ is a regular triangulation of Ω of mesh size h . Then for every $f \in W_q^{d+1}(\Omega)$, there exists a spline $s \in \mathcal{S}_d^r(\Delta_h)$ such that

$$\begin{aligned} \|D_1^\alpha D_2^\beta (u - s_u)\|_{L^p(\Omega)} &\leq Kh^{d+1-\alpha-\beta} |u|_{d+1,p} \\ \forall 0 \leq \alpha + \beta &\leq d \end{aligned} \quad (16)$$

If Ω is convex, then the constant K depends only on r, d and the smallest angle on Δ ; otherwise K also depends on the Lipschitz constant of the boundary of Ω .

Galerkin approximation with smooth splines

In this section, we describe how we arrive at a family of continuous bivariate spline functions that approximate the minimizer of the functional

$$E_\lambda^f(u) := \lambda J(u) + \frac{1}{2} \int_{\Omega} |u - f|^2 dx \quad (17)$$

Before undertaking the analysis of our approximation method, let us briefly explain why the ROF model is well posed. In fact, Proposition 3 implies that for $f \in L^2(\Omega)$ and $\lambda > 0$ fixed, the ROF functional E_λ^f is strictly convex and lower semi-continuous on $BV(\Omega) \cap L^2(\Omega)$ for the norm of $L^2(\Omega)$. Therefore, the ROF model (1) has a unique solution and the problem is well posed as illustrated by the following result.

Theorem 9. *Let $u_\lambda^f \in BV(\Omega)$ be the minimizer of the ROF functional $E_\lambda^f(u)$. Then for any $v \in BV(\Omega)$, there holds*

$$\|v - u_\lambda^f\|_{L^2}^2 \leq 2 \left(E_\lambda^f(v) - E_\lambda^f(u_\lambda^f) \right) \quad (18)$$

and

$$\inf_{x \in \Omega} f(x) \leq u_\lambda^f(x) \leq \sup_{x \in \Omega} f(x) \quad \text{for a.e. } x \in \Omega \quad (19)$$

Moreover, if u_λ^g is the minimizer of $E_\lambda^g(u)$, then

$$\|u_\lambda^f - u_\lambda^g\|_{L^2} \leq \|f - g\|_{L^2} \quad (20)$$

Proof. The proof is a simple exercise of convex analysis and uses the characterization of the minimizer of a convex functional using subdifferentials. The complete proof is found in Matamba Messi.²⁹ \square

The approximation of the minimizer of the ROF model by continuous spline functions is possible because the space $S_d^0(\Delta)$ possesses very good approximation power in high-order Sobolev spaces as illustrated by Theorem 5 and a function of bounded variation can be approximated by smooth functions. In using Theorem 5, we will need to control the norm of high order derivatives of the mollification of a BV function. This is done as in the lemma below.

Lemma 10. *Let $u \in BV(\mathbb{R}^2)$ be fixed. Then for any integer $m \geq 0$, any pair of nonnegative integer (α, β) such that $\alpha + \beta = m + 1$, and any $\epsilon > 0$, we have*

$$\|D_1^\alpha D_2^\beta(\eta_\epsilon * u)\|_{L^1(\Omega)} \leq \frac{C}{\epsilon^m} |Du|(\mathbb{R}^2) \quad (21)$$

where C is a constant depending only on m and Ω .

Proof. Let $\varphi \in C_c^1(\Omega)$ be given. Let α and β be two nonnegative integers such that $\alpha + \beta = m + 1$; we may assume without loss of generality that $\alpha \geq 1$. Then, with $\tilde{\eta}_\epsilon m(x) = D_1^{\alpha-1} D_2^\beta \eta_\epsilon(-x)$ we have

$$\begin{aligned} \int_{\Omega} D_1^\alpha D_2^\beta(\eta_\epsilon * u) \varphi dx &= - \int_{\mathbb{R}^2} D_1^{\alpha-1} D_2^\beta(\eta_\epsilon * u) \frac{\partial \varphi}{\partial x_1} dx \\ &= - \int_{\mathbb{R}^2} D_1^{\alpha-1} D_2^\beta \eta_\epsilon * u \frac{\partial \varphi}{\partial x_1} dx \\ &= - \int_{\mathbb{R}^2} u \tilde{\eta}_\epsilon m * \frac{\partial \varphi}{\partial x_1} dx \\ &= - \int_{\mathbb{R}^2} u \frac{\partial}{\partial x_1} [\tilde{\eta}_\epsilon m * \varphi] dx \end{aligned}$$

Thus

$$\int_{\Omega} D_1^\alpha D_2^\beta(\eta_\epsilon * u) \varphi dx \leq \|\tilde{\eta}_\epsilon m * \varphi\|_{\infty} |Du|(\mathbb{R}^2)$$

Now by Hölder's inequality we have

$$\|\tilde{\eta}_\epsilon m * \varphi\|_{\infty} \leq \|\tilde{\eta}_\epsilon m\|_{L^2(\mathbb{R}^2)} \|\varphi\|_{L^2(\Omega)}$$

a simple computation shows that

$$\begin{aligned} \|\tilde{\eta}_\epsilon m\|_{L^2(\mathbb{R}^2)}^2 &\leq \frac{\sqrt{\pi}}{\epsilon^m} \|D_1^{\alpha-1} D_2^\beta \eta\|_{\infty}^{1/2} \\ \text{and } \|\varphi\|_{L^2(\Omega)} &\leq \sqrt{|\Omega|} \|\varphi\|_{\infty} \end{aligned}$$

where $|\Omega|$ is the Lebesgue measure of Ω . Consequently

$$\int_{\Omega} D_1^\alpha D_2^\beta(\eta_\epsilon * u) \varphi dx \leq \frac{C(m, \eta)}{\epsilon^m} \|\varphi\|_{\infty} |Du|(\mathbb{R}^2) \quad (22)$$

where

$$C(m, \eta) = \sqrt{\pi |\Omega|} \max_{\alpha+\beta=m} \|D_1^\alpha D_2^\beta \eta\|_{\infty}^{1/2} \quad (23)$$

Taking the supremum in equation (22) over all $\varphi \in C_c^1(\Omega)$ such that $\|\varphi\|_{\infty} \leq 1$, we obtain by duality and a denseness argument that

$$\|D_1^\alpha D_2^\beta(\eta_\epsilon * u)\|_{L^1(\Omega)} \leq \frac{C(m, \eta)}{\epsilon^m} |Du|(\mathbb{R}^2)$$

\square

Suppose that Ω is endowed with a regular triangulation Δ_h of size h , and let $d \in \mathbb{N}$ be given. As a finite dimensional space, $\mathcal{S}_d^0(\Delta_h)$ is a closed and convex subset of $L^2(\Omega)$. Thus, the ROF functional has a unique minimizer in $\mathcal{S}_d^0(\Delta_h)$. Let $s_h^d(f)$ be the spline function defined by

$$s_h^d(f) = \arg \min_{u \in \mathcal{S}_d^0(\Delta_h)} \lambda J(u) + \frac{1}{2} \int_{\Omega} |u-f|^2 dx \quad (24)$$

We are ready to prove that our construction of minimum splines above yields a minimizing sequence for the ROF functional. Let h_n be a monotonically decreasing sequence of real numbers such that $h_n \searrow 0$ as $n \rightarrow \infty$. Let Δ_n be a quasi-regular triangulation of Ω with mesh size h_n and smallest angle θ_n . We have the following result:

Theorem 11. Suppose that the sequence of regular triangulations $\{\Delta_n\}_n$ is such that

$$\inf_{n \in \mathbb{N}} \theta_n > \theta > 0 \quad (25)$$

Given $d \in \mathbb{N}$, the sequence $\{s_n^d(f)\}_n$ defined by

$$s_n^d(f) = \arg \min_{u \in \mathcal{S}_d^0(\Delta_n)} \lambda J(u) + \frac{1}{2} \int_{\Omega} |u-f|^2 dx \quad (26)$$

is minimizing for the ROF functional $E_{\lambda}^f(u)$.

Proof. Let $T : BV(\Omega) \rightarrow BV(\mathbb{R}^2)$ be the extension operator associated with the $1-$ neighborhood of Ω , the existence of which is guaranteed by Theorem 2. We recall that T is also a bounded linear operator from $W^{1,1}(\Omega)$ into $W^{1,1}(\mathbb{R}^2)$, and for any $u \in BV(\Omega)$, Tu is supported on the the $1-$ neighborhood of Ω .

Let $0 < \epsilon < 1$ and $d \in \mathbb{N}$ be fixed. Let u_{λ}^f be the minimizer of $E_{\lambda}^f(u)$ and put $u_{\epsilon}^f = \eta_{\epsilon} * Tu_{\lambda}^f$. Let $s_{\epsilon}^f \in \mathcal{S}_d^0(\Delta_n)$ be as in Theorem 5. Then by Lemma 10, we have

$$\|u_{\epsilon}^f - s_{\epsilon}^f\|_{W^{1,1}(\Omega)} \leq C(d, \theta) \left(\frac{h_n}{\epsilon}\right)^d \quad (27)$$

where C depends solely on d and θ . Moreover, since $T : W^{1,1}(\Omega) \rightarrow W^{1,1}(\mathbb{R}^2)$ is bounded, and Tu is compactly supported for every u , it follows from the Poincaré inequality (equation (9)) that

$$\begin{aligned} \|u_{\epsilon}^f - s_{\epsilon}^f\|_{L^2(\Omega)} &\leq \|T(u_{\epsilon}^f - s_{\epsilon}^f)\|_{L^2(\mathbb{R}^2)} \\ &\leq C \int_{\mathbb{R}^2} |\nabla(T(u_{\epsilon}^f - s_{\epsilon}^f))| dx \\ &\leq C \|T(u_{\epsilon}^f - s_{\epsilon}^f)\|_{W^{1,1}(\mathbb{R}^2)} \\ &\leq C \|T\|_* \|u_{\epsilon}^f - s_{\epsilon}^f\|_{W^{1,1}(\Omega)} \end{aligned} \quad (28)$$

with C a universal constant depending only on Ω_1 the $1-$ neighborhood of Ω , and $\|T\|_*$ is the operator norm of T .

We now show that by choosing a suitable regularization scale ϵ , we achieve the convergence of $E_{\lambda}^f(s_n^d(f))$ to $E_{\lambda}^f(u_{\lambda}^f)$ as $n \rightarrow \infty$. In fact for any $\epsilon > 0$, we have

$$\begin{aligned} &E_{\lambda}^f(s_n^d(f)) - E_{\lambda}^f(u_{\lambda}^f) \\ &= \underbrace{E_{\lambda}^f(s_n^d(f)) - E_{\lambda}^f(s_{\epsilon}^f)}_{\leq 0} + E_{\lambda}^f(s_{\epsilon}^f) - E_{\lambda}^f(u_{\epsilon}^f) \\ &\quad + E_{\lambda}^f(u_{\epsilon}^f) - E_{\lambda}^f(u_{\lambda}^f) \\ &\leq E_{\lambda}^f(s_{\epsilon}^f) - E_{\lambda}^f(u_{\epsilon}^f) + E_{\lambda}^f(u_{\epsilon}^f) - E_{\lambda}^f(u_{\lambda}^f) \end{aligned}$$

So to finish the proof, it suffices to show that

$$E_{\lambda}^f(u_{\epsilon}^f) \rightarrow E_{\lambda}^f(u_{\lambda}^f)$$

and

$$E_{\lambda}^f(s_{\epsilon}^f) \rightarrow E_{\lambda}^f(u_{\lambda}^f)$$

as $n \rightarrow \infty$ for a suitable choice of ϵ . First, we observe that the convergence of $E_{\lambda}^f(u_{\epsilon}^f)$ to $E_{\lambda}^f(u_{\lambda}^f)$ follows from the fact that

$$u_{\epsilon}^f \xrightarrow[\epsilon \rightarrow 0]{L^2(\Omega)} u_{\lambda}^f$$

and by Lemma 4 applied to

$$Tu_{\lambda}^f : |Du_{\epsilon}^f|(\Omega) \xrightarrow[\epsilon \rightarrow 0]{} |DTu_{\lambda}^f|(\bar{\Omega}) = J(u)$$

Second, by the triangle inequality and direct calculation we have

$$\begin{aligned} &|E_{\lambda}^f(s_{\epsilon}^f) - E_{\lambda}^f(u_{\epsilon}^f)| \\ &\leq \lambda \int_{\Omega} |\nabla(s_{\epsilon}^f - u_{\epsilon}^f)| dx + \frac{1}{2} [\|s_{\epsilon}^f - u_{\epsilon}^f\|_{L^2(\Omega)}^2 \\ &\quad + 2\|u_{\epsilon}^f - f\|_{L^2(\Omega)} \|u_{\epsilon}^f - s_{\epsilon}^f\|_{L^2(\Omega)}] \end{aligned}$$

We continue to rearrange the second term to have

$$\begin{aligned} &|E_{\lambda}^f(s_{\epsilon}^f) - E_{\lambda}^f(u_{\epsilon}^f)| \\ &\leq \lambda \int_{\Omega} |\nabla(s_{\epsilon}^f - u_{\epsilon}^f)| dx + \frac{1}{2} \|s_{\epsilon}^f - u_{\epsilon}^f\|_{L^2(\Omega)} \\ &\quad \cdot (\|u_{\epsilon}^f - s_{\epsilon}^f\|_{L^2(\Omega)} + 2\|u_{\epsilon}^f - f\|_{L^2(\Omega)}) \\ &\leq [\lambda + \frac{1}{2} \|u_{\epsilon}^f - s_{\epsilon}^f\|_{L^2(\Omega)} + \|u_{\epsilon}^f - f\|_{L^2(\Omega)}] \end{aligned}$$

$$\begin{aligned}
& \cdot [\|u'_\epsilon - s'_\epsilon\|_{W^{1,1}(\Omega)} + \|u'_\epsilon - s'_\epsilon\|_{L^2(\Omega)}] \\
& \leq (1 + C\|T\|_*)[\lambda + \frac{C\|T\|_*}{2}\|u'_\epsilon - s'_\epsilon\|_{W^{1,1}(\Omega)} \\
& \quad + \|u'_\epsilon - f\|_{L^2(\Omega)}]\|u'_\epsilon - s'_\epsilon\|_{W^{1,1}(\Omega)}
\end{aligned}$$

where we have used the estimate (equation (28)).

Now, using the estimate (equation (27)) and letting $\epsilon = h_n^{1/4d}$, we infer from the latter inequality that

$$\begin{aligned}
|E'_\lambda(s'_\epsilon) - E'_\lambda(u'_\epsilon)| & \leq (1 + C\|T\|_*)C(d, \theta) \\
& \times [\lambda + C(d, \theta, T)h_n^{d-1/4} + C(f, u'_\lambda)]h_n^{d-1/4}
\end{aligned}$$

where

$$C(f, u'_\lambda) = \|f\|_{L^2(\Omega)} \sup_{0 < \epsilon < 1} \|u'_\epsilon\|_{L^2(\Omega)}$$

$$C(d, \theta, T) := \frac{C\|T\|_*C(d, \theta)}{2}$$

Thus, $E'_\lambda(s_n(f)) \rightarrow E'_\lambda(u'_\lambda)$ as $h_n \rightarrow 0$ and the proof is complete. \square

Remark 12. It is easy to construct a sequence of triangulation with vanishing mesh sizes for which condition (25) is satisfied. Starting from a triangulation Δ_0 of Ω with smallest angle θ_0 and mesh size h_0 , a sequence of triangulations Δ_n is generated via successive refinements as follows: Given Δ_n , we obtain Δ_{n+1} by subdividing each triangle $t \in \Delta_n$ into four triangles by connecting the midpoints of the edges of t . The resulting triangulation Δ_{n+1} has mesh size $h_0 2^{-n-1}$ and smallest angle θ_0 .

Corollary 13. *Under the assumptions of Theorem 11, the sequence $\{s_n^d(f)\}_n$ satisfies the following two properties:*

$$s_n^d(f) \xrightarrow{L^p(\Omega)} u'_\lambda \text{ as } n \rightarrow \infty, \quad \text{for any } p \in [1, 2] \quad (29)$$

and

$$J(s_n^d(f)) \rightarrow J(u) \quad \text{as } n \rightarrow \infty \quad (30)$$

Proof. Since Ω is a bounded domain it suffices to establish (29) for $p=2$. The result for $1 \leq p < 2$ follows from the fact that $L^2(\Omega)$ is canonically embedded into $L^p(\Omega)$. The case $p=2$ follows easily from Theorem 9 and Theorem 11. Indeed, owing to equation (18), we have for all $n \in \mathbb{N}$

$$\|s_n^d(f) - u'_\lambda\|_{L^2(\Omega)}^2 \leq 2(E'_\lambda(s_n^d(f)) - E'_\lambda(u'_\lambda))$$

thus by Theorem 11 above, we have

$$\|s_n^d(f) - u'_\lambda\|_{L^2(\Omega)}^2 \rightarrow 0$$

as $n \rightarrow \infty$. Finally, since

$$\begin{aligned}
& J(s_n^d(f)) - J(u) \\
& = \frac{1}{\lambda}[E'_\lambda(s_n^d(f)) - E'_\lambda(u'_\lambda)] + \frac{1}{2}\|u'_\lambda - f\|_{L^2}^2 \\
& \quad - \frac{1}{2}\|s_n^d(f) - f\|_{L^2}^2 \leq \frac{1}{\lambda}[E'_\lambda(s_n^d(f)) - E'_\lambda(u'_\lambda)] \\
& \quad + \frac{1}{2}\|u'_\lambda - s_n^d(f)\|_2\|u'_\lambda + s_n^d(f) - 2f\|_2 \\
& \leq \frac{1}{\lambda}(E'_\lambda(s_n^d(f)) - E'_\lambda(u'_\lambda))^{1/2} \\
& \quad \cdot [E'_\lambda(s_n^d(f)) - E'_\lambda(u'_\lambda) + \|u'_\lambda + s_n^d(f) - 2f\|_2]
\end{aligned}$$

and the sequence $\{\|s_n^d(f)\|_2\}_n$ is bounded, thanks to Theorem 11 taking the limit of the latter identity as $n \rightarrow \infty$ yields (30) and the proof is complete. \square

Remark 14. Bartels [5] establish Corollary 13 for the case $d=1$. Our result generalizes and is applicable to higher order finite elements under h -refinement for which property (equation (14)) holds with infinitely differentiable functions.

Remark 15. The results of Theorem 20 and Corollary 13 hold if we replace $S_d^0(\Delta_h)$ with $S_d^r(\Delta_h)$ in the definition of the spline minimizer $s_h^d(f)$ provided that the hypotheses of Theorem 8 hold. In particular, we must have $d \geq 3r + 2$ and a family of regular triangulations.

Fixed point relaxation algorithm

The challenge in computing with the ROF model stems from the fact that the objective functional E'_λ is not Gâteaux differentiable; so the solution cannot be described in terms of the first variation of E'_λ . The reason why we cannot use the standard machinery of calculus of variation to solve equation (24) is that the associated Lagrangian

$$L(\mathbf{p}, z, x) = \lambda|\mathbf{p}| + \frac{1}{2}(z - f)^2, \quad \forall (\mathbf{p}, z, x) \in \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^2$$

is not differentiable with respect to \mathbf{p} at the origin $\mathbf{p} = \mathbf{0}$. One way to mitigate this difficulty is to find a differentiable relaxation of the Lagrangian L such that the corresponding energy functional is a *perturbation* of $E'_\lambda(u)$; that approach has been successfully used on at least three occasions in the literature.^{8,9,27}

Following Chambolle and Lions,⁹ we let Φ_ϵ be the real-valued function defined on \mathbb{R}^2 by

$$\Phi_\epsilon(x) = \begin{cases} \frac{|x|^2}{2\epsilon} + \frac{\epsilon}{2} & \text{if } |x| \leq \epsilon \\ |x| & \text{if } |x| > \epsilon \end{cases} \quad (31)$$

and consider the optimization problem

$$\arg \min_{u \in \mathcal{S}_d^0(\Delta_h)} \left\{ E_{\lambda,\epsilon}^f(u) := \lambda \int_{\Omega} \Phi_\epsilon(\nabla u) dx + \frac{1}{2} \int_{\Omega} |u-f|^2 dx \right\} \quad (32)$$

It is easy to check that the functional $E_{\lambda,\epsilon}^f$ is strictly convex and lower semicontinuous on $\mathcal{S}_d^0(\Delta_h)$ with respect to the L^2 -norm. Consequently, the minimization problem (equation (32)) has a unique solution that we denote $s_h^d(f, \epsilon)$. We now show that the family of functional $E_{\lambda,\epsilon}^f$ converges uniformly to E_λ^f and the minimizers $s_h^d(f, \epsilon)$ converge to $s_h^d(f)$ as ϵ goes to zero.

Proposition 16. *The family of functionals $E_{\lambda,\epsilon}^f(u)$ converges uniformly in $\mathcal{S}_d^0(\Delta_h)$ to $E_\lambda^f(u)$ and $s_h^d(f, \epsilon) \xrightarrow{L^2(\Omega)} s_h^d(f)$ as $\epsilon \searrow 0$. Furthermore, under the assumptions of Theorem 11, we have $s_h^d(f, h^{1/4d}) \xrightarrow{L^2(\Omega)} u_\lambda^f$ as h goes to 0.*

Proof. Let Φ be the continuous function defined on \mathbb{R}^2 by $\Phi(x) = |x| = \sqrt{x_1^2 + x_2^2}$. It is easy to show that

$$\sup_{x \in \mathbb{R}^2} |\Phi_\epsilon(x) - \Phi(x)| \leq \epsilon$$

Therefore, for any $u \in \mathcal{S}_d^0(\Delta_h)$ we have the estimate

$$|E_{\lambda,\epsilon}^f(u) - E_\lambda^f(u)| \leq \lambda \int_{\Omega} |\Phi_\epsilon(\nabla u) - \Phi(\nabla u)| dx \leq \lambda |\Omega| \epsilon$$

and it follows that $E_{\lambda,\epsilon}^f$ converges uniformly in $\mathcal{S}_d^0(\Delta_h)$ to E_λ^f .

Next, we note that Theorem 9 remains true on $\mathcal{S}_d^0(\Delta_h)$. Therefore, rewriting equation (18) in $\mathcal{S}_d^0(\Delta_h)$ for $s_h^d(f)$, we obtain

$$\begin{aligned} & \|s_h^d(f, \epsilon) - s_h^d(f)\|_{L^2(\Omega)}^2 \\ & \leq 2(E_\lambda^f(s_h^d(f, \epsilon)) - E_\lambda^f(s_h^d(f))) \\ & \leq 2(E_{\lambda,\epsilon}^f(s_h^d(f, \epsilon)) - E_{\lambda,\epsilon}^f(s_h^d(f, \epsilon))) \\ & \quad + 2(E_{\lambda,\epsilon}^f(s_h^d(f, \epsilon)) - E_\lambda^f(s_h^d(f))) \\ & \leq 2(\lambda |\Omega| \epsilon + \underbrace{E_{\lambda,\epsilon}^f(s_h^d(f, \epsilon)) - E_{\lambda,\epsilon}^f(s_h^d(f))}_{\leq 0}) \\ & \quad + E_{\lambda,\epsilon}^f(s_h^d(f)) - E_\lambda^f(s_h^d(f)) \\ & \leq 4\lambda\epsilon |\Omega| \end{aligned}$$

Thus, $\|s_h^d(f, \epsilon) - s_h^d(f)\|_{L^2(\Omega)} \leq 2\sqrt{\lambda|\Omega|}\epsilon$, and it follows that $s_h^d(f, \epsilon)$ converges to $s_h^d(f)$ in $L^2(\Omega)$ as ϵ goes to 0. Finally, by the triangle inequality we have

$$\|s_h^d(f, h^{1/4d}) - u_\lambda^f\|_{L^2(\Omega)} \leq 2\sqrt{\lambda|\Omega|}h^{1/4d} + \|s_h^d(f) - u_\lambda^f\|_{L^2(\Omega)}$$

Taking the limit of the latter inequality as h goes to 0 and using Corollary 3.5, it follows that $s_h^d(f, h^{1/4d})$ converges to u_λ^f in $L^2(\Omega)$ as h goes to 0. \square

Unlike the spline ROF model (24) for which we cannot use the first variation to characterize the solution, the functional $E_{\lambda,\epsilon}^f$ associated with the relaxation problem (equation (32)) is Gâteaux differentiable. Therefore, the spline function $s_h^d(f)$ is characterized by:

Proposition 17. *A function $u \in \mathcal{S}_d^0(\Delta_h)$ is the minimizer of the functional $E_{\lambda,\epsilon}^f$ in $\mathcal{S}_d^0(\Delta_h)$ if and only if u satisfies the variational equation*

$$\lambda \int_{\Omega} \frac{1}{\epsilon \vee |\nabla u|} \nabla u \cdot \nabla s dx + \int_{\Omega} (u-f)s dx = 0 \quad (33)$$

for all $s \in \mathcal{S}_d^0(\Delta_h)$, where $a \vee b := \max(a, b)$.

Proof. First, we observe that $E_{\lambda,\epsilon}^f(u)$ is Gâteaux differentiable with directional derivatives at any point $u \in \mathcal{S}_d^0(\Delta_h)$ given by

$$\begin{aligned} \langle dE_{\lambda,\epsilon}^f(u), s \rangle &= \lambda \int_{\Omega} \frac{1}{\epsilon \vee |\nabla u|} \nabla u \cdot \nabla s dx + \int_{\Omega} (u-f)s dx \\ & \quad \forall s \in \mathcal{S}_d^0(\Delta_h) \end{aligned} \quad (34)$$

Therefore, u is a minimizer of $E_{\lambda,\epsilon}^f(u)$ in $\mathcal{S}_d^0(\Delta_h)$ if and only if $dE_{\lambda,\epsilon}^f(u) = 0$, i.e. (equation (33)) holds. \square

We now set up the variational equation (33) as a fixed point equation for a nonlinear operator on $\mathcal{S}_d^0(\Delta_h)$. Fix $u \in \mathcal{S}_d^0(\Delta_h)$ and define the bilinear form $L[u, \lambda]$ on $\mathcal{S}_d^0(\Delta_h)$ by

$$L[u, \lambda](v, w) := \lambda \int_{\Omega} \frac{1}{\epsilon \vee |\nabla u|} \nabla v \cdot \nabla w dx + \int_{\Omega} vw dx$$

By Markov Inequality (Theorem 6) $L[u, \lambda]$ is a continuous, coercive bilinear form on the Hilbert space $\mathcal{S}_d^0(\Delta_h)$ as a subspace of $L^2(\Omega)$. Thus, by Lax–Milgram Theorem⁵⁰ (Theorem 2.7.7), for any $f \in L^2(\Omega)$ the equation

$$L[u, \lambda](v, w) = \int_{\Omega} f w dx \quad \forall w \in \mathcal{S}_d^0(\Delta_h)$$

has a unique solution in $\mathcal{S}_d^0(\Delta_h)$ that we denote by $L[u, \lambda]f$. Moreover, since $L[u, \lambda]$ is symmetric $L[u, \lambda]f$ is characterized by

$$L[u, \lambda]f = \arg \min_{v \in \mathcal{S}_d^0(\Delta_h)} \lambda \int_{\Omega} \frac{|\nabla v|^2 dx}{\epsilon \vee |\nabla u|} + \int_{\Omega} |v-f|^2 dx \quad (35)$$

Hence for $f \in L^2(\Omega)$ fixed, we define the nonlinear operator

$$F: \begin{array}{ccc} \mathcal{S}_d^0(\Delta_h) & \rightarrow & \mathcal{S}_d^0(\Delta_h) \\ u & \mapsto & L[u, \lambda]f \end{array}$$

It is easily shown using the characterizing equation (35) for $L[u, \lambda]f$ and the dominated convergence theorem that the operator F is continuous. Furthermore, Proposition 17 above defines $s_{h,\epsilon}^d(f)$ as a fixed point of F . So we may compute $s_{h,\epsilon}^d(f)$ using the fixed point iterative scheme presented and analyzed below.

Algorithm 18. Start from any bounded nonnegative function $v^0 \in \mathcal{S}_d^0(\Delta_h)$ and for $n \geq 0$, let

$$u_{n+1} = \arg \min_{u \in \mathcal{S}_d^0(\Delta_h)} \lambda \int_{\Omega} v_n |\nabla u|^2 dx + \int_{\Omega} |u-f|^2 dx \quad (36a)$$

$$v_{n+1} := \arg \min_{0 < v \leq 1/\epsilon} \int_{\Omega} v |\nabla u_{n+1}|^2 + \frac{1}{v} dx = \frac{1}{\epsilon \vee |\nabla u_{n+1}|} \quad (36b)$$

A standard argument using Lax-Milgram Theorem (see Brezis, 39 Corollary 5.8 p. 140) shows that u_{n+1} is characterized by the variational equation

$$\lambda \int_{\Omega} v_n \nabla u_{n+1} \cdot \nabla s dx + \int_{\Omega} (u_{n+1} - f)s dx = 0 \quad (37)$$

for all $s \in \mathcal{S}_d^0(\Delta_h)$. The existence and uniqueness of u_{n+1} follows by observing that the bilinear form

$$L[u_n](u, v) := \int_{\Omega} \lambda v_n \nabla u \cdot \nabla v + uv dx$$

is continuous—thanks to Theorem 6—and coercive on $\mathcal{S}_d^0(\Delta_h) \times \mathcal{S}_d^0(\Delta_h)$ with respect to the L^2 -norm. Consequently, the equation (37) has a unique solution.

We fix $\epsilon > 0$ and for the sake of notation conciseness, consider the functional E defined by

$$E(u, v) = \int_{\Omega} \lambda \left(v |\nabla u|^2 + \frac{1}{v} \right) dx + \int_{\Omega} |u-f|^2 dx \quad (38)$$

It is easy to check that

$$u_{n+1} = \arg \min_{u \in \mathcal{S}_d^0(\Delta_h)} E(u, v_n) \quad (39a)$$

$$v_{n+1} = \arg \min_{0 < v \leq 1/\epsilon} E(u_{n+1}, v) \quad (39b)$$

Lemma 19. *The sequence $\{u_n\}_n$ is bounded in $H^1(\Omega)$ and satisfies for all $n \in \mathbb{N}$, and any $s \in \mathcal{S}_d^0(\Delta_h)$*

$$\|s - u_n\|_{L^2(\Omega)}^2 \leq E(s, v_{n-1}) - E(u_n, v_{n-1}) \quad (40)$$

In particular, we have for all $n \in \mathbb{N}$

$$\|u_{n+1} - u_n\|_{L^2(\Omega)}^2 \leq E(u_n, v_n) - E(u_{n+1}, v_{n+1}) \quad (41)$$

Proof. We observe that in view of Theorem 6, proving the boundedness of $\{u_n\}$ in $H^1(\Omega)$ is equivalent to proving its boundedness in $L^2(\Omega)$. Let $n \in \mathbb{N}$ be given. Then by definition of u_n , we have

$$\begin{aligned} E(u_n, v_{n-1}) &\leq E(0, v_{n-1}) \\ &= \|f\|_{L^2}^2 + \int_{\Omega} \frac{1}{v_{n-1}} dx \leq \|f\|_{L^2}^2 + \frac{|\Omega|}{\epsilon} \end{aligned}$$

Consequently, we get $\|u_n - f\|_{L^2(\Omega)}^2 \leq \|f\|_{L^2(\Omega)}^2 + \frac{|\Omega|}{\epsilon}$, and deduce by the triangle inequality that

$$\|u_n\|_{L^2(\Omega)} \leq 2\|f\|_{L^2(\Omega)} + \sqrt{\frac{|\Omega|}{\epsilon}}$$

We now show that $\|u_n - s\|_{L^2(\Omega)}^2 \leq E(s, v_{n-1}) - E(u_n, v_{n-1})$. For any $s \in \mathcal{S}_d^0(\Delta_h)$, we have

$$\begin{aligned} &E(s, v_{n-1}) - E(u_n, v_{n-1}) \\ &= \int_{\Omega} \lambda v_{n-1} (|\nabla s|^2 - |\nabla u_n|^2) + (|s-f|^2 - |u_n-f|^2) dx \\ &= \int_{\Omega} \lambda v_{n-1} |\nabla(s - u_n)|^2 + |s - u_n|^2 dx \\ &\quad + 2 \underbrace{\int_{\Omega} \lambda v_{n-1} \nabla u_n \cdot \nabla(s - u_n) + (u_n - f)(s - u_n) dx}_{=0 \text{ by (37)}} \\ &= \int_{\Omega} \lambda v_{n-1} |\nabla(s - u_n)|^2 + |s - u_n|^2 dx \\ &\geq \|s - u_n\|_{L^2(\Omega)}^2 \quad \text{since } v_{n-1} \geq 0 \end{aligned}$$

In particular for any $n \in \mathbb{N}$,

$$\begin{aligned} \|u_{n+1} - u_n\|_{L^2(\Omega)}^2 &\leq E(u_n, v_n) - E(u_{n+1}, v_n) \\ &= E(u_n, v_n) - E(u_{n+1}, v_{n+1}) \\ &\quad + \underbrace{E(u_{n+1}, v_{n+1}) - E(u_{n+1}, v_n)}_{\leq 0 \text{ by (39)}} \\ &\leq E(u_n, v_n) - E(u_{n+1}, v_{n+1}) \end{aligned}$$

Thus, the sequence $\{E(u_n, v_n)\}_n$ is monotone non-increasing and $\|u_n - u_{n+1}\|_{L^2(\Omega)} \rightarrow 0$. \square

Theorem 20. *The sequence $\{u_n\}_n$ constructed in Algorithm 4.3 converges in $L^2(\Omega)$ to the minimizer $s_h^d(f, \epsilon)$ of $E_{\lambda, \epsilon}^f(u)$.*

Proof. In view of Proposition 17, it suffices to show that any cluster point u of the sequence $\{u_n\}_n$ with respect to the L^2 -norm satisfies the Euler-Lagrange equation (33). To begin, we note that the sequence $\{u_n\}_n$ has at least one cluster point as a bounded sequence in a finite dimensional normed vector space.

Let u be any cluster point of $\{u_n\}_n$ in $L^2(\Omega)$ and $\{u_{n_k}\}_k$ a subsequence such that $u_{n_k} \xrightarrow{L^2(\Omega)} u$. Since

$$\|u_{n_{k+1}} - u_{n_k}\|_{L^2(\Omega)} \rightarrow 0$$

it follows that $u_{n_{k+1}} \xrightarrow{L^2(\Omega)} u$ as well. By Markov inequality—Theorem 6—we also have

$$u_{n_k} \xrightarrow[k \rightarrow \infty]{H^1(\Omega)} u, \quad \text{and} \quad u_{n_{k+1}} \xrightarrow[k \rightarrow \infty]{H^1(\Omega)} u$$

Therefore, by Lebesgue dominated convergence theorem, we get

$$v_{n_k} = \frac{1}{|\nabla u_{n_k}|} \wedge \frac{1}{\epsilon} \xrightarrow[k \rightarrow \infty]{L^2(\Omega)} \frac{1}{|\nabla u|} \wedge \frac{1}{\epsilon} = \frac{1}{\epsilon \vee |\nabla u|}$$

Next, we establish that u satisfies the variational equation

$$\lambda \int_{\Omega} \frac{1}{\epsilon \vee |\nabla u|} \nabla u \cdot \nabla s \, dx + \int_{\Omega} (u - f) s \, dx = 0 \quad (42)$$

for all $s \in \mathcal{S}_d^0(\Delta_h)$. Indeed by definition of $u_{n_{k+1}}$, there holds

$$\lambda \int_{\Omega} v_{n_k} \nabla s \cdot \nabla u_{n_{k+1}} \, dx + \int_{\Omega} (u_{n_{k+1}} - f) s \, dx = 0 \quad (43)$$

for all $k \in \mathbb{N}$. Since $\nabla u_{n_{k+1}}$ converges strongly to ∇u in $L^2(\Omega) \times L^2(\Omega)$ and $v_{n_k} \nabla s$ converges strongly to $\nabla s / (\epsilon \vee |\nabla u|)$, it follows that

$$\int_{\Omega} v_{n_k} \nabla s \cdot \nabla u_{n_{k+1}} \, dx \rightarrow \int_{\Omega} \frac{1}{\epsilon \vee |\nabla u|} \nabla u \cdot \nabla s \, dx \quad (44)$$

as $k \rightarrow \infty$. Similarly, as $u_{n_{k+1}}$ converges strongly to u in $L^2(\Omega)$, we infer that

$$\int_{\Omega} (u_{n_{k+1}} - f) s \, dx \rightarrow \int_{\Omega} (u - f) s \, dx \quad \text{as } k \rightarrow \infty \quad (45)$$

On passing to the limit as $k \rightarrow \infty$ in (43) and taking into account equations (44) and (45), we obtain equation (42) and the proof is complete. \square

Remark 21. Many choices of the function $\Phi_{\epsilon} \in C^1(\mathbb{R}^2)$ for constructing a relaxation of the ROF functional are possible; presumably any choice of a family of convex continuously differentiable functions that yields a uniform approximation of the Euclidian norm should do the trick. A common choice seen in the literature and introduced by Acar and Vogel⁸ is the function

$$\Phi_{\epsilon}(x) = \sqrt{\epsilon + |x|^2}$$

Remark 22. For $d=1$ the relaxation algorithm is not necessary as a direct algorithm for computing the minimizers is readily available. Indeed, in this case the objective functional reads

$$E_{\lambda}^f(s) = \sup_{\substack{q \in \mathcal{S}_0^{-1}(\Delta_h) \times \mathcal{S}_0^{-1}(\Delta_h) \\ |q|_T \leq 1, T \in \Delta_h}} \left\{ \lambda \int_{\Omega} \nabla s \cdot q \, dx + \frac{1}{2} \int_{\Omega} |s - f|^2 \, dx \right\} \quad (46)$$

and the ROF model over the continuous affine functions turns into the saddle-point problem

$$u_{\lambda}^f = \arg \min_{s \in \mathcal{S}_1^0(\Delta_h)} \sup_{\substack{q \in \mathcal{S}_0^{-1}(\Delta_h) \times \mathcal{S}_0^{-1}(\Delta_h) \\ |q|_T \leq 1, T \in \Delta_h}} \left\{ \lambda \int_{\Omega} \nabla s \cdot q \, dx + \frac{1}{2} \int_{\Omega} |s - f|^2 \, dx \right\} \quad (47)$$

One can then solve the latter problem using the first-order primal dual algorithm studied by Chambolle and Pock.⁵⁰ Indeed, Bartels⁵² has studied it in details and provided ample evidence of convergence.

Applications to digital image processing

In this section, we report the results of some numerical experiments done using the algorithm described above on digital images. It is well known (some of these observations have been confirmed by theory) that:

1. the ROF model is excellent on piecewise constant images up to a reduction in contrast;
2. finite difference algorithms for the ROF model are vulnerable to the staircase effect, whereby smooth regions are recovered as mosaics of piecewise constant subregions;
3. total variation based image enhancement methods are ineffective in discriminating textures from noise at well mixed scales.

Two examples illustrating the issues raised above are provided, using both the finite difference and spline methods. However, we will see that the staircase effect is tamed by the spline method.

A digital total variation spline model

A digital image is a quantization of a light intensity field. For our purposes here, we model digital images as samples or quantization of functions defined on a domain Ω . For example, an image with resolution $M \times N$ could be thought of as the evaluation on the grid $\{(i, j) : 1 \leq i \leq M, 1 \leq j \leq N\}$ of a function f defined on $\Omega = (1/2, M + 1/2) \times (1/2, N + 1/2)$; alternatively we could think of it as a sample of local averages of the function f on squares centered at (i, j) .

The algorithm described in the previous section assumes that f is a function on the continuum domain Ω , however, digital images are merely samples of such functions. Therefore for processing digital images with the ROF model on spline spaces, we should try to estimate the function f from its samples $\{f_i : 1 \leq i \leq P\}$. One way to do these is to use any of the penalized spline fitting method introduced by Awanou et al.⁵² The problem with this approach is that the preliminary estimation step significantly modifies the input data. When the estimated function is fed to the ROF model, we cannot easily discriminate the contribution of the total variation smoothing procedure on the final output.

In order to clearly illustrate the effect of the total variation smoothing procedure in digital image processing, we solve the following variant of the spline minimization problem (equation (24))

$$\arg \min_{s \in \mathcal{S}_d^r(\Delta_h)} \lambda \int_{\Omega} |\nabla s| dx + \frac{1}{2} \sum_{T \in \Delta_h} \sum_{x_i \in T} |s(x_i) - f_i|^2 \quad (48)$$

where P is the total number of pixels and $s(x_i)$ is the value of the spline function s at the pixel location $x_i \in \Omega$. We have simply replaced the continuum L^2 fidelity term with a discrete counterpart based on the available data. In general, the optimization problem (50) may not have a solution unless the pixel locations $\mathcal{D} = \{x_i \in \Omega : 1 \leq i \leq P\}$ are well distributed across the triangles of Δ_h .

Theorem 23. Suppose that the pixel locations

$$\mathcal{D} = \{x_i \in \Omega : 1 \leq i \leq P\}$$

are such that, the mapping

$$N_{\mathcal{D}}(s) = \left(\sum_{T \in \Delta_h} \sum_{x_i \in T} s(x_i)^2 \right)^{1/2} \quad (49)$$

is a norm on $\mathcal{S}_d^r(\Delta_h)$. Then there exists a unique spline function $s_h \in \mathcal{S}_d^r(\Delta_h)$ such that

$$s_h = \arg \min_{s \in \mathcal{S}_d^r(\Delta_h)} E_d(s)$$

where

$$E_d(s) := \lambda \int_{\Omega} |\nabla s| dx + \frac{1}{2} \sum_{T \in \Delta_h} \sum_{x_i \in T} |s(x_i) - f_i|^2$$

Proof. Notice that $E_d(s)$ is strictly convex and continuous on $\mathcal{S}_d^r(\Delta_h)$; therefore $E_d(s)$ has at most one minimizer in $\mathcal{S}_d^r(\Delta_h)$. Let $\{s_n\}_n$ be a minimizing sequence of E_d , i.e. $E_d(s_n)$ converges to $\inf_{s \in \mathcal{S}_d^r(\Delta_h)} E_d(s)$. The sequence $\{N_{\mathcal{D}}(s_n)\}_n$ is bounded, and since we assume that $N_{\mathcal{D}}$ is a norm and $\mathcal{S}_d^r(\Delta_h)$ is finite dimensional, it follows that any subsequence of $\{s_n\}_n$ has a convergent subsequence with respect to the norm $N_{\mathcal{D}}$. Now, if s^* is the limit of a subsequence of $\{s_n\}_n$, then by continuity of E_d we have $E_d(s^*) = \inf_{s \in \mathcal{S}_d^r(\Delta_h)} E_d(s)$ and s^* is a minimizer of E_d . Thus, the set of minimizers of E_d is non-empty. Finally, since E_d is strictly convex and limit points of $\{s_n\}_n$ are minimizers of E_d , we infer that the minimizing sequence $\{s_n\}_n$ converges to the unique minimizer s_h of E_d . \square

Remark 24. The condition (49) is equivalent to saying that the pixel locations set \mathcal{D} is determining for the space of piecewise polynomials $\mathcal{S}_d^{-1}(\Delta_h)$, that is every element $s \in \mathcal{S}_d^{-1}(\Delta_h)$ is uniquely determined by the values of $s|_T$ at the pixel locations $\mathcal{D}_T = \mathcal{D} \cap T$ for every $T \in \Delta_h$. Consequently, each triangle T should contain at least $(d+2)(d+1)/2$ pixel locations.

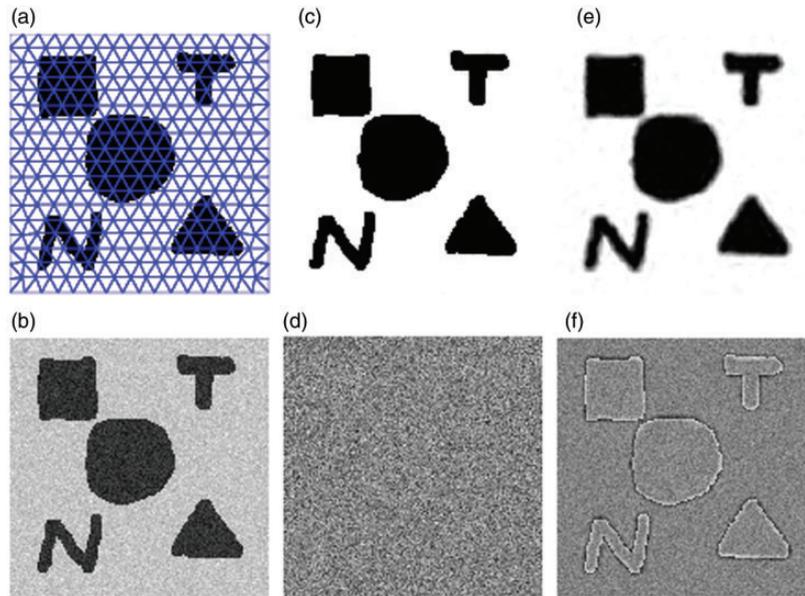


Figure 1. (a) The original cartoon image overlaid with a triangulation made of 562 triangles and 316 vertices. (b) The noised image obtained by adding a white noise with $\sigma = 25$ to the cartoon image. (c) The image recovered with the projected gradient algorithm. (d) The difference between the noised and recovered images. (e) The image recovered by fitting a continuous cubic spline over the triangulation in image (a). (f) The difference between the spline and noised images.

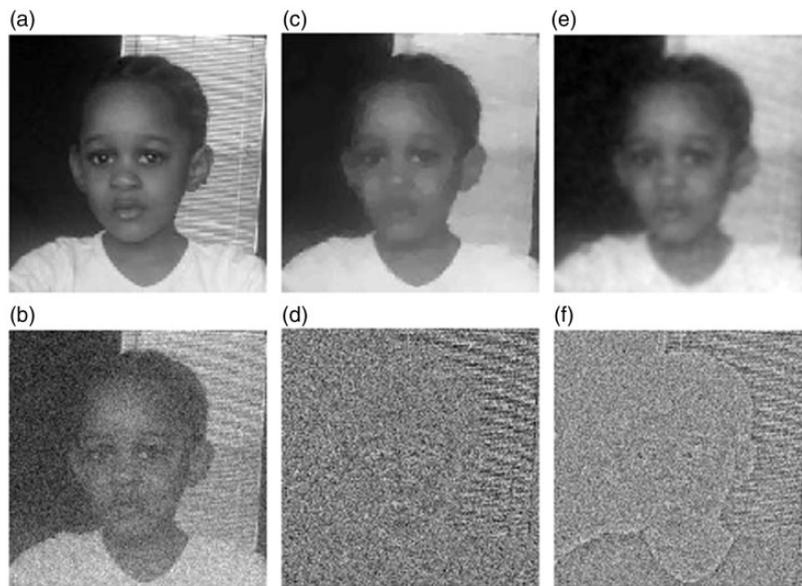


Figure 2. (a) A toddler portrait. (b) The noised image obtained by adding a white noise with $\sigma = 25$ to the image in (a). (c) The image recovered with the projected gradient algorithm. (d) The difference between the noised and recovered images. (e) The image recovered by fitting a continuous cubic spline over a mesh made of 1271 triangles. (f) The residual image of the spline algorithm.



Figure 3. The spline method is less vulnerable to the staircase effect. (a) Portion of the noised image in Figure 1 (b). (b) The image recovered using the projected gradient algorithm. (c) The image recovered with the spline method.

Remark 25. In practice the pixel locations are fixed by the image processing task. Therefore, given a choice of the degree d , condition (49) restricts our options of triangulations as well as the shape of the individual triangles as well. For example when denoising a $M \times N$ image, we may not use a triangulation containing more than $\frac{2MN}{(d+2)(d+1)}$ triangles.

Following section of Fixed point relaxation algorithm above, the actual computation is done by iteratively solving the sequence of quadratic programs

$$s_{n+1} = \arg \min_{s \in S'_d(\Delta_n)} \lambda \int_{\Omega} v_n |\nabla s|^2 dx + \sum_{T \in \Delta_n} \sum_{x_i \in T} |s(x_i) - f_i|^2 \quad (50)$$

where

$$v_n = \frac{1}{\epsilon \vee |\nabla s_n|} \quad \text{or} \quad v_n = \frac{1}{\sqrt{\epsilon + |\nabla s_n|^2}}$$

These expressions of v_n correspond to the relaxation derived from the functions

$$\Phi_{\epsilon}(x) = \begin{cases} \frac{|x|^2}{2\epsilon} + \frac{\epsilon}{2} & \text{if } |x| \leq \epsilon \\ |x| & \text{if } |x| > \epsilon \end{cases}$$

$$\text{and } \Phi_{\epsilon}(x) = \sqrt{\epsilon + |x|^2}, \text{ respectively.}$$

We shall refer to the relaxation associated to $\Phi_{\epsilon}(x) = \sqrt{\epsilon + |x|^2}$, the *minimal surface approximation* (MSA) spline method.

Image processing experiments

In this section we present the result of the numerical experiments for digital image processing situations. We consider denoising, in painting and image resizing. Each of these task can be done using total variation smoothing and can be formulated in the form given in equation (50).

Experiment 1 (Denoising a cartoon image).

In this test, we use the ROF model to clean up a Gaussian noise added to the a binary image made of five geometric shapes. For comparison purposes, we ran the spline algorithm 4.3 and the finite difference algorithm studied by the authors in Lai and Messi.²⁸ The spline algorithm is less capable to accurately resolve the edges than the finite difference algorithm, as seen by visually comparing panel (d) and panel (f) in Figure 1. The performance of the spline algorithm may be improved by choosing a triangulation that is

adapted to the edges in the image. However, generating such triangulations augment the computational cost of the algorithm as we would first identify the edges and the performance of the edge detection algorithm may be hampered by the noise in the data.

Experiment 2 (Denoising a natural image).

We now show the performance of the spline algorithm on a natural image with minor textures. $\epsilon = 1/20$, and $\tau = 1/8$. Both the projected gradient algorithm and the spline method effectively reduce the noise. The finite difference method produces sharper edges than the spline method, see shirt collar in panel (d) and panel (f) in Figure 2. However, the finite difference method results in an image with more blocky

Table 1. The PSNRs of various regions in Figure 4, Panel (b).

| | Perona–Malik | MSA spline |
|-----------------------|--------------|------------|
| Region 1 (red, left) | 29.9161 | 29.9109 |
| Region 2 (green) | 28.8445 | 28.9952 |
| Region 3 (yellow) | 31.6307 | 31.6989 |
| Region 4 (purple) | 33.0744 | 33.1167 |
| Region 5 (blue) | 33.3464 | 33.5832 |
| Region 6 (red, right) | 32.4490 | 32.5882 |

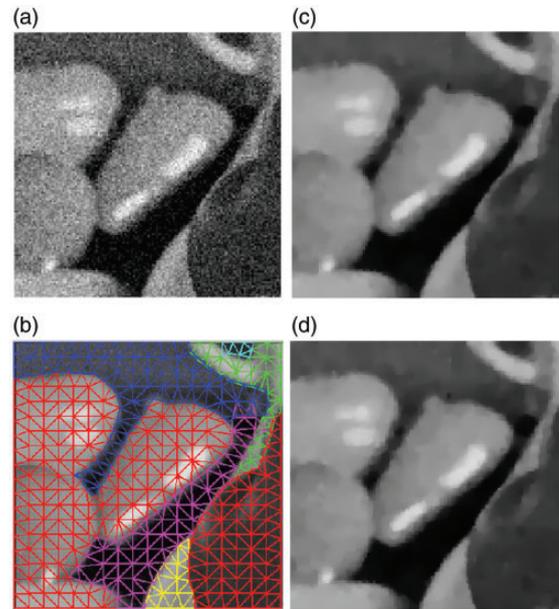


Figure 4. (a) The image to be cleaned was obtained by addition of a sample from a white noise with $\sigma = 20$ onto the ground truth image in panel (b). (b) Edge adapted triangulation of the image domain. (c) Recovered image using the MSA spline approximation with $\epsilon = 1$, $d = 6$ and $r = 1$. (d) Recovered image using Perona–Malik finite difference denoising. Original MSA Bicubic Bilinear.



Figure 5. Two images are scaled by a factor of 10 using the MSA spline method with $d=7$ and $r=1$. The results are compared to the bicubic and bilinear interpolation methods implemented in the MATLAB®'s function `imresize`.⁵³

regions than the one recovered by the spline method, see panel (b) and panel (c) in Figure 3.

Experiment 3 (Denoising a natural image with an edge-adapted triangulation). We show that by using a triangulation that aligns with the edges in the image, we make our approach very competitive with the finite difference method as measured by peak signal-to-noise ratio, see Table 1. In this experiment, the MSA spline method is run separately on the regions identified by a segmentation of the image, see Figure 4 panel (b).

Experiment 4 (Total variation image resizing). Image resizing consists in increasing/decreasing the resolution of a given digital image. We achieve this easily by fitting a spline function to the available image using (50) and evaluating the resulting spline over the pixel locations for the new resolution of the image. We illustrate this approach with two examples in Figure 5; each image in the first column is resized by a factor of 10, that is, the resulting images have 10 times more pixels along each direction than the original images. The first original image is of dimension 18×28 pixels, while the second one has a resolution of 15×24 pixels. One can notice that both the bicubic and bilinear interpolation methods lead to the Gibbs discontinuity effect at the edges of the two images, while our approach barely show any such effect.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

Jingyue Wang acknowledges supports from the National Natural Science Foundation of China under Grant 11771084, and from the Natural Science Foundation of Fujian under Grant 2016J01016.

Leopold Matamba Messi acknowledges support from the U. S. National Science Foundation under Grant DMS-0931642 to the Mathematical Biosciences Institute.

ORCID iD

Jingyue Wang  <http://orcid.org/0000-0002-3896-5137>

References

1. Rudin L, Osher S and Fatemi E. Nonlinear total variation based noise removal algorithms. *Physica D Nonlinear Phenomena* 1992; 60: 259–268.
2. Chambolle A, Caselles V, Cremers D, et al. An introduction to total variation for image analysis. In *Theoretical foundations and numerical methods for sparse recovery, volume 9 of Radon Ser. Comput. Appl. Math.* Berlin: Walter de Gruyter, 2010, pp. 263–340.
3. Caselles V, Chambolle A and Novaga M. Regularity for solutions of the total variation denoising problem. *Rev Mat Iberoam* 2011; 27: 233–252.
4. Allard WK. Total variation regularization for image denoising, I. Geometric theory. *Siam J Math Anal* 2007; 39: 1150–1190.
5. Allard WK. Total variation regularization for image denoising, II. Examples. *Siam J Imaging Sci* 2008; 1: 400–417.
6. Caselles V, Chambolle A and Novaga M. The discontinuity set of solutions of the TV denoising problem and some extensions. *Multiscale Model Simul* 2007; 6: 879–894.
7. Chambolle A, Duval V, Peyre G, et al. Geometric properties of solutions to the total variation denoising problem. *Inverse Prob* 2017; 33: 015002.
8. Acar R and Vogel CR. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Prob* 1994; 10: 1217–1229.
9. Chambolle A and Lions P-L. Image recovery via total variation minimization and related problems. *Numer Math* 1997; 76: 167–188.
10. Chambolle A. An algorithm for total variation minimization and applications. *J Math Imag Vision* 2004; 20: 89–97. Special issue on mathematics and image analysis.
11. Chambolle A. Total variation minimization and a class of binary MRF models. In A Rangarajan, B Vemuri and A Yuille (eds), *Energy minimization methods in computer vision and pattern recognition, volume 3757 of lecture notes in computer science.* Berlin/Heidelberg: Springer, 2005, pp. 136–152.

12. Aujol J-F and Dossal C. Stability of over-relaxations for the forward-backward algorithm, application to Fista. *Siam J Optim* 2015; 25: 2408–2433.
13. Beck A and Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J Imaging Sci* 2009; 2: 183–202.
14. Beck A. On the convergence of alternating minimization for convex programming with applications to iteratively reweighted least squares and decomposition schemes. *Siam J Optim* 2015; 25: 185–209.
15. Combettes PL and Wajs VR. Signal recovery by proximal forward-backward splitting. *Multiscale Model Simul* 2005; 4: 1168–1200.
16. Esser E, Zhang X and Chan TF. A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science. *Siam J Imaging Sci* 2010; 3: 1015–1046.
17. Goldstein T and Osher S. The split Bregman method for L1-regularized problems. *SIAM J Imaging Sci* 2009; 2: 323–343.
18. Zhu M and Chan TF. An efficient primal dual hybrid gradient algorithm for total variation image restoration. Technical Report 08-34, UCLA, Center for Applied Math, 2008.
19. Zhu M, Wright SJ and Chan TF. Duality-based algorithms for total-variation-regularized image restoration. *Comput Optim Appl* 2010; 47: 377–400.
20. Chambolle A and Pock T. An introduction to continuous optimization for imaging. *Acta Numerica* 2016; 25: 161–319.
21. Boykov Y, Veksler O and Zabih R. Fast approximate energy minimization via graph cuts. *IEEE Trans Pattern Anal Machine Intell* 2001; 23: 1222–1239.
22. Darbon J and Sigelle M. Image restoration with discrete constrained total variation—part I: fast and exact optimization. *J Math Imaging Vis* 2006; 26: 261–276.
23. Darbon J and Sigelle M. Image restoration with discrete constrained total variation—part II: Levelable functions, convex priors and non-convex cases. *J Math Imaging Vis* 2006; 26: 277–291. December
24. Goldfarb D and Yin W. Parametric maximum flow algorithms for fast total variation minimization. *SIAM J Sci Comput* 2009; 31: 3712–3743.
25. Lai M, Lucier B and Wang J. The convergence of a central-difference discretization of Rudin-Osher-Fatemi model for image denoising. *Scale Space Variat Methods Comput Vision* 2009; 5567: 514–526.
26. Wang J and Lucier BJ. Error bounds for finite-difference methods for Rudin-Osher-Fatemi image smoothing. *SIAM J Numer Anal* 2011; 49: 845–868.
27. Dobson DC and Vogel CR. Convergence of an iterative method for total variation denoising. *SIAM J Numer Anal* 1997; 34: 1779–1791.
28. Lai M-J and Messi LM. Piecewise linear approximation of the continuous Rudin-Osher-Fatemi model for image denoising. *SIAM J Numer Anal* 2012; 50: 2446–2466.
29. Matamba Messi L. *Theoretical and numerical approximation of the Rudin-Osher-Fatemi model for image denoising in the continuous setting*. PhD thesis, The University of Georgia, Athens, GA, 2012.
30. Bartels S. Broken Sobolev space iteration for total variation regularized minimization problems. *IMA J Numer Anal* 2016; 36: 493–502.
31. Bartels S. Error control and adaptivity for a variational model problem defined on functions of bounded variation. *Math Comp* 2015; 84: 1217–1240.
32. Bartels S and Milicevic M. Iterative finite element solution of a constrained total variation regularized model problem. *Discrete Cont Dynamical Syst Series S* 2017; 10: 1207–1232.
33. Chen K and Tai X-C. A nonlinear multigrid method for total variation minimization from image restoration. *J Sci Comput* 2007; 33: 115–138.
34. Feng X and Prohl A. Analysis of total variation flow and its finite element approximations. *Esaim: M2an* 2003; 37: 533–556.
35. Feng X, von Oehsen M and Prohl A. Rate of convergence of regularization procedures and finite element approximations for the total variation flow. *Numer Math* 2005; 100: 441–456.
36. Litvinov W, Rahman T and Tai X. A modified TV-stokes model for image processing. *SIAM J Sci Comput* 2011; 33: 1574–1597.
37. Stamm B and Wihler T. A total variation discontinuous Galerkin approach for image restoration. *Int J Numer Anal Model* 2015; 12: 81–93.
38. Tian WYi and Yuan X. Convergence analysis of primal-dual based methods for total variation minimization with finite element approximation. *J Sci Comput* 2018; 76: 243–274.
39. Xu J, Tai X-C and Wang L-L. A two-level domain decomposition method for image restoration. *IPI* 2010; 4: 523–545.
40. Brezis H. *Functional analysis, sobolev spaces and partial differential equations*. Universitext, New York: Springer, 2011. [Database][Mismatch]
41. Ambrosio L, Fusco N and Pallara D. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs, New York: The Clarendon Press, Oxford University Press, 2000.
42. Giusti E. *Minimal surfaces and functions of bounded variation, volume 80 of Monographs in Mathematics*. Birkhäuser Basel, New York, 1984.
43. Lai M-J and Schumaker LL. *Spline functions on triangulations, volume 110 of encyclopedia of mathematics and its applications*. Cambridge: Cambridge University Press, 2007.
44. Lai M-J, Liu C and Wenston P. Bivariate spline method for numerical solution of time evolution Navier-Stokes equations over polygons in stream function formulation. *Numer Methods Partial Differential Eq* 2003; 19: 776–827.
45. Lai M-J, Liu C and Wenston P. Numerical simulations on two nonlinear biharmonic evolution equations. *Appl Anal* 2004; 83: 563–577.
46. Lai M-J and Wenston P. Bivariate spline method for numerical solution of steady state Navier-Stokes

- equations over polygons in stream function formulation. *Numer Methods Partial Different Eq* 2000; 16: 147–183.
47. Lai M-J and Wenston P. Trivariate C^1 cubic splines for numerical solution of biharmonic equations. In *Trends in approximation theory (Nashville, TN, 2000)*, Innov. Appl. Math. Vanderbilt Univ. Press, 2001, pp. 225–234.
 48. Lai M-J, Wenston P and Ying LA. Bivariate C^1 cubic splines for exterior biharmonic equations. In *Approximation theory, X (St. Louis, MO, 2001)*, Innov. Appl. Math. Vanderbilt Univ. Press, Nashville, TN, 2002, pp. 385–404.
 49. Awanou G. Robustness of a spline element method with constraints. *J Sci Comput* 2008; 36: 421–432.
 50. Brenner SC and Scott LR. *The mathematical theory of finite element methods, volume 15 of texts in applied mathematics*. New York: Springer, 3rd edition, 2008.
 51. Chambolle A and Pock T. A first-order primal-dual algorithm for convex problems with applications to imaging. *J Math Imaging Vis* 2011; 40: 120–145.
 52. Bartels S. Total variation minimization with finite elements: convergence and iterative solution. *SIAM J Numer Anal* 2012; 50: 1162–1180.
 53. Awanou G, Lai M-J and Wenston P. The multivariate spline method for scattered data fitting and numerical solution of partial differential equations. In *Wavelets and splines: Athens 2005*, Mod. Methods Math., pp. 24–74. Brentwood, TN: Nashboro Press, 2006.
 54. The MathWorks, Inc. *MATLAB^[textregistered] (R2011a)*. Natick, Massachusetts, April 2011.