THE RAYLEIGH-RITZ METHOD FOR TOTAL VARIATION MINIMIZATION USING BIVARIATE SPLINE FUNCTIONS ON TRIANGULATIONS

QIANYING HONG*, MING-JUN LAI†, AND LEOPOLD MATAMBA MESSI‡

Abstract. Total variation smoothing methods have proven 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, the Rayleigh-Ritz method was applied to the total variation with L^2 penalty denoising model with smooth bivariate spline functions on triangulations as approximating spaces. Using the extension property of functions of bounded variation on Lipschitz domains, a convergent minimizing sequence of continuous bivariate spline functions of fixed degree for the TV-L^2 energy functional was constructed. An algorithm for computing spline minimizers was developed and its convergence studied.

Key words. Total variation; Rayleigh-Ritz method; Bivariate spline functions; Function with bounded variation.

AMS subject classifications. 65N06, 65N22, 97N50

1. Introduction. Rudin, Osher and Fatemi [28] proposed a constrained total variation minimization method for image enhancement. Suppose we have an image f filled with artifacts and we want to enhance the quality of our image while preserving its salient details as much as possible. Assuming that the image f is a function defined on a domain $\Omega \subset \mathbb{R}^2$, Rudin, Osher and Fatemi's approach is to solve the following penalized total variation minimization problem

$$\arg\min_{u\in L^2(\Omega)} \lambda\, J(u) + \frac{1}{2}\int_\Omega |u-f|^2\, dx, \tag{1.1}$$

where $J(u)$ is the total variation of u on Ω, and λ is a positive parameter controlling the fidelity of the recovered image to the initial image f. We shall refer to the minimization problem (1.1) as the ROF model, and denote its objective functional by

$$E^f_\lambda(u) := \lambda J(u) + \frac{1}{2}\int_\Omega |u-f|^2\, dx. \tag{1.2}$$

Notice that a minimal condition for the ROF model to be defined is that f be a square integrable function over Ω; thus the domain of $E^f_\lambda(u)$ is $BV(\Omega)\cap L^2(\Omega)$, where $BV(\Omega)$ stands for the space of functions of bounded variation over Ω.

The ROF model has been extensively investigated in the past two decades with most efforts going towards the development of efficient algorithms for digital images. On the theoretical side, the exact Euler-Lagrange partial differential equation was derived [11] and regularity results were obtained [8]. More precisely, Caselles et al. [8] proved that if f has modulus of continuity ω, then so does the minimizer of $E^f_\lambda(u)$ provided that Ω is convex. It is also known that regardless of the order of smoothness of the data function f, the solution of the ROF model is at best Lipschitz continuous. On the computational side of the model, an exact yet highly efficient algorithm [9, 10] was developed and its error rate derived [29, 18]. Dobson and Vogel [15, Theorem 2.2, p. 1782] gave a sufficient condition for the convergence of a Rayleigh-Ritz scheme for the ROF model. However, they also observed that the

*Department of Mathematics, University of Kansas, Lawrence, KS 66045 (anbrida@gmail.com)

†Department of Mathematics, The University of Georgia,Athens, GA 30602 (mjlai@math.uga.edu)

‡Mathematical Biosciences Institute, The Ohio State University, Columbus, OH 43210 (matambamessi.1@mbi.osu.edu)

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.

This work addresses that question by constructing a convergent Rayleigh-Ritz scheme regardless of the regularity of the solution of the ROF model. We point out that Bartels [6] had proved that a conforming Rayleigh-Ritz scheme based on continuous piecewise affine elements converge with no assumption on the regularity of the ROF minimizer. Also, Lai and Matamba [21, 27] studied a nonconforming Rayleigh-Ritz scheme with continuous piecewise affine elements for the ROF model that converged if $f \in \mathrm{Lip}(\beta, L^2(\Omega))$. Here, we show that the conforming scheme used by Bartels [6] can be extended to any continuous finite elements consisting of piecewise polynomials, and propose a new nonconforming scheme for the approximation of the ROF model. We develop a fixed point algorithm for computing the terms of the nonconforming approximation and provided numerical evidence that finite elements methods can successfully be used in digital image processing.

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 on triangulations. In section 3, we review relevant properties of the ROF model and prove the main results of this paper. Section 4 is devoted to study of a fixed-point algorithm for the nonconforming scheme. The last section 5 reports numerical experiments on various digital image processing task performed using the fixed-point algorithm analyzed in section 4; we make a case for the use of finite elements methods in digital image processing.

2. Preliminaries. 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.

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

$$J(u) := \sup \left\{ \int_\Omega u \,\mathrm{div}(\varphi) dx : \varphi \in C_c^1(\Omega, \mathbb{R}^2), |\varphi(x)| \leq 1, \forall x \in \Omega \right\} \tag{2.1}$$

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

$$J(u) = \int_\Omega |\nabla u| dx. \tag{2.2}$$

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

$$\|u\|_{BV} := \|u\|_{L^1} + J(u). \tag{2.3}$$

Furthermore, if $u \in BV(\Omega)$ then its distributional derivative [7], Du, is a finite vector-valued Radon measure on Ω, and its total variation $|Du|$ induces a Borel measure on Ω known as the total variation measure of u. We shall denote the total variation of u over of Borel set $B \subseteq \Omega$ by $\int_B |Du|$.

The following result asserts that a function u defined on a domain Ω of $\mathbb{R}^2$ with zero total variation must be constant.

THEOREM 2.1 (Poincaré Inequality [2, 16]). *Suppose that Ω is a bounded Lipschitz domain of $\mathbb{R}^2$. Then there exists a constant C depending only on Ω such that*

$$\|u - u_\Omega\|_{L^2(\Omega)} \leq C \int_\Omega |Du|, \quad \forall\, u \in BV(\Omega), \tag{2.4}$$

where $u_\Omega = \dfrac{1}{|\Omega|}\displaystyle\int_\Omega u(x)dx$ *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* $u \in BV(\mathbb{R}^2)$

$$\|u\|_{L^2(\mathbb{R}^2)} \leq C \int_{\mathbb{R}^2} |Du|. \tag{2.5}$$

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 2.2 ([2, 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* $\mathcal{A} : BV(\Omega) \to BV(\mathbb{R}^2)$ *such that the following hold:*

(a) For any $u \in BV(\Omega)$, $\displaystyle\int_\Gamma |D\mathcal{A}u| = 0$ *and the support of* $\mathcal{A}u$ *is contained in* A*.*

(b) The restriction of $\mathcal{A}$ *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 operator [7] from $\mathrm{W}^{1,1}(\Omega)$ to $W^{1,1}(\mathbb{R}^2)$ using a partition of unity argument. A complete proof is found in [2]. □

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

PROPOSITION 2.3. *The total variation functional* $J : L^1(\Omega) \to [0, +\infty]$ *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 semi-continuous,* i.e*, if* (u_n) *is a sequence which converges in* $L^1(\Omega)$ *to* u*, then*

$$J(u) \leq \liminf_{n\to\infty} J(u_n). \tag{2.6}$$

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. □

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 (2.6). By exploiting the extension property of functions of bounded variation (see Theorem 2.2 above), we will use the standard technique of convolution and the following lemma to achieve this goal.

LEMMA 2.4 ([16, Proposition 1.15]). *Suppose* $u \in BV(\Omega)$*. If* $A \subset\subset \Omega$ *is a relatively compact open subset of* Ω *such that*

$$\int_{\partial A} |Du| = 0, \tag{2.7}$$

then

$$\int_A |Du| = \lim_{\epsilon\to 0} \int_A |D(u * \eta_\epsilon)|, \tag{2.8}$$

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

We will use convolution to construct a sequence of smooth functions that converge to the minimizer of the ROF model and use the spline approximation theorems 2.6 and 2.7 to

derive the spline approximants. In so doing 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 2.5. *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*

$$\left\| D_1^\alpha D_2^\beta (\eta_\epsilon * u) \right\|_{L^1(\Omega)} \leq \frac{C}{\epsilon^m} |Du|(\mathbb{R}^2), \tag{2.9}$$

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, we have

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

Thus

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

Now by Hölder's inequality we have $\|\check{\eta}_\epsilon^m * \varphi\|_\infty \leq \|\check{\eta}_\epsilon^m\|_{L^2(\mathbb{R}^2)} \|\varphi\|_{L^2(\Omega)}$; a simple computation shows that

$$\|\check{\eta}_\epsilon^m\|_{L^2(\mathbb{R}^2)}^2 \leq \frac{\sqrt{\pi}}{\epsilon^m} \left\| D_1^{\alpha-1} D_2^\beta \eta \right\|_\infty^{1/2} \text{ and } \|\varphi\|_{L^2(\Omega)} \leq \sqrt{|\Omega|} \, \|\varphi\|_\infty,$$

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

$$\int_\Omega D_1^\alpha D_2^\beta (\eta_\epsilon * u)\varphi dx \leq \frac{C(m, \eta)}{\epsilon^m} \|\varphi\|_\infty |Du|(\mathbb{R}^2) \tag{2.10}$$

where

$$C(m, \eta) = \sqrt{\pi |\Omega|} \max_{\alpha+\beta=m} \left\| D_1^\alpha D_2^\beta \eta \right\|_\infty^{1/2}. \tag{2.11}$$

Taking the supremum in (2.10) over all $\varphi \in C_c^1(\Omega)$ such that $\|\varphi\|_\infty \leq 1$, we obtain by duality and a denseness argument that

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

□

2.2. Bivariate spline functions on triangulations. 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

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

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

$$\mathcal{S}_d^r(\Delta) = C^r(\Omega) \cap \mathcal{S}_d^{-1}(\Delta) = \{s \in C^r(\Omega) \colon s|_t \in \mathbb{P}_d, \ \forall T \in \Delta\}.$$

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 form [22]

$$\mathcal{S}_d^r(\Delta) \equiv \left\{\boldsymbol{c} \in \mathbb{R}^N \colon A_r \boldsymbol{c} = \mathbf{0}\right\}, \tag{2.12}$$

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 Setting up linux-libc-dev (3.2.0-61.92) ... 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 methods [20, 19, 25, 24, 26] and more recently for the numerical simulation of the Darcy-Stokes equation [4]. In general, spline functions may be utilized as approximation spaces to study some classes of variational equations using the Rayleigh-Ritz method. Their appeal to us in this work is twofold. Firstly, bivariate spline functions possess good approximation power in the Sobolev spaces $\mathrm{W}^{m,p}(\Omega)$ as illustrated by the following theorem.

THEOREM 2.6 ([22, 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*

$$\|D_1^\alpha D_2^\beta (u - s_u)\|_{L^p(\Omega)} \leq K h^{d+1-\alpha-\beta} |u|_{d+1,p} \ \forall\, 0 \leq \alpha + \beta \leq d, \tag{2.13}$$

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

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

Natural images contain structural information in the form of edges and textures. Resolving these entities with continuous spline functions will require a combination of fine triangulations and high degree spline functions 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 of low degree. The following result gives the relationship between the degree and the order of the spline function to guarantee suitable approximation in Sobolev spaces.

THEOREM 2.7 ([22, Theorem 10.10]). *Let r and d be two nonnegative integers such that $d \geq 3r+2$, 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*

$$\|D_1^\alpha D_2^\beta (u - s_u)\|_{L^p(\Omega)} \leq K h^{d+1-\alpha-\beta} |u|_{d+1,p} \ \ \forall\, 0 \leq \alpha+\beta \leq d, \tag{2.14}$$

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.

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. We will use it in section 4 to prove the existence and uniqueness of the solution of each step of the fixed point algorithm.

THEOREM 2.8 (Markov inequality [22, 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*

$$\|D_1^\alpha D_2^\beta s\|_{L^p(\Omega)} \leq \frac{K}{\rho^{\alpha+\beta}} \|s\|_{L^p(\Omega)}, \ \ \forall s \in \mathcal{S}_d^{-1}(\Delta), \tag{2.15}$$

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

3. Rayleigh-Ritz approximation with splines. In this section, we describe how we arrive at a family of continuous bivariate spline functions that approximate the minimizer of the functional

$$E_\lambda^f(u) := \lambda J(u) + \frac{1}{2}\int_\Omega |u-f|^2 dx. \tag{3.1}$$

Before undertaking the analysis of our approximation method, let us briefly explain why the ROF model is well posed. In fact, Proposition 2.3 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 $L^2(\Omega)$ for the norm of $L^2(\Omega)$. Therefore, the ROF model (1.1) has a unique solution and the problem is well posed as illustrated by the following result.

THEOREM 3.1. *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*

$$\left\| v - u_\lambda^f \right\|_{L^2}^2 \leq 2\left(E_\lambda^f(v) - E_\lambda^f(u_\lambda^f) \right) \tag{3.2}$$

and

$$\inf_{x\in\Omega} f(x) \leq u_\lambda^f(x) \leq \sup_{x\in\Omega} f(x) \quad \textit{for a.e. } x \in \Omega. \tag{3.3}$$

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

$$\|u_\lambda^f - u_\lambda^g\|_{L^2} \leq \|f-g\|_{L^2}. \tag{3.4}$$

Proof. The proof of (3.2) is a simple exercise of convex analysis and uses the characterization of the minimizer of a convex functional using subdifferentials; the inequality (3.4)

is a consequence of the inequality (3.2). Finally, the inequality (3.3) follows from the definition of the total variation of a function as the integral of the perimeters of its level-sets. A complete proof is found in [27]. □

The approximation of the minimizer of the ROF model by spline functions is possible because a function of bounded variation can be approximated by smooth functions and smooth functions are in turn well approximated by spline functions.

3.1. The conforming method. 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

$$s_h^d(f) = \underset{u \in \mathcal{S}_d^0(\Delta_h)}{\arg\min}\ \lambda J(u) + \frac{1}{2}\int_\Omega |u-f|^2 dx. \tag{3.5}$$

We 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 $n \to \infty$. 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 3.2. *Suppose that the sequence of quasi-regular triangulations $\{\Delta_n\}_n$ is such that*

$$\inf_{n\in\mathbb{N}} \theta_n > \theta > 0. \tag{3.6}$$

Given $d \in \mathbb{N}$, the sequence $\{s_n^d(f)\}_n$ defined by (3.5) *converges to the minimizer u_λ^f of the ROF functional E_λ^f on $L^2(\Omega)$.*

Proof. Choose an open neighborhood $\mathcal{O}$ of $\bar{\Omega}$ and let $T\colon BV(\Omega) \to BV(\mathbb{R}^2)$ be the extension operator associated with the $\mathcal{O}$, the existence of which is guaranteed by Theorem 2.2. We recall that T is also a bounded linear operator from $W^{1,1}(\Omega)$ into $W^{1,1}(\mathbb{R}^2)$. Moreover, for any $u \in BV(\Omega)$, Tu is supported on $\mathcal{O}$ and $DTu(\bar{\Omega}) = J(u)$.

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 2.6. Then by Lemma 2.5, we have

$$\|u_\epsilon^f - s_\epsilon^f\|_{W^{1,1}(\Omega)} \le C(d,\theta)\left(\frac{h_n}{\epsilon}\right)^d, \tag{3.7}$$

where C depends solely on d and θ. Moreover, since $T : W^{1,1}(\Omega) \to W^{1,1}(\mathbb{R}^2)$ is bounded, and Tu is compactly supported for every u, it follows from the Poincaré inequality (2.5) that

$$\begin{aligned}\|u_\epsilon^f - s_\epsilon^f\|_{L^2(\Omega)} &\le \|T(u_\epsilon^f - s_\epsilon^f)\|_{L^2(\mathbb{R}^2)} \le C\int_{\mathbb{R}^2} |\nabla(T(u_\epsilon^f - s_\epsilon^f))|dx \\ &\le C\|T(u_\epsilon^f - s_\epsilon^f)\|_{W^{1,1}(\mathbb{R}^2)} \le C\|T\|_*\|u_\epsilon^f - s_\epsilon^f\|_{W^{1,1}(\Omega)},\end{aligned} \tag{3.8}$$

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 $n \to \infty$. In fact for any $\epsilon > 0$, we have

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

So to finish the proof, it suffices to show that $E^f_\lambda(u^f_\epsilon) \to E^f_\lambda(u^f_\lambda)$ and $E^f_\lambda(s^f_\epsilon) \to E^f_\lambda(u^f_\epsilon)$ as $n \to \infty$ for a suitable choice of ϵ. First, we observe that the convergence of $E^f_\lambda(u^f_\epsilon)$ to $E^f_\lambda(u^f_\lambda)$ follows from the fact that $u^f_\epsilon \xrightarrow[\epsilon\to 0]{L^2(\Omega)} u^f_\lambda$ and by Lemma 2.4 applied to Tu^f_λ: $|Du^f_\epsilon|(\Omega) \xrightarrow[\epsilon\to 0]{} |DTu^f_\lambda|(\overline{\Omega}) = J(u)$. Second, by the triangle inequality we have

$$\begin{aligned}
|E^f_\lambda(s^f_\epsilon) - E^f_\lambda(u^f_\epsilon)| &= \left|\lambda\left[\int_\Omega |\nabla s^f_\epsilon| dx - \int_\Omega |\nabla u^f_\epsilon| dx\right] + \frac{1}{2}\left[\|s^f_\epsilon - f\|^2_{L^2} - \|u^f_\epsilon - f\|^2_{L^2}\right]\right| \\
&\le \lambda \int_\Omega |\nabla(s^f_\epsilon - u^f_\epsilon)| dx + \frac{1}{2}\left[\|s^f_\epsilon - u^f_\epsilon\|^2_{L^2(\Omega)} + 2\|u^f_\epsilon - f\|_{L^2(\Omega)}\|u^f_\epsilon - s^f_\epsilon\|_{L^2(\Omega)}\right] \\
&\le \lambda \int_\Omega \left|\nabla(s^f_\epsilon - u^f_\epsilon)\right| dx + \frac{1}{2}\|s^f_\epsilon - u^f_\epsilon\|_{L^2(\Omega)}(\|u^f_\epsilon - s^f_\epsilon\|_{L^2(\Omega)} + 2\|u^f_\epsilon - f\|_{L^2(\Omega)}) \\
&\le \left[\lambda + \frac{1}{2}\|u^f_\epsilon - s^f_\epsilon\|_{L^2(\Omega)} + \|u^f_\epsilon - f\|_{L^2(\Omega)}\right]\left[\|u^f_\epsilon - s^f_\epsilon\|_{W^{1,1}(\Omega)} + \|u^f_\epsilon - s^f_\epsilon\|_{L^2(\Omega)}\right] \\
&\le (1 + C\|T\|_*)\left[\lambda + \frac{C\|T\|_*}{2}\|u^f_\epsilon - s^f_\epsilon\|_{W^{1,1}(\Omega)} + \|u^f_\epsilon - f\|_{L^2}\right]\|u^f_\epsilon - s^f_\epsilon\|_{W^{1,1}(\Omega)},
\end{aligned}$$

where we have used the estimate (3.8).

Now, using the estimate (3.7) and letting $\epsilon = h_n^{1/4d}$, we infer from the latter inequality that

$$|E^f_\lambda(s^f_\epsilon) - E^f_\lambda(u^f_\epsilon)| \le (1 + C\|T\|_*)C(d,\theta)\left[\lambda + C(d,\theta,T)h_n^{d-1/4} + C(f,u^f_\lambda)\right]h_n^{d-1/4},$$

where

$$C(f,u^f_\lambda) = \|f\|_{L^2(\Omega)} \sup_{0<\epsilon<1} \|u^f_\epsilon\|_{L^2(\Omega)} \text{ and } C(d,\theta,T) := \frac{C\|T\|_* C(d,\theta)}{2}.$$

Thus, $E^f_\lambda(s_n(f)) \to E^f_\lambda(u^f_\lambda)$ as $h_n \to 0$ and the proof is complete. □

REMARK 3.3. *It is easy to construct a sequence of triangulation with vanishing mesh sizes for which condition* (3.6) *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 3.4. *Under the assumptions of Theorem 3.2, the sequence $\{s^d_n(f)\}_n$ satisfies the following two properties:*

$$s^d_n(f) \xrightarrow{L^p(\Omega)} u^f_\lambda \text{ as } n \to \infty, \text{ for any } p \in [1,2], \tag{3.9}$$

and

$$J(s^d_n(f)) \to J(u) \text{ as } n \to \infty. \tag{3.10}$$

Proof. Since Ω is a bounded domain it suffices to establish (3.9) for $p = 2$. The result for $1 \le p < 2$ follows from the fact that $L^2(\Omega)$ is canonically embedded into $L^p(\Omega)$. The case $p = 2$ follows easily from Theorem 3.1 and Theorem 3.2. Indeed, owing to equation (3.2), we have

$$\forall\, n \in \mathbb{N}, \quad \|s^d_n(f) - u^f_\lambda\|^2_{L^2(\Omega)} \le 2\left(E^f_\lambda(s^d_n(f)) - E^f_\lambda(u^f_\lambda)\right);$$

thus by Theorem 3.2 above, we have $\|s_n^d(f) - u_\lambda^f\|_{L^2(\Omega)}^2 \to 0$ as $n \to \infty$. Finally, since

$$\begin{aligned} J(s_n^d(f)) - J(u) &= \frac{1}{\lambda}\left[E_\lambda^f(s_n^d(f)) - E_\lambda^f(u_\lambda^f) + \frac{1}{2}\|u_\lambda^f - f\|_{L^2}^2 - \frac{1}{2}\|s_n^d(f) - f\|_{L^2}^2\right] \\ &\leq \frac{1}{\lambda}\left[E_\lambda^f(s_n^d(f)) - E_\lambda^f(u_\lambda^f) + \frac{1}{2}\|u_\lambda^f - s_n^d(f)\|_2\|u_\lambda^f + s_n^d(f) - 2f\|_2\right] \\ &\leq \frac{1}{\lambda}\big(E_\lambda^f(s_n^d(f)) - E_\lambda^f(u_\lambda^f)\big)^{1/2}\left[E_\lambda^f(s_n^d(f)) - E_\lambda^f(u_\lambda^f) + \|u_\lambda^f + s_n^d(f) - 2f\|_2\right] \end{aligned}$$

and the sequence $\{\|s_n^d(f)\|_2\}_n$ is bounded, thanks to Theorem 3.2 taking the limit of the latter identity as $n \to \infty$ yields (3.10) and the proof is complete. □

REMARK 3.5. *Bartels [6] established equation* (3.9) *of Corollary 3.4 for the case* $d = 1$ *and* $p = 2$*. Our result generalizes and is applicable to higher order finite elements under* h*-refinement for which property* (2.13) *holds with infinitely differentiable functions.*

REMARK 3.6. *The results of Theorem 4.3 and Corollary 3.4 hold if we replace* $\mathcal{S}_d^0(\Delta_h)$ *with* $\mathcal{S}_d^r(\Delta_h)$ *in the definition of the spline minimizer* $s_h^d(f)$ *provided that the hypotheses of Theorem 2.7 hold. In particular, we must have* $d \geq 3r + 2$ *and a family of regular triangulations.*

3.2. The nonconforming method. The challenge in computing with the ROF model stems from the fact that the objective functional E_λ^f is not Gâteaux differentiable; so the solution cannot be characterized by the first variation. The reason being that the associated Lagrangian

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

is not differentiable with respect to $\boldsymbol{p}$ at the origin $\boldsymbol{p} = \mathbf{0}$. 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 $E_\lambda^f(u)$; that approach has been successfully implemented for the ROF model on at least three occasions in the literature [12, 1, 15].

Following Chambolle and Lions [12], we let Φ_ϵ be the real-valued function defined on $\mathbb{R}^2$ by

$$\Phi_\epsilon(x) = \begin{cases} \dfrac{|x|^2}{2\epsilon} + \dfrac{\epsilon}{2} & \text{if } |x| \leq \epsilon, \\ |x| & \text{if } |x| > \epsilon, \end{cases} \tag{3.11}$$

and consider the optimization problem

$$\underset{u \in \mathcal{S}_d^0(\Delta_h)}{\arg\min}\left\{E_{\lambda,\epsilon}^f(u) := \lambda\int_\Omega \Phi_\epsilon(\nabla u)dx + \frac{1}{2}\int_\Omega |u - f|^2 dx\right\}. \tag{3.12}$$

The functional $E_{\lambda,\epsilon}^f$ is strictly convex and lower semicontinuous on $\mathcal{S}_d^0(\Delta_h)$ for the L^2-norm. Consequently, the minimization problem (3.12) 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. Moreover by choosing a suitable function $\epsilon(h)$, we show that $s_h^d(f, \epsilon(h))$ converge to u_λ^f as h goes to zero.

PROPOSITION 3.7. *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 3.2, 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

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

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

Next, we note that Theorem 3.1 remains true on $\mathcal{S}_d^0(\Delta_h)$. Therefore, rewriting equation (3.2) in $\mathcal{S}_d^0(\Delta_h)$ for $s_h^d(f)$, we obtain

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

Thus, $\big\|s_h^d(f,\epsilon) - s_h^d(f)\big\|_{L^2(\Omega)} \leq 2\sqrt{\lambda|\Omega|\epsilon}$, 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

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

Taking the limit of the latter inequality as h goes to 0 and using Corollary 3.4, it follows that $s_h^d(f,h^{1/4d})$ converges to u^f_λ in $L^2(\Omega)$ as h goes to 0. □

We close this section with a variational characterization of the nonconforming spline minimizer $s^d_{h,\epsilon}$. We note that the functional $E^f_{\lambda,\epsilon}$ associated with the relaxation problem (3.12) is Gâteaux differentiable; therefore the spline function $s_h^d(f,\epsilon)$ is characterized by

PROPOSITION 3.8. *A function $u \in \mathcal{S}_d^0(\Delta_h)$ is the minimizer of the functional $E^f_{\lambda,\epsilon}$ in $\mathcal{S}_d^0(\Delta_h)$ if and only if u satisfies the variational equation*

$$\lambda \int_\Omega \frac{1}{\epsilon \vee |\nabla u|} \nabla u \cdot \nabla s\, dx + \int_\Omega (u - f)s\, dx = 0 \quad \forall\, s \in \mathcal{S}_d^0(\Delta_h), \tag{3.13}$$

where $a \vee b := \max(a, b)$.

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

$$\langle dE^f_{\lambda,\epsilon}(u), s\rangle = \lambda \int_\Omega \frac{1}{\epsilon \vee |\nabla u|} \nabla u \cdot \nabla s\, dx + \int_\Omega (u - f)s\, dx \quad \forall s \in \mathcal{S}_d^0(\Delta_h). \tag{3.14}$$

Therefore, u is a minimizer of $E^f_{\lambda,\epsilon}(u)$ in $\mathcal{S}_d^0(\Delta_h)$ if and only if $dE^f_{\lambda,\epsilon}(u) = 0$, *i.e* (3.13) holds. □

4. Fixed point relaxation algorithm. In this section, we first establish that the spline minimizer $s_h^d(f,\epsilon)$ is a fixed-point of a nonlinear operator on $\mathcal{S}_d^r(\Delta_h)$ and derive and algorithm for computing $s_h^d(f,\epsilon)$.

Let $u \in \mathcal{S}_d^0(\Delta_h)$ be fixed 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 2.8), $L[u,\lambda]$ is a continuous and coercive bilinear form on the Hilbert space $\mathcal{S}_d^0(\Delta_h)$ as a topological subspace of $L^2(\Omega)$. Thus by Lax-Milgram Theorem, for any $f \in L^2(\Omega)$, the equation

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

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

$$L[u,\lambda]f = \operatorname*{arg\,min}_{v\in\mathcal{S}_d^0(\Delta_h)} E_{\lambda,\epsilon,u}(v) := \lambda \int_\Omega \frac{1}{\epsilon \vee |\nabla u|} \nabla v \cdot \nabla v \, dx + \int_\Omega |v-f|^2 dx. \tag{4.1}$$

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

$$\begin{aligned} F_\lambda : \quad \mathcal{S}_d^0(\Delta_h) &\to \mathcal{S}_d^0(\Delta_h) \\ u &\mapsto L[u,\lambda]f. \end{aligned}$$

We claim that F_λ is continuous. Indeed, if $\{u_n\} \subset \mathcal{S}_d^0(\Delta_h)$ is a sequence that converges to $u \in \mathcal{S}_d^0(\Delta_h)$, then E_{λ,ϵ,u_n} converges pointwise to $E_{\lambda,\epsilon,u}$ as $n \to \infty$. Next, since $E_{\lambda,\epsilon,u}$ is lower semicontinuous, it follows that E_{λ,ϵ,u_n} Γ-converges to $E_{\lambda,\epsilon,u}$ as $n \to \infty$; consequently $F_\lambda(u_n)$ converges to $F_\lambda(u)$ as $n \to \infty$ and F_λ is continuous. Furthermore, Proposition 3.8 above defines $s_h^d(f,\epsilon)$ as a fixed point of F. So we may compute $s_h^d(f,\epsilon)$ using a fixed point iteration.

ALGORITHM 4.1. *Start from any nonnegative function $v_0 \in \mathcal{S}_d^0(\Delta_h)$ and let*

$$u_{n+1} = \operatorname*{arg\,min}_{u\in S_d^0(\Delta_h)} \lambda \int_\Omega v_n |\nabla u|^2 \, dx + \int_\Omega |u-f|^2 \, dx \quad \forall n \geq 0, \tag{4.2a}$$

$$v_{n+1} := \operatorname*{arg\,min}_{0<v\leq 1/\epsilon} \int_\Omega v|\nabla u_{n+1}|^2 + \frac{1}{v} dx = \frac{1}{\epsilon \vee |\nabla u_{n+1}|}. \tag{4.2b}$$

A standard argument using Lax-Milgram Theorem (see [7, Corollary 5.8 p. 140]) shows that u_{n+1} is characterized by the variational equation

$$\lambda \int_\Omega v_n \nabla u_{n+1} \cdot \nabla s \, dx + \int_\Omega (u_{n+1} - f) s \, dx = 0, \quad \forall \, s \in \mathcal{S}_d^0(\Delta_h). \tag{4.3}$$

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 2.8– 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 (4.3) has a unique solution.

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

$$E(u,v) = \int_\Omega \lambda(v|\nabla u|^2 + \frac{1}{v})\,dx + \int_\Omega |u-f|^2\,dx. \tag{4.4}$$

It is easy to check that

$$u_{n+1} = \underset{u\in\mathcal{S}_d^0(\Delta_h)}{\arg\min}\, E(u,v_n) \text{ and } v_{n+1} = \underset{0<v\leq 1/\epsilon}{\arg\min} E(u_{n+1},v). \tag{4.5}$$

LEMMA 4.2. *The sequence $\{u_n\}_n$ is bounded in $H^1(\Omega)$ and satisfies*

$$\forall\, n\in\mathbb{N},\ \forall s\in\mathcal{S}_d^0(\Delta_h),\ \ \|s-u_n\|^2_{L^2(\Omega)} \leq E(s,v_{n-1}) - E(u_n,v_{n-1}). \tag{4.6}$$

In particular, we have

$$\forall n\in\mathbb{N},\quad \|u_{n+1}-u_n\|^2_{L^2(\Omega)} \leq E(u_n,v_n) - E(u_{n+1},v_{n+1}). \tag{4.7}$$

Proof. We observe that in view of Theorem 2.8, 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

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

Consequently, we get $\|u_n - f\|^2_{L^2(\Omega)} \leq \|f\|^2_{L^2(\Omega)} + \frac{|\Omega|}{\epsilon}$, 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 $\|u_n - s\|^2_{L^2(\Omega)} \leq E(s,v_{n-1}) - E(u_n,v_{n-1})$. For any $s\in\mathcal{S}_d^0(\Delta_h)$, we have

$$\begin{aligned}
E(s,v_{n-1}) - E(u_n,v_{n-1}) &= \int_\Omega \lambda v_{n-1}(|\nabla s|^2 - |\nabla u_n|^2) + (|s-f|^2 - |u_n-f|^2)\,dx \\
&= \int_\Omega \lambda v_{n-1}|\nabla(s-u_n)|^2 + |s-u_n|^2\,dx + \\
&\qquad \underbrace{2\int_\Omega \lambda v_{n-1}\nabla u_n\cdot\nabla(s-u_n) + (u_n-f)(s-u_n)\,dx}_{=0 \text{ by (4.3)}} \\
&= \int_\Omega \lambda v_{n-1}|\nabla(s-u_n)|^2 + |s-u_n|^2\,dx \\
&\geq \|s-u_n\|^2_{L^2(\Omega)} \quad \text{since } v_{n-1}\geq 0.
\end{aligned}$$

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

$$\begin{aligned}
\|u_{n+1}-u_n\|^2_{L^2(\Omega)} &\leq E(u_n,v_n) - E(u_{n+1},v_n) \\
&= E(u_n,v_n) - E(u_{n+1},v_{n+1}) + \underbrace{E(u_{n+1},v_{n+1}) - E(u_{n+1},v_n)}_{\leq 0 \text{ by (4.5)}} \\
&\leq E(u_n,v_n) - E(u_{n+1},v_{n+1}).
\end{aligned}$$

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

THEOREM 4.3. *The sequence $\{u_n\}_n$ constructed in Algorithm 4.1 converges in $L^2(\Omega)$ to the minimizer $s_h^d(f,\epsilon)$ of $E^f_{\lambda,\epsilon}(u)$.*

Proof. In view of Proposition 3.8, 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 (3.13). 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 $\|u_{n_k+1} - u_{n_k}\|_{L^2(\Omega)} \to 0$, it follows that $u_{n_k+1} \xrightarrow{L^2(\Omega)} u$ as well. By Markov inequality – Theorem 2.8 – we also have

$$u_{n_k} \xrightarrow[k\to\infty]{H^1(\Omega)} u \text{ and } u_{n_k+1} \xrightarrow[k\to\infty]{H^1(\Omega)} u.$$

Therefore, by Lebesgue dominated convergence theorem, we get

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

Next, we establish that u satisfies the variational equation

$$\lambda \int_\Omega \frac{1}{\epsilon \vee |\nabla u|} \nabla u \cdot \nabla s \, dx + \int_\Omega (u - f) s \, dx = 0, \quad \forall s \in \mathcal{S}_d^0(\Delta_h). \tag{4.8}$$

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

$$\lambda \int_\Omega v_{n_k} \nabla s \cdot \nabla u_{n_k+1} \, dx + \int_\Omega (u_{n_k+1} - f) s \, dx = 0, \quad \forall k \in \mathbb{N}. \tag{4.9}$$

Since ∇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 $\dfrac{\nabla s}{\epsilon \vee |\nabla u|}$, it follows that

$$\int_\Omega v_{n_k} \nabla s \cdot \nabla u_{n_k+1} \, dx \longrightarrow \int_\Omega \frac{1}{\epsilon \vee |\nabla u|} \nabla u \cdot \nabla s \, dx \text{ as } k \to \infty. \tag{4.10}$$

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

$$\int_\Omega (u_{n_k+1} - f) s \, dx \longrightarrow \int_\Omega (u - f) s \, dx \text{ as } k \to \infty. \tag{4.11}$$

On passing to the limit as $k \to \infty$ in (4.9) and taking into account (4.10) and (4.11), we obtain (4.8) and the proof is complete. □

REMARK 4.4. *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 Vogel [1] is the function*

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

See [15] for a detail analysis of an iterative method for this choice of Φ_ϵ, and the first author's dissertation [17] for a detail study of a fixed-point algorithm with respect to spline spaces.

REMARK 4.5. *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*

$$E_\lambda^f(s) = \sup_{\substack{q\in\mathcal{S}_0^{-1}(\Delta_h)\times\mathcal{S}_0^{-1}(\Delta_h)\\ \left|q\right|_T\Big|_2\leq 1,\ T\in\Delta_h}} \lambda\int_\Omega \nabla s\cdot q dx + \frac{1}{2}\int_\Omega |s-f|^2 dx, \tag{4.12}$$

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

$$u_\lambda^f = \arg\min_{s\in\mathcal{S}_1^0(\Delta_h)} \sup_{\substack{q\in\mathcal{S}_0^{-1}(\Delta_h)\times\mathcal{S}_0^{-1}(\Delta_h)\\ |q_T|_2\leq 1,\ T\in\Delta_h}} \lambda\int_\Omega \nabla s\cdot q dx + \frac{1}{2}\int_\Omega |s-f|^2 dx. \tag{4.13}$$

One can then solve the latter problem using the first-order primal dual algorithm studied by Chambolle and Pock [13]. Indeed, Bartels [6] has studied it in details and provided ample evidence of convergence.

5. 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 differences algorithms for the ROF model ar 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.

We will present examples addressing these issues. We expect the staircase effect to be reduced in the spline solution while the edges should not be well resolved due to the continuity of the spline function. Moreover, we anticipate the issue with textures to more pronounced in our method.

5.1. A semi-discrete total variation spline model. The algorithm described in the previous section assumes that f is a function on the continuum domain Ω; however, digital images are mere samples of such functions. Therefore, for processing digital images with the ROF model on spline spaces, we should estimate the function f from its samples $\{f_i\colon 1\leq i\leq P\}$. This could be done using any of the spline fitting method introduced by Awanou, Lai and Wenston [5]. The problem with that approach is that the preliminary estimation step significantly modifies the input data. When the estimated function is fed to the ROF model, we cannot discriminate the contribution of the total variation smoothing procedure on the final output.

In order to clearly illustrate the performance of total variation smoothing of digital images using spline functions, we solve the following variant of the spline minimization problem (3.5)

$$\arg\min_{s\in\mathcal{S}_d^r(\Delta_h)} \lambda\int_\Omega |\nabla s| dx + \frac{1}{2}\sum_{T\in\Delta_h}\sum_{x_i\in T}\left|s(x_i)-f_i\right|^2, \tag{5.1}$$

where $\mathcal{D} = \{x_i\in\Omega\colon 1\leq i\leq P\}$ are the loacations of known pixels values and $s(x_i)$ is the value of the spline function s at the pixel location $x_i\in\Omega$.

In general, the optimization problem (5.1) may not have a unique solution unless the pixel locations $\mathcal{D} = \{x_i \in \Omega\colon 1 \le i \le P\}$ are well distributed across the triangles of Δ_h in the sense of the following theorem.

THEOREM 5.1. *Suppose that the pixel locations $\mathcal{D} = \{x_i \in \Omega\colon 1 \le i \le P\}$ are such that*

$$\text{the mapping } N_D(s) = \left(\sum_{T\in\Delta_h} \sum_{x_i\in T} s(x_i)^2 \right)^{1/2} \text{ is a norm on } \mathcal{S}_d^r(\Delta_h). \tag{5.2}$$

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

$$s_h = \operatorname*{arg\,min}_{s\in\mathcal{S}_d^r(\Delta_h)} E_d(s) := \lambda \int_\Omega |\nabla s| dx + \frac{1}{2} \sum_{T\in\Delta_h} \sum_{x_i\in T} \big|s(x_i) - f_i\big|^2.$$

Proof. We note that in general the functional E_d is merely convex and continuous on $\mathcal{S}_d^r(\Delta_h)$; therefore E_d has at least one minimizer in $\mathcal{S}_d^r(\Delta_h)$. We claim that under the assumption (5.2), one such minimizer is the limit of a minimizing sequence. 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 $\{N_D(s_n)\}_n$ is bounded, and since we assumed that N_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_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. ∎

REMARK 5.2. *The condition* (5.2) *is equivalent to saying that the collection of pixel locations D is determining for the spline space $\mathcal{S}_d^r(\Delta_h)$, that is every element $s \in \mathcal{S}_d^{-1}(\Delta_h)$ is uniquely determined by the values of $s\big|_T$ 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. Therefore, given a choice of the degree d, condition* (5.2) *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 $\dfrac{2MN}{(d+2)(d+1)}$ triangles.*

Following section 4 above, the actual computation is done by iteratively solving the sequence of quadratic programs

$$s_{n+1} = \operatorname*{arg\,min}_{s\in\mathcal{S}_d^r(\Delta_h)} \lambda \int_\Omega v_n \big|\nabla s\big|^2 dx + \sum_{T\in\Delta_h} \sum_{x_i\in T} \big|s(x_i) - f_i\big|^2, \tag{5.3}$$

where

$$v_n = \frac{1}{\epsilon \vee |\nabla s_n|}.$$

This expression of v_n correspond to the relaxation derived from the function

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

5.2. Implementation of the algorithm for (5.3)**.** First, we identify the space $\mathcal{S}_d^{-1}(\Delta_h)$ of piecewise polynomial funttions of degree d on Δ_h to

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

where $\mathbb{P}_d^T$ is the vector space of polynomial of degree less than or equal to d with basis the Bernstein-Bezier polynomials $\mathcal{B}^T := \left\{B_{pqr}^{T,d} : p+q+r=d\right\}$ relative to T. In this setting, a spline function s is represented by its coefficients $\boldsymbol{c} = (\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_t)$ such that for each i, $\boldsymbol{c}_i$ are the coefficients of $s\big|_{T_i}$ with respect to $\mathcal{B}^T$. Furthermore, the spline space $\mathcal{S}_d^r(\Delta_h)$ is the kernel of some linear operator A_r on $\mathcal{S}_d^{-1}(\Delta_h)$.

Given an enumeration $\Delta_h = \{T_1, T_2, \ldots, T_t\}$ of the triangulation, the quadratic program (5.3) reduces to the coordinates constrained quadratic program

$$\underset{\boldsymbol{c}\in\mathbb{R}^{t\,m}}{\arg\min}\ \lambda\,\boldsymbol{c}^T\,K_n\,\boldsymbol{c} + \|M\boldsymbol{c}-\boldsymbol{f}\|_2^2 \quad \text{subject to } A_r\boldsymbol{c}=\boldsymbol{0}, \tag{5.4}$$

where K is a block diagonal matrix with blocks

$$K_{n,i} = \left[\int_{T_i} v_n \nabla B_\mu \cdot \nabla B_\nu\, dx\right]_{1\le\mu,\nu\le m}, \quad 1\le i\le t,$$

M is a $P\times t\,m$ matrix such that the ℓ-th row of M is the evaluation of the Bernstein-Bezier polynomials at x_ℓ for every triangle containing x_ℓ, and $\boldsymbol{f}$ is a column vector of length P containing the input image. In our implementation of the algorithm, the entries of $K_{n,i}$ are computed using a zero-th order quadrature formula.

We solve the constrained optimization problem (5.4) using the augmented Lagrangian method, leading us to the saddle-point system

$$\begin{bmatrix} \lambda K_n + M^TM & A_r^T \\ A_r & 0 \end{bmatrix}\begin{bmatrix}\boldsymbol{c}\\ \boldsymbol{\mu}\end{bmatrix} = \begin{bmatrix} M^T\boldsymbol{f}\\ \boldsymbol{0}\end{bmatrix}, \tag{5.5}$$

where $\boldsymbol{\mu}$ is the vector of Lagrange multipliers associated to the constraints $A_r\boldsymbol{c}=\boldsymbol{0}$. Notice that under the assumptions of Theorem 5.1, the matrix $\lambda K_n + M^TM$ is positive definite. Therefore, any off-the-shelf algorithm for solving saddle point matrix equations can be used to solve (5.5). If the assumption of Theorem 5.1 is not satisfied, we compute a least-square solution of (5.5).

We note that the size of the linear systems (5.5) grows quickly with the degree and the number of triangles. For example for a triangulation with 10000 triangles and for degree 5, MATLAB displays an *out-of-memory* message on a typical consumer laptop. To circumvent that in the experiments that we report in this paper, we used the domain decomposition method introduced by Lai and Schumaker in [23].

5.3. Image processing experiments. We present the results of numerical experiments for various digital image processing situations. Our goal here is to demonstrate that finite element methods could be used effectively for total variation based image processing. We consider denoising, inpainting and image resizing. Each of these tasks can be done using total variation smoothing and may be formulated as in equation (5.1). All the numerical results reported below are obtained using 10 iterations of the corresponding algorithm. No attempt was made to produce the best PSNR by tuning any of the parameters λ, the degree d of the spline, the triangulation, or the number of iterations.

Denoising a cartoon image. We clean up a realization of Gaussian noise added to an image made of five geometric shapes. For comparison purposes, we run the spline algorithm 4.1 and the finite difference algorithm studied by the authors in [21]. The spline algorithm recovers a smoother image than the finite difference algorithm (see the last two panels in the first row of Figure 5.1), and surprisingly resolves the edges better, too. This is illustrated in the method noise panels of Figure 5.1. The contours of the shapes are more pronounced in the method noise of the finite difference algorithm than in that of the spline algorithm. In Table 5.1 we provide more data documenting the competitiveness of the method when using splines of various degrees.

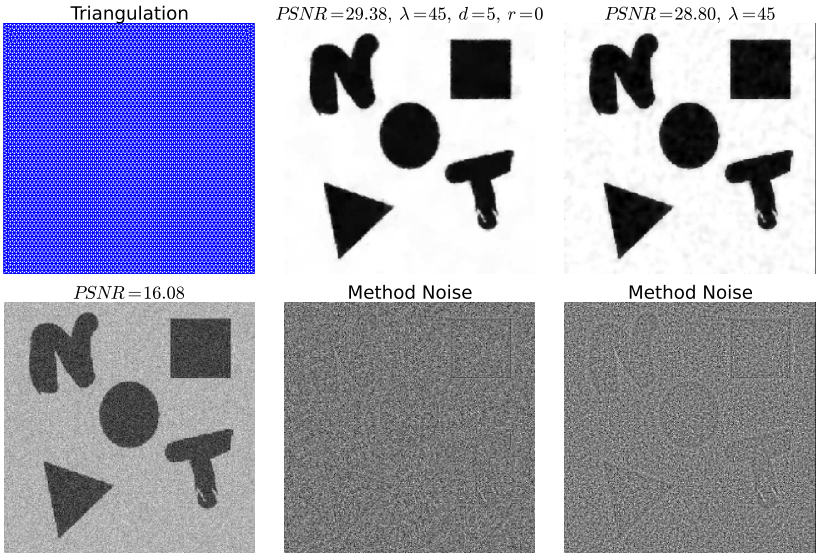

FIG. 5.1. *Bottom left: A noisy cartoon image to be cleaned* (256×256)*. Upper-left: the triangulation (4775 vertexes and 9275 triangles) used with the spline algorithm to clean the image. Top-middle: the image recovered using continuous quintic splines. Upper-right: the image recovered using the finite-difference algorithm. Method noise images are computed as the difference between the input image and the recovered image for each method.*

Denoising a natural image. We now show the performance of the spline algorithm on a natural image. Both the finite difference algorithm and the spline method effectively reduce the noise, see top row, second and third column of Figure 5.2. Also see Table 5.2 for more PSNR values obtained using splines of various degree.

Denoising a texture-rich image. This experiment aim is to demonstrate that the spline method studied in this paper remains competive even on images that are rich in textures. The input image is a noisy image of a woman. The head scarf, the table cloth, and the pants contain most of the textures. Both the spline and finite difference algorithms perform similarly on this test, see Table 5.3 for more data comparing the two methods.

Total variation based 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 (5.1) and evaluating the resulting spline at the pixel locations for the new resolution of the image. We illustrate this approach in Figure 5.4. An image (inset of each panel) is successively downsampled and upsampled. As expected, the total variation based rescaling suffer the same travails of any interpolation method for downsampling, see Figure 5.4 (Top panels).

d \ λ	15	25	45	5
	Spline Algorithm with $r = 0$			
2	27.91	27.03	26.28	29.14
3	30.46	28.78	26.73	32.79
4	31.18	29.23	27.95	34.84
5	32.46	30.46	29.38	35.41
6	33.05	30.86	30.14	35.45
7	32.68	30.44	29.87	34.71
8	31.80	29.65	29.71	33.89
	Finite Difference Algorithm			
	33.19	30.49	28.8	35.75
	PSNR of noisy input image			
	22.14	18.58	16.08	28.14

TABLE 5.1

Table of PSNR values for various degrees of spline functions. The triangulation used in all cases is the one shown in the upper-left panel of Figure 5.1. Even with an unstructured mesh like the one used here, the spline algorithm is very competitive with the finite difference algorithm.

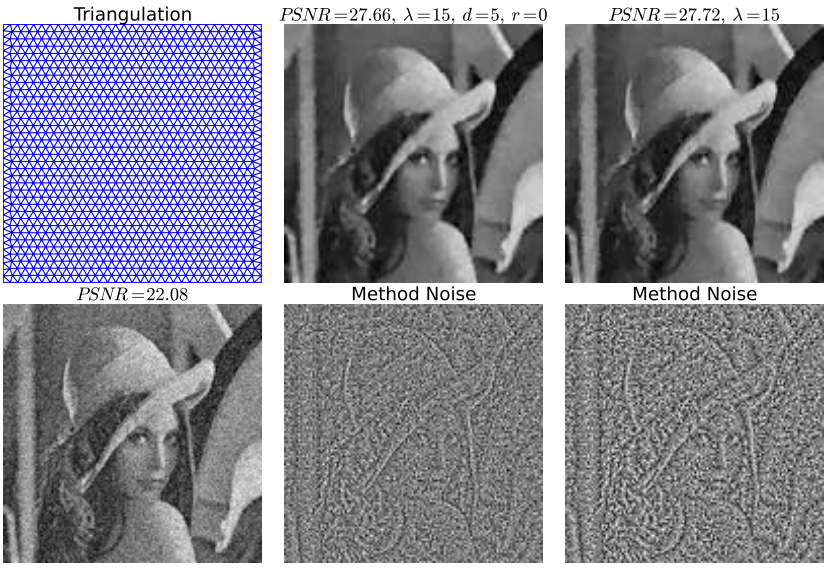

FIG. 5.2. *Bottom left: A noisy natural image (*128×128*) with* $\mathrm{PSNR} = 22.08$ *dB. Upper-left: the triangulation (1205 vertexes and 2270 triangles) used to clean the image with continuous quintic splines. Top-middle: the image recovered using continuous quintic splines. Upper-right: the image recovered using the finite-difference algorithm. Method noise images are computed as the difference between the input image and the recovered image for each method (bottom row middle and right).*

Total variation Image Inpainting. Image inpainting is an image recovery procedure by which missing or destroyed portion of the image are reconstructed. Many models of image inpainting have been proposed in the literature [3] including a total variation minimization model [14]. We use the spline method to do inpainting on a given image in the following three cases: (a) missing pixel values at random; (b) text removal from an image; (c) occlusion

d \ λ	15	25	45	5
	Spline Algorithm with $r = 0$			
2	24.91	23.60	21.91	27.12
3	25.89	24.13	22.26	29.28
4	26.72	25.2	23.45	30.23
5	27.66	25.79	23.80	31.26
6	27.98	25.97	24.07	31.71
7	27.87	25.87	24.11	31.46
8	27.66	25.71	24.25	31.15
	Finite Difference Algorithm			
	27.72	25.64	23.74	31.56
	PSNR of noisy input image			
	22.08	18.57	16.16	28.03

TABLE 5.2

Table of PSNR values for various degrees of spline functions. The triangulation used in all cases is the one shown in the upper-left panel of Figure 5.2. Even with an unstructured mesh like the one used here, the spline algorithm is very competitive with the finite difference algorithm. However, the spline method is very expensive computationally compared with the finite difference algorithm.

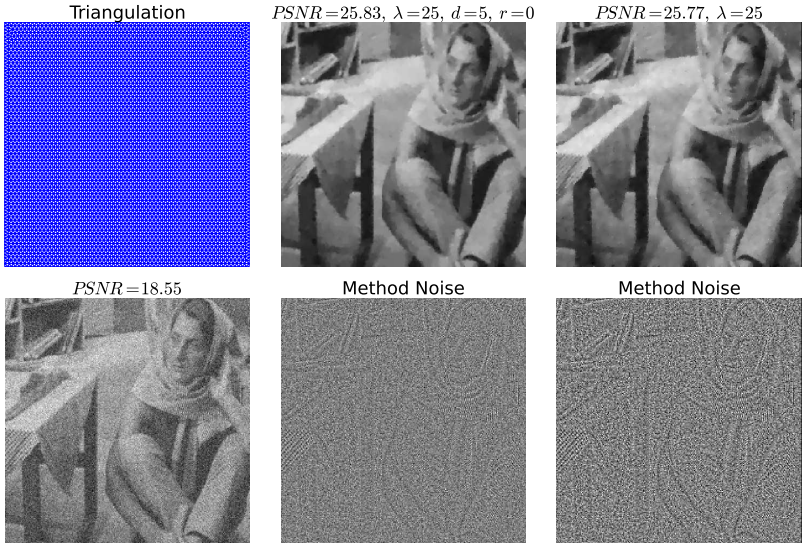

FIG. 5.3. Bottom left*: A noisy natural image with* PSNR $= 18.55$ *dB.* Top left*: the triangulation (4775 vertexes and 9275 triangles) used to clean the image with continuous quintic splines.* Top middle*: the image recovered using continuous quintic splines.* Top right*: the image recovered using the finite-difference algorithm. Method noise images are computed as the difference between the input image and the recovered image for each method (bottom row middle and right).*

removal. Total variation based inpainting does not correctly restore objects that are disconnected by the inpainting domain. This effect is more pronounced when the inpainting domain covers two heterogeneous region of the image, see the last row in Figure 5.6.

6. Conclusion. In this paper we studied the approximation of the total variation with L^2 penalty (TV-L^2) model for image denoising. The study was motivated by a recent result in

d \ λ	15	25	45	5
	Spline Algorithm with $r = 0$			
2	25.86	24.89	23.66	27.34
3	26.13	25.01	23.85	28.18
4	26.67	25.61	24.56	28.73
5	27.11	25.83	24.74	29.72
6	27.36	25.89	24.84	30.69
7	27.42	25.87	24.87	30.99
8	27.36	25.73	24.89	30.92
	Finite Difference Algorithm			
	27.30	25.77	24.6	30.88
	PSNR of input image			
	22.10	18.55	16.12	28.16

TABLE 5.3

Table of PSNR values for various degrees of spline functions. The triangulation used in all cases is the one shown in the upper-left panel of Figure 5.3. Even with an unstructured mesh like the one used here, the spline algorithm is very competitive with the finite difference algorithm.

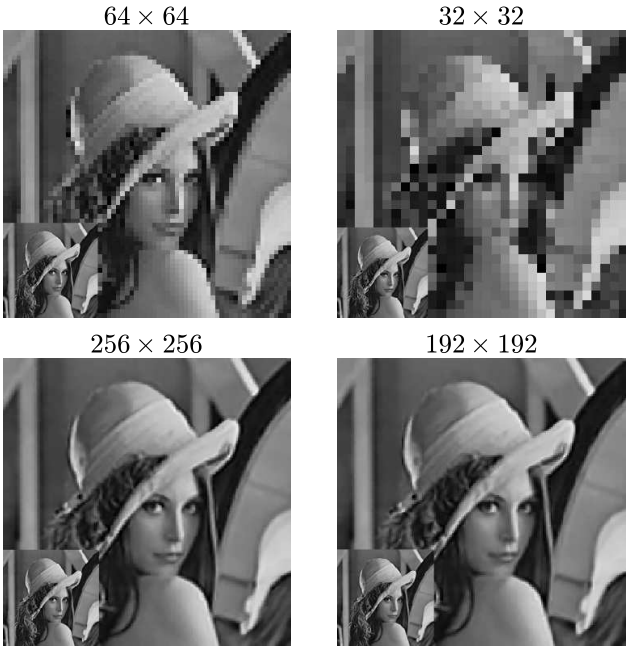

FIG. 5.4. *The inset image in each panel represents the original image. Its resolution is* 128×128 *pixels.* Top row*: From left to right the original image (inset) is downsampled to* 64×64 *and* 32×32*, respectively.* Bottom row*: the inset image is upsampled to* 256×256 *and* 192×192*, respectively.*

[8] demonstrating that the TV-L^2 model preserves the modulus of continuity of functions, and the earlier work [15] establishing a sufficient condition for the convergence of Rayleigh-Ritz approximation of the TV-L^2 model. Using the extension property of functions with bounded variation, we showed that the Rayleigh-Ritz method on bivariate spline spaces leads to a min-

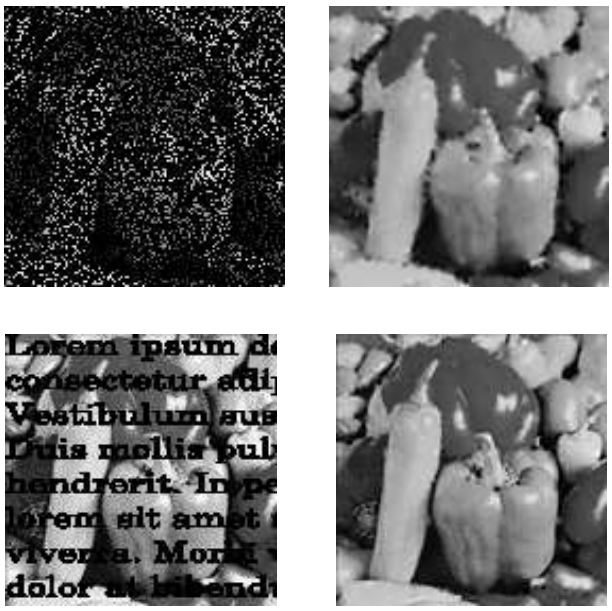

FIG. 5.5. First column*: An image with 80% of the pixels missing at random and another image overlayed with text.* Second column*: Both images are recovered using continuous quintic splines with* $\lambda = 1$.

imizing sequence of smooth spline functions of fixed degree for the TV-L^2 model. We formulated a relaxation algorithm for approximating the terms of the minimizing sequence of spline functions and studied its convergence. Numerous examples demonstrating the competitiveness of this algorithm on three distinct digital image processing tasks(denoising, rescaling, and inpainting) were provided.

Acknowledgements.. LMM acknowledges support from the U.S. National Science Foundation under Grant DMS-0931642 to the Mathematical Biosciences Institute.

REFERENCES

[1] R. Acar and C. R. Vogel. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Problems*, 10(6):1217–1229, 1994.
[2] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs, 2000.
[3] G. Aubert and P. Kornprobst. *Mathematical problems in image processing: partial differential equations and the calculus of variations*, volume 147 of *Applied Mathematical Sciences*. Springer-Verlag, New York, Second edition, 2006.
[4] G. Awanou. Robustness of a spline element method with constraints. *J. Sci. Comput.*, 36(3):421–432, 2008.
[5] G. Awanou, M.-J. Lai, and P. Wenston. The multivariate spline method for scattered data fitting and numerical solution of partial differential equations. In *Wavelets and splines: Athens 2005*, Mod. Methods Math., pages 24–74. Nashboro Press, Brentwood, TN, 2006.
[6] S. Bartels. Total variation minimization with finite elements: convergence and iterative solution. *SIAM J. Numer. Anal.*, 50(3):1162–1180, 2012.
[7] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.
[8] V. Caselles, A. Chambolle, and M. Novaga. Regularity for solutions of the total variation denoising problem. *Rev. Mat. Iberoam.*, 27(1):233–252, 2010.

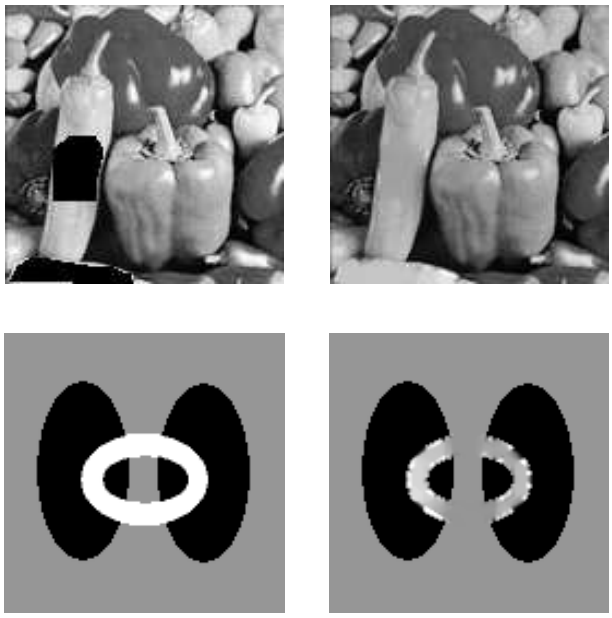

FIG. 5.6. First column*: An image with portions occluded by two black patches and another image occludes by a white annulus, respectively.* Second column*: We recovered the images using continuous quintic splines with* $\lambda = 1$*. A human would have thought of the image in the third column as two black ellipses on a gray background occluded by a white annulus. However, the total variation based inpainting does not agree with that interpretation.*

[9] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vision*, 20(1-2):89–97, 2004. Special issue on mathematics and image analysis.
[10] A. Chambolle. Total Variation Minimization and a Class of Binary MRF Models. In A. Rangarajan, B. Vemuri, and A. Yuille, editors, *Energy Minimization Methods in Computer Vision and Pattern Recognition*, volume 3757 of *Lecture Notes in Computer Science*, pages 136–152. Springer Berlin / Heidelberg, 2005.
[11] A. Chambolle, V. Caselles, D. Cremers, M. Novaga, and T. Pock. 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.*, pages 263–340. Walter de Gruyter, Berlin, 2010.
[12] A. Chambolle and P.-L. Lions. Image recovery via total variation minimization and related problems. *Numer. Math.*, 76(2):167–188, 1997.
[13] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vision*, 40(1):120–145, 2011.
[14] T. F. Chan and J. J. Shen. Variational image inpainting. *Communications on Pure and Applied Mathematics*, 58(5):579–619, 2005.
[15] D. C. Dobson and C. R. Vogel. Convergence of an iterative method for total variation denoising. *SIAM J. Numer. Anal.*, 34(5):1779–1791, 1997.
[16] E. Giusti. *Minimal Surfaces and Functions of Bounded Variation*, volume 80 of *Monographs in Mathematics*. Birkhäuser, 1984.
[17] Q. Hong. *Bivariate Splines for Image Enhancements based on Variational Models*. PhD thesis, The University of Georgia, Athens, GA 30602, 2011.
[18] M. Lai, B. Lucier, and J. Wang. The convergence of a central-difference discretization of rudin-osher-fatemi model for image denoising. *Scale Space and Variational Methods in Computer Vision*, pages 514–526, 2009.
[19] M.-J. Lai, C. Liu, and P. Wenston. Bivariate spline method for numerical solution of time evolution Navier-Stokes equations over polygons in stream function formulation. *Numer. Methods Partial Differential Equations*, 19(6):776–827, 2003.
[20] M.-J. Lai, C. Liu, and P. Wenston. Numerical simulations on two nonlinear biharmonic evolution equations. *Appl. Anal.*, 83(6):563–577, 2004.
[21] M.-J. Lai and L. M. Messi. Piecewise Linear Approximation of the Continuous Rudin-Osher-Fatemi Model

for Image Denoising. *SIAM J. Numerical Analysis*, 50(5):2446–2466, 2012.

[22] M.-J. Lai and L. L. Schumaker. *Spline functions on triangulations*, volume 110 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2007.

[23] M.-J. Lai and L. L. Schumaker. A domain decomposition method for comptuing bivariate spline fits of scattered data. *SIAM J. Numer. Anal.*, 47(2):911–928, 2009

[24] M.-J. Lai and P. Wenston. Bivariate spline method for numerical solution of steady state Navier-Stokes equations over polygons in stream function formulation. *Numer. Methods Partial Differential Equations*, 16(2):147–183, 2000.

[25] M.-J. Lai and P. Wenston. Trivariate C^1 cubic splines for numerical solution of biharmonic equations. In *Trends in approximation theory (Nashville, TN, 2000)*, Innov. Appl. Math., pages 225–234. Vanderbilt Univ. Press, 2001.

[26] M.-J. Lai, P. Wenston, and L.-A. Ying. Bivariate C^1 cubic splines for exterior biharmonic equations. In *Approximation theory, X (St. Louis, MO, 2001)*, Innov. Appl. Math., pages 385–404. Vanderbilt Univ. Press, Nashville, TN, 2002.

[27] L. Matamba Messi. *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 30602, 2012.

[28] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4):259–268, 1992.

[29] J. Wang and B. J. Lucier. Error Bounds for Finite-Difference Methods for Rudin-Osher-Fatemi Image Smoothing. *SIAM J. Numer. Anal.*, 49(2):845–868, 2011.