PIECEWISE LINEAR APPROXIMATION OF THE CONTINUOUS RUDIN-OSHER-FATEMI MODEL FOR IMAGE DENOISING

MING-JUN LAI †‡ AND LEOPOLD MATAMBA MESSI†§

Abstract. This paper is concerned with the numerical approximation of the minimizer of the continuous Rudin-Osher-Fatemi (ROF) model for image denoising. A new discrete total variation is proposed and the associated Hilbertian total variation denoising model is used to construct continuous piecewise linear functions that approximates the minimizer of the ROF model in the strong topology of $L^2(\Omega)$, provided that the data function is bounded and weakly regular in the sense of $\mathrm{Lip}(\alpha, \mathrm{L}^2(\Omega))$.

Key words. Total variation regularization; Variational methods; Finite-difference methods; Polynomial interpolation.

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

1. Introduction.

Since the seminal work of Rudin, Osher, and Fatemi[19] total variation based models for image restoration have received a great deal of attention. They are now used in image denoising, image deblurring, and image inpainting. For image denoising, the problem reads as follows:

$$\underset{u\in BV(\Omega)}{\operatorname{argmin}} \left\{J(u) := |Du|(\Omega)\right\}. \text{ subject to the constraints} \tag{1.1}$$

$$\int_\Omega u(x)dx = \int_\Omega f(x)dx \text{ and } \int_\Omega |u(x)-f(x)|^2 dx \le \sigma^2|\Omega|, \tag{1.2}$$

where $BV(\Omega)$ is the Banach space of functions of bounded variation, and the constraints indicate that the noise in f has mean zero and variance less than or equal to σ^2. It was shown in [11] that problem (1.1)–(1.2) is equivalent to the following unconstrained minimization

$$\underset{u\in BV(\Omega)}{\operatorname{argmin}} \left\{E_\lambda^f(u) := J(u) + \frac{1}{2\lambda}\int_\Omega (u-f)^2\, dx\right\} \tag{1.3}$$

where $\lambda > 0$ is a Lagrange multiplier. The existence and uniqueness of the minimizer of the unconstrained problem was established in [1] and [11].

To find "good" numerical approximations of the solution of (1.3), one has to devise an equally "good" discrete scheme for the total variation term $|Du|(\Omega)$. All the finite-difference methods proposed in the literature are based on the observation that (see for example [2, 14] for details)

$$|Du|(\Omega) = \int_\Omega |\nabla u| dx, \quad \forall u \in \mathrm{W}^{1,1}(\Omega). \tag{1.4}$$

As such, the total variation is then approximated using a combination of a quadrature formula for the integral, and a first order finite-difference approximation of the derivative.

The discrete analogue of the minimization problem (1.3) has received a lot of attention, and it has been shown that the best algorithm based on finite-difference schemes for denoising a digital image using the above model has convergence rate $\mathscr{O}(1/k^2)$, where k is the number of iterations. However, the numerical approximation of the solution of (1.3) has not seen much effort. The first works in that direction were published in [16] and [20].

†*Current Address*: Department of Mathematics, The University of Georgia, Athens, GA 30602.
‡*Email Address*: mjlai@math.uga.edu – *Tel*: 706-542-2065 – *Fax*: 706-542-2573
§*Email Address*: lmatamba@math.uga.edu – *Tel*: 706-542-5807 – *Fax*: 706-542-5907

In this paper, we were interested in constructing a convergent piecewise linear approximation of the solution of (1.3). To this end, we needed to introduce a novel finite-difference discretization of the total variation. We remind the reader that such discretizations in general do no make sense as functions of bounded variation are defined up to a set of measure zero, and in general do not have a continuous representative. Consequently, a finite-difference discretization of the functional $E_\lambda^f(u)$ requires a suitable discretization of the function u and the data function f.

Our contributions are the following: We propose a new discrete ROF model and formulate two algorithms for computing its solution. A piecewise linear approximation of the solution of the continuous model is then obtained by interpolating the solution of the discrete model. One advantage of our new discretized ROF model is that the continuous piecewise linear interpolation of the discrete solution converges to the solution of the continuous ROF model (1.3). Thus, the numerical solution of the new model is a reliable approximation of the solution of ROF model in the continuous setting. Not many discretized ROF models with this convergence property are available in the literature. To the authors' knowledge, only the schemes in [20], [16], and [10] are proven to have such a convergence property.

It is widely accepted in the literature that with a suitable discretization of L^1 functions, one can construct a discrete approximation of E_λ^f that Γ-converges to $E_\lambda^f(u)$ in $BV(\Omega)$ for an appropriate topology. The minimizers of such approximating functionals will then converge to the solution of the ROF model. However, no such construction has been proposed to our knowledge, nor has the $\Gamma-$convergence of the available schemes been established. This work is the first that constructs an approximation of the minimizer of (1.3) by continuous functions, and proves the convergence of the approximations.

The remainder of the paper is structured as follows: In section 2 we established the notations, recall relevant facts about functions of bounded variations, and present the new discrete total variation that we shall use to guarantee convergence of our approximations. Section 3 contains the main result of the paper and its proof. Finally, in section 4 we present numerical evidence of the convergence of our approximation in the special case where f is the characteristic function of a circle inside Ω. In this case the closed form of the minimizer is known, thus allowing us to demonstrate the convergence of our method numerically.

2. Preliminaries and notations. In this section we give preliminary results and introduce the notations that we shall use in the paper. Throughout the paper Ω shall denote the open set $(0,1)\times(0,1)$ unless otherwise noted, and Ω_m the open set $(-m,m)\times(-m,m)$, where m is a natural number.

2.1. Basic notations. The characteristic function of Ω is defined by

$$\mathbb{1}_\Omega(x) = \begin{cases} 1, & x \in \Omega, \\ 0, & x \notin \Omega. \end{cases}$$

For a given $\eta \in \mathbb{R}^2$, we shall denote by $\tau_\eta\Omega$ the image of the set Ω under the translation with the vector η, i.e

$$\tau_\eta\Omega := \{x+\eta : x \in \Omega\}.$$

For a function $u : \Omega \to \mathbb{R}$, we denote by $\tau_\eta u$ the function whose domain is $\tau_{-\eta}\Omega$ and is defined by

$$\tau_\eta u(x) = u(x+\eta), \quad x \in \tau_{-\eta}\Omega.$$

It is well known that the translation operator τ_η is a bounded linear operator from $L^p(\Omega)$ into $L^p(\tau_{-\eta}\Omega)$.

Let $h > 0$ be given. The $p-$modulus of continuity of order h, of a function $u \in L^p(\Omega)$, is defined by

$$\omega(u,h)_p = \sup_{|\eta|\leq h} \|\tau_\eta u - u\|_{L^p(\Omega\cap\tau_{-\eta}\Omega)}, \tag{2.1}$$

where $|\eta|$ stands for the Euclidean norm of η.

Let $u \in L^p_{\text{loc}}(\mathbb{R}^2)$ and $A \subset\subset \mathbb{R}^2$ a relatively compact open subset of $\mathbb{R}^2$. The $p-$modulus of continuity of u of order h with respect to A denoted $\omega_p(u,h)_A$, is defined by

$$\omega(u,h)_{p,A} = \omega(u\mathbb{1}_A,h)_p. \tag{2.2}$$

Let $0 < \alpha \leq 1$, we denote $\text{Lip}(\alpha, L^p(\Omega))$ the subspace of $L^p(\Omega)$ defined by

$$\text{Lip}(\alpha, L^p(\Omega)) := \left\{ u \in L^p(\Omega) : \sup_{0<h<1} h^{-\alpha}\omega(u,h)_p < \infty \right\}.$$

2.2. Functions of bounded variation. In this section, we recall without proofs the relevant facts about functions of bounded variation that shall be used in this paper. We follow the reference [14]; the interested reader is encouraged to refer to the above book and [2] for a thorough treatment of the concept of functions of bounded variation.

Let Ω be a bounded Lipschitz region in $\mathbb{R}^2$. A function $u\colon \Omega \to \mathbb{R}$ is said to be of *bounded variation* if $u \in L^1(\Omega)$ and

$$|Du|(\Omega) := \sup\left\{ \int_\Omega u \operatorname{div}(\varphi)\, dx : \varphi \in \mathscr{D}(\Omega,\mathbb{R}^2),\, |\varphi(x)| \leq 1, \forall x \in \Omega \right\}$$

is finite. The quantity $|Du|(\Omega)$ is called the *total variation* of u on Ω. The set of functions of bounded variation on Ω, denoted $BV(\Omega)$, is a Banach space for the norm

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

We now explain how one extends the total variation of a function $u \in BV(\Omega)$ into a finite positive Borel measure over Ω. Let $u \in BV(\Omega)$ be fixed. The total variation of u with respect to an open subset $A \subset \Omega$ is naturally given by

$$|Du|(A) = \sup\left\{ \int_\Omega u \operatorname{div}(\varphi)\, dx : \varphi \in \mathscr{D}(A,\mathbb{R}^n),\, |\varphi(x)| \leq 1, \forall x \in A \right\}. \tag{2.3}$$

Furthermore, if B is a general Borel subset of Ω, then we define the total variation of u over B by

$$|Du|(B) := \inf\{|Du|(O) : O \supset B \text{ and } O \text{ open}\}. \tag{2.4}$$

It can be shown that under the definition (2.4), $|Du|$ is a positive Borel measure on Ω which will be called *the total variation measure of u*. Consequently, by additivity of measures, the following identity holds for all Borel subset $K \subseteq \Omega$

$$|Du|(\Omega) = |Du|(\Omega \setminus K) + |Du|(K). \tag{2.5}$$

We recall relevant properties of functions of bounded variation. We start with a result asserting that a function of bounded variation has a trace on the boundary.

THEOREM 2.1 (Trace on the boundary, [14, Theorem 2.10]). *Let Ω be a bounded Lipschitz region in $\mathbb{R}^2$, and $u \in BV(\Omega)$. Then there exists a function $\gamma_0(u) \in L^1(\partial\Omega)$ such that for $\mathscr{H}^1$-almost all $x \in \partial\Omega$,*

$$\lim_{r\to 0} \frac{1}{r^2} \int_{\{z\in\Omega:\ |z-x|<r\}} |u(z) - \gamma_0(u)(x)| dz = 0. \tag{2.6}$$

Furthermore, for every $g \in C^1(\bar{\Omega}, \mathbb{R}^2)$

$$\int_\Omega u \operatorname{div}(g) dx = -\int_\Omega \langle g, Du\rangle + \int_{\partial\Omega} \gamma_0(u)\langle g, \nu\rangle d\mathscr{H}^1, \tag{2.7}$$

where ν is the unit outer normal to $\partial\Omega$, and $\mathscr{H}^1$ is the $1-$dimensional Hausdorff measure on $\mathbb{R}^2$.

The trace $\gamma_0(u)$ of a function $u \in BV(\Omega)$ is uniquely defined by the equation (2.6) and for $\mathscr{H}^1$-almost every $x \in \partial\Omega$

$$\gamma_0(u)(x) = \lim_{r\to 0} \frac{1}{|C(x,r)|} \int_{C(x,r)} u(z) dz, \tag{2.8}$$

where $C(x,r) = \{z \in \Omega\colon\ |z-x| < r\}$ and $|C(x,r)|$ is the Lebesgue measure of $C(x,r)$.

The next result allows to define extensions beyond Ω of functions of bounded variation on Ω. We will use it later in our work to define an extension via successive reflections of a function of bounded variation without creating new oscillations at the boundary of Ω.

LEMMA 2.2 (Pasting Lemma [14, Proposition 2.8]). *Let O be an open set such that $\Omega \subset\subset O$. Let $u_1 \in BV(\Omega)$, and $u_2 \in BV(O \setminus \bar{\Omega})$ be given. Then the function $u : O \to \mathbb{R}$ defined by*

$$u(x) = \begin{cases} u_1(x), & x \in \Omega \\ u_2(x), & x \notin \bar{\Omega} \end{cases}$$

is an element of $BV(O)$ and

$$|Du|(O) = |Du_1|(\Omega) + |Du_2|(O \setminus \bar{\Omega}) + \int_{\partial\Omega} |\gamma_0(u_1) - \gamma_0(u_2)| d\mathscr{H}^1. \tag{2.9}$$

Moreover, the total variation of u over the boundary of Ω is given by

$$|Du|(\partial\Omega) = \int_{\partial\Omega} |\gamma_0(u_1) - \gamma_0(u_2)| d\mathscr{H}^1. \tag{2.10}$$

Finally, we recall an alternate formula for the total variation of u over Ω that shall be instrumental in establishing a maximum principle like property for the minimizer of the ROF functional.

THEOREM 2.3 (Coarea formula [14, Theorem 1.23]). *Let a function $u \in BV(\Omega)$ be given and define for every $t \in \mathbb{R}$ the sub-level set of u at level t by*

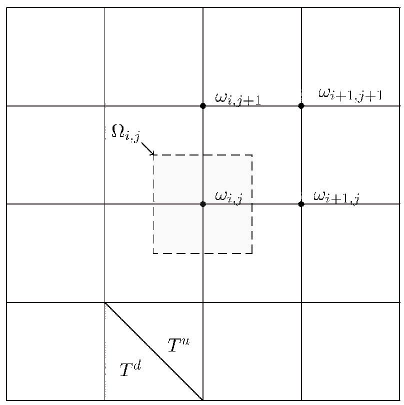
$$U_t := \{x \in \Omega\colon\ u(x) < t\}. \tag{2.11}$$

Then, the following identity holds

$$|Du|(\Omega) = \int_{-\infty}^{\infty} |D\mathbb{1}_{U_t}|(\Omega) dt. \tag{2.12}$$

2.3. A discrete approximation of the ROF functional. We propose a discrete approximation of the continuous ROF functional $E_\lambda^f(u)$ defined in (1.3). The closure $\bar{\Omega}$ of Ω is subdivided into N^2 square sub-domains of side length h yielding a uniform quadrangulation $\Box_h$. A triangulation Δ_h of Ω is constructed from $\Box_h$ by dividing each rectangle into two triangles using the Northwest-Southeast diagonal as shown in Figure 2.1.

Fig. 2.1: A type I triangulation of Ω: $T_{i,j}^u$ is the triangle with vertexes $\langle \omega_{i+1,j},\, \omega_{i+1,j+1},\, \omega_{i,j+1}\rangle$ and $T_{i,j}^d$ is the triangle with vertexes $\langle \omega_{i,j},\, \omega_{i+1,j},\, \omega_{i,j+1}\rangle$. $\Omega_{i,j}$ is used to discretize functions in $L^1(\Omega)$.



Let $\omega_{1,1}$ be lower left corner of $\bar{\Omega}$. We denote the set of vertexes of the triangulation Δ_h by

$$\mathscr{V}_h = \bar{\Omega} \cap \left\{\omega_{1,1} + h\mathbb{Z}^2\right\} := \{\omega_{i,j} : 1 \le i, j \le N\},$$

so that the (i,j)-th sub-square $\Omega_{i,j}$ is given by

$$\Omega_{i,j} := \Omega \cap \left(\omega_{i,j} + [-h/2, h/2]^2\right).$$

We are interested in devising a numerical scheme for computing an approximation of the minimizer of $E_\lambda^f(u)$. However, since Du is a measure, discrete approximation of $E_\lambda^f(u)$ solely based on u are a delicate matter. In the discrete setting, the situation is much simpler. The popular discrete counterpart of $E_\lambda^f(u))$ is inspired by the closed form of $|Du|(\Omega)$ given in (1.4). Assuming that a satisfactory discrete approximation of u has been designed, we introduce the discrete gradient operators $\nabla_+ = (\nabla_+^x, \nabla_+^y)$ and $\nabla_- = (\nabla_-^x, \nabla_-^y)$, defined by

$$\begin{aligned}
(\nabla_+^x u)_{i,j} &= \begin{cases} 0, & \text{if } i = N \text{ or } j = N \\ \dfrac{u_{i+1,j} - u_{i,j}}{h} & \text{otherwise;} \end{cases} \\
(\nabla_+^y u)_{i,j} &= \begin{cases} 0, & \text{if } i = N \text{ or } j = N \\ \dfrac{u_{i,j+1} - u_{i,j}}{h} & \text{otherwise;} \end{cases}
\end{aligned} \tag{2.13}$$

and

$$\begin{aligned}(\nabla_-^x u)_{i,j} &= \begin{cases} 0, & \text{if } i=1 \text{ or } j=1 \\ \dfrac{u_{i,j}-u_{i-1,j}}{h} & \text{otherwise;} \end{cases} \\ (\nabla_-^y u)_{i,j} &= \begin{cases} 0, & \text{if } i=1 \text{ or } j=1 \\ \dfrac{u_{i,j}-u_{i,j-1}}{h} & \text{otherwise.} \end{cases}\end{aligned} \tag{2.14}$$

We propose to compute the discrete total variation as follows

$$J_h(u) := \frac{h^2}{2} \sum_{1\le i,j\le N} |(\nabla_+ u)_{i,j}| + \frac{h^2}{2} \sum_{1\le i,j\le N} |\nabla_-(u)_{i,j})|, \tag{2.15}$$

where $|\cdot|$ is the Euclidean norm in $\mathbb{R}^2$. The advantage of this discrete total variation will be explained later after introducing more notation. The resulting approximation of the ROF-functional $E_\lambda^f(u)$ is given by

$$E_h^f(u) := J_h(u) + \frac{h^2}{2\lambda} \sum_{1\le i,j\le N} \left|u_{i,j} - f_{i,j}\right|^2 \quad u \in L^2(\Omega), \tag{2.16}$$

where $f_{i,j}$ is a suitable discrete approximation of the data function f.

We now describe a method for discretizing functions in $L^1_{\text{loc}}(\Omega)$. We also explain how one may construct an element of the space $L^p_{\text{loc}}(\Omega)$ from a discrete function on the lattice $\{\omega_{i,j} \colon 1 \le i,j \le N\}$.

2.4. Embedding and Projection Operators. In the previous section, we gave a formal discrete approximation of the ROF functional. We now clarify how this discretization is obtained.

On one hand, a function $u \in L^1_{\text{loc}}(\Omega)$ is discretized using the sampling operator, $Q_h : L^1_{\text{loc}}(\Omega) \to \mathbb{R}^{N\times N}$, defined by

$$(Q_h f)_{i,j} := \frac{1}{|\Omega_{i,j}|} \int_{\Omega_{i,j}} f(x)dx. \tag{2.17}$$

Q_h shall also denote the projection of $L^1_{\text{loc}}(\Omega)$ onto the space of piecewise constant functions with respect to the partition $\{\Omega_{i,j} \colon 1 \le i,j \le N\}$ of Ω, in which case Q_h is defined by

$$Q_h f(x) = \frac{1}{|\Omega_{i,j}|} \int_{\Omega_{i,j}} f(y)dy \quad \text{for all } x \in \overset{\circ}{\Omega}_{i,j}, \tag{2.18}$$

where $\overset{\circ}{\Omega}_{i,j}$ stands for the interior of $\Omega_{i,j}$.

On the other hand, a digital image $u \in \mathbb{R}^{N\times N}$ may be extended into a function $C_h u \in L^1_{\text{loc}}(\Omega)$ in a natural way as a piecewise constant function with respect to the quadrangulation $\{\Omega_{i,j} \colon 1 \le i,j \le N\}$ as follows:

$$C_h u(y) = u_{i,j}, \text{ if } y \in \overset{\circ}{\Omega}_{i,j}. \tag{2.19}$$

We shall also need a continuous interpolation of a digital image $u \in \mathbb{R}^{N\times N}$ which we define as the continuous piecewise linear function on Ω defined by

$$P_h u(y) = \sum_{1\le i,j\le N} u_{i,j} \phi_{i,j}(y), \tag{2.20}$$

with $\phi_{i,j} : \Omega \to \mathbb{R}$ the continuous piecewise linear function such that

$$\phi_{i,j}(\omega_{i,j}) = 1, \text{ and } \phi_{i,j}(\omega) = 0,\ \omega \in \mathscr{V}_h \setminus \{\omega_{i,j}\}. \tag{2.21}$$

LEMMA 2.4. *Suppose that Ω is endowed with the triangulation Δ_h. Then for all $u \in \mathbb{R}^{N\times N}$, there holds*

$$\|P_h u\|_{L^2}^2 \leq \|C_h u\|_{L^2}^2 + \frac{h^2}{12}(u_{1,N}^2 + u_{N,1}^2). \tag{2.22}$$

Proof. Let $u \in \mathbb{R}^{N\times N}$ be fixed. We first observe that $P_h u$ is the continuous bivariate spline of degree 1 over the triangulation Δ_h whose coefficients in the Bernstein-Bézier representation are $\{u_{i,j},\ 1 \leq i, j \leq N\}$. Therefore, using the closed form formula for the inner product of splines in Bernstein-Bézier form [17, Theorem 2.34], we get

$$\int_{T_{i,j}^u} P_h u(y)^2 dy = \frac{h^2}{24}\left(u_{i+1,j}^2 + u_{i+1,j+1}^2 + u_{i,j+1}^2 + (u_{i+1,j} + u_{i+1,j+1} + u_{i,j+1})^2\right),$$

and

$$\int_{T_{i,j}^d} P_h u(y)^2 dy = \frac{h^2}{24}\left(u_{i,j+1}^2 + u_{i,j}^2 + u_{i+1,j}^2 + (u_{i,j+1} + u_{i,j} + u_{i+1,j})^2\right).$$

Consequently, by the multinomial theorem and the elementary inequality $2ab \leq a^2 + b^2$, we have

$$\int_{T_{i,j}^u} P_h u(y)^2 dy \leq \frac{h^2}{6}\left(u_{i+1,j}^2 + u_{i+1,j+1}^2 + u_{i,j+1}^2\right) \tag{2.23}$$

and

$$\int_{T_{i,j}^d} P_h u(y)^2 dy \leq \frac{h^2}{6}\left(u_{i,j+1}^2 + u_{i,j}^2 + u_{i+1,j}^2\right). \tag{2.24}$$

Furthermore, a direct computation gives

$$\|C_h u\|_{L^2}^2 = h^2 \sum_{i,j=2}^{N-1} u_{i,j}^2 + \frac{h^2}{2} \sum_{\substack{i=2 \\ j\in\{1,N\}}}^{N-1} (u_{j,i}^2 + u_{i,j}^2) + \frac{h^2}{4} \sum_{i,j\in\{1,N\}} u_{i,j}^2. \tag{2.25}$$

Therefore, we have

$$\begin{aligned}
\|P_h u\|_{L^2}^2 &= \sum_{1\leq i,j\leq N-1} \int_{T_{i,j}^u} P_h u(y)^2 dy + \int_{T_{i,j}^d} P_h u(y)^2 dy \\
&\leq \frac{h^2}{3} \sum_{1\leq i,j<N} \left(u_{i,j+1}^2 + u_{i+1,j}^2\right) + \frac{h^2}{6} \sum_{1\leq i,j<N} \left(u_{i,j}^2 + u_{i+1,j+1}^2\right) \\
&= h^2 \sum_{i,j=2}^{N-1} u_{i,j}^2 + \frac{h^2}{2} \sum_{i=2}^{N-1} (u_{1,i}^2 + u_{i,1}^2 + u_{N,i}^2 + u_{i,N}^2) + \\
&\qquad\qquad + \frac{h^2}{6}(u_{1,1}^2 + u_{N,N}^2 + 2u_{1,N}^2 + 2u_{N,1}^2) \\
&\leq \|C_h u\|_{L^2}^2 + \frac{h^2}{12}(u_{1,N}^2 + u_{N,1}^2).
\end{aligned}$$

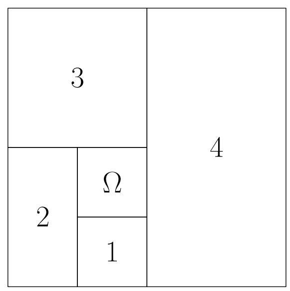
□

2.5. An extension of functions of bounded variation. In this section, we construct an extension of a function $u \in BV(\Omega)$ into a function $X[u] \in BV(\mathbb{R}^2)$ such that

$$|DX[u]|(\partial\Omega) = 0.$$

Let $u \in BV(\Omega)$ be given. The extension $X[u]$ of u to all of $\mathbb{R}^2$ is the function that equals zero outside the open set $\Omega_0 = \{x \in \mathbb{R}^2 : -1 < x_1, x_2 < 3\}$, and is defined on Ω_0 using four successive reflections of the function u in steps 1 through 4 starting across one side of Ω as shown in Figure 2.2.

Fig. 2.2: Extension of a function $u \in L^1(\Omega)$ by successive reflections starting across the boundary of Ω in four steps.



Let ρ be the radially symmetric function defined by

$$\rho(x) = \begin{cases} c\exp\left(\dfrac{1}{|x|^2-1}\right), & |x| < 1 \\ 0, & \text{otherwise,} \end{cases} \tag{2.26}$$

where the constant c is chosen such that $\displaystyle\int_{\mathbb{R}^2} \rho(x)dx = 1$. Let $\rho_\varepsilon(x) = \varepsilon^{-2}\rho\left(\dfrac{x}{\varepsilon}\right)$ be the corresponding family of mollifiers.

THEOREM 2.5. *Let $u \in BV(\Omega)$ be given. Then we have*

(a) $\left|DX[u]\right|(\partial\Omega) = 0$

(b) $\lim\limits_{\varepsilon\to 0} |D(X[u]*\rho_\varepsilon)|(\Omega) = |Du|(\Omega)$

Proof. Let $u \in BV(\Omega)$ be given, and u_0 the restriction of $X[u]$ to $O = \mathbb{R}^2 \setminus \bar{\Omega}$. Clearly $\partial O = \partial\Omega$ and it is easy to check that the trace of $\gamma_0(u_0) = \gamma_0(u)$. Since $X[u]$ is obtained by pasting u and u_0, it follows from the pasting Lemma 2.2 that $\left|DX[u]\right|(\partial\Omega) = 0$, and (a) is proved.

Next we show that (b) holds. Let $0 < \varepsilon < \operatorname{dist}(\bar{\Omega}, \partial\Omega_0)$ be fixed. It is easy to show from the definition of the total variation that

$$\left|D(X[u]*\rho_\varepsilon)\right|(\Omega) \le \left|DX[u]\right|(\Omega_\varepsilon) \text{ where } \Omega_\varepsilon := \{x \in \mathbb{R}^2 : \operatorname{dist}(x,\Omega) < \varepsilon\}.$$

Since $\Omega_\varepsilon \to \bar{\Omega}$, we infer from the latter inequality that

$$\limsup_{\varepsilon\to 0}\left|D(X[u]*\rho_\varepsilon)\right|(\Omega) \le \left|DX[u]\right|(\bar{\Omega}) = |Du|(\Omega) \text{ by part (a).}$$

Furthermore, we have $X[u] \xrightarrow{L^1(\Omega)} u$; thus, by lower semi-continuity of the total variation, we get

$$\liminf_{\varepsilon\to 0}\left|D(X[u]*\rho_\varepsilon)\right|(\Omega) \ge |Du|(\Omega).$$

Combining the last two inequalities above, we obtain $\lim_{\varepsilon\to 0}|D(X[u]*\rho_\varepsilon)|(\Omega) = |Du|(\Omega)$. □

REMARK 2.6. *It is known that property (b) of Theorem 2.5 above may not hold for the zero extension of u. A straightforward consequence of (b) above is the convergence of* $E_\lambda^f(X[u]*\rho_\varepsilon)$ *to* $E_\lambda^f(u)$ *as* ε *goes to zero.*

PROPOSITION 2.7. *Let* $f\in L^2(\Omega)$ *be fixed. Then for any* $0<h\ll 1$, *we have*

$$\omega(X[f],h)_{2,\Omega_{1,2}} \le 4\sqrt{2}\,\omega(f,h)_2, \tag{2.27}$$

where $\Omega_{1,2} = (-1,2)\times(-1,2)$.

Proof. Let $f\in L^2(\Omega)$ be given, and $\eta\in\mathbb{R}^2$ be fixed with $|\eta|\le h$.

$$\begin{aligned}
\|\tau_\eta(X[f]) - X[f]\|^2_{L^2(\Omega_{1,2}\cap\tau_{-\eta}\Omega_{1,2})} &= \int_{\Omega_{1,2}\cap\tau_{-\eta}\Omega_{1,2}} |X[f](x+\eta)-X[f](x)|^2dx \\
&\le \sum_{\substack{-1\le m,n\le 2\\ -1\le i,j\le 2}} \sum_{\substack{|m-i|=1\\ |n-j|=1}} \int_{\tau_{m,n}\Omega\cap\tau_{(i,j)-\eta}\Omega} |X[f](x+\eta)-X[f](x)|^2dx \\
&\le 2\sum_{-1\le i,j\le 2}\int_{\tau_{i,j}(\Omega\cap\tau_{-\eta}\Omega)} |X[f](x+\eta)-X[f](x)|^2dx \\
&\le 32\int_{\Omega\cap\tau_{-\eta}\Omega}|f(x+\eta)-f(x)|^2dx = 32\|\tau_\eta f - f\|_2^2.
\end{aligned}$$

Thus, we have $\omega(X[f],h)_{2,\Omega_{1,2}} \le 4\sqrt{2}\,\omega(f,h)_2$. □

LEMMA 2.8. *For any* $f\in L^2(\Omega)$ *and* $0<h\ll 1$, *there holds*

$$\|f - C_hQ_hf\|_2 \le K_1\,\omega(f,h)_2 \tag{2.28}$$

and

$$\|P_hQ_hf - C_hQ_hf\|_2 \le K_2\,\omega(f,h)_2, \tag{2.29}$$

where K_1 *and* K_2 *are positive constants independent of h.*

Proof. By definition of the operators Q_h (see (2.17)) and C_h (see (2.19)), we have

$$\begin{aligned}
\|f - C_hQ_hf\|_2^2 &= \sum_{1\le i,j\le N}\int_{\Omega_{i,j}} \left| f(x) - \frac{1}{|\Omega_{i,j}|}\int_{\Omega_{i,j}} f(y)dy\right|^2 dx \\
&\le \sum_{1\le i,j\le N}\int_{\Omega_{i,j}} \left(\frac{1}{|\Omega_{i,j}|}\int_{\Omega_{i,j}}|f(x)-f(y)|\,dy\right)^2 dx \\
&\le \sum_{1\le i,j\le N}\int_{\Omega_{i,j}} \left(\frac{4}{h^2}\int_{\{z:\,|z|\le\sqrt{2}h\}}|X[f](x)-X[f](x+z)|\,dz\right)^2 dx \\
&= \int_\Omega \left(\frac{4}{h^2}\int_{\{z:\,|z|\le\sqrt{2}h\}}|X[f](x)-X[f](x+z)|\,dz\right)^2 dx \\
&\le \frac{4}{h^2}\int_{\{z:\,|z|\le\sqrt{2}h\}}\int_\Omega |X[f](x)-X[f](x+z)|^2\,dxdz,
\end{aligned}$$

where we have used Cauchy-Schwarz inequality, and Fubini Theorem to swap the order of integration.

Now, we observe that for $h \ll 1$, for any $x \in \Omega$ and any $z \in \mathbb{R}^2$ such that $|z| \leq \sqrt{2}h$, we have $\{x, x+z\} \subset \Omega_{1,2}$; so that

$$\int_{\Omega} |X[f](x) - X[f](x+z)|^2\, dx \leq (\omega(X[f], \sqrt{2}h)_{2,\Omega_{1,2}})^2.$$

Therefore,

$$\begin{aligned}
\|f - C_h Q_h f\|_2^2 &\leq \frac{4}{h^2} \int_{\{z:\ |z| \leq \sqrt{2}h\}} \int_{\Omega} |X[f](x) - X[f](x+z)|^2\, dxdz \\
&\leq (\omega(X[f], h)_{2,\Omega_{1,2}})^2 \frac{4}{h^2} \int_{\{z:\ |z| \leq \sqrt{2}h\}} dz \\
&\leq 8\pi\, (\omega(X[f], \sqrt{2}h)_{2,\Omega_{1,2}})^2 \\
&\leq 32\pi\, (\omega(X[f], h)_{2,\Omega_{1,2}})^2 \text{ since } \omega(X[f], \sqrt{2}h)_{2,\Omega_{1,2}} \leq 2\,\omega(X[f], h)_{2,\Omega_{1,2}} \\
&\leq \pi\, (32\,\omega(f, h)_2)^2 \text{ by (2.27);}
\end{aligned}$$

hence the inequality (2.28) holds with $K_1 = 32\sqrt{\pi}$.

We now prove the inequality (2.29). By definition of the operators P_h, Q_h, and C_h, we have

$$\begin{aligned}
\|P_h Q_h f - C_h Q_h f\|_2^2 &= \sum_{1 \leq i,j \leq N} \int_{\Omega_{i,j}} |P_h Q_h f(x) - (Q_h f)_{i,j}|^2 dx \\
&\leq 2 \sum_{1 \leq i,j \leq N} \int_{\Omega_{i,j}} \sum_{-1 \leq k,l \leq 1} \left| \left((Q_h f)_{i+l,j+k} - (Q_h f)_{i,j}\right) \phi_{i+l,j+k}(x) \right|^2 dx \\
&\leq 2 \sum_{-1 \leq l,k \leq 1} \sum_{\substack{1 \leq i+l \leq N \\ 1 \leq j+k \leq N}} h^2 \left| (Q_h f)_{i+l,j+k} - (Q_h f)_{i,j} \right|^2 \\
&\leq 2 \sum_{-1 \leq l,k \leq 1} \sum_{\substack{1 \leq i+l \leq N \\ 1 \leq j+k \leq N}} \int_{\Omega_{i,j}} |f(x + (lh, kh)) - f(x)|^2 dx \\
&\leq 18 (\omega(f, \sqrt{2}h)_{2,\Omega_{1,2}})^2.
\end{aligned}$$

Thus,

$$\begin{aligned}
\|P_h Q_h f - C_h Q_h f\|_2 &\leq 3\sqrt{2}\omega(f, \sqrt{2}h)_{2,\Omega_{1,2}} \\
&\leq 6\sqrt{2}\omega(f, h)_{2,\Omega_{1,2}} \text{ since } \omega(f, \sqrt{2}h)_{2,\Omega_{1,2}} \leq 2\omega(f, h)_{2,\Omega_{1,2}} \\
&\leq 48\,\omega(f, h)_2 \text{ by (2.27).}
\end{aligned}$$

Hence (2.29) holds with $K_2 = 48$, and the proof is complete. $\square$

3. Piecewise linear approximation of the continuous ROF model. In this section, we construct continuous piecewise linear functions and prove their convergence to the minimizer of the ROF model. Let $f \in L^2(\Omega)$ be fixed and $Q_h f$ the discretization of f with respect to the quadrangulation $\square_h$. Let $z^{f,h}$ be the minimizer of the functional

$$E_h^f(u) = J_h(u) + \frac{h^2}{2\lambda} \sum_{1 \leq i,j \leq N} |u_{i,j} - (Q_h f)_{i,j}|^2, \tag{3.1}$$

over $\mathbb{R}^{N\times N}$ with $J_h(u)$ defined in (2.15). We denote the minimizer of the ROF model in the continuous setting by

$$u^f = \operatorname*{argmin}_{u\in BV(\Omega)} E^f_\lambda(u), \tag{3.2}$$

where $E^f_\lambda(u)$ is defined in (1.3). We present two important properties of the ROF model that are at the foundation of the convergence analysis carried in this section.

THEOREM 3.1. *Let $u^f \in BV(\Omega)$ be the minimizer of the ROF functional $E^f_\lambda(u)$. Then, for any $v \in BV(\Omega)$, there holds*

$$\left\|v-u^f\right\|_2^2 \le 2\lambda\left(E^f_\lambda(v)-E^f_\lambda(u^f)\right). \tag{3.3}$$

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

$$\|u^f-u^g\|_2 \le \|f-g\|_2. \tag{3.4}$$

Proof. The proof is straightforward and relies on the fact that $E^f_\lambda(u)$ is subdifferentiable with respect to the topology of $L^2(\Omega)$. □

The inequalities (3.3) and (3.4) were exploited by Wang and Lucier [20] to study the error bound of a piecewise constant approximation of the continuous ROF model.

THEOREM 3.2 (Maximum principle). *Suppose that $f\in L^\infty(\Omega)$ and let u^f be the minimizer of $E^f_\lambda(u)$ on $BV(\Omega)$. Then, $u \in L^\infty(\Omega)$ and*

$$\|u^f\|_\infty \le \|f\|_\infty. \tag{3.5}$$

More precisely, we have

$$\inf_{x\in\Omega} f(x) \le u(x) \le \sup_{x\in\Omega} f(x) \quad \textit{for a.e } x\in\Omega. \tag{3.6}$$

Proof. Let $u\in BV(\Omega)$ be fixed. Let $M=\|f\|_\infty$ and set

$$u^M(x) = \begin{cases} u(x), & |u(x)|\le M \\ \operatorname{sign}(u(x))M, & |u(x)|>M. \end{cases}$$

The sub-level set of the function u^M are

$$U^M_t = \begin{cases} \Omega & t>M, \\ U_t & |t|\le M, \\ \emptyset, & t<-M, \end{cases}$$

where U_t is the sublevel set of u at level t. On the one hand, since $|D\mathbb{1}_\Omega|(\Omega)=0$ and $|D\mathbb{1}_\emptyset|(\Omega)=0$, it follows from the coarea formula that

$$|Du^M|(\Omega) = \int_{-M}^{M} |D\mathbb{1}_{U_t}|(\Omega)dt \le |Du|(\Omega).$$

On the other hand, it is easy to check that $|u(x)-f(x)|\ge|f(x)-u^M(x)|$ for a.e. $x\in\Omega$. Thus, we have $E^f_\lambda(u)\ge E^f_\lambda(u^M)$ and it follows that if u^f is the minimizer of E^f_λ, then by uniqueness of the minimizer it must be the case that $u^f=(u^f)^M$. Thus,

$$|u^f(x)|\le M, \quad \text{for a.e } x\in\Omega,$$

and (3.5) holds. A similar truncation technique shows that (3.6) holds as well. □

We now construct a continuous piecewise linear approximation of u^f and show that it converges to u^f for a special class of functions f. Let $P_h z^{f,h}$ be the continuous piecewise linear interpolation of the discrete minimizer $z^{f,h}$ over the triangulation Δ_h. By the estimate (3.3), we have

$$\left\|P_h z^{f,h} - u^f\right\|_2^2 \leq 2\lambda \left(E_\lambda^f(P_h z^{f,h}) - E_\lambda^f(u^f)\right).$$

Therefore, it suffices to show that $|E_\lambda^f(P_h z^{f,h}) - E_\lambda^f(u^f)| \to 0$ as $h \to 0$ to get that the continuous piecewise linear functions $P_h z^{f,h}$ approximate the solution of the ROF model. To this aim, we shall compare both $E_\lambda^f(P_h z^{f,h})$ and $E_\lambda^f(u^f)$ to the discrete energy $E_h^f(z^{f,h})$.

LEMMA 3.3. *Let $z^{f,h}$ be the minimizer of the functional $E_h^f(u)$ with respect to $\mathbb{R}^{N\times N}$. Then*

$$E_\lambda^f(P_h z^{f,h}) \leq E_h^f(z^{f,h}) + \frac{1}{2\lambda} C\omega(f,h)_2 \left(C\omega(f,h)_2 + 8\|f\|_2\right) \tag{3.7}$$

where C are positive constant depending only on f.

Proof. Since $J_h(z^{f,h}) = |DP_h z^{f,h}|(\Omega)$ which is the reason we so defined discrete total variation in (2.15), we have

$$\begin{aligned}
2\lambda(E_\lambda^f(P_h z^{f,h}) - E_h^f(z^{f,h})) &= \|P_h z^{f,h} - f\|_2^2 - \sum_{1\leq i,j\leq N} h^2 |z_{i,j}^{f,h} - Q_h f_{i,j}|^2 \\
&\leq \|P_h Q_h f - f\|_2 (\|P_h Q_h f - f\|_2 + 2\|P_h(z^{f,h} - Q_h f)\|_2) + \\
&\quad + \|P_h(z^{f,h} - Q_h f)\|_2^2 - \sum_{1\leq i,j\leq N} h^2 |z_{i,j}^{f,h} - Q_h f_{i,j}|^2 \\
&\leq \|P_h Q_h f - f\|_2 (\|P_h Q_h f - f\|_2 + 2\|P_h(z^{f,h} - Q_h f)\|_2),
\end{aligned} \tag{3.8}$$

where the last inequality above follows from the proof of Lemma 2.4.

To finish the proof, it suffices to show that

$$\|P_h Q_h f - f\|_2 \leq C\omega(f,h)_2 \text{ and } \|P_h(z^{f,h} - Q_h f)\|_2 \leq 4\|f\|_2. \tag{3.9}$$

First, from the proof of Lemma 2.4 it is easy to show that

$$\begin{aligned}
\|P_h(z^{f,h} - Q_h f)\|_2^2 &\leq \sum_{1\leq i,j\leq N} h^2 |z_{i,j}^{f,h} - (Q_h f)_{i,j}|^2 \\
&\leq 2\lambda E_h^f(0) \\
&= \sum_{1\leq i,j\leq N} h^2 |(Q_h f)_{i,j}|^2 \leq 4\|f\|_{L^2}^2.
\end{aligned}$$

Next, by Lemma 2.8, we have

$$\begin{aligned}
\|P_h Q_h f - f\|_2 &\leq \|P_h Q_h f - C_h Q_h f\|_2 + \|C_h Q_h f - f\|_2 \\
&\leq (K_1 + K_2)\,\omega(f,h)_2 && \text{by (2.28) and (2.29)}
\end{aligned}$$

Therefore, the inequalities (3.9) holds with $C = K_1 + K_2$ and the proof is complete. □

LEMMA 3.4. *Let $z^{f,h}$ be the solution of* (3.1)*. For $0 < \varepsilon \ll 1$, set $u_\varepsilon^f = X[u^f] * \rho_\varepsilon$. If $f \in L^\infty(\Omega)$, then*

$$E_h^f(z^{f,h}) \leq E_\lambda^f(u_\varepsilon^f) + 16\|f\|_\infty^2 h + \mathscr{O}(h/\varepsilon^2). \tag{3.10}$$

Proof. With a slight abuse of notation, we also let u_ε^f be the element of $\mathbb{R}^{N\times N}$ obtained by evaluating u_ε^f at the grid points $\omega_{i,j}$. Since $z^{f,h}$ is the minimizer of $E_h^f(u)$, we have

$$\begin{aligned} E_h^f(z^{f,h}) &\le E_h^f(u_\varepsilon^f) = J_h(P_h u_\varepsilon^f) + \frac{1}{2\lambda}\sum_{i,j=1}^N h^2|u_\varepsilon^f(\omega_{i,j}) - (Q_h f)_{i,j}|^2 \\ &\le \int_\Omega |\nabla(P_h u_\varepsilon^f)|dx + \frac{1}{2\lambda}\sum_{i,j=1}^N h^2|u_\varepsilon^f(\omega_{i,j}) - (Q_h f)_{i,j}|^2 \\ &\le \int_\Omega |\nabla u_\varepsilon^f|dx + \int_\Omega |\nabla(P_h u_\varepsilon^f - u_\varepsilon^f)|dx + \frac{1}{2\lambda}\sum_{i,j=1}^N h^2|u_\varepsilon^f(\omega_{i,j}) - (Q_h f)_{i,j}|^2. \end{aligned} \tag{3.11}$$

Next, for each $1 \le i,j \le N$, we have

$$\begin{aligned} |u_\varepsilon^f(\omega_{i,j}) - (Q_h f)_{i,j}|^2 = |u_\varepsilon^f(\omega_{i,j}) - (Q_h u_\varepsilon^f)_{i,j}|^2 + |(Q_h u_\varepsilon^f - Q_h f)_{i,j}|^2 + \\ + 2|u_\varepsilon^f(\omega_{i,j}) - (Q_h u_\varepsilon^f)_{i,j}| \cdot |(Q_h u_\varepsilon^f - Q_h f)_{i,j}| \end{aligned} \tag{3.12}$$

and by the mean value theorem

$$\begin{aligned} |u_\varepsilon^f(\omega_{i,j}) - (Q_h u_\varepsilon^f)_{i,j}|^2 &\le \frac{1}{|\Omega_{i,j}|}\int_{\Omega_{i,j}} |u_\varepsilon^f(\omega_{i,j}) - u_\varepsilon^f(x)|^2 dx \\ &\le \frac{1}{|\Omega_{i,j}|}\sup_{x\in\bar{\Omega}_{i,j}} |\nabla u_\varepsilon^f(x)|^2 \int_{\Omega_{i,j}} |x-\omega_{i,j}|^2 dx \\ &\le \frac{C}{\varepsilon^2}|\Omega_{i,j}| \text{ where we used } |x-\omega_{i,j}|^2/|\Omega_{ij}| \le 8 \end{aligned} \tag{3.13}$$

where C is a positive constant depending only on u through its L^1- norm. Thus

$$\sum_{i,j=1}^N h^2|u_\varepsilon^f(\omega_{i,j}) - (Q_h f)_{i,j}|^2 \le C|\Omega|\frac{h^2}{\varepsilon^2} + \sum_{i,j=1}^N h^2|(Q_h u_\varepsilon^f - Q_h f)_{i,j}|^2 + C'\frac{h}{\varepsilon}, \tag{3.14}$$

where C, C' are positive constants depending on f, u^f, and Ω. Now, we establish an upper bound for the second term on the right in the inequality (3.14). By definition of the operator Q_h, the Cauchy-Schwarz inequality, and Theorem 3.2, we have

$$\sum_{i,j=1}^N h^2|Q_h(u_\varepsilon^f - f)_{i,j}|^2 \le \|u_\varepsilon^f - f\|^2_{L^2(\Omega)} + 16\|f\|^2_\infty h. \tag{3.15}$$

Taking into account (3.15) and (3.14) in the inequality (3.11), we obtain

$$E_h^f(z^{f,h}) \le E_\lambda^f(u_\varepsilon^f) + 16\|f\|^2_\infty h + C|\Omega|\frac{h^2}{\varepsilon^2} + C'\frac{h}{\varepsilon} + \|P_h u_\varepsilon^f - u_\varepsilon^f\|_{W^{1,1}(\Omega)}. \tag{3.16}$$

Since the rectangular domain is endowed with a type I triangulation, we have (see [5, Theorem 4.4.20, p. 121])

$$\|P_h u_\varepsilon^f - u_\varepsilon^f\|_{W^{1,1}(\Omega)} \le Ch\sum_{|\alpha|=2}\|D^\alpha u_\varepsilon^f\|_{L^1(\Omega)} \le C''\frac{h}{\varepsilon^2}, \tag{3.17}$$

where C'' is a constant that depends on $\|u\|_{L^1(\Omega)}$. Thus, the estimate (3.16) becomes

$$E_h^f(z^{f,h}) \leq E_\lambda^f(u_\varepsilon^f) + 16\|f\|_\infty^2 h + C\frac{h}{\varepsilon^2},$$

where we have used the fact that $x^2 < x$ for any $0 < x < 1$. □

We now prove the main result of this paper.

THEOREM 3.5. *Suppose that $f \in \mathrm{Lip}\left(\alpha, L^2(\Omega)\right) \cap L^\infty(\Omega)$ for some $\alpha \in (0,1]$. Let $z^{f,h}$ be the minimizer of the functional $E_h^f(u)$ in $\mathbb{R}^{N\times N}$ and u^f be defined by (3.2). Then $P_h z^{f,h}$ converges in $L^2(\Omega)$ to u^f as $h \to 0$.*

Proof. For any $0 < h \ll 1$ and any $\varepsilon > 0$, we have

$$\begin{aligned}\|P_h z^{f,h} - u^f\|_{L^2(\Omega)}^2 &\leq 2\lambda\left[E_\lambda^f(P_h z^{f,h}) - E_\lambda^f(u^f)\right] && \text{by (3.3)}\\ &\leq 2\lambda\left[E_\lambda^f(P_h z^{f,h}) - E_h^f(z^{f,h}) + E_h^f(z^{f,h}) - E_\lambda^f(u^f)\right].\end{aligned}$$

Next, by equation (3.7) in Lemma 3.3, we have

$$E_\lambda^f(P_h z^{f,h}) - E_h^f(z^{f,h}) \leq \frac{1}{2\lambda}\omega(f,h)_2\left(\omega(f,h)_2 + C\|f\|_2\right)$$

and equation (3.10) in Lemma 3.4 yields

$$E_h^f(z^{f,h}) - E_\lambda^f(u^f) \leq E_\lambda^f(u_\varepsilon^f) - E_\lambda^f(u^f) + 16\|f\|_\infty^2 h + C\frac{h}{\varepsilon^2}.$$

Thus,

$$\begin{aligned}\|P_h z^{f,h} - u^f\|_{L^2(\Omega)}^2 \leq{}& \omega(f,h)_2\left(\omega(f,h)_2 + C\|f\|_2\right) + \\ &+ 32\lambda\|f\|_\infty^2 h + 2C\lambda\frac{h}{\varepsilon^2} + 2\lambda\left(E_\lambda^f(u_\varepsilon^f) - E_\lambda^f(u^f)\right).\end{aligned} \tag{3.18}$$

Since $f \in \mathrm{Lip}\left(\alpha, L^2(\Omega)\right)$ we have $\omega(f,h)_2 \leq \mathscr{O}(h^\alpha)$. Letting $\varepsilon = h^{1/2(\alpha+1)}$, we infer from inequality (3.18) that

$$\|P_h z^{f,h} - u^f\|_{L^2(\Omega)}^2 \leq C h^{\alpha/(\alpha+1)} + 2\lambda\left(E_\lambda^f(u_\varepsilon^f) - E_\lambda^f(u^f)\right), \tag{3.19}$$

where we have used the fact that the function $x \mapsto a^x$ is decreasing when $0 < a < 1$.

Since $u_\varepsilon^f \to u^f$ in $L^2(\Omega)$ as $\varepsilon \to 0$, it follows from Theorem 2.5 (b) that for our choice of $\varepsilon = h^{1/2(\alpha+1)}$, $E_\lambda^f(u_\varepsilon^f) - E_\lambda^f(u^f) \to 0$ as $h \to 0$. Thus, taking the limit as $h \to 0$ in (3.19), we conclude that $\|P_h z^{f,h} - u^f\|_{L^2(\Omega)} \to 0$ as $h \to 0$ and the proof is complete. □

COROLLARY 3.6. *Under the assumptions of Theorem 3.5, we have*

$$J_h(P_h z^{f,h}) \to J(u^f), \quad \text{when } h \to 0.$$

Proof. This is a direct consequence of the convergence of $E_h^f(P_h z^{f,h})$ to $E_\lambda^f(u^f)$ as $h \to 0$. □

REMARK 3.7. *It transpires from the proof above that to establish a convergence rate of the proposed piecewise linear approximation, one will need a convergence rate of $E_\lambda^f(u_\varepsilon^f)$ to $E_\lambda^f(u^f)$ which we have not been able to establish at this point. Moreover the optimal convergence rate, if one could be derived, should be of the order of $\mathscr{O}(h^\beta)$ with $0 < \beta \leq 1/2$. The convergence is slower for smaller values of β and one would need very small values of h to get significant evidence of the convergence when doing numerical simulations.*

4. Numerical experiments. In this section we formulate two algorithms for computing the discrete solution $z^{f,h}$ and show numerical evidence to support the convergence result established in Theorem 3.5.

4.1. The algorithms. The objective functional of which $z^{f,h}$ is a special case of proximal operator [12]; thus the proximal forward-backward splitting algorithm [12] may be adapted to this problem. One could also adapt the split Bregman method [15] to obtain a fast algorithm for computing $z^{f,h}$. However, in this paper we use the duality method to derive our algorithms.

Let $X := \mathbb{R}^{N\times N}$ and $Y = X\times X$. To study iterative algorithms for computing the discrete minimizer $z^{f,h}$, we introduce the discrete divergence operators $\mathrm{div}_+ : Y \to X$ and $\mathrm{div}_- : Y \to X$ associated with the discrete gradients ∇_+ and ∇_-, respectively, and defined by

$$\begin{aligned} \mathrm{div}_+(p)_{i,j} = &\begin{cases} 0 & \text{if } i=N \text{ or } j=N \\ \dfrac{p^1_{i,j}}{h} & \text{otherwise} \end{cases} - \begin{cases} 0 & \text{if } i=1 \text{ or } j=N \\ \dfrac{p^1_{i-1,j}}{h} & \text{otherwise} \end{cases} \\ &+ \begin{cases} 0 & \text{if } i=N \text{ or } j=N \\ \dfrac{p^2_{i,j}}{h} & \text{otherwise} \end{cases} - \begin{cases} 0 & \text{if } i=N \text{ or } j=1 \\ \dfrac{p^2_{i,j-1}}{h} & \text{otherwise.} \end{cases} \end{aligned} \tag{4.1}$$

and

$$\begin{aligned} \mathrm{div}_-(p)_{i,j} = &\begin{cases} 0 & \text{if } i=N \text{ or } j=1 \\ \dfrac{p^1_{i+1,j}}{h} & \text{otherwise} \end{cases} - \begin{cases} 0 & \text{if } i=1 \text{ or } j=1 \\ \dfrac{p^1_{i,j}}{h} & \text{otherwise} \end{cases} \\ &+ \begin{cases} 0 & \text{if } i=1 \text{ or } j=N \\ \dfrac{p^2_{i,j+1}}{h} & \text{otherwise} \end{cases} - \begin{cases} 0 & \text{if } i=1 \text{ or } j=1 \\ \dfrac{p^2_{i,j}}{h} & \text{otherwise} \end{cases} \end{aligned} \tag{4.2}$$

It can be shown that the discrete divergence operators div_+ and div_- are the negative of the adjoint operators of the discrete gradient ∇_+ and ∇_-, respectively.

By standard duality arguments similar to the one used in [6, 7], we show that the minimizer $z^{f,h}$ is given by

$$z^{f,h} = Q_h f + \frac{\lambda}{2}\left(\mathrm{div}_+(\bar{p}) + \mathrm{div}_-(\bar{q})\right), \tag{4.3}$$

where

$$(\bar{p},\bar{q}) \in \underset{p,q\in B_Y}{\mathrm{argmin}} \left|\lambda\left(\mathrm{div}_+(p) + \mathrm{div}_-(\bar{q})\right) + 2Q_h f\right|^2, \tag{4.4}$$

with $|p|_\infty = \max(|p_{i,j}| : 1 \le i,j \le N)$ and $B_Y = \{p \in Y : |p|_\infty \le 1\}$.

By observing that the point $(\bar{p},\bar{q})$ defined above is characterized by

$$\forall\, \tau > 0, \quad \begin{cases} \bar{p} = P_{B_Y}\left(\bar{p} + \tau\nabla_+\left[\mathrm{div}_+(\bar{p}) + \mathrm{div}_-(\bar{q}) + 2f/\lambda\right]\right), \\ \bar{q} = P_{B_Y}\left(\bar{q} + \tau\nabla_-\left[\mathrm{div}_+(\bar{p}) + \mathrm{div}_-(\bar{q}) + 2f/\lambda\right]\right), \end{cases} \tag{4.5}$$

where

$$P_{B_Y}(p)_{i,j} = \left(\frac{p^1_{i,j}}{\max(1,|p_{i,j}|)}, \frac{p^2_{i,j}}{\max(1,|p_{i,j}|)}\right), \quad 1 \le i,j \le N$$

is the orthogonal projection of p onto B_Y, we get the following algorithm for computing the solution $z^{f,h}$.

ALGORITHM 4.1 (Dual Projected-Gradient). *Let $\tau > 0$ be fixed. For $n = 0$, let $p_0 = q_0 = 0$.*

Step 1: *Compute u_n*

$$u_n = Q_h f + \frac{\lambda}{2}\left[\mathrm{div}_+(p_n) + \mathrm{div}_-(q_n)\right]. \tag{4.6}$$

Step 2: *Compute p_{n+1} and q_{n+1}*

$$\begin{aligned} p_{n+1} &= P_{B_Y}\left(p_n + \frac{2\tau}{\lambda}\nabla_+(u_n)\right), \\ q_{n+1} &= P_{B_Y}\left(q_n + \frac{2\tau}{\lambda}\nabla_-(u_n)\right). \end{aligned} \tag{4.7}$$

Step 3: *Until the stopping criterion is met, let $n \leftarrow n+1$ and return to step 1.*

The algorithm 4.1 is a special case of Bermúdez-Moreno Algorithm [4] and its convergence can be obtained as in [3]. Specifically we have the following theorem.

THEOREM 4.2. *If $0 < \tau < h^2/8$, then Algorithm 4.1 converges. More precisely, given $p_0, q_0 \in B_Y$, there exists a point $(\bar{p}_0, \bar{q}_0)$ satisfying* (4.4) *such that the sequence (p_n, q_n) defined by* (4.7) *converges to $(\bar{p}_0, \bar{q}_0)$ and the sequence u_n defined by* (4.6) *converges to $z^{f,h}$.*

Proof. A direct proof is obtained by modifying and completing the argument in [13] and may be found in [18] . □

An alternating version of Algorithm 4.1 is obtained by using p_{n+1} to compute q_{n+1} thus resulting in the following algorithm

ALGORITHM 4.3 (Alternating Dual Projected-Gradient). *Let $\tau > 0$ be fixed and choose $p_0, q_0 \in B_Y$.*

Step 1: *Compute u_n*

$$u_n = Q_h f + \frac{\lambda}{2}\left[\mathrm{div}_+(p_n) + \mathrm{div}_-(q_n)\right]. \tag{4.8}$$

Step 2: *Compute p_{n+1} and q_{n+1}*

$$\begin{cases} p_{n+1} = P_{B_Y}\left(p_n + \tau\nabla_+\left[\mathrm{div}_+(p_n) + \mathrm{div}_-(q_n) + 2f/\lambda\right]\right), \\ q_{n+1} = P_{B_Y}\left(q_n + \tau\nabla_-\left[\mathrm{div}_+(p_{n+1}) + \mathrm{div}_-(q_n) + 2f/\lambda\right]\right). \end{cases} \tag{4.9}$$

Step 3: *Until the stopping criterion is met, let $n \leftarrow n+1$ and return to step 1.*

While the proof of the convergence of Algorithm 4.3 is still eluding us, the numerical experiments suggest that one should be able to prove a convergence result for $0 < \tau < 1/4$.

4.2. Numerical tests. We report the results of two numerical experiments with Algorithm 4.1. First, we compare the performance of the algorithms proposed above to Chambolle's fixed-point algorithm [7], and the projected-gradient algorithm proposed in [8]. The noised images are obtained by adding a realization of a zero mean Gaussian random variable to the images in Figure 4.1. It should be noted that in our tests, we did not attempt to choose the parameters τ and λ for optimal performance of the algorithms.

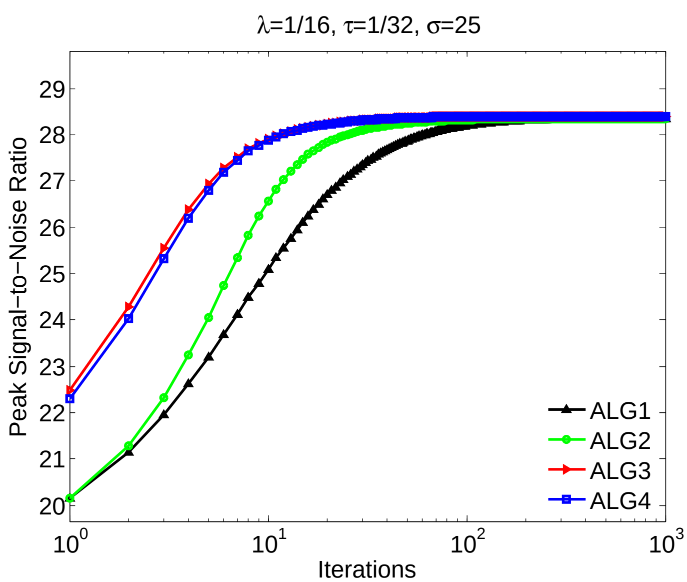
We shall use the following abbreviations to identify the algorithms under consideration here.

ALG1: The dual fixed point iterative algorithm described in [7].
ALG2: The dual projected-gradient algorithm described in [8, 13].

Fig. 4.1: The images used in the numerical experiments below.

(a) Lena 256×256. (b) Peppers 256×256. (c) Boats 512×512.

Fig. 4.2: Convergence of the four algorithms for the image in Figure 4.1a with $\sigma = 25$, $\tau = 1/32$, and $\lambda = 1/16$. The PSNR is computed relative to the ground truth in Figure 4.1a. The algorithm are terminated after 1000 iterations or when the mean square error (MSE) to the ground truth is less than 10^{-8} or the absolute change in MSE at consecutive iterations is less than 10^{-13}.



ALG3: The dual projected-gradient algorithm presented in Algorithm 4.1.

ALG4: The alternating dual projected-gradient algorithm presented in Algorithm 4.3.

The algorithms that we proposed have the best convergence at the onset (see Figure 4.2); thus would be appropriate for use in situations where one needs to clean an image as a preprocessing step of an image analysis task.

Table 4.1 and Table 4.2 below show the capability of Algorithm 4.1 to remove noise for various noise levels. The inputs for all four algorithms are obtained by adding a zero mean Gaussian noise with standard deviation σ to the images in Figure 4.1. Our experiments show that the two projected gradient algorithms that we developed are effective and performs equally with the algorithms ALG1 and ALG2. Moreover, our algorithms get to a viable solution within the first ten iterations, making them favorable tools when denoising is required merely as a preprocessing step in the image analysis.

Table 4.1: Comparison of the Algorithms using the image in Figure 4.1b. We report the results in the format $a(b,c)$, where a is the PSNR, b and c are the number of iterations and the CPU time used in reaching the PSNR value, respectively. The algorithm are terminated after 1000 iterations or when the MSE is less than 10^{-8} or the absolute change in MSE at consecutive iterations is less than 10^{-13}.

τ	λ	σ	ALG1	ALG2	ALG3	ALG4
		15.00	$31.79(10^3,10s)$	$31.78(10^3,15s)$	$31.81(10^3,31s)$	$31.87(43,2s)$
$\frac{1}{48}$	$\frac{1}{24}$	20.00	$29.37(10^3,10s)$	$29.37(10^3,15s)$	$29.40(10^3,31s)$	$29.41(86,3s)$
		25.00	$26.72(10^3,10s)$	$26.72(10^3,15s)$	$26.75(10^3,31s)$	$26.75(110,4s)$
		30.00	$24.21(10^3,10s)$	$24.21(10^3,15s)$	$24.23(205,7s)$	$24.23(115,4s)$
		15.00	$31.31(10^3,10s)$	$31.28(10^3,15s)$	$31.29(10^3,31s)$	$31.30(348,12s)$
		20.00	$30.34(10^3,10s)$	$30.32(10^3,15s)$	$30.35(10^3,31s)$	$30.36(307,11s)$
$\frac{1}{32}$	$\frac{1}{16}$	25.00	$28.81(10^3,10s)$	$28.80(10^3,15s)$	$28.85(10^3,31s)$	$28.86(76,3s)$
		30.00	$26.73(10^3,10s)$	$26.72(10^3,15s)$	$26.77(10^3,5s)$	$26.77(116,4s)$

Table 4.2: Comparison of the Algorithms using the image in Figure 4.1c. We report the results in the form $a(b,c)$, where a is the PSNR, b and c are the number of iterations and the CPU time for reaching the PSNR value, respectively. The algorithms are terminated after 1000 iterations or when the MSE is less than 10^{-8} or the absolute change in MSE at consecutive iterations is less than 10^{-13}.

τ	λ	σ	ALG1	ALG2	ALG3	ALG4
		15.00	$30.49(10^3,60s)$	$30.48(10^3,60s)$	$30.52(218,32s)$	$30.51(115,22s)$
$\frac{1}{48}$	$\frac{1}{24}$	20.00	$28.72(10^3,60s)$	$28.72(10^3,72s)$	$28.77(379,52s)$	$28.77(156,30s)$
		25.00	$26.37(10^3,60s)$	$26.37(566,41s)$	$26.42(319,44s)$	$26.42(144,28s)$
		30.00	$24.11(10^3,60s)$	$24.11(10^3,73s)$	$24.14(291,40s)$	$24.14(133,26s)$
		15.00	$29.58(10^3,60s)$	$29.56(10^3,72s)$	$29.56(10^3,136s)$	$29.57(240,45s)$
		20.00	$29.12(10^3,60s)$	$29.11(10^3,72s)$	$29.13(157,21s)$	$29.13(108,21s)$
$\frac{1}{32}$	$\frac{1}{16}$	25.00	$28.04(10^3,60s)$	$28.03(10^3,72s)$	$28.09(276,38s)$	$28.09(166,32s)$
		30.00	$26.38(10^3,60s)$	$26.38(960,71s)$	$26.44(371,53s)$	$26.44(171,34s)$

We then used the iterative algorithm 4.1 to obtain numerical evidence confirming the theoretical result in Theorem 3.5 with the function

$$f = 255\mathbb{1}_C,$$

where C is the disk centered at $(1/2, 1/2)$ with radius $R = 1/4$. We draw the attention of the reader to the fact [9] that in this case the minimizer u^f is given by

$$u^f = 255\max(1 - 2\lambda/r, 0)\mathbb{1}_C, \quad \forall\lambda > 0.$$

Note that the size of the discrete data grows as $1/h^2$ as $h \to 0$, therefore we will only show the result of moderate size data. See Table 4.3 demonstrating the convergence of the piecewise linear interpolation to the minimizer of $E_\lambda^f(u)$ as $h \to 0$.

Table 4.3: The $L^2(\Omega)$ distance between u^f and $P_h u_{100}$ where u_{100} is the approximation of $z^{f,h}$ computed using Algorithm 4.1. It is already apparent that the distance is decreasing with h even though we are only using an approximation of the discrete minimizer $z^{f,h}$.

	λ/R			
h	2^{-3}	2^{-5}	2^{-7}	2^{-9}
2^{-5}	25.0682	18.4406	17.9938	17.9808
2^{-6}	26.1967	13.8377	11.5935	11.3495
2^{-7}	21.0148	14.1954	9.1324	8.5836
2^{-8}	17.8916	14.1036	7.3424	6.0095
2^{-9}	16.1267	10.2853	7.3082	4.5298
2^{-10}	15.1462	7.6813	7.1739	3.6942

5. Conclusion. In this paper, we were interested in the numerical computations of the minimizer of the Rudin-Osher-Fatemi model for image denoising:

$$\underset{u \in BV(\Omega)}{\operatorname{argmin}} \left\{ \int_\Omega |Du|(\Omega) + \frac{1}{2\lambda} \int_\Omega |u - f|^2 dx \right\}.$$

Although this model was introduced for its practical purpose of digital image enhancement, its mathematical analysis is important on its own right and has generated lots of interesting literature. The theory guarantees the existence of the solution of the ROF model for any $f \in L^2(\Omega)$ and any $\lambda > 0$, however, in general we do not have analytical formulæ of the solutions. It is also known that if $f \in BV(\Omega)$ is continuous, then the minimizer will be continuous as well. The research done in this paper gives us a tool to visualize such solutions in the absence of their analytical formulæ. We constructed piecewise linear interpolations of the minimizers of discrete functionals derived by discretizing the data function f and show that the family of piecewise linear polynomials thus generated converges to the solution u^f. We also described an algorithm for computing the discrete solution and showed its convergence. This is the first attempt to compute numerical approximation of the ROF minimizer using continuous functions.

The researchers in [20] have used similar techniques to approximate the minimizer of the ROF functional using piecewise constant functions and established a convergence rate under weaker assumptions on the data function f. We extended their analysis to study the approximation with continuous piecewise linear functions, and simplified the proof of the comparison Lemma 3.3. Moreover, unlike these authors we studied an iterative algorithm for computing the discrete solution directly from the discrete model used to establish the convergence. However, we did not obtain an error rate as we have not been able to obtain error rates for the convergence of $J(u_\varepsilon^f)$ to $J(u^f)$ and the compatibility of u^f with the translations of the data function f.

Acknowledgements. We would like to thank the two anonymous referees for the thorough review of the manuscript. Their comments and suggestions have greatly improved the readability of this paper.

REFERENCES

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

[2] L. AMBROSIO, N. FUSCO, AND D. PALLARA, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, 2000.

[3] J.-F. AUJOL, *Some first-order algorithms for total variation based image restoration*, J. Math. Imaging Vision, 34 (2009), pp. 307–327.

[4] A. BERMÚDEZ AND C. MORENO, *Duality methods for solving variational inequalities*, Comput. Math. Appl., 7 (1981), pp. 43–58.

[5] S. C. BRENNER AND L. R. SCOTT, *The mathematical theory of finite element methods*, vol. 15 of Texts in Applied Mathematics, Springer, New York, third ed., 2008.

[6] J. L. CARTER, *Dual methods for total variation-based image restoration*, PhD thesis, University of California, Los Angeles, 2001.

[7] A. CHAMBOLLE, *An algorithm for total variation minimization and applications*, J. Math. Imaging Vision, 20 (2004), pp. 89–97. Special issue on mathematics and image analysis.

[8] ———, *Total variation minimization and a class of binary mrf models*, in Energy Minimization Methods in Computer Vision and Pattern Recognition, A. Rangarajan, B. Vemuri, and A. Yuille, eds., vol. 3757 of Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2005, pp. 136–152.

[9] 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, M. Fromasier, ed., vol. 9 of Radon Ser. Comput. Appl. Math., Walter de Gruyter, Berlin, 2010, pp. 263–340.

[10] A. CHAMBOLLE, S. LEVINE, AND B. J. LUCIER, *An upwind finite-difference method for total variation-based image smoothing.*, SIAM J. Imaging Sciences, 4 (2011), pp. 277–299.

[11] A. CHAMBOLLE AND P.-L. LIONS, *Image recovery via total variation minimization and related problems*, Numer. Math., 76 (1997), pp. 167–188.

[12] P. L. COMBETTES AND V. R. WAJS, *Signal recovery by proximal forward-backward splitting*, Multiscale Model. Simul., 4 (2005), pp. 1168–1200 (electronic).

[13] V. DUVAL, J.-F. AUJOL, AND L. A. VESE, *Mathematical modeling of textures: Application to color image decomposition with a projected gradient algorithm*, J Math Imaging Vis, 37 (2010), pp. 232–248.

[14] E. GIUSTI, *Minimal Surfaces and Functions of Bounded Variation*, vol. 80 of Monographs in Mathematics, Birkhäuser, 1984.

[15] T. GOLDSTEIN AND S. OSHER, *The split Bregman method for L1-regularized problems*, SIAM J. Imaging Sci., 2 (2009), pp. 323–343.

[16] 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, (2009), pp. 514–526.

[17] M.-J. LAI AND L. L. SCHUMAKER, *Spline functions on triangulations*, vol. 110 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 2007.

[18] 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.

[19] L. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Physica D: Nonlinear Phenomena, 60 (1992), pp. 259–268.

[20] J. WANG AND B. J. LUCIER, *Error bounds for finite-difference methods for rudin-osher-fatemi image smoothing.*, SIAM J. Numerical Analysis, 49 (2011), pp. 845–868.