Noname manuscript No.
(will be inserted by the editor)

A New Projected Gradient Algorithm for Total Variation Based Image Denoising

Ming-Jun Lai · Leopold Matamba Messi

the date of receipt and acceptance should be inserted later

Abstract This paper is concerned with the numerical approximation of the minimizer of the continuous Rudin-Osher-Fatemi (ROF) model. A new projected gradient algorithm for computing a numerical solution of the ROF model in the discrete setting is proposed. We show that the new algorithm converges. The solution of our discrete algorithm is then used to construct a continuous piecewise linear spline which approximates the solution of the continuous ROF model in the L^2-norm. We show that if the noised image is in the space $\mathrm{Lip}(\alpha, \mathrm{L}^2(\Omega)),\ 0 < \alpha \leq 1$, our piecewise linear approximations converge to the solution of the ROF model. Finally, we demonstrate with numerical experiments that the new algorithm is as good as the traditional projected gradient algorithm and Chambolle's fixed point algorithm.

1 Introduction

Since the seminal work of Rudin, Osher, and Fatemi[13] total variation based model for image restoration have received a great deal of attention. They are now used in image denoising, image deblurring, and image inpainting. In each case, the problem is formulated as a minimization of a functional of the form

$$\operatorname{argmin} |Du|(\Omega), \text{ subject to the constraints } F(u), \tag{1.1}$$

where $|Du|(\Omega)$ is the total-variation of the function u on Ω, and $F(u)$ is a suitable set of constraints satisfied by u, and $\Omega \subset \mathbb{R}^2$ is the domain of the image u. For image de-noising, the problem was formulated as follows: recover the true image u from a contaminated version $f = u + n$, where n is a white noise with mean 0 and standard deviation σ. The corresponding minimization problem is

$$\operatorname{argmin} |Du|(\Omega) \text{ subject to the constraints} \tag{1.2}$$

$$\int_\Omega u(x)dx = \int_\Omega f(x)dx \text{ and } \int_\Omega |u(x) - f(x)|^2 dx \leq \sigma^2. \tag{1.3}$$

It was shown (cf. [6, 13]) that problem (1.2) is equivalent to the following unconstrained minimization

$$\operatorname*{argmin}_{u \in BV(\Omega)} |Du|(\Omega) + \frac{1}{2\lambda} \int_\Omega (u - f)^2 dx \tag{1.4}$$

where $\lambda > 0$ is a Lagrange multiplier, and $BV(\Omega)$ is the Banach space of functions of bounded variation. The existence and uniqueness of the minimizer of the above problem was established in [1] and [6].

Ming-Jun Lai
Department of Mathematics
University of Georgia
Athens, GA 30602, USA
Tel.: +1706-542-2065
Fax: +1706-542-5907
E-mail: mjlai@math.uga.edu

Leopold Matamba Messi
Department of Mathematics
University of Georgia
Athens, GA 30602, USA
Tel.: +1706-542-5961
Fax: +1706-542-5907
E-mail: lmatamba@math.uga.edu

To find numerical approximations of the solution of (1.4), one has to understand how to discretize the total variation term $|Du|(\Omega)$. All the finite difference methods proposed in the literature are based on the following fact (see for example [2,10] for details)

$$|Du|(\Omega) = \int_{\Omega} |\nabla u| dx = \int_{\Omega} \sqrt{\left(\frac{\partial u}{\partial x_1}\right)^2 + \left(\frac{\partial u}{\partial x_2}\right)^2} dx, \quad \forall\, u \in \mathrm{W}^{1,1}(\Omega). \tag{1.5}$$

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.

Several algorithms to find approximations of the solution of problem (1.4) have been proposed and studied in the literature. See for example [3,6,13]. Indeed, a time dependent nonlinear PDE associated with the Euler-Lagrange equation for (1.4) was used to approximate the minimizer of (1.4) and then solved by finite difference methods (cf. [13]) or finite elements methods (cf. [7]).

In the discrete setting, a primal-dual algorithm (a popular approach in optimization) based on a straightforward discretization of the TV term and L^2 term was introduced by Chambolle in [**?**] for image denoising. Recently, Duval et al [8] proposed a projected-gradient algorithm for the discrete L^2-TV regularization to compute an approximation of the minimizer. A numerical algorithm based on completely symmetric discretization of the TV term was studied in [15]. In particular, the researchers in [15] were able to show that their numerical solution approximates the minimizer of problem (1.4) in L^2 norm. Furthermore, the convergence of the discrete solution of the finite difference method for time dependent nonlinear PDE was recently established in [11]. These motivate us to study the convergence of the discrete solutions of Chambolle algorithm and projected gradient algorithm to the solution of problem (1.4). After studying the convergence issue of these two algorithms, we realize that one has to use a new discretization of the total variation term which leads to a new version of the projected-gradient algorithm. With this new discretization, we are able to use the solution in the discrete setting to construct a piecewise linear interpolatory spline which can be shown to approximate the solution of the original problem (1.4).

Our contributions in this paper are the following: (**1**) we first prove the existence and uniqueness of the solution based on the new discrete ROF functional (see Theorem 1); (**2**) We formulate a projected gradient algorithm for the new minimization problem (cf. Algorithm 31) and prove its convergence (cf. Theorem 3); (**3**) We demonstrate numerically that our new version of the projected-gradient algorithm performs slightly better than the original projected-gradient algorithm in [5,8], and is at least as good as Chambolle's fixed point iterative algorithm[4], as measured by the Peak Signal to Noise Ratio (PSNR) (see Section 5); (**4**) We obtain an error bound in $\mathrm{L}^2(\Omega)$ between the piecewise linear interpolatory spline of the discrete solution and the solution of the continuous ROF model (see Theorem 2).

2 Preliminaries and Notations

In this section, we give preliminary results, and introduce the notations that we shall use in the paper.

2.1 Basic notations

Let Ω be an open subset of $\mathbb{R}^2$, we denote the indicator function of the set Ω 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

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

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

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

Next, we recall the definitions of some functional spaces that are used in this work. Let $1 \le p < \infty$ be fixed. $\mathrm{L}^p(\Omega)$ is the standard Banach space of $p-$integrable functions

$$\mathrm{L}^p(\Omega) := \left\{ u : \Omega \to \mathbb{R} : \int_{\Omega} |u(x)|^p dx < \infty \right\}.$$

$\mathrm{L}^p_{\mathrm{loc}}(\Omega)$ is the set of functions u that are locally $p-$integrable, i.e

$$\mathrm{L}^p_{\mathrm{loc}}(\Omega) = \left\{ u : \Omega \to \mathbb{R} : \int_{K} |u(x)|^p dx < \infty, \ \forall\, K \subset \Omega \text{ compact} \right\}.$$

The norm of an element $u \in \mathrm{L}^p(\Omega)$ is given by

$$\|u\|_p = \left(\int_\Omega |u(x)|^p dx\right)^{1/p}.$$

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

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

$$\omega_p(u,h) = \sup_{|\eta|\le h} \|\tau_\eta u - u\|_p, \tag{2.1}$$

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

For $u \in \mathrm{L}^p_{\mathrm{loc}}(\Omega)$ and A a subset of Ω whose closure $\bar{A} \subset \Omega$ is a compact, the $p-$modulus of continuity of u with respect to A, denoted $\omega_p(u,h)_A$, is given by

$$\omega_p(u,h)_A = \omega_p(u\mathbb{1}_A, h). \tag{2.2}$$

Let $0 < \alpha \le 1$. We denote $\mathrm{Lip}(\alpha, \mathrm{L}^2(\Omega))$ the subspace of $\mathrm{L}^2(\Omega)$ defined by

$$\mathrm{Lip}(\alpha, \mathrm{L}^2(\Omega)) = \left\{ u \in \mathrm{L}^2(\Omega) \,:\, \sup_{h>0} h^{-\alpha}\omega_2(u,h) < \infty \right\}.$$

2.2 A discretization of the ROF functional.

In the sequel Ω shall denote the unit square $[0,1]^2$ of $\mathbb{R}^2$. We subdivide Ω into $(N-1)^2$ square sub-domains of side length h to get a uniform quadrangulation $\square_h$. A triangulation Δ_h is then obtained from $\square_h$ by splitting each square into two triangles using the Northwest-Southeast diagonal as shown in Figure 1.

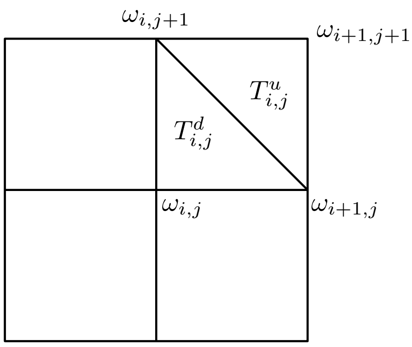


Fig. 1: Type I triangulation of the domain Ω with vertexes $\omega_{i,j}$. $T^u_{i,j}$ is the triangle with vertexes $\langle \omega_{i+1,j}, \omega_{i+1,j+1}, \omega_{i,j+1}\rangle$ and $T^d_{i,j}$ is the triangle with vertexes $\langle \omega_{i,j}, \omega_{i+1,j}, \omega_{i,j+1}\rangle$.

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

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

and define a partition $\{\Omega_{i,j}\}$ of Ω subordinate to $\mathcal{V}_h$ by

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

Let

$$E^f_\lambda(u) := |Du|(\Omega) + \frac{1}{2\lambda}\int_\Omega (u-f)^2 dx. \tag{2.3}$$

Throughout the paper $E_\lambda^f(u)$ will be referred to as the continuous ROF functional. 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, discretizing $E_\lambda^f(u)$ solely based on u is a delicate matter. Many researchers [5,12] have used the closed form of $|Du|(\Omega)$ given in (1.5) to designed discretization scheme of the ROF functional $E_\lambda^f(u)$ that combine a formal finite difference approximation of ∇u with a quadrature approximation of the integral. Our discretization of the ROF functional $E_\lambda^f(u)$ has a similar flavor. We first discretize the total variation as follows

$$J_h(u) := \sum_{1\le i,j<N} \frac{h^2}{2}\sqrt{\left|\frac{u_{i+1,j}-u_{i,j}}{h}\right|^2+\left|\frac{u_{i,j+1}-u_{i,j}}{h}\right|^2} + \sum_{1\le i,j<N} \frac{h^2}{2}\sqrt{\left|\frac{u_{i+1,j+1}-u_{i,j+1}}{h}\right|^2+\left|\frac{u_{i+1,j+1}-u_{i+1,j}}{h}\right|^2}, \quad (2.4)$$

where $u_{i,j}$ is a suitable discretization of the function u at the vertexes of the triangulation Δ_h. For example if u is a continuous piecewise linear spline with respect to Δ_h, on can easily show that $J_h(u) = |Du|(\Omega)$; this is the reason why we use (2.4) to discretize $|Du|(\Omega)$. The resulting discrete ROF-functional $E_{\lambda,h}^f(u)$ is given by

$$E_{\lambda,h}^f(u) := J_h(u) + \frac{1}{2\lambda}\sum_{1\le i,j\le N} h^2\left|u_{i,j}-f_{i,j}\right|^2, \quad u\in\mathbb{R}^{N\times N} \quad (2.5)$$

where $f_{i,j} := f(\omega_{i,j})$ is a suitable discretization on Δ_h of the datum f. From this point onwards, $E_{\lambda,h}^f(u)$ will be referred to as the discrete ROF functional and the associated miniization problem

$$\underset{u\in\mathbb{R}^{N\times N}}{\operatorname{argmin}}\, E_{\lambda,h}^f(u) \quad (2.6)$$

shall be called the discrete ROF model.

2.3 Embedding and Projection Operators

We introduce various operators that will help us construct a piecewise linear approximation of the solution to the continuous L^2-ROF model. In the previous section, we gave a formal discrete approximation of the ROF functional. We now clarify how this discretization is obtained.

The sampling operator, $Q_h : \mathrm{L}^2(\Omega) \to \mathbb{R}^{N\times N}$, is defined by

$$(Q_h f)_{i,j} := \frac{1}{h^2}\int_{\Omega_{i,j}} f(x)dx. \quad (2.7)$$

Q_h will also denote the projection of $\mathrm{L}^2(\Omega)$ onto the linear space of piecewise constant function with respect to the partition $\{\bar\Omega_{i,j} : 1\le i,j\le N\}$ of $\bar\Omega$, in which case Q_h is defined by

$$Q_h f(y) := (Q_h f)_{i,j} = \frac{1}{h^2}\int_{\Omega_{i,j}} f(x)dx, \quad \text{for all } y\in\Omega_{i,j}. \quad (2.8)$$

Next, we denote by $\ell^2(\mathcal{V}_h)$ the vector space $\mathbb{R}^{N\times N}$ endowed with the inner product

$$\langle u,v\rangle_h := \sum_{i,j=1}^N h^2 u_{i,j}v_{i,j}, \quad (2.9)$$

and the corresponding norm $|u|_h := \langle u,u\rangle_h^{1/2}$. The following properties of Q_h follows from Jensen's Inequality.

Lemma 1 *The sampling operator Q_h defined above has the following properties*

$$|Q_h f|_h \le \|f\|_2, \text{ and } \|Q_h f\|_2 \le \|f\|_2, \quad \text{for all } f\in\mathrm{L}^2(\Omega). \quad (2.10)$$

Next, we explain two methods of constructing a function of bounded variation from an element $u \in \ell^2(\mathcal{V}_h)$. First, we create a piecewise constant function defined on Ω via the partition $\{\Omega_{i,j} : 1 \leq i,j \leq N\}$ as follows

$$C_h u(y) = u_{i,j}, \text{ if } y \in \Omega_{i,j}. \tag{2.11}$$

We denote by $C_h \colon \ell^2(\mathcal{V}_h) \to BV(\Omega)$ the piecewise constant embedding of $\ell^2(\mathcal{V}_h)$ into $BV(\Omega)$ defined by (2.11). Second, we define a piecewise linear interpolatory operator. $P_h \colon \ell^2(\mathcal{V}_h) \to BV(\Omega)$. For $u \in \ell^2(\mathcal{V}_h)$, $P_h u$ is the continuous piecewise linear polynomial defined by

$$P_h u(y) = \sum_{1 \leq i,j \leq N} u_{i,j} \phi_{i,j}(y), \tag{2.12}$$

with $\phi_{i,j}$ the continuous piecewise linear function such that

$$\phi_{i,j}(\omega_{i,j}) = 1, \text{ and } \phi_{i,j}(\omega) = 0, \ \omega \in \mathcal{V}_h \setminus \omega_{i,j}. \tag{2.13}$$

With a slight abuse of notation, P_h will also denote the piecewise linear interpolation of functions of class $\mathcal{C}^k(\bar{\Omega}),\ k \geq 0$ over the set $\mathcal{V}_h$; in which case it will be defined by

$$P_h f(y) = \sum_{1 \leq i,j \leq N} f(\omega_{i,j}) \phi_{i,j}(y), \quad y \in \Omega,\ f \in \mathcal{C}^k(\bar{\Omega}).$$

Lemma 2 *Suppose that Ω is endowed with the triangulation Δ_h. Then for all $u \in \ell^2(\mathcal{V}_h)$, there holds*

$$\|C_h u\|_2 = |u|_{\ell^2(\mathcal{V}_h)} \text{ and } \|P_h u\|_2 \leq 3|u|_{\ell^2(\mathcal{V}_h)}, \tag{2.14}$$

$$\|P_h u - C_h u\|_2 \leq \omega(u,1)_{\ell^2(\mathcal{V}_h)} \tag{2.15}$$

where $\omega(u,m)_{\ell^2(\mathcal{V}_h)}$ denotes the discrete modulus of smoothness of order $m \in \mathbb{N}$ and is defined by

$$\omega(u,m)^2_{\ell^2(\mathcal{V}_h)} = \sup_{\substack{\alpha=(\alpha_1,\alpha_2)\in\mathbb{Z}^2,\\ \max(|\alpha_1|,|\alpha_2|)\leq m}} \sum_{\substack{1\leq i,i+\alpha_1\leq N\\ 1\leq j,j+\alpha_2\leq N}} h^2 \left| u_{i+\alpha_1,j+\alpha_2} - u_{i,j} \right|^2. \tag{2.16}$$

Proof Let $u \in \ell^2(\mathcal{V}_h)$ be given. The inequalities (2.14) follow from the definition of the Euclidean norm on $\ell^2(\mathcal{V}_h)$ and the fact that the area of the support of $\phi_{i,j}$ is at most $3|\Omega_{i,j}|$. Next we prove the inequality (2.15). By definition, we have

$$\begin{aligned}
\|P_h u - C_h u\|_2^2 &= \sum_{1\leq i,j\leq N} \int_{\Omega_{i,j}} |P_h u(y) - C_h u(y)|^2 \, dy \\
&\leq \sum_{1\leq i,j\leq N} \int_{\Omega_{i,j}} \left| \sum_{1\leq l,k\leq N} \left(u_{l,k} - u_{i,j}\right) \phi_{l,k}(y) \right|^2 dy \\
&\leq \sum_{1\leq i,j\leq N} \sum_{1\leq l,k\leq N} \left|u_{l,k} - u_{i,j}\right|^2 \int_{\Omega_{i,j}} \phi^2_{l,k}(y) dy \\
&\leq \sum_{1\leq i,j\leq N} \sum_{\substack{1\leq l\leq N,\, |l-i|=1\\ 1\leq k\leq N,\, |k-j|=1}} |\Omega_{i,j}| \left|u_{l,k} - u_{i,j}\right|^2 \leq \omega(u,1)^2_{\ell^2(\mathcal{V}_h)},
\end{aligned}$$

which completes the proof. □

Remark 1 Let $u \in \mathbb{R}^{N\times N}$ be fixed, and let $\omega(u,m)$ be the modulus of continuity of $u \in \mathbb{R}^{N\times N}$ with respect to the Euclidian norm $|u| = \sqrt{\langle u,u\rangle}$. It is easy to see that

$$\omega(u,m)_{\ell^2(\mathcal{V}_h)} \leq h\omega(u,m). \tag{2.17}$$

2.4 Periodic Extension operators

In this section, we construct a periodic extension of $u \in \mathbb{R}^{N\times N}$ to $\mathbb{Z}^2$. We also present the construction of a periodic extension to $\mathbb{R}^2$ of a function defined on Ω.

The discrete periodic extension Let $u \in \mathbb{R}^{N\times N}$ be fixed. We construct a discrete function $\operatorname{Ext}(u)\colon \mathbb{Z}^2 \to \mathbb{R}$ defined in two steps as follows:

1) First, we extend u into an element $u^{2N} \in \mathbb{R}^{2N\times 2N}$ as follows

$$u^{2N}_{i,j} = \begin{cases} u_{i,j} & \text{if } 1 \le i,j \le N, \\ u_{2N-i+1,j} & \text{if } i > N, 1 \le j \le N, \\ u_{i,2N-j+1} & \text{if } 1 \le i \le N, j > N, \\ u_{2N-i+1,2N-j+1} & \text{else.} \end{cases} \tag{2.18}$$

2) $\operatorname{Ext}(u)$ is the periodic extension of u^{2N} to all of $\mathbb{Z}^2$.

In the sequel, we will denote by u^{2N} the restriction of the function $\operatorname{Ext}(u)$ to the set

$$\Psi_{2N} := \{(i,j) \in \mathbb{Z}\times\mathbb{Z}\colon 1 \le i,j \le 2N\}.$$

For $\alpha = (\alpha_1,\alpha_2) \in \mathbb{Z}^2$ and $u := (u_{i,j})_{i,j\in\mathbb{Z}}$ a discrete function, we define the translation $\tau_\alpha u$ of u as follows:

$$(\tau_\alpha u)_{i,j} = u_{i+\alpha_1,j+\alpha_2}, \quad i,j \in \mathbb{Z}.$$

The restriction of $\tau_\alpha u$ to a set of the form $\Psi_M := \{(i,j)\colon 1 \le i,j \le M\}$ shall be denoted u^M_α.

The continuous periodic extension The construction of the periodic extension that follows was presented in [16,15]. Let $u : \Omega \to \mathbb{R}$ be given. First, we extend u to $u^{2\Omega} : [0,2]^2 \to \mathbb{R}$ by the formula

$$\begin{aligned} u^{2\Omega}(x_1,x_2) &= u(x_1,x_2)\mathbb{1}_\Omega(x_1,x_2) + u(2-x_1,x_2)\mathbb{1}_{\tau_{1,0}\Omega}(x_1,x_2) \\ &\quad + u(x_1,2-x_2)\mathbb{1}_{\tau_{0,1}\Omega}(x_1,x_2) + u(2-x_1,2-x_2)\mathbb{1}_{\tau_{1,1}\Omega}(x_1,x_2), \end{aligned} \tag{2.19}$$

where $\mathbb{1}_A$ denotes the indicator function of the set A, and $\tau_{m,n}\Omega = \{(x_1+m, x_2+n)\colon (x_1,x_2) \in \Omega\}$. The extension of u is the periodic function $\operatorname{Ext}(u) : \mathbb{R}^2 \to \mathbb{R}$ with period $[0,2]^2$ that coincides with $u^{2\Omega}$ on $[0,2]^2$. Since Ω has compact closure, it is easy to show that the continous periodic extension operator, Ext, maps $\mathrm{L}^p(\Omega)$ into $\mathrm{L}^p_{\mathrm{loc}}(\mathbb{R}^2)$.

Proposition 1 *Suppose that $f \in \mathrm{L}^2(\Omega)$. Then for any $0 < h \ll 1$, we have*

$$\omega_2(\operatorname{Ext}(f),h)_{\Omega^1} \le C\omega_2(f,h), \tag{2.20}$$

where $C > 0$ is a constant independent of f and $\Omega^1 = \bigcup_{\substack{(m,n)\in\mathbb{Z}^2 \\ |(m,n)|_\infty \le 1}} \tau_{(m,n)}\Omega.$

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

$$\begin{aligned} \|\tau_\eta(\operatorname{Ext}(f)\mathbb{1}_{\Omega^1}) - \operatorname{Ext}(f)\mathbb{1}_{\Omega^1}\|_2^2 &= \int_{\Omega^1\cap\tau_{-\eta}\Omega^1} |\operatorname{Ext}(f)(x+\eta) - \operatorname{Ext}(f)(x)|^2 dx \\ &\le \sum_{-1\le i,j\le 1} \sum_{\substack{-1\le m\le 1 \\ |m-i|=1 \\ -1\le n\le 1 \\ |n-j|=1}} \int_{\tau_{(m,n)}\Omega\cap\tau_{(i,j)-\eta}\Omega} |\operatorname{Ext}(f)(x+\eta) - \operatorname{Ext}(f)(x)|^2 dx \\ &\le 2 \sum_{-1\le i,j\le 1} \int_{\tau_{(i,j)}(\Omega\cap\tau_{-\eta}\Omega)} |\operatorname{Ext}(f)(x+\eta) - \operatorname{Ext}(f)(x)|^2 dx \\ &\le 18 \int_{\Omega\cap\tau_{-\eta}\Omega} |f(x+\eta) - f(x)|^2 dx = 18\|\tau_\eta f - f\|_2^2. \end{aligned}$$

As a consequence, we have $\omega_2(\operatorname{Ext}(f),h)_{\Omega^1} \le 3\sqrt{2}\,\omega_2(f,h)$. □

Lemma 3 *For any $f \in \mathrm{L}^2(\Omega)$ and $0 < h \ll 1$, there holds*

$$\|f - C_h Q_h f\|_2 \le K\,\omega_2(f,h), \tag{2.21}$$

where $K > 0$ is a constant independent of h.

Proof By definition, we have

$$
\begin{aligned}
\|f - C_h Q_h f\|_2 &= \left(\sum_{1\le i,j\le N} \int_{\Omega_{i,j}} \left| f(x) - \frac{1}{h^2} \int_{\Omega_{i,j}} f(y)dy \right|^2 dx \right)^{1/2} \\
&\le \left(\sum_{1\le i,j\le N} \int_{\Omega_{i,j}} \left(\frac{1}{h^2} \int_{\Omega_{i,j}} |f(x) - f(y)|\, dy \right)^2 dx \right)^{1/2} \\
&\le \left(\sum_{1\le i,j\le N} \int_{\Omega_{i,j}} \left(\frac{4}{h^2} \int_{\{z:\,|z|\le\sqrt{2}h\}} |\mathrm{Ext}(f)(x) - \mathrm{Ext}(f)(x+z)|\, dz \right)^2 dx \right)^{1/2} \\
&= \left(\int_{\Omega} \left(\frac{4}{h^2} \int_{\{z:\,|z|\le\sqrt{2}h\}} |\mathrm{Ext}(f)(x) - \mathrm{Ext}(f)(x+z)|\, dz \right)^2 dx \right)^{1/2} \\
&\le \frac{4}{h^2} \int_{\{z:\,|z|\le\sqrt{2}h\}} \left(\int_{\Omega} |\mathrm{Ext}(f)(x) - \mathrm{Ext}(f)(x+z)|^2\, dx \right)^{1/2} dz \\
\|f - C_h Q_h f\|_2 &\le 8\pi\omega_2(\mathrm{Ext}(f), \sqrt{2}h)_{\Omega^1}.
\end{aligned}
$$

Inequality (2.21) follows from the latter inequality thanks to Proposition 1, and the fact that

$$
\omega_2(\mathrm{Ext}(f), \sqrt{2}h)_{\Omega^1} \le 2\,\omega_2(\mathrm{Ext}(f), h)_{\Omega^1}.
$$

□

3 A total variation based model for digital image denoising

In the sequel we treat gray-scale images of size $N \times N$ as rectangular matrices of dimension $N \times N$ and denote by $X := \mathbb{R}^{N\times N}$ the manifold of gray-scale images of size $N \times N$.

3.1 The model and its properties

To compute the total variation of elements of X, we introduce two discrete gradient operators $\nabla_+ = (\nabla_+^x, \nabla_+^y)$ and $\nabla_- = (\nabla_-^x, \nabla_-^y)$, which are linear operators from X into $Y := X \times X$, defined by

$$
(\nabla_+^x u)_{i,j} = \begin{cases} 0, & \text{if } i = N \text{ or } j = N \\ u_{i+1,j} - u_{i,j} & \text{otherwise;} \end{cases} \tag{3.1}
$$

$$
(\nabla_+^y u)_{i,j} = \begin{cases} 0, & \text{if } i = N \text{ or } j = N \\ u_{i,j+1} - u_{i,j} & \text{otherwise;} \end{cases} \tag{3.2}
$$

and

$$
(\nabla_-^x u)_{i,j} = \begin{cases} 0, & \text{if } i = 1 \text{ or } j = 1 \\ u_{i,j} - u_{i-1,j} & \text{otherwise;} \end{cases} \tag{3.3}
$$

$$
(\nabla_-^y u)_{i,j} = \begin{cases} 0, & \text{if } i = 1 \text{ or } j = 1 \\ u_{i,j} - u_{i,j-1} & \text{otherwise.} \end{cases} \tag{3.4}
$$

We associate to the discrete gradient operators ∇_+ and ∇_- the discrete divergence operators, $\mathrm{div}_+ := -\nabla_+^* : Y \to X$ and $\mathrm{div}_- := -\nabla_-^* : Y \to X$, defined respectively by

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

and

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

The total variation of a gray-scale image $u \in X$ is then defined by

$$J(u) := \frac{1}{2} \sum_{\substack{1\le i\le N \\ 1\le j\le N}} \left(\left|(\nabla_+ u)_{i,j}\right| + \left|(\nabla_- u)_{i,j}\right| \right), \tag{3.5}$$

where for an element $p := (p_{i,j}) := (p^1_{i,j}, p^2_{i,j}) \in Y$, $\left|p_{i,j}\right|$ is the Euclidean norm of $p_{ij} := (p^1_{i,j}, p^2_{i,j})$ in $\mathbb{R}^2$. Consequently, the digital ROF-functional is given by

$$E^f_{d,\lambda}(u) := J(u) + \frac{1}{2\lambda}|u-f|^2, \tag{3.6}$$

where $\|\cdot\|_X$ is the Euclidean norm on $X = \mathbb{R}^{N\times N}$, and $f \in X$.

Since the functional $E^f_{d,\lambda}(u)$ is strictly convex and coercive, it follows (cf. [9, Proposition 1.2, page 35]) that the minimization problem $\operatorname*{argmin}_{u\in X} E^f_{d,\lambda}(u)$ has a unique solution. Henceforth, we will refer to the problem

$$\operatorname*{argmin}_{u\in X} E^f_{d,\lambda}(u) \tag{3.7}$$

as the digital ROF-model.

The next result, similar to the one proved in [15, 16], shows that problem (3.7) is stable under small perturbations in the datum f.

Lemma 4 (Stability) *Let u_f be the minimizer of $E^f_{d,\lambda}(u)$ and u_g the minimizer of $E^g_{d,\lambda}(u)$. Then*

$$|u_f - u_g| \le |f-g|. \tag{3.8}$$

We now prove a technical lemma that asserts that the discrete ROF model is compatible with translation of the datum f.

Lemma 5 *Let $f \in \mathbb{R}^{N\times N}$ be fixed and u_f be the minimizer of the functional $E^f_{d,\lambda}(u)$. Then, u^{2N}_f is a minimizer of $E^{f^{2N}}_{d,\lambda}(u)$. Moreover, for any $\alpha \in \mathbb{Z}^2$ such that $|\alpha|_\infty = 1$, the minimizer of $E^{f^{2N}_\alpha}_{d,\lambda}(u)$ is $u^{2N}_{f,\alpha}$, the restriction of $\tau_\alpha u_f$ to Ψ_{2N}.*

Proof We first show that $u^{2N}_f = \operatorname*{argmin}_{u\in\mathbb{R}^{2N\times 2N}} E^{f^{2N}}_{d,\lambda}(u)$. It is easy to see that for all $u \in \mathbb{R}^{N\times N}$, we have

$$E^{f^{2N}}_{d,\lambda}(u^{2N}) = 4E^f_{d,\lambda}(u)$$

Furthermore, for every $u \in \mathbb{R}^{2N\times 2N}$, there exists $x_u \in \mathbb{R}^{N\times N}$ such that $E^{f^{2N}}_{d,\lambda}(u) \ge E^{f^{2N}}_{d,\lambda}(x^{2N}_u)$. In effect, letting x_u be the constant vector that equals u_{i_0,j_0} with $(i_0,j_0) \in \operatorname*{argmin}_{1\le i,j\le N} |u_{i,j} - f_{i,j}|$, it is easy to verify that $E^{f^{2N}}_{d,\lambda}(u) \ge E^{f^{2N}}_{d,\lambda}(x^{2N}_u)$. Consequently,

$$\min_{u\in\mathbb{R}^{2N\times 2N}} E^{f^{2N}}_{d,\lambda}(u) = \min_{u\in\mathbb{R}^{N\times N}} E^{f^{2N}}_{d,\lambda}(u^{2N}) = 4 \min_{u\in\mathbb{R}^{N\times N}} E^f_{d,\lambda}(u).$$

Hence, u^{2N}_f is the minimizer of $E^{f^{2N}}_{d,\lambda}(u)$ with respect to $\mathbb{R}^{2N\times 2N}$.

Next let $\alpha \in \mathbb{Z}^2$ with $|\alpha|_\infty = 1$. For clarity of the argument, we fix $\alpha = (1,0)$. An argument identical to the one above shows that

$$\min_{u\in\mathbb{R}^{2N\times 2N}} E^{f^{2N}_\alpha}_{d,\lambda}(u) = \min_{u\in\mathbb{R}^{N\times N}} E^{f^{2N}_\alpha}_{d,\lambda}(u^{2N}_\alpha).$$

Let $k \in \mathbb{N}$ be fixed. Then, for any periodic discrete function g with period $[1,2N]\times[1,2N]$, we have

$$k^2 E^{g^{2N}}_{d,\lambda}(u^{2N}_g) = E^{g^{2Nk}}_{d,\lambda}(u^{2Nk}_g) = \min_{u\in\mathbb{R}^{2Nk\times 2Nk}} E^{g^{2Nk}}_{d,\lambda}(u), \tag{3.9}$$

where u_g is the minimizer of $E_{d,\lambda}^{g^{2N}}(u)$, and u_g^{2Nk} is the restriction of the periodic (with period $[1,2N]\times[1,2N]$) extension of u_g to $[1,2Nk]\times[1,2Nk]$. Moreover, for any $u\in\mathbb{R}^{N\times N}$, there holds

$$E_{d,\lambda}^{g_\alpha^{2Nk}}(u_\alpha^{2Nk}) = E_{d,\lambda}^{g^{2Nk}}(u^{2Nk}) - kD(u), \tag{3.10}$$

where

$$D(u) = \sum_{1\leq j<N} \left(|\nabla_+(u)_{1,j}| + |\nabla_-(u)_{1,j}|\right). \tag{3.11}$$

Let $y\in\mathbb{R}^{N\times N}$ be such that $y_\alpha^{2N} = \underset{u\in\mathbb{R}^{2N\times 2N}}{\operatorname{argmin}} E_{d,\lambda}^{f_\alpha^{2N}}(u)$. We want to prove that $y_\alpha^{2N} = u_{f,\alpha}^{2N}$, i.e $E_{d,\lambda}^{f_\alpha^{2N}}(u_{f,\alpha}^{2N}) \leq E_{d,\lambda}^{f_\alpha^{2N}}(y_\alpha^{2N})$. Let us consider the minimization problem associated to f_α^{2Nk}. Since u_f^{2N} is the minimizer of $E_\lambda^{f^{2N}}(u)$, we have

$$\begin{aligned}
k^2 E_{d,\lambda}^{f_\alpha^{2N}}(y_\alpha^{2N}) &= E_{d,\lambda}^{f_\alpha^{2Nk}}(y_\alpha^{2Nk})\\
&= E_{d,\lambda}^{f^{2Nk}}(y^{2Nk}) - kD(y) \text{ by equation (3.10)}\\
&\geq E_{d,\lambda}^{f^{2Nk}}(u_f^{2Nk}) - kD(y) \text{ by equation (3.9)}\\
&\geq E_{d,\lambda}^{f_\alpha^{2Nk}}(u_{f,\alpha}^{2Nk}) + k(D(u_f) - D(y)) \text{ by equation (3.10)}\\
k^2 E_{d,\lambda}^{f_\alpha^{2N}}(y_\alpha^{2N}) &\geq k^2 E_{d,\lambda}^{f_\alpha^{2N}}(u_{f,\alpha}^{2N}) + k(D(u_f) - D(y)) \text{ by equation (3.9).}
\end{aligned}$$

Dividing the last inequality above by k^2, we obtain

$$E_{d,\lambda}^{f_\alpha^{2N}}(y_\alpha^{2N}) \geq E_{d,\lambda}^{f_\alpha^{2N}}(u_{f,\alpha}^{2N}) + \frac{(D(u_f) - D(y))}{k}, \quad \forall k\in\mathbb{N},\, k\geq 1. \tag{3.12}$$

Passing to the limit as $k\to\infty$ in (3.12) yields $E_{d,\lambda}^{f_\alpha^{2N}}(y_\alpha^{2N}) \geq E_{d,\lambda}^{f_\alpha^{2N}}(u_{f,\alpha}^{2N})$. Thus, $u_{f,\alpha}^{2N} = \underset{u\in\mathbb{R}^{N\times N}}{\operatorname{argmin}} E_{d,\lambda}^{f_\alpha^{2N}}(u_\alpha^{2N})$.

Our argument above works mutatis mutandis for any $\alpha := (\alpha_1,\alpha_2)$ with one of the components being zero. For $\alpha\in\{(1,1),(1,-1),(-1,1),(-1,-1)\}$, the proof follows from the previous case by observing that $\tau_\alpha = \tau_{(\alpha_1,0)}\circ\tau_{(0,\alpha_2)}$. □

Remark 2 We observe that for any $u\in\mathbb{R}^{N\times N}$

$$E_{\lambda,h}^f(u) = h^2 E_{d,\lambda/h^2}^{f/h}(u/h). \tag{3.13}$$

Therefore, Lemma 4 and Lemma 5 remain valid for the functional $E_{\lambda,h}^f(u)$. The equation (3.13) gives the relation between the discrete ROF functional (2.5) and the digital ROF functional (3.6).

3.2 Primal-Dual formulation

We establish the primal-dual formulation of the minimization problem associated to the functional $E_{d,\lambda}^f(u)$. We also prove the existence and uniqueness of the minimizer. First, we establish an alternate formula for the discrete total variation $J(u)$ defined in (3.5). By Riesz representation Theorem, we can rewrite the discrete total variation $J(u)$ above in the following form:

$$\begin{aligned}
J(u) &= \frac{1}{2}\sum_{\substack{1\leq i\leq N\\1\leq j\leq N}} \sup_{p_{i,j}\in\mathbb{R}^2,\,|p_{i,j}|\leq 1} \langle(\nabla_+ u)_{i,j}, p_{i,j}\rangle + \sup_{q_{i,j}\in\mathbb{R}^2,\,|q_{i,j}|\leq 1} \langle(\nabla_- u)_{i,j}, q_{i,j}\rangle\\
&= \frac{1}{2}\sup_{p\in Y,\,|p|_\infty\leq 1}\langle\nabla_+ u, p\rangle_Y + \frac{1}{2}\sup_{q\in Y,\,|q|_\infty\leq 1}\langle\nabla_- u, q\rangle_Y\\
&= \frac{1}{2}\sup_{p\in Y,\,|p|_\infty\leq 1}\langle u, -\operatorname{div}_+(p)\rangle_X + \frac{1}{2}\sup_{q\in Y,\,|q|_\infty\leq 1}\langle u, -\operatorname{div}_-(q)\rangle_X
\end{aligned}$$

$$J(u) = \sup_{\substack{p\in Y,\,|p|_\infty\leq 1\\ q\in Y,\,|q|_\infty\leq 1}} -\langle u, \frac{1}{2}\operatorname{div}_+(p) + \frac{1}{2}\operatorname{div}_-(q)\rangle_X \tag{3.14}$$

where for $p = (p^1, p^2)$ and $q = (q^1, q^2)$ in Y,

$$\langle p, q\rangle_Y = \sum_{1\le i,j\le N} p^1_{i,j} q^1_{i,j} + p^2_{i,j} q^2_{i,j}, \text{ and } |p|_\infty = \max_{i,j} \sqrt{(p^1_{ij})^2 + (p^2_{ij})^2}.$$

Therefore, the minimization problem (3.7) is equivalent to the following saddle-point problem

$$\underset{u\in X}{\operatorname{argmin}} \sup_{\substack{p\in Y,\, |p|_\infty\le 1\\ q\in Y,\, |q|_\infty \le 1}} -\langle u, \frac{1}{2}\operatorname{div}_+(p) + \frac{1}{2}\operatorname{div}_-(q)\rangle_X + \frac{1}{2\lambda}|u-f|^2. \tag{3.15}$$

The saddle-point problem (3.15) above is referred to as the primal-dual formulation of problem (3.7), with primal variable u and dual variable p.

Let

$$B_Y = \{p \in Y\colon |p|_\infty \le 1\},$$

and $\mathcal{L}(u : p, q)$ be the functional defined on $X \times Y \times Y$ by

$$\mathcal{L}(u; p, q) := -\langle u, \frac{1}{2}\operatorname{div}_+(p) + \frac{1}{2}\operatorname{div}_-(q)\rangle + \frac{1}{2\lambda}|u-f|^2. \tag{3.16}$$

We have the following result for the existence of a solution to problem (3.15).

Lemma 6 *The functional $\mathcal{L} : X \times Y \times Y \to \mathbb{R}$ defined by* (3.16) *has a saddle point $(\bar u; \bar p, \bar q)$ in the set $X \times B_Y \times B_Y$. Furthermore the component $\bar u$ of any saddle point of $\mathcal{L}$ is the solution of the minimization problem* (3.7).

Proof We observe that $\mathcal{L}(u; p, q)$ is a strictly convex quadratic function in u and linear in (p, q). Moreover, both the mappings $u \mapsto \mathcal{L}(u; p, q)$ and $(p, q) \mapsto \mathcal{L}(u; p, q)$ are Gâteaux-differentiable, with derivatives

$$\partial_u \mathcal{L}(u; p, q) = -\frac{1}{2}\left(\operatorname{div}_+(p) + \operatorname{div}_-(q)\right) - \frac{1}{\lambda}\left(u - f\right), \tag{3.17}$$

$$\partial_{p,q}\mathcal{L}(u; p, q) = \left[\frac{1}{2}\nabla_+ u, \frac{1}{2}\nabla_- u\right]. \tag{3.18}$$

Let

$$\bar u = f + \frac{\lambda}{2}\left(\operatorname{div}_+(\bar p) + \operatorname{div}_-(\bar q)\right), \tag{3.19}$$

with

$$(\bar p, \bar q) \in \underset{p,q\in B_Y}{\operatorname{argmin}} |\lambda\left(\operatorname{div}_+(p) + \operatorname{div}_-(q)\right) + 2f|^2. \tag{3.20}$$

We show that $(\bar u; \bar p, \bar q)$ is a saddle-point for $\mathcal{L}(u; p, q)$ with respect to the set $X \times B_Y \times B_Y$. To this end, it suffices to check that

$$\langle \partial_u \mathcal{L}(\bar u; \bar p, \bar q), u - \bar u\rangle \ge 0, \quad \forall u \in X$$
$$\langle \partial_{p,q}\mathcal{L}(\bar u; \bar p, \bar q), (p - \bar p, q - \bar q)\rangle \le 0, \quad \forall\, p, q \in B_Y.$$

which follows by definition of $\bar u$ and $(\bar p, \bar q)$. In effect, by definition of $\bar u$, we have $\partial_u \mathcal{L}(\bar u; \bar p, \bar q) = 0$. Thus,

$$\langle \partial_u \mathcal{L}(\bar u; \bar p, \bar q), u - \bar u\rangle = 0, \quad \forall u \in X.$$

Next, it is easy to see that for all $p, q \in Y$ we have

$$\langle \partial_{p,q}\mathcal{L}(\bar u; \bar p, \bar q), (p - \bar p, q - \bar q)\rangle = -\frac{1}{2\lambda}\langle \lambda(\operatorname{div}_+ \bar p + \operatorname{div}_- \bar q) + 2f, \lambda(\operatorname{div}_+(p - \bar p) + \operatorname{div}_-(q - \bar q))\rangle.$$

Now, we recall that $\lambda(\operatorname{div}_+(\bar p) + \operatorname{div}_-(\bar q))$ is the orthogonal projection of $-2f$ onto the closed convex set λK. Consequently, by the characterization of the orthogonal projection, we have

$$\langle \lambda(\operatorname{div}_+ \bar p + \operatorname{div}_- \bar q) + 2f, \lambda(\operatorname{div}_+(p - \bar p) + \operatorname{div}_-(q - \bar q))\rangle \ge 0, \quad \forall\, p, q \in B_Y.$$

Hence, by (3.18)

$$\partial_{p,q}\mathcal{L}(\bar u; \bar p, \bar q), (p - \bar p, q - \bar q)\rangle \le 0 \text{ for all } p, q \in B_Y.$$

□

Remark 3 If $(\bar{u};\bar{p},\bar{q})$ is a saddle-point of $\mathcal{L}(u;p,q)$ with respect to the set $X \times B_Y \times B_Y$, then

$$\begin{aligned}\mathcal{L}(\bar{u};\bar{p},\bar{q}) &= \inf_{u\in X} J(u) + \frac{1}{2\lambda}|u-f|^2 = J(\bar{u}) + \frac{1}{2\lambda}|\bar{u}-f|^2 \\ &= \max_{p,q\in B_Y} \frac{2}{\lambda}|f|^2 - \frac{\lambda}{8}\left|\mathrm{div}_+(p) + \mathrm{div}_-(q) + \frac{2f}{\lambda}\right|^2.\end{aligned}$$

Hence, for any saddle-point $(\bar{u};\bar{p},\bar{q})$ of $\mathcal{L}(u;p,q)$, we have

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

We have proved the following characterization of a minimizer of the discrete ROF functional $E^f_{d,\lambda}$.

Theorem 1 *A vector $u \in X$ is the solution of problem* (3.7) *if and only if $2\,(u-f)$ is the orthogonal projection of $-2f$ onto the closed convex set*

$$\lambda K := \{\lambda\,[\mathrm{div}_+(p) + \mathrm{div}_-(q)]\ :\ p,q \in B_Y\}\,. \tag{3.21}$$

Furthermore, the solution u is given by

$$u = f + \frac{\lambda}{2}\,(\mathrm{div}_+(\bar{p}) + \mathrm{div}_-(\bar{q}))\,,$$

with $(\bar{p},\bar{q}) \in \underset{p,q\in B_Y}{\mathrm{argmin}}\,|\lambda\,(\mathrm{div}_+(p) + \mathrm{div}_-(q)) + 2f|^2$.

For a closed convex subset A of a Euclidean space, we denote by π_A the orthogonal projection onto A. In particular for $A = B_Y$, it can be shown that

$$\pi_{B_Y}(p)_{i,j} = \left(\frac{p^1_{i,j}}{\max(1,|p_{i,j}|)}, \frac{p^2_{i,j}}{\max(1,|p_{i,j}|)}\right), \quad 1 \le i,j \le N. \tag{3.22}$$

The formula above follows from the observation that B_Y is the cartesian product of unit disks of $\mathbb{R}^2$, and $\mathbb{R}^2$ can be isometrically embedded in Y, with both spaces endowed with their natural Euclidian norms. Therefore, for $1 \le i,j \le N$, $\pi_{B_Y}(p)_{i,j}$ is the orthogonal projection of $p_{i,j} = (p^1_{i,j}, p^2_{i,j})$ onto the unit disk of $\mathbb{R}^2$, given by (3.22).

The following result is a straightforward consequence of the characterization of minimizers of a Gâteaux-differentiable function and forms the basis of our projected-gradient algorithm.

Proposition 2 *If* $(\bar{p},\bar{q}) \in \underset{p,q\in B_Y}{\mathrm{argmin}}\,|\lambda\,(\mathrm{div}_+(p) + \mathrm{div}_-(q)) + 2f|^2$*, then for any $\tau > 0$, there holds*

$$\begin{cases}\bar{p} = \pi_{B_Y}\left(\bar{p} + \tau\nabla_+\left[\mathrm{div}_+(\bar{p}) + \mathrm{div}_-(\bar{q}) + 2f/\lambda\right]\right), \\ \bar{q} = \pi_{B_Y}\left(\bar{q} + \tau\nabla_-\left[\mathrm{div}_+(\bar{p}) + \mathrm{div}_-(\bar{q}) + 2f/\lambda\right]\right).\end{cases} \tag{3.23}$$

Proof Let $(\bar{p},\bar{q}) \in \underset{p,q\in B_Y}{\mathrm{argmin}}|\lambda(\mathrm{div}_+(p) + \mathrm{div}_-(q) + 2f)|^2$ be fixed. Then,

$$\begin{aligned}\bar{p} &\in \underset{p\in B_Y}{\mathrm{argmin}}\,|\,\mathrm{div}_+(p) + \mathrm{div}_-(\bar{q}) + 2f/\lambda|^2, \\ \bar{q} &\in \underset{q\in B_Y}{\mathrm{argmin}}\,|\,\mathrm{div}_+(\bar{p}) + \mathrm{div}_-(q) + 2f/\lambda)|^2.\end{aligned}$$

We notice that the mappings $p \mapsto |\,\mathrm{div}_+(p) + \mathrm{div}_-(\bar{q}) + 2f/\lambda|^2$ and $q \mapsto |\,\mathrm{div}_+(\bar{p}) + \mathrm{div}_-(q) + 2f/\lambda|^2$ are Gâteaux-differentiable with differentials

$$-\nabla_+(\mathrm{div}_+(p) + \mathrm{div}_-(\bar{q}) + 2f/\lambda) \text{ and } -\nabla_-(\mathrm{div}_+(\bar{p}) + \mathrm{div}_-(q) + 2f/\lambda),$$

respectively. Therefore (see [9, Proposition 2.1, page 37]), we have

$$\begin{aligned}\langle -\nabla_+(\mathrm{div}_+(\bar{p}) + \mathrm{div}_-(\bar{q}) + 2f/\lambda), p - \bar{p}\rangle \ge 0, &\quad \forall p \in B_Y, \\ \langle -\nabla_-(\mathrm{div}_+(\bar{p}) + \mathrm{div}_-(\bar{q}) + 2f/\lambda), q - \bar{q}\rangle \ge 0, &\quad \forall q \in B_Y.\end{aligned}$$

It then follows that for any $\tau > 0$

$$\begin{aligned}\langle \bar{p} - [\bar{p} + \tau\nabla_+(\mathrm{div}_+(\bar{p}) + \mathrm{div}_-(\bar{q}) + 2f/\lambda)]\,, p - \bar{p}\rangle \ge 0, &\quad \forall\, p \in B_Y, \\ \langle \bar{q} - [\bar{q} + \tau\nabla_-(\mathrm{div}_+(\bar{p}) + \mathrm{div}_-(\bar{q}) + 2f/\lambda)]\,, q - \bar{q}\rangle \ge 0, &\quad \forall\, q \in B_Y,\end{aligned}$$

which is equivalent to (3.23). □

3.3 A projected gradient algorithm

By Theorem 1, computing the minimizer of the discrete ROF functional $E^f_{d,\lambda}(u)$ is equivalent to projecting $-2f$ onto the set λK defined by (3.21). In view of Proposition 2, we propose a projected gradient algorithm for computing the projection of $-2f$ onto λK. We prove the convergence of the proposed algorithm by adapting the argument used in [8].

Algorithm 31 (Projected-Gradient) *Let $\tau > 0$ be fixed. For $n = 0$, let $p_0 = q_0 = 0$.*

Step 1: *Compute u_n*

$$u_n = f + \frac{\lambda}{2}\left[\operatorname{div}_+(p_n) + \operatorname{div}_-(q_n)\right]. \tag{3.24}$$

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

$$\begin{cases} p_{n+1} = \pi_{B_Y}\left(p_n + \tau\nabla_+\left[\operatorname{div}_+(p_n) + \operatorname{div}_-(q_n) + 2f/\lambda\right]\right), \\ q_{n+1} = \pi_{B_Y}\left(q_n + \tau\nabla_-\left[\operatorname{div}_+(p_n) + \operatorname{div}_-(q_n) + 2f/\lambda\right]\right), \end{cases} \tag{3.25}$$

with π_{B_Y} given by the equation (3.22).

Step 3: *Increment n by 1, then go back to step 1.*

We have the following convergence result for Algorithm 31.

Proposition 3 *If $0 < \tau < 1/8$, then Algorithm 31 converges. More precisely, for any $(\bar p, \bar q)$ satisfying (3.20), we have $\lim\limits_{n\to\infty}(\operatorname{div}_+(p_n) + \operatorname{div}_-(q_n)) = \operatorname{div}_+(\bar p) + \operatorname{div}_-(\bar q)$.*

Proof Let $(\bar p, \bar q)$ satisfying (3.23) be fixed and $\tau > 0$ be given. Since the projection mapping π_{B_Y} is non-expansive, we infer from (3.23) that

$$|p_{n+1} - \bar p|^2 + |q_{n+1} - \bar q|^2 \le \|Id - \tau A\|(|p_n - \bar p|^2 + |q_n - \bar q|^2),$$

where $A : Y \times Y \to Y \times Y$ is the linear operator

$$A := \begin{pmatrix} -\nabla_+ \operatorname{div}_+ & -\nabla_+ \operatorname{div}_- \\ -\nabla_- \operatorname{div}_+ & -\nabla_- \operatorname{div}_- \end{pmatrix}.$$

Consequently, as long as $\tau > 0$ is such that $\kappa := \|Id - \tau A\| = 1$, we get

$$\left|\begin{pmatrix} p_{n+1} - \bar p \\ q_{n+1} - \bar q \end{pmatrix}\right| \le \left|\begin{pmatrix} p_n - \bar p \\ q_n - \bar q \end{pmatrix}\right|, \quad \forall n \ge 0. \tag{3.26}$$

Next, we show that $\kappa = 1$ for $0 < \tau < 1/8$. We note to begin that A is self-adjoint and nonnegative definite. Indeed for $p,\ q \in Y$, we have

$$\begin{aligned} \left\langle A\begin{pmatrix} p \\ q \end{pmatrix}, \begin{pmatrix} p \\ q \end{pmatrix}\right\rangle_{Y\times Y} &= \langle -\nabla_+(\operatorname{div}_+(p) + \operatorname{div}_-(q)), p\rangle_Y + \langle -\nabla_-(\operatorname{div}_+(p) + \operatorname{div}_-(q)), q\rangle_Y \\ &= \langle \operatorname{div}_+(p) + \operatorname{div}_-(q), \operatorname{div}_+(p)\rangle_X + \langle \operatorname{div}_+(p) + \operatorname{div}_-(q), \operatorname{div}_-(q)\rangle_X \\ &= \langle \operatorname{div}_+(p) + \operatorname{div}_-(q), \operatorname{div}_+(p) + \operatorname{div}_-(q)\rangle_X \ge 0. \end{aligned}$$

Thus, $Y \times Y = \ker(A) \overset{\perp}{\oplus} F$,where F is the closure if $\operatorname{Im}(A)$. Moreover, all of the eigenvalues of A are nonnegative and

$$\kappa = \max(1, |1 - \tau\|A\||),$$

where $\|A\|$ denotes the spectral norm of A (the largest eigenvalue of A). So, for any τ such that $0 < \tau \le 2/\|A\|$, we have $\|Id - \tau A\| = 1$. Now, we compute an upper-bound for $\|A\|$.

We recall that by definition

$$\|A\|^2 = \sup_{(p,q)\neq(0,0)} \frac{\left\langle A\begin{pmatrix} p \\ q \end{pmatrix}, A\begin{pmatrix} p \\ q \end{pmatrix}\right\rangle_{Y\times Y}}{|p|^2 + |q|^2}.$$

By Cauchy-Schwarz inequality and the definition of the norm of a linear operator, we have

$$\begin{aligned}
\|A\|^2 &= \sup_{(p,q)\neq(0,0)} \frac{\left\langle A\begin{pmatrix}p\\q\end{pmatrix}, A\begin{pmatrix}p\\q\end{pmatrix}\right\rangle_{Y\times Y}}{|p|^2+|q|^2} \\
&\leq \sup_{(p,q)\neq(0,0)} \frac{(\|\nabla_+\|^2+\|\nabla_-\|^2)|\operatorname{div}_+(p)+\operatorname{div}_-(q)|^2}{|p|^2+|q|^2} \\
&\leq (\|\nabla_+\|^2+\|\nabla_-\|^2) \sup_{(p,q)\neq(0,0)} \frac{\|\operatorname{div}_+\|^2|p|^2 + 2\|\operatorname{div}_+\|\|\operatorname{div}_-\||p||q| + \|\operatorname{div}_-\|^2|q|^2}{|p|^2+|q|^2} \\
&\leq 2(\|\nabla_+\|^2+\|\nabla_-\|^2)\max(\|\operatorname{div}_+\|^2, \|\operatorname{div}_-\|^2)
\end{aligned}$$

Thus, it is the case that

$$\|A\| \leq \sqrt{2(\|\nabla_+\|^2+\|\nabla_-\|^2)\max(\|\operatorname{div}_+\|^2, \|\operatorname{div}_-\|^2)}. \tag{3.27}$$

By definition of div_+, for all $u \in X$ and all $p \in Y$

$$\begin{aligned}
\langle \nabla_+ u, \nabla_+ u\rangle_Y &= \langle u, -\operatorname{div}_+(\nabla_+ u)\rangle_X \\
\langle \operatorname{div}_+(p), \operatorname{div}_+(p)\rangle_X &= \langle p, -\nabla_+(\operatorname{div}_+(p))\rangle_Y.
\end{aligned}$$

Consequently, by Cauchy-Schwarz inequality and the definition of the norm of a linear operator, we obtain

$$\|\nabla_+\|^2 \leq \|\operatorname{div}_+\| \cdot \|\nabla_+\| \text{ and } \|\operatorname{div}_+\|^2 \leq \|\nabla_+\| \cdot \|\operatorname{div}_+\|$$

which shows that $\|\nabla_+\| = \|\operatorname{div}_+\|$. A similar argument shows that $\|\nabla_-\| = \|\operatorname{div}_-\|$. Now, for $u \in X$,

$$\begin{aligned}
\|\nabla_+ u\|^2 &= \sum_{1\leq i,j\leq N-1} (u_{i+1,j}-u_{i,j})^2 + (u_{i,j+1}-u_{i,j})^2 \\
&\leq 2\sum_{1\leq i,j\leq N-1} u_{i+1,j}^2 + 2u_{i,j}^2 + u_{i,j+1}^2 \leq 8\|u\|^2.
\end{aligned}$$

Hence, $\|\nabla_+\|^2 \leq 8$. A similar argument shows that $\|\operatorname{div}_-\|^2 = \|\nabla_-\|^2 \leq 8$. Thus, by (3.27) we have that

$$\|A\| \leq 16. \tag{3.28}$$

Therefore, for any $0 < \tau < 1/8$, there holds the inequality $\kappa = \|Id - \tau A\| = 1$.

We now show that the sequence $(\operatorname{div}_+(p_n) + \operatorname{div}_-(q_n))_{n\geq 0}$ converges by showing that all of its convergent subsequences have the same limit. Let $(\operatorname{div}_+(p_{n_k}) + \operatorname{div}_-(q_{n_k}))_{k\geq 0}$ be a convergent subsequence. There exists a further subsequence, which we denote again by n_k, such that $p_{n_k} \to \tilde{p}$ and $q_{n_k} \to \tilde{q}$. Since the projection π_{B_Y} is non-expansive, it follows from (3.25) and (3.26) that $p_{n_k+1} \to \hat{p}$, $q_{n_k+1} \to \hat{q}$, and

$$\left|\begin{pmatrix}\hat{p}-\bar{p}\\ \hat{q}-\bar{q}\end{pmatrix}\right| = \left|\begin{pmatrix}\tilde{p}-\bar{p}\\ \tilde{q}-\bar{q}\end{pmatrix}\right|,$$

where $|\cdot|$ is the Euclidean norm on $Y \times Y$ induced by the Euclidean norm of Y. As a consequence of the above equality and the equation

$$\begin{cases}\hat{p} = \pi_{B_Y}\left(\tilde{p} + \tau\nabla_+(\operatorname{div}_+(\tilde{p}) + \operatorname{div}_-(\tilde{q}))\right), \\ \hat{q} = \pi_{B_Y}\left(\tilde{q} + \tau\nabla_-(\operatorname{div}_+(\tilde{p}) + \operatorname{div}_-(\tilde{q}))\right),\end{cases}$$

we have

$$\begin{aligned}
\left|\begin{pmatrix}\tilde{p}-\bar{p}\\ \tilde{q}-\bar{q}\end{pmatrix}\right|^2 = \left|\begin{pmatrix}\hat{p}-\bar{p}\\ \hat{q}-\bar{q}\end{pmatrix}\right|^2 &\leq \left|(Id-\tau A)\begin{pmatrix}\tilde{p}-\bar{p}\\ \tilde{q}-\bar{q}\end{pmatrix}\right|^2 \\
&\leq \left|\begin{pmatrix}\tilde{p}-\bar{p}\\ \tilde{q}-\bar{q}\end{pmatrix}_{\ker(A)}\right|^2 + |1-\tau\|A\||^2\left|\begin{pmatrix}\tilde{p}-\bar{p}\\ \tilde{q}-\bar{q}\end{pmatrix}_F\right|^2 \\
&< \left|\begin{pmatrix}\tilde{p}-\bar{p}\\ \tilde{q}-\bar{q}\end{pmatrix}\right|^2 \text{ if } \begin{pmatrix}\tilde{p}-\bar{p}\\ \tilde{q}-\bar{q}\end{pmatrix} \notin \ker(A) \text{ and } 0 < \tau < 1/8.
\end{aligned}$$

But the latter inequality is nonsense. Thus, $\begin{pmatrix}\tilde{p}-\bar{p}\\ \tilde{q}-\bar{q}\end{pmatrix} \in \ker(A)$, or equivalently,

$$\operatorname{div}_+(\tilde{p}-\bar{p}) + \operatorname{div}_-(\tilde{q}-\bar{q}) \in \ker(\nabla_+) \cap \ker(\nabla_-)$$

Indeed, by definition of the divergence operators

$$\begin{aligned}|\operatorname{div}_+(\tilde p-\bar p)+\operatorname{div}_-(\tilde q-\bar q)|^2 &= -\langle \tilde p-\bar p, \nabla_+\left(\operatorname{div}_+(\tilde p-\bar p)+\operatorname{div}_-(\tilde q-\bar q)\right)\rangle_Y\\ &\quad -\langle \tilde q-\bar q, \nabla_-\left(\operatorname{div}_+(\tilde p-\bar p)+\operatorname{div}_-(\tilde q-\bar q)\right)\rangle_Y\\ &= -\langle \tilde p-\bar p, 0\rangle_Y - \langle \tilde q-\bar q, 0\rangle_Y = 0.\end{aligned}$$

Hence, $\operatorname{div}_+(\tilde p-\bar p)+\operatorname{div}_-(\tilde q-\bar q)=0$, or equivalently, $\operatorname{div}_+(\tilde p)+\operatorname{div}_-(\tilde q)=\operatorname{div}_+(\bar p)+\operatorname{div}_-(\bar q)$. We have shown that any convergent subsequence of $(\operatorname{div}_+(p_n)+\operatorname{div}_-(q_n))_{n\geq 0}$ converges to $\operatorname{div}_+(\bar p)+\operatorname{div}_-(\bar q)$, thus the sequence $(\operatorname{div}_+(p_n)+\operatorname{div}_-(q_n))_{n\geq 0}$ converges to $\operatorname{div}_+(\bar p)+\operatorname{div}_-(\bar q)$. □

Remark 4 One may obtain an alternating version of Algorithm 31 by modifying the second step as follows

$$\begin{cases} p_{n+1} = \pi_{B_Y}\left(p_n+\tau\nabla_+\left[\operatorname{div}_+(p_n)+\operatorname{div}_-(q_n)+2f/\lambda\right]\right),\\ q_{n+1} = \pi_{B_Y}\left(q_n+\tau\nabla_-\left[\operatorname{div}_+(p_{n+1})+\operatorname{div}_-(q_n)+2f/\lambda\right]\right).\end{cases} \tag{3.29}$$

While the proof of convergence of the alternating version of Algorithm 31 is still eluding us, the numerical experiments suggest that one should be able to prove a convergence result for $0<\tau<1/4$.

4 Numerical Approximation of the continuous ROF model

In this section, we study how to approximate the minimizer of the continuous ROF model

$$u_f = \operatorname*{argmin}_{u\in BV(\Omega)} E_\lambda^f(u), \tag{4.1}$$

where $E_\lambda^f(u)$ is defined in (2.3)

Our approach consists in extending the minimizer $z_{f,h}$ of the discrete ROF functional (2.5) associated with $Q_h f$ into a continuous piecewise linear polynomial $P_h z_{f,h}$ with respect to the triangulation Δ_h. We shall obtain an error bound in $\mathrm{L}^2(\Omega)$ between $P_h z_{f,h}$ and u_f.

To begin, we recall the following standard fact about the ROF functional (cf. [16,15]), and provide a proof for convenience.

Lemma 7 *Let $u_f\in BV(\Omega)$ be the minimizer of the functional $E_\lambda^f(u)$ defined by* (2.3). *Then, for any $v\in BV(\Omega)$, there holds*

$$\left\|v-u_f\right\|_2^2 \leq 2\lambda\left(E_\lambda^f(v)-E_\lambda^f(u_f)\right). \tag{4.2}$$

Moreover,

$$E_\lambda^f(u_f) = \frac{1}{2\lambda}(\|f\|_2^2-\|u_f\|_2^2). \tag{4.3}$$

Proof Since u_f is the minimizer of $E_\lambda^f(u)$ and E_λ^f is convex, it follows that $0\in\partial E_\lambda^f(u_f)$. Since, $\partial E_\lambda^f(u_f) = \partial|Du_f|(\Omega)+(u_f-f)/\lambda$, we infer that $(f-u_f)/\lambda\in\partial|Du_f|(\Omega)$. Thus, for any $v\in BV(\Omega)$, we have

$$|Dv|(\Omega)-\left|Du_f\right|(\Omega) \geq \langle\frac{f-u_f}{\lambda}, v-u_f\rangle. \tag{4.4}$$

As a consequence, we have

$$\begin{aligned} E_\lambda^f(v)-E_\lambda^f(u_f) &= |Dv|(\Omega)-\left|Du_f\right|(\Omega)+\frac{1}{2\lambda}\left(\|v-f\|_2^2-\|u_f-f\|_2^2\right)\\ &\geq \langle\frac{f-u_f}{\lambda}, v-u_f\rangle+\langle\frac{u_f-f}{\lambda}, v-u_f\rangle+\frac{1}{2\lambda}\|v-u_f\|_2^2\\ E_\lambda^f(v)-E_\lambda^f(u_f) &\geq \frac{1}{2\lambda}\|v-u_f\|_2^2.\end{aligned}$$

To show the equality (4.3), we choose v in (4.4) to be equal to 0 and $2u_f$, respectively. Using the fact that the total variation functional is positively 1-homogeneous, we infer that $|Du_f|(\Omega)=\langle\frac{f-u_f}{\lambda}, u_f\rangle$. Therefore

$$E_\lambda^f(u_f) = \langle\frac{f-u_f}{\lambda}, u_f\rangle+\langle\frac{f-u_f}{\lambda}, \frac{f-u_f}{2}\rangle = \frac{1}{2\lambda}(\|f\|_2^2-\|u_f\|_2^2).$$

□

The next result is an upper bound of the difference between the energy of the discrete minimizer and the energy of its piecewise linear interpolation.

Lemma 8 *Let $f \in \mathrm{L}^2(\Omega)$ be given, and $z_{f,h}$ the minimizer of the functional $E^f_{\lambda,h}(u)$. Then*

$$E^f_\lambda(P_h z_{f,h}) - E^f_{\lambda,h}(z_{f,h}) \le \frac{C}{2\lambda}\left[h\omega(z_{f,h},1) + \omega_2(f,h)\right], \tag{4.5}$$

where C is a positive constant depending on f and λ.

Proof Since $J_h(z_{f,h}) = |DP_h z_{f,h}|(\Omega)$, we have

$$\begin{aligned}2\lambda\left(E^f_\lambda(P_h z_{f,h}) - E^f_{\lambda,h}(z_{f,h})\right) &= \|P_h z_{f,h} - f\|_2^2 - \|C_h z_{f,h} - C_h Q_h f\|_2^2\\ &\le \|P_h z_{f,h} - C_h z_{f,h}\|\left(\|P_h z_{f,h} - C_h z_{f,h}\|_2 + 2\|C_h z_{f,h} - f\|_2\right)\\ &\quad + \|C_h z_{f,h} - f\|^2 - \|C_h z_{f,h} - C_h Q_h f\|_2^2\end{aligned}$$

Now, let us find an upper bound for the difference $\|C_h z_{f,h} - f\|_2^2 - \|C_h z_{f,h} - C_h Q_h f\|_2^2$.

$$\begin{aligned}\|C_h z_{f,h} - f\|_2^2 - \|C_h z_{f,h} - C_h Q_h f\|_2^2 &= \int_\Omega \left(|C_h z_{f,h}(x) - f(x)|^2 - |C_h z_{f,h}(x) - C_h Q_h f(x)|^2\right)dx\\ &= \int_\Omega (C_h Q_h f(x) - f(x))\left(2C_h z_{f,h}(x) - C_h Q_h f(x) - f(x)\right)dx\\ &\le \|C_h Q_h f - f\|_2 \|2C_h z_{f,h} - C_h Q_h f - f\|_2.\end{aligned}$$

The inequality (4.5) follows from Lemma 2 together with (2.17), and Lemma 3 by observing that

$$\|2C_h z_{f,h} - C_h Q_h f - f\|_2 \le 4\|f\|_2 \text{ and } \|C_h z_{f,h} - f\|_2 \le 2\|f\|_2.$$

□

Let $\Omega_1 := \{x \in \mathbb{R}^2 : \mathrm{dist}(x,\Omega) < 1\}$. We denote by $\bar u_f$ an extension (see [2, Proposition 3.21 page 131]) of u_f to a function in $BV(\mathbb{R}^2)$ such that $\bar u_f = 0$ in $\mathbb{R}^2 \setminus \Omega_1$, and $|D\bar u_f|(\partial\Omega) = 0$. Let

$$\rho(x) = c\exp\left(\frac{1}{|x|^2-1}\right)\mathbb{1}_{\{|x|\le 1\}}(x), \quad x \in \mathbb{R}^2$$

be the standard convolution kernel, and $\{\rho_\epsilon\}_{\epsilon>0}$ the ensuing family of mollifiers. The constant $c > 0$ in the definition of ρ above is chosen such that $\int \rho(x)dx = 1$. For $\epsilon > 0$ fixed, we denote by $\bar u^\epsilon_f := \rho_\epsilon * \bar u_f$ the ϵ-mollification of $\bar u_f$. Then, it is known that (see [2, 10])

$$|D\bar u^\epsilon_f|(\Omega) \le |D\bar u_f|(\bar\Omega) = |D\bar u_f|(\Omega) = |Du_f|(\Omega). \tag{4.6}$$

Lemma 9 *Suppose $f \in \mathrm{L}^2(\Omega)$. Let u_f be the minimizer of the ROF-functional $E^f_\lambda(u)$, and $z_{f,h}$ the minimizer of the discrete ROF-functional $E^f_{\lambda,h}(u)$. Then, for any $0 < \alpha < 1$ and for $h \ll 1$*

$$E^f_{\lambda,h}(z_{f,h}) \le E^f_\lambda(u_f) + K_1 h^{(1-\alpha)/2} + K_2\|\bar u^h_f - \bar u_f\|_2 \tag{4.7}$$

where $\bar u^h_f$ is the $h^{\alpha/4}$–mollification of $\bar u_f$, K_1 and K_2 are positive constants that depend only on λ, Ω, and f.

Proof Let $\epsilon > 0$ be fixed. By definition of $z_{f,h}$, we have

$$E^f_{\lambda,h}(z_{f,h}) \le E^f_{\lambda,h}(Q_h\bar u^\epsilon_f) = J_h(Q_h\bar u^\epsilon_f) + \frac{1}{2\lambda}\|C_h Q_h \bar u^\epsilon_f - C_h Q_h f\|_2^2. \tag{4.8}$$

Moreover by definition of the operators C_h and Q_h, it is easy to establish that

$$\begin{aligned}\|C_h Q_h\bar u^\epsilon_f - C_h Q_h f\|_2^2 &\le \|\bar u^\epsilon_f - f\|_2^2\\ &\le \|u_f - f\|_2^2 + \|\bar u^\epsilon_f - \bar u_f\|_2\left(2\|u_f - f\|_2 + \|\bar u^\epsilon_f - \bar u_f\|_2\right)\\ &\le \|u_f - f\|_2^2 + 6\|f\|_2\|\bar u^\epsilon_f - \bar u_f\|_2 \text{ using (4.3).}\end{aligned} \tag{4.9}$$

Next, we obtain an upper bound on $J_h(Q_h\bar u^\epsilon_f)$. By definition

$$\begin{aligned}J_h(Q_h\bar u^\epsilon_f) = &\sum_{1\le i,j<N}\frac{h^2}{2}\sqrt{\left|\frac{(Q_h\bar u^\epsilon_f)_{i+1,j} - (Q_h\bar u^\epsilon_f)_{i,j}}{h}\right|^2 + \left|\frac{(Q_h\bar u^\epsilon_f)_{i,j+1} - (Q_h\bar u^\epsilon_f)_{i,j}}{h}\right|^2}\\ &+ \sum_{1\le i,j<N}\frac{h^2}{2}\sqrt{\left|\frac{(Q_h\bar u^\epsilon_f)_{i+1,j+1} - (Q_h\bar u^\epsilon_f)_{i,j+1}}{h}\right|^2 + \left|\frac{(Q_h\bar u^\epsilon_f)_{i+1,j+1} - (Q_h\bar u^\epsilon_f)_{i+1,j}}{h}\right|^2}.\end{aligned} \tag{4.10}$$

Now, let $1 \leq i, j < N$ be fixed. Then, by the Taylor formula, we have

$$\frac{(Q_h \bar{u}_f^\epsilon)_{i+1,j} - (Q_h \bar{u}_f^\epsilon)_{i,j}}{h} = \frac{1}{h^2} \int_{\Omega_{i,j}} \frac{\bar{u}_f^\epsilon(x_1 + h, x_2) - \bar{u}_f^\epsilon(x_1, x_2)}{h} dx$$
$$= \frac{1}{h^2} \int_{\Omega_{i,j}} \left(\frac{\partial \bar{u}_f^\epsilon}{\partial x_1}(x) + \frac{h}{2} \int_0^1 \frac{\partial^2 \bar{u}_f^\epsilon}{\partial x_1^2}(x_1 + th, x_2) dt \right) dx.$$

Thanks to Jensen's inequality, we infer from the latter equation that

$$\left| \frac{(Q_h \bar{u}_f^\epsilon)_{i+1,j} - (Q_h \bar{u}_f^\epsilon)_{i,j}}{h} \right|^2 \leq \frac{1}{h^2} \int_{\Omega_{i,j}} \left| \frac{\partial \bar{u}_f^\epsilon}{\partial x_1}(x) \right|^2 dx + B_{ij}^1(h), \tag{4.11}$$

where $B_{i,j}^1(h)$ is given by

$$B_{i,j}^1(h) = \frac{1}{h} \int_{\Omega_{i,j}} \int_0^1 \left| \frac{\partial \bar{u}_f^\epsilon}{\partial x_1}(x) \frac{\partial^2 \bar{u}_f^\epsilon}{\partial x_1^2}(x_1 + th, x_2) \right| dt dx + \frac{1}{4} \int_{\Omega_{i,j}} \int_0^1 \left| \frac{\partial^2 \bar{u}_f^\epsilon}{\partial x_1^2}(x_1 + th, x_2) \right|^2 dt dx. \tag{4.12}$$

Exchanging the order of integration and using Cauchy-Schwarz inequality, we obtain an upper bound on $B_{i,j}^1(h)$ as follows:

$$B_{i,j}^1(h) \leq \frac{1}{h} \left(\int_{\Omega_{i,j}} \left| \frac{\partial \bar{u}_f^\epsilon}{\partial x_1}(x) \right|^2 dx + \frac{2+h}{4} \int_{\Omega_{i,j} \cup \Omega_{i+1,j}} \left| \frac{\partial^2 \bar{u}_f^\epsilon}{\partial x_1^2}(x) \right|^2 dx \right). \tag{4.13}$$

Similarly, we can show that

$$\left| \frac{(Q_h \bar{u}_f^\epsilon)_{i,j+1} - (Q_h \bar{u}_f^\epsilon)_{i,j}}{h} \right|^2 \leq \frac{1}{h^2} \int_{\Omega_{i,j}} \left| \frac{\partial \bar{u}_f^\epsilon}{\partial x_2}(x) \right|^2 dx + B_{ij}^2(h), \tag{4.14}$$

with $B_{ij}^2(h)$ defined by

$$B_{i,j}^2(h) = \frac{1}{h} \int_{\Omega_{i,j}} \int_0^1 \left| \frac{\partial \bar{u}_f^\epsilon}{\partial x_2}(x) \frac{\partial^2 \bar{u}_f^\epsilon}{\partial x_2^2}(x_1, x_2 + th) \right| dt dx + \frac{1}{4} \int_{\Omega_{i,j}} \int_0^1 \left| \frac{\partial^2 \bar{u}_f^\epsilon}{\partial x_2^2}(x_1, x_2 + th) \right|^2 dt dx, \tag{4.15}$$

and satisfying the inequality

$$B_{i,j}^2(h) \leq \frac{1}{h} \left(\int_{\Omega_{i,j}} \left| \frac{\partial \bar{u}_f^\epsilon}{\partial x_2} \right|^2 dx + \frac{2+h}{4} \int_{\Omega_{i,j} \cup \Omega_{i,j+1}} \left| \frac{\partial^2 \bar{u}_f^\epsilon}{\partial x_2^2} \right|^2 dx \right). \tag{4.16}$$

Combining (4.11) and (4.14) in equation (4.10), and using Jensen's inequality, we obtain for $h \ll 1$

$$J_h(Q_h \bar{u}_f^\epsilon) \leq |D\bar{u}_f^\epsilon|(\Omega) + h^{1/2} \left(|D\bar{u}_f^\epsilon|(\Omega) + 2 \int_\Omega \sqrt{\left| \frac{\partial^2 \bar{u}_f^\epsilon}{\partial x_1^2} \right|^2 + \left| \frac{\partial^2 \bar{u}_f^\epsilon}{\partial x_2^2} \right|^2} dx \right)$$
$$\leq |Du_f|(\Omega) + h^{1/2} \left(|Du_f|(\Omega) + 2 \int_\Omega \left| \frac{\partial^2 \bar{u}_f^\epsilon}{\partial x_1^2} \right| + \left| \frac{\partial^2 \bar{u}_f^\epsilon}{\partial x_2^2} \right| dx \right) \text{ by (4.6).} \tag{4.17}$$

Now, using the isometric embedding of $\mathrm{L}^1(\Omega)$ into the space of Radon measure, we have for $i = 1, 2$

$$\int_\Omega \left| \frac{\partial^2 \bar{u}_f^\epsilon}{\partial x_i^2} \right| dx = \sup_{\substack{\phi \in \mathcal{C}(\bar{\Omega}) \\ \|\phi\|_\infty \leq 1}} \int_\Omega \frac{\partial^2 \bar{u}_f^\epsilon}{\partial x_i^2} \phi \, dx = \sup_{\phi \in \mathcal{C}(\bar{\Omega}) \|\phi\|_\infty \leq 1} \int_\Omega \bar{u}_f \left(\frac{\partial^2 \rho_\epsilon}{\partial x_i^2} * \phi \right) dx$$
$$\leq \sup_{\substack{\phi \in \mathcal{C}(\bar{\Omega}) \\ \|\phi\|_\infty \leq 1}} \|u_f\|_2 \left\| \frac{\partial^2 \rho_\epsilon}{\partial x_i^2} * \phi \right\|_2 \leq \frac{K}{\epsilon^2}, \quad \text{with } K = \|u_f\|_2 \left\| \frac{\partial^2 \rho}{\partial x_i^2} \right\|_2. \tag{4.18}$$

Substituting the latter inequality into (4.17) gives

$$J_h(Q_h \bar{u}_f^\epsilon) \leq (1 + \sqrt{h}) |Du_f|(\Omega) + K \frac{\sqrt{h}}{\epsilon^2}. \tag{4.19}$$

Finally, letting $\epsilon = h^{\alpha/4}$ with $\alpha < 1$ and combining (4.19) and (4.9), we get

$$E_{\lambda,h}^f(z_{f,h}) \leq E_\lambda^f(u_f) + K_1 h^{(1-\alpha)/2} + K_2 \|\bar{u}_f^h - \bar{u}_f\|_2,$$

where $\bar{u}_f^h$ is the $h^{\alpha/4}-$mollification of $\bar{u}_f$, and the K_i's are positive constants depending only on Ω, λ, and f. □

We are now ready to establish an error bound for the L^2-norm of $P_h z_{f,h} - u_f$.

Theorem 2 *Let u_f be the minimizer of the ROF functional $E_\lambda^f(u)$ in $BV(\Omega)$, and $z_{f,h}$ the minimizer of the discrete ROF functional $E_{\lambda,h}^f(u)$. Then, for any $0<\alpha<1$ and $0<h\ll 1$*

$$\|P_h z_{f,h} - u_f\|_2^2 \le K_3\,\omega_2(f,h) + K_4 h^{\frac{1-\alpha}{2}} + K_5\|\bar u_f^h - \bar u_f\|_2, \tag{4.20}$$

where $K_3,\ K_4$, and K_5 are positive constants independent of h.

Proof Using (4.2), we have

$$\begin{aligned}\left\|P_h z_{f,h} - u_f\right\|_2^2 &\le 2\lambda\left(E_\lambda^f(P_h z_{f,h}) - E_\lambda^f(u_f)\right)\\ &= 2\lambda\left(E_\lambda^f(P_h z_{f,h}) - E_{\lambda,h}^f(z_{f,h}) + E_{\lambda,h}^f(z_{f,h}) - E_\lambda^f(u_f)\right)\end{aligned}$$

Now, thanks to the Lemmas 8 and 9, i.e., (4.5) and (4.7), we have that

$$\left\|P_h z_{f,h} - u_f\right\|_2^2 \le C(h\omega(z_{f,h},1) + \omega_2(f,h)) + 2\lambda(K_1 h^{(1-\alpha)/2} + K_2\|\bar u_f^h - \bar u_f\|_2). \tag{4.21}$$

Next, we notice that by Lemmas 4 and 5, we have

$$\begin{aligned}\omega(z_{f,h},1)^2 &= \sup_{\substack{\alpha=(\alpha_1,\alpha_2)\\ |\alpha|_\infty=1}} \sum_{\substack{1\le i+\alpha_1\le N\\ 1\le j+\alpha_2\le N}} |(z_{f,h})_{i+\alpha_1,j+\alpha_2} - (z_{f,h})_{i,j}|^2\\ &\le \sup_{\substack{\alpha=(\alpha_1,\alpha_2)\\ |\alpha|_\infty=1}} |z_{f,h,\alpha}^{2N} - z_{f,h}^{2N}|^2\\ &\le \sup_{\substack{\alpha=(\alpha_1,\alpha_2)\\ |\alpha|_\infty=1}} |(Q_h f)_\alpha^{2N} - (Q_h f)^{2N}|^2, \text{ by (3.8)}\\ &\le 2 \sup_{\substack{\alpha=(\alpha_1,\alpha_2)\\ |\alpha_1|+|\alpha_2|=1}} |(Q_h f)_\alpha^{2N} - (Q_h f)^{2N}|^2\\ \omega(z_{f,h},1)^2 &\le 8 \sup_{\substack{\alpha=(\alpha_1,\alpha_2)\\ |\alpha_1|+|\alpha_2|=1}} \sum_{\substack{1\le i+\alpha_1\le N\\ 1\le j+\alpha_2\le N}} |(Q_h f)_{i+\alpha_1,j+\alpha_2} - (Q_h f)_{i,j}|^2.\end{aligned}$$

From the latter inequality, the definition of $Q_h f$, and Jensen's inequality, we obtain

$$\begin{aligned}\omega(z_{f,h},1)^2 &\le 8 \sup_{\substack{\alpha=(\alpha_1,\alpha_2)\\ |\alpha_1|+|\alpha_2|=1}} \sum_{\substack{1\le i+\alpha_1\le N\\ 1\le j+\alpha_2\le N}} |(Q_h f)_{i+\alpha_1,j+\alpha_2} - (Q_h f)_{i,j}|^2\\ &\le 8 \sup_{\substack{\alpha=(\alpha_1,\alpha_2)\\ |\alpha_1|+|\alpha_2|=1}} \sum_{\substack{1\le i+\alpha_1\le N\\ 1\le j+\alpha_2\le N}} \frac{1}{h^2}\int_{\Omega_{i,j}} |f(x+\alpha h) - f(x)|^2 dx\\ &\le \frac{8}{h^2} \sup_{\substack{\alpha=(\alpha_1,\alpha_2)\\ |\alpha_1|+|\alpha_2|=1}} \int_{\Omega_\alpha} |f(x+\alpha h) - f(x)|^2 dx \le \frac{8}{h^2}\,\omega_2(f,h)^2.\end{aligned}$$

Thus, we have

$$h\omega(z_{f,h},1) \le 4\,\omega_2(f,h). \tag{4.22}$$

Finally, we combine the inequalities (4.21) and (4.22), to obtain inequality (4.20) with $K_3 = 5C,\ K_4 = 2\lambda K_1$, and $K_5 = 2\lambda K_2$. □

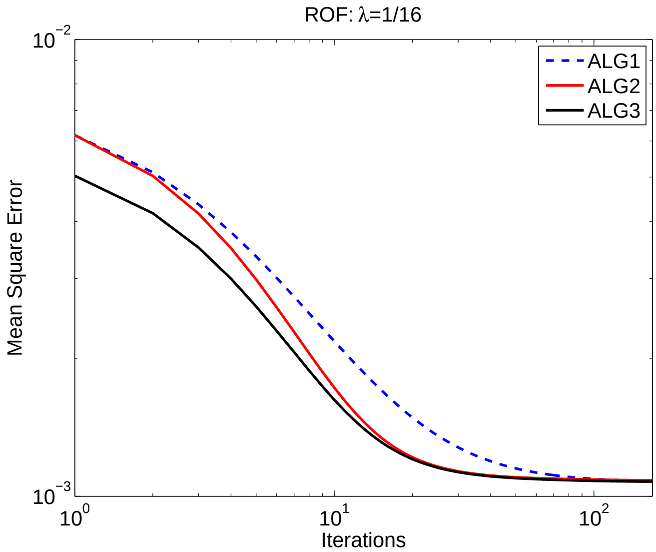
Corollary 1 *Let $0<\beta\le 1$ be given. If $f\in \mathrm{Lip}(\beta, \mathrm{L}^2(\Omega))$, then $\|P_h z f, h - u_f\|_2 \to 0$ as $h\to 0$.*

Proof Suppose that $f\in \mathrm{Lip}(\beta, \mathrm{L}^2(\Omega))$. There exists $M_\beta>0$ such that for any $0<h\ll 1$, $h^{-\beta}\omega_2(f,h)\le M_\beta$. As a consequence, (4.20) becomes with $\alpha=\beta/2$

$$\|P_h z_{f,h} - u_f\|_2^2 \le K_3 M_\beta h^\beta + K_4 h^{(2-\beta)/4} + K_5\|\bar u_f^h - \bar u_f\|_2.$$

Since $\|\bar u_f^h - \bar u_f\|_2 \to 0$ as $h\to 0$, we infer that $\|P_h z_{f,h} - u_f\|_2 \to 0$ as $h\to 0$, and the proof is complete. □

Fig. 2: Convergence of ALG1, ALG2 and ALG3 for the image in Figure 3a with $\sigma = 25$ and $\lambda = 1/13$. In algorithms ALG1 and ALG2 we set $\tau = 0.025$, while $\tau = 0.0125$ in ALG3. In this case we used the value of λ that gave the best PSNRs for our choices of τ.



5 Numerical experiments

In this section, we report the result of the numerical experiments with Algorithm 31. We compare the performance of the algorithm proposed above to Chambolle's fixed-point algorithm proposed in [4], and the projected-gradient algorithm proposed in [5]. The test images used are presented in Figure 3. We shall use the following abbreviations to identify the three algorithms under consideration here.

ALG1: The fixed point iterative algorithm described in [4].
ALG2: The projected-gradient algorithm described in [5,8].
ALG3: The projected-gradient algorithm presented in Algorithm 31.

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.

Table 1 through Table 4 below show the capability of Algorithm 31 to remove noise for various noise level. The inputs for all three algorithms are obtained by adding a zero mean Gaussian noise with standard deviation σ to the images in Figure 3. Our experiments show that the new projected gradient algorithm, Algorithm 31, is slightly more efficient than ALG2. Finally in Figure 2, we show the asymptotic behavior of the three algorithms: ALG1, ALG2, and ALG3. All of the three algorithms display the same asymptotic behavior as the number of iterations goes to infinity. Of course, this agreement of asymptotic behaviors is expected since all three algorithms – in the discrete setting – were proved to converge to the solution of the ROF model.

References

1. R. Acar and C.R. Vogel: Analysis of bounded variation penalty methods for ill-posed problems, *Inverse Problems* **10** (1994), 1217–1229.
2. L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford University Press, USA (2000).
3. T. F. Chan, G. H. Golub, and P. Mulet, A nonlinear primal-dual method for total variation-based image restoration, *SIAM J. Sci. Comput.* **20** (1999), no. 6, 1964–1977.
4. A. Chambolle, An algorithm for total variation minimization and applications. Special issue on mathematics and image analysis, *J. Math. Imaging Vision* **20** (2004), no. 1-2, pp. 89 – 97.
5. A. Chambolle, Total variation minimization and a class of binary MRF models, in Energy minimization methods in computer vision and pattern recognition, *Lecture Notes in Computer Science*, Vol. 3757/2005 (2005), pp. 136–152.
6. A. Chambolle and P.-L. Lions, Image recovery via total variation minimization and related problems, *Numer. Math.* **76** (1997), no. 2, 167–188.
7. D. C. Dobson and C. R. Vogel, Convergence of an iterative method for total variation denoising, *SIAM J. Numer. Anal.* **34** (1997), 1779–1791.
8. 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), 232–248.
9. I. Ekeland and R. Temam, *Convex analysis and variational problems*, Society for Industrial Mathematics, Philadelphia, PA, USA (1987).
10. E. Giusti, *Minimal surfaces and functions of bounded variation*, Monographs in Mathematics, 80, Birkhäuser, Basel (1984).

Fig. 3: The images used in the numerical experiments. The images in the top row are of size 256×256, while those in the bottom row have resolution 512×512.

(a) Lena (b) Peppers

(c) Bank. (d) Boats

11. Q. Hong, M. J. Lai, and J. Wang, Convergence Analysis of a Finite Difference Scheme for the Gradient Flow associated with the ROF Model, submitted (2010).
12. M. -J. Lai, B. Lucier, and J. Wang, The convergence of a central-difference discretization of Rudin-Osher-Fatemi model for image denoising, in Scale Space and Variational Methods in Computer Vision, *Lecture Notes in Computer Science*, Vol. 5567/2009 (2009), 514-526.
13. L.I. Rudin, S. Osher, and E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D: Nonlinear Phenomena* **60** (1992), no. 1-4, 259–268.
14. C. R. Vogel and M. E. Oman, Iterative methods for total variation denoising, *SIAM J. Sci. Comput.* **17** (1996), no. 1, 227–238.
15. J. Wang and Lucier, B., Error bounds for finite-difference methods for Rudin-Osher-Fatemi image smoothing, to appear in SIAM J. Numer. Analysis (2010).
16. J. Wang, Error bound for numerical methods for the ROF image smoothing model, PhD Dissertation, Purdue University, IN, USA (2008).

Table 1: Comparison of the algorithms using the image in Figure 3a. Each algorithm is terminated when the difference between the Mean Square Error (MSE) at two consecutive steps is less than 10^{-8}.

σ	λ	τ	ALG1	ALG2	τ	ALG3
15	1/24	0.250	31.1792	31.1607	0.125	31.1947
20	1/16		29.6780	29.6486		29.6798
15	1/24	0.200	31.1790	31.1620	0.100	31.1953
20	1/16		29.6778	29.6502		29.6812
15	1/24	0.150	31.1789	31.1621	0.075	31.1953
20	1/16		29.6777	29.6504		29.6815
15	1/24	0.100	31.1786	31.1622	0.500	31.1954
20	1/16		29.6775	29.6507		29.6817
15	1/24	0.050	31.1782	31.1621	0.025	31.1952
20	1/16		29.6772	29.6507		29.6819

Table 2: Comparison of the Algorithms using the image in Figure 3b. Each algorithm is terminated when the difference between the Mean Square Error (MSE) at two consecutive steps is less than 10^{-8}.

σ	λ	τ	ALG1	ALG2	τ	ALG3
15	1/24	0.250	31.8309	31.8136	0.125	31.8461
20	1/16		30.4052	30.3592		30.3937
15	1/24	0.200	31.8305	31.8456	0.100	31.8695
20	1/16		30.4049	30.3952		30.4219
15	1/24	0.150	31.8300	31.8430	0.075	31.8676
20	1/16		30.4043	30.3960		30.4219
15	1/24	0.100	31.8294	31.8424	0.500	31.8666
20	1/16		30.4035	30.3966		30.4223
15	1/24	0.050	31.8287	31.8414	0.025	31.8655
20	1/16		30.4027	30.3971		30.4225

Table 3: Comparison of the Algorithms using the image in Figure 3c. Each algorithm is terminated when the difference between the Mean Square Error (MSE) at two consecutive steps is less than 10^{-8}.

σ	λ	τ	ALG1	ALG2	τ	ALG3
15	1/24	0.250	31.3389	31.3582	0.125	31.4086
20	1/16		29.8443	29.8624		29.8998
15	1/24	0.200	31.3222	31.3561	0.100	31.4052
20	1/16		29.8332	29.8615		29.8980
15	1/24	0.150	31.3076	31.3445	0.075	31.3929
20	1/16		29.8126	29.8484		29.8846
15	1/24	0.100	31.2741	31.3234	0.050	31.3708
20	1/16		29.7792	29.8224		29.8620
15	1/24	0.050	31.1943	31.2709	0.025	31.3161
20	1/16		29.7071	29.7708		29.8089

Table 4: Comparison of the Algorithms using the image in Figure 3d. Each algorithm is terminated when the difference between the Mean Square Error (MSE) at two consecutive steps is less than 10^{-8}.

σ	λ	τ	ALG1	ALG2	τ	ALG3
15	1/24		30.5134	30.3148		30.3527
20	1/16	0.250	29.1672	28.9448	0.125	28.9745
15	1/24		30.5132	30.4803		30.5157
20	1/16	0.200	29.1663	29.1092	0.100	29.1317
15	1/24		30.5128	30.4795		30.5157
20	1/16	0.150	29.1657	29.1088	0.075	29.1317
15	1/24		30.5125	30.4795		30.5154
20	1/16	0.100	29.1650	29.1084	0.050	29.1317
15	1/24		30.5121	30.4791		30.5148
20	1/16	0.050	29.1642	29.1078	0.025	29.1312