Continuous Primal-Dual methods for Image Processing

M. Goldman[*]

July 30, 2011

Abstract

In this article we study a continuous Primal-Dual method proposed by Appleton and Talbot and generalize it to other problems in image processing. We interpret it as an Arrow-Hurwicz method which leads to a better description of the system of PDEs obtained. We show existence and uniqueness of solutions and get a convergence result for the denoising problem. Our analysis also yields new *a posteriori* estimates.

1 Introduction

In imaging, duality has been recognized as a fundamental ingredient for designing numerical schemes solving variational problems involving a total variation term. Primal-Dual methods were introduced in the field by Chan, Golub and Mulet in [12]. Afterwards, Chan and Zhu [20] proposed to rewrite the discrete minimization problem as a min-max and solve it using an Arrow-Hurwicz [5] algorithm which is a gradient ascent in one direction and a gradient descent in the other. Just as for the simple gradient descent, one can

---

[*]CMAP, CNRS UMR 7641, Ecole Polytechnique, 91128 Palaiseau Cedex, France. e-mail: goldman@cmap.polytechnique.fr

think of extending this method to the continuous framework. This is in fact what does the algorithm previously proposed by Appleton and Talbot in [4] derived by analogy with discrete graph cuts techniques. The first to notice the link between their method and Primal-Dual schemes were Chambolle et al. in [9].

Besides its intrinsic theoretical interest, considering the continuous framework has also pratical motivations. Indeed, as illustrated by Appleton and Talbot in [4], this approach leads to higher quality results compared with fully discrete schemes such as those proposed by Chan and Zhu. We will numerically illustrate this in the final part of this paper.

This paper proposes to study the continuous Primal-Dual algorithm following the philosophy of the work done for the gradient flow by Caselles and its collaborators (see the book of Andreu et al. [2] and the references therein). We give a rigorous definition of the system of PDEs which is obtained and show existence and uniqueness of a solution to the Cauchy problem. We prove strong L^2 convergence to the minimizer for the Rudin-Osher-Fatemi model and derive some *a posteriori* estimates. As a byproduct of our analysis we also obtain *a posteriori* estimates for the numerical scheme proposed by Chan and Zhu.

1.1 Presentation of the problem

Many problems in image processing can be seen as minimizing in $BV \cap L^2$ an energy of the form

$$J(u) = \int_\Omega |Du| + G(u) + \int_{\partial\Omega_D} |u - \varphi|. \tag{1}$$

The notation $\int_\Omega |Du|$ stands for the total variation of the function u and is rigourously defined in Definition 3.1. We assume that Ω is a bounded Lipschitz open set of $\mathbb{R}^d$ (in applications for image processing, usually $d = 2$ or $d = 3$) and that $\partial\Omega_D$ is a subset of $\partial\Omega$. The function φ being given in $L^1(\partial\Omega_D)$, the term $\int_{\partial\Omega_D} |u - \varphi|$ is a Dirichlet condition on $\partial\Omega_D$. We call $\partial\Omega_N$ the complement of $\partial\Omega_D$ in $\partial\Omega$ and assume that G is convex and continuous

2

in L^2 with
$$G(u) \leq C(1 + |u|_2^p) \qquad \text{with } 1 \leq p \leq +\infty.$$

In this paper we note $|u|_2$ the L^2 norm of u. According to Giaquinta et al. [16] we have,

Proposition 1.1. *The functional J is convex and lower-semi-continuous (lsc) in L^2.*

In the following, we also assume that J attains its minimum in $BV \cap L^2$. This is for example true if G satisfies some coercivity hypothesis or if G is non negative.

Two fundamental applications of our method are image denoising via total variation regularization and segmentation with geodesic active contours.

In the first problem, one starts with a corrupted image $f = \bar{u} + n$ and wants to find the clean image $\bar{u}$. Rudin, Osher and Fatemi proposed to look for an approximation of $\bar{u}$ by minimizing
$$\int_\Omega |Du| + \frac{\lambda}{2} \int_\Omega (u - f)^2.$$

This corresponds to $G(u) = \frac{\lambda}{2} \int_\Omega (u - f)^2$ and $\partial\Omega_D = \emptyset$ in (1). For a comprehensive introduction to this subject, we refer to the lecture notes of Chambolle et al. [8]. Figure 1 shows the result of denoising using the algorithm of Chan and Zhu.

Figure 1: Denoising using the ROF model

The issue in the second problem is to extract automatically the boundaries of an object within an image. We suppose that we are given two subsets S and T of $\partial\Omega$ such that S lies inside the object that we want to segment and T lies outside. Caselles et al. proposed in [7] to associate a positive function g to the image in a way that g is high where the gradient of the image is low and vice versa. The object is then segmented by minimizing

$$\min_{E \supset S,\, E^c \supset T} \int_{\partial E} g(s)ds. \tag{2}$$

In order to simplify the notations, we will only deal with $g = 1$ in the following. It is however straightforward to extend our discussion to general (continuous) g. The energy we want to minimize is thus $\int_\Omega |D\chi_E|$. This functional is non convex but by the coarea formula (see Ambrosio-Fusco-Pallara [1]), it can be relaxed to functions $u \in [0, 1]$.

Let $\varphi = 1$ on S and $\varphi = 0$ on T. Letting $\partial\Omega_D = S \cup T$, and f be a L^2 function, our problem can be seen as a special case of the prescribed mean curvature problem (in our original segmentation problem, $f = 0$),

$$\inf_{\substack{0 \leq u \leq 1 \\ u = \varphi \text{ in } \partial\Omega_D}} \int_\Omega |Du| + \int_\Omega fu. \tag{3}$$

If u is a solution of (3), a minimizer E of (2) is then given by any superlevel of u, namely $E = \{u > s\}$ for any $s \in]0, 1[$. This convexification argument is somewhat classical but more details can be found in the lecture notes [8] Section 3.2.2.

It is however well known that in general the infimum is not attained because of the lack of compactness for the boundary conditions in BV. Following the ideas of Giaquinta et al. [16] we have to relax the boundary conditions by adding a Dirichlet term $\int_{\partial\Omega_D} |u - \varphi|$ to the functional. We also have to deal with the hard constraint, $0 \leq u \leq 1$. This last issue will be discussed afterwards but it brings some mathematical difficulties that we were not able to solve. Fortunately, our problem is equivalent (see [9]) to the minimization of the unsconstrained problem

$$J(u) = \inf_{u \in BV(\Omega)} \int_\Omega |Du| + \int_{\partial\Omega_D} |u - \varphi| + \int_\Omega f^+|u| + \int_\Omega f^-|1 - u|.$$

Here $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$.

We give in Figure 2 the result of this segmentation on yeasts. The small

square is the set S and the set T is taken to be the image boundary. The study of this problem was in fact our first motivation for this work.
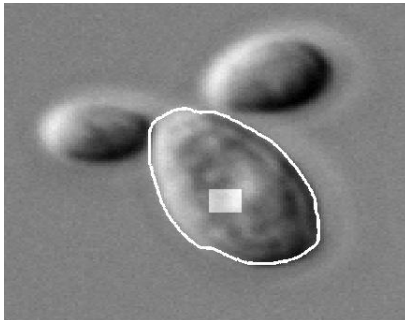

Figure 2: Yeast segmentation

1.2 Idea of the Primal-Dual method

Formally, the idea behind the Primal-Dual method is using the definition of $\int_\Omega |Du|$ (see Definition 3.1) in order to write J as

$$J(u) = \sup_{\substack{\xi \in \mathcal{C}_c^1(\Omega) \\ |\xi|_\infty \leq 1}} K(u, \xi).$$

Where $K(u, \xi) = -\int_\Omega u \operatorname{div}(\xi) + \int_{\partial\Omega_D} |u - \varphi| + G(u)$. Then, finding a minimum of J is equivalent to finding a saddle point of K. This is done by a gradient descent in u and a gradient ascent in ξ.

Let $I_{B(0,1)}(\xi)$ be the indicator function of the unit ball in L^∞ (it takes the value 0 if $|\xi|_\infty \leq 1$ and $+\infty$ otherwise) and ∂ denotes the subdifferential (see Ekeland-Temam [14] for the definition). As

$$K(u, \xi) = -\int_\Omega u \operatorname{div}(\xi) + \int_{\partial\Omega_D} |u - \varphi| + G(u) - I_{B(0,1)}(\xi).$$

we have $\nabla_u K \simeq -\operatorname{div} \xi + \partial G(u)$ and $\nabla_\xi K \simeq Du - \partial I_{B(0,1)}(\xi)$. We are thus led to solve the system of PDEs:

$$\begin{cases} \partial_t u = \operatorname{div}(\xi) - \partial G(u) \\[2mm] \partial_t \xi = Du - \partial I_{B(0,1)}(\xi) \\[2mm] + \text{ boundary conditions.} \end{cases} \tag{4}$$

5

This system is almost the one proposed by Appleton and Talbot in [4] for the segmentation problem.

Let us remark that, at least formally, the differential operator
$A(u,\xi) = \begin{pmatrix} -\operatorname{div}\xi + \partial G(u) \\ -Du + \partial I_{B(0,1)}(\xi) \end{pmatrix}$ verifies by Green's formula and the monotonicity of the subdifferential (see Proposition 2.3),

$$\langle A(u,\xi),(u,\xi)\rangle = \langle \partial G(u), u\rangle + \langle \partial I_{B(0,1)}(\xi),\xi\rangle \geq 0$$

which means that A is monotone (see Definition 2.2).
In the next section we recall some facts about the theory of maximal monotone operators and its applications for finding saddle points. In the last section we use it to give a rigourous meaning to the hyperbolic system (4) together with existence and uniqueness of solutions of the Cauchy problem.

2   Maximal Monotone Operators

Following Brézis [6], we present briefly in the first part of this section the theory of maximal monotone operators. In the second part we show how this theory sheds light on the general Arrow-Hurwicz method. We mainly give results found in Rockafellar's paper [19].

2.1   Definitions and first properties of maximal monotone operators

Definition 2.1. *Let X be an Hilbert space. An operator is a multivaluated mapping A from X into $\mathcal{P}(X)$. We call $D(A) = \{x \in X \,/\, A(x) \neq \varnothing\}$ the domain of A and $R(A) = \bigcup_{x \in X} A(x)$ its range. We identify A and its graph in $X \times X$.*

Definition 2.2. *An operator A is monotone if :*

$$\forall x_1, x_2 \in D(A), \qquad \langle A(x_1) - A(x_2), x_1 - x_2\rangle \geq 0$$

or more precisely if for all $x_1^* \in A(x_1)$ and $x_2^* \in A(x_2)$,

$$\langle x_1^* - x_2^*, x_1 - x_2\rangle \geq 0.$$

6

It is maximal monotone if it is maximal in the set of monotone operators. The maximality is to be understood in the sense of graph inclusion.

One of the essential results for us is the maximal monotonicity of the sub-gradient for convex functions.

Proposition 2.3. *[6] Let φ be a proper lower-semi-continuous convex function on X then $\partial\varphi$ is a maximal monotone operator.*

Before stating the main theorem of this theory, namely the existence of solutions of the Cauchy problem $-u' \in A(u(t))$ we need one last definition.

Definition 2.4. *Let A be maximal monotone. For $x \in D(A)$ we call $A^\circ(x)$ the projection of 0 on $A(x)$ (it exists since $A(x)$ is closed and convex, see Brézis [6] p. 20).*

We now turn to the theorem.

Theorem 2.5. *[6] Let A be maximal monotone then for all $u_0 \in D(A)$, there exists a unique function $u(t)$ from $[0, +\infty[$ into X such that*

- *$u(t) \in D(A)$ for all $t > 0$,*

- *$u(t)$ is Lipschitz continous on $[0, +\infty[$, i.e $u' \in L^\infty(0, +\infty; X)$ (in the sense of distributions) and*

$$|u'|_{L^\infty(0,+\infty;X)} \leq |A^\circ(u_0)|,$$

- *$-u'(t) \in A(u(t))$ for almost every t,*

- *$u(0) = u_0$.*

Moreover u verifies,

- *u has a right derivative for every $t \in [0, +\infty[$ and $-\dfrac{d^+u}{dt} \in A^\circ(u(t))$,*

- *the function $t \to A^\circ(u(t))$ is right continuous and $t \to |A^\circ(u(t))|$ is non increasing,*

- *if u and $\hat{u}$ are two solutions then $|u(t) - \hat{u}(t)| \leq |u(0) - \hat{u}(0)|$.*

2.2 Application to Arrow-Hurwicz methods

Let us now see how this theory can be applied for tracking saddle points. As mentioned before, we follow here [19]. We start with some definitions.

Definition 2.6. *Let $X = Y \oplus Z$ where Y and Z are two Hilbert spaces. A proper saddle function on X is a function K such that :*

- *for all $y \in Y$, the function $K(y, \cdot)$ is convex,*

- *for all $z \in Z$, the function $K(\cdot, z)$ is concave,*

- *there exists $x = (y, z)$ such that $K(y, z') < +\infty$ for all $z' \in Z$ and $K(y', z) > -\infty$ for all $y' \in Y$. The set of x for which it holds, is called the effective domain of K and is noted $\operatorname{dom} K$.*

Definition 2.7. *A point $(y, z) \in X$ is called a saddle point of K if*

$$K(y, z') \leq K(y, z) \leq K(y', z) \qquad \forall y' \in Y, \forall z' \in Z.$$

We then have,

Proposition 2.8. *A point (y, z) is a saddle point of a saddle function K, if and only if*

$$K(y, z) = \sup_{z' \in Z} \inf_{y' \in Y} K(y', z') = \inf_{y' \in Y} \sup_{z' \in Z} K(y', z').$$

The proof of this proposition is easy and can be found in Rockafellar's book [18] p.380.
The next theorem shows that the Arrow-Hurwicz method always provides a monotone operator.

Theorem 2.9. *[19] Let K be a proper saddle function. For $x = (y, z)$ let*

$$T(x) = \left\{ (y^*, z^*) \in Y \oplus Z / \begin{array}{l} y^* \text{ is a subgradient of } K(\cdot, z) \text{ in } y \\ z^* \text{ is a subgradient of } -K(y, \cdot) \text{ in } z \end{array} \right\}.$$

Then T is a monotone operator with $D(T) \subset \operatorname{dom} K$.

We can now characterize the saddle points of K using the operator T.

Proposition 2.10. *[19] Let K be a proper saddle function then a point x is a saddle point of K if and only if $0 \in T(x)$.*

Remark . This property is to be compared with the minimality condition $0 \in \partial f(x)$ for convex functions f.

The next theorem shows that for regular enough saddle functions, the corresponding operator T is maximal.

Theorem 2.11. [19] *Let K be a proper saddle function on X. Suppose that K is lsc in y and upper-semi-continuous in z then T is maximal monotone.*

Proof. We just sketch the proof because it will inspire us in the following. The idea is to use the equivalent theorem for convex functions. For this we "invert" the operator T in the second variable. Let

$$H(y, z^*) = \sup_{z \in X} \langle z^*, z \rangle + K(y, z).$$

The proof is then based on the following lemma :

Lemma 2.12. [19] *H is a convex lsc function on X and*

$$(y^*, z^*) \in T(y, z) \Leftrightarrow (y^*, z) \in \partial H(y, z^*).$$

It is then not too hard to prove that T is maximal.

$\square$

3 Study of the Primal-Dual Method

In this section, unless otherly stated, everything holds for general functionals J of the type (1).

Before starting the study of the Primal-Dual method, let us remind some facts about functions with bounded variation and pairings between measures and bounded functions.

Definition 3.1. *Let $BV(\Omega)$ be the space of functions u in L^1 for which*

$$\int_\Omega |Du| := \sup_{\substack{\xi \in \mathcal{C}^1_c(\Omega) \\ |\xi|_\infty \leq 1}} \int_\Omega u \operatorname{div} \xi < +\infty.$$

With the norm $|u|_{BV} = \int_\Omega |Du| + |u|_{L^1}$ it is a Banach space. We note the functional space $BV^2 = BV(\Omega) \cap L^2$.

9

Proposition 3.2. [1] *Let $u \in L^1(\Omega)$ then $u \in BV(\Omega)$ if and only if its distributional derivative Du is a finite Radon measure. Moreover the total variation of Du is equal to $\displaystyle\int_\Omega |Du|$.*

More informations about functions with bounded variation, can be found in the books [1] or [17].
Following Anzellotti [3], we define $\int_\Omega [\xi, Du]$ which has to be understood as $\int_\Omega \xi \cdot Du$, for functions u with bounded variation and bounded functions ξ with divergence in L^2.

Definition 3.3.   • *Let $X^2 = \left\{ \xi \in (L^\infty(\Omega))^d \,/\, \operatorname{div}\xi \in L^2(\Omega) \right\}$.*

• *For $(u, \xi) \in BV^2 \times X^2$ we define the distribution $[\xi, Du]$ by*

$$\langle [\xi, Du], \varphi \rangle = - \int_\Omega u\varphi \operatorname{div}(\xi) - \int_\Omega u\, \xi \cdot \nabla\varphi \qquad \forall \varphi \in \mathcal{C}_c^\infty(\Omega).$$

Theorem 3.4. [3] *The distribution $[\xi, Du]$ is a bounded Radon measure on Ω and if ν is the outward unit normal to Ω, we have Green's formula,*

$$\int_\Omega [\xi, Du] = - \int_\Omega u \operatorname{div}(\xi) + \int_{\partial\Omega} (\xi \cdot \nu)u.$$

We now prove a useful technical lemma which seems to be quite known but to our best knowledge nowhere explicitly proved.

Proposition 3.5. *Let $u \in BV(\Omega)$ then*

$$\int_\Omega |Du| = \sup_{\substack{\xi \in X^2 \\ |\xi|_\infty \leq 1}} \int_\Omega [\xi, Du].$$

Proof. By the definition of the total variation,

$$\int_\Omega |Du| \leq \sup_{\substack{\xi \in X^2 \\ |\xi|_\infty \leq 1}} \int_\Omega [\xi, Du].$$

We thus only have to prove the opposite inequality.
Let $\mathcal{C}(\Omega)$ be the space of continuous functions on Ω then by Proposition 1.47

p.41 of the book [1],

$$\int_\Omega |Du| = \sup_{\substack{\xi \in \mathcal{C}(\Omega) \\ |\xi|_\infty \leq 1}} \int_\Omega \xi \cdot Du$$

$$\geq \sup_{\substack{\xi \in \mathcal{C}(\Omega) \cap X^2 \\ |\xi|_\infty \leq 1}} \int_\Omega [\xi, Du].$$

In the second inequality, the fact that $\int_\Omega \xi \cdot Du = \int_\Omega [\xi, Du]$ comes from Proposition 2.3 of [3]. Let us also note that in the original Proposition 1.47 cited above, the supremum is taken over functions in $\mathcal{C}_c(\Omega)$ but a quick look to the proof shows that it can be enlarged to functions whose support is not compact.

We now want to show that for every ξ in X^2 with $|\xi|_\infty \leq 1$, there exists a sequence ξ_n in $X^2 \cap \mathcal{C}(\Omega)$ with $|\xi_n|_\infty \leq 1$ such that $\int_\Omega [\xi_n, Du]$ tends to $\int_\Omega [\xi, Du]$, which would end the proof.

By Lemma 2.2 and Proposition 2.1 of [3], for every $\xi \in X^2$ with $|\xi|_\infty \leq 1$, we can find $\xi_n \in X^2 \cap \mathcal{C}(\Omega)$ with $|\xi_n|_\infty \leq 1$ and $[\xi_n, Du]$ tending to $[\xi, Du]$ in the sense of weak convergence of measures.
The final step is now very similar to the proof of Theorem 4.2 of [3].

Let $\varepsilon > 0$ be given . There exists a number $\delta = \delta(\varepsilon) > 0$ such that if we let $\Omega_\delta = \{x \in \Omega \mid dist(x, \partial\Omega) > \delta\}$

$$\int_{\Omega \setminus \Omega_\delta} |Du| \leq \varepsilon.$$

Take η a function of $\mathcal{C}_c(\Omega)$ with $\eta = 1$ on Ω_δ and $|\eta|_\infty \leq 1$, then

$$\int_\Omega [\xi_n, Du] - \int_\Omega [\xi, Du] =$$
$$\left[\int_\Omega [\xi_n, Du]\eta - \int_\Omega [\xi, Du]\eta \right] + \left[\int_\Omega [\xi_n, Du](1 - \eta) - \int_\Omega [\xi, Du](1 - \eta) \right]$$

11

The first term in brackets goes to zero because of the weak convergence of $[\xi_n, Du]$ to $[\xi, Du]$. The second term can be bounded by

$$2|\xi_n|_\infty \int_{\Omega \backslash \Omega_\delta} |Du| + 2|\xi|_\infty \int_{\Omega \backslash \Omega_\delta} |Du| \le 4\varepsilon.$$

This shows the desired result. □

The next proposition gives a characterization of the minimizers of the functional J.

Proposition 3.6. *Let* $J(u) = \int_\Omega |Du| + G(u) + \int_{\partial\Omega_D} |u - \varphi|$ *then* u *is a minimizer of* J *in* BV^2 *if and only if there exists* $\xi \in X^2$ *with* $|\xi|_\infty \le 1$ *such that*

$$\begin{cases} \operatorname{div}(\xi) \in \partial G(u) \\[2mm] \int_\Omega |Du| = \int_\Omega [\xi, Du] \\[2mm] \xi \cdot \nu = 0 \ \text{in} \ \partial\Omega_N \ \ \text{and} \ \ (\xi \cdot \nu) \in \operatorname{sign}(\varphi - u) \ \text{in} \ \partial\Omega_D. \end{cases}$$

We do not give the proof of this proposition here since it can be either found in Andreu et al. [2] p.143 or derived more directly using the techniques we used in Proposition 3.7 and Proposition 3.8.

With these few propositions in mind we can turn back to the analysis of the Primal-Dual method. As noticed in the introduction, finding a minimizer of J is equivalent to finding a saddle point of

$$K(u, \xi) = \int_\Omega [Du, \xi] + G(u) + \int_{\partial\Omega_D} |u - \varphi| - I_{B(0,1)}(\xi).$$

The saddle function K does not fulfill the assumptions of Theorem 2.11 since it is not lsc in u. However staying in the spirit of Lemma 2.12, we set

$$H(u, \xi^*) = \sup_{\substack{\xi \in X^2 \\ |\xi|_\infty \leq 1}} \langle \xi, \xi^* \rangle + K(u, \xi)$$

$$= \sup_{\substack{\xi \in X^2 \\ |\xi|_\infty \leq 1}} \langle \xi, \xi^* \rangle + \int_\Omega [Du, \xi] + G(u) + \int_{\partial \Omega_D} |u - \varphi|$$

$$= \int_\Omega |Du + \xi^*| + G(u) + \int_{\partial \Omega_D} |u - \varphi|.$$

Where the last equality is obtained as in Proposition 3.5. The function H is then a convex lsc function on $L^2 \times (L^2)^d$ hence ∂H is maximal monotone. We are now able to define a maximal monotone operator T by

$$T(u, \xi) = \{(u^*, \xi^*) \, / \, (u^*, \xi) \in \partial H(u, \xi^*)\}.$$

In order to compute ∂H, which gives the expression of T, we use the characterization of the subdifferential

$$(u^*, \xi) \in \partial H(u, \xi^*) \iff \langle u^*, u \rangle + \langle \xi^*, \xi \rangle = H(u, \xi^*) + H^*(u^*, \xi).$$

A first step is thus to determine what H^* is.

Proposition 3.7. *We have*

$$D(H^*) = \left\{ (u^*, \xi) \, / \, u^* \in L^2(\Omega) \text{ and } \xi \in X^2 \, , \, \xi \cdot \nu = 0 \text{ in } \partial \Omega_N \, , \, |\xi|_\infty \leq 1 \right\}$$

and

$$H^*(u^*, \xi) = G^*(u^* + \mathrm{div}(\xi)) - \int_{\partial \Omega_D} (\xi \cdot \nu) \varphi.$$

Proof. We start by computing the domain of H^*.
If $(u^*, \xi) \in D(H^*)$ then there exists a constant C such that for every $(u, \xi^*) \in BV^2 \times (L^2)^d$,

$$\langle u^*, u \rangle + \langle \xi^*, \xi \rangle - H(u, \xi^*) \leq C.$$

Restraining to $u \in H^1(\Omega)$ with $u_{|\partial \Omega_D} = 0$ and $\xi^* \in (L^2)^d$, we find that

$$\langle u^*, u \rangle + \langle \xi^*, \xi \rangle - \int_\Omega |\nabla u + \xi^*| - G(u) \leq C$$

13

from which

$$\langle \nabla u + \xi^*, \xi \rangle - \langle \nabla u, \xi \rangle + \langle u^*, u \rangle - \int_\Omega |\nabla u + \xi^*| - G(u) \le C.$$

Setting $\xi' = \nabla u + \xi^*$ and taking the supremum over all $\xi' \in (L^2)^d$ we have that $|\xi|_\infty \le 1$ and for all $u \in H^1(\Omega)$ with $u_{|\partial\Omega_D} = 0$,

$$-\langle \nabla u, \xi \rangle + \langle u^*, u \rangle \le C + G(u).$$

Taking now $\tilde u = \lambda u$ with λ positive and reminding the form of G, it can be shown letting λ tending to infinity, that for every $u \in H^1$ with $u_{|\partial\Omega_D} = 0$,

$$-\langle \nabla u, \xi \rangle + \langle u^*, u \rangle \le C|u|_2.$$

This implies that $u^* + \operatorname{div}\xi \in L^2$ hence $\operatorname{div}\xi \in L^2$. Then by Green's formula in $H^1(\operatorname{div})$ (see Dautray-Lions [13] p.205) we have $\xi \cdot \nu = 0$ in $\partial\Omega_N$.

Let us now compute H^*.
Let $(u^*, \xi) \in D(H^*)$,

$$H^*(u^*, \xi) = \sup_{\xi^* \in L^2} \sup_{u \in BV^2} \left\{ \langle u^*, u \rangle + \langle \xi^*, \xi \rangle - \int_\Omega |Du + \xi^*| - G(u) - \int_{\partial\Omega_D} |u - \varphi| \right\}.$$

Let $\xi^* \in L^2$ be fixed. Then by Lemma 5.2 p.316 of Anzellotti's paper [3], for every $u \in BV^2$ there exists $u_n \in \mathcal{C}^\infty \cap BV^2$ such that

$$u_n \xrightarrow{L^2} u, \quad (u_n)_{|\partial\Omega_D} = u_{|\partial\Omega_D} \qquad \text{and}$$
$$\int_\Omega |Du_n + \xi^*| \to \int_\Omega |Du + \xi^*|.$$

We can thus restrict the supremum to functions u of class $\mathcal{C}^\infty(\Omega)$. We then have

$$H^*(u^*, \xi) = \sup_{u \in BV^2 \cap \mathcal{C}^\infty} \sup_{\xi \in L^2} \left\{ \langle u^*, u \rangle + \langle \xi^*, \xi \rangle - \int_\Omega |Du + \xi^*| - G(u) - \int_{\partial \Omega_D} |u - \varphi| \right\}$$

$$= \sup_{u \in BV^2 \cap \mathcal{C}^\infty} \left\{ \langle u^*, u \rangle - \langle \nabla u, \xi \rangle - G(u) - \int_{\partial \Omega_D} |u - \varphi| \right\}$$

$$= \sup_{u \in BV^2} \left\{ \langle u^*, u \rangle - \int_\Omega [Du, \xi] - G(u) - \int_{\partial \Omega_D} |u - \varphi| \right\}$$

$$= \sup_{u \in BV^2} \left\{ \langle u, u^* + \operatorname{div} \xi \rangle - G(u) - \int_{\partial \Omega_D} \{|u - \varphi| + (\xi \cdot \nu) u\} \right\}.$$

Beware that $u \in BV^2 \cap \mathcal{C}^\infty$ implies that $\nabla u \in L^1$ and not $\nabla u \in L^2$ but the density of L^2 in L^1 allows us to pass from the first equality to the second. The third equality follows from Lemma 1.8 of [3]. We now have to show that we can take separately the supremum in the interior of Ω and on the boundary $\partial \Omega_D$.

Let f be in $L^1(\partial \Omega)$ and v be in $L^2(\Omega)$. We want to find $u_\varepsilon \in BV^2$ converging to v in L^2 and such that $(u_\varepsilon)_{|\partial \Omega_D} = f$.

By Lemma 5.5 of [3] there is a $w_\varepsilon \in W^{1,1}$ with $(w_\varepsilon)_{|\partial \Omega_D} = f$ and $|w_\varepsilon|_2 \leq \varepsilon$. By density of $\mathcal{C}_c^\infty(\Omega)$ in L^2 we can find $v_\varepsilon \in \mathcal{C}_c^\infty(\Omega)$ with $|v_\varepsilon - v|_2 \leq \varepsilon$ We can then take $u_\varepsilon = v_\varepsilon + w_\varepsilon$.

This shows that

$$H^*(u^*, \xi) = \sup_{u \in L^2(\Omega)} \left\{ \langle u, u^* + \operatorname{div} \xi \rangle - G(u) \right\} - \inf_{u \in L^1} \int_{\partial \Omega_D} \{|u - \varphi| + (\xi \cdot \nu) u\}$$

$$= G^*(u^* + \operatorname{div}(\xi)) - \int_{\partial \Omega_D} (\xi \cdot \nu) \varphi.$$

$\square$

We can now compute T

Proposition 3.8. *Let $(u, \xi) \in BV^2 \times X^2$ then, $(u^*, \xi^*) \in T(u, \xi)$ if and only*

if

$$\begin{cases} u^* + \mathrm{div}(\xi) \in \partial G(u) \\[2mm] \int_\Omega |\xi^* + Du| = \langle \xi^*, \xi \rangle + \int_\Omega [\xi, Du] \\[2mm] \xi \cdot \nu = 0 \ \text{in} \ \partial\Omega_N \ \ \text{and} \ \ (\xi \cdot \nu) \in sign(\varphi - u) \ \text{in} \ \partial\Omega_D. \end{cases}$$

Proof. Let us first note that,

$$G(u) + G^*(u^* + \mathrm{div}(\xi)) \geq \langle u, u^* + \mathrm{div}(\xi) \rangle \tag{5}$$

$$\int_\Omega |Du + \xi^*| \geq \int_\Omega [\xi, Du] + \int_\Omega \xi^* \xi \tag{6}$$

$$|u - \varphi| \geq (\xi \cdot \nu)(\varphi - u) \tag{7}$$

where the second inequality is obtained arguing as in Proposition 3.5.
By definition, $(u^*, \xi^*) \in T(u, \xi)$ if and only if

$$\langle u, u^* \rangle + \langle \xi, \xi^* \rangle = H(u, \xi^*) + H^*(u^*, \xi)$$
$$= \int_\Omega |Du + \xi^*| + G(u) + \int_{\partial\Omega_D} |u - \varphi|$$
$$+ G^*(u^* + \mathrm{div}(\xi)) - \int_{\partial\Omega_D} (\xi \cdot \nu)\varphi.$$

This shows that (5), (6) and (7) must be equalities which is exactly

$$\begin{cases} u^* + \mathrm{div}(\xi) \in \partial G(u) \\[2mm] \int_\Omega |\xi^* + Du| = \langle \xi^*, \xi \rangle + \int_\Omega [\xi, Du] \\[2mm] (\xi \cdot \nu) \in sign(\varphi - u) \ \text{in} \ \partial\Omega_D. \end{cases}$$

Moreover, $\xi \cdot \nu = 0$ in $\partial\Omega_N$ because $(u, \xi) \in D(T)$. $\qquad\square$

Remark .

- The condition $(\xi \cdot \nu) \in sign(\varphi - u)$ in $\partial\Omega_D$ is equivalent to

$$\int_{\partial\Omega_D} |u - \varphi| + (\xi \cdot \nu)u = \inf_v \int_{\partial\Omega_D} |v - \varphi| + (\xi \cdot \nu)v$$

  because inequality (7) holds true for every v and is an equality for u.

- Whenever it has a meaning, it can be shown that the condition

$$\int_\Omega |\xi^* + Du| = \langle \xi^*, \xi \rangle + \int_\Omega [\xi, Du]$$

  is equivalent to

$$\xi^* + Du \in \partial I_{B(0,1)}(\xi)$$

  so that we will not distinguish between these two notations.

- This analysis shows why the constraint $u \in [0, 1]$ is hard to deal with. In fact, it imposes that $\mathrm{div}(\xi)$ is a measure but not necessarily a L^2 function. It is not easy to give a meaning to $\int_\Omega Du \cdot \xi$ or to $(\xi \cdot \nu)$ on the boundary for such functions. However, when dealing with numerical implementations, it is better to keep the constraint on u.

We can summarize those results in the following theorem which says that the Primal-Dual Method is well-posed.

Theorem 3.9. *For all $(u_0, \xi_0) \in \mathrm{dom}(T)$, there exists a unique $(u(t), \xi(t))$ such that*

$$\begin{cases} \partial_t u \in \mathrm{div}(\xi) - \partial G(u) \\[2mm] \partial_t \xi \in Du - \partial I_{B(0,1)}(\xi) \\[2mm] (\xi \cdot \nu) \in sign(\varphi - u) \ in \ \partial\Omega_D \qquad \xi \cdot \nu = 0 \ in \ \partial\Omega_N \\[2mm] (u(0), \xi(0)) = (u_0, \xi_0). \end{cases} \qquad (8)$$

Moreover, the energy $|\dfrac{d^+ u}{dt}|_2^2 + |\dfrac{d^+ \xi}{dt}|_2^2$ is non increasing and if $(\bar{u}, \bar{\xi})$ is a saddle point of K, $|u - \bar{u}|_2^2 + |\xi - \bar{\xi}|_2^2$ is also non increasing.

Proof. The operator T is maximal monotone hence Theorem 2.5 applies and gives the result. $\square$

Remark . This theorem also shows that whenever J has a minimizer, K has saddle points. This is because stationnary points of the system (8) are minimizers of J (verifying the Euler-Lagrange equation for J, remind Proposition 3.6).

For the Rudin-Osher-Fatemi model, one can show that there is convergence of u to the minimizer of the functional J and obtain *a posteriori* estimates.

Proposition 3.10. *Let* $G = \dfrac{\lambda}{2} \displaystyle\int_\Omega (u - f)^2$ *and* $\partial\Omega_D = \emptyset$. *Then if* $\bar{u}$ *is the minimizer of* J, *every solution of (8) converges in* L^2 *to* $\bar{u}$. *Furthermore,*

$$
|u - \bar{u}|_2 \leq \frac{1}{2}\left(\frac{1}{\lambda}|\partial_t u|_2 + \sqrt{ \frac{|\partial_t u|_2^2}{\lambda^2} + \frac{8|\Omega|^{\frac{1}{2}}}{\lambda}|\partial_t \xi|_2 } \right).
$$

Proof. Let $(\bar{u}, \bar{\xi})$ be such that $0 \in T(\bar{u}, \bar{\xi})$. Let $e(t) = |u(t) - \bar{u}|_2^2$ and $g(t) = |\xi(t) - \bar{\xi}|_2^2$. We show that

$$
\frac{1}{2}(e + g)' \leq -\lambda e. \tag{9}
$$

Indeed, by definition of the flow,

$$
\int_\Omega [\xi, Du] - \langle \xi, \partial_t \xi \rangle \geq \int_\Omega [\bar{\xi}, Du] - \langle \bar{\xi}, \partial_t \xi \rangle \qquad \text{and}
$$

$$
\int_\Omega [\bar{\xi}, D\bar{u}] - \langle \bar{\xi}, \partial_t \bar{\xi} \rangle \geq \int_\Omega [\xi, D\bar{u}] - \langle \xi, \partial_t \bar{\xi} \rangle.
$$

Summing these two we find,

$$
\int_\Omega [\xi - \bar{\xi}, D(u - \bar{u})] \geq \langle \xi - \bar{\xi}, \partial_t \xi - \partial_t \bar{\xi} \rangle.
$$

We thus have

$$
\begin{aligned}
\frac{1}{2}(e + g)' &= \langle u - \bar{u}, \partial_t u - \partial_t \bar{u} \rangle + \langle \xi - \bar{\xi}, \partial_t \xi - \partial_t \bar{\xi} \rangle \\
&\leq \langle u - \bar{u}, \operatorname{div}(\xi - \bar{\xi}) - \lambda(u - \bar{u}) \rangle + \int_\Omega [\xi - \bar{\xi}, D(u - \bar{u})] \\
&= -\lambda e.
\end{aligned}
$$

The functions e and g are Lipschitz continuous. Let L be the Lipschitz constant of e and let $h = e + g$.

Let us show by contradiction that e tends to zero when t tends to infinity.

18

Suppose that there exists $\alpha > 0$ and $T > 0$ such that $e \geq \alpha$ for all $t > T$, then we would have $h' \leq -\lambda\alpha$ and h would tend to minus infinity which is impossible by positivity of h. Hence

$$\forall \alpha > 0 \ \forall T > 0 \ \exists t \geq T \qquad e(t) \leq \alpha.$$

Suppose now the existence of $\varepsilon > 0$ such that for all $T \geq 0$ there exists $t \geq T$ with $e(t) \geq \varepsilon$.
By continuity of e, there exists a sequence $(t_n)_{n \in \mathbb{N}}$ with $\lim\limits_{n \to +\infty} t_n = +\infty$ such that

$$e(t_{2n}) = \frac{\varepsilon}{2} \qquad e(t_{2n+1}) = \varepsilon.$$

Moreover, on $[t_{2n-1}, t_{2n}]$, we have $e(t) \geq \frac{\varepsilon}{2}$. We then find that

$$|e(t_{2n}) - e(t_{2n-1})| \leq L(t_{2n} - t_{2n-1}) \qquad \text{so}$$
$$\frac{\varepsilon}{2L} \leq t_{2n} - t_{2n-1}.$$

From which we see that,

$$h(t_{2n+2}) = h(t_{2n+1}) + \int_{t_{2n+1}}^{t_{2n+2}} h'(t)\, dt$$
$$\leq h(t_{2n+1}) - \varepsilon\lambda(t_{2n+2} - t_{2n+1})$$
$$\leq h(t_{2n}) - \frac{\lambda\varepsilon^2}{2L}.$$

This shows that $\lim\limits_{t \to +\infty} e(t) = 0$.

We now prove the *a posteriori* error estimate.

We have that

$$u = f + \frac{1}{\lambda}(\operatorname{div} \xi - \partial_t u)$$
$$\bar{u} = f + \frac{1}{\lambda} \operatorname{div} \bar{\xi}.$$

19

Which leads to

$$|u - \bar{u}|_2^2 = \frac{1}{\lambda}\langle \operatorname{div}(\xi - \bar{\xi}) - \partial_t u, u - \bar{u}\rangle$$

$$= \frac{1}{\lambda}\left[\langle \operatorname{div}(\xi - \bar{\xi}), u - \bar{u}\rangle - \langle \partial_t u, u - \bar{u}\rangle\right]$$

$$= \frac{1}{\lambda}\left[-\langle \xi - \bar{\xi}, Du - D\bar{u}\rangle - \langle \partial_t u, u - \bar{u}\rangle\right]$$

$$\leq \frac{1}{\lambda}\left[\int_\Omega |Du| - \int_\Omega [\xi, Du] + |\partial_t u|_2 |u - \bar{u}|_2\right]$$

Where the last inequality follows from $\int_\Omega [\bar{\xi}, Du] \leq \int_\Omega |Du|$ and $\int_\Omega \bar{\xi} \cdot D\bar{u} = \int_\Omega |D\bar{u}| \geq 0$.

Studying the inequality $X^2 \leq A + BX$, we can deduce that

$$|u - \bar{u}|_2 \leq \frac{1}{2}\left(\frac{1}{\lambda}|\partial_t u|_2 + \sqrt{\frac{|\partial_t u|_2^2}{\lambda^2} + \frac{4}{\lambda}(\int_\Omega |Du| - \int_\Omega [\xi, Du])}\right).$$

The estimate follows from the fact that

$$\int_\Omega |-\partial_t \xi + Du| = \int_\Omega [\xi, Du] - \int_\Omega \partial_t \xi \cdot \xi \quad \text{thus}$$

$$\int_\Omega |Du| - \int_\Omega |\partial_t \xi| \leq \int_\Omega [\xi, Du] - \int_\Omega \partial_t \xi \cdot \xi \quad \text{hence}$$

$$\int_\Omega |Du| - \int_\Omega [\xi, Du] \leq 2\int_\Omega |\partial_t \xi| \leq 2|\Omega|^{\frac{1}{2}}|\partial_t \xi|_2.$$

$$\square$$

Following the same lines, we can show *a posteriori* error estimates for general finite difference scheme. Indeed if ∇^h is any discretization of the gradient and if div^h is defined as $-(\nabla^h)^*$, the associated algorithm is

$$\begin{cases} \xi^n = P_{B(0,1)}(\xi^{n-1} + \delta\tau^n \nabla^h u^{n-1}) \\ u^n = u^{n-1} + \delta t^n(\operatorname{div}^h \xi^n - \lambda(u^{n-1} - f)). \end{cases} \tag{10}$$

Where $P_{B(0,1)}(\xi)_{i,j} = \dfrac{\xi_{i,j}}{\max(|\xi_{i,j}|, 1)}$ is the componentwise projection of ξ on the unit ball. This algorithm is exactly the one proposed by Chan and Zhu in [20]. We can associate to this system a discrete energy,

$$J_h(u) = \sum_{i,j} |\nabla^h u|_{i,j} + \frac{\lambda}{2} \sum_{i,j} |u_{i,j} - f_{i,j}|^2.$$

The algorithm (10) could have been directly derived from this discrete energy using the method of Chan and Zhu [20] (which is just the discrete counterpart of our continuous method). Hence, the next proposition gives a stopping criterion for their algorithm.

Proposition 3.11. *Let $N \times M$ be the size of the discretization grid and $\bar{u}$ be the minimizer of J_h then*

$$|u^n - \bar{u}|_2 \leq \frac{1}{2} \left(\frac{1}{\lambda} |\partial_t u^n|_2 + \sqrt{\frac{|\partial_t u^n|_2^2}{\lambda^2} + \frac{8\sqrt{N \times M}}{\lambda} |\partial_t \xi^n|_2} \right).$$

Where $\partial_t u^n = \dfrac{u^{n+1} - u^n}{\delta t^{n+1}}$ and $\partial_t \xi^n = \dfrac{\xi^{n+1} - \xi^n}{\delta \tau^{n+1}}$.

The proof of this discrete estimate is almost the same as for the continuous one. We give it in the appendix.

Remark . In opposition to the continuous framework where we were able to prove a convergence result, no fully satisfactory statement is known in the discrete framework. For some partial results we refer to Esser et al. [15] and to Chambolle and Pock [11]. These works mainly focus on slight modifications of the Primal-Dual algorithm (10) but they also show that in some restricted cases the algorithm of Chan and Zhu converges.

For the general problem, there is no uniqueness for the minimizer (for example in the segmentation problem) and hence convergence may not occur or be hard to prove. Indeed, even when uniqueness holds, we can have non vanishing oscillations. For example in the simpler one dimensional problem

$$\min_{u \in BV([0,1])} \int_0^1 |u'|$$

the unique minimizer is $u = 0$ but $u(t,x) = \frac{1}{2}\cos(\pi x)\sin(\pi t)$ and $\xi(t,x) = \frac{1}{2}\sin(\pi x)\cos(\pi t)$ gives a solution to the associated PDE system which does not converge to a saddle point. In this example, the energy is constant hence not converging to zero. We can however show general *a posteriori* estimates for the energy.

Proposition 3.12. *For every saddle point $(\bar{u}, \bar{\xi})$ and every (u_0, ξ_0), the solution $(u(t), \xi(t))$ of (8) satisfies*

$$|J(u) - J(\bar{u})| \leq \left(\sqrt{|u_0 - \bar{u}|_2^2 + |\xi_0 - \bar{\xi}|_2^2}\right)|\partial_t u|_2 + 2|\Omega|^{\frac{1}{2}}|\partial_t \xi|_2.$$

Proof. Let $(\bar{u}, \bar{\xi})$ be a saddle point and $(u(t), \xi(t))$ be a solution of (8).

$$J(u) - J(\bar{u}) = \int_\Omega |Du| + \int_{\partial\Omega_D} |u - \varphi| - \int_\Omega |D\bar{u}| - \int_{\partial\Omega_D} |\bar{u} - \varphi| + G(u) - G(\bar{u}).$$

By definition of the operator T we have

$$\int_\Omega [\xi, Du] - \int_\Omega \partial_t \xi \cdot \xi = \int_\Omega |Du - \partial_t \xi|$$

$$\geq \int_\Omega |Du| - \int_\Omega |\partial_t \xi|.$$

This shows that

$$\int_\Omega |Du| \leq \int_\Omega [\xi, Du] + 2\int_\Omega |\partial_t \xi|. \qquad (11)$$

On the other hand,

$$\int_\Omega [\xi, Du] + \int_{\partial\Omega_D} |u - \varphi| = -\int_\Omega u \operatorname{div}\xi + \int_{\partial\Omega_D} \{(\xi \cdot \nu)u + |u - \varphi|\}.$$

Applying $\int_{\partial\Omega_D} \{(\xi \cdot \nu)u + |u - \varphi|\} = \inf_v \int_{\partial\Omega_D} \{(\xi \cdot \nu)v + |v - \varphi|\}$ (remember the Remarks after Proposition 3.8) to $v = \bar{u}$ we have

$$\int_\Omega [\xi, Du] + \int_{\partial\Omega_D} |u - \varphi| - \int_{\partial\Omega_D} |\bar{u} - \varphi| \leq -\int_\Omega u \operatorname{div}\xi + \int_{\partial\Omega_D} (\xi \cdot \nu)\bar{u}$$

$$= -\int_\Omega u \operatorname{div}\xi + \int_\Omega \bar{u} \operatorname{div}\xi + \int_\Omega [\xi, D\bar{u}]$$

$$= \int_\Omega (\bar{u} - u)\operatorname{div}\xi + \int_\Omega [\xi, D\bar{u}].$$

22

This and (11) show that

$$J(u) - J(\bar{u}) \leq \int_\Omega (\bar{u} - u) \operatorname{div} \xi + \int_\Omega [\xi, D\bar{u}] + 2 \int_\Omega |\partial_t \xi| - \int_\Omega |D\bar{u}| + G(u) - G(\bar{u}).$$

If we now use the definition of the subgradient to get

$$G(u) - G(\bar{u}) \leq \langle \operatorname{div}(\xi) - \partial_t u, u - \bar{u} \rangle$$

we find with Cauchy-Schwarz's inequality,

$$J(u) - J(\bar{u}) \leq 2|\Omega|^{\frac{1}{2}} |\partial_t \xi|_2 + \int_\Omega (\bar{u} - u)\partial_t u + \int_\Omega [\xi, D\bar{u}] - \int_\Omega |D\bar{u}|$$

$$\leq 2|\Omega|^{\frac{1}{2}} |\partial_t \xi|_2 + |\bar{u} - u|_2 |\partial_t u|_2$$

which gives the estimate reminding that $\sqrt{|u - \bar{u}|_2^2 + |\xi - \bar{\xi}|_2^2}$ is non increasing. $\qquad\square$

Remark .
Supported by numerical evidence, we can conjecture that whenever the constraint on ξ is saturated somewhere, convergence of u occurs. It might however be also necessary to add the constraint $u \in [0, 1]$ in order to have this convergence.

Considering a finite difference scheme, just as for the Rudin-Osher-Fatemi model, we can define a discrete energy J_h and show the corresponding *a posteriori* estimate.

Proposition 3.13. *If $\bar{u}$ is a minimizer of J_h and (u^n, ξ^n) is defined by*

$$\begin{cases} \xi^n = P_{B(0,1)}(\xi^{n-1} + \delta\tau^n \nabla^h u^{n-1}) \\ u^n = u^{n-1} + \delta t^n (\operatorname{div}^h \xi^n - p^n). \end{cases}$$

with $p^n \in \partial G^h(u^{n-1})$ then

$$|J_h(u^n) - J_h(\bar{u})| \leq 2\sqrt{N \times M} |\partial_t \xi^n| + |\partial_t u^n||u^{n-1} - \bar{u}|.$$

We omit the proof because it is exactly the same as for Proposition 3.12.

Remark .

- The boundary conditions are hidden here in the operator ∇^h.

- In the discrete framework, the estimate involves $|u^n - \bar{u}|$ which can not be easily bounded by the initial error.

- In this still more general framework than in (10) no good convergence result is known. For some partial results for variants of this algorithm, we refer again to Chambolle and Pock [11].

4 Numerical Experiments

To illustrate the relevance of our *a posteriori* estimates, we first consider the simple example of denoising a rectangle (see Figure 3). We then compare the *a posteriori* error bound with the "true" error. We use the relative L^2 error defined as $\dfrac{|u^n - \bar{u}|}{|\bar{u}|}$ and ran the algorithm of Chan and Zhu with $\lambda = 0.005$ and fixed time steps verifying $\lambda \delta t = 1$ and $\delta \tau = \frac{\lambda}{5}$. With this choice of parameters convergence is guaranteed by the work of Esser et al. [15]. The minimizer $\bar{u}$ is computed by the algorithm after 50000 iterations. Figure 4 shows that the *a posteriori* bound is quite sharp.

Figure 3: Denoising of a rectangle using the ROF model

The second experiment is performed on the yeast segmentation of Figure 2. The solution was computed with the algorithm of Chan and Zhu using as weight function g the one proposed by Appleton and Talbot [4]. We used this time the error $|J_h(u^n) - J_h(\bar{u})|$ and ran the algorithm with $\delta t = 0.2$ and $\delta \tau = 0.2$. For this problem there is no proof of convergence of the algorithm.
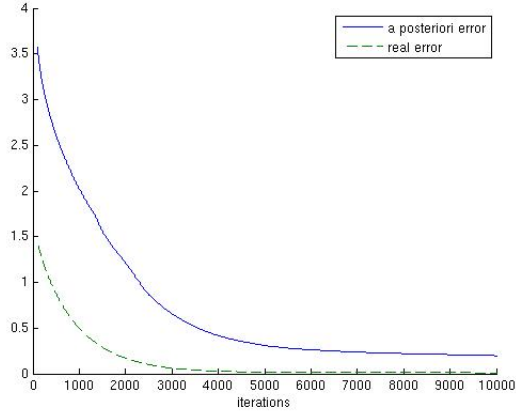
Figure 4: Comparison of the relative L^2 error with the predicted *a posteriori* bound.

The minimizer $\bar{u}$ is computed by the algorithm after 50000 iterations. We can see on Figure 5 that for this problem, the *a posteriori* estimate is not so sharp. We must also notice that in general we do not know $\bar{u}$.
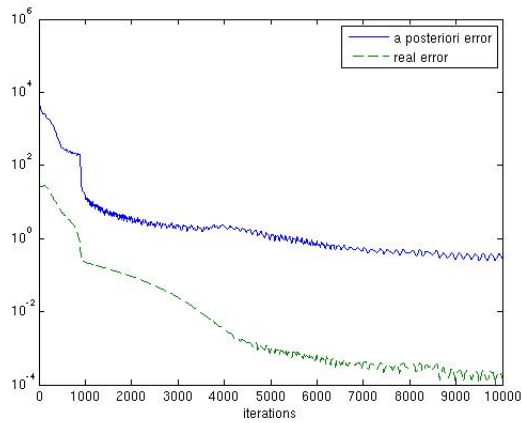

Figure 5: Comparison for the segmentation problem.

In the last numerical example, we compare the results obtained by the algorithm of Appleton and Talbot (see [4]) with those obtained by a classical discretization of the total variation. In Figure 6, we can see the denoising of

a disk with these two methods for $\lambda = 0.003$. We used the algorithm of Chan and Zhu [20] to compute the minimization of the discrete total variation. Looking at the top right corner (see Figure 7), we can see that the result is more accurate and less anisotropical for the algorithm of Appleton and Talbot than for the scheme of Chan and Zhu. These results are to be compared with those obtained by Chambolle et al. for the so-called "upwind" discrete BV norm in [10].
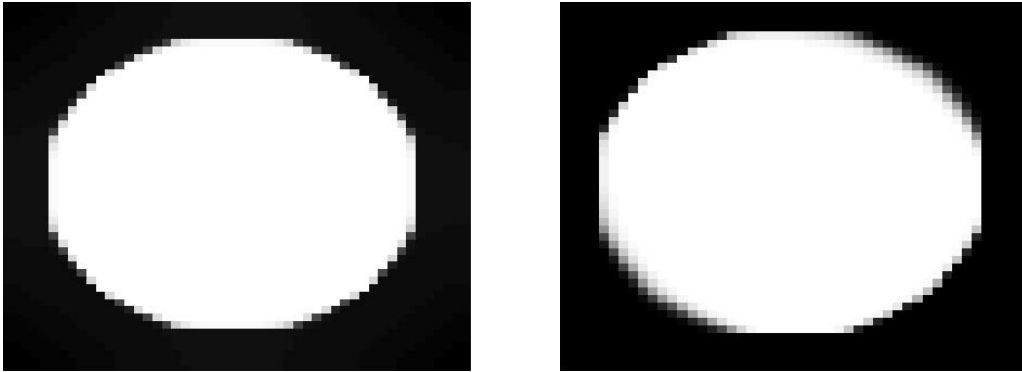

Figure 6: Denoising of a disk using the algorithm of Appleton-Talbot (left) and Chan-Zhu (right)

Figure 7: Top right corner of the denoised disk, Appleton-Talbot (left) and Chan-Zhu (right)

5   Conclusion

In this article we have shown the well posedness of the continuous Primal-Dual method proposed by Appleton and Talbot for solving problems arising

in imaging. We have also proved for the ROF model, that in the continuous setting there is convergence towards the minimizer. We then derived some *a posteriori* estimates. Numerical experiments have illustrated that if these estimates are quiet sharp for the ROF model, they should be improved for applications to other problems.

The continuous framework leaves the way open to a wide variety of numerical schemes, ranging from finite differences to finite volumes. Indeed, by designing algorithms solving the system of PDEs (8) one can expect to find accurate algorithms for computing solutions of variational problems involving a total variation term.

A   Proof of Proposition 3.11

For notational convenience, we present the proof for $\lambda = 1$. Let $\bar{u}$ be the minimizer of J_h then there exists $\bar{\xi}$ such that $|\bar{\xi}|_\infty \leq 1$ and

$$
\begin{cases}
\sum_{i,j} |\nabla^h \bar{u}|_{i,j} = \langle \nabla^h \bar{u}, \bar{\xi} \rangle \\
\bar{u} = \operatorname{div}^h \bar{\xi} + f.
\end{cases}
$$

Reminding that $u^n = f + \operatorname{div}^h \xi^{n+1} - \partial_t u^n$ we get

$$
\begin{aligned}
|u^n - \bar{u}|^2 &= \langle \operatorname{div}^h(\xi^{n+1} - \bar{\xi}) - \partial_t u^n, u^n - \bar{u} \rangle \\
&= -\langle \xi^{n+1} - \bar{\xi}, \nabla^h u^n - \nabla^h \bar{u} \rangle - \langle \partial_t u^n, u^n - \bar{u} \rangle \\
&\leq \langle \bar{\xi} - \xi^{n+1}, \nabla^h u^n \rangle + |\partial_t u^n||u^n - \bar{u}|.
\end{aligned}
$$

We have that $\xi^{n+1} = P_{B(0,1)}(\xi^n + \delta\tau^{n+1}\nabla^h u^n)$ hence by definition of the projection,

$$
\forall \bar{\xi} \in B(0,1) \qquad \langle \xi^{n+1} - (\xi^n + \delta\tau^{n+1}\nabla^h u^n), \bar{\xi} - \xi^{n+1} \rangle \geq 0.
$$

This gives us

$$
\langle \nabla^h u^n, \bar{\xi} - \xi^{n+1} \rangle \leq \langle \partial_t \xi^n, \bar{\xi} - \xi^n \rangle.
$$

Combining this with $\langle \partial_t \xi^n, \bar{\xi} \rangle - \langle \partial_t \xi^n, \xi^n \rangle \leq 2\sqrt{N \times M}|\partial_t \xi^n|$ (which holds by Cauchy-Schwarz's inequality, $|\bar{\xi}|_\infty \leq 1$ and $|\xi^n|_\infty \leq 1$), we find that

$$
|u^n - \bar{u}|^2 \leq 2\sqrt{N \times M}|\partial_t \xi^n| + |\partial_t u^n||u^n - \bar{u}|.
$$

The announced inequality easily follows.

References

[1] L. AMBROSIO, N. FUSCO AND D. PALLARA, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford Science Publications, 2000.

[2] F. ANDREU-VAILLO, V. CASELLES AND J.M. MAZÒN, *Parabolic Quasilinear Equations Minimizing Linear Growth Functionals*, Birkhäuser, collection "Progress in Mathematics", no.223, 2004.

[3] G. ANZELLOTTI, *Pairings between measures and bounded functions and compensated compactness*, Annali di Matematica Pura ed Applicata, Vol.135, No.1, p.293-318, 1983.

[4] B. APPLETON AND H. TALBOT, *Globally Minimal Surfaces by Continuous Maximal Flows*, IEEE Trans. on Pattern Anal. and Mach. Intelligence, vol.28, no.1, p.106-118, 2006.

[5] K. J. ARROW, L. HURWICZ AND H. UZAWA (editors), *Studies in Linear and Nonlinear Programming*, Stanford Univ. Press, Stanford, 1958.

[6] H. BRÉZIS, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, Amsterdam-London: North Holland. 1973.

[7] V. CASELLES, R. KIMMEL AND G. SHAPIRO, *Geodesic Active Contours*, Int. J. Computer Vision, vol.22, no.1, p.61-79, 1997.

[8] A. CHAMBOLLE, V. CASELLES, M. NOVAGA, D. CREMERS AND T. POCK, *An introduction to Total Variation for Image Analysis*, hal-00437581.

[9] A. CHAMBOLLE, D. CREMERS AND T. POCK, *A convex approach for computing minimal partitions*, preprint CMAP, November 2008.

[10] A. CHAMBOLLE, S. E. LEVINE AND B. J. LUCIER, *Some Variations on Total Variation-Based Image Smoothing*, preprint 2009.

[11] A. CHAMBOLLE AND T. POCK, *A First Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging*, preprint CMAP, May 2010.

[12] 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, p.1964-1977, 1999.

[13] R. DAUTRAY AND J.L. LIONS, *Mathematical Analysis and Numerical Methods for Science and Technology*, vol.3, Spectral Theory and Applications, Springer, 1990.

[14] I. EKELAND AND R. TEMAM, *Analyse convexe et problèmes variationnels*, Dunod Gauthiers-Villars, collection "Études mathématiques", 1974.

[15] E. ESSER, X. ZHANG AND T. CHAN, *A General Framework for a Class of First Order Primal-Dual Algorithms for TV Minimization*, CAM Reports 09-67, UCLA, Center for Applied Math., 2009.

[16] M. GIAQUINTA, G. MODICA AND J. SOUČEK, *Functionals with linear growth in the calculus of variations*, Com. Math. Uni. Carolinae, 20, p.143-171, 1979.

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

[18] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, 1970.

[19] R. T. ROCKAFELLAR, *Monotone Operators Associated with Saddle-Functions and Minimax Problems*, Nonlinear Functional Analysis, F.E. Browder ed., Proc. of Symp. in Pure Math. of the Am. Math. Society, vol. XVIII, part. 1, p.241-250, 1968.

[20] M. ZHU AND T. CHAN, *An Efficient Primal-Dual Hybrid Gradient Algorithm for Total Variation Image Restoration*, cam report (08-34), May 2008.