



HAL
open science

A penalization approach for tomographic reconstruction of binary axially symmetric objects

Romain Abraham, Maïtine Bergounioux, Emmanuel Trélat

► **To cite this version:**

Romain Abraham, Maïtine Bergounioux, Emmanuel Trélat. A penalization approach for tomographic reconstruction of binary axially symmetric objects. *Applied Mathematics and Optimization*, 2008, 58 (3), pp.345-371. 10.1007/s00245-008-9039-8 . hal-00139494v2

HAL Id: hal-00139494

<https://hal.science/hal-00139494v2>

Submitted on 28 Mar 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A penalization approach for tomographic reconstruction of binary axially symmetric objects

R. Abraham M. Bergounioux E. Trélat *

Abstract

We propose a variational method for tomographic reconstruction of blurred and noised binary images based on a penalization process of a minimization problem settled in the space of bounded variation functions. We prove existence and/or uniqueness results and derive a penalized optimality system. Numerical simulations are provided to demonstrate the relevance of the approach.

Keywords: Tomography, Optimization, Penalization.

1 Introduction

In this article, we focus on a specific application of tomographic reconstruction for a physical experiment whose goal is to study the behavior of a material under a shock. The experiment, depicted on Figure 1, consists in causing the implosion of the hull of some material (usually, a metal) whose features are well known, using surrounding explosives. The problem is to determine the shape of the interior interface at a specific moment of the implosion. For this purpose, a single X-ray radiography is performed, and the shape of the object must then be reconstructed using a tomographic approach.

When enough projections of the object, taken from different angles, are available, several techniques exist for tomographic reconstruction, providing an analytic formula for the solution (see for instance [15] or [12]). There is a huge literature about theoretical and practical aspects of the problem of reconstruction from projections, the applications of which concern medicine, optics, material science, astronomy, geophysics, and magnetic resonance imaging (see [6]). An important application is the problem of medical transmission X-ray tomography (see [16]), in which X-rays are fired from many angles through a single cross section of the body, measuring line integrals of the linear attenuation coefficient of the object. The resulting collection of projections then permits to reconstruct the 3D body.

*Université d'Orléans, UFR Sciences, Math., Labo. MAPMO, UMR 6628, Route de Chartres, BP 6759, 45067 Orléans cedex 2, France. E-mail: romain.abraham@univ-orleans.fr, maitine.bergounioux@univ-orleans.fr, emmanuel.trelat@univ-orleans.fr

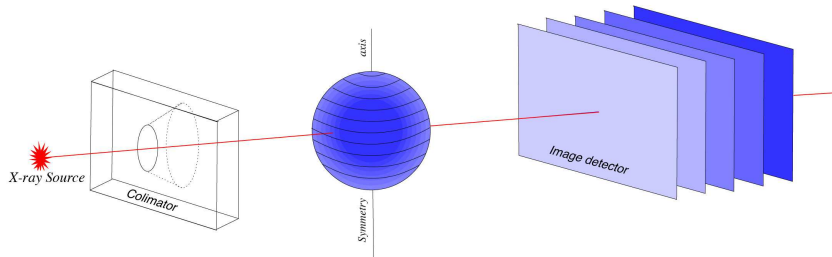


Figure 1: Experimental setup.

When only few projections are known, these methods cannot be used directly, and some alternative methods have been proposed to reconstruct partially the densities (see for instance [11]). In our experiment, all components of the initial physical setup (object, hull, explosives, etc) are assumed to be axially symmetric, and are assumed to remain as such during the implosion process. High speed image capture provides a snapshot of the deformation of an object by X-ray radiography. Since this object is assumed to be axially symmetric, a single radiograph of the cross section suffices in theory to reconstruct the 3D object. For objects having nearly circular symmetry, such an approach, using a single radiograph taken with a radiographic axis perpendicular to the symmetry axis of the object, offers significant benefits as an image analysis tool, and has been applied to many problems in industrial tomography (see [14]) or in astronomy (see [18]).

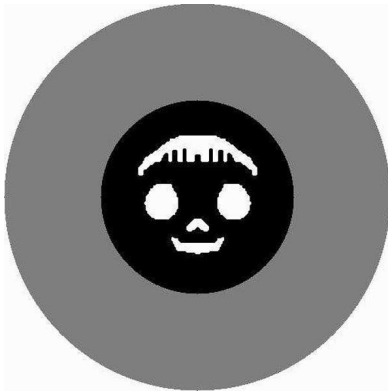
As in any tomographic reconstruction process, this problem leads to an ill-posed inverse problem. Since we only have one radiograph at our disposal, data are not redundant and the ill-posed character is even more accurate. Moreover, the flash has to be very brief (several nanoseconds) due to the imploding movement of the hull. Such X-rays cannot be provided by standard sources, and hence numerous drawbacks appear, for instance the X-rays beam is not well focused and the X-rays source is not punctual. This causes a blur on the radiograph. Furthermore, contrarily to medical radiography where photons are absorbed by bones, here X-rays must cross a very dense object and therefore must be of high energy. Most of the photons are actually absorbed by the object and only a few number of them arrive at the detector. It is therefore necessary to add some amplification devices and very sensitive detectors, which cause a high noise level and another blur.

In our experiment, X-rays are assumed to be parallel, “horizontal” slices of the object are independent and are treated separately. Hence, usual regularization techniques for tomography (such as filtered backprojection) are not adapted, since they deal with one particular slice. Here, because of the axial symmetry, slices are composed of concentric annuli and do not need any regularization. At the contrary, some regularity between the different slices is required, and only few results in that direction are known (see [8, 11]).

Another difficulty is that we deal with *binary objects* composed of one homogeneous material (drawn in black) and of some holes (in white). Our working example, drawn

on Figure 2, represents a synthetic object containing all standard difficulties that may appear, such as:

- several disconnected holes;
- a small hole located on the symmetry axis (where details are expected to be difficult to recover because the noise variance is maximal around the symmetry axis after reconstruction);
- smaller details on the boundary of the top hole, serving as a test for lower bound detection.



(a) Slice of a binary axially symmetric object by a plane containing the symmetry axis.



(b) Zoom on the interior of the object of Figure 2(a); the homogeneous material is drawn in black and the holes in white.

Figure 2: Working example.

Figure 2(a) shapes an object composed of concentric shells of homogeneous materials (called the “exterior” in what follows) surrounding a ball (called the “interior”) of another homogeneous material containing empty holes. It can be viewed as the slice of a axially symmetric 3D object by a plane containing the symmetry axis of that object. A rotation of the image of Figure 2(a) around the z axis must be performed in order to recover the 3D-object, in which, for instance, the two white holes generate a torus. Since the object is mainly featured in the shape of the holes, in the sequel we will focus on the interior of the object (see Figure 2(b)).

Based on a single X-ray radiograph which is at our disposal, it is our aim to perform a tomographic reconstruction to reconstruct the whole axially symmetric object. We propose in this article a variational method adapted to the tomographic reconstruction of blurred and noised binary images, based on a minimization problem. Note that our

approach is global, contrarily to usual methods of reconstruction rather dealing with a line-by-line analysis carried out on each line of the radiograph (see [14]).

The article is structured as follows.

In Section 2, the experimental setup described above is modeled mathematically, the projection operator is defined and some properties are derived. We show that the method which consists in inverting the projection operator is not suitable.

To deal with this ill-posed problem, we introduce in Section 3 a variational approach for tomographic reconstruction. We define a minimization problem in the space of bounded variation functions, using the concept of total variation, prove existence and uniqueness results. The binary structure of the material under consideration is modeled as a binary constraint: the intensity function is either equal to 0 or 255 (normalized to 0 and 1). Due to this binary constraint, deriving an optimality system is not straightforward, and we propose a penalization method for which we establish some properties and derive an optimality system.

Section 4 is devoted to the numerical implementation of these conditions. We propose a fixed-point type algorithm based on a projected gradient method and on an algorithm of [10] to handle the total variation which is nonsmooth. Simulation results are provided and commented for several benches of parameters, concerning our working example of Figure 2. We do not perform an exhaustive comparison with existing numerical methods in tomography, however we provide results obtained with the usual back-filtered projection, which happen to be of bad quality, as expected.

2 Modelization and projection operator

Radiography measures the attenuation of X-rays through the object. A point on the radiograph is determined by Cartesian coordinates (y, z) , where the z axis is the symmetry axis. Let I_0 denote the intensity of the incident X-rays flux. Then, the measured flux $I(y, z)$ at a point (y, z) is given by

$$I = I_0 e^{-\int \mu(r, \theta, z) dl},$$

where the integral operates along the ray that reaches the point (y, z) of the detector, dl is the infinitesimal element of length along the ray, and μ is the linear attenuation coefficient. Considering the Neperian logarithm of this attenuation permits to deal rather with linear operators, and the linear mapping

$$\mu \longmapsto \int \mu dl$$

is called the *projection operator*.

In practice, the ratio of the diameter of the object by the distance “X-ray source - object” is less than 1/100, and it is therefore assumed throughout the article, for the sake

of simplicity, that the rays are parallel, and orthogonal to the symmetry axis. It follows that horizontal slices of the object can be considered separately to perform the projection. In these conditions, for a 3D object represented by a linear attenuation coefficient $\tilde{u}(x, y, z)$ (with compact support) in Cartesian coordinates, the projection operator H_0 can be written

$$(H_0\tilde{u})(y, z) = \int_{\mathbb{R}} \tilde{u}(x, y, z) dx. \quad (1)$$

Since the objects under study in this article are bounded and axially symmetric, it is relevant to make use of cylindrical coordinates (r, θ, z) , where the z -axis designates the symmetry axis. Such objects are represented by a linear attenuation coefficient $u(r, z)$, where u denotes a function of $L^\infty(\mathbb{R}_+ \times \mathbb{R})$ with compact support. In the sequel, all such functions are assumed to have a compact support contained in the subset $\Omega = [0, a) \times (-a, a)$ of \mathbb{R}^2 , where $a > 0$ is fixed.

For a 3D axially symmetric bounded object represented by a function $u(r, z)$ in cylindrical coordinates, where u denotes a function of $L^\infty(\mathbb{R}_+ \times \mathbb{R})$ with compact support contained in Ω , one defines $H_0u = H_0\tilde{u}$, where $\tilde{u}(x, y, z) = u(\sqrt{x^2 + y^2}, z)$ for all $x, y, z \in \mathbb{R}$, and where $H_0\tilde{u}$ is defined by (1). An obvious change of variable leads to

$$(H_0u)(y, z) = 2 \int_{|y|}^{+\infty} u(r, z) \frac{r}{\sqrt{r^2 - y^2}} dr, \quad (2)$$

for almost all $y, z \in \mathbb{R}$.

It is clear that the function H_0u is of compact support contained in $\tilde{\Omega} = (-a, a)^2$. In what follows, functions of compact support contained in Ω (resp., in $\tilde{\Omega}$) and their restriction to Ω (resp., in $\tilde{\Omega}$) are denoted similarly.

Lemma 1. *The linear operator H_0 extends to a linear operator (still denoted H_0) on $L^p(\Omega)$, for every $p \in [1, +\infty]$. Moreover,*

$$\|H_0u\|_{L^s(\tilde{\Omega})} \leq \left(\frac{4}{2-q}\right)^{\frac{1}{q}} a^{\frac{2}{q}-1} \|u\|_{L^p(\Omega)},$$

for every $u \in L^p(\Omega)$, and all $p \in [1, +\infty]$, $q \in [1, 2)$, with s such that $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{s}$.

In particular, $H_0 : L^p(\Omega) \rightarrow L^s(\tilde{\Omega})$ is a continuous linear operator, for every $p \in [1, +\infty]$ and every s such that $s \in [p, \frac{2p}{2-p})$ whenever $1 \leq p \leq 2$, and $s \in [p, +\infty]$ whenever $p > 2$.

Proof. Let $u \in L^\infty(\Omega)$. Note that $(H_0u)(y, z) = (H_0u)(-y, z)$, for almost all $y, z \in \mathbb{R}$, and that $(H_0u)(y, z) = 0$ whenever $|y| \geq a$ or $|z| \geq a$. For $y, z \in (-a, a)$, there holds

$$|(H_0u)(y, z)| \leq 2 \int_{|y|}^a |u(r, z)| \frac{r}{\sqrt{r+|y|}} \frac{1}{\sqrt{r-|y|}} dr \leq 2\sqrt{a} \int_{|y|}^a |u(r, z)| \frac{1}{\sqrt{r-|y|}} dr \leq f_z * g(|y|)$$

where the (nonnegative) functions f_z and g are defined by

$$f_z(r) = 2\sqrt{a}|u(r, z)|1_{(-a, a)}(r) \quad \text{and} \quad g(r) = \frac{1}{\sqrt{-r}}1_{(0, a)}(-r),$$

and where 1_A denotes the usual characteristic function of a Borelian subset $A \subset \mathbb{R}$, and $f_z * g$ denotes the standard convolution $f_z * g(y) = \int_{\mathbb{R}} f_z(r)g(y - r)dr$. It is clear that $g \in L^q(\mathbb{R})$, for every $q \in [1, 2)$, and $\|g\|_{L^q(\mathbb{R})} = a^{\frac{1}{q}-\frac{1}{2}}/(1 - \frac{q}{2})^{\frac{1}{q}}$. Since $u \in L^p(\Omega)$ for every $p \in [1, +\infty]$, it follows from Young's inequality (see for instance [3, Corollary 2.25 page 34] that the function $y \mapsto (H_0u)(y, z)$ belongs to $L^s(\mathbb{R})$ (and, actually, to $L^s(-a, a)$ since it is of compact support), for every $z \in (-a, a)$, where s is such that $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{s}$, and, moreover,

$$\int_{-a}^a |(H_0u)(y, z)|^s dy \leq \|f_z\|_{L^p(\mathbb{R})}^s \|g\|_{L^q(\mathbb{R})}^s,$$

for every $z \in (-a, a)$. Integrating with respect to z leads to

$$\|H_0u\|_{L^s(\tilde{\Omega})}^s \leq (2\sqrt{a})^s \frac{a^{\frac{1}{q}-\frac{1}{2}}}{(1 - \frac{q}{2})^{\frac{1}{q}}} \int_{-a}^a \left(\int_0^a |u(r, z)|^p dr \right)^{\frac{s}{p}} dz.$$

Using Hölder's inequality, the right-hand side of this inequality is bounded by

$$\int_{-a}^a \left(\int_0^a |u(r, z)|^p dr \right)^{\frac{s}{p}} dz \leq \|u\|_{L^p(\Omega)}^p (2a)^{1-\frac{s}{p}},$$

and the conclusion follows. \square

In particular, $H_0 : L^2(\Omega) \rightarrow L^2(\tilde{\Omega})$ is a continuous linear operator. Let $H_0^* : L^2(\tilde{\Omega}) \rightarrow L^2(\Omega)$ denote the adjoint operator of H_0 (i.e., the back-projection operator), for the pivot space L^2 , i.e., $\langle H_0u, v \rangle_{L^2(\tilde{\Omega})} = \langle u, H_0^*v \rangle_{L^2(\Omega)}$, for every $u \in L^2(\Omega)$ and every $v \in L^2(\tilde{\Omega})$. An obvious computation gives

$$(H_0^*v)(r, z) = 2 \int_{-r}^r v(y, z) \frac{r}{\sqrt{r^2 - y^2}} dy, \quad (3)$$

for every $(r, z) \in \Omega$. As previously, the next lemma holds.

Lemma 2. *The operator H_0^* extends to a continuous linear operator $H_0^* : L^{s'}(\tilde{\Omega}) \rightarrow L^{p'}(\Omega)$, for every $s' \in [1, +\infty]$ and every p' such that $p' \in [s', \frac{2s'}{2-s'})$ whenever $1 \leq s' \leq 2$, and $p' \in [s', +\infty]$ whenever $s' > 2$.*

The operator H_0 features the Radon transform of the object. Practically, one radiograph suffices to reconstruct the object. From the theoretical point of view, inverting the operator H_0 requires further differentiability. More precisely, the next lemma holds.

Lemma 3. Let $g \in L^2(\tilde{\Omega})$ such that $g(y, z) = g(-y, z)$ for a.e. $(y, z) \in \tilde{\Omega}$ and such that $\frac{\partial g}{\partial y}$ exists almost everywhere on $\tilde{\Omega}$ and is measurable and bounded. Then, there exists a unique $u \in L^\infty(\Omega)$ such that $H_0 u = g$, and

$$u(r, z) = -\frac{1}{\pi} \int_r^a \frac{\partial g}{\partial y}(y, z) \frac{1}{\sqrt{y^2 - r^2}} dy,$$

for every $(r, z) \in \Omega$.

Proof. Checking this formula is an easy exercise, noticing that, for all real numbers x and y such that $|y| \leq x$,

$$\int_{|y|}^x \frac{2r}{\sqrt{x^2 - r^2} \sqrt{r^2 - y^2}} dr = \int_0^1 \left(\sqrt{\frac{t}{1-t}} + \sqrt{\frac{1-t}{t}} \right) dt = \pi.$$

□

Without any ambiguity, we denote $u = H_0^{-1}g$, and this defines the linear operator H_0^{-1} . The arguments of the proof of Lemma 1 apply, and it is clear that H_0^{-1} extends to a continuous linear operator $H_0^{-1} : W^{1,p}(\tilde{\Omega}) \rightarrow L^s(\Omega)$, for every $p \in [1, +\infty]$ and every s such that $s \in [p, \frac{2p}{2-p})$ whenever $1 \leq p \leq 2$, and $s \in [p, +\infty]$ whenever $p > 2$.

Because of the derivative term, the operator H_0^{-1} cannot be extended as a continuous linear operator from $L^p(\Omega)$ to $L^q(\Omega)$ for suitable p and q . Concretely, this means that a small variation of the measure induces significant errors on the reconstruction. Since the radiographs at our disposal are strongly perturbed, applying H_0^{-1} thus provides a deficient and imperfect reconstruction of the original image. Moreover, due to the experimental setup, there are two additional main perturbations:

- A *blur*, due to the detector response and the X-ray source spot size. To simplify, it is assumed that the effect B of the blur is linear, and writes

$$B u_d = K * u_d, \tag{4}$$

where $*$ is the usual convolution operation, u_d is the projected image, and K is a positive symmetric kernel with compact support and such that $\int K d\mu = 1$.

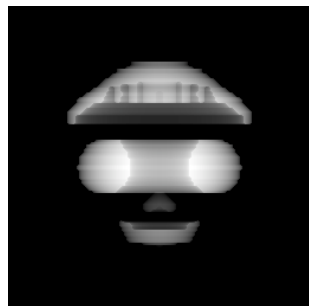
- A *noise*, assumed to be an additive Gaussian white noise, denoted τ , of zero mean and of standard deviation σ_τ .

Others perturbations, such as scattered field or motion blur, are not taken into account in our study. With these assumptions, the projection of an object u is

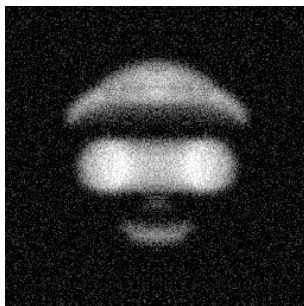
$$v_d = B H_0 u + \tau.$$

A comparison between the theoretical projection $B H_0 u$ and the perturbed one is provided on Figures 3(a) and 3(b). The real object u is drawn on Figure 3(c). The reconstruction

using the inverse operator H_0^{-1} applied to v_d is drawn on Figure 3(d). The purpose of the experiment is to separate the material from the empty holes and thus to determine precisely the boundary between the two areas. This task is difficult to perform on the reconstruction $H_0^{-1}v_d$, and an inspection of Figure 3(d) shows that the use of the inverse operator is not suitable.



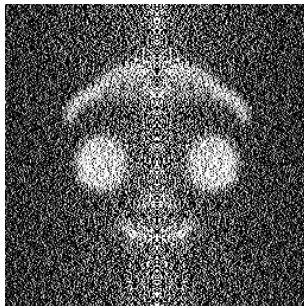
(a) Theoretical projection BH_0u of the object of Figure 2.



(b) Real projection $v_d = BH_0u + \tau$ of the same object with realistic noise and blur.



(c) The real object u .



(d) Reconstruction $H_0^{-1}v_d$ computed with H_0^{-1} applied to the real projection.

Figure 3: Comparison of u , BH_0u , $v_d = BH_0u + \tau$, $H_0^{-1}v_d$.

A useful tool to deal with ill-posed problems is a regularization process based on optimization methods. In the next section, we propose a variational approach for tomographic reconstruction, based on a minimization problem settled in the space of bounded variation functions. A penalization procedure permits to derive an optimality system. These conditions are then implemented, and numerical simulations are provided in Section 4.

3 A variational approach for tomographic reconstruction

3.1 Minimization problem in $BV(\Omega)$

The most suitable functional space used in image restoration is the space $BV(\Omega)$ of bounded variation functions (see [5]), defined by

$$BV(\Omega) = \{u \in L^1(\Omega) \mid \Phi(u) < +\infty\},$$

where

$$\Phi(u) = \sup \left\{ \int_{\Omega} u(x) \operatorname{div} \xi(x) dx \mid \xi \in \mathcal{C}_c^1(\Omega), \|\xi\|_{\infty} \leq 1 \right\}. \quad (5)$$

The space $BV(\Omega)$, endowed with the norm $\|u\|_{BV(\Omega)} = \|u\|_{L^1} + \Phi(u)$, is a Banach space. The derivative in the sense of the distributions of every $u \in BV(\Omega)$ is a bounded Radon measure, denoted Du , and $\Phi(u) = \int_{\Omega} |Du|$ is the total variation of Du . We next recall standard properties of bounded variation functions (see [2, 4]).

Proposition 1. [2, 4] *Let Ω be an open subset of \mathbb{R}^2 with Lipschitzian boundary.*

1. *For every $u \in BV(\Omega)$, the Radon measure Du can be decomposed into $Du = Du dx + D^s u$, where $Du dx$ is the absolutely continuous part of Du with respect of the Lebesgue measure and $D^s u$ is the singular part.*
2. *The mapping $u \mapsto \Phi(u)$ is lower semi-continuous (denoted in short *lsc*) from $BV(\Omega)$ to \mathbb{R}^+ for the $L^1(\Omega)$ topology.*
3. *$BV(\Omega) \subset L^2(\Omega)$ with continuous embedding.*
4. *$BV(\Omega) \subset L^p(\Omega)$ with compact embedding, for every $p \in [1, 2)$.*

Assume that the kernel K modelling the blur has a compact support contained in $\tilde{\Omega}$. Then, the operator B defined by (4) is a continuous linear operator from $L^p(\tilde{\Omega})$ to $L^q(2\tilde{\Omega})$, for all $p, q \in [1, +\infty]$, where $2\tilde{\Omega} = (-2a, 2a)^2$.

Let $v_d \in L^2(\tilde{\Omega})$ be the projected image (observed data), and let $\alpha > 0$. Define $H = BH_0$, and consider the minimization problem

$$(\mathcal{P}) \quad \begin{cases} \min F(u), & \text{with } F(u) = \frac{1}{2} \|Hu - v_d\|_2^2 + \alpha \Phi(u), \\ u \in BV(\Omega), \\ u(x) \in \{0, 1\} \text{ a.e. on } \Omega, \end{cases}$$

where $\|\cdot\|_2$ stands for the $L^2(\tilde{\Omega})$ norm.

Remark 1. A similar problem has been studied in [9] with a smoother projection operator and *convex* constraints. Here, the pointwise constraint, $u(x) \in \{0, 1\}$ a.e. on Ω , is a very hard constraint. The constraint set is not convex and its interior is empty for most usual topologies.

Theorem 1. *The minimization problem (\mathcal{P}) admits at least a solution.*

Proof. The proof is straightforward. Let (u_n) be a minimizing sequence of $BV(\Omega)$, satisfying $u_n(x) \in \{0, 1\}$ a.e. on Ω . Then, the sequence $(\Phi(u_n))$ is bounded. Due to the constraint, and since Ω is bounded, it follows that the sequence (u_n) is bounded in $BV(\Omega)$, and hence, up to a subsequence, it converges to some $u \in BV(\Omega)$ for the weak-star topology. The compact embedding property recalled in Proposition 1 implies that the sequence (u_n) converges strongly to u in $L^p(\Omega)$, for every $p \in [1, 2)$. It follows from Lemma 1 that $H : L^{3/2}(\Omega) \rightarrow L^2(2\tilde{\Omega})$ is continuous, and hence the sequence (Hu_n) converges strongly to Hu in $L^2(2\tilde{\Omega})$. Moreover, since (u_n) converges strongly to u in $L^1(\Omega)$ and since Φ is lsc with respect to the $L^1(\Omega)$ topology, it follows that

$$\Phi(u) \leq \liminf_{n \rightarrow \infty} \Phi(u_n).$$

Hence,

$$\inf F = \lim_{n \rightarrow +\infty} \left(\frac{1}{2} \|Hu_n - v_d\|_{L^2}^2 + \alpha \Phi(u_n) \right) \geq \|Hu - v_d\|_{L^2}^2 + \alpha \Phi(u) = F(u).$$

Finally, (u_n) converges to u in $L^1(\Omega)$, and thus, converges almost everywhere (up to a subsequence) to u . Hence, the pointwise constraint $u(x) \in \{0, 1\}$ is satisfied almost everywhere, and therefore u is a solution of (\mathcal{P}) . \square

3.2 Penalization of the minimization problem (\mathcal{P})

To deal with the binarity constraint, we use a penalization method. Let $\varepsilon > 0$, $\beta \geq 0$, and let \bar{u} be a solution of (\mathcal{P}) . Define

$$J_\varepsilon(u) = \frac{1}{2} \|Hu - v_d\|_2^2 + \frac{1}{2\varepsilon} \|u - u^2\|_2^2 + \frac{\beta}{2} \|u - \bar{u}\|_2^2,$$

and

$$F_\varepsilon(u) = F(u) + \frac{1}{2\varepsilon} \|u - u^2\|_2^2 + \frac{\beta}{2} \|u - \bar{u}\|_2^2.$$

Remark 2. The term $\beta \|u - \bar{u}\|_2^2$ is an additional penalization term permitting to focus on a particular solution \bar{u} of (\mathcal{P}) . In practice, the solution \bar{u} is of course not known and we choose $\beta = 0$.

It follows from Lemma 1 that $Hu \in L^2(2\tilde{\Omega})$ whenever $u \in L^{1+s}(\Omega)$, for every $s > 0$. In addition, for $F_\varepsilon(u)$ to be well defined, it is required that $u \in L^4(\Omega)$. We consider the penalized problem

$$(\mathcal{P}_\varepsilon) \quad \begin{cases} \min F_\varepsilon(u) \\ u \in BV(\Omega) \cap \mathcal{B}_R \end{cases}$$

where $R > 1$ is a fixed real number, and

$$\mathcal{B}_R = \{u \in L^\infty(\Omega) \mid \|u\|_\infty \leq R\}.$$

In the sequel we do not mention the dependence of $(\mathcal{P}_\varepsilon)$ with respect to the real number R that can be chosen as large as desired but is fixed. A contrario the parameter ε will tend to 0. Note that the constraint $u \in \mathcal{B}_R$ is required theoretically to ensure convergence properties, however it does not affect the numerical process. The following lemma will be useful.

Lemma 4. *Let $(u_n)_{n \in \mathbb{N}}$ be a bounded sequence of $BV(\Omega) \cap \mathcal{B}_R$. Then, there exist $u^* \in BV(\Omega) \cap \mathcal{B}_R$ and a subsequence of $(u_n)_{n \in \mathbb{N}}$ converging to u^* for the strong $L^p(\Omega)$ topology, for every $p \in [1, +\infty)$.*

Proof. Since the sequence $(u_n)_{n \in \mathbb{N}}$ is bounded in $BV(\Omega)$, it converges, up to a subsequence, to some $u^* \in BV(\Omega)$ for the weak-star topology, and hence for the strong $L^1(\Omega)$ topology. Since it is moreover bounded (by R) in $L^\infty(\Omega)$, the conclusion follows from the Lebesgue dominated convergence theorem. In addition, \mathcal{B}_R is compact for the L^∞ weak star topology, hence $u^* \in \mathcal{B}_R$. \square

Theorem 2. *The minimization problem $(\mathcal{P}_\varepsilon)$ has at least a solution $u_\varepsilon \in BV(\Omega) \cap \mathcal{B}_R$.*

Proof. To prove existence, consider a minimizing sequence $(u_n)_n$ of $BV(\Omega) \cap \mathcal{B}_R$. Lemma 4 yields that (u_n) converges (up to a subsequence) to some $\tilde{u} \in BV(\Omega) \cap \mathcal{B}_R$, weakly in $BV(\Omega)$ and strongly in $L^p(\Omega)$ for every $p \in [1, +\infty)$. Since J_ε is continuous for the L^2 topology, and Φ is lsc for the L^1 topology, one gets

$$F_\varepsilon(\tilde{u}) \leq \liminf_n F_\varepsilon(u_n) \leq \inf F_\varepsilon,$$

and the conclusion follows. \square

Theorem 3. *1. Every cluster point u^* in $BV(\Omega) \cap L^p(\Omega)$ (for $p \in [1, +\infty)$) of the family (u_ε) at $\varepsilon = 0$ is a solution of (\mathcal{P}) . If moreover $\beta > 0$ then $u^* = \bar{u}$.*

2. There holds $\lim_{\varepsilon \rightarrow 0} F_\varepsilon(u_\varepsilon) = \inf F$, and $\lim_{\varepsilon \rightarrow 0} \int_\Omega |Du_\varepsilon| = \int_\Omega |Du|$.

Proof. Since $\bar{u} \in BV(\Omega)$ is a solution of (\mathcal{P}) , one has

$$F_\varepsilon(u_\varepsilon) \leq F_\varepsilon(\bar{u}) = F(\bar{u}) = \inf F,$$

for every $\varepsilon > 0$. Therefore, the family (u_ε) is bounded in $BV(\Omega) \cap \mathcal{B}_R$, and $\|u_\varepsilon - u_\varepsilon^2\|_2 \rightarrow 0$. Let u^* be a (strong) cluster point of (u_ε) in $L^p(\Omega)$ (for $p \in [1, +\infty)$). Then,

$$\|u^* - u^{*2}\|_2 \leq \liminf_{\varepsilon \rightarrow 0} \|u_\varepsilon - u_\varepsilon^2\|_2 = 0,$$

so that $u^*(1 - u^*) = 0$ a.e. on Ω . Since

$$F(u_\varepsilon) + \beta \|u_\varepsilon - \bar{u}\|_2^2 \leq F_\varepsilon(u_\varepsilon) \leq \inf F,$$

one gets

$$F(u^*) \leq F(u^*) + \beta \|u^* - \bar{u}\|_2^2 \leq \inf F.$$

Therefore u^* is a solution of (\mathcal{P}) . In addition, if $\beta > 0$, then $u^* = \bar{u}$. Finally, since

$$\inf F = F(u^*) \leq F_\varepsilon(u^*) \leq \liminf F_\varepsilon(u_\varepsilon) \leq \inf F,$$

and $F_\varepsilon(u_\varepsilon) \leq \inf F$, it follows that $\lim_{\varepsilon \rightarrow 0} F_\varepsilon(u_\varepsilon) = \inf F$. Moreover, writing

$$\liminf F(u_\varepsilon) + \limsup \frac{1}{\varepsilon} \|u_\varepsilon - u_\varepsilon^2\|_2^2 \leq \lim (F(u_\varepsilon) + \frac{1}{\varepsilon} \|u_\varepsilon - u_\varepsilon^2\|_2^2) = F(u^*) \leq \liminf F(u_\varepsilon),$$

it follows that

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \|u_\varepsilon - u_\varepsilon^2\|_2^2 = 0,$$

and then that

$$\lim_{\varepsilon \rightarrow 0} F(u_\varepsilon) = F(u^*).$$

The conclusion follows then from the continuity properties of H . \square

3.3 Optimality system of the penalized minimization problem $(\mathcal{P}_\varepsilon)$

We first recall a result of [9] useful to derive optimality conditions for the penalized problem $(\mathcal{P}_\varepsilon)$.

Theorem 4 ([9, Theorem 2.3]). *Let Ω be a Borelian subset of \mathbb{R}^n . Let $\bar{u} \in K \cap BV(\Omega)$ be the solution of*

$$\begin{cases} \min \mathcal{J}(u) + \alpha \int_{\Omega} |Du|, \\ u \in K \cap BV(\Omega), \end{cases}$$

where K is a closed convex subset of $L^p(\Omega)$ and \mathcal{J} is continuous and Gâteaux differentiable from $L^p(\Omega)$ to \mathbb{R} ($1 \leq p < +\infty$), and either K is bounded or \mathcal{J} is coercive. Then, there exists $\bar{\lambda} \in (\mathcal{M}(\Omega)^n)'$ (the dual space of Radon measures) such that

$$\forall u \in K \cap BV(\Omega) \quad \langle \mathcal{J}'(\bar{u}) - \alpha \operatorname{div} \bar{\lambda}, u - \bar{u} \rangle \geq 0, \quad (6)$$

$$\forall \mu \in (\mathcal{M}(\Omega))^n \quad \langle \bar{\lambda}, \mu - D\bar{u} \rangle + \int_{\Omega} |D\bar{u}| \leq \int_{\Omega} |\mu|, \quad (7)$$

where $D : BV(\Omega) \rightarrow (\mathcal{M}(\Omega))^n$ and

$$\forall u \in BV(\Omega) \quad \langle \operatorname{div} \bar{\lambda}, u \rangle = - \langle \bar{\lambda}, Du \rangle. \quad (8)$$

This result cannot be applied to the original problem since the constraints set is not convex, but can be used to handle the penalized problem. It yields the existence of $\lambda_\varepsilon \in (\mathcal{M}(\Omega)^2)'$ such that

$$\forall u \in BV(\Omega) \cap \mathcal{B}_R \quad \langle J'_\varepsilon(u_\varepsilon) - \alpha \operatorname{div} \lambda_\varepsilon, u - u_\varepsilon \rangle \geq 0, \quad (9)$$

and

$$\forall \mu \in (\mathcal{M}(\Omega))^2 \quad \langle \lambda_\varepsilon, \mu - Du_\varepsilon \rangle + \int_\Omega |Du_\varepsilon| \leq \int_\Omega |\mu|. \quad (10)$$

The functional J_ε is differentiable in $L^2(\Omega)$, and

$$J'_\varepsilon(u_\varepsilon) = H^*(Hu_\varepsilon - v_d) + q_\varepsilon,$$

where

$$q_\varepsilon = \frac{1}{\varepsilon} (2u_\varepsilon - 1)(u_\varepsilon^2 - u_\varepsilon) + \beta(u_\varepsilon - \bar{u}). \quad (11)$$

Since $u_\varepsilon \in L^\infty(\Omega)$, there holds $q_\varepsilon \in L^\infty(\Omega)$. Considering $\mu = Dv$ with $v \in BV(\Omega)$ in equation (10) leads to

$$\forall v \in BV(\Omega) \quad \langle \lambda_\varepsilon, D(v - u_\varepsilon) \rangle + \int_\Omega |Du_\varepsilon| \leq \int_\Omega |Dv|,$$

i.e.,

$$\forall v \in BV(\Omega) \quad \Phi(v) \geq \Phi(u_\varepsilon) - \langle \operatorname{div} \lambda_\varepsilon, v - u_\varepsilon \rangle,$$

which is equivalent to

$$\mu_\varepsilon \in \partial\Phi(u_\varepsilon),$$

where $\mu_\varepsilon = -\operatorname{div} \lambda_\varepsilon$.

Theorem 5. *Let u_ε be a solution of $(\mathcal{P}_\varepsilon)$. Then there exist $\lambda_\varepsilon \in (\mathcal{M}(\Omega)^2)'$, $q_\varepsilon \in L^\infty(\Omega)$ and $\mu_\varepsilon = -\operatorname{div} \lambda_\varepsilon$ such that*

$$\forall u \in BV(\Omega) \cap \mathcal{B}_R \quad \langle H^*(Hu_\varepsilon - v_d) + q_\varepsilon + \alpha\mu_\varepsilon, u - u_\varepsilon \rangle \geq 0, \quad (12a)$$

$$\mu_\varepsilon \in \partial\Phi(u_\varepsilon), \quad (12b)$$

3.4 Optimality system of the minimization problem (\mathcal{P})

In order to derive an optimality system for the minimization problem (\mathcal{P}) , it is natural to attempt to pass to the limit in (12a)-(12b). There is however a difficulty explained next.

Let u^* be a cluster point in $BV(\Omega) \cap L^p(\Omega)$ (for $p \in [1, +\infty)$) of the family (u_ε) at $\varepsilon = 0$. From Theorem 3, u^* is a solution of (\mathcal{P}) .

The family (λ_ε) shares nice asymptotic properties, as shown in the following lemma.

Lemma 5. *The family (λ_ε) is uniformly bounded in $(BV(\Omega) \cap L^r(\Omega))^2$ with respect to ε , for every $r \in [1, +\infty]$. The family $(\operatorname{div} \lambda_\varepsilon)$ is uniformly bounded in $(BV(\Omega))'$ (and thus in $H^{-1}(\Omega)$) with respect to ε . Moreover, for every weak cluster point λ^* of (λ_ε) at $\varepsilon = 0$ in $(BV(\Omega) \cap L^r(\Omega))^2$, $\operatorname{div} \lambda^*$ is a weak star cluster point of $(\operatorname{div} \lambda_\varepsilon)$ at $\varepsilon = 0$ in $(BV(\Omega))'$ (and in $H^{-1}(\Omega)$).*

Proof. Let $v \in BV(\Omega)$ and $\mu = D(v + u_\varepsilon) \in \mathcal{M}(\Omega)^2$. From (10), we infer that

$$\langle \lambda_\varepsilon, Dv \rangle + \int_{\Omega} |Du_\varepsilon| \leq \int_{\Omega} |Dv| + \int_{\Omega} |Du_\varepsilon|,$$

and thus, $-\langle \operatorname{div} \lambda_\varepsilon, v \rangle \leq \int_{\Omega} |Dv| \leq \|v\|_{BV(\Omega)}$. It follows that

$$\sup_{v \in BV(\Omega), \|v\|_{BV(\Omega)} \leq 1} \langle \operatorname{div} \lambda_\varepsilon, v \rangle \leq 1.$$

Therefore, $(\operatorname{div} \lambda_\varepsilon)$ is uniformly bounded in the dual of $BV(\Omega)$ and (λ_ε) is uniformly bounded in $(BV(\Omega))^2$ with respect to ε .

Now, let $r \in (1, +\infty]$, and let $\varphi \in (L^{r'})^2$ with $r' = \frac{r}{r-1} \in [1, +\infty)$. Choose $v = \varphi + Du_\varepsilon \in \mathcal{M}(\Omega)^2$. Then, we infer from (10) that

$$\langle \lambda_\varepsilon, \varphi \rangle_{r,r'} \leq \|\varphi\|_1 \leq C \|\varphi\|_{r'},$$

and hence (λ_ε) is uniformly bounded in $L^r(\Omega)^2$. In particular it is bounded in $L^1(\Omega)$.

Finally, it follows from the relation

$$\forall u \in BV(\Omega) \quad \langle \operatorname{div} \lambda_\varepsilon, u \rangle = -\langle \lambda_\varepsilon, Du \rangle,$$

that, if λ^* is a weak cluster point of (λ_ε) at $\varepsilon = 0$ in $(BV(\Omega) \cap L^r(\Omega))^2$, then $\operatorname{div} \lambda^*$ is a weak star cluster point of $(\operatorname{div} \lambda_\varepsilon)$ at $\varepsilon = 0$ in $(BV(\Omega))'$ (and in $H^{-1}(\Omega)$). \square

The difficulty is that it is not possible to bound the family (q_ε) uniformly (in $H^{-1}(\Omega)$) with respect to ε . Indeed if it were possible then one would be able to derive a ‘‘classical’’ optimality system, formally written as

$$u^*(x) \in \{0, 1\} \text{ a.e. in } \Omega, \tag{13a}$$

$$H^*(Hu^* - v_d) + (2u^* - 1)s^* - \alpha \operatorname{div} \lambda^* = 0 \text{ in } H^{-1}(\Omega), \tag{13b}$$

$$\forall \mu \in (\mathcal{M}(\Omega))^2 \quad \langle \lambda^*, \mu - Du^* \rangle + \int_{\Omega} |Du^*| \leq \int_{\Omega} |\mu|. \tag{13c}$$

In particular, $s^* := \frac{q^*}{2u^* - 1}$ would be a Lagrange multiplier associated to the constraint $u^*(x) \in \{0, 1\}$. This is not possible since no qualification condition is available for this kind of constraints (see [7] where counterexamples are provided).

Nevertheless, though it would be theoretically satisfying to get such a limit optimality system, it is not directly useful from the numerical point of view and we use the penalized one.

4 Numerical simulations

4.1 Resolution of the penalized minimization problem (\mathcal{P}_ε)

We use the penalized optimality system (12a-12b), and have to deal with two parameters ε and α . Since the solution of (\mathcal{P}) is unknown, we choose $\beta = 0$. Noticing that equation (12a) is formally equivalent to

$$u_\varepsilon = P_{\mathcal{B}_R}(u_\varepsilon - H^*(Hu_\varepsilon - v_d) - q_\varepsilon - \alpha\mu_\varepsilon),$$

where P_K is a projection on K (with respect to a topology to be precised), we propose the following fixed-point type algorithm.

Algorithm (ε fixed)

1. Initialization : $n = 0$, choose u^0 .
2. Solve the variational inequality (12b) to find $\mu^{n+1} \in \partial\Phi(u^n)$.
3. Solve

$$u^{n+1} = P_{\mathcal{B}_R}(u^n - H^*(Hu^n - v_d) - q(u^n) + \alpha\mu^{n+1}),$$

with

$$q(u) = \frac{1}{\varepsilon} (2u - 1)(u^2 - u).$$

4. Stopping criterion : stop or set $n = n + 1$ and goto 2.

The discretization process is standard (see for instance [5]). The discretized image is represented by a $N \times N$ array identified with a N^2 vector. Due to the symmetry, it suffices to deal with half an image (of size $N \times N/2$). Denote $X = \mathbb{R}^{N \times N}$ and $Y = X \times X$, endowed with the usual scalar product $(u, v)_X = \sum_{1 \leq i, j \leq N} u_{ij}v_{ij}$. For $g = (g^1, g^2) \in Y$, denote

$$|g_{i,j}| = \sqrt{(g_{i,j}^1)^2 + (g_{i,j}^2)^2}.$$

The Radon measure Du is approximated as follows. For $u \in X$, Du is identified with a vector of Y of coordinates $(Du)_{i,j} = ((Du)_{i,j}^1, (Du)_{i,j}^2)$ defined by

$$(Du)_{i,j}^1 = \begin{cases} u_{i+1,j} - u_{i,j} & \text{if } i < N \\ 0 & \text{if } i = N \end{cases} \quad (Du)_{i,j}^2 = \begin{cases} u_{1,j+1} - u_{i,j} & \text{if } j < N \\ 0 & \text{if } j = N \end{cases}$$

The total variation is then approximated by $\Phi(u) = \sum_{1 \leq i, j \leq N} |(Du)_{i,j}|$.

Concerning the divergence operator $\text{div} = -D^*$, where D^* is the adjoint operator of D , one has

$$\forall p \in Y, \forall u \in X \quad (-\text{div } p, u)_X = (p, Du)_Y = (p^1, D^1u)_X + (p^2, D^2u)_X,$$

It is standard to discretize $\operatorname{div} p$ with

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

Resolution of Step 2. The choice of $\mu \in \partial\Phi(u)$ follows Chambolle's method (see [10]). It is known that the Fenchel-Legendre conjugate function Φ^* of Φ is the indicatrix function $1_{\mathcal{K}}$ of

$$\mathcal{K} = \{\operatorname{div} g \mid g \in Y, |g_{i,j}| \leq 1, \forall i, j\}.$$

Moreover,

$$\mu \in \partial\Phi(u) \Leftrightarrow u \in \partial 1_{\mathcal{K}}(\mu) \Leftrightarrow \mu = \Pi_{\mathcal{K}}(\mu + u)$$

(see for instance [10]), where $\Pi_{\mathcal{K}}$ denotes the orthogonal projection on \mathcal{K} . Therefore, μ can be computed with the successive approximation process

$$\mu_k = \Pi_{\mathcal{K}}(\mu_{k-1} + u),$$

or with a semi-smooth Newton method. The projected element

$$\Pi_{\mathcal{K}}(v) = \operatorname{argmin} \{\|\operatorname{div} p - v\|_X^2 \mid p_{i,j} \leq 1, i, j = 1, \dots, N\},$$

can be computed as in [10], using the iteration process

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n + \rho (D(\operatorname{div} p^n - v))_{i,j}}{1 + \rho |(D(\operatorname{div} p^n - v))_{i,j}|}.$$

If $\rho \leq 1/8$ then $\operatorname{div} p^n \rightarrow \Pi_{\mathcal{K}}(v)$.

Resolution of Step 3. Newton methods are not adapted to solve Step 3. Indeed, the matrix H^*H is ill conditioned and the polynomial function

$$\psi : t \mapsto (2t - 1)(t^2 - t) \tag{14}$$

has an attractive zero at $t = 0.5$ that must be avoided. We rather use a projected gradient method, which is however quite slow. The use of an optimal step strategy does not seem to improve significantly the results. In addition we note that u^n should converge to the solution which is "close" to 0 or 1, so that the projection inequality of Step 3 is expected to be an equality. We use the following algorithm where the projection step is decomposed.

Algorithm (ε fixed)

1. Initialization: $n = 0$, choose $u^0 = H^{-1}(v_d)$.

2. Find $\mu^{n+1} \in \partial\Phi(u^n)$.
3. Set $q^{n+1} = -H^*(Hu^n - v_d) - \alpha\mu^{n+1}$,
4. Solve $\psi(\tilde{u}^n) = \varepsilon q^{n+1}$ and set $u_{n+1} = P_{[0,1]}(\tilde{u}^n)$ as the solution, with a projected gradient method, where ψ is given by (14).
5. Stopping criterion: stop or set $n = n + 1$ and goto 2.

We did not perform any convergence analysis of this algorithm, neither within the continuous framework nor in the discrete one. This issue will be investigated in some future work. Denoting similarly the discretized and the continuous unknowns, we provide below an elementary result asserting that, if the discretization algorithm converges, then the limit is the optimal solution.

Lemma 6. *Let $\varepsilon > 0$ fixed. If the sequence (u^n) of the above algorithm converges to some $u^\varepsilon \in \mathbb{R}^N$, then u^ε is solution of the discretized optimality system of $(\mathcal{P}_\varepsilon)$.*

Proof. Assume that (u^n) converges to u^ε in \mathbb{R}^N (where N is the dimension of the discretization space, i.e., the size of the image). From the projection step, one has $0 \leq u^\varepsilon \leq 1$. By continuity of ψ and step 3, q^n converges to some q^ε such that $\psi(u^\varepsilon) = \varepsilon q^\varepsilon$. Similarly, step 2 yields the convergence of μ^n to μ^ε such that $q^\varepsilon = -H^*(Hu^\varepsilon - v_d) - \alpha\mu^\varepsilon$. Since

$$\mu^{n+1} \in \partial\Phi(u^n) \Leftrightarrow \mu^{n+1} = \Pi_{\mathcal{K}}(\mu^{n+1} + u^n),$$

passing to the limit as $n \rightarrow +\infty$ yields $\mu^\varepsilon = \Pi_{\mathcal{K}}(\mu^\varepsilon + u^\varepsilon)$, i.e., $\mu^\varepsilon \in \partial\Phi(u^\varepsilon)$, with $H^*(Hu^\varepsilon - v_d) + \alpha\mu^\varepsilon + q^\varepsilon = 0$. \square

4.2 Numerical results

We present numerical results in two cases:

- first case: with no blur, $B = I$;
- second case: there is a blur, modeled by a centered Gaussian filter with standard deviation $\sigma = 3$.

In both cases, the projected image (observed data) is perturbed with a Gaussian noise τ with standard deviation $\sigma_\tau = 0.2$ (the image has been rescaled between 0 and 1),

$$\tau(x) = \frac{1}{\sqrt{2\pi}\sigma_\tau} e^{-\frac{|x|^2}{2\sigma_\tau^2}}.$$

In the second case, the projected image is moreover perturbed with a Gaussian blur with standard deviation $\sigma_B = 3$, so that the observed data is

$$v_d = (K * H_0)(u_{orig}) + \tau,$$

where

$$K(x) = Ce^{-\frac{|x|^2}{2\sigma_B^2}} 1_{\tilde{\Omega}}(x).$$

where C is a normalizing constant so that $\int K d\mu = 1$.

The polynomial function ψ has three zeros 0, 1/2, and 1 which have the same multiplicity. However this function is quite "flat" so that the Newton method fails but the projected gradient method works well with a small descent step. This part of our algorithm may however be improved.

The descent step of the gradient method was set to 10^{-5} . With this value, it happens that convergence occurs in our numerical tests for any α and ε . The step is small, in accordance with the constraint $u \in [0, 1]$. Numerical refinements such as the determination of an optimal step do not seem to improve significantly the results.

The gradient algorithm was limited to $it_{max} = 2000$ iterations. The convergence is quite slow and it is necessary to perform enough iterations. Nevertheless it is not necessary to compute the exact solution at each step. Table 3 provides some numerical results with several values of the step ρ of the gradient method. The algorithm consists in two imbricated loops: the main loop consists in computing u^n , and the inside loop concerns the gradient method (with at most it_{max} iterations) in step 4.

Concerning the stopping criterion, a first possibility is to stop the algorithm as soon as $\|u^{n+1} - u^n\|_\infty$ is small enough. However, the convergence happens to be nonmonotonic in the sense that $\|u^{n+1} - u^n\|_\infty$ may oscillate in some cases while the cost function is decreasing. Therefore, we chose the stopping criterion

$$\delta_\varepsilon^n := \left| \frac{\mathbf{F}_\varepsilon(u^{n+1}) - \mathbf{F}_\varepsilon(u^n)}{\mathbf{F}_\varepsilon(u^n)} \right| \leq \text{tol},$$

where \mathbf{F}_ε is the discretized version of F_ε , and tol was set to 10^{-3} . This stopping criterion is based on the relative error between two consecutive values of the cost functional. Results are provided on Table 2 for different values of tol .

The resolution of our test images was 256×256 pixels, and computations were performed using MATLAB[®] software on a MAC-G4 computer (1.5 GHz).

Remark 3. It is certainly possible to improve the performances of our algorithm with numerical refinements. The purpose of this article is however not to provide sharp optimized numerical algorithms, but to test an approach based on a penalization method and investigate its numerical efficiency.

The algorithm behaves like a descent method (see Figures 4 and 5), but the convergence of u^n is not monotonic.

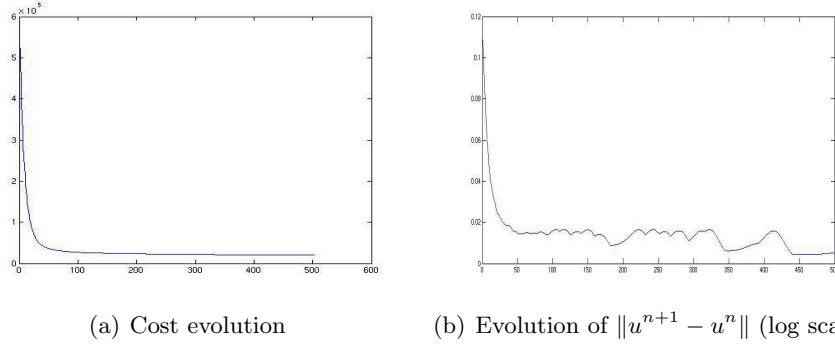


Figure 4: Case without blur: $\varepsilon = 0.5$, $\alpha = 15$, $\text{tol} = 10^{-4}$.

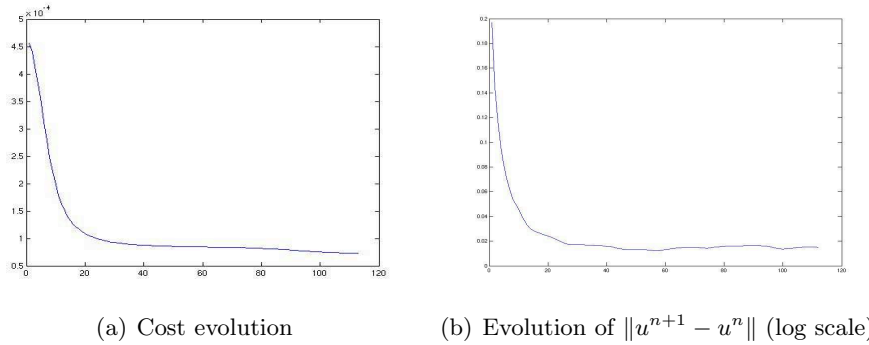


Figure 5: Case with blur: $\varepsilon = 0.5$, $\alpha = 30$, $\text{tol} = 10^{-3}$.

On these figures, one can observe that the functional F_ε strongly decreases at small times and then decreases slowly. Actually, the image is considerably improved after few iterations, far from the symmetry axis (see Figure 8). This is due to the fact that the outermost pixels of the image carry more information than the innermost pixels. Indeed, since the object is axially symmetric, an outer pixel generates, when rotating around the symmetry axis z , a torus, which projects onto a strip on a plane containing the axis z . This strip is of course longer for an outer pixel than for an inner one, thus carrying more weight in the functional to be minimized in the iteration process. This is the reason why the convergence is slow for pixels around the axis, but very fast for outermost pixels.

Figure 6 represents, in the case without blur, the original image, the observed image, and the initial point of the gradient method.

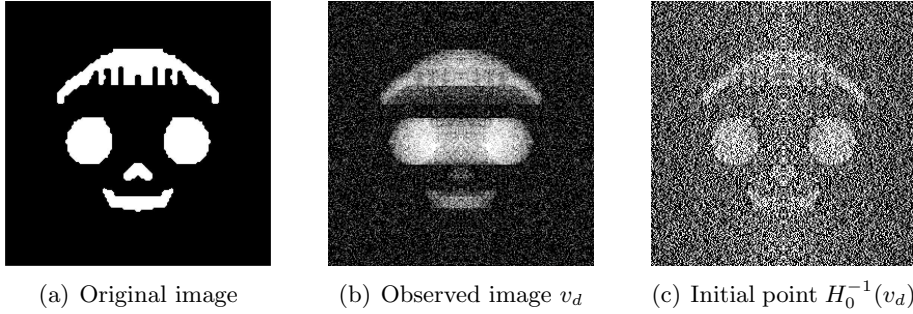


Figure 6: Case without blur: original image, observed image, and initial guess.

Table 1 and Figure 7 provide different results with respect to ε and α , in the case without blur.

ε	α	# iterations	F_ε	$\ u^{n+1} - u^n\ _\infty$	δ_ε^n	CPU time (s)
0.1	1	16	9.341691e-01	6.277e-02	9.658e-04	1 972
0.1	10	157	1.789839	5.666e-02	9.883e-04	17 580
0.5	1	136	3.733188e-01	1.212e-02	9.903e-04	17 190
0.5	5	259	2.051840e-01	1.444e-02	9.947e-04	26 570
0.5	10	195	2.861554e-01	1.445e-02	9.918e-04	18 060
0.5	20	207	4.128942e-01	1.608e-02	9.955e-04	18 240
1	1	292	1.964433e-01	6.244e-03	9.942e-04	38 990
1	10	184	2.861960e-01	7.003e-03	9.931e-04	14 300

Table 1: Case without blur: sensitivity with respect to ε and α , with $tol = 1e - 03$ and $it_{max} = 2000$.

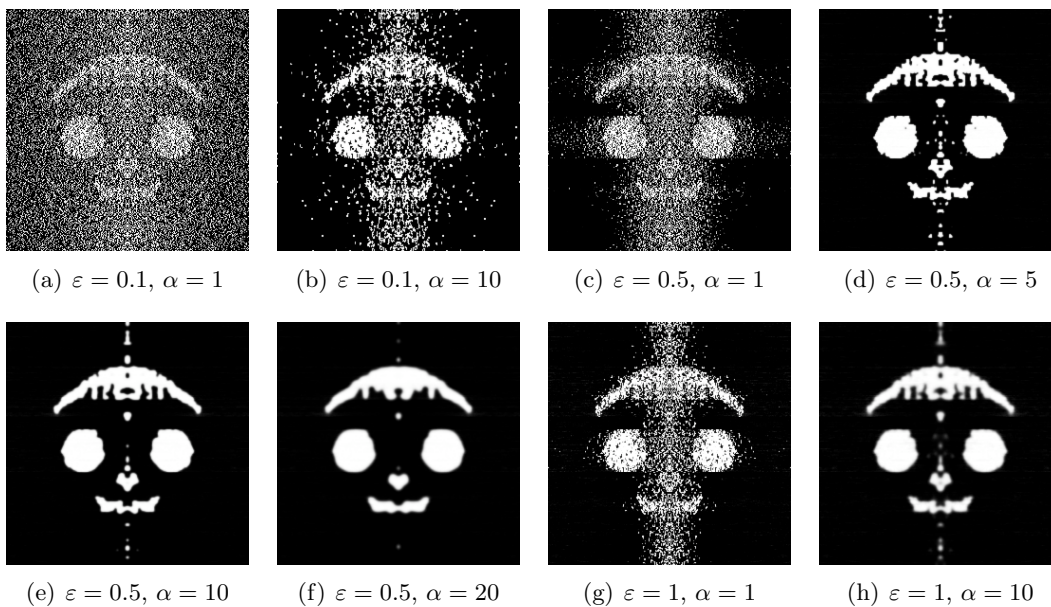


Figure 7: Case without blur, for different values of ε and α .

Table 2 and Figure 8 provide different results with respect to tol , in the case without blur.

tol	# iterations	F_ε	$\ u^{n+1} - u^n\ _\infty$	δ_ε^n	CPU time (s)
1e-01	6	4.852204	7.388e-02	9.952e-02	569
1e-02	56	5.197695e-01	1.453e-02	9.967e-03	4398
1e-03	173	3.659651e-01	1.327e-02	9.930e-04	14 395
1.33e-04	500	3.1302e-01	5.2026e-03	1.3314e-04	50 615

Table 2: Case without blur: sensitivity with respect to tol , with $\varepsilon = 0.5$, $\alpha = 15$, and $it_{max} = 2000$.

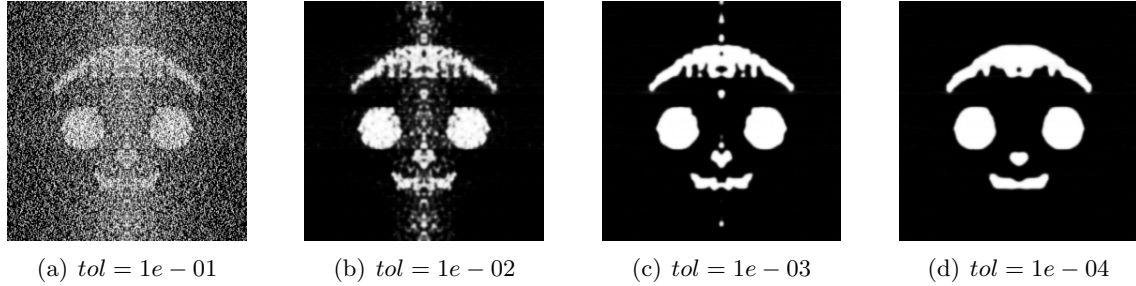


Figure 8: Case without blur, for different values of tol , with $\varepsilon = 0.5$, $\alpha = 15$, and $it_{max} = 2000$.

Table 3 and Figure 9 provide numerical results in the case without blur, for different values of the gradient method step ρ .

ρ	# iterations	F_ε	$\ u^{n+1} - u^n\ _\infty$	δ_ε^n	CPU time
1e-03	5	1.806698e+01	1	8.636e-03	672
5e-04	6	1.795947e+01	1	8.082e-03	831
1e-04	18	3.129822e+00	1	7.835e-03	2 589
5e-05	23	3.839284e-01	7.396e-02	9.695e-03	1 834
1e-05	56	5.197695e-01	1.453e-02	9.967e-03	4 398

Table 3: Case without blur: sensitivity with respect to the gradient method step ρ , with $\varepsilon = 0.5$, $\alpha = 15$, $tol = 1e - 02$, and $it_{max} = 2000$.

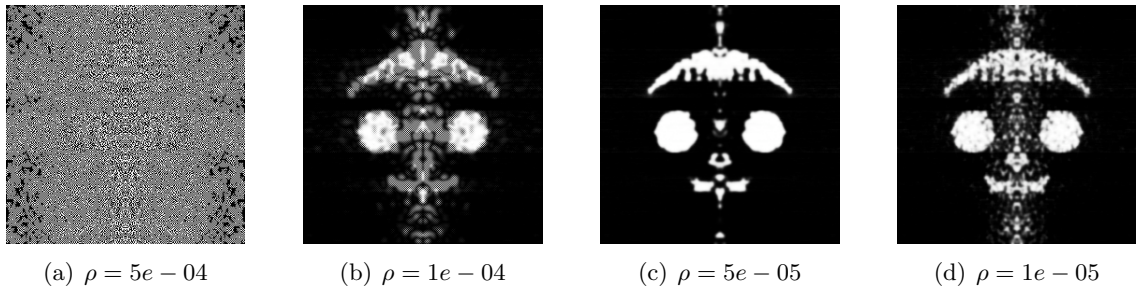


Figure 9: Case without blur, for different values of ρ , with $\varepsilon = 0.5$, $\alpha = 15$, $tol = 1e - 02$, and $it_{max} = 2000$.

Table 4 reports numerical results in the case without blur, for different values of the gradient method maximal iteration number it_{max} . In the simulations, the maximal number of iterations it_{max} of the gradient method is always attained. We report in the table the number of iterations of the main loop.

it_{max}	# iterations	F_ε	$\ u^{n+1} - u^n\ _\infty$	δ_ε^n	CPU time
100	2	1.155450e+01	6.817e-03	6.619e-03	220
500	122	9.053952e-01	4.977e-03	9.999e-03	6 751
1000	85	6.359336e-01	8.705e-03	9.803e-03	5 243
2000	56	5.197695e-01	1.453e-02	9.967e-03	4 398
3000	44	4.758873e-01	2.268e-02	9.755e-03	4 470

Table 4: Case without blur: sensitivity with respect to the gradient method maximum iteration number it_{max} , with $\varepsilon = 0.5$, $\alpha = 15$, and $tol = 1e - 02$.

When adding a Gaussian blur, the observed data is the projected image perturbed with a Gaussian blur with standard deviation $\sigma_B = 3$ and the previous Gaussian noise τ with $\sigma_\tau = 0.1$,

$$u_d = (K * H_o)(u_{orig}) + \tau ,$$

where

$$K(x) = \frac{1}{\sqrt{2\pi}\sigma_B} e^{-\frac{|x|^2}{2\sigma_B^2}} .$$

The tolerance is 10^{-3} and the maximal number of gradient iterations is 2000. Results are reported on Table 5 and Figure 10.

ε	α	# iterations	F	F_ε	$\ u^{n+1} - u^n\ _\infty$	δ_ε^n	CPU time (s)
0.5	1	170	8.778e-02	1.788e-01	1.331e-02	9.903e-04	16 426
0.5	5	235	6.047e-02	1.234e-01	1.208e-02	9.855e-04	21 815
0.5	10	125	1.005e-01	2.050e-01	1.097e-02	9.840e-04	10 406
0.5	20	108	1.692e-01	3.434e-01	1.651e-02	9.919e-04	7 903
0.5	30	114	2.262e-01	4.576e-01	1.485e-02	9.918e-04	7 206
0.1	10	163	4.325e-01	8.766e-01	5.488e-02	9.858e-04	17 981
0.5	10	125	1.005e-01	2.050e-01	1.097e-02	9.840e-04	10 406
1	10	122	8.936e-02	1.823e-01	7.734e-03	9.847e-04	10 348
5	10	102	8.102e-02	1.634e-01	4.442e-03	9.466e-04	9 612
10	10	100	7.974e-02	1.602e-01	4.207e-03	9.935e-04	9 803

Table 5: Case with blur: sensitivity with respect to α and ε , with $tol = 1e - 03$ and $it_{max} = 2000$.

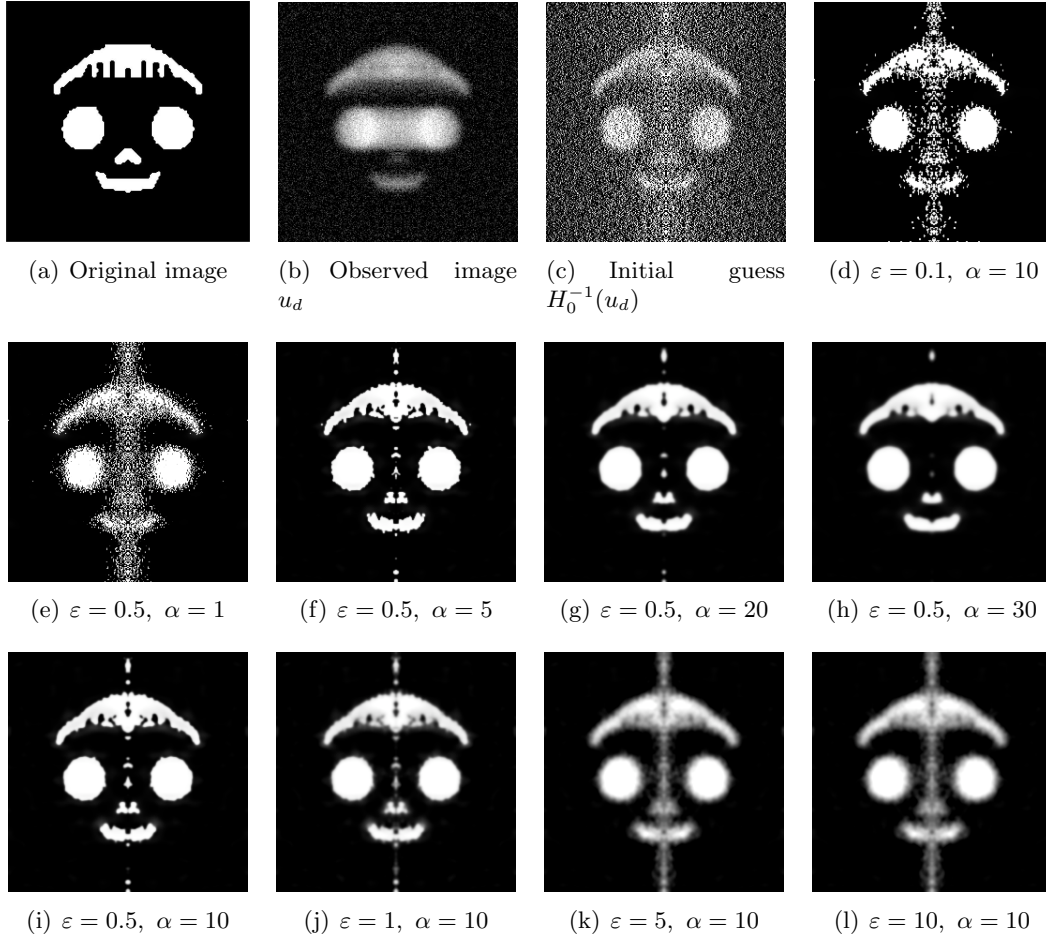


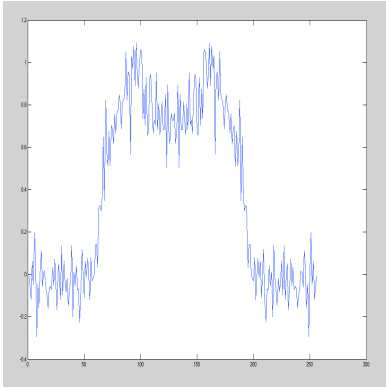
Figure 10: Case with blur, for different values of α and ε .

Our algorithm depends on the parameters $\varepsilon, \alpha, \rho, tol, it_{max}$. Whereas it is quite easy and standard to tune adequately the parameters ρ, tol and it_{max} , we observe on the above simulations that the results strongly depend on the values chosen for the parameters ε and α . This is one of the drawbacks of the method to be very sensitive to parameters adjustment. The parameter α is attached to the total variation. As expected, if α is too large, the image resulting from the simulation is too smooth, see e.g. Figure 10(h). At the contrary if α is chosen too small, the algorithm does not permit to denoise sufficiently the image, see e.g. Figure 7(a,c,g) and Figure 10(e). The parameter ε is the penalization parameter attached to the binary constraint. If ε is chosen too large, then the binary features of the object are not well recovered and details are too much smoothed, see e.g. Figure 10(k,l).

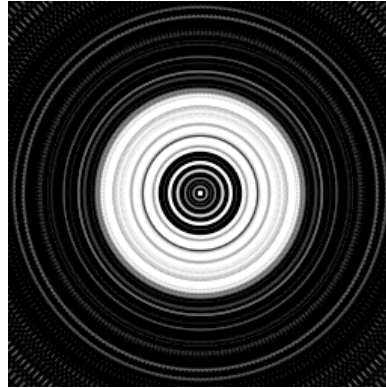
For our specific data, in the case with blur (see Figure 10), it seems that the best choice corresponds to $\varepsilon = 0.5$ and $\alpha = 10$ or 20 . In the general case the question of an optimal tuning seems is open. One of our research directions is to investigate, for α fixed, ways to choose an optimal penalization parameter ε at each step of the iteration.

As mentioned formerly, at this stage the code is quite basic and one of our next issues is to improve the performances of the algorithm. In this version, the step 4 is quite delicate to solve and for the moment time-consuming. Convergence results and convergence rates are to be derived. Moreover, sharp comparisons with existing methods must be performed. What can be said however is that the method presented here is far more efficient than the method presented in [1], based on a Hamilton-Jacobi approach applied to the same problem, and which requires far more computation times. Note that, in our method, we have a well defined stopping criterion, contrarily to [1].

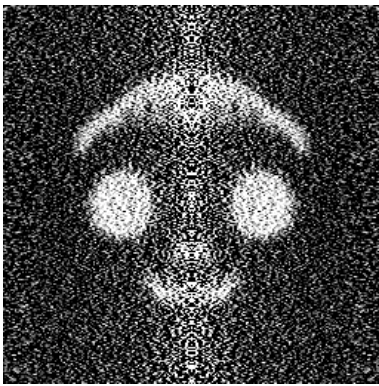
Finally, Figure 11 presents numerical results obtained with the back-filtered projection routine `iradon` of MATLAB[®] applied to slices generated by the rotation of object lines. These results are of bad quality, as expected.



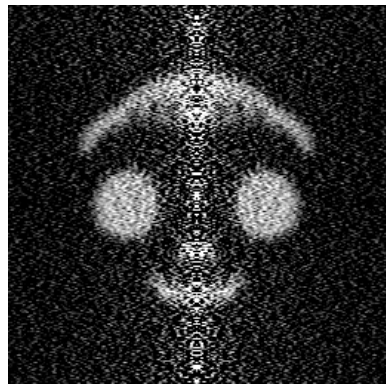
(a) Median line of the object ($i = 128$).



(b) Largest slice of the binary axially symmetric object by a plane orthogonal to the symmetry axis.



(c) Result using the cropped Ram-Lak filter.



(d) Result using the Ram-Lak filter with a Hamming window.

Figure 11: Back-filtered projection results

Movies showing the evolution of the image during the execution of the algorithm are available at <http://www.univ-orleans.fr/mapmo/membres/maitine/Movies>. The subdirectory `Blurred` contains movies of cases with blur.

5 Conclusion

In this article, we have described and implemented a variational method for tomographic reconstruction of blurred and noised binary images, using a minimization procedure in the space of bounded variation functions. The binary constraint is handled with a penalization process. Our approach is global, contrarily to usual methods based on a line-by-line analysis of the radiograph.

We proved existence and/or uniqueness results for the minimization problem under consideration, and derived a penalized optimality system. Then, we implemented the procedure and provided numerical simulations.

Much progress remains to be done, in particular to improve the robustness and the performances of the algorithm. The main difficulty consists in tuning the parameters of the procedure, in particular, the parameter α , which is the weight of the total variation in the optimization criterion, and the parameter ε , related to the penalization of the binary constraint. An important open question is to develop an automatic procedure that would permit to tune adequately these parameters. It is all the more important that our algorithm is very sensitive to parameters adjustment. For instance, it seems reasonable to try to develop a strategy permitting to choose the parameter ε , at each step of the iteration, in function of α .

Up to now, no sharp comparisons were led with other existing procedures of tomographic reconstruction for binary objects, however our algorithm is clearly more efficient (in particular, less time-consuming) than the strategy described in [1], based on a Hamilton-Jacobi approach. Simulations are still quite slow but the algorithm was not optimized.

Finally, a refined functional analysis of the Radon transform actually permits to derive stronger regularity properties of the operator H_0 defined by (2). From these results, knowing precisely the image of the space $BV(\Omega)$ by the operator H , it may be relevant to consider other norms for the definition of the minimization criterion $F(u)$ of the problem (\mathcal{P}). Here a natural choice was the standard L^2 norm, which has the double advantage of being simple to differentiate (in order to derive the optimality system) and simple to compute numerically. Other choices are however possible and it is an open interesting direction to investigate the quality of resulting simulations.

References

- [1] I. Abraham, R. Abraham, M. Bergounioux, *An active curve approach for tomographic reconstruction of binary axially symmetric objects*, preprint HAL csd-00084855 (2006).
- [2] R. Acar, C.R. Vogel, *Analysis of bounded variation penalty methods for ill-posed problems*, Inverse Problems 10 (1994), no. 6, 1217–1229.
- [3] R.A. Adams, J.J.F. Fournier, *Sobolev Spaces*, Second Edition, Pure and Applied Mathematics, Volume 140, Hardcover (2003).
- [4] L. Ambrosio, N. Fusco, D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford mathematical monographs, Oxford University Press (2000).
- [5] G. Aubert, P. Kornprobst, *Mathematical Problems in Image Processing, Partial Differential Equations and the Calculus of Variations*, Applied Mathematical Sciences 147, Springer Verlag (2006).

- [6] R.H.T. Bates, K.L. Garden, T.M. Peters, *Overview of computerized tomography with emphasis on future developments*, Proc. IEEE **71**, 3 (1983), 356–297.
- [7] M. Bergounioux, F. Mignot, *Control of Variational Inequalities and Lagrange Multipliers*, ESAIM, COCV, **5** (2000), 45–70.
- [8] J.P. Bruandet, F. Peyrin, J.M. Dinten, M. Barlaud, *3D tomographic reconstruction of binary images from cone-beam projections: a fast level-set approach*, IEEE International Symposium on Biomedical Imaging (2002), 677–680.
- [9] E. Casas, K. Kunisch, C. Pola, *Regularization by Functions of Bounded Variation and Applications to Image Enhancement*, Applied Mathematics and Optimization **40** (1999), 229–257.
- [10] A. Chambolle, *An algorithm for total variation minimization and applications*, Journal of Mathematical Imaging and Vision **20**, 1-2 (2004), 89–97.
- [11] J.-M. Dinten, *Tomographie à partir d'un nombre limité de projections : régularisation par champs markovien*, PHD thesis, Université Paris-Sud (1990).
- [12] N.J. Dusaussoy, *Image reconstruction from projections*, SPIE's international symposium on optics, imaging and instrumentation, San Diego (1994).
- [13] H. Feng, W. Karl, D. Castanon, *A curve evolution approach to object-based tomographic reconstruction*, IEEE Trans. Image Proc. **12** (2003), 44–57.
- [14] K. Hanson, *Tomographic reconstruction of axially symmetric objects from a single radiograph*, High Speed Photography **491** (1984).
- [15] G. Herman, *Image reconstruction from projections: the fundamentals of computerized tomography*, Academic Press (1980).
- [16] A.C. Kak, M. Slaney, *Principles of Computerized Tomographic Imaging*, New York, IEEE Press (1988).
- [17] J.-M. Lagrange, *Reconstruction tomographique à partir d'un petit nombre de vues*, PHD thesis, ENS Cachan (1998).
- [18] M. Magnor, G. Kindlmann, C. Hansen, N. Duric, *Reconstruction and visualization of planetary nebulae*, IEEE Trans. Visualization Computer Graphics **11**, 5 (2005), 485–496.
- [19] D. Mumford, J. Shah, *Optimal approximations by piecewise smooth functions and associated variational problem*, Comm. Pure Appl. Math. **42** (1989), 577–685.
- [20] J.A. Sethian, *Theory, algorithms and applications of level set methods for propagating interfaces*, Acta Numerica **5** (1996), 309–395, Iserles, A. (ed.), Cambridge University Press.