HAL
open science

A second-order model for image denoising

Maïtine Bergounioux, Loïc Piffet

To cite this version:

Maïtine Bergounioux, Loïc Piffet. A second-order model for image denoising. Set Valued Analysis, 2010, 18 (3-4), pp.277-306. 10.1007/s11228-010-0156-6 . hal-00440872v2

HAL Id: hal-00440872

https://hal.science/hal-00440872v2

Submitted on 7 Jun 2010

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.

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

A second-order model for image denoising

M. Bergounioux · L. Piffet

the date of receipt and acceptance should be inserted later

Abstract We present a variational model for image denoising and/or texture identification. Noise and textures may be modelled as oscillating components of images. The model involves a L^2- data fitting term and a Tychonov-like regularization term. We choose the BV^2 norm instead of the classical BV norm. Here BV^2 is the bounded hessian function space that we define and describe. The main improvement is that we do not observe staircasing effects any longer, during denoising process. Moreover, texture extraction can be performed with the same method. We give existence results and present a discretized problem. An algorithm close to the one set by Chambolle [9] is used: we prove convergence and present numerical tests.

Keywords Second order total variation · image reconstruction · denoising · texture · variational method

Mathematics Subject Classification (2000) 65D18 · 68U10 · 65K10

1 Introduction

Variational models in image processing have been extensively studied during the past decade. There are used for segmentation processes (geodesic or geometric contours) and restoration purpose as well. We are mainly interested in the last item which involves denoising or deblurring methods and textures extraction as well. Roughly speaking image restoration problems are severely ill posed and a Tychonov-like regularization process is needed. The general form of such models consists in the minimization of an "energy" functional :

$$\mathcal{F}(u) = \|u - u_d\|_X + \mathcal{R}(u) \ , u \in Y \subset X \ ,$$

where X, Y are (real) Banach spaces, $\mathcal{R}$ is a regularization operator, u_d is the observed (or measured) image and u is the image to recover or to denoise. The first term is the

Université d'Orléans, UFR Sciences
Math., Labo. MAPMO, UMR 6628
Route de Chartres, BP 6759
45067 Orléans cedex 2
E-mail: maitine.bergounioux@univ-orleans.fr, loic.piffet@univ-orleans.fr

fitting data term and the second one permits to get a problem which is no longer ill posed via a regularization process. The most famous model is the Rudin-Osher-Fatemi denoising model (see [1], [20]). This model involves a regularization term that preserves discontinuities, what a classical H^1 -Tychonov regularization method does not. The observed image to recover is split in two parts $u_d = u + v$ where u represents the oscillating component (noise or texture) and v is the smooth part (often called the *cartoon* component). So we look for the solution as $u + v$ with $v \in BV(\Omega)$ and $u \in L^2(\Omega)$, where $BV(\Omega)$ is the functions of bounded variation space defined on an open set Ω ([2,4]). The regularization term involves only the cartoon component v, while the remainder term $u_d - v$ represents the noise to be minimized. We get

$$\min_{v \in BV(\Omega)} \mu |v|_{BV(\Omega)} + \frac{1}{2} \|u_d - v\|^2_{L^2(\Omega)}, \qquad (\mathcal{P}_{ROF})$$

where $\mu > 0$. This problem has a unique solution in $BV(\Omega)$. This functional space is the good one to deal with discontinuous functions that imply that the derivative (in the distribution sense) may be a measure (for example a Dirac measure).

This model is used for denoising purpose. However, the use of the BV norm implies numerical perturbations. The computed solution turns to be piecewise constant which is called the "staircasing effect". Therefore, though noise can be successfully removed, the solution is not satisfactory. This variational model has been improved using different functional spaces, for the data fitting term or the regularizing term.

Recently people considered that an image can be decomposed into many components, each component describing a particular property of the image ([5,6,21–23] and references therein for example). We shall assume as well that the image we want to recover from the data u_d can be decomposed as $f = u + v$ where u and v are functions that characterize different parts of f (see [6,23,26] for example).

Components u and v belong to different functional spaces: u is the non regular part and belongs to $L^2(\Omega)$ while v is a more regular part and belongs to $BV^2(\Omega)$ (that we define in the sequel). The remainder term $u = f - v$ involves the oscillating component (as noise and/or texture) and possibly contours. Such decompositions have been already performed [5–7] using the so called Meyer-space of oscillating functions G [19] instead of $BV^2(\Omega)$. So far, the modelling we propose is not the same: the oscillating component will be a priori included in the non regular remainder term part $u := f - v$ while v involves the cartoon. Our philosophy is different than the one used in Meyer approach. Here we perform a second order analysis to get sharper result so that we look for the smooth (cartoon) part. The oscillating part, texture and/or noise and possible contours will be part of the remainder term and are not modelled a priori.

The paper is organized as follows : in the next section, we present the functional framework and the space $BV^2(\Omega)$ with useful properties. Section 3 is devoted to the variational model. In section 4, we focus on the discretization process. We present a Chambolle [9] like algorithm in section 5 and numerical tests are reported in the last section.

2 The space $BV^2(\Omega)$

Let Ω be an open bounded subset of $\mathbb{R}^n$, $n \geq 2$ (practically $n = 2$) smooth enough (with the cone property and $\mathcal{C}^1$ for example). Following Demengel [10], we define the space of bounded hessian functions that we call $BV^2(\Omega)$. We first recall the definition

and the main properties of the space $BV(\Omega)$ of functions of bounded variation (see [2, 4,7] for example), defined by

$$BV(\Omega) = \{u \in L^1(\Omega) \mid \Phi_1(u) < +\infty\},$$

where

$$\Phi_1(u) := \sup\left\{\int_\Omega u(x)\,\mathrm{div}\,\xi(x)\,dx \mid \xi \in \mathcal{C}_c^1(\Omega),\ \|\xi\|_\infty \leqslant 1\right\}. \tag{1}$$

The space $BV(\Omega)$, endowed with the norm $\|u\|_{BV(\Omega)} = \|u\|_{L^1} + \Phi_1(u)$, is a Banach space. The derivative in the sense of distributions of every $u \in BV(\Omega)$ is a bounded Radon measure, denoted Du, and $\Phi_1(u) = \int_\Omega |Du|$ is the total variation of u. We next recall standard properties of functions of bounded variation [2,4].

Proposition 1 *Let Ω be an open subset of $\mathbb{R}^n$ with Lipschitz boundary.*

1. *For every $u \in BV(\Omega)$, the Radon measure Du can be decomposed into $Du = \nabla u\,dx + D^s u$, where $\nabla u\,dx$ is the absolutely continuous part of Du with respect of the Lebesgue measure and $D^s u$ is the singular part.*
2. *The mapping $u \mapsto \Phi_1(u)$ 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, if $n = 2$.*
4. *$BV(\Omega) \subset L^p(\Omega)$ with compact embedding, for every $p \in [1, 2)$, if $n = 2$.*

Now we extend this definition to the second derivative (in the distributional sense). Recall that the Sobolev space is defined as

$$W^{1,1}(\Omega) = \{\ u \in L^1(\Omega) \mid \nabla u \in L^1(\Omega)\ \}$$

where ∇u stands for the first order derivative of u (in the sense of distributions).

Definition 1 *A function $u \in W^{1,1}(\Omega)$ is Hessian bounded if*

$$\Phi_2(u) := \sup\left\{\int_\Omega \langle \nabla u, div(\xi)\rangle_{\mathbb{R}^n} \mid \xi \in \mathcal{C}_c^2(\Omega, \mathbb{R}^{n\times n}),\ \|\xi\|_\infty \leqslant 1\right\} < \infty,$$

where

$$div(\xi) = (div(\xi_1), div(\xi_2), \ldots, div(\xi_n)), \tag{2}$$

with

$$\forall i,\ \xi_i = (\xi_i^1, \xi_i^2, \ldots, \xi_i^n) \in \mathbb{R}^n \quad \textit{and} \quad div(\xi_i) = \sum_{k=1}^n \frac{\partial \xi_i^k}{\partial x_k}.$$

$BV^2(\Omega)$ is defined as the space of $W^{1,1}(\Omega)$ functions such that $\Phi_2(u) < +\infty$.

Remark 1 *If $V = \mathbb{R}^{n\times n}$, $\|\xi\|_\infty = \sup_{x\in\Omega} \sqrt{\sum_{i,j=1}^n \left|\xi_i^j(x)\right|^2}$.*

We give thereafter many useful properties of $BV^2(\Omega)$ (proofs can be found in [10,24]).

Theorem 1 *The space $BV^2(\Omega)$ endowed with the following norm*

$$\|f\|_{BV^2(\Omega)} := \|f\|_{W^{1,1}(\Omega)} + \Phi_2(f) = \|f\|_{L^1} + \|\nabla f\|_{L^1} + \Phi_2(f), \tag{3}$$

where Φ_2 is given by (2), is a Banach space.

Proposition 2 *A function u belongs to $BV^2(\Omega)$ if and only if $u \in W^{1,1}(\Omega)$ and $\dfrac{\partial u}{\partial x_i} \in BV(\Omega)$ for $i \in \{1, \dots, n\}$. In particular*

$$\Phi_2(u) \leqslant \sum_{i=1}^{n} \Phi_1\left(\frac{\partial u}{\partial x_i}\right) \leqslant n\ \Phi_2(u).$$

Remark 2 *The previous result shows that*

$$BV^2(\Omega) = \left\{ u \in W^{1,1}(\Omega) \mid \forall i \in \{1, \dots, n\},\ \frac{\partial u}{\partial x_i} \in BV(\Omega) \right\}.$$

We get a lower semi-continuity result for the semi-norm Φ_2 as well.

Theorem 2 *The operator Φ_2 is lower semi-continuous from $BV^2(\Omega)$ endowed with the strong topology of $W^{1,1}(\Omega)$ to $\mathbb{R}$. More precisely, if $\{u_k\}_{k\in\mathbb{N}}$ is a sequence of $BV^2(\Omega)$ that strongly converges to u in $W^{1,1}(\Omega)$ then*

$$\Phi_2(u) \leqslant \liminf_{k\to\infty} \Phi_2(u_k).$$

Remark 3 *In particular, if $\liminf\limits_{k\to\infty} \Phi_2(u_k) < \infty$, then $u \in BV^2(\Omega)$.*

We have embedding results as well:

Theorem 3 (Demengel [10]) *Assume $n \geq 2$. Then*

$$BV^2(\Omega) \hookrightarrow W^{1,q}(\Omega) \quad \text{with} \quad q \leqslant \frac{n}{n-1}, \tag{4}$$

with continuous embedding. Moreover the embedding is compact if $q < \frac{n}{n-1}$. In particular

$$BV^2(\Omega) \hookrightarrow L^q(\Omega) \quad \text{for} \quad q \leqslant \frac{n}{n-2} \quad \text{if} \quad n > 2 \tag{5}$$

$$BV^2(\Omega) \hookrightarrow L^q(\Omega), \quad \forall q \in [1, \infty[, \quad \text{if} \quad n = 2. \tag{6}$$

In the sequel, we set $n = 2$ and Ω is a subset of $\mathbb{R}^2$, so that $BV^2(\Omega) \subset H^1(\Omega)$ with continuous embedding and $BV^2(\Omega) \subset W^{1,1}(\Omega)$ with compact embedding.

3 The variational model

We now assume that the image we want to recover from the data u_d can be decomposed as $f = u + v$ where u and v are functions that characterize different parts of f (see [6,23,26] for example). Components u and v belong to different functional spaces: v is the (smooth) second order part and belongs to $BV^2(\Omega)$ while u is the remainder term that should involve noise and/or texture.
We consider the following cost functional defined on $BV^2(\Omega)$:

$$F(v) = \frac{1}{2}\|u_d - v\|^2_{L^2(\Omega)} + \lambda\Phi_2(v) + \delta\|v\|_{W^{1,1}(\Omega)}, \tag{7}$$

where $\lambda, \delta \geq 0$. We are looking for a solution to the optimization problem

$$\inf\{\ F(v) \mid v \in BV^2(\Omega)\ \} \tag{$\mathcal{P}$}$$

The first term $\|u_d - v\|^2_{L^2(\Omega)}$ of F is the fitting data term. Here we have chosen the L^2-norm for simplicity but any L^p norm can be used ($p \in [2, +\infty)$). We shall investigate in a future work the very case where $p = 1$; indeed we need to develop an approximation process to deal with this additional non differentiable term. Other terms are Tychonov-like regularization terms. The term $\lambda\Phi_2(v) + \delta\|v\|_{W^{1,1}(\Omega)}$ is nothing else that the $BV^2(\Omega)$ norm of v. However, we have split it because the δ-part is not useful from the modelling point of view. It is only necessary to prove existence of solutions. We shall choose $\delta = 0$ for numerical tests.

If the image is noisy, the noise is considered as a texture and will be involved in the remainder term $u := u_d - v$: more precisely v will be the part of the image without the oscillating component, that is the denoised part and u is expected to involve noise, contours (and part of texture). In the sequel we shall focus on the denoising process. Such an approach has already been used by Hinterberger and Scherzer [15] with the $BV^2(\Omega)$ space. Their algorithm is different from the one we use here. Note that if we decide to split the function we look for in more than two components ($f = u + v + w$, where u is the texture part, v the cartoon and w the noise for example, with appropriate models) then the minimization problem is a structured optimization problem to which standard decomposition methods can be applied (for example alternating minimization (Gauss Seidel) or parallel methods). Therefore, the study of ($\mathcal{P}$) is quite significant from this point of view. For parallel methods one may consult [3].
First we give a general existence and uniqueness result for problem ($\mathcal{P}$).

Theorem 4 *Assume that $\lambda > 0$ and $\delta > 0$. Problem ($\mathcal{P}$) has a unique solution v.*

Proof.- We first prove existence by using the direct method in calculus of variations. Let $v_n \in BV^2(\Omega)$ be a minimizing sequence, i.e.

$$\lim_{n\to+\infty} F(v_n) = \inf\{\ F(v) \mid v \in BV^2(\Omega)\ \} < +\infty.$$

The sequence $(v_n)_{n\in\mathbb{N}}$ is bounded in $BV^2(\Omega)$. With the compactness result of Theorem 3, this yields that $(v_n)_{n\in\mathbb{N}}$ strongly converges (up to a subsequence) in $W^{1,1}(\Omega)$ to $v^* \in BV^2(\Omega)$. Theorem 2 gives the following:

$$\Phi_2(v^*) \leqslant \liminf_{n\to+\infty} \Phi_2(v_n).$$

So

$$F(v^*) \leqslant \liminf_{n\to+\infty} F(v_n) = \min_{v\in BV^2(\Omega)} F(v),$$

and v^* is a solution to $(\mathcal{P})$. Uniqueness is straightforward with the strict convexity of F due to the term $\|u_d - v\|^2_{L^2(\Omega)}$.

■

Remark 4 *It is an open question to know if we really need $\delta > 0$. We could expect a Poincaré-Wirtinger inequality in the $BV^2(\Omega)$- space which is not very difficult to prove using appropriate "density" results. However, the boundary conditions we have to add are not clear (we have to deal with "second order" Neumann boundary conditions) . It is not straightforward to get the suitable boundary conditions that allow to assert that the semi-norm Φ_2 is equivalent to the $BV^2(\Omega)$-norm. If we assume $\delta = 0$ we are not sure to prove existence of a solution without additional assumptions and we lose uniqueness.*

4 The discretized problem

Problem $(\mathcal{P})$ can be written equivalently as

$$\inf_{v\in BV^2(\Omega)} \frac{\|u_d - v\|^2_{L^2(\Omega)}}{2\lambda} + \Phi_2(v) + \delta\|v\|_{W^{1,1}(\Omega)}, \qquad (\tilde{\mathcal{P}})$$

where δ has been replaced by $\frac{\delta}{\lambda}$. We are going to compute the solution numerically. We first present the discretization process.

4.1 Discretization of problem $\tilde{\mathcal{P}}$

We assume for simplicity that the image is squared with size $N \times N$. We note $X := \mathbb{R}^{N\times N} \simeq \mathbb{R}^{N^2}$ endowed with the usual inner product and the associated Euclidean norm

$$\langle u, v\rangle_X := \sum_{1\leqslant i,j\leqslant N} u_{i,j}v_{i,j}, \quad \|u\|_X := \sqrt{\sum_{1\leqslant i,j\leqslant N} u^2_{i,j}}\ . \qquad (8)$$

We set $Y = X \times X$. It is classical to define the discrete total variation as following (see for example [7]): the discrete gradient of the numerical image $u \in X$ is $\nabla u \in Y$:

$$(\nabla u)_{i,j} = \left((\nabla u)^1_{i,j}, (\nabla u)^2_{i,j}\right), \qquad (9)$$

where $(\nabla u)^1_{i,j} = \begin{cases} u_{i+1,j} - u_{i,j} & \text{if } i < N \\ 0 & \text{if } i = N, \end{cases}$ and $(\nabla u)^2_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j} & \text{if } j < N \\ 0 & \text{if } j = N. \end{cases}$
The (discrete) total variation corresponding to $\Phi_1(u)$ is given by

$$J_1(u) = \sum_{1\leqslant i,j\leqslant N} \left\|(\nabla u)_{i,j}\right\|_{\mathbb{R}^2}, \qquad (10)$$

where $\left\|(\nabla u)_{i,j}\right\|_{\mathbb{R}^2} = \left\|\left((\nabla u)^1_{i,j}, (\nabla u)^2_{i,j}\right)\right\|_{\mathbb{R}^2} = \sqrt{\left((\nabla u)^1_{i,j}\right)^2 + \left((\nabla u)^2_{i,j}\right)^2}$.
The discrete divergence operator div is the opposite of the adjoint operator of the gradient operator ∇:

$$\forall (p,u) \in Y \times X, \quad \langle -\text{div } p, u\rangle_X = \langle p, \nabla u\rangle_Y ,$$

so that

$$(\text{div } p)_{i,j} = \begin{cases} p^1_{i,j} - p^1_{i-1,j} & \text{if } 1 < i < N \\ p^1_{i,j} & \text{if } i = 1 \\ -p^1_{i-1,j} & \text{if } i = N \end{cases} + \begin{cases} p^1_{i,j} - p^2_{i,j-1} & \text{if } 1 < j < N \\ p^2_{i,j} & \text{if } j = 1 \\ -p^1_{i,j-1} & \text{if } i = N. \end{cases} \tag{11}$$

To define a discrete version of the second order total variation Φ_2 we have to introduce the discrete Hessian operator. For any $v \in X$, the Hessian matrix of v, denoted Hv is identified to a X^4 vector: $(Hv)_{i,j} = \left((Hv)^{11}_{i,j}, (Hv)^{12}_{i,j}, (Hv)^{21}_{i,j}, (Hv)^{22}_{i,j}\right)$, with, for every $i, j = 1, \ldots, N$,

$$(Hv)^{11}_{i,j} = \begin{cases} v_{i+1,j} - 2v_{i,j} + v_{i-1,j} & \text{if } 1 < i < N, \\ v_{i+1,j} - v_{i,j} & \text{if } i = 1, \\ v_{i-1,j} - v_{i,j} & \text{if } i = N, \end{cases}$$

$$(Hv)^{12}_{i,j} = \begin{cases} v_{i,j+1} - v_{i,j} - v_{i-1,j+1} + v_{i-1,j} & \text{if } 1 < i \leqslant N,\ 1 \leqslant j < N, \\ 0 & \text{if } i = 1, \\ 0 & \text{if } i = N, \end{cases}$$

$$(Hv)^{21}_{i,j} = \begin{cases} v_{i+1,j} - v_{i,j} - v_{i+1,j-1} + v_{i,j-1} & \text{if } 1 \leqslant i < N,\ 1 < j \leqslant N, \\ 0 & \text{if } i = 1, \\ 0 & \text{if } i = N, \end{cases}$$

$$(Hv)^{22}_{i,j} = \begin{cases} v_{i,j+1} - 2v_{i,j} + v_{i,j-1} & \text{if } 1 < j < N, \\ v_{i,j+1} - v_{i,j} & \text{if } j = 1, \\ v_{i,j-1} - v_{i,j} & \text{if } j = N. \end{cases}$$

The discrete second order total variation corresponding to $\Phi_2(v)$ is defined as

$$J_2(v) = \sum_{1 \leqslant i,j \leqslant N} \left\|(Hv)_{i,j}\right\|_{\mathbb{R}^4}. \tag{12}$$

The discretized problem stands

$$\inf_{v \in X} \frac{\|u_d - v\|^2_X}{2\lambda} + J_2(v) + \delta(|v| + J_1(v)), \tag{$\mathcal{P}_d$}$$

where

$$|v| := \sum_{1 \leqslant i,j \leqslant N} |v_{i,j}| .$$

In the finite dimensional case we have an existence result with $\delta = 0$.

Theorem 5 *Problem $\mathcal{P}_d$ has a unique solution for every $\delta \geq 0$ and $\lambda > 0$.*

Proof.- The cost functional

$$F_\delta := \frac{\|u_d - v\|^2_X}{2\lambda} + J_2(v) + \delta(|v| + J_1(v)) ,$$

is continuous and coercive because of the term $\|u_d - v\|^2_X$. In addition it is strictly convex so that we get the result.

■

For numerical purpose we shall set $\delta = 0$. In fact, we have performed tests with $\delta = 0$ and very small $\delta \neq 0$ (as required by the theory to get a solution to problem $\tilde{\mathcal{P}}$): results were identical.

In the sequel, we adapt the method by Chambolle in the $BV(\Omega)$-case, to the second order framework. We briefly recall the original result of [9]. Consider the finite dimensional optimization problem derived from the discretization of the ROF model :

$$\inf_{u \in X} \left(J_1(u) + \frac{1}{2\lambda} \|u_d - u\|_X^2 \right). \tag{$\mathcal{P}_d^1$}$$

The following result holds :

Proposition 3 *The solution to ($\mathcal{P}_d^1$) is given by*

$$u = u_d - P_{\lambda K_1}(u_d), \tag{13}$$

where

$$K_1 = \{ div\ g \mid g \in Y,\ \left\|g_{i,j}\right\|_{\mathbb{R}^2} \leqslant 1 \quad \forall i,j = 1, \cdots, N \},$$

and $P_{\lambda K_1}$ is the orthogonal projector operator on λK_1.

We have a very similar result that we present in next sections. We first describe the conjugate function of J_2

4.2 The J_2 Legendre-Fenchel conjugate function

As the function J_2 is positively homogeneous, the Legendre-Fenchel conjugate

$$J_2^*(v) = \sup_u \langle u, v \rangle_X - J_2(u),$$

is the characteristic function of a closed, convex set K

$$J_2^*(v) = \mathbf{1}_K(v) = \begin{cases} 0 & \text{if } v \in K \\ +\infty & \text{else.} \end{cases}$$

As $J_2^{**} = J_2$, we get $J_2(v) = \sup_{u \in K} \langle v, u \rangle_X$. We use the inner scalar product of X^4 :

$$\langle p, q \rangle_{X^4} = \sum_{1 \leqslant i,j \leqslant N} \left(p_{i,j}^1 q_{i,j}^1 + p_{i,j}^2 q_{i,j}^2 + p_{i,j}^3 q_{i,j}^3 + p_{i,j}^4 q_{i,j}^4 \right),$$

for every $p = \left(p^1, p^2, p^3, p^4\right)$, $q = \left(q^1, q^2, q^3, q^4\right) \in X^4$. So, for every $v \in X$,

$$J_2(v) = \sup_{p \in \mathcal{C}} \langle p, Hv \rangle_{X^4}, \tag{14}$$

where the feasible set is

$$\mathcal{C} := \{ p \in X^4 \mid \left\|p_{i,j}\right\|_{\mathbb{R}^4} \leqslant 1,\ \forall\, 1 \leq i,j \leq N \}\ .$$

Let us compute the adjoint operator of H (which is the discretized "second divergence" operator) :

$$\forall p \in X^4,\ \forall v \in X \qquad \langle H^* p, v \rangle_X = \langle p, Hv \rangle_{X^4}.$$

We verify that $H^* : X^4 \to X$ satisfies for every $p = (p^{11}, p^{12}, p^{21}, p^{22}) \in X^4$

$$
\begin{aligned}
(H^*p)_{i,j} = &\begin{cases} p^{11}_{i-1,j} - 2p^{11}_{i,j} + p^{11}_{i+1,j} & \text{if } \ 1 < i < N \\ p^{11}_{i+1,j} - p^{11}_{i,j} & \text{if } \ i = 1, \\ p^{11}_{i-1,j} - p^{11}_{i,j} & \text{if } \ i = N, \end{cases} \\
+ &\begin{cases} p^{22}_{i,j-1} - 2p^{22}_{i,j} + p^{22}_{i,j+1} & \text{if } \ 1 < j < N, \\ p^{22}_{i,j+1} - p^{22}_{i,j} & \text{if } \ j = 1, \\ p^{22}_{i,j-1} - p^{22}_{i,j} & \text{if } \ j = N, \end{cases} \\
+ &\begin{cases} p^{12}_{i,j-1} - p^{12}_{i,j} - p^{12}_{i+1,j-1} + p^{12}_{i+1,j} & \text{if } \ 1 < i, j < N, \\ p^{12}_{i+1,j} - p^{12}_{i+1,j-1} & \text{if } \ i = 1, \ 1 < j < N, \\ p^{12}_{i,j-1} - p^{12}_{i,j} & \text{if } \ i = N, \ 1 < j < N, \\ p^{12}_{i+1,j} - p^{12}_{i,j} & \text{if } \ 1 < i < N, \ j = 1, \\ p^{12}_{i,j-1} - p^{12}_{i+1,j-1} & \text{if } \ 1 < i < N, \ j = N, \\ p^{12}_{i+1,j} & \text{if } \ i = 1, \ j = 1, \\ -p^{12}_{i+1,j-1} & \text{if } \ i = 1, \ j = N, \\ -p^{12}_{i,j} & \text{if } \ i = N, \ j = 1, \\ p^{12}_{i,j-1} & \text{if } \ i = N, \ j = N, \end{cases} \\
+ &\begin{cases} p^{21}_{i-1,j} - p^{21}_{i,j} - p^{21}_{i-1,j+1} + p^{21}_{i,j+1} & \text{if } \ 1 < i, j < N, \\ p^{21}_{i,j+1} - p^{21}_{i,j} & \text{if } \ i = 1, \ 1 < j < N, \\ p^{21}_{i-1,j} - p^{21}_{i-1,j+1} & \text{if } \ i = N, \ 1 < j < N, \\ p^{21}_{i,j+1} - p^{21}_{i-1,j+1} & \text{if } \ 1 < i < N, \ j = 1, \\ p^{21}_{i-1,j} - p^{21}_{i,j} & \text{if } \ 1 < i < N, \ j = N, \\ p^{21}_{i,j+1} & \text{if } \ i = 1, \ j = 1, \\ -p^{21}_{i,j} & \text{if } \ i = 1, \ j = N, \\ -p^{21}_{i-1,j+1} & \text{if } \ i = N, \ j = 1, \\ p^{21}_{i-1,j} & \text{if } \ i = N, \ j = N, \end{cases}
\end{aligned}
\tag{15}
$$

Theorem 6 *The Legendre-Fenchel conjugate of J_2 is $J_2^* = \mathbf{1}_{K_2}$ where*

$$K_2 := \{H^*p \mid p \in X^4, \ \|p_{i,j}\|_{\mathbb{R}^4} \leqslant 1, \ \forall i, j = 1, \ldots, N\} \subset X. \tag{16}$$

Proof.- We have already mentioned that $J_2^* = \mathbf{1}_K$ where K is of a non empty, closed, convex set subset of X.
We first show that $K_2 \subset K$. Let be $v \in K_2$. The discretized version of the definition of the second order total variation gives

$$J_2(v) = \sup_{\xi \in K_2} \langle \xi, v \rangle_X ,$$

and $\langle \xi, v \rangle - J_2(v) \leqslant 0$ for every $\xi \in K_2$ and $v \in X$. This gives for every $v^* \in K_2$

$$J_2^*(v^*) = \sup_{v \in K_2} \langle v^*, v \rangle - J_2(v) \leqslant 0. \tag{17}$$

As J_2^* takes only one finite value, we get $J_2^*(v^*) = 0$, which yields that $v^* \in K$. Therefore, $K_2 \subset K$ and as K is closed we finally obtain

$$\bar{K_2} \subset K.$$

In particular

$$J_2(v) = \sup_{\xi \in K_2} \langle v, \xi \rangle \leqslant \sup_{\xi \in \bar{K_2}} \langle v, \xi \rangle \leqslant \sup_{\xi \in K} \langle v, \xi \rangle = \sup_{\xi \in K} \langle v, \xi \rangle - J_2^*(\xi) = J_2^{**}(v).$$

As $J_2^{**} = J_2$, we have

$$\sup_{\xi \in K_2} \langle v, \xi \rangle = \sup_{\xi \in \bar{K_2}} \langle v, \xi \rangle = \sup_{\xi \in K} \langle v, \xi \rangle . \tag{18}$$

Now, let us assume there exists $v^* \in K$ such that $v^* \notin \bar{K_2}$. On may strictly separate v^* and the closed convex set $\bar{K_2}$. There exists $\alpha \in \mathbb{R}$ and v_0 such that

$$\left\langle v_0, v^* \right\rangle > \alpha \geqslant \sup_{u \in \bar{K_2}} \langle v_0, u \rangle .$$

We obtain

$$\sup_{\xi \in K} \langle v_0, \xi \rangle \geqslant \left\langle v_0, v^* \right\rangle > \alpha \geqslant \sup_{u \in \bar{K_2}} \langle v_0, u \rangle = \sup_{u \in K} \langle v_0, u \rangle ,$$

that gives a contradiction. We have proved that $K = \bar{K_2}$. As K_2 is closed we get $K = K_2$.

■

Remark 5 *We proved the previous result for the convenience of the reader. Indeed, one may note that J_2 is the support function of K_2 which is the conjugate function of the indicator function 1_{K_2} of K_2 (see [12] p. 19). Therefore, as K_2 is closed and convex and J_2 is continuous we get $J_2^* = 1_{K_2}^{**} = 1_{K_2}$.*

4.3 Case where $\delta = 0$

Now we focus on the case where $\delta = 0$. Indeed, the δ-term in definition 7 was needed as a tool in the infinite dimensional framework to prove existence of solutions to $(\mathcal{P})$. However, we have seen that the finite dimensional problem has a solution even if $\delta = 0$, which is the most interesting case. The problem we have to solve turns to be

$$\min_{v \in X} \frac{\|u_d - v\|_X^2}{2\lambda} + J_2(v). \tag{$\mathcal{P}_d^2$}$$

As in the BV-case (proposition 3) we have a characterization of the solution.

Theorem 7 *The solution v of $(\mathcal{P}_d^2)$ verifies*

$$v = u_d - P_{\lambda K_2}(u_d),$$

where $P_{\lambda K_2}$ is the orthogonal projector operator on λK_2.

Proof.- The proof is similar to the one by Chambolle [9] but we give it for convenience. The solution v to $(\mathcal{P}_d^2)$ is characterized by

$$0 \in \partial \left(J_2(v) + \frac{1}{2\lambda}\|v - u_d\|_2^2 \right) = \frac{v - u_d}{\lambda} + \partial J_2(v),$$

that is $\dfrac{u_d - v}{\lambda} \in \partial J_2(v)$. As J_2 is proper, convex and continuous, then

$$v^* \in \partial J_2(v) \Longleftrightarrow v \in \partial J_2^*(v^*).$$

So

$$v \in \partial J_2^*(\frac{u_d - v}{\lambda}) \iff 0 \in -v + \partial J_2^*(\frac{u_d - v}{\lambda}) \iff 0 \in \frac{u_d - v}{\lambda} - \frac{u_d}{\lambda} + \frac{1}{\lambda}\partial J_2^*(\frac{u_d - v}{\lambda}).$$

This means that $w = \dfrac{u_d - v}{\lambda}$ is the solution to

$$\min_w \frac{1}{2}\left\| w - \frac{u_d}{\lambda} \right\|_X^2 + \frac{1}{\lambda}\partial J_2^*(w).$$

As $J_2^* = \mathbf{1}_{K_2}$ this implies that $\dfrac{u_d - v}{\lambda}$ is the orthogonal projection of $\dfrac{u_d}{\lambda}$ on K_2 :

$$\frac{u_d - v}{\lambda} = P_{K_2}(\frac{u_d}{\lambda}) \ .$$

As $P_{K_2}(\frac{u_d}{\lambda}) = \frac{1}{\lambda}P_{\lambda K_2}(u_d)$, this ends the proof.

■

5 A fixed-point algorithm to compute ∂J_2

In [9] Chambolle proposed a fixed-point algorithm to compute $P_{\lambda K_1}(f)$ (and $\partial J_1(f)$). Let us briefly recall the main idea that we use again in the BV^2 context. To compute $P_{\lambda K_1}(f)$ we have to solve

$$\min\{\|\lambda \text{div } p - u_d\|_X^2 \mid \|p_{i,j}\|_{\mathbb{R}^2} \leqslant 1 \ \forall i,j = 1,\dots,N\}. \tag{19}$$

that can be solved with a fixed-point method :

$$p^0 = 0 \tag{20a}$$

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n + \tau(\nabla(\text{div } p^n - u_d/\lambda))_{i,j}}{1 + \tau\|(\nabla(\text{div } p^n - u_d/\lambda))_{i,j}\|_{\mathbb{R}^2}} \tag{20b}$$

In addition, a convergence result is provided :

Theorem 8 ([9], Theorem 3.1) *Assume that τ satisfies $\tau \leqslant 1/8$. Then λdiv p^n converges to $P_{\lambda K_1}(u_d)$ as $n \to +\infty$.*

Remark 6 *In [8] J.F. Aujol proves that the modified algorithm :*

$$p^0 = 0 \tag{21a}$$

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n + \tau(\nabla(div\, p^n - u_d/\lambda))_{i,j}}{\max\left(1, \left\| p_{i,j}^n + \tau(\nabla(div\, p^n - u_d/\lambda))_{i,j} \right\|_{\mathbb{R}^2}\right)} \tag{21b}$$

converges if $\tau \leqslant 1/4$.

We extend this result to the second-order case. To compute $P_{\lambda K_2}(u_d)$ we have to solve

$$\min\left\{ \left\|\lambda H^* p - u_d\right\|_X^2 \mid p \in X^4,\ \left\|p_{i,j}\right\|_{\mathbb{R}^4}^2 - 1 \leqslant 0,\ i,j = 1,\dots,N \right\}. \tag{$\mathcal{P}'$}$$

Let us denote $F(p) = \|\lambda H^* p - u_d\|_X^2$ and

$$g_{i,j}(p) = \left\|p_{i,j}\right\|_{\mathbb{R}^4}^2 - 1 = (p_{i,j}^{11})^2 + (p_{i,j}^{12})^2 + (p_{i,j}^{21})^2 + (p_{i,j}^{22})^2 - 1.$$

First order optimality conditions give the existence of Lagrange multipliers $\alpha_{i,j}$, $(i,j) \in \{1,\dots,N\}^2$, such that

$$\nabla F(p) + \sum_{i,j=1}^{N} \alpha_{i,j} \nabla g_{i,j}(p) = 0, \tag{22a}$$

$$\alpha_{i,j} \geqslant 0 \ \text{ and } \ \alpha_{i,j} g_{i,j}(p) = 0,\ (i,j) \in \{1,\dots,N\}^2. \tag{22b}$$

It is easy to see that $\nabla F(p) = 2\lambda H\left[\lambda H^* p - u_d\right]$ and that

$$\sum_{i,j=1}^{N} \alpha_{i,j} \nabla g_{i,j}(p) = 2\alpha_{i,j} \left((p_{i,j}^{11}, p_{i,j}^{22}, p_{i,j}^{12}, p_{i,j}^{21}) \right)_{1 \leqslant i,j \leqslant N}.$$

Therefore relations (22) are equivalent to

$$\forall 1 \le i,j \le N \qquad \left(H\left[\lambda H^* p - u_d\right]\right)_{i,j} + \alpha_{i,j} p_{i,j} = 0, \tag{23a}$$

$$\forall 1 \le i,j \le N \qquad \alpha_{i,j} \geqslant 0 \ \text{ and } \ \alpha_{i,j} g_{i,j}(p) = 0. \tag{23b}$$

Let us compute the multipliers $\alpha_{i,j}$ more precisely :

- If $\alpha_{i,j} > 0$ then $\left\|p_{i,j}\right\|_{\mathbb{R}^4} = 1$.
- If $\alpha_{i,j} = 0$ then $\left(H\left[\lambda H^* p - u_d\right]\right)_{i,j} = 0$.

In both cases we get

$$\forall 1 \le i,j \le N \qquad \alpha_{i,j} = \left\| \left(H\left[\lambda H^* p - u_d\right]\right)_{i,j} \right\|_{\mathbb{R}^4}$$

and we finally obtain the following equality : $\forall (i,j) \in \{1,\dots,N\}^2$,

$$\left(H\left[\lambda H^* p - u_d\right]\right)_{i,j} + \left\| \left(H\left[\lambda H^* p - u_d\right]\right)_{i,j} \right\|_{\mathbb{R}^4} p_{i,j} = 0. \tag{24}$$

We use a semi-implicit gradient method to solve these equations, namely :
Choose $\tau > 0$, and

1. Set $p^0 = 0,\ n = 0$

2. p^n is known. For $(i,j) \in \{1,\dots,N\}^2$, compute $p_{i,j}^{n+1}$ with

$$p_{i,j}^n = p_{i,j}^{n+1} + \tau\left(\left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j} + \left\|\left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4} p_{i,j}^{n+1}\right).$$

This is equivalent to

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n - \tau\left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j}}{1 + \tau\left\|\left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}}. \tag{25}$$

The algorithm step τ is related to the adjoint operator H^* norm that we call κ in the sequel and we first give a κ estimate:

Lemma 1 *The adjoint operator H^* norm, κ satisfies $\kappa \leq 8$.*

Proof.- The definition of κ reads $\kappa = \sup_{\|p\|_{X^4} \leqslant 1} \left\|H^*p\right\|$.
As $\left\|H^*p\right\|_X = \sup_{q \in \vec{\mathcal{B}}_X(0,1)} \left\langle H^*p, q\right\rangle_X$ and

$$\forall q \in X^4 \qquad \left\langle H^*p, q\right\rangle_X = \langle p, Hq\rangle_{X^4} \leqslant \|Hq\|_{X^4}\|p\|_{X^4},$$

we get

$$\left\|H^*p\right\|_X \leqslant |||H|||\ \|p\|_{X^4}. \tag{26}$$

For any $q \in X^4$

$$\begin{aligned}
\|Hq\|_{X^4}^2 = & \sum_{1 \leqslant i,j \leqslant N} \Big[\left(q_{i+1,j} - 2q_{i,j} + q_{i-1,j}\right)^2 + \left(q_{i,j+1} - q_{i,j} - q_{i-1,j+1} + q_{i-1,j}\right)^2 \\
& + \left(q_{i+1,j} - q_{i,j} - q_{i+1,j-1} + q_{i,j-1}\right)^2 + \left(q_{i,j+1} - 2q_{i,j} + q_{i,j-1}\right)^2\Big] \\
\leqslant & \ 4 \sum_{1 \leqslant i,j \leqslant N} \Big[q_{i+1,j}^2 + q_{i,j}^2 + q_{i,j}^2 + q_{i-1,j}^2 + q_{i,j+1}^2 + q_{i,j}^2 + q_{i-1,j+1}^2 + q_{i-1,j}^2 \\
& + q_{i+1,j}^2 + q_{i,j}^2 + q_{i+1,j-1}^2 + q_{i,j-1}^2 + q_{i,j+1}^2 + q_{i,j}^2 + q_{i,j}^2 + q_{i,j-1}^2\Big] \\
\leqslant & \ 4 \times 16\ \|q\|_X^2.
\end{aligned}$$

Finally $|||H||| \leqslant 8$, and with relation (26), $\|H^*p\|_X \leqslant 8\ \|p\|_{X^4}$.
We deduce that $\kappa \leqslant 8$ (and $\kappa^2 \leqslant 64$).

■

Theorem 9 *Let be $\tau \leqslant 1/64$. Then $\lambda\left(H^*p^n\right)_n$ converges to $P_{\lambda K_2}(u_d)$ as $n \to +\infty$.*

Proof.- It is easy to check (by induction) that for every $n \geqslant 0$ and i,j, $\left\|p_{i,j}^n\right\|_{\mathbb{R}^4} \leqslant 1$.
Set $n > 0$ and let be $\eta^n = (p^n - p^{n+1})/\tau$. Denote by $(U_n)_n$ the sequence defined by :

$$U_n = \left\|H^*p^n - u_d/\lambda\right\|_X^2, \quad \forall n \geqslant 0.$$

The outline of the proof is the following : we first prove that $(U_n)_n$ is a (nonnegative) non increasing sequence and therefore it is convergent. The proof provides an equality (29) that allows to show that the sequences $(p^n)_n$ and $(p^{n+1})_n$ have the same cluster points which are the unique solution to a projection problem. Therefore the whole sequence $(p^n)_n$ converges.

• We first prove that $(U_n)_n$ is a non increasing sequence if $\tau \leqslant 1/\kappa^2$ where κ is the norm of H^*. We get

$$\begin{aligned}
U_{n+1} &= \left\|H^*(p^n - \tau\eta^n) - u_d/\lambda\right\|_X^2 \\
&= \left\langle -\tau H^*\eta^n + \left(H^*p^n - u_d/\lambda\right), -\tau H^*\eta^n + \left(H^*p^n - u_d/\lambda\right)\right\rangle_X \\
&= U_n - 2\tau\left\langle H^*\eta^n, H^*p^n - u_d/\lambda\right\rangle_X + \tau^2\left\|H^*\eta^n\right\|_X^2 \\
&= U_n - 2\tau\left\langle \eta^n, H\left[H^*p^n - u_d/\lambda\right]\right\rangle_{X^4} + \tau^2\left\|H^*\eta^n\right\|_X^2 \\
&\leqslant U_n - \tau\left[2\left\langle \eta^n, H\left[H^*p^n - u_d/\lambda\right]\right\rangle_{X^4} - \kappa^2\tau\left\|\eta^n\right\|_{X^4}^2\right] \\
&\leqslant U_n - \tau\left[\sum_{i,j=1}^N \left\langle 2\eta_{i,j}^n, \left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j}\right\rangle_{\mathbb{R}^4} - \kappa^2\tau\left\|\eta_{i,j}^n\right\|_{\mathbb{R}^4}^2\right]
\end{aligned}$$

As

$$\eta_{i,j}^n = \left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j} + \rho_{i,j}^n,$$

with

$$\rho_{i,j}^n = \left\|\left(H\left[\lambda H^*p^n - u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4} p_{i,j}^{n+1};$$

we obtain

$$\begin{aligned}
&\left\langle 2\eta_{i,j}^n, \left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j}\right\rangle_{\mathbb{R}^4} - \kappa^2\tau\left\|\eta_{i,j}^n\right\|_{\mathbb{R}^4}^2 \\
&= \left\langle \eta_{i,j}^n, \left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j}\right\rangle_{\mathbb{R}^4} + \left\langle \eta_{i,j}^n, \left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j}\right\rangle_{\mathbb{R}^4} - \kappa^2\tau\left\|\eta_{i,j}^n\right\|_{\mathbb{R}^4}^2 \\
&= \left\langle \eta_{i,j}^n, \eta_{i,j}^n - \rho_{i,j}^n\right\rangle_{\mathbb{R}^4} \\
&\qquad + \left\langle \left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j} + \rho_{i,j}^n, \left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j}\right\rangle_{\mathbb{R}^4} - \kappa^2\tau\left\|\eta_{i,j}^n\right\|_{\mathbb{R}^4}^2 \\
&= \left\|\eta_{i,j}^n\right\|_{\mathbb{R}^4}^2 - \left\langle \eta_{i,j}^n, \rho_{i,j}^n\right\rangle_{\mathbb{R}^4} + \left\|\left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}^2 \\
&\qquad + \left\langle \rho_{i,j}^n, \left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j}\right\rangle_{\mathbb{R}^4} - \kappa^2\tau\left\|\eta_{i,j}^n\right\|_{\mathbb{R}^4}^2 \\
&= \left\|\eta_{i,j}^n\right\|_{\mathbb{R}^4}^2 - \left\langle \eta_{i,j}^n, \rho_{i,j}^n\right\rangle_{\mathbb{R}^4} + \left\|\left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}^2 \\
&\qquad + \left\langle \rho_{i,j}^n, \eta_{i,j}^n - \rho_{i,j}^n\right\rangle_{\mathbb{R}^4} - \kappa^2\tau\left\|\eta_{i,j}^n\right\|_{\mathbb{R}^4}^2 \\
&= (1-\kappa^2\tau)\left\|\eta_{i,j}^n\right\|_{\mathbb{R}^4}^2 + \left(\left\|\left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}^2 - \left\|\rho_{i,j}^n\right\|_{\mathbb{R}^4}^2\right).
\end{aligned}$$

Finally

$$U_{n+1} \leqslant U_n - \tau\left[\sum_{i,j=1}^N (1-\kappa^2\tau)\left\|\eta_{i,j}^n\right\|_{\mathbb{R}^4}^2 + \left(\left\|\left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}^2 - \left\|\rho_{i,j}^n\right\|_{\mathbb{R}^4}^2\right)\right] \quad (27)$$

As $\left\|p_{i,j}^{n+1}\right\| \leqslant 1$, $\left\|\rho_{i,j}^n\right\|_{\mathbb{R}^4} \leqslant \left\|(H\left[\lambda H^*p^n - u_d/\lambda\right])_{i,j}\right\|_{\mathbb{R}^4}$ this yields that if $\tau \leqslant 1/\kappa^2$, then the sequence (U_n) is non increasing. Thus, the sequence $(U_n)_n$ is convergent to some m. Moreover, relation (27) gives

$$0 \leq \sum_{i,j=1}^N (1-\kappa^2\tau)\left\|\eta_{i,j}^n\right\|_{\mathbb{R}^4}^2 + \left(\left\|\left(H\left[H^*p^n - u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}^2 - \left\|\rho_{i,j}^n\right\|_{\mathbb{R}^4}^2\right) \leq \frac{U_n - U_{n+1}}{\tau}. \quad (28)$$

Passing to the limit in relation (28) gives

$$\lim_{n\to+\infty}\sum_{i,j=1}^{N}\left[(1-\kappa^2\tau)\|\eta^n_{i,j}\|^2_{\mathbb{R}^4}+\left(\left\|\left(H\left[H^*p^n-u_d/\lambda\right]\right)_{i,j}\right\|^2_{\mathbb{R}^4}-\|\rho^n_{i,j}\|^2_{\mathbb{R}^4}\right)\right]=0. \quad (29)$$

• Let us prove that we may extract a subsequence $(p^{n_k})_k$ of $(p^n)_n$ such that $(p^{n_k})_k$ and $(p^{n_k+1})_k$ have the same limit.
As $(p^n)_n$ is bounded, we call $\bar{p}$ one cluster point and $(p^{n_k})_k$ the corresponding subsequence of $(p^n)_n$ such that $p^{n_k}\to\bar{p}$, as $k\to+\infty$. Let us call $\bar{p}'$ the limit (up to a subsequence) of p^{n_k+1}. We obtain with relation (25)

$$\bar{p}'_{i,j}=\frac{\bar{p}_{i,j}-\tau\left(H\left[H^*\bar{p}-u_d/\lambda\right]\right)_{i,j}}{1+\tau\left\|\left(H\left[H^*\bar{p}-u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}}. \quad (30)$$

We note $\bar{\rho}$ et $\bar{\eta}$ the respective limits of ρ^{n_k} et η^{n_k} when k tends to $+\infty$. With (29), we obtain

$$\sum_{i,j=1}^{N}\left[(1-\kappa^2\tau)\|\bar{\eta}_{i,j}\|^2_{\mathbb{R}^4}+\left(\left\|\left(H\left[H^*\bar{p}-u_d/\lambda\right]\right)_{i,j}\right\|^2_{\mathbb{R}^4}-\|\bar{\rho}_{i,j}\|^2_{\mathbb{R}^4}\right)\right]=0.$$

As the terms of the sum are nonnegative, we get

$$\forall i,j\quad (1-\kappa^2\tau)\|\bar{\eta}_{i,j}\|^2_{\mathbb{R}^4}=0\ \text{ and }\ \left\|\left(H\left[H^*\bar{p}-u_d/\lambda\right]\right)_{i,j}\right\|^2_{\mathbb{R}^4}-\|\bar{\rho}_{i,j}\|^2_{\mathbb{R}^4}=0.$$

- If $\kappa^2\tau<1$, then $\bar{\eta}_{i,j}=0$ for all i,j, and so $\bar{p}'=\bar{p}$.
- If $\kappa^2\tau=1$, then, for all i,j, $\|\bar{\rho}_{i,j}\|_{\mathbb{R}^4}=\left\|\left(H\left[H^*\bar{p}-u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}$, that is to say

$$\left\|\left(H\left[H^*\bar{p}-u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}\|\bar{p}'_{i,j}\|_{\mathbb{R}^4}=\left\|\left(H\left[H^*\bar{p}-u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}.$$

  This implies

$$\|\bar{p}'_{i,j}\|_{\mathbb{R}^4}=1\ \text{ or }\ \left\|\left(H\left[H^*\bar{p}-u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}=0.$$

  - If $\left\|\left(H\left[H^*\bar{p}-u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}=0$, then relation (30) gives $\bar{p}'_{i,j}=\bar{p}_{i,j}$.
  - If $\|\bar{p}'_{i,j}\|_{\mathbb{R}^4}=1$, then

$$1=\frac{\left\|\bar{p}_{i,j}-\tau\left(H\left[H^*\bar{p}-u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}}{1+\tau\left\|\left(H\left[H^*\bar{p}-u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}}\leqslant\frac{\|\bar{p}_{i,j}\|_{\mathbb{R}^4}+\left\|\tau\left(H\left[H^*\bar{p}-u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}}{1+\tau\left\|\left(H\left[H^*\bar{p}-u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}},$$

    therefore $\|\bar{p}_{i,j}\|_{\mathbb{R}^4}\geqslant 1$ and (together with $\|\bar{p}_{i,j}\|_{\mathbb{R}^4}\leqslant 1$) $\|\bar{p}_{i,j}\|_{\mathbb{R}^4}=1$. We deduce that

$$\left\|\bar{p}_{i,j}-\tau\left(H\left[H^*\bar{p}-u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}=\|\bar{p}_{i,j}\|_{\mathbb{R}^4}+\left\|\tau\left(H\left[H^*\bar{p}-u_d/\lambda\right]\right)_{i,j}\right\|_{\mathbb{R}^4}.$$

    Since triangular inequality turns to be an equality, there exists $\beta\in\mathbb{R}^*$ so that

$$\tau\left(H\left[H^*\bar{p}-u_d/\lambda\right]\right)_{i,j}=\beta\bar{p}_{i,j}.$$

    As $\left|\bar{p}'_{i,j}\right|=1$, relation (30) implies $\bar{p}'_{i,j}=\dfrac{1-\beta}{1+|\beta|}\bar{p}_{i,j}$, so that $\left|\dfrac{1-\beta}{1+|\beta|}\right|=1$; this yields $\beta\leqslant 0$ and $\bar{p}'_{i,j}=\bar{p}_{i,j}$.

Finally, $\bar{p} = \bar{p}'$, and

$$\forall i,j, \quad \left(H\left[\lambda H^* \bar{p} - u_d\right]\right)_{i,j} + \left\|\left(H\left[\lambda H^* \bar{p} - u_d\right]\right)_{i,j}\right\|_{\mathbb{R}^4} \bar{p}_{i,j} = 0.$$

This is the Euler equation for $(\mathcal{P}')$. Therefore $\bar{p}$ is a solution to $(\mathcal{P}')$. With uniqueness of the projection, we deduce that the sequence $(\lambda H^* p^n)_n$ converges to $P_{\lambda K_2}(u_d)$. We conclude with Lemma 1 since $\tau \leqslant 1/64$ then $\tau \leq 1/\kappa^2$.

■

6 Numerical results

In this section we briefly present numerical tests for the denoising process. A full comparison with existing methods will be performed in a forthcoming paper.

6.1 Examples

Throughout this section, we consider the following images that are degraded with a white Gaussian noise with standard deviation σ:

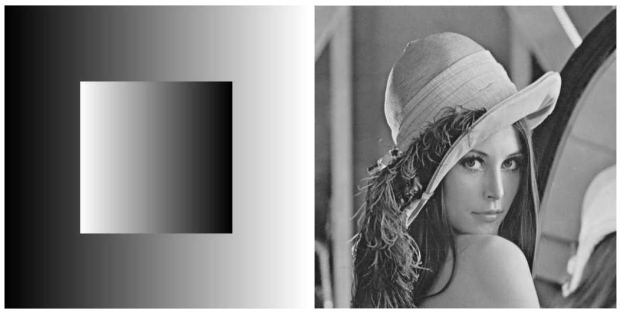


Fig. 1 Test images ("Shapes" (size : 740 × 740) and "Lena" (size : 512 × 512))

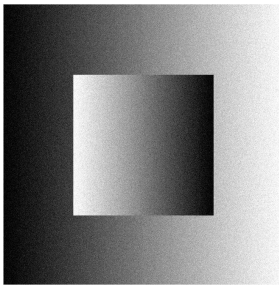
We perform numerical tests with different values of σ. In any case, we observe that the model is quite efficient for image restoration. Moreover, we note that we lose details information when parameter λ increases, what was expected. However, especially when the "observed" image is very noisy, we have a blurriness (subjective) feeling, that we do not have when restoration is performed with the standard ROF model. Checking what happens precisely on slices (lines) of the image (Figure 8 for example), we remark that the BV^2-model keeps contour information pretty well, anyway better than expected watching the image.

Numerical tests have been performed on different machines so that we cannot report precisely on the CPU time. However, the result is quite satisfactory after few iterations so that the process is not too slow. In what follows, the stopping criterion has been set to a maximal number of iterations `itmax` that can be chosen arbitrary large. The

algorithms have been implemented with MATLAB © software. We give al so the Signal to Noise Ratio (SNR)[1] : $SNR(v) = 20\log_{10}\left(\frac{\|u\|_{L^2}}{\|u-v\|_{L^2}}\right)$, where u is the expected image and v is the restored one. $SNR(u_d)$ gives the observed SNR (with the noisy input image).

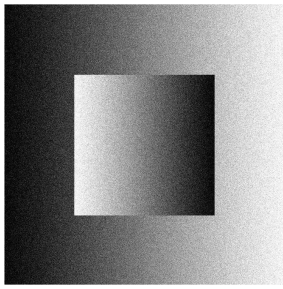
We have performed tests for two σ values. In the first example $\sigma = 0.15$ and we stopped after `itmax`=5000 iterations (Figure 3) and in the second case $\sigma = 0.25$ (the noise is more important). The noisy images are below:

(a) $\sigma = 0.15$ (SNR=21.28)



(b) $\sigma = 0.15$ (SNR=23.08)

(c) $\sigma = 0.25$ (SNR=16.90)

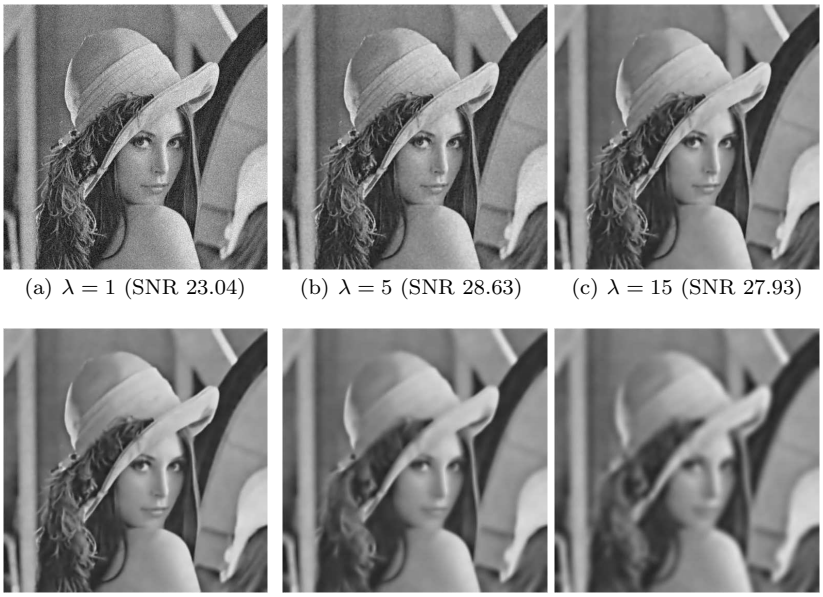


(d) $\sigma = 0.25$ (SNR=18.82)

Fig. 2 Noisy images - white Gaussian noise with standard deviation σ.

[1] There are different ways to compute this quantity with MATLAB © . We used the following syntax : `SNR = 20*log10(norm(u(:))/norm(u(:)-v(:)))`

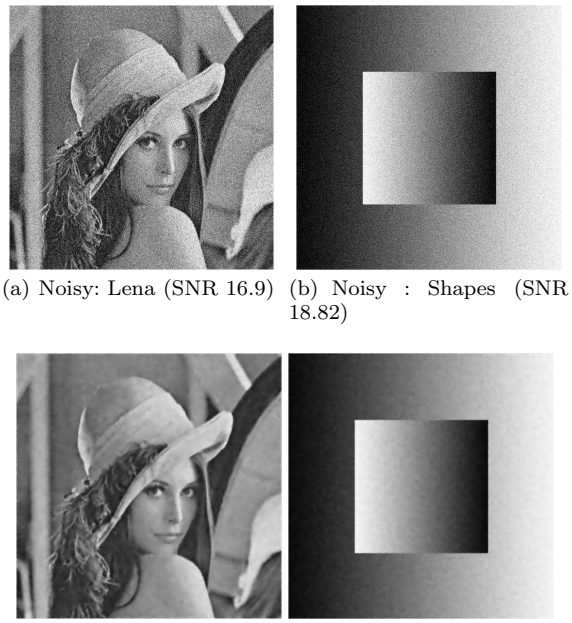
6.2 Sensibility with respect to λ parameter



(a) $\lambda = 1$ (SNR 23.04) (b) $\lambda = 5$ (SNR 28.63) (c) $\lambda = 15$ (SNR 27.93)

(d) $\lambda = 25$ (SNR 25.77) (e) $\lambda = 100$ (SNR 21.02) (f) $\lambda = 500$ (SNR 18.08)

Fig. 3 Solution v - Standard deviation $\sigma = 0.15$ - 5 000 iterations



(a) Noisy: Lena (SNR 16.9) (b) Noisy : Shapes (SNR 18.82)

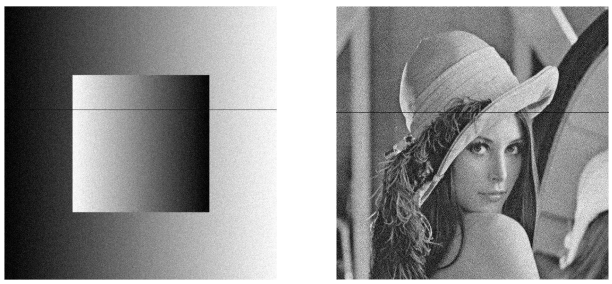
(c) SNR : 25.5 (d) SNR : 31.04

Fig. 4 Solution v - Standard deviation $\sigma = 0.25$ and $\lambda = 25$ - 5 000 iterations

As expected, we see on Figure 3 that the smoothing process is more efficient when λ is large. For both images, the result is satisfactory for $\lambda \simeq 10$.

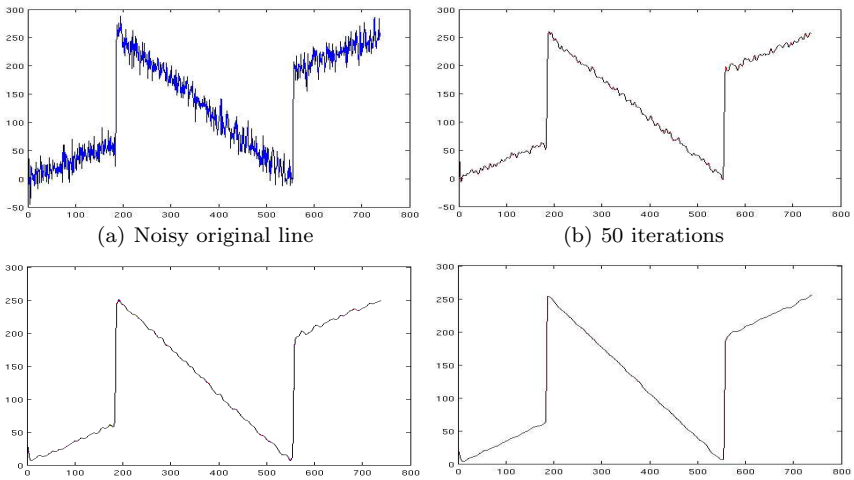
6.3 Sensitivity with respect to iterations number `itmax`

We fix $\lambda = 15$ and $\sigma = 0.15$. Figures 6 and 7 give the behavior of a slice (line) during iterations (we can see more easily how noise is removed). The algorithm converges well: the quality of restoration is improved as the number of iterations grows. Noise is removed and contours are preserved.



(a) Shape (b) Lena

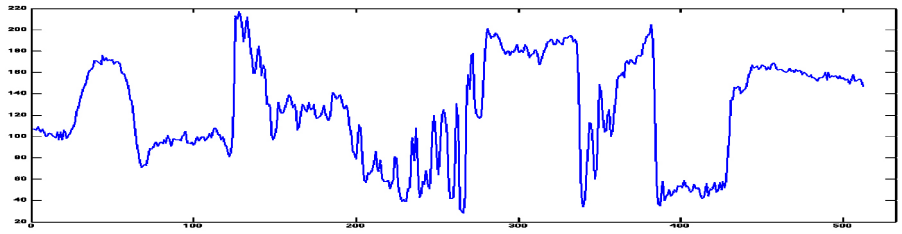
Fig. 5 Noisy images with line - $\sigma = 0.15$



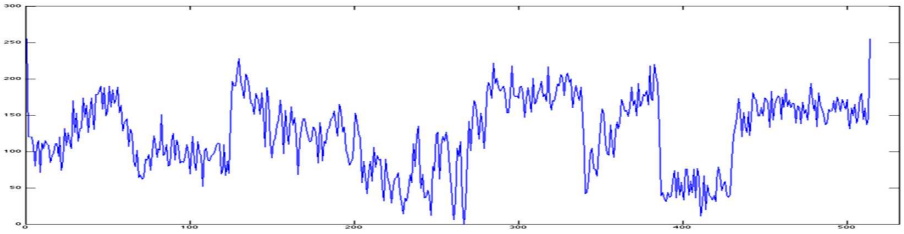
(a) Noisy original line (b) 50 iterations

(c) 500 iterations (d) 5000 iterations

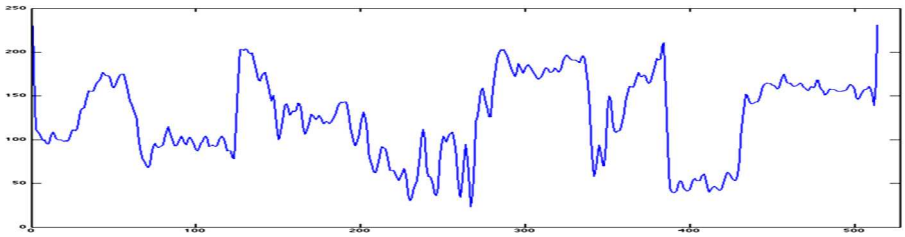
Fig. 6 Sensitivity with respect to the number of iterations - $\sigma = 0.15$, $\lambda = 50$ - Slice of "Shapes" image



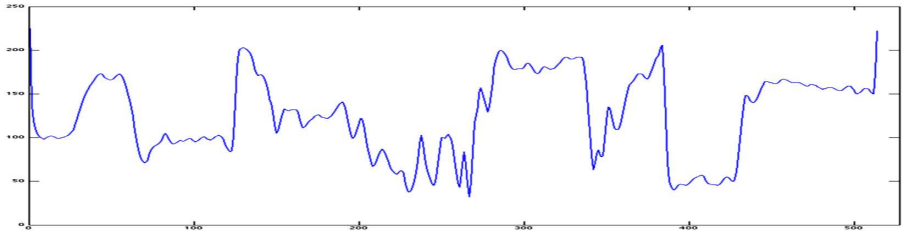
(a) Original slice



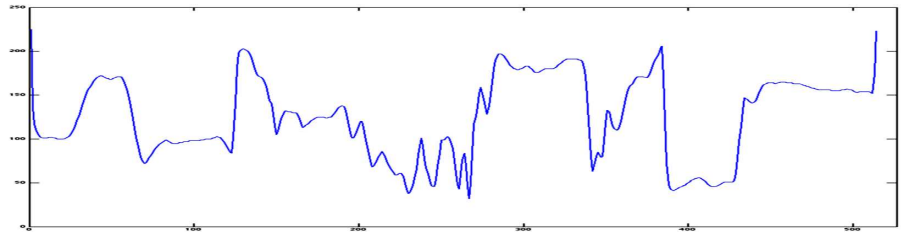
(b) Noisy slice



(c) 10 iterations



(d) 100 iterations



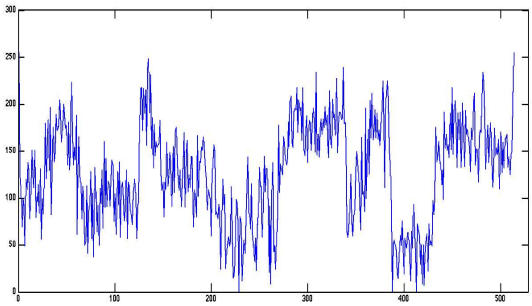
(e) 500 iterations

Fig. 7 Sensitivity with respect to the number of iterations - $\sigma = 0.15$, $\lambda = 15$ - Slice of "Lena" image

6.4 Comparison with the classical Rudin-Osher-Fatemi model

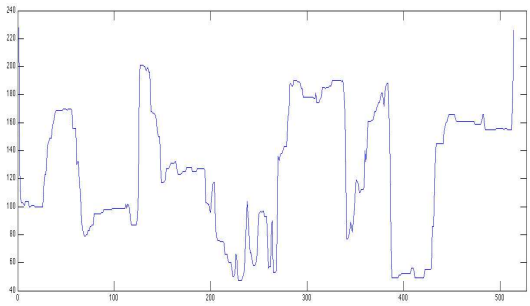
It is well known that the ROF model makes staircasing effect appear, since the resulting image is piecewise constant on large areas. We first compare the two models on test images that are not very noisy. In Figure 9 we see that piecewise constant areas appear with ROF, which is not the case with the BV^2 model. To focus on this fact, we have selected a line of the image that meets contours.

(a) Noisy image (SNR 16.9)



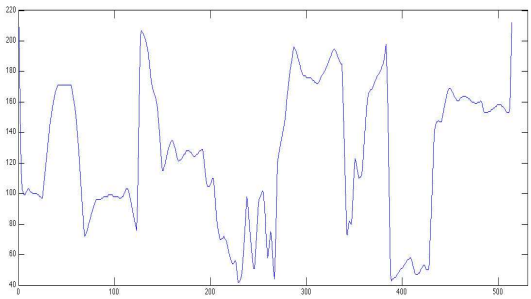
(b) Noisy slice

(c) ROF model solution (SNR 23.82)



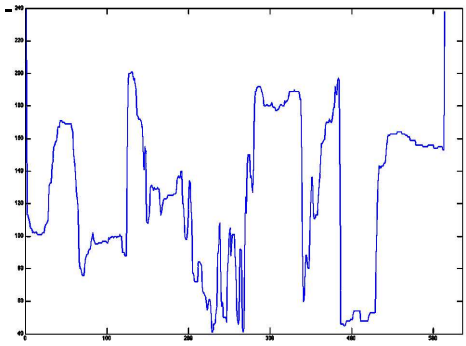
(d) ROF model slice

(e) BV^2 model solution - (SNR 25.5)

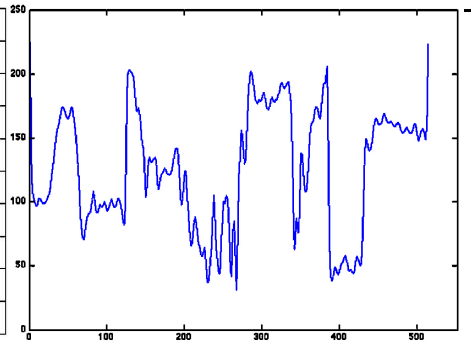


(f) BV^2 model slice

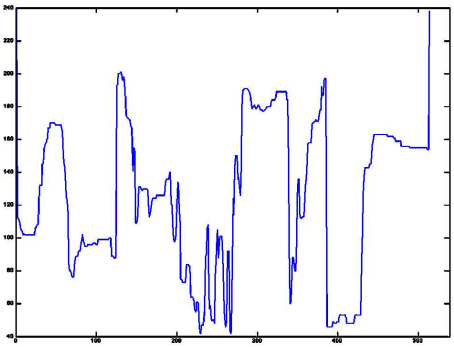
Fig. 8 Comparison between ROF and BV^2 models - $\sigma = 0.25$, $\lambda = 25$, 5 000 iterations.



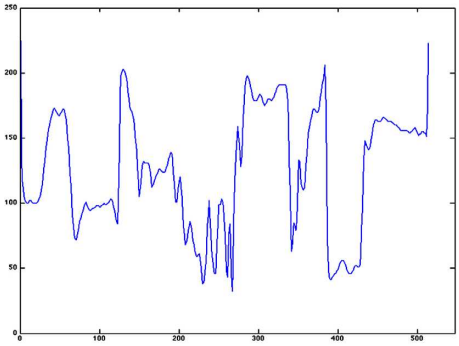
(a) ROF model - 50 iterations



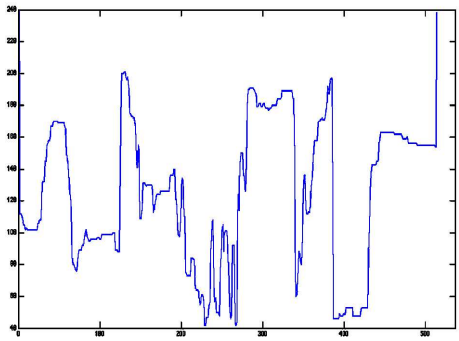
(b) BV^2 model- 50 iterations



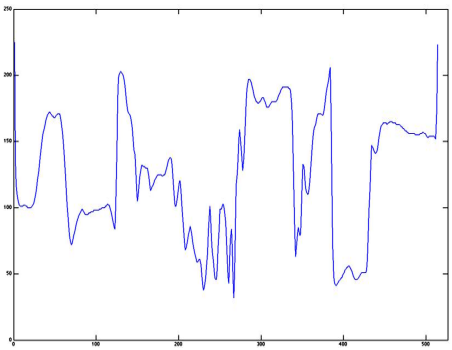
(c) ROF model - 300 iterations



(d) BV^2 model- 300 iterations

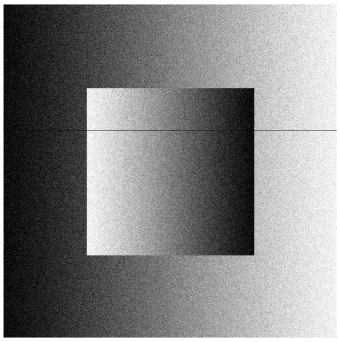


(e) ROF model - 500 iterations

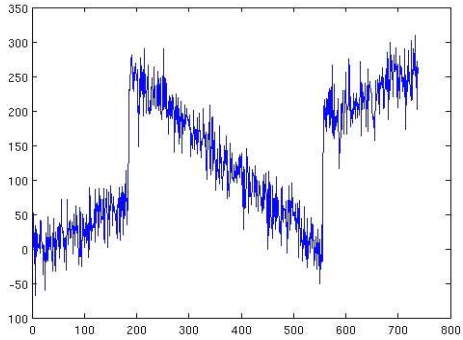


(f) BV^2 model- 500 iterations

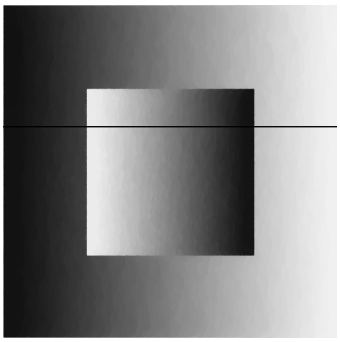
Fig. 9 Comparison between ROF and BV^2 models - $\sigma = 0.15$, $\lambda = 15$ (Lena slice).



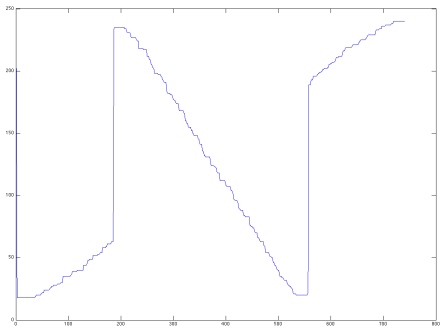
(a) Noisy image



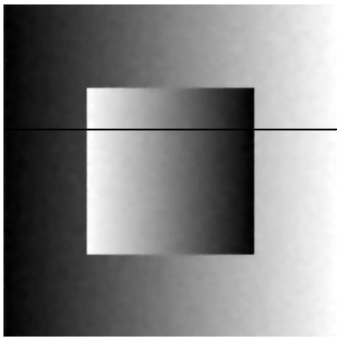
(b) Noisy slice



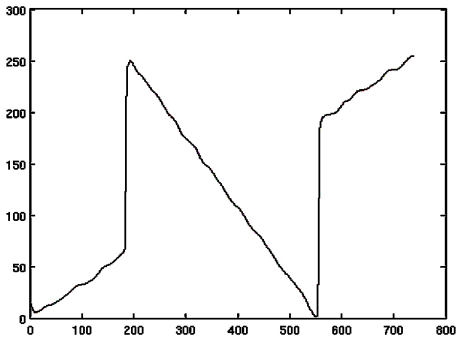
(c) ROF model image



(d) ROF model slice



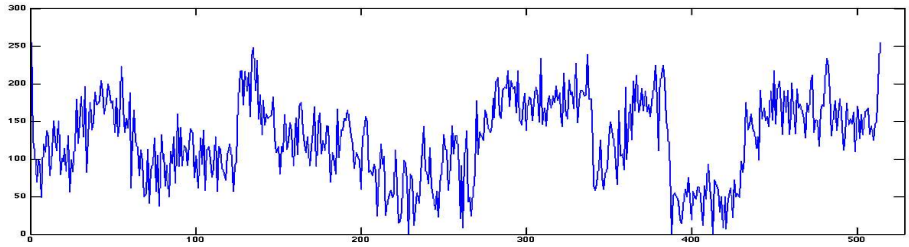
(e) BV^2 model image



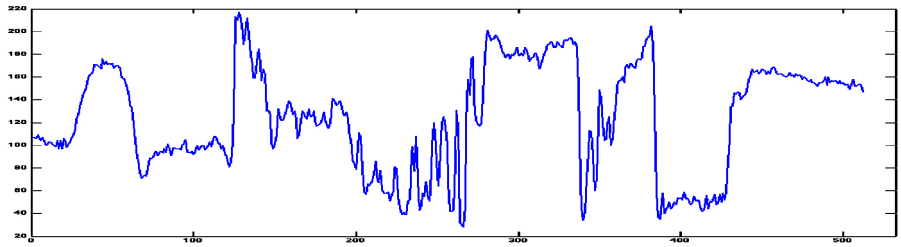
(f) BV^2 model slice

Fig. 10 Comparison between ROF and BV^2 models - $\sigma = 0.25$, $\lambda = 50$, 10 000 iterations

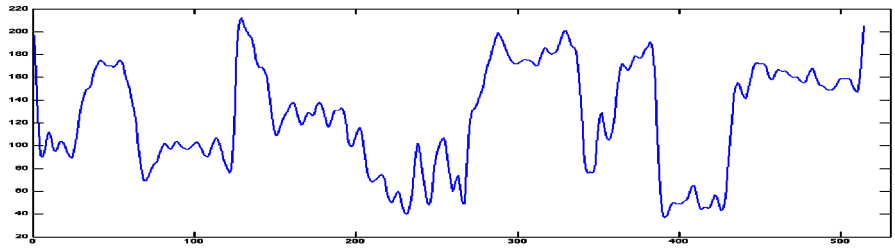
Figures 8 and 11 are obtained for $\lambda = 25$ and $\lambda = 50$ respectively and 5 000 iterations. Figures 10 and 12 are obtained for 10 000 iterations and $\lambda = 50$ and show precisely what happens: the image restored with ROF is clearly piecewise constant, and the BV^2 restored one seems to be blurred (Figure 12). However, this is an optical effect: considering a slice shows that the BV^2 model removes noise significantly and contours are better preserved: the amplitude of high peaks that correspond to contours is not changed, which is not the case in ROF-model (Figure 11).



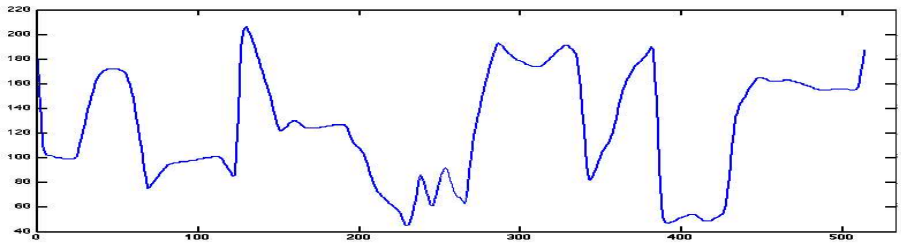
(a) Noisy slice



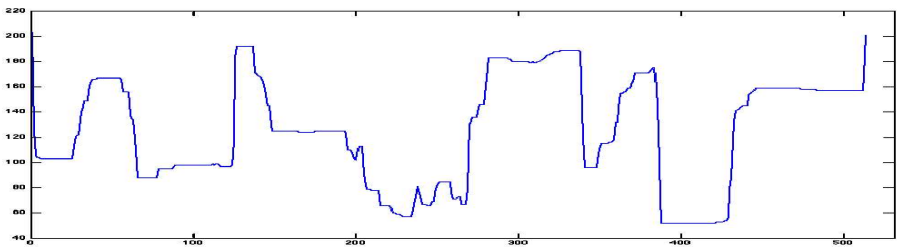
(b) Original slice



(c) BV^2 model - 50 iterations.



(d) BV^2 model - 5 000 iterations.



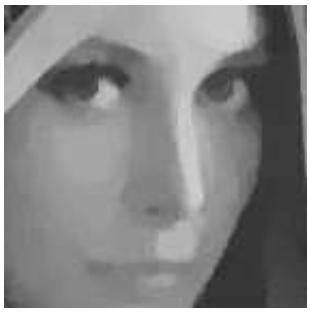
(e) ROF model - 5 000 iterations.

Fig. 11 Zoom on "Lena" slices- $\sigma = 0.25$, $\lambda = 50$

(a) Noisy image

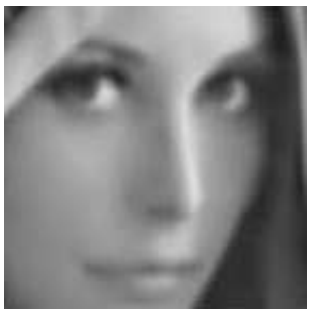
(b) Zoom

(c) ROF model image



(d) Zoom

(e) BV^2 model image

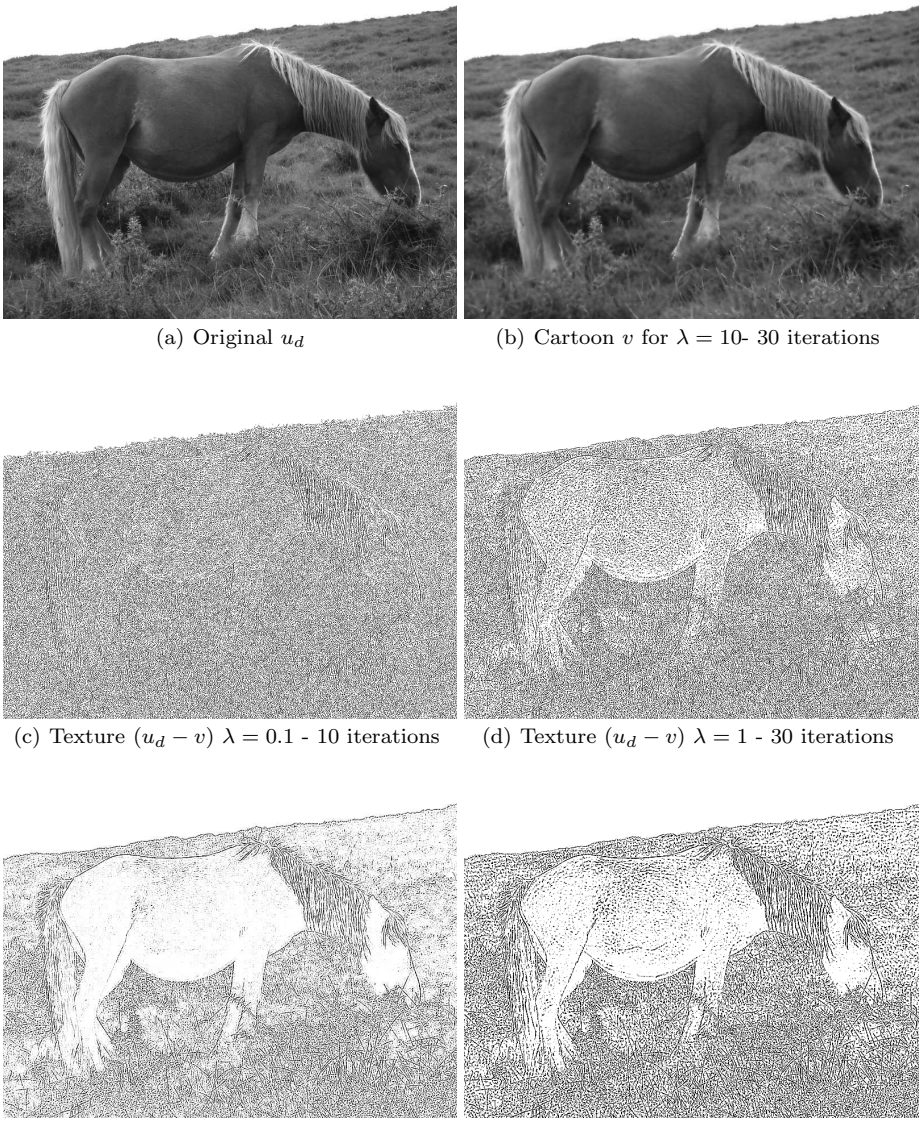


(f) Zoom

Fig. 12 Staircasing effect - $\sigma = 0.25$, $\lambda = 50$, 10 000 iterations.

6.5 Texture extraction

We do not report much on texture extraction process. The parameter tuning is slightly different but the algorithm behaves similarly (see [11]). Many variational methods have been developed for texture extraction (see [5,6] and the references therein). We shall precisely compare the BV^2 method to the existing ones in a forthcoming paper. We present an example in Figures 13 and 14.



(a) Original u_d

(b) Cartoon v for $\lambda = 10$- 30 iterations

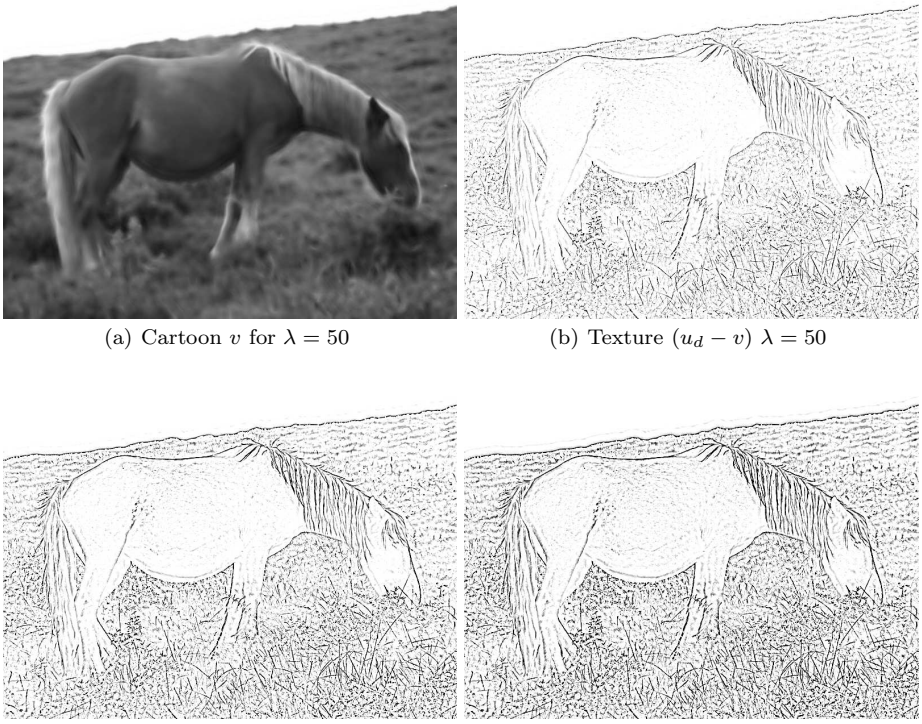
(c) Texture $(u_d - v)$ $\lambda = 0.1$ - 10 iterations

(d) Texture $(u_d - v)$ $\lambda = 1$ - 30 iterations

(e) Texture $(u_d - v)$ $\lambda = 5$ - 10 iterations

(f) Texture $(u_d - v)$ $\lambda = 5$ - 30 iterations

Fig. 13 Texture extraction - Rescaled texture image $u_d - v$ at 10 and 15 iterations



(a) Cartoon v for $\lambda = 50$ (b) Texture $(u_d - v)$ $\lambda = 50$

(c) Texture $(u_d - v)$ $\lambda = 10$ (d) Texture $(u_d - v)$ $\lambda = 5$

Fig. 14 Texture extraction - Rescaled texture image $u_d - v$ at 100 iterations

7 Conclusion

The second order approach via the BV^2 space seems promising. Many improvements have to be performed. The algorithm is still slow (though we get acceptable results for quite few iterations $\simeq 30$). We currently investigate the Nesterov- techniques to speed up the method (see Weiss et al. [25] or Fadili-Peyré [13]). Moreover, we have to look for modifications of the variational model using different norms (for example the L^1 norm) for the fitting data term. Furthermore, coupling existing techniques for texture extraction with the second order approach should give quite performing results. This will be done in future works.

References

1. Acar,R., Vogel, C.R.: Analysis of bounded variation penalty methods for ill-posed problems. Inverse Problems 10, no. 6, 1217–1229 (1994)
2. Ambrosio, L., Fusco, N., Pallara, D.: Functions of bounded variation and free discontinuity problems. Oxford mathematical monographs, Oxford University Press (2000).
3. Attouch, H., Briceño-Arias, L.M, Combettes, P.L.: A Parallel Splitting Method for Coupled Monotone Inclusions, SIAM J. Control Optim. 48 3246 (2010)
4. Attouch, H., Buttazzo, Michaille, G.: Variational analysis in Sobolev and BV spaces : applications to PDEs and optimization. MPS-SIAM series on optimization, 2006

5. Aubert, G., Aujol, J.F.: Modeling very oscillating signals, application to image processing. Applied Mathematics and Optimization, 51, no 2, 163–182 (2005)
6. Aubert, G., Aujol, J.F., Blanc-Feraud, L., Chambolle, A.:Image decomposition into a bounded variation component and an oscillating component. Journal of Mathematical Imaging and Vision 22, no 1, 71–88 (2005)
7. Aubert, G., Kornprobst, P.: Mathematical Problems in Image Processing, Partial Differential Equations and the Calculus of Variations. Applied Mathematical Sciences 147, Springer Verlag (2006).
8. Aujol, J.F.: Some First-Order Algorithms for Total Variation Based Image Restoration. J Math Imaging Vis (2009) 34: 307–327
9. Chambolle, A.: An algorithm for total variation minimization and applications. Journal of Mathematical Imaging and Vision, 20 89–97 (2004)
10. Demengel, F.:Fonctions à hessien borné. Annales de l'institut Fourier, tome 34, no 2, pp. 155-190 (1984)
11. Echegut, R., Piffet, L. : A variational model for image texture identification. `http://hal.archives-ouvertes.fr/hal-00439431/fr/`, submitted.
12. Ekeland, I., Temam, R. : Convex Analysis and Variational problems, SIAM Classic in Applied Mathematics, 28, 1999,
13. Fadili, J., Peyré, G. :Total Variation Projection with First Order Schemes, Preprint, `http://hal.archives-ouvertes.fr/hal-00380491/fr/`
14. Garnett, J. B., Le, Triet M., Meyer, Y. Vese, L. A.: Image decompositions using bounded variation and generalized homogeneous Besov spaces. Appl. Comput. Harmon. Anal., 23, no. 1, 25–56 (2007)
15. Hinterberger, W., Scherzer, O. :Variational methods on the space of functions of bounded Hessian for convexification and denoising. Computing, 76, no. 1-2, 109–133 (2006)
16. Hofmann, B., Kaltenbacher, B., Pöschl, C., Scherzer, O.: A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators. Inverse Problems, 23, no. 3, 987–1010 (2007)
17. Le, Triet M., Vese, L. A. : Image decomposition using total variation and div(BMO). Multiscale Model. Simul. 4, no. 2, 390–423 (2005)
18. Lieu, Linh H., Vese, L. A.: Image restoration and decomposition via bounded total variation and negative Hilbert-Sobolev spaces. Appl. Math. Optim. 58 (2008), no. 2, 167–193 (2008)
19. Meyer, Y.: Oscillating patterns in image processing and nonlinear evolution equations, Vol. 22 of University Lecture Series, AMS, 2002.
20. Osher, S., Fatemi, E, Rudin L..: Nonlinear total variation based noise removal algorithms. Physica D 60, 259–268 (1992)
21. Osher, S., Sole, A., Vese L.:, Image decomposition and restoration using total variation minimization and the H^1 norm. SIAM Journal on Multiscale Modeling and Simulation, 1-3, 349–370 (2003)
22. Osher, S.,Vese, L.: Modeling textures with total variation minimization and oscillating patterns in image processing. Journal of Scientific Computing 19, no 1-3, 553–572 (2003)
23. Osher, S. J, Vese, L. A., : Image denoising and decomposition with total variation minimization and oscillatory functions. Special issue on mathematics and image analysis. J. Math. Imaging Vision, 20, no. 1-2, 7–18 (2004)
24. Piffet, L. : Modèles variationnels pour l'extraction de textures 2D, PhD Thesis, Orléans, 2010
25. Weiss, P., Blanc-Féraud, L., Aubert, G.: Efficient schemes for total variation minimization under constraints in image processing, SIAM Journal on Scientific Computing, 2009, to appear.
26. Yin, W., Goldfarb, D., Osher, S.: A comparison of three total variation based texture extraction models. J. Vis. Commun. Image, 18, 240–252 (2007)