Mathematical and numerical analysis of a variational second order image decomposition model

Maïtine Bergounioux

▶ **To cite this version:**

Maïtine Bergounioux. Mathematical and numerical analysis of a variational second order image decomposition model. 2014. hal-01002958v1

HAL Id: hal-01002958

https://hal.science/hal-01002958v1

Preprint submitted on 7 Jun 2014 (v1), last revised 3 Apr 2015 (v3)

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.

Mathematical and numerical analysis of a variational second order image decomposition model

M. Bergounioux *

Abstract. We deal with a second order image decomposition model to perform denoising and texture extraction that was previously presented in [9]. We look for the decomposition $f = u + v + w$ where u is a first order term, v a second order term and w the remainder term (0 order). For highly textured images the model gives a two-scale texture decomposition: u can be viewed as a *macro-texture* (larger scale) which oscillations are not too large and w is the *micro-texture* (very oscillating) that contains the noise. Here, we perform mathematical analysis of the model and give qualitative properties of solutions using the dual problem and inf-convolution formulation. Then we perform numerical experiments, discuss the behavior of the model and investigate the lack of uniqueness.

Key words. Second order total variation, image decomposition, variational method, inf-convolution

AMS subject classifications. 65D18, 68U10, 65K10

1. Introduction . The most famous variational denoising model is the Rudin-Osher-Fatemi one ([1, 26]). 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 $u_d = w + u$ where w represents the oscillating component (noise or texture) and u is the *smooth* part. So we look for a solution u such that $u_d = w + u$ with $u \in BV(\Omega)$ and $w \in L^2(\Omega)$, where $BV(\Omega)$ is the functions of bounded variation space defined on an open subset $\Omega \subset \mathbb{R}^d$ ([3, 4, 19]). The regularization term involves only the so-called *cartoon* component u, while the remainder term $w := u_d - u$ represents the noise to be minimized.

A lot of people have investigated such decomposition models based on variational formulation, considering that an image can be decomposed into many components, each component describing a particular property of the image ([5, 7, 22, 23, 24, 28] and references therein for example).

In [9, 10] we have presented second order models where the (first order) classical total variation term has been replaced by a second order total variation term with the appropriate functional framework, namely the space of functions with bounded hessian introduced as $BH(\Omega)$ in [17] (and denoted $BV^2(\Omega)$ in [8, 10, 9]). The use of such a model allows to get rid of the *staircasing effect* that appears with the ROF model in denoising processes. However we had to involve a penalization term in the continuous setting to get existence results for the minimization problems.

Second order models have been investigated in the context of segmentation and inpainting problems with Mumford-Shah types functionals (see [2, 13, 14] for example). The functional framework is the so called $GSBV$ space composed of functions u whose truncated forms $(\min(-N, \max(u, N))$ belong to SBV_{loc} for every $N \in \mathbb{N})$. The definition of $GSBV^2$ is

* Université d'Orléans, UFR Sciences, Math., Labo. MAPMO, UMR 7349, Route de Chartres, BP 6759, 45067 Orléans cedex 2, maitine.bergounioux@univ-orleans.fr

slightly different from the one we consider since

$$GSBV^2(\Omega = \{u \in GSBV(\Omega) \mid \nabla u \in [GSBV(\Omega)]^d \} .$$

The aim of this paper is to give an existence result without any additional penalization term as in [9] and to perform a qualitative analysis of the model. Uniqueness et regularity issues will be also addressed.

More precisely, we assume that an image (in $L^2(\Omega)$) can be split in three components: a smooth (continuous) part v, a *cartoon* (piecewise constant) part u and an oscillating part w that should involve noise and/or fine textures. Such decompositions have already been investigated by Aujol and al. [5, 7]. These authors use the Meyer space of oscillating functions [21] rather than the $BH(\Omega)$ space (we shall present these spaces in the sequel). The model we propose here is different: the oscillating part of the image is not penalized but a priori included in the remainder term $w = u_d - u - v$, while v is the smooth part (in $BH(\Omega)$) and u belongs to $BV(\Omega)$: we hope u to be piecewise constant so that its jump set gives the image contours. For highly textured images, the model provides a two-scale texture decomposition: u can be viewed as a *macro-texture* (large scale) whose oscillations are not too large and w is the *micro-texture* (mmuch more oscillating) that contains the noise.

Therefore, we look for components u, v and w that belong to different spaces: u belongs to $BV(\Omega)$ (and if possible not to $W^{1,1}(\Omega)$), $v \in BH(\Omega)$ and $w \in L^2(\Omega)$. This last component $w = u_d - u - v$ lies in the same space as the observed image u_d.

The paper is organized as follows. We first present the functional framework and perform a quick comparison between the second-order total variation we use and the one defined by Bredies et al. in [11]. In section 3, we present the variational model, give existence result and an equivalent formulation with inf-convolution. This allows to compute the dual problem. Next section is devoted to giving qualitative properties of the solutions . We end with numerical experiments.

2. Functional framework for second order variational analysis .

2.1. Spaces $BV(\Omega)$ and $BH(\Omega)$. In the whole paper, Ω is an open bounded subset of $\mathbb{R}^d$ (practically $d = 2$) smooth enough (with the cone property and Lipschitz for example). More precisely, if $d = 2$, Ω may satisfy next assumption

$$\left\{ \begin{array}{c} \Omega \text{ is a bounded connected open set, strongly Lipschitz such that} \\ \partial\Omega \text{ is the union of finitely many } \mathcal{C}^2 \text{ curves} \end{array} \right. \tag{2.1}$$

Following [3, 4, 6] and [10, 17], we recall the definitions and main properties of the spaces of functions of first and second order bounded variation. The space $BV(\Omega)$ is the classical Banach space of functions of bounded variation defined by

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

where $TV(u)$ is the total variation of u

$$TV(u) := \sup \left\{ \int_\Omega u(x) \,\mathrm{div}\, \xi(x)\, dx \mid \xi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^d),\ \|\xi\|_\infty \leq 1 \right\}, \tag{2.2}$$

endowed with the norm $\|u\|_{BV(\Omega)} = \|u\|_{L^1} + TV(u)$.
We say that a sequence $(u_n)_{n\in\mathbb{N}}$ of $BV(\Omega)$ converges to some $u \in BV(\Omega)$ for the *intermediate* (or *strict*) convergence if u_n strongly converges to u for the $L^1(\Omega)$ topology and $TV(u_n)$ converges to $TV(u)$ (in $\mathbb{R}$) (see [3, 4, 29]).
The space of functions with bounded hessian has been introduced by Demengel [17] (where it was denoted $BH(\Omega)$). It is the space of $W^{1,1}(\Omega)$ functions such that $TV^2(u) < +\infty$, where

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

∇u stands for the first order derivative of u in the sense of distributions and

$$TV^2(u) := \sup\left\{ \int_\Omega \langle \nabla u, div(\xi)\rangle_{\mathbb{R}^d} \mid \xi \in \mathcal{C}_c^2(\Omega, \mathbb{R}^{d\times d}),\ \|\xi\|_\infty \le 1 \right\} < \infty, \tag{2.3}$$

is the second order total variation of u . Here, $div(\xi) = (div(\xi_1), div(\xi_2), \dots, div(\xi_d))$, and

$$\forall i,\ \xi_i = (\xi_{i,1}, \xi_{i,2}, \dots, \xi_{i,d}) \in \mathbb{R}^d, \qquad div(\xi_i) = \sum_{j=1}^{d} \frac{\partial \xi_{i,j}}{\partial x_j}.$$

The space $BH(\Omega)$ endowed with the following norm

$$\|f\|_{BH(\Omega)} := \|f\|_{W^{1,1}(\Omega)} + TV^2(f) = \|f\|_{L^1} + \|\nabla f\|_{L^1} + TV^2(f), \tag{2.4}$$

where TV^2 is given by (2.3) is a Banach space. Note that a function u belongs to $BH(\Omega)$ if and only if $u \in W^{1,1}(\Omega)$ and $\dfrac{\partial u}{\partial x_i} \in BV(\Omega)$ for $i \in \{1, \dots, d\}$. In particular

$$TV^2(u) \le \sum_{i=1}^{d} TV\left(\frac{\partial u}{\partial x_i}\right) \le d\ TV^2(u).$$

We give thereafter important properties of these spaces which proofs can be found in [3, 4, 10, 12, 17] for example.

Theorem 2.1. *[Semi-continuity of total variation]*

i. The mapping $u \mapsto TV(u)$ is lower semi-continuous (denoted in short lsc) from $BV(\Omega)$ to $\mathbb{R}^+$ for the $L^1(\Omega)$ topology.

ii. The operator TV^2 is lower semi-continuous from $BH(\Omega)$ endowed with the strong topology of $W^{1,1}(\Omega)$ to $\mathbb{R}$.

Theorem 2.2. *[Embedding results]Assume $d \ge 2$. Then*

i. $BH(\Omega) \hookrightarrow W^{1,q}(\Omega)$ with $q \le \dfrac{d}{d-1}$, with continuous embedding. Moreover the embedding is compact if $q < \frac{n}{n-1}$. In particular

$$BH(\Omega) \hookrightarrow L^q(\Omega), \quad \forall q \in [1, \infty[, \quad \text{if } d = 2.$$

ii. If $d = 2$

– $BV(\Omega) \subset L^2(\Omega)$ with continuous embedding.

– $BV(\Omega) \subset L^p(\Omega)$ *with compact embedding, for every* $p \in [1, 2)$.

iii. If $d = 2$ *and if* Ω *satisfies assumption* (2.1) *then* $BH(\Omega) \subset \mathcal{C}^0(\bar{\Omega})$.

So $BH(\Omega) \subset H^1(\Omega)$ with continuous embedding and $BH(\Omega) \subset W^{1,1}(\Omega)$ with compact embedding. Let us define the space $BV_0(\Omega)$ as the space of functions of bounded variation that vanish on the boundary $\partial\Omega$ of Ω. More precisely as Ω is bounded and $\partial\Omega$ is Lipschitz, functions of $BV(\Omega)$ have a trace of class L^1 on $\partial\Omega$ [3, 4, 29], and the trace mapping $T : BV(\Omega) \to L^1(\partial\Omega)$ is linear, continuous from $BV(\Omega)$ equipped with the intermediate convergence to $L^1(\partial\Omega)$ endowed with the strong topology ([4] Theorem 10.2.2 p 386). The space $BV_0(\Omega)$ is then defined as the kernel of T. It is a Banach space, endowed with the induced norm:

$$BV_0(\Omega) := \{u \in BV(\Omega) \mid u_{|\partial\Omega} = 0 \} .$$

In addition , if $u \in BH(\Omega)$ we may define the trace $u_{|\partial\Omega} \in W^{1,1}(\partial\Omega)$ and the normal derivative $\dfrac{\partial u}{\partial n} \in L^1(\partial\Omega)$ (Theorem 2.9 [12]). So we may define similarly

$$BH_0(\Omega) := \{u \in BH(\Omega) \mid \frac{\partial u}{\partial x_i} = 0 \text{ on } \partial\Omega,\ i = 1, \cdots, d\} .$$

We set also

$$BV_m(\Omega) := \{u \in BV(\Omega) \mid \int_\Omega u(x)\, dx = 0\ i = 1, \cdots, n\} ,$$

and

$$BH_m(\Omega) := \{u \in BH(\Omega) \mid \int_\Omega \frac{\partial u}{\partial x_i}\, dx = 0\ i = 1, \cdots, d\} .$$

The Ostrograski formula gives

$$\int_\Omega \frac{\partial u}{\partial x_i} dx = -\int_{\partial\Omega} u_i\, n_i \ ,$$

where u_i is the i^{th} partial function with respect to the i^{th} coordinate and $n = (n_1, \cdots, n_d)$ is the outer normal vector. In particular, if $u = 0$ on $\partial\Omega$, then $u \in BH_m(\Omega)$. At last we shall use the following result of [8]:

Lemma 2.3. *Let* $\Omega \subset \mathbb{R}^n$ *be an open Lipschitz bounded set.There exist generic constants only depending on* Ω, $C_1, C_2 > 0$ *such that*

$$\forall u \in BV_m(\Omega) \qquad \|u\|_{L^1(\Omega)} \leq C_1 TV(u), \tag{2.5}$$

$$\forall u \in BH_0(\Omega) \cup BH_m(\Omega) \qquad TV(u) \leq C_2 TV^2(u) \tag{2.6}$$

2.2. Comparison with BGV^2. Another definition for second-order total variation spaces has been set in [11]. The main difference lies in the choice of the test functions for the weak variational formulation. The authors define the *Total Generalized Variation* $TGV^2(u)$ as the supremum of the duality product between u and symmetric tests functions that are bounded

together with their derivative. First, we note that we may define $TV^2(u)$ in a equivalent way as following: for any $\xi \in \mathcal{C}_c^2(\Omega, \mathbb{R}^{d\times d})$ recall that

$$\forall i,\ \xi_i = (\xi_{i,1}, \xi_{i,2}, \dots, \xi_{i,d}) \in \mathbb{R}^d, \qquad div(\xi_i) = \sum_{j=1}^{d} \frac{\partial \xi_{i,j}}{\partial x_j}$$

and define as in [11]

$$\mathrm{div}^2 \xi := \sum_{i,j=1}^{d} \frac{\partial^2 \xi_{i,j}}{\partial x_i \partial x_j} \ .$$

Let us call

$$\mathcal{B} := \left\{ \xi \in \mathcal{C}_c^2(\Omega, \mathbb{R}^{d\times d}),\ \|\xi\|_\infty \le 1 \right\}.$$

Then, for every function $u \in W^{1,1}(\Omega)$

$$TV^2(u) := \sup \left\{ \int_\Omega u \,\mathrm{div}^2 \xi \, dx, \xi \in \mathcal{B} \right\}, \tag{2.7}$$

Indeed, an integration by parts gives

$$\int_\Omega u \,\mathrm{div}^2 \xi \, dx = -\int_\Omega (\nabla u, \mathrm{div}\ \xi)_{\mathbb{R}^d} \, dx \ .$$

Let be $\alpha = (\alpha_0, \alpha_1) > 0$, we call

$$TGV_\alpha^2(u) = \sup \left\{ \int_\Omega u \,\mathrm{div}^2 \xi \, dx, \xi \in \mathcal{B}_\alpha \right\},$$

where

$$\mathcal{B}_\alpha := \{ \xi \in \mathcal{K},\ \xi_{ij} = \xi_{ji} \ \ \forall i,j,\ \|\xi\|_\infty \le \alpha_0,\ \|\mathrm{div}\ \xi\|_\infty \le \alpha_1 \ \}.$$

We may define ([11])

$$BGV_\alpha^2(\Omega) = \left\{ u \in L^1(\Omega) \ , TGV_\alpha^2(u) < +\infty \ \right\}. \tag{2.8}$$

Proposition 2.1. *Let be $\alpha = (\alpha_0, \alpha_1) > 0$. For every function u in $W^{1,1}(\Omega)$ we get*

$$TGV_\alpha^2(u) \le \alpha_0 TV^2(u) \ .$$

Therefore

$$\forall \alpha > 0 \qquad BH(\Omega) \subset BGV_\alpha^2(\Omega)$$

with continuous embedding.

Proof. As $\mathcal{B}_\alpha \subset \alpha_0 \mathcal{B}$ the first relation is obvious. Moreover if $u \in BH(\Omega)$, then $u \in W^{1,1}(\Omega)$ and $TGV_\alpha^2(u) < +\infty$. In addition

$$\|u\|_{BVG_\alpha^2} = \|u\|_{L^1} + TGV_\alpha^2(u) \le \|u\|_{W^{1,1}} + \alpha_0 TV^2(u) \le \max(1, \alpha_0)\|u\|_{BH} \ ,$$

which gives the continuity of the embedding mapping. ■

Corollary 2.1. *For any $u \in BH(\Omega)$, $TV^2(u) = 0$ if and only if u is a polynomial function of order 1.*

Proof. For any $u \in BH(\Omega)$, $TV^2(u) = 0 \Longrightarrow TGV^2_\alpha(u) = 0$. Then we use Proposition 3.3 of [11]. ■

The main difference between the two approaches concerns the functions regularity. The $BGV^2(\Omega)$ functions do not necessarily belong to $L^1(\Omega)$. In particular, the indicator function of smooth open sets belong to $BGV^2(\Omega)$ and not to $BH(\Omega)$. On the other hand, we cannot have Sobolev-type embeddings for $BGV^2(\Omega)$.

3. A second-order variational model for image decomposition .

3.1. Presentation of the model. We have already presented this model in [9] so that we do no detail so much. However we provide here an existence result that was expected but only proved in the finite dimensional case. We now assume that the u_d belongs to $L^2(\Omega)$ and that the image we want to recover can be decomposed as $u_d = w + u + v$ where u, v and w are functions that characterize different parts of u_d. Components belong to different functional spaces: v is the (smooth) second order part and belongs to $BH(\Omega)$, u is a $BV(\Omega)$ component and $w \in L^2(\Omega)$ is the remainder term. We consider the following cost functional defined on $BV(\Omega) \times BH(\Omega)$:

$$\mathcal{F}_{\lambda,\mu}(u,v) = \frac{1}{2}\|u_d - u - v\|^2_{L^2(\Omega)} + \lambda TV(u) + \mu TV^2(v), \tag{3.1}$$

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

$$\inf\{\ \mathcal{F}_{\lambda,\mu}(u,v) \mid (u,v) \in BV(\Omega) \times BH_0(\Omega)\ \} \tag{$\mathcal{P}_{\lambda,\mu}$}$$

Remark 3.1. *We decide to look for the minima of $\mathcal{F}_{\lambda,\mu}$ on $BV(\Omega) \times BH_0(\Omega)$ and not $BV(\Omega) \times BH(\Omega)$ to get an existence result. This will cause troubles to set the dual problem because of the computation of Legendre-Fenchel conjugate functions. Nevertheless, the constraint $v \in BH_0(\Omega)$ (that is $\dfrac{\partial v}{\partial n} = 0$ on $\partial\Omega$) is a usual one in image processing and the difficulty will be overcome in the discrete setting.*

We expect v to be the smooth *colored* part of the image, u to be a $BV(\Omega)\backslash BH(\Omega)$ function which derivative is a measure supported by the contours and $w := u_d - u - v \in L^2$ is the noise and/or small textures (we shall detail this point later). First, we give an existence result for problem $(\mathcal{P}_{\lambda,\mu})$.

Theorem 3.1 (Existence). *The problem $(\mathcal{P}_{\lambda,\mu})$ has at least an optimal solution (u^*, v^*) in $BV(\Omega) \times BH_0(\Omega)$.*

Proof. We first prove that the auxiliary problem

$$\inf\{\ \mathcal{F}_{\lambda,\mu}(u,v) \mid (u,v) \in BV_m(\Omega) \times BH_0(\Omega)\ \} \tag{3.2}$$

has an optimal solution. Let $(u_n, v_n) \in BV_m(\Omega) \times BH_0(\Omega)$ be a minimizing sequence, i.e.

$$\lim_{n\to+\infty} \mathcal{F}_{\lambda,\mu}(u_n, v_n) = \inf\{\ \mathcal{F}_{\lambda,\mu}(v) \mid (u,v) \in BV_m(\Omega) \times BH_0(\Omega)\ \} < +\infty.$$

Therefore

- $TV^2(v_n)$ is bounded and with lemma 2.3, $\|\nabla v_n\|_{L^1}$ is bounded as well.
- $TV(u_n)$ is bounded. Using once again lemma 2.3 this yields that u_n is bounded in $L^1(\Omega$. Therefore the sequence u_n is bounded in $BV(\Omega)$.
- As $u_n + v_n$ is L^2 -bounded, it is L^1 -bounded as well so that v_n is L^1 bounded. As $\|\nabla v_n\|_{L^1}$ and $TV^2(v_n)$ are bounded this means that the sequence v_n is bounded in $BH(\Omega)$.

With the compactness result of Theorem 2.2, this yields that $(v_n)_{n\in\mathbb{N}}$ strongly converges (up to a subsequence) in $W^{1,1}(\Omega)$ to $v^* \in BH(\Omega)$. Moreover, $v^* \in BH_0(\Omega)$ because the trace operator is continuous [12, 17]. Similarly $(u_n)_{n\in\mathbb{N}}$ strongly converges (up to a subsequence) in $L^1(\Omega)$ to $u^* \in BV_m(\Omega)$. Moreover $u_n + v_n$ weakly converges to $u^* + v^*$ in $L^2(\Omega)$. With theorem 2.1 we get

$$TV(u^*) \leq \liminf_{n\to+\infty} TV(u_n),\ TV^2(v^*) \leq \liminf_{n\to+\infty} TV^2(v_n).$$

So

$$\mathcal{F}_{\lambda,\mu}(u^*, v^*) \leq \liminf_{n\to+\infty} \mathcal{F}_{\lambda,\mu}(u_n, v_n) = \min_{(u,v)\in BV_m(\Omega)\times BH_0(\Omega)} \mathcal{F}_{\lambda,\mu}(u, v),$$

and (u^*, v^*) is a solution to (3.2).
For every $(u, v) \in BV(\Omega) \times BH_0(\Omega)$, we have $(u - \bar{u}, v + \bar{u}) \in BV_m(\Omega) \times BH_0(\Omega)$ where $\bar{u} = \dfrac{1}{|\Omega|}\displaystyle\int_\Omega u$ is the mean value of u. Moreover

$$\mathcal{F}_{\lambda,\mu}(u, v) = \mathcal{F}_{\lambda,\mu}(u - \bar{u}, v + \bar{v}) \geq \mathcal{F}_{\lambda,\mu}(u^*, v^*).$$

Therefore (u^*, v^*) is an optimal solution to $(\mathcal{P}_{\lambda,\mu})$. ■

Remark 3.2. *Uniqueness of the solution is challenging. We shall prove partial results in section 4.2.*

3.2. Optimality conditions. In what follows, we fix $\lambda > 0$ and $\mu > 0$ and set for any $u \in L^2(\Omega) : \mathcal{N}(u) = \dfrac{1}{2}\|u\|_2^2$,

$$\Phi^1_\lambda(u) = \begin{cases} \lambda TV(u) & \text{if } u \in BV(\Omega) \\ +\infty & \text{else.} \end{cases}$$

and

$$\Phi^2_\mu(v) = \begin{cases} \mu TV^2(v) & \text{if } v \in BH_0(\Omega) \\ +\infty & \text{else.} \end{cases}$$

It is easy to see that $(\bar{u}, \bar{v})$ is a solution to $(\mathcal{P}_{\lambda,\mu})$ if and only if

$$\bar{u} = \operatorname{argmin}\left\{ \frac{1}{2}\|u_d - \bar{v} - u\|^2 + \Phi^1_\lambda(u),\ u \in L^2(\Omega) \right\}, \tag{3.3}$$

$$\bar{v} = \operatorname{argmin}\left\{ \frac{1}{2}\|u_d - \bar{u} - v\|^2 + \Phi^2_\mu(v),\ v \in L^2(\Omega) \right\}. \tag{3.4}$$

and we may derive optimality conditions in a standard way :

Theorem 3.2. *$(\bar u, \bar v)$ is a solution to $(\mathcal{P}_{\lambda,\mu})$ if and only if*

$$\bar w := u_d - \bar u - \bar v \in \partial\Phi^1_\lambda(\bar u) \cap \partial\Phi^2_\mu(\bar v). \tag{3.5}$$

The proof is obvious. Here $\partial f(u)$ stands for the subdifferential of f at u where $f : V \to \mathbb{R}$:

$$\partial f(u) = \{u^* \in V' \mid \forall v \in V \quad f(v) - f(u) \geq \langle u^*, v - u\rangle \},$$

and $\langle\cdot,\cdot\rangle$ is the duality product between V et V'.

3.3. Inf-convolution formulation . We are going to interpret $(\mathcal{P}_{\lambda,\mu})$ as successive inf-convolution processes. Recall that the inf-convolution ([4] p 324) is defined as

$$(f\#g)(v) = \inf\{f(u) + g(v-u),\ u \in V\}\ ,$$

where $f, g : V \to \mathbb{R} \cup \{+\infty\}$.

Lemma 3.3. *The functionals $\mathcal{N}\#\Phi^1_\lambda$ and $\mathcal{N}\#\Phi^2_\mu$ are convex, continuous from $L^2(\Omega)$ to $L^2(\Omega)$.*

Proof. In the sequel we set $\Phi = \Phi^1_\lambda$ or Φ^2_μ indifferently. As Φ and $\mathcal{N}$ are convex so is $\mathcal{N}\#\Phi$ (see [20] for example). Let be $u \in L^2(\Omega)$:

$$(\mathcal{N}\#\Phi)(u) = \inf_{v\in L^2(\Omega)} \frac{1}{2}\|u-v\|_2^2 + \Phi(v) \leq \frac{1}{2}\|u\|_2^2 + \Phi(0) = \frac{1}{2}\|u\|_2^2\ .$$

As $(\mathcal{N}\#\Phi)(0) = 0$ this gives the $\mathcal{N}\#\Phi$ continuity at 0 and its boundedness in a neighborhood of 0. As it is convex, it continuous on its whole domain $L^2(\Omega)$(see [18] for example). ■

Note that problem (3.3) is equivalent to $\bar u \in \mathcal{N}\#\Phi^1_\lambda(u_d - \bar v)$ and (3.4) is equivalent to $\bar v \in \mathcal{N}\#\Phi^2_\mu(u_d - \bar u)$. In fact, problem $(\mathcal{P}_{\lambda,\mu})$ can be written as successive inf-convolution processes. More precisely we have

Theorem 3.4. *Let $(\bar u, \bar v) \in BV(\Omega) \times BH_0(\Omega)$ be a solution to $(\mathcal{P}_{\lambda,\mu})$ and $\bar m := \inf(\mathcal{P}_{\lambda,\mu})$. Then*

$$\begin{aligned}\bar m &= \mathcal{N}(u_d - \bar u - \bar v) + \Phi^1_\lambda(\bar u) + \Phi^2_\mu(\bar v)\\ &= (\mathcal{N}\#\Phi^1_\lambda))(u_d - \bar v) + \Phi^2_\mu(\bar v) = (\mathcal{N}\#\Phi^2_\mu))(u_d - \bar u) + \Phi^1_\lambda(\bar u)\\ &= (\Phi^1_\lambda\#(\mathcal{N}\#\Phi^2_\mu))(u_d) = (\Phi^2_\mu\#(\mathcal{N}\#\Phi^1_\lambda))(u_d).\end{aligned}$$

Proof. Let $(\bar u, \bar v) \in BV(\Omega) \times BH_0(\Omega)$ be a solution to $(\mathcal{P}_{\lambda,\mu})$.
Then, for every $(u,\ v) \in BV(\Omega) \times BH_0(\Omega)$, we get

$$\bar m = \mathcal{N}(u_d - \bar u - \bar v) + \Phi^1_\lambda(\bar u) + \Phi^2_\mu(\bar v) \leq \mathcal{N}(u_d - u - v) + \Phi^1_\lambda(\bar u) + \Phi^2_\mu(v)\ . \tag{3.6}$$

This gives, for every $v \in BH_0(\Omega)$

$$\bar m \leq \inf_{u\in L^2(\Omega)} \mathcal{N}(u_d - u - v) + \Phi^1_\lambda(\bar u) + \Phi^2_\mu(v) = (\mathcal{N}\#\Phi^1_\lambda)(u_d - v) + \Phi^2_\mu(v)$$

so that

$$\bar m \leq \inf_{v\in BH_0(\Omega)} (\mathcal{N}\#\Phi^1_\lambda)(u_d - v) + \Phi^2_\mu(v) \leq (\Phi^2_\mu\#(\mathcal{N}\#\Phi^1_\lambda))(u_d)\ .$$

Similarly

$$\bar{m} \leq (\Phi_\lambda^1 \# (\mathcal{N} \# \Phi_\mu^2))(u_d) \ .$$

Conversely, by definition of inf-convolution, we get for every $(u,\ v) \in BV(\Omega) \times BH_0(\Omega)$

$$(\Phi_\mu^2 \# (\mathcal{N} \# \Phi_\lambda^1))(u_d) \leq (\mathcal{N} \# \Phi_\lambda^1)(u_d - v) + \Phi_\mu^2(v) \leq \mathcal{N}(u_d - v - u) + \Phi_\lambda^1(u) + \Phi_\mu^2(v) \ ,$$

so that $(\Phi_\mu^2 \# (\mathcal{N} \# \Phi_\lambda^1))(u_d) \leq \bar{m}$.

We finally obtain $\bar{m} = (\Phi_\lambda^1 \# (\mathcal{N} \# \Phi_\mu^2))(u_d) = (\Phi_\mu^2 \# (\mathcal{N} \# \Phi_\lambda^1))(u_d)$. ■

3.4. Computing Fenchel conjugate function. We are going to write the dual problem of $(\mathcal{P}_{\lambda,\mu})$ and we need to compute the conjugate functions of Φ_λ^1 and Φ_μ^2 and $\tilde{f} : u \mapsto f(u_d + u)$. We recall that if $f : V \to \mathbb{R} \cup \{+\infty\}$, the Legendre-Fenchel conjugate f^* is defined on V' as

$$\forall u^* \in V' \qquad f^*(u^*) = \sup_{u \in V} \langle u^*, u \rangle - f(u) \ .$$

We obviously have

$$\forall \lambda > 0,\ \forall u^* \in V' \qquad (\lambda f)^*(u^*) = \lambda f^*(\frac{u^*}{\lambda}) \ ,$$

and the following useful result:

Proposition 3.1. *[4] Let V be a normed space and $f : V \to \mathbb{R} \cup \{+\infty\}$ a closed, convex, proper function. then*

$$u^* \in \partial f(u) \iff u \in \partial f^*(u^*) \iff f(u) + f^*(u^*) = \langle u^*, u \rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the duality $V - V'$product.

Lemma 3.5. *Let be $f : L^2(\Omega) \to \mathbb{R} \cup \{+\infty\}$ and $\tilde{f}$ such that $\tilde{f}(u) = f(u_d + u)$. Then $\tilde{f}$ conjugate function writes*

$$\forall u^* \in L^2(\Omega) \qquad (\tilde{f})^*(u^*) = f^*(u^*) - (u^*, u_d)_2 \ ,$$

where $(\cdot, \cdot)_2$ denotes the $L^2(\Omega)$ inner product.

Proof. Let be $u^* \in L^2(\Omega)$. We have

$$\begin{aligned}(\tilde{f})^*(u^*) &= \sup_{u \in L^2(\Omega)} (u, u^*)_2 - f(u_d + u) = \sup_{w \in L^2(\Omega)} (w - u_d, u^*)_2 - f(w) \\ &= \sup_{w \in L^2(\Omega)} (w, u^*)_2 - f(w) - (u_d, u^*)_2 = f^*(u^*) - (u^*, u_d)_2.\end{aligned}$$

■

In the sequel $\mathbf{1}_C$ denotes the indicator function of the set C :

$$\mathbf{1}_C(u) = \begin{cases} 0 & \text{if } u \in C \\ +\infty & \text{else.} \end{cases}$$

Lemma 3.6. *The conjugate function of Φ_λ^1 is $(\Phi_\lambda^1)^* = \lambda \mathbf{1}_{\lambda \mathcal{K}_1}$, where $\mathcal{K}_1 = \overline{\mathbf{K}_1}$ is the L^2-closure of*

$$\mathbf{K}_1 := \left\{ \xi = div\ \varphi \mid \varphi \in \mathcal{C}_c^1(\Omega),\ \|\varphi\|_\infty \leq 1 \right\}. \tag{3.7}$$

The conjugate function of Φ_μ^2 *is* $(\Phi_\mu^2)^* = \mu \mathbf{1}_{\mu\mathcal{K}_2}$, *where* $\mathcal{K}_2 \supset \overline{\mathbf{K}_2}$ *and* $\overline{\mathbf{K}_2}$ *is the* L^2*-closure of*

$$\mathbf{K}_2 := \left\{ \xi = div^2 \psi \ | \psi \in \mathcal{C}_c^2(\Omega, \mathbb{R}^{d\times d}),\ \|\psi\|_\infty \leq 1 \right\}. \tag{3.8}$$

Proof. It is known that the conjugate TV^* of TV is the indicator function of $\overline{\mathcal{K}_1}$ (see [6, 16] for example). As $\Phi_\lambda^1 = \lambda TV$ (or $+\infty$ outside $BV(\Omega)$) then $(\Phi_\lambda^1)^*(u^*) = \lambda TV^*\left(\dfrac{u^*}{\lambda}\right)$. This gives the result.
The result is not exactly the same since Φ_μ^2 is equal to μTV^2 on $BH_0(\Omega)$ and $+\infty$ outside (and in particular on $BH(\Omega)\backslash BH_0(\Omega)$). Therefore the conjugate of Φ_μ^2 is not the same as the conjugate of μTV^2. We know that the conjugate function of TV^2 is $\mathbf{1}_{\overline{\mathbf{K}_2}}$ (see [10]); as Φ_1^2 is positively homogeneous ($\mu = 1$), it is the indicator function of a closed subset $\mathcal{K}_2$ of $L^2(\Omega)$. Moreover

$$\mathbf{1}_{\mathcal{K}_2}(v^*) = (\Phi_1^2)^*(v^*) = \sup_{v\in BH_0} \langle v^*, v\rangle - TV^2(v) \leq \sup_{v\in BH} \langle v^*, v\rangle - TV^2(v) = \mathbf{1}_{\overline{\mathbf{K}_2}}(v^*)\ .$$

This implies that $\overline{\mathbf{K}_2} \subset \mathcal{K}_2$ but we cannot prove the converse inclusion (for example). We end the proof with the same argument as in the BV case. ■

Eventually it is easy to see that $\mathcal{N}^* = \mathcal{N}$.

3.5. Dual problem to $(\mathcal{P}_{\lambda,\mu})$. In the present subsection we use convex duality tools that we recall thereafter (see [4] for example).

Theorem 3.7. *[[4] p 366] Let* V *be a banach space,* $f, g : V \to \mathbb{R}\cup\{+\infty\}$ *lower semi-continuous convex functions and* A *a linear continuous operator from* V *to* V*. Assume there exists* $u_o \in$ *dom* g *and* f *continuous at* Au_o*. Then*

$$\inf_{u\in V} \left(f(Au) + g(u)\right) = \max_{u^*\in V'} \left(-f^*(u^*) - g^*(-A^*u^*)\right).$$

Moreover, if $\bar{u}$ *is a solution to the primal problem and* $\bar{u}^*$ *is a solution to the dual one then*

$$\bar{u}^* \in \partial f(A\bar{u}) \text{ and } -A^*\bar{u}^* \in \partial g(\bar{u})\ ,$$

where $\partial f(u)$ *stands for the subdifferential of* f *at* u*.*

Theorem 3.8. *[[4] p 328] Let* V *be a banach space and* $f, g : V \to \mathbb{R}\cup\{+\infty\}$ *proper functions. Then*

$$(f\#g)^* = f^* + g^*\ .$$

In addition if f *and* g *satisfy the assumptions of Theorem 3.7, then*

$$(f+g)^* = f^*\#g^*\ .$$

Now we may compute the dual problem to $(\mathcal{P}_{\lambda,\mu})$ and get the following

Theorem 3.9. *The dual problem to* $(\mathcal{P}_{\lambda,\mu})$ *writes*

$$\inf_{w\in\lambda\mathcal{K}_1\cap\mu\mathcal{K}_2} \frac{1}{2}\|u_d - w\|_2^2. \tag{3.9}$$

The unique solution w^* is the L^2-projection of u_d on the closed convex set $\lambda\mathcal{K}_1 \cap \mu\mathcal{K}_2$:

$$w^* = \Pi_{\lambda\mathcal{K}_1 \cap \mu\mathcal{K}_2}(u_d) \ .$$

Proof. Solving problem $(\mathcal{P}_{\lambda,\mu})$ is equivalent to solving

$$\inf_{u\in L^2(\Omega)} (\mathcal{N}\#\Phi^2_\mu)(u_d - u) + \Phi^1_\lambda(u) = \inf_{u\in L^2(\Omega)} \widetilde{(\mathcal{N}\#\Phi^2_\mu)}(Au) + \Phi^1_\lambda(u)$$

with $Au = -u$. It clear that $A^* = A$. Moreover, Φ^1_λ is lsc with respect to the L^1- topology and thus for L^2- topology since Ω is bounded. As $\mathcal{N}\#\Phi^2_\mu$, $\widetilde{\mathcal{N}\#\Phi^2_\mu}$, A and Φ^1_λ fulfill assumptions of Theorem 3.7, the dual problem of $(\mathcal{P}_{\lambda,\mu})$ writes

$$(\mathcal{P}^*) \qquad \max_{w\in L^2(\Omega)} -\widetilde{(\mathcal{N}\#\Phi^2_\mu)}^*(w) - (\Phi^1_\lambda)^*(w) \ ,$$

where $\widetilde{(\mathcal{N}\#\Phi^2_\mu)}(w) = (\mathcal{N}\#\Phi^2_\mu)(u_d + w)$. Using Lemma 3.5 and Theorem 3.8 it easy to see that

$$\begin{aligned}
\widetilde{(\mathcal{N}\#\Phi^2_\mu)}^*(w) &= -(u_d, w)_2 + (\mathcal{N}\#\Phi^2_\mu)^*(w) \\
&= -(u_d, w)_2 + \mathcal{N}^*(w) + (\Phi^2_\mu)^*(w) \\
&= -(u_d, w)_2 + \mathcal{N}(w) + (\Phi^2_\mu)^*(w) \ .
\end{aligned}$$

Therefore, $(\mathcal{P}^*)$ writes

$$\max_{w\in L^2(\Omega)} (u_d, w)_2 - \mathcal{N}(w) - (\Phi^1_\lambda)^*(w) - (\Phi^2_\mu)^*(w) \ ,$$

that is

$$-\min_{w\in L^2(\Omega)} -(u_d, w)_2 + \mathcal{N}(w) + (\Phi^1_\lambda)^*(w) + (\Phi^2_\mu)^*(w) \ .$$

Finally, $(\mathcal{P}^*)$ is equivalent to

$$\min_{w\in\lambda\mathcal{K}_1\cap\mu\mathcal{K}_2} \frac{1}{2}\|u_d - w\|_2^2, \tag{3.10}$$

The dual problem has obviously a unique solution w^* which is the L^2 projection of u_d on the closed convex set $\lambda\mathcal{K}_1 \cap \mu\mathcal{K}_2$. ■

Next we have a relation between the solutions to $(\mathcal{P}_{\lambda,\mu})$ and the (unique) solution of the dual problem.

Theorem 3.10. *1. Let w^* be the (unique) solution to the dual problem $(\mathcal{P}_{\lambda,\mu})^*$:*

$$w^* = \Pi_{\lambda\mathcal{K}_1 \cap \mu\mathcal{K}_2}(u_d) \ .$$

Then there exists $(\bar{u}, \bar{v}) \in BV(\Omega) \times BH_0(\Omega)$ an optimal solution to $(\mathcal{P}_{\lambda,\mu})$ such that

$$w^* = u_d - \bar{u} - \bar{v} \text{ and } w^* \in \partial\Phi^2_\mu(\bar{v}) \cap \partial\Phi^1_\lambda(\bar{u}) \ .$$

2. Conversely, if $(\bar u,\bar v)\in BV(\Omega)\times BH_0(\Omega)$ *is any solution to* $(\mathcal{P}_{\lambda,\mu})$ *then*

$$\bar w = u_d-\bar u-\bar v = \Pi_{\lambda\mathcal{K}_1\cap\mu\mathcal{K}_2}(u_d)\ . \tag{3.11}$$

Proof. Let $(\bar u,\tilde v)$ be a solution to $(\mathcal{P}_{\lambda,\mu})$. A direct consequence of Theorem 3.7 is

$$w^*\in\partial\Phi^1_\lambda(\bar u)\ \text{ and }\ w^*\in\partial(\widetilde{\mathcal{N}\#\Phi^2_\mu})(-\bar u).$$

A simple calculus shows that

$$\partial(\widetilde{\mathcal{N}\#\Phi^2_\mu})(-\bar u)=\partial(\mathcal{N}\#\Phi^2_\mu)(u_d-\bar u)$$

so that

$$w^*\in\partial\Phi^1_\lambda(\bar u)\cap\partial(\mathcal{N}\#\Phi^2_\mu)(u_d-\bar u)\ .$$

As

$$(\mathcal{N}\#\Phi^2_\mu)(u_d-\bar u)=\mathcal{N}(u_d-\bar u-\tilde v)+\Phi^2_\mu(\tilde v)=\ \operatorname{argmin}_{v\in L^2(\Omega)}\frac{1}{2}\|v+\bar u-u_d\|^2+\Phi^2_\mu(v)\ ,$$

then

$$u_d-\tilde v-\bar u\in\partial\Phi^2_\mu(\tilde v);$$

so, the inf-convolution is exact and we get ([20])

$$\partial(\mathcal{N}\#\Phi^2_\mu)(u_d-\bar u)=\bigcup_{v\in L^2(\Omega)}\partial\mathcal{N}(u_d-\bar u-v)\cap\partial\Phi^2_\mu(v)\ .$$

As $\partial\mathcal{N}(u_d-\bar u-v)=\{u_d-\bar u-v\}$ this means that there exists $\bar v\in L^2(\Omega)$ such that

$$w^*=u_d-\bar u-\bar v\in\partial\Phi^2_\mu(\bar v)\ .$$

So

$$w^*\in\partial\Phi^2_\mu(\bar v)\cap\partial\Phi^1_\lambda(\bar u)\ ,$$

with $\bar v=u_d-\bar u-w^*$. This prove that $(\bar u,\bar v)$ is a solution to $(\mathcal{P}_{\lambda,\mu})$ as well: we use Theorem 3.2 with $\bar w=w^*$ to conclude.
Let us prove the converse property. Let $(\bar u,\bar v)\in BV(\Omega)\times BH_0(\Omega)$ be a solution to $(\mathcal{P}_{\lambda,\mu})$ and $\bar w=u_d-\bar u-\bar v$. Theorem 3.2 yields

$$\bar w\in\partial\Phi^1_\lambda(\bar u)\cap\partial\Phi^2_\mu(\bar v)\ ,$$

that is

$$\bar u\in\partial(\Phi^1_\lambda)^*(\bar w)\text{ and }\bar v\in\partial(\Phi^2_\mu)^*(\bar w)\ .$$

With the previous computations this gives

$$\bar u\in\partial\lambda\mathbf{1}_{\lambda\mathcal{K}_1}(\bar w)\text{ and }\bar v\in\partial\lambda\mathbf{1}_{\mu\mathcal{K}_2}(\bar w)\ .$$

Therefore

$$\forall w \in \lambda\mathcal{K}_1 \cap \mu\mathcal{K}_2 \qquad \langle \bar{u}, w - \bar{w} \rangle \leq 0 \text{ and } \langle \bar{v}, w - \bar{w} \rangle \leq 0 \ .$$

Adding the above inequalities gives

$$\forall w \in \lambda\mathcal{K}_1 \cap \mu\mathcal{K}_2 \qquad \langle \bar{u} + \bar{v}, w - \bar{w} \rangle = \langle u_d - \bar{w}, w - \bar{w} \rangle \leq 0 \ .$$

This is equivalent to (3.11). ■

Corollary 3.1. *If $(\bar{u}, \bar{v})$ is a solution to $(\mathcal{P}_{\lambda,\mu})$, then $\bar{w} = \bar{u} + \bar{v}$ is unique. In particular, there is a unique solution to $(\mathcal{P}_{\lambda,\mu})$ such that $\bar{u} = 0$ almost everywhere.*

Remark 3.3. *We cannot permute the role of Φ^1_λ and Φ^2_μ in the previous proof because Φ^2_μ is not lower semi-continuous with respect to the L^2 topology. Indeed $L^2(\Omega)$ is not embedded in $W^{1,1}(\Omega)$.*

4. Solution properties ($d \leq 2$).

4.1. Structure of the solutions. Recall (see [5, 21]) that the *Meyer space* $G(\Omega)$ is defined as

$$G(\Omega) := \{ \ f \in L^2(\Omega) \mid \exists \varphi = (\varphi_1, \varphi_2) \in L^\infty(\Omega, \mathbb{R}^2) \ f = \text{div } \varphi \ \text{ and } \varphi \cdot n = 0 \text{ on } \partial\Omega \ \}$$

where n is the outer normal vector to $\partial\Omega$. The space G is endowed with a norm denoted by $\| \cdot \|_G$ and defined as

$$\|f\|_G = \inf\{ \| \sqrt{\varphi_1^2 + \varphi_2^2} \|_\infty \mid f = \text{div } \varphi \ , \varphi \cdot n = 0 \text{ on } \partial\Omega \ \} \ .$$

We shall need the

Lemma 4.1 ([5], Lemma 2.1). *For every $u \in BV(\Omega)$ and $g \in G(\Omega)$ then*

$$\left| \int_\Omega u(x)\, g(x)\, dx \right| \leq TV(u)\, \|g\|_G \ .$$

We may now precise the structure of a generic solution.

Theorem 4.2. *Let us denote by $(\bar{u}, \bar{v}) \in BV(\Omega) \times BH_0(\Omega)$ a solution to problem $(\mathcal{P}_{\lambda,\mu})$ (for any fixed λ and μ) and set $\bar{w} = u_d - \bar{u} - \bar{v}$.*

i. $\bar{w} = u_d - \bar{u} - \bar{v} \in G(\Omega)$

ii. If $d = 2$ and Ω satisfies assumption (2.1), $\bar{v}$ is continuous on $\bar{\Omega}$.

iii. If $d = 2$, Ω satisfies assumption (2.1) and $u_d \in BV(\Omega) \cap L^\infty(\Omega)$, then the jump set of $\bar{u}$ is included in the jump set of u_d.

Proof. (i) This is a direct consequence of Theorem 3.10. Indeed $\bar{w} \in \lambda\mathcal{K}_1$. Therefore, there exists a sequence $\varphi_n \in \mathcal{C}^1_c(\Omega, \mathbb{R}^2)$ with $\|\varphi_n\|_\infty \leq 1$ such that $w_n = \lambda \text{div}(\varphi_n)$ L^2-converges to $\bar{w}$. As $\|\varphi_n\|_\infty \leq 1$ on may extract a weak-star subsequence that converges to $\bar{\varphi}$ in $L^\infty(\Omega)$. Therefore $\bar{\varphi} \in L^\infty(\Omega)$ and $\bar{\varphi}.n = 0$ on $\partial\Omega$. So, we get :

$$\forall u \in \mathcal{D}(\Omega) \qquad (w_n, u)_{L^2} = \lambda \int_\Omega \text{div}\varphi_n u = -\lambda \int_\Omega \varphi_n \nabla u \to -\lambda \int_\Omega \bar{\varphi} \nabla u.$$

As $(w_n, u)_{L^2} \to (\bar{w}, u)_{L^2}$ this gives

$$(\bar{w}, u) = -\lambda \langle \bar{\varphi}, \nabla u \rangle = \lambda \langle \text{div } \bar{\varphi}, u \rangle ,$$

in the distributional sense. Therefore $\bar{w} = \text{div } (\lambda \bar{\varphi})$. Moreover, $\bar{\varphi} \cdot n = 0$ on $\partial\Omega$ since φ_n as compact support. This proves that $\bar{w} \in G(\Omega)$.

(ii) Assumption (2.1) yields that $\bar{v} \in BH(\Omega)$ is continuous (Theorem 2.2, (iii)).
(iii) With (ii), the jump discontinuity set of u_d is the same as the one of $u_d - \bar{v}$. Moreover $\bar{u}$ is a solution to

$$\min_{u \in BV(\Omega)} \frac{1}{2} \|u_d - \bar{v} - u\|^2 + \lambda TV(u).,$$

Therefore, following [[15],Theorem 3.3] we get the result. ■

Remark 4.1. *The point (i) means that $\bar{w}$ is an oscillating function: this is consistent with the fact that we expect $\bar{w}$ to be the noise and/or micro-textures.*
The continuity of $\bar{v}$ still hold if $d \geq 2$. Assumptions on Ω are slightly different (see [12, 17]).

Corollary 4.1. *Let us denote by $(\bar{u}, \bar{v}) \in BV(\Omega) \times BH_0(\Omega)$ a solution to problem $(\mathcal{P}_{\lambda,\mu})$ (for any fixed λ and μ) and set $\bar{w} = u_d - \bar{u} - \bar{v}$. Then*

$$\int_\Omega \bar{w}(x)\, dx = 0 \ .$$

Proof. This is a direct consequence of proposition 2.1 of [5].■

The previous theorem deals with the case where $u_d \in BV(\Omega)$. This is not the case if u_d is noisy for example. In the case where $u_d \notin BV(\Omega)$ we have the following results due to W. Ring [25].
We first consider the 1D case where $\Omega = (a, b)$. Following Proposition 4 of [25], if we assume that

$$(\mathcal{H}_1) \qquad \begin{array}{c} \forall \mathcal{U} \text{ open subset of } (a,b) \text{ with positive Lebesgue measure} \\ u_d \text{ does not coincide on } U \text{with some function } u \in BV(a,b). \end{array}$$

then $u_d - \bar{v}$ satisfies $\mathcal{H}_1$ and we get $D_a \bar{u} = 0$ where $D_a u$ is the absolutely continuous part of the measure Du. Let Γ be the support of the singular part of $D\bar{u}$. Therefore $\bar{u}$ is piecewise constant on $(a, b)\backslash\Gamma$.
We have also a similar result for the 2D-case . Assume that

$$(\mathcal{H}_2) \qquad \forall \mathcal{U} \text{ open subset of } \Omega, \ u_{d|\mathcal{U}} \text{ is not equal not a } BV(\Omega) \text{ function.}$$

then $u_d - \bar{v}$ satisfies $(\mathcal{H}_2)$ as well (since $\bar{v} \in W^{1,1}(\Omega)$). Following Proposition 6 of [25], there is no open subset ω of Ω on which both components $\dfrac{\partial \bar{u}}{\partial x_i}, i = 1, 2$ have constant, non-zero sign.

4.2. Uniqueness. The functional $\mathcal{F}_{\lambda,\mu}$ is convex but not strictly convex, because of the degenerating direction $u+v=0$. It is obvious that if (u^*,v^*) is a solution then (u^*+c,v^*-c), where c is constant, is a solution as well. Let us call

$$\mathbf{C}(\Omega) := \{(u,v) \in BV(\Omega)\times BH_0(\Omega) \,|\exists c\in\mathbb{R}\ u=c \text{ and } v=-c \text{ a.e on } \Omega\ \}. \tag{4.1}$$

The question of uniqueness reduces to uniqueness up to $\mathbf{C}(\Omega)$ functions. In other words, if (u_1,v_1) and (u_2,v_2) are two optimal solutions of $(\mathcal{P}_{\lambda,\mu})$ can we show that $u_2=u_1+c$ and $v_2=v_1-c$ where c is a constant function? It is still an open problem for the 2D case. We shall discuss this point more precisely in the numerical section. Nevertheless we may give partial results:

Proposition 4.1. *Assume (u_1,v_1) and (u_2,v_2) are two optimal solutions of $(\mathcal{P}_{\lambda,\mu})$. Then there exists $\varphi\in BV(\Omega)\cap BH_0(\Omega)$ such that $u_2=u_1-\varphi$ and $v_2=v_1+\varphi$.*

Proof. Set $u=u_2-u_1(\in BV(\Omega))$ and $v=v_2-v_1(\in BH_0(\Omega))$. As $u_d-u_1-v_1=u_d-u_2-v_2$ (this is the unique solution of the dual problem), then $u+v=0$. This yields that $u=-v\in BV(\Omega)\cap BH_0(\Omega)$ and we get the result. ■

Lemma 4.3. *The only solutions $(\bar u,\bar v)$ to $(\mathcal{P}_{\lambda,\mu})$ that satisfies $\bar u+\bar v=0$ are functions of $\mathbf{C}(\Omega)$.*

Proof. Assume that $\bar u+\bar v=0$ then $\bar u\in BH_0(\Omega)$ and $\Phi^2(\bar v)=\Phi^2(-\bar u)=\Phi^2(\bar u)$. As $\mathcal{F}_{\lambda,\mu}(\bar u,\bar v)\le\mathcal{F}_{\lambda,\mu}(u,v)$, for every $(u,v)\in BV(\Omega)\times BH_0(\Omega)$ this yields

$$\|u_d\|^2_{L^2(\Omega)}+2\lambda TV(\bar u)+2\mu TV^2(\bar u)\le\|u_d-u-v\|^2_{L^2(\Omega)}+2\lambda TV(u)+2\mu TV^2(v)\ .$$

Taking $u=v=0$ gives

$$\|u_d\|^2_{L^2(\Omega)}+2\lambda TV(\bar u)+2\mu TV^2(\bar u)\le\|u_d\|^2_{L^2(\Omega)}.$$

So we get $\lambda TV(\bar u)+\mu TV^2(\bar u)=0$. This implies that $TV(\bar u)=0$ and that $\bar u$ is a constant function. ■

Theorem 4.4. *Let be (λ,μ) nonnegative real numbers such that $\lambda\ge\|u_d\|_G$ and $\mu\ge C_2\lambda$ where C_2 is the constant of Lemma 2.3. Then the $\mathbf{C}(\Omega)$ functions are the only solutions to $(\mathcal{P}_{\lambda,\mu})$.*

Proof. Let us assume that $\lambda\ge\|u_d\|_G$ and $\mu\ge C_2\lambda$ where C_2 is the constant of Lemma 2.3. Lemma 4.1 gives

$$\forall(u,v)\in BV(\Omega)\times BH_0(\Omega)\qquad |(u_d,u+v)_2|\le\lambda TV(u+v)$$

since $u_d\in L^2(\Omega)$ and $BH_0(\Omega)\subset BV(\Omega)$. Then

$$|(u_d,u+v)_2|\le\lambda TV(u)+\lambda TV(v)\ .$$

Lemma 2.3 gives a constant C_2 (only depending on Ω) such that

$$\forall v\in BH_0(\Omega)\qquad TV(v)\le C_2TV^2(v)\ ,$$

so that $\forall(u,v)\in BV(\Omega)\times BH_0(\Omega)$

$$|(u_d,u+v)_2|\le\lambda TV(u)+C_2\lambda TV^2(v)\le\lambda TV(u)+\mu TV^2(v). \tag{4.2}$$

Finally, we get for every $(u,v) \in BV(\Omega) \times BH_0(\Omega)$

$$\frac{1}{2}\|u_d\|^2 = \frac{1}{2}\|u_d - u - v\|^2 - \frac{1}{2}\|u+v\|^2 + (u_d, u+v)_2 \leq \frac{1}{2}\|u_d - u - v\|^2 + \lambda TV(u) + \mu TV^2(v).$$

This means that $\mathcal{F}_{\lambda,\mu}(0,0) \leq \mathcal{F}_{\lambda,\mu}(u,v)$: so $(0,0)$ is a solution to $(\mathcal{P}_{\lambda,\mu})$. Let $(\bar{u},\bar{v}) \in BV(\Omega) \times BH_0(\Omega)$ be another solution to $\mathcal{P}_{\lambda,\mu}$. With proposition 4.1, we get $\bar{u} + \bar{v} = 0$ and lemma 4.3 gives $(\bar{u},\bar{v}) \in \mathbf{C}(\Omega)$. This ends the proof. ■

Remark 4.2. *The previous theorem tells that if* $\dfrac{\mu}{\lambda}$ *and* λ *are large enough then the set of solutions is* $\mathbf{C}(\Omega)$ *. In addition, if we impose (for example) that* $u \in G$ *(that is* u *has a null mean value), then the unique solution is* $(0,0)$ *since* $\mathbf{C}(\Omega) \cap (G \times BH_0(\Omega)) = \{(0,0)\}$.

Eventually, we have a uniqueness result for the 1D case:

Theorem 4.5. *Assume* $n = 1$, $\Omega =]a,b[$ *and that* u_d *satisfies assumption* $(\mathcal{H}_1)$. *Then, for every* $\lambda > 0, \mu > 0$ *problem* $(\mathcal{P}_{\lambda,\mu})$ *has a unique solution up to a* $\mathbf{C}(\Omega)$ *function.*
More precisely, if (u_1,v_1) *and* (u_2,v_2) *are two optimal solutions of* $(\mathcal{P}_{\lambda,\mu})$ *then* $\varphi := u_2 - u_1 = v_2 - v_1$ *is a constant function. In particular, problem* $(\mathcal{P}_{\lambda,\mu})$ *has a unique solution* (u^*,v^*) *such that* u^* *has a null mean value.*

Proof. Let (u_1,v_1) and (u_2,v_2) be two optimal solutions of $(\mathcal{P}_{\lambda,\mu})$. Then, with proposition 4.1, there exists $\varphi \in BV(\Omega) \cap BH_0(\Omega)$ such that $\varphi = u_2 - u_1 = v_2 - v_1$. If u_d satisfies $(\mathcal{H}_1)$ then $u_d - v_i, i = 1,2$ obviousy satisfies this assumption as well. As $u_i, i = 1,2$ is solution to the ROF problem

$$u_i = \operatorname{argmin}\left\{ \frac{1}{2}\|u_d - v_i - u\|^2 + \Phi^1_\lambda(u),\ u \in L^2(\Omega) \right\},\ i = 1,2.$$

then, u_1, u_2 and φ are piecewise constant on Ω. In addition $\varphi = v_2 - v_1 \in BH(\Omega) \subset W^{1,1}(\Omega)$. This implies that φ is continuous and proves that φ is constant. ■

5. Numerical aspects.

5.1. Discretized problem and algorithm. This section is devoted to numerical computation (see [9]). We assume that the image is rectangular with size $N \times M$. We note $X := \mathbb{R}^{N\times M} \simeq \mathbb{R}^{NM}$ endowed with the usual (normalized) inner product and the associated Euclidean norm

$$\langle u,v\rangle_X := \frac{1}{NM}\sum_{1\leq i\leq N}\sum_{1\leq j\leq M} u_{i,j}v_{i,j}, \quad \|u\|_X := \sqrt{\frac{1}{NM}\sum_{1\leq i\leq N}\sum_{1\leq j\leq M} u_{i,j}^2}\,. \tag{5.1}$$

We set $Y = X \times X$. It is classical to define the discrete total variation with finite difference schemes as following (see for example [6]): the discrete gradient of the numerical image $u \in X$ is $\nabla u \in Y$ and may be computed by the following forward scheme for instance:

$$(\nabla u)_{i,j} = \left((\nabla u)^1_{i,j}, (\nabla u)^2_{i,j} \right), \tag{5.2}$$

where

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

Note that the constraint $\dfrac{\partial u}{\partial n} = 0$ is involved in the discretization process of the gradient. Therefore, in a discrete setting, the sets $\mathbf{K_2}$ and $\mathcal{K}_2$ coincide. The (discrete) total variation corresponding to $\Phi_1(u)$ is given by

$$J_1(u) = \frac{1}{NM} \sum_{1\le i\le N} \sum_{1\le j\le M} \left\| (\nabla u)_{i,j} \right\|_{\mathbb{R}^2}, \tag{5.3}$$

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 adjoint operator of the gradient operator ∇:

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

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:

$$(Hv)_{i,j} = \left((Hv)^{11}_{i,j}, (Hv)^{12}_{i,j}, (Hv)^{21}_{i,j}, (Hv)^{22}_{i,j} \right).$$

We refer to [10, 9] for the detailed expressions of these quantities. The discrete second order total variation corresponding to $\Phi_2(v)$ writes

$$J_2(v) = \frac{1}{NM} \sum_{1\le i\le N} \sum_{1\le j\le M} \|(Hv)_{i,j}\|_{\mathbb{R}^4}, \tag{5.4}$$

with

$$\|(Hv)_{i,j}\|_{\mathbb{R}^4} = \sqrt{(Hv^{11}_{i,j})^2 + (Hv^{12}_{i,j})^2 + (Hv^{21}_{i,j})^2 + (Hv^{22}_{i,j})^2} .$$

The discretized problem stands

$$\inf_{(u,v)\in X\times X} F_{\lambda,\mu} := \frac{1}{2}\|u_d - u - v\|^2_X + \lambda J_1(u) + \mu J_2(v). \tag{5.5}$$

Problem (5.5) has obviously a solution $\tilde{u}$ and $\tilde{v}$ that satisfies the following necessary and sufficient optimality conditions

$$\tilde{u} = u_d - \tilde{v} - \Pi_{\lambda K_1} (u_d - \tilde{v}) \ , \tag{5.6a}$$

$$\tilde{v} = u_d - \tilde{u} - \Pi_{\mu K_2} (u_d - \tilde{u}) , \tag{5.6b}$$

where K_1 and K_2 are the following convex closed subsets :

$$K_1 = \{\text{div } p \mid p \in X^2, \ \|p_{i,j}\|_{\mathbb{R}^2} \le 1 \ \forall i = 1,\dots,N, j = 1,\dots,M\}, \tag{5.7a}$$

$$K_2 = \{H^* p \mid p \in X^4, \ \|p_{i,j}\|_{\mathbb{R}^4} \le 1, \ \forall i = 1,\dots,N, j = 1,\dots,M\}, \tag{5.7b}$$

and Π_{K_i} denotes the orthogonal projection on K_i. These projections are computed with a Nesterov-type scheme as in [27]. We refer to [9] for more details. This leads to the following fixed-point algorithm :

Algorithme 1

Initialization step. Choose $u_0,\ v_0$, set $0 < \alpha < 1/2$ and $n = 1$.

Iteration. Define the sequences $((u_n, v_n))_n$ as

$$\begin{cases} u_{n+1} = u_n + \alpha\left(u_d - u_n - v_n - \Pi_{\lambda K_1}\left(u_d - v_n\right)\right) \\ v_{n+1} = v_n + \alpha\left(u_d - u_n - v_n - \Pi_{\mu K_2}\left(u_d - u_n\right)\right). \end{cases}$$

Stopping test. If $\max(\|u_{n+1} - u_n\|_{L^2}, \|v_{n+1} - v_n\|_{L^2}) \leq \varepsilon$ where $\varepsilon > 0$ is a prescribed tolerance, or if the iterations number is larger than a prescribed maximum number itmax, then STOP.

For any $\alpha \in (0, 1/2)$, the sequence generated by the algorithm converges to a stationary point, solution of (5.6) that we generically denote (u^*, v^*) in the sequel. The tolerance was set to $\varepsilon = 10^{-2}$ so that the stopping criterion is *de facto* the maximum number of iterations itmax. In the sequel, we have set itmax = 10 000 for the 1D case and itmax = 400 for the 2D case. We do not report on CPU time since all tests have been done with MATLAB© and the code is not optimized. A parallelized C^{++} is version is written that reduces the computational time significantly.

5.2. Examples. We use 1D and 2D examples.

For the first (1D) example we set $s = s_0 + s_1 + s_2$ on [0,1] with s_0 a white gaussian noise with standard deviation $\sigma = 0.02$ and

$$s_1 = \begin{cases} 0.4 & \text{on } [\frac{3}{10}, \frac{6}{10}] \\ 0 & \text{elsewhere} \end{cases}, \qquad s_2(x) = \begin{cases} 0.8\,x + 0.2 & \text{on } [0, \frac{1}{2}] \\ -1.2\ (x-1) & \text{elsewhere.} \end{cases}$$

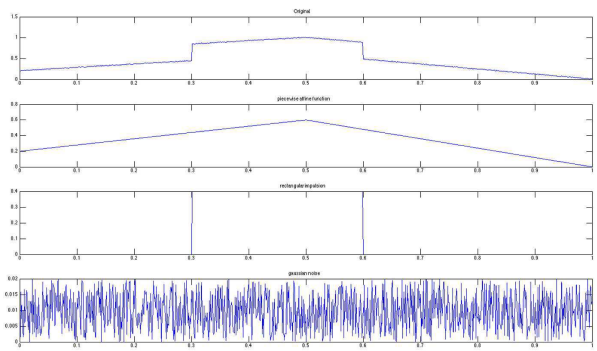


Figure 5.1. *1D example - 1000 points*

The second example is a 2D picture of a *butterfly* and the third one an highly textured image (old *wall*). We used geometrical images as well but we do not report on them.

(a) Test 2D - Butterfly (b) Test 2D - Wall

Figure 5.2. *2D examples*

We present some results and comments in the next subsections[1].

5.3. Initialization process. We have tested many initialization choices for algorithm. Indeed, we have not proved uniqueness (though we conjecture it). So the computed solution is only a stationary point. As we may have many, we may think that the initialization process has a significant influence on the generated sequence.
More precisely, we used

- $u_0 = 0,\ v_0 = u_d$, that we call initialization (a) in the sequel,
- $u_0 = u_d,\ v_0 = 0$ that we call initialization (a') in the sequel,
- $u_0 = 0,\ v_0 = 0$: initialization (b),
- randomized initializations around u_d mean value.

Initialization (a) (resp. (a')) provides a stationary pair (u^*, v^*) such that u^* (resp. v^*) has null mean value.

Proposition 5.1. *Assume $u_0 = 0$ and $v_0 = u_d$. Then any solution (u^*, v^*) given by the algorithm satisfies $\int_\Omega u^* = 0$. Similarly, if $u_0 = u_d$ and $v_0 = 0$, the pair (u^*, v^*) given by the algorithm satisfies $\int_\Omega v^* = 0$.*

Proof. Though we consider a discrete setting we use a continuous setting notation (using for example a piecewise affine approximation). We first note that

$$w \in K_1 \cup K_2 \Longrightarrow \int_\Omega w = 0 \ .$$

We prove the first assertion. Assume that $u_0 = 0$ and $v_0 = u_d$. It is easy to see by induction that

$$\forall n \in \mathbb{N} \qquad \int_\Omega u_n = 0 \text{ and } \int_\Omega (v_n - u_d) = 0 \ . \tag{5.8}$$

[1]Complete results (text files, movies, other examples) and MATLAB© code, are available at http://maitinebergounioux.net/PagePro/Movies.html

using

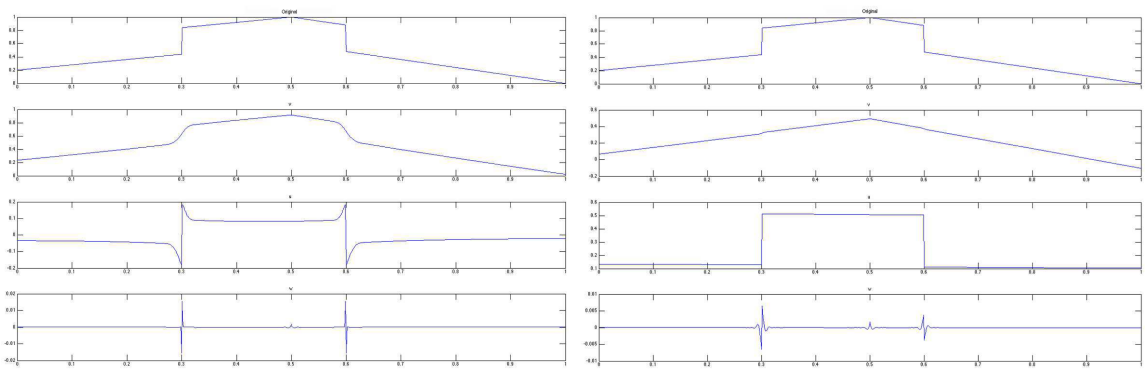
$$\begin{cases} u_{n+1} = u_n + \alpha \left(u_d - u_n - v_n - \Pi_{\lambda K_1} \left(u_d - v_n \right) \right) \\ v_{n+1} = v_n + \alpha \left(u_d - u_n - v_n - \Pi_{\mu K_2} \left(u_d - u_n \right) \right) . \end{cases}$$

Passing to the limit we get

$$\int_\Omega u^* = 0 \text{ and } \int_\Omega (v^* - u_d) = 0 \ .$$

The second assertion is proved similarly. ■

Proposition 5.1 yields that the BV- part u^* (or the BH- part v^*) belongs to the discrete Meyer space G (see [7]) if we perform the appropriate initialization step. This means it is an oscillating function. More precisely, choosing $u_0 = u_d,\ v_0 = 0$ gives a BH- part that belongs to G. This is not what we want, since the BH- part should not be oscillating. Therefore, we shall never use such an initialization.
Initializations (a) and (a') seem to give different results from initialization (b). We shall see in the sequel that the difference is *small* if the iteration number is large enough. Therefore, we think that the initial guess has no influence on the result, but only on the convergence speed.
We can see on Figure 5.3 (1D example) the oscillating effect of initialization $u_0 = 0,\ v_0 = u_d$:



(a) Initialization $u_0 = 0,\ v_0 = u_d$ (b) Initialization $u_0 = 0,\ v_0 = 0$

Figure 5.3. *Example 1D without noise, $\lambda = 10^{-2}$, $\mu = 5\ 10^{-2}$ and different initializations. Both u^* and w^* have null mean value for init (a). We recover the original decomposition with init (b).*

Figure 5.4 and Table 1 gives the computed pairs with initializations (a) (a') (b) and a randomized initialization around the mean value of u_d.

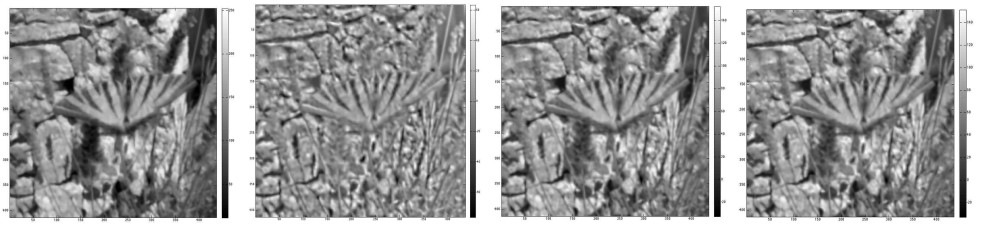
Initialization	$F_{\lambda,\mu}(u^*, v^*)$	$\|w^*\|_{L^2}$	$TV(u^*)$	$TV^2(v^*)$	Error	# it.
$\lambda = 1,\ \mu = 10$						
$u_0 = 0,\ v_0 = u_d$	23.68	1.04	12.70	1.04	2.35	400
$u_0 = u_d,\ v_0 = 0$	18.29	1	13.43	0.43	0.73	400
$u_0 = 0,\ v_0 = 0$	20.36	1.03	12.87	0.69	0.89	400
Random	20.39	1.03	12.88	0.69	0.87	400
$\lambda = 2,\ \mu = 0.1$						
$u_0 = 0,\ v_0 = u_d$	1.5414	2.24 e-01	3.64 e-04	15.15	8.48 e-03	22
$u_0 = u_d,\ v_0 = 0$	8.0239	2.76 e-01	3.31	13.52	4.35	400
$u_0 = 0,\ v_0 = 0$	3.3335	2.45 e-01	0.92	14.65	3.12	400
Random	3.5384	3.18 e-01	1.02	14.62	3.30	400
$\lambda = 5,\ \mu = 7$						
$u_0 = 0,\ v_0 = u_d$	61.7005	4.22	5.71	3.45	1.67	400
$u_0 = u_d,\ v_0 = 0$	62.7803	4.02	7.29	2.60	3.25	400
$u_0 = 0,\ v_0 = 0$	61.6248	4.15	6.34	3.04	1.50	400
Random	61.6331	4.15	6.35	3.04	1.55	400
$\lambda = 7,\ \mu = 7$						
$u_0 = 0,\ v_0 = u_d$	69.6775	5.23	2.29	5.69	9.76 e-01	400
$u_0 = u_d,\ v_0 = 0$	72.6262	4.96	4.09	4.51	4.74	400
$u_0 = 0,\ v_0 = 0$	70.2957	5.13	2.97	5.19	2.45	400
Random	70.3114	5.12	2.98	5.18	2.52	400
$\lambda = 7,\ \mu = 9$						
$u_0 = 0,\ v_0 = u_d$	79.7064	5.42	4.10	4.03	1.33	400
$u_0 = u_d,\ v_0 = 0$	80.1229	5.18	5.40	3.33	4.09	400
$u_0 = 0,\ v_0 = 0$	79.8224	5.33	4.58	3.72	1.86	400
Random	79.8297	5.33	4.58	3.72	1.89	400
$\lambda = 10,\ \mu = 15$						
$u_0 = 0,\ v_0 = u_d$	116.2130	7.04	3.59	3.69	1.39	400
$u_0 = u_d,\ v_0 = 0$	116.9598	6.79	4.33	3.36	4.67	400
$u_0 = 0,\ v_0 = 0$	116.0822	6.95	3.83	3.56	2.02	400
Random	116.0918	6.95	3.84	3.56	2.10	400

Table 1

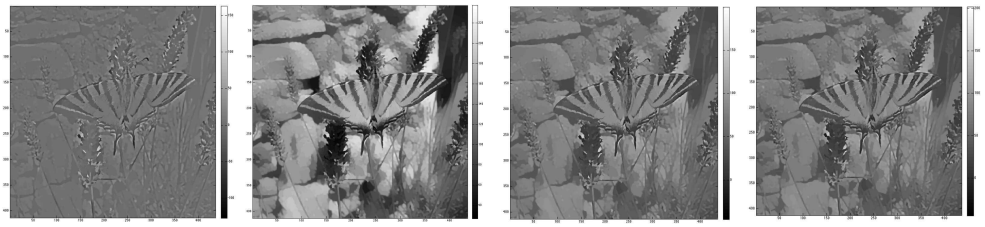
Comparison of different initializations (Butterfly)- itmax=400 -The stationary pair is denoted (u^*, v^*) *and* $w^* = u_d - u^* - v^*$.

The blue (grey) lines of Table 1 show the *optimal* solution, that is the computed pair whose cost functional value is the lowest. We observed that

- the randomized initialization gives the same result as initialization (b),
- the component $w^* = u_d - u^* - v^*$ is always the same, which is consistent with the theoretical result of uniqueness,
- the values of the cost functional may be quite close and the computed pairs quite different: see for example $\lambda = 5, \mu = 7$ (and figure 5.4),
- initialization (b) gives a pair (u_b, v_b) such that neither u_b nor v_b has null mean value.



(a) $u_0 = 0,\ v_0 = u_d$ (b) $u_0 = u_d,\ v_0 = 0$ (c) $u_0 = 0,\ v_0 = 0$ (d) Random

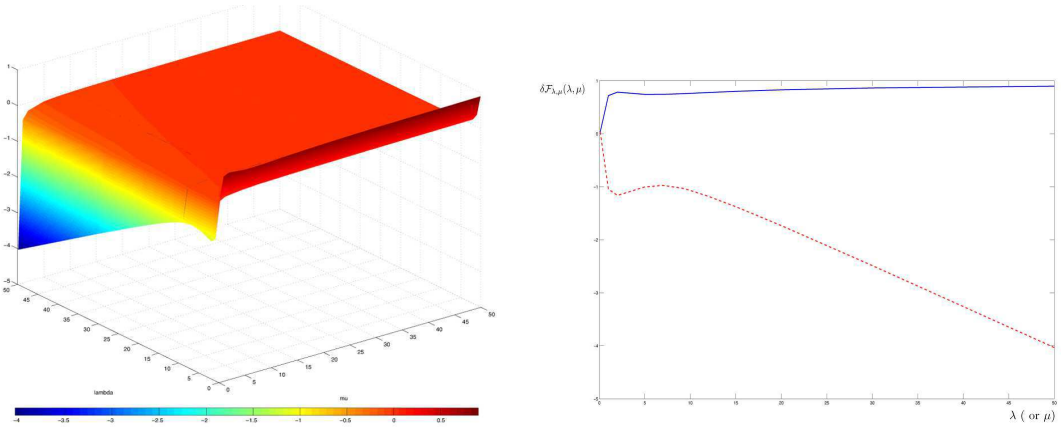


(e) $u_0 = 0,\ v_0 = u_d$ (f) $u_0 = u_d,\ v_0 = 0$ (g) $u_0 = 0,\ v_0 = 0$ (h) Random

Figure 5.4. *BH-part v (first line) and BV-part u (second line) given by initializations (a), (a'), (b) and random for* $\lambda = 5,\ \mu = 7$ *- Butterfly example with 400 iterations*

In the sequel, (u_a, v_a) denotes the pair given by the algorithm with initialization (a) and (u_b, v_b) the one given by the algorithm with initialization (b). Moreover, we set the signed relative error as

$$\delta\mathcal{F}_{\lambda,\mu} = \frac{F_{\lambda,\mu}(u_a, v_a) - F_{\lambda,\mu}(u_b, v_b)}{\min(F_{\lambda,\mu}(u_a, v_a), F_{\lambda,\mu}(u_b, v_b))}. \tag{5.9}$$



(a) $\delta\mathcal{F}_{\lambda,\mu}(\lambda, \mu)$ (b) $\delta\mathcal{F}_{\lambda,\mu}(0.1, \mu)$ (red dotted line) - $\delta\mathcal{F}_{\lambda,\mu}(\lambda, 0.1)$(blue solid line)

Figure 5.5. *Behavior of* $\delta\mathcal{F}_{\lambda,\mu}$ *for 400 iterations (Butterfly example). If* λ *and* μ *are large enough (*$\lambda > 0.1$ *and* $\mu > 0.1$ *for example), both optimal values are very close.*

Figure 5.5 shows the behavior of $\delta\mathcal{F}_{\lambda,\mu}$ with respect to λ and μ.

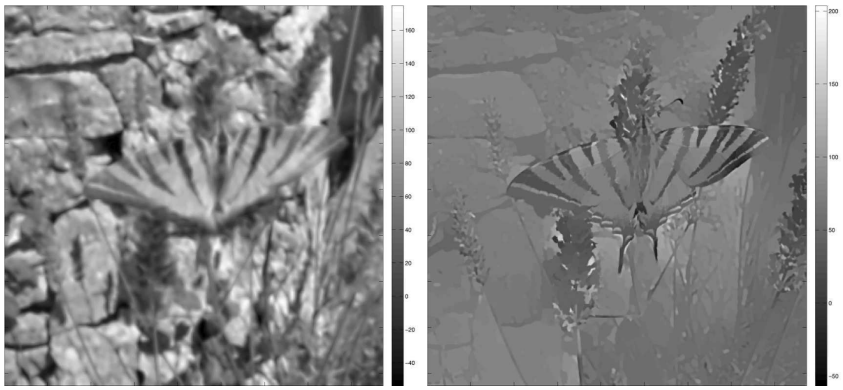
# it.	$F_{\lambda,\mu}(u_a, v_a)$	$F_{\lambda,\mu}(u_b, v_b)$	$\lvert\delta\mathcal{F}_{\lambda,\mu}\rvert$
50	82.38439	81.69328	8 e-03
100	80.9555	80.9579	3 e-05
200	80.18443	80.35509	2 e-03
400	79.83481	79.94497	1.3 e-03
600	79.73564	79.80224	8 e-04
800	79.68948	79.73411	5.6 e-04
1000	79.66213	79.69571	4.2 e-04
1200	79.64396	79.67121	3.4 e-04
1500	79.62567	79.64659	2.6 e-04
5000	79.5738	79.5718	2.5 e-05

# it.	$TV(u_a)$	$TV(u_b)$	$TV(\varphi)$	$TV^2(v_a)$	$TV^2(v_b)$	$TV^2(\varphi)$	Error (a)	Error (b)
50	2.45	5.62	3.57	5.40	3.20	4.05	7.41	7.14
100	3.29	5.31	2.53	4.69	3.33	2.82	4.88	5.03
200	3.85	4.93	1.66	4.23	3.52	1.87	2.87	3.26
400	4.12	4.60	1.03	4.01	3.70	1.18	1.34	1.87
600	4.19	4.47	0.771	3.95	3.77	0.896	1.02	1.32
800	4.22	4.40	0.628	3.93	3.81	0.736	0.855	1.03
1000	4.23	4.37	0.536	3.92	3.83	0.632	0.735	0.845
1200	4.24	4.35	0.470	3.91	3.84	0.556	0.642	0.723
1500	4.24	4.33	0.396	3.90	3.86	0.472	0.535	0.595
5000	4.28	4.26	0.148	3.88	3.89	0.180	0.207	0.208

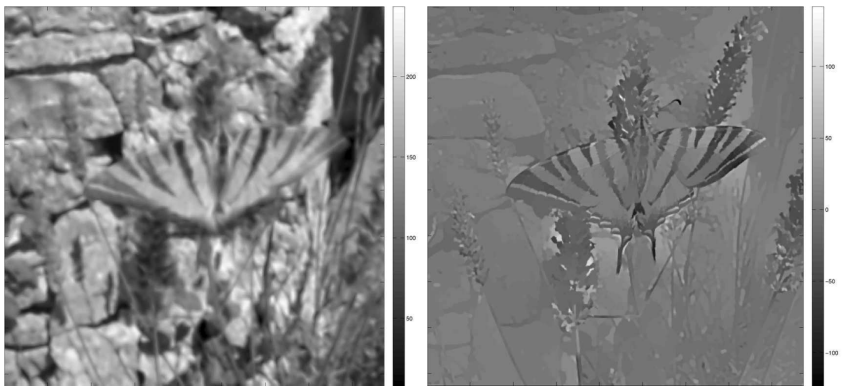
Table 2

Cost functional, TV and TV^2 for pairs given by initializations (a) and (b) and $\lambda = 7$, $\mu = 9$, as the number of iterations increases. Here $\varphi = u_b - u_a = v_b - v_a$ and the error is given by the stopping criterion of Algorithm.

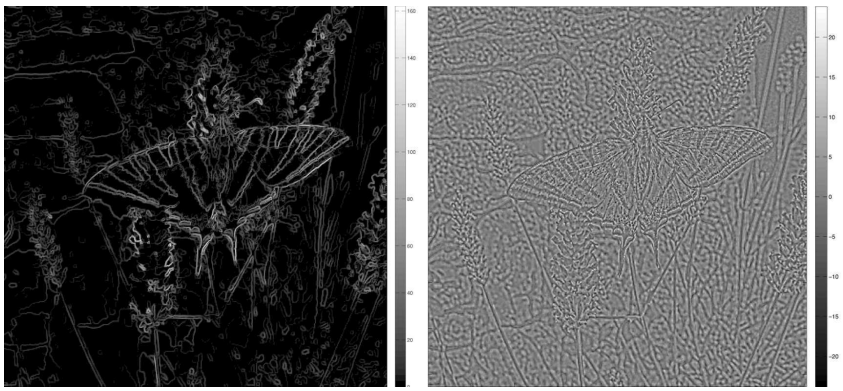
Though $F_{\lambda,\mu}(u_a, v_a) \simeq F_{\lambda,\mu}(u_b, v_b)$ the pairs (u_a, v_a) and (u_b, v_b) may be very different. More precisely, we have $u_b = u_a - \varphi$ and $v_b = v_a + \varphi$. Though the computed function φ_k at iteration k is not a constant function (see Figure 5.6), we infer that φ_k converges to a constant function as the iteration number increases. Indeed, we have numerically observed (see Table 2) that both $TV(\varphi)$ and $TV^2(\varphi)$ decreases to 0 as the iteration number increases. Nevertheless, we can perform only a limited number of iterations. So the computed solutions differ from a (small) piecewise constant function (see Figure 5.6). In addition, it is numerically confirmed that $w_a = u_d - u_a - v_a = w_b$ (what was theoretically proved).



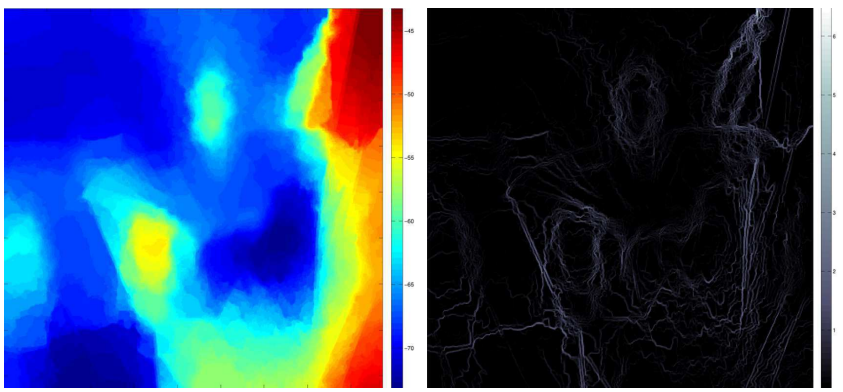
(a) v_a - init : $u_0 = 0,\ v_0 = u_d$ (b) u_a - init : $u_0 = 0,\ v_0 = u_d$



(c) v_b - init : $u_0 = 0,\ v_0 = 0$ (d) u_b - init : $u_0 = 0,\ v_0 = 0$



(e) Norm of u_b gradient (f) $w_b = w_a$



(g) $\varphi = u_b - u_a = v_a - v_b$ (h) Norm of φ gradient

Figure 5.6. *Difference between the solutions given by initializations (a) and (b) for* $\lambda = 7,\ \mu = 9$ *- 5000 iterations .* $\|\varphi\|_2 = 0.1518,\ TV(\varphi) = 0.1484,\ TV^2(\varphi) = 0.1803$. *The function* φ *seems to be piecewise constant as we see it on the gradient norm.*

5.4. Convergence. We chose $\alpha = 0.25$ in the fixed point algorithm and we always observed convergence. We set the maximal number of iterations quite large but we noticed that the solution is satisfactory with less iterations (400 for 2D case and 1000 for 1D case).

# it.	$F_{\lambda,\mu}(u_a, v_a)$	$F_{\lambda,\mu}(u_b, v_b)$	$\lvert\delta\mathcal{F}_{\lambda,\mu}\rvert$
$\lambda = 1,\ \mu = 10$			
50	39.132	25.513	5 e-01
100	31.727	23.113	3.7 e-01
200	26.907	21.440	2.5 e-01
400	23.711	20.377	1.6 e-01
600	22.410	19.978	1.2 e-01
800	21.688	19.774	9.6 e-02
$\lambda = 10,\ \mu = 15$			
50	119.448	117.102	2 e-02
100	117.578	116.601	8.3 e-03
200	116.612	116.257	3 e-03
400	116.215	116.083	1.1 e-03
600	116.106	116.031	6.5 e-04
800	116.052	116.006	4 e-04
$\lambda = 10,\ \mu = 2$			
50	25.90989	39.586	5.2 e-01
100	25.91003	33.501	2.9 e-01
200	25.91008	29.512	1.4 e-01
400	25.91009	27.558	6.3 e-02
600	25.91009	26.986	4.1 e-02
800	25.91009	26.699	3 e-02

# it.	$TV(u_a)$	$TV(u_b)$	$TV(\varphi)$	$TV^2(v_a)$	$TV^2(v_b)$	$TV^2(\varphi)$	Error (a)	Error (b)
$\lambda = 1,\ \mu = 10$								
50	10.73	12.12	3.98	2.74	1.27	2.19	14.58	6.51
100	11.81	12.51	3.13	1.92	1	1.50	8.30	3.33
200	12.39	12.74	2.47	1.39	0.81	1.09	4.39	1.72
400	12.70	12.87	1.96	1.04	0.69	0.82	2.37	0. 90
600	12.80	12.92	1.70	0.9	0.65	0.70	1.53	0.58
800	12.85	12.94	1.53	0.83	0.63	0.63	1.09	0.42
$\lambda = 10,\ \mu = 15$								
50	2.59	4.83	2.74	4.42	3.09	3.14	9.04	9.33
100	3.19	4.44	1.85	3.98	3.27	2.14	5.29	6.17
200	3.48	4.07	1.18	3.77	3.45	1.39	2.49	3.77
400	3.59	3.83	0.73	3.69	3.56	0.87	1.39	2.03
600	3.61	3.76	0.55	3.67	3.60	0.67	1.08	1.40
800	3.62	3.72	0.45	3.67	3.62	0.54	0.87	1.07
$\lambda = 10,\ \mu = 2$								
50	8.937 e-03	1.619	1.619	11.214	9.88	1.80	1.15 e-02	18.18
100	8.951 e-03	8.948 e-01	8.94 e-01	11.214	10.566	9.62 e-01	4.72 e-03	12.92
200	8.956 e-03	4.116 e-01	4.11 e-01	11.214	11.010	3.35 e-01	1.93 e-03	8.92
400	8.957 e-03	1.804 e-01	1.80 e-01	11.214	11.183	6.05 e-02	9.30 e-04	4.48
600	8.957 e-03	1.180 e-01	1.17 e-01	11.214	11.207	1.54 e-02	6.23 e-04	2.54
800	8.957 e-03	8.841 e-02	8.77 e-02	11.214	11.212	6.4 e-03	4.42 e-04	1.77

Table 3

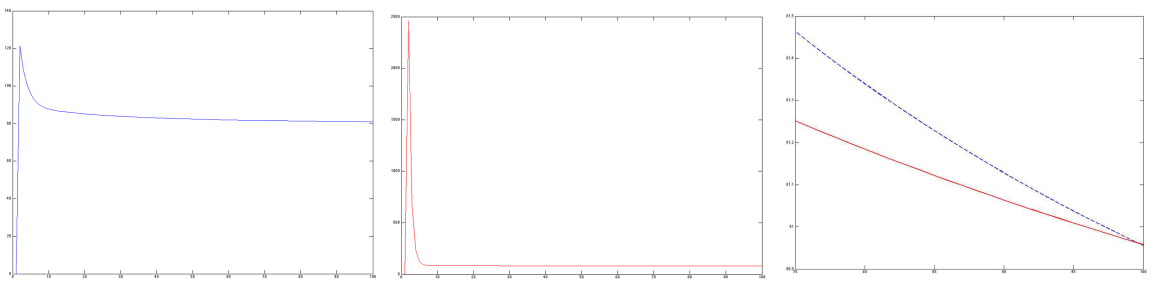
Sensitivity with respect to number of iterations. Here $\varphi = u_b - u_a = v_b - v_a$ and the error is given by the stopping criterion of Algorithm. On can refer to Table 2 as well.

(a) BV part - $\lambda = 1,\ \mu = 10$ (b) BV part - $\lambda = 10,\ \mu = 15$ (c) BV part -$\lambda = 10,\ \mu = 2$. In this case $u \simeq 0$

(d) φ - $\lambda = 1,\ \mu = 10$ (e) φ - $\lambda = 10,\ \mu = 15$ (f) φ -$\lambda = 10,\ \mu = 2$

Figure 5.7. *BV component u_a and φcorresponding to Table 3 - 800 iterations*

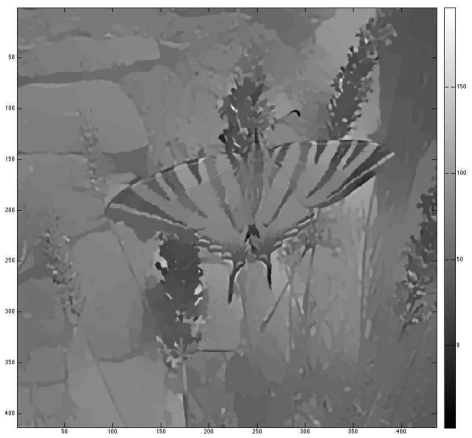
Figure 5.8 illustrates the generic behavior of the cost-functional $\mathcal{F}_{\lambda,\mu}$.[2]



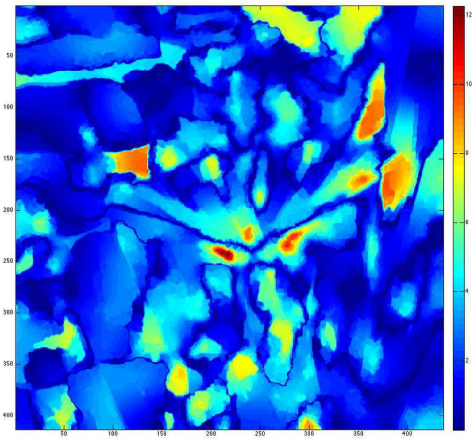
(a) Init (a) : $u_0 = 0,\ v_0 = u_d$ (b) Init (b) : $u_0 = 0,\ v_0 = 0$ (c) Zoom 75-100 iterations

Figure 5.8. *Behavior of the cost functional for $\lambda = 7,\ \mu = 9$, 100 iterations- Dotted (blue) line is initialization (a) and solid (red) line is initialization (b)*

[2]One can look at http://maitinebergounioux.net/PagePro/Movies.html to see the convergence process.

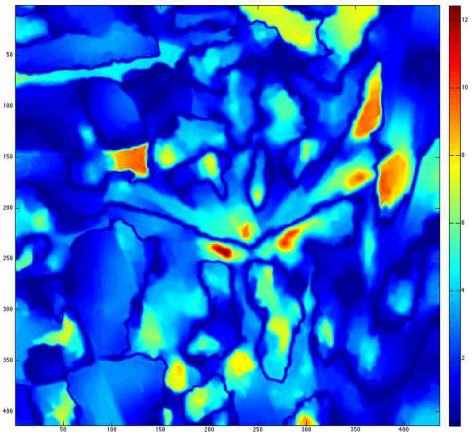


(a) BV part u - 800 iterations

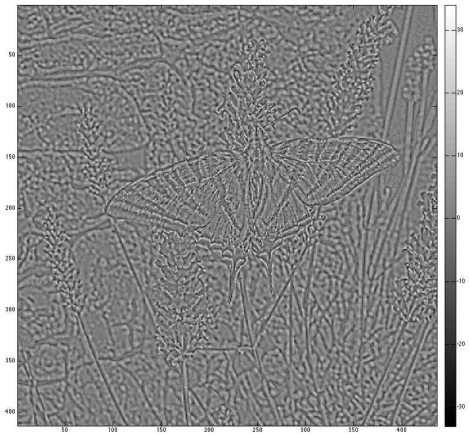


(b) Absolute difference of BV parts between it 400 and it 800

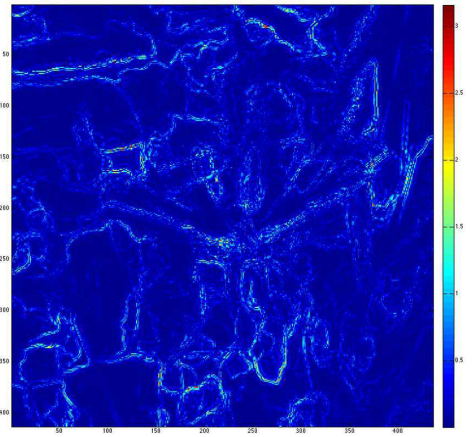
(c) BH part v - 800 iterations



(d) Absolute difference of BH parts between it 400 and it 800



(e) L^2- part w - 800 iterations



(f) Absolute difference of L^2- parts between it 400 and it 800

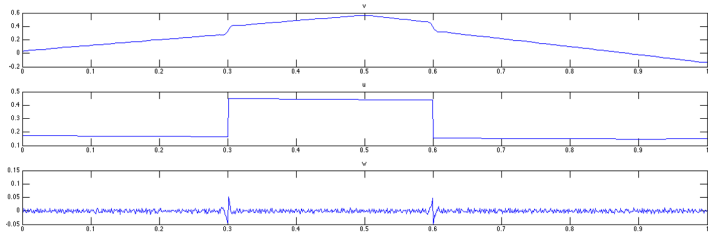
Figure 5.9. *Test 2D - Initialization (b) for* $\lambda = 10$, $\mu = 15$- *Difference between the computed pairs at iteration 400 and iteration 800.*

5.5. Sensitivity with respect to sampling and quantification. Table 4 and Figure 5.10 show that the model is robust with respect to sampling. Here, we have discretized the analogical signal of example 1D with 10^3, 10^4 and 10^5 points respectively.

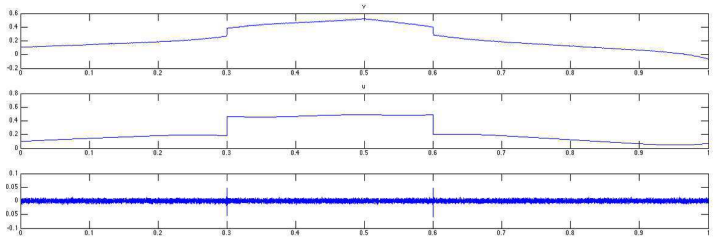
λ	μ	$\mathcal{F}_{\lambda,\mu}(u_b, v_b)$	$\|w_b\|_{L^2}$	$TV(u_b)$	$TV^2(v_b)$
		10^3 points 10^4 points 10^5 points	10^3 points 10^4 points 10^5 points	10^3 points 10^4 points 10^5 points	10^3 points 10^4 points 10^5 points
1e -03	1 e-02	6.47 e-06 5.78 e-06 5.73 e-06	1.43 e-03 1.47e-03 1.5 e-03	5.01 e-03 4.37 e-03 4.34 e-03	4.28 e-05 3.19 e-05 3.09 e-05
1e-03	1	4.86 e-05 3.73 e-05 3.63 e-05	1.43 e-03 1.47 e-03 1.47 e-03	5.01 e-03 4.37 e-03 4.34 e-03	4.25 e-05 3.18 e-05 3.09 e-05
1e -02	1e-01	3.27 e-05 2.62 e-05 2.52 e-05	5.07 e-03 5.16 e-03 5.24 e-03	8.80 e-04 1.81 e-04 7.26 e-05	1.10 e-04 1.10 e-04 1.08 e-04
1e -02	1	1.32 e-04 1.26 e-04 1.22 e-04	5.07 e-03 5.16 e-03 5.24 e-03	8.80 e-04 1.81 e-04 7.26 e-05	1.10 e-04 1.10 e-04 1.08 e-04
1e-01	1e-01	1.01 e-04 3.71 e-05 2.73 e-05	6.90 e-03 5.53 e-03 5.38 e-03	5.43 e-04 8.74 e-05 1.20 e-05	2.32 e-04 1.30 e-04 1.16 e-04

Table 4

Test 1D (with noise) - sensitivity with respect to sampling - Initialization (b) ($u_0 = v_0 = 0$) and 10 000 iterations



(a) 10^3 points



(b) 10^4 points

Figure 5.10. *Test 1D (with noise) - Pair given by initialization (b) for $\lambda = 0.1$, $\mu = 1$, 10 000 iterations and different samplings.*

We now investigate the sensitivity of the model with respect to quantification. Let u_d a data (with values in [0, 255] for example). Let (λ, μ) be chosen parameters and $(u_{\lambda,\mu}, v_{\lambda,\mu})$ the corresponding computed pair (with the appropriate initialization). Let $\alpha > 0$ and consider the new data αu_d. This is the case, for example, if we get 16 bits images and convert them to 8 bits : in this case $\alpha = (2^8 - 1)/(2^{16} - 1)$. We may want to normalize the data as well: in this case $\alpha = 1/\max(u_d)$. The question is to know what new parameters $(\tilde{\lambda}, \tilde{\mu})$ must be chosen to get $u_{\tilde{\lambda},\tilde{\mu}} = \alpha u_{\lambda,\mu}$ and $v_{\tilde{\lambda},\tilde{\mu}} = \alpha v_{\lambda,\mu}$. For any $(u_{\lambda,\mu}, v_{\lambda,\mu})$ solution to $(\mathcal{P}_{\lambda,\mu})$, we get

$$\begin{aligned}
F_{\lambda,\mu}(u_{\lambda,\mu}, v_{\lambda,\mu}) &= \frac{1}{2}\|u_d - u_{\lambda,\mu} - v_{\lambda,\mu}\|^2 + \lambda TV(u_{\lambda,\mu}) + \mu TV^2(v_{\lambda,\mu}) \\
&= \frac{1}{2\alpha^2}\|\alpha u_d - \alpha u_{\lambda,\mu} - \alpha v_{\lambda,\mu}\|^2 + \frac{\lambda}{\alpha} TV(\alpha u_{\lambda,\mu}) + \frac{\mu}{\alpha} TV^2(\alpha v_{\lambda,\mu}) \\
&= \frac{1}{\alpha^2}\left(\frac{1}{2}\|\alpha u_d - u_{\tilde{\lambda},\tilde{\mu}} - v_{\tilde{\lambda},\tilde{\mu}}\|^2 + \alpha\lambda TV(u_{\tilde{\lambda},\tilde{\mu}}) + \alpha\mu TV^2(v_{\tilde{\lambda},\tilde{\mu}})\right) \\
&= \frac{1}{\alpha^2} F_{\tilde{\lambda},\tilde{\mu}}(u_{\tilde{\lambda},\tilde{\mu}}, v_{\tilde{\lambda},\tilde{\mu}}) \text{ with}
\end{aligned}$$

$$\tilde{\lambda} = \alpha\lambda \text{ and } \tilde{\mu} = \alpha\mu.$$

α	1/255	100
$\lambda = 7,\ \mu = 9$ Initialization (a)		
$\|u_{\alpha\lambda,\alpha\mu} - \alpha u_{\lambda,\mu}\|_\infty/\alpha$	3.0291e-01	3.0291e-01
$\|v_{\alpha\lambda,\alpha\mu} - \alpha v_{\lambda,\mu}\|_\infty/\alpha$	3.1291e-01	3.1291e-01
$\lambda = 7,\ \mu = 9$ Initialization (b)		
$\|u_{\alpha\lambda,\alpha\mu} - \alpha u_{\lambda,\mu}\|_\infty/\alpha$	1.8006e-01	1.8006e-01
$\|v_{\alpha\lambda,\alpha\mu} - \alpha v_{\lambda,\mu}\|_\infty/\alpha$	1.7924e-01	1.7924e-01
$\lambda = 10,\ \mu = 2$ Initialization (a)		
$\|u_{\alpha\lambda,\alpha\mu} - \alpha u_{\lambda,\mu}\|_\infty/\alpha$	8.0280e-15	8.4421e-15
$\|v_{\alpha\lambda,\alpha\mu} - \alpha v_{\lambda,\mu}\|_\infty/\alpha$	1.1324e-13	1.4552e-13
$\lambda = 10,\ \mu = 2$ Initialization (b)		
$\|u_{\alpha\lambda,\alpha\mu} - \alpha u_{\lambda,\mu}\|_\infty/\alpha$	3.9216e-03	4.5475e-14
$\|v_{\alpha\lambda,\alpha\mu} - \alpha v_{\lambda,\mu}\|_\infty/\alpha$	1.1324e-13	1.0914e-13

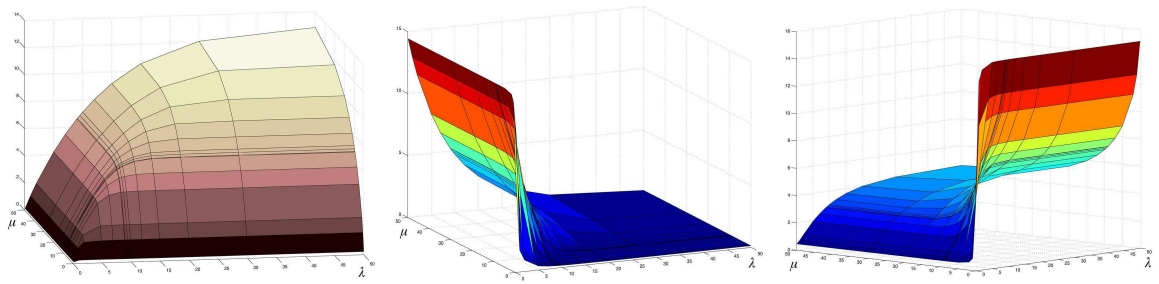
Table 5

Sensitivity with respect to quantification- Initialization (b) - itmax = 400

5.6. Sensitivity with respect to parameters. As mentioned before the computed stationary pair depends on the initialization guess via the convergence speed. We consider three cases and we illustrate them on test 2D (*Butterfly*).

- If $\mu << \lambda$, then initialization (a) : $u_0 = 0$ and $v_0 = u_d$ is the best choice to make the algorithm converge quickly. So we use this initialization to get the *solution* (u^*, v^*). In

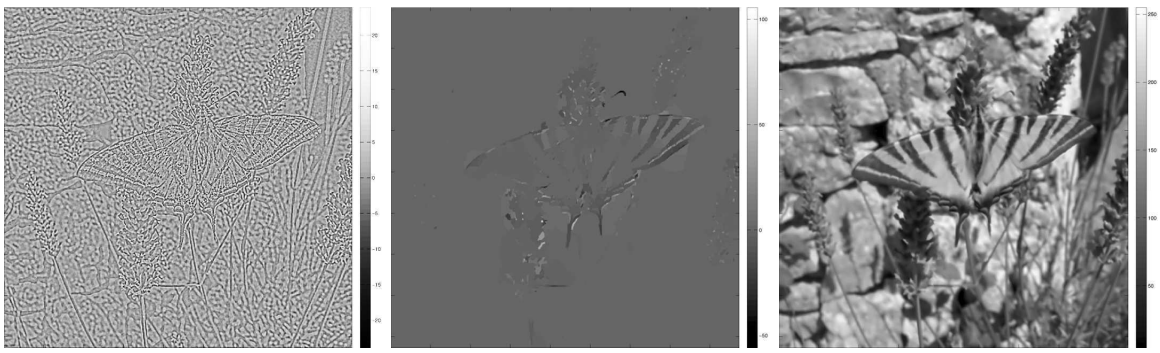
this case, the BV part is close to 0. However, we note that if we fix μ then $TV(u^*(\lambda,\mu))$ decreases to 0 and $TV^2(v^*(\lambda,\mu))$ increases to become constant (see Figure 5.11) when $\lambda \to +\infty$. This means that if λ is large then u^* is constant. As we know that u^* has a null mean value, then $u^* = 0$. This is consistent with theorem 4.4.



(a) $\|w^*\|_{L^2}$ (b) $TV(u^*)$ (c) $TV^2(v^*)$

Figure 5.11. *Generic L^2- norm, TV and TV^2 behavior (μ fixed) 400 iterations - Example 2D (Butterfly).*

On can see an example on Figure 5.7 for $\lambda = 10,\ \mu = 2$ and Figure 5.12.



(a) L^2 part (b) BV part (c) BH part

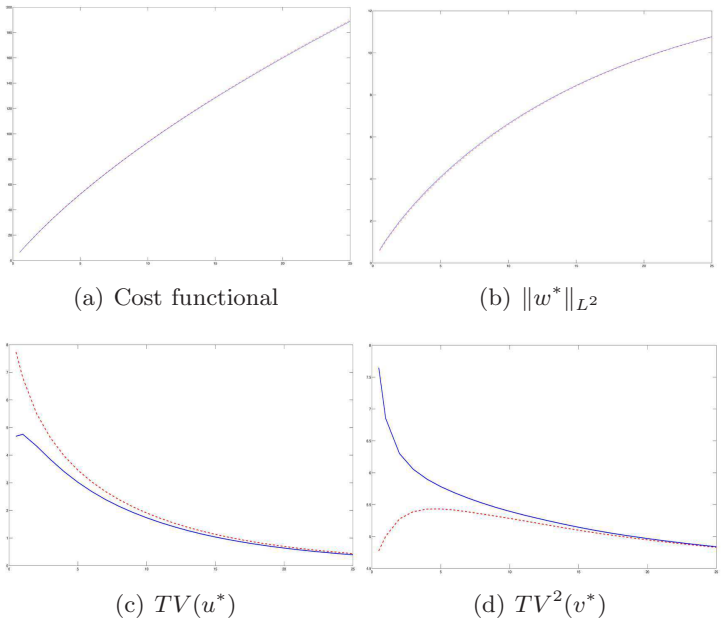
Figure 5.12. $\lambda = 7,\ \mu = 5$ - *initialization $u_0 = 0$ and $v_0 = u_d$, 400 iterations*

- If $\mu \simeq \lambda$, both initializations seem equivalent. For the *Butterfly* test, init (a) remains slightly faster (in this case the minimum value of cost functional is achieved first) while it is the converse for the *Wall* test and small values of λ. Figures 5.13 and 5.14 show the behavior of the cost functional, L^2- norm, TV and TV^2 for both initializations and $\lambda = \mu \in [0.5, 1, 2, 3 \cdots 25]$. We report the behavior of cost functional, L^2- norm, TV and TV^2 in Table 6

$\lambda = \mu$	$\mathcal{F}_{\lambda,\lambda}$	$\|w_a\|_2$	$TV(u_a)$	$TV2(v_a)$	Error
0.5	6.3577	1.459 e-03	4.685	7.646	4.13 e-01
1	12.2448	2.655 e-03	4.756	6.853	5.19 e-01
5	52.4043	9.654 e-03	3.020	5.782	5.83 e-01
10	93.2718	1.562 e-02	1.733	5.395	5.59 e-01
13	114.9198	1.835 e-02	1.268	5.237	5.11 e-01
17	141.3794	2.128 e-02	8.506 e-01	5.066	4.78 e-01
21	165.9126	2.357 e-02	5.768 e-01	4.940	4.36 e-01
25	188.8569	2.537 e-02	3.941 e-01	4.840	3.86 e-01

Table 6

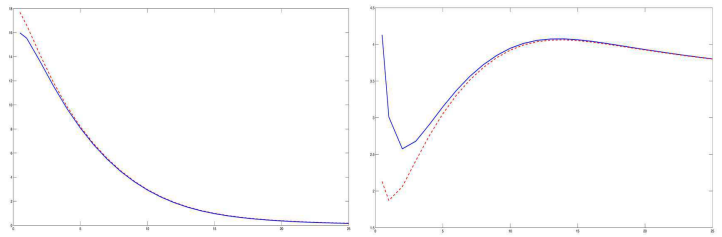
Cost functional, L^2- norm, TV and TV^2 for $\lambda = \mu = 0.5, 1, 2, 3 \cdots 25$- init (a) - 800 iterations - Butterfly



(a) Cost functional (b) $\|w^*\|_{L^2}$

(c) $TV(u^*)$ (d) $TV^2(v^*)$

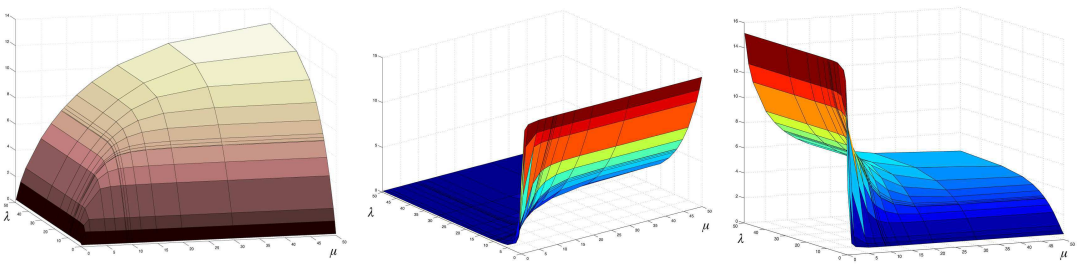
Figure 5.13. *Cost functional, L^2- norm, TV and TV^2 for $\lambda = \mu = 0.5, 1, 2, 3 \cdots 25$ - Dotted (blue) line is initialization (a) and solid (red) line is initialization (b) - 800 iterations - Butterfly test*



(a) $TV(u^*)$ (b) $TV^2(v^*)$

Figure 5.14. *TV and TV^2 for $\lambda = \mu = 0.5, 1, 2, 3 \cdots 25$ - Dotted (blue) line is initialization (a) and solid (red) line is initialization (b) - 800 iterations - Wall test*

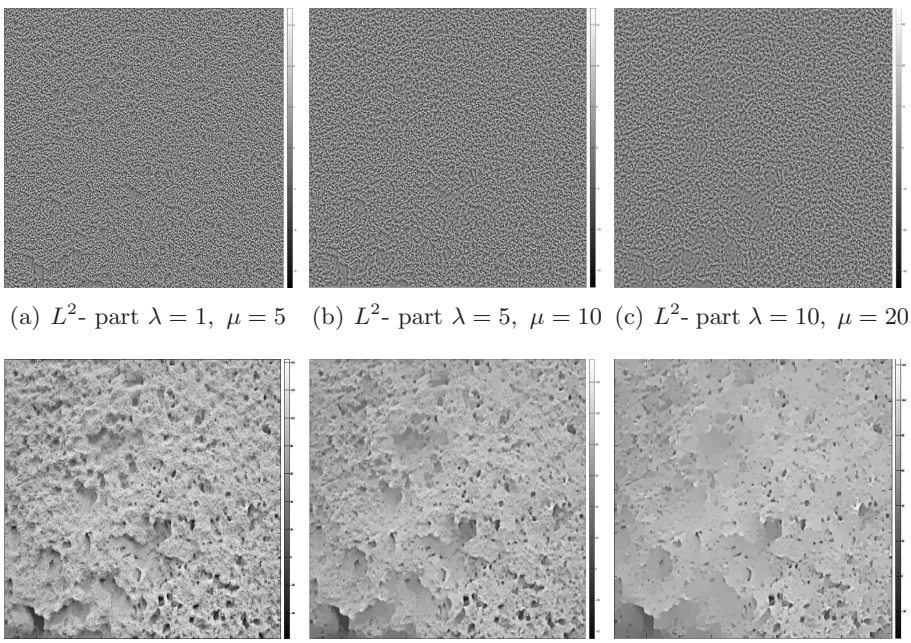
- If $\lambda << \mu$, then we choose initialization (b) : $u_0 = 0$ and $v_0 = 0$ to get the *solution*. The behavior is similar to the case $\mu < \lambda$: if we fix λ, then $TV(u^*(\lambda, \mu))$ increases to a constant value and $TV^2(v^*(\lambda, \mu))$ converges to 0 as $\mu \to +\infty$ (see figure 5.15). This means that if μ is large enough then solution is always the same : v^* is an affine fonction.



(a) $\|w^*\|_{L^2}$ (b) $TV(u^*)$ (c) $TV^2(v^*)$

Figure 5.15. *Generic L^2- norm, TV and TV^2 behavior - (λ fixed) 400 iterations - Example 2D (Butterfly).*

Examples of solutions are given in Figures 5.4, 5.7 and 5.9. We give another example below on a textured image:



(a) L^2- part $\lambda = 1,\ \mu = 5$ (b) L^2- part $\lambda = 5,\ \mu = 10$ (c) L^2- part $\lambda = 10,\ \mu = 20$

(d) BV part $\lambda = 1,\ \mu = 5$ (e) BV part $\lambda = 5,\ \mu = 10$ (f) BV part $\lambda = 10,\ \mu = 20$

Figure 5.16. *BV and L^2 components with $\lambda < \mu$ - 800 iterations - Wall example*

λ	μ	$\mathcal{F}_{\lambda,\mu}(u^*, v^*)$	$\|w^*\|_{L^2}$	$TV(u^*)$	$TV^2(v^*)$
1	5	21.6334	4.270 e-03	18.453	3.868 e-01
5	10	87.2937	1.759 e-02	10.402	1.413
10	20	152.5461	2.854 e-02	4.922	2.382

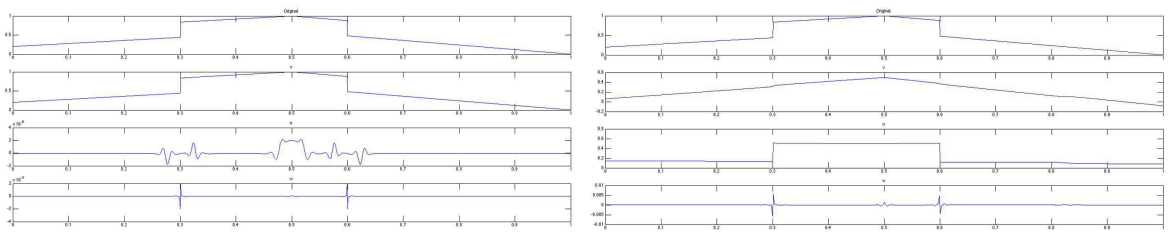
Table 7

Wall- example, initialization $u_0 = 0,\ v_0 = 0$ *- 800 iterations*

6. Conclusion . The model is well adapted to texture extraction. In the case, where the data is noiseless and/or is not too much textured, the decomposition given par $\lambda \precsim \mu$ and initialization $u_0 = v_0 = 0$, gives a cartoon part which is piecewise constant as expected. This means that $u = \sum_i u_i \mathbf{1}_{\Gamma_i}$ where $\bigcup_i \Gamma_i$ is the contour set. In this case, the remainder L^2 term is the texture and/or noise. The decomposition is robust with respect to quantification, sampling and is always the same for any $\mu >> \lambda$, once λ has been chosen.

In the case where the image is highly textured the model provides a two-scale decomposition. The TV part represents the macro-texture and the L^2 part the micro-texture and/or noise. The scaling is tuned via the ratio $\rho = \dfrac{\lambda}{\mu}$.
The notion of *highly textured* may be quantified par the G-norm. In our 2D examples, the *butterfly* G norm was $\simeq 7.71$ and the *wall* one was $\simeq 4.92$.

Figure 6.17 shows the behavior of the different components with respect to λ and μ. We have chosen the 1D noiseless case, to see the multi-scale effect on components u and w when $\mu < \lambda$ (init (a)).



(a) $\lambda = 5\ 10^{-3},\ \mu = 10^{-3}$- (init (a)) (b) $\lambda = 5\ 10^{-3},\ \mu = 10^{-2}$- (init (b))

Figure 6.17. *Test 1D without noise (1000 points)*

Moreover, the initialization process has no influence on the solution (up to a constant function) but rather on the algorithm speed. The choice has to be made with respect to the parameters: roughly speaking, if $\lambda < \mu$ we choose $u_0 = 0,\ v_0 = 0$ and if $\lambda \geq \mu$ we choose $u_0 = 0,\ v_0 = u_d$. Finally, we have observed (numerically) that the L^2-component w is unique.

Our next issue is to speed up the algorithm and set an automatic parameter tuning with respect to data properties (G norm, Signal to Noise Ratio , and so on.) From the theoretical point of view, we infer that problem $(\mathcal{P}_{\lambda,\mu})$ has a unique solution (up to $\mathbf{C}(\Omega)$) functions but the question is still open.

REFERENCES

[1] R. Acar and C.R. Vogel, *Analysis of bounded variation penalty methods for ill-posed problems*, Inverse Problems 10, (1994).

[2] L. Ambrosio, L. Faina, and R. March, *Variational approximation of a second order free discontinuity problem in computer vision*, SIAM J. Math. Anal., 32 (2001), pp. 1171–1197.

[3] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 2000.

[4] H. Attouch, G. Buttazzo, and G. Michaille, *Variational analysis in Sobolev and BV spaces*, vol. 6 of MPS/SIAM Series on Optimization, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006. Applications to PDEs and optimization.

[5] G. Aubert and J.-F. Aujol, *Modeling very oscillating signals. Application to image processing*, Appl. Math. Optim., 51 (2005), pp. 163–182.

[6] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing, Partial Differential Equations and the Calculus of Variations*, vol. 147 of Applied Mathematical Sciences, Springer Verlag, 2006.

[7] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle, *Image decomposition into a bounded variation component and an oscillating component*, J. Math. Imaging Vision, 22 (2005), pp. 71–88.

[8] M. Bergounioux, *On Poincaré-Wirtinger inequalities in* BV *- spaces*, Control & Cybernetics, 4 (2011), pp. 921–929.

[9] M. Bergounioux and L. Piffet, *A second-order model for image denoising*, Set-Valued Var. Anal., 18 (2010), pp. 277–306.

[10] ———, *A full second order variational model for multiscale texture analysis*, Computational Optimization and Applications, 54 (2013), pp. 215–237.

[11] K. Bredies, K. Kunisch, and T. Pock, *Total generalized variation*, SIAM J. Imaging Sci., 3 (2010), pp. 492–526.

[12] M. Carriero, A. Leaci, and F. Tomarelli, *Special bounded hessian and elatsic-plastic plate*, Rend. Ac. Naz delle Scienze, XVI (1992), pp. 223–258.

[13] ———, *Uniform density estimates for the blake & zisserman functional*, Discrete and Continuous Dynamical Systems, 31 (2011), pp. 1129–1150.

[14] ———, *Free gradient disconitnuity and image inpainting*, J. Math. Sciences, 181 (2012), pp. 805–819.

[15] V. Caselles, A. Chambolle, and M. Novaga, *The discontinuity set of solutions of the tv denoising problem and some extensions*, SIAM Multiscale Model. Simul., 6 (2007), pp. 879–894.

[16] A. Chambolle, *An algorithm for total variation minimization and applications*, Journal of Mathematical Imaging and Vision, 20 (2004), pp. 89–97.

[17] F. Demengel, *Fonctions à hessien borné*, Annales de l'institut Fourier, 34 (1984), pp. 155–190.

[18] I. Ekeland and R. Temam, *Convex Analysis and Variational problems*, SIAM Classic in Applied Mathematics, 28, 1999.

[19] L.C. Evans and R. Gariepy, *Measure theory and fine properties of functions*, CRC Press, 1992.

[20] J.-B. Hiriart-Urruty and R.R. Phelps, *Subdifferential calculus using ε -subdifferentials*, Journal of Functional Analysis, 118 (1993), pp. 150–166.

[21] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, vol. 22 of University Lecture Series, AMS, 2001.

[22] S. Osher, A. Sole, and Vese L., *Image decomposition and restoration using total variation minimization and the* H^1 *norm*, SIAM Journal on Multiscale Modeling and Simulation, 1-3 (2003).

[23] S. Osher and L. Vese, *Modeling textures with total variation minimization and oscillating patterns in image processing*, Journal of Scientific Computing, 19 (2003), pp. 553–572.

[24] ———, *Image denoising and decomposition with total variation minimization and oscillatory functions. special issue on mathematics and image analysis*, J. Math. Imaging Vision,, 20 (2004), pp. 7–18.

[25] W. Ring, *Structural properties of solutions of total variation regularization problems*, ESAIM, Math Modelling and Numerical Analysis, 34 (2000), pp. 799 –840.

[26] L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D, 60 (1992), pp. 259–268.

[27] P. Weiss, G. Aubert, and L. Blanc-Féraud, *Efficient schemes for total variation minimization under constraints in image processing*, SIAM Journal on Scientific Computing, 31 (2009), pp. 2047–2080.

[28] W. Yin, D. Goldfarb, and S. Osher, *A comparison of three total variation based texture extraction models*, J. Vis. Commun. Image, 18 (2007), pp. 240–252.

[29] W.P. Ziemer, *Weakly Differentiable Functions - Sobolev Space and Functions of Bounded Variation*, Springer, 1980.