Reconstruction of noisy signals by minimization of non-convex functionals

Boris Mederos[a,*], Ramón A. Mollineda[b], Julián Antolín Camarena[c]

[a]*Departamento de Física y Matemática Universidad Autónoma de Ciudad Juárez, Mexico boris.mederos@uacj.mx*
[b]*Institute of New Imaging Technologies, Universitat Jaume I, Castelló de la Plana, Spain mollined@uji.es*
[c]*University of New Mexico (UNM), USA jantolin@unm.edu*

Abstract

Non-convex functionals have shown sharper results in signal reconstruction as compared to convex ones, although the existence of a minimum has not been established in general. This paper addresses the study of a general class of either convex or non-convex functionals for denoising signals which combines two general terms for fitting and smoothing purposes, respectively. The first one measures how close a signal is to the original noisy signal. The second term aims at removing noise while preserving some expected characteristics in the true signal such as edges and fine details. A theoretical proof of the existence of a minimum for functionals of this class is presented. The main merit of this result is to show the existence of minimizer for a large family of non-convex functionals.

Keywords: Non-convex funtional, Signal denoising, Minimizer, Calculus of variations.

1. Introduction

Despite the many recent research efforts in the field of signal restoration, the design of robust solutions to the problem of signal denoising is still an open challenge. Traditionally, non-convex variational functionals has been successfully applied to this problem without any theoretical guarantee [1, Section 3.2.6].

---

*Corresponding author.

The general denoising task can be modeled as a variational problem that consists in finding a global minimizer of a functional known as *energy* or *cost* function [1, 2, 3, 4]. It is composed of two terms. The first one (the *fitting* or *fidelity* term) is intended to reward the closeness of a signal to the original noisy signal. The second one (the *potential* or *regularization* term) aims at removing noise while preserving some prior expected characteristics such as borders, details, among others. To this end, the potential term should favor the smoothing of regions with small or moderate gradients, usually corresponding to noises, while it should not penalize regions with strong gradients, usually corresponding to edges. The combination of these two terms can be understood as a trade-off between approximation accuracy and noise removal. Signals very close to the original noisy signal yield low cost values of the fitting term and high cost values in the regularization term. Otherwise, smooth signals can be substantially different from the original one, thus inducing a high cost value in the fitting term, but a low potential value.

There has been much work in the optimization of both convex and non-convex cost functionals. Convex functionals are very popular as they are easy to work with and they are guaranteed to have a minimizer. On the other hand, it has been systematically and empirically proved [1, 5] that non-convex regularization provides more chances for image restoration with well-defined edges as compared to convex methods. However, the existence of a minimum for non-convex functionals can not be established in general in continuous domains.

One of the pioneering works in the variational framework is a non-convex functional proposed by Geman and McClure [3], which has been extensively used for image restoration. After proving the non-existence of minimum for this functional, Chipot et al. [6] showed that perturbing it by adding a quadratic term yields a new non-convex functional with minimizer in one-dimensional signals, that also presents a good performance preserving edges. Formally, the proposal by Chipot et al. [6] was defined as follows:

$$E(u) = \int_a^b [u(x) - g(x)]^2 \, dx + \int_a^b \phi(u'(x)) dx \qquad (1)$$

where $\phi(t) = \frac{t^2}{\alpha^2 + t^2} + \gamma t^2$.

The first term in (1) measures how much a function $u(x)$ fits the original noisy signal $g(x)$ in the least-squares sense, and it is particularly effective

when additive Gaussian noise is present.

The second term (a regularization term) rewards more the smoothing of regions of moderate gradients (noise removal) than the smoothing of regions with large gradients (edge blurring). Therefore, it promotes clean signals with neat edges.

March et al. [7] proved the existence of a minimum for a generalized form of (1) in order to show the existence of a solution to the Perona-Malik equation [8] in the sense of BV [9]. This generalized form keeps unchanged the quadratic fitting term and introduces a general smooth robust function as a second term. However, this model does not consider non-smooth potentials which could provide better restoration results for certain problems. In this regard, Nikolova et al. [10] have shown that non-convex non-smooth potentials are better to preserve piecewise constant signals.

The works [6, 7] guarantee the existence of a minimizer for certain non-convex denoising functionals in only one-dimensional signals in continuous domains. Recently, Harjulehto et al. [11] have demonstrated the existence of a minimum of the functional (1) on arbitrary dimensions, although this result does not hold for general functionals with arbitrary non-convex potential and fidelity terms. Non-convex fitting based on M-estimators [12] has shown to be quite robust in the presence of outliers caused by impulsive noise [13, 14].

The two terms involved in the functionals proposed in [6, 11, 7] can be viewed as specific fitting and denoising solutions, respectively. In particular, the fidelity term has a quadratic form, which makes these functionals particular for Gaussian noise attenuation. However, these methods fail to remove Impulsive noise from classes of signals that arise naturally in various fields of applied sciences, for instance piecewise constant (PWC) and piecewise smooth (PWS) signals.

A theoretical study on the existence of minimizers for a family of non-convex functionals in the discrete setting is shown in [10], while [15] provides empirical evidence on the usefulness of a particular non-convex functional for image denoising and deblurring in the presence of impulsive noise. However, no proof on the existence of minimizers for nonconvex variational methods with both non-quadratic fidelity terms and edge-preserving potential terms in the continuous domain has been reported.

In this paper we present a generalization of the models discussed above

Table 1: Examples of functions $h(\cdot)$. The initials C and NC stand for Convex and Non-convex, respectively.

Name	Function	Type
Linear $\mathcal{L}^1$	t	C
Quadratic $\mathcal{L}^2$	t^2	C
$\mathcal{L}^\alpha$	$t^\alpha, 1 < \alpha < 2$	C
Quantile regression	$Q_\beta(t) = \begin{cases} (1-\beta)t & t > 0 \\ \beta(-t) & t \leq 0 \end{cases}$	C
Hubert	$H_\alpha(t) = \begin{cases} \frac{t^2}{2} & t \leq \alpha \\ \alpha t - \frac{\alpha^2}{2} & t > \alpha \end{cases}$	C
Leclerc	$\phi(t) = 1 - e^{-\alpha t^2}$	NC
Geman and McClure	$\phi(t) = \frac{\alpha t^2}{1+\alpha t^2}$	NC
LogCauchy	$LC_\gamma(t) = \log(\gamma^2 + t^2)$	NC
Lorentzian	$L_\sigma(t) = \log(1 + 1/2(t/\sigma)^2)$	NC

to accommodate more general fidelity and regularization terms:

$$\mathcal{F}(u) = \int_a^b h(|u - g|)dx + \int_a^b \phi(u')dx \tag{2}$$

The first term is formulated by means of a wide range of functions $h(\cdot)$ to measure how well a signal u fits the original signal g under different types of additive noises (e.g. Gaussian, "salt and pepper", Laplacian), provided that the signal g satisfies a condition similar to the one introduced in [7], referred in this paper to as ϕ-compatible condition (see Definition 1). Table 1 illustrates some particular functions $h(\cdot)$, comprising convex and non-convex examples. The second general term comprises a family of either smooth or non-smooth and convex or non-convex functions $\phi(\cdot)$ that are intended to preserve high-quality edges, including those that result from adding a small quadratic perturbation to functions derived from a robust estimation framework. Some illustrative examples of $\phi(\cdot)$ are listed in Tab. 2. According to [16], the robust statistical approach provides means for detecting boundaries (edges) and preserving fine details in piecewise smooth regions of a

Table 2: Examples of functions $\phi(\cdot)$ with quadratic growth $(\gamma \neq 0)$.

Functions	Type
$\phi(t) = \frac{\alpha t^2}{1+\alpha t^2} + \gamma t^2$	Smooth
$\phi(t) = 1 - e^{-\alpha t^2} + \gamma t^2$	Smooth
$\phi(t) = \min\{1, \alpha t^2\} + \gamma t^2$	Smooth
$\phi(t) = \log(\alpha t^2 + 1) + \gamma t^2$	Smooth
$\phi(t) = \frac{-1}{1+\alpha t^2} + \gamma t^2$	Smooth
$\phi(t) = \log(\alpha|t| + 1) + \gamma t^2$	Non-smooth
$\phi(t) = 1 - e^{-\alpha|t|} + \gamma t^2$	Non-smooth
$\phi(t) = \frac{\alpha|t|}{1+\alpha|t|} + \gamma t^2$	Non-smooth
$\phi(t) = \frac{-1}{1+\alpha|t|} + \gamma t^2$	Non-smooth

signal. Meanwhile, the addition of the small quadratic perturbation $(\gamma \neq 0)$, as mentioned above, guarantees the existence of the minimum for the general model without significantly change the behavior of the perturbed term. The families of functions $h(\cdot)$ and $\phi(\cdot)$ will be precisely defined in the next section through specific properties. Without loss of generality, and to ease the comprehension of the manuscript, the scope of the words *fitting* (or *fidelity*) and *potential* (or *regularization*) will be extended to denote the functions $h(\cdot)$ and $\phi(\cdot)$, respectively, apart from the two terms of $\mathcal{F}(\cdot)$ in (2) as stated before.

Besides the trend-setting works of Geman and McClure [4] and Chipot et al. [6], the practical effectiveness of non-convex functionals have also been shown in other studies. In signal/image denoising, the researches of Aubert et al. [1], Nikolova [17] and Selesnick et al. [18] illustrate by concrete examples the better performance of some non-convex functionals as compared to convex ones. Vese [19] shows the ability of several non-convex potentials to reconstruct signals, being particularly accurate at restoring edges and linear regions. In Charbonnier et al. [20], different non-convex potentials were successfully used to reconstruct 2-D single photon emission tomography (SPECT) images with both synthetic and real data. Other challenging applications on images that have benefited from the use of non-convex functionals are pixel classification [21] and segmentation [1, 22].

Numerical solutions to the minimization of functionals of the class (2) could produce apparently discontinuous discrete functions [17], which have

proven to be satisfactory, computationally feasible, and good visual quality results. In spite of their suitability, these functionals are defined on a domain $W^{1,p}(\Omega)$, $p > 1$, which does not include functions with discontinuities. To cope with such cases, a new class of functionals is defined in Sec. 5 as a generalization of (2) in the spirit of the Mumford-Shah [23] functional.

The formulation (2) represents a class of either convex or non-convex denoising functionals, from which the energy function (1) is a particular case. The main objective of this work is to prove the existence of a minimum for all functionals of this class. However, the primary motivation is to study the subclass of non-convex functionals, because of the problem of finding a minimizer for the complementary convex cases is well studied. As far as we know, no work has analytically studied the existence of a minimum for a family of non-convex denoising functionals as general as the one encompassed by the class $\mathcal{F}(\cdot)$. To have theoretical guarantees for the existence of that minimum would provide support to search for other well-posedness non-convex functionals besides (1), which could potentially produce better results at preserving edges and fine signal features in specific tasks.

In addition, it is shown that the denoising proposal $\mathcal{F}(\cdot)$ is a suitable solution to reconstruct signals of two popular application domains. A first problem domain is determined by PWC signals contaminated with Impulsive noise, White Gaussian noise, and mixtures of them. A second domain is defined by PWS signals corrupted with White Gaussian noise, both alone and jointly with Impulsive noise. These denoising problems might result from scientific and engineering areas such as cross-hybridization of DNA [24], molecular bioscience [25], reconstruction of brain stimuli [26], among others. From a practical point of view, impulsive noises may appear due to sensor malfunctioning, while White Gaussian noise is mainly caused by poor illumination, high temperature, transmission errors, among others.

The high quality of the minimization results for some functionals of the class $\mathcal{F}(\cdot)$ is asserted in the works of Nikolova [17] and Bar et al. [15], which successfully deals with both PWC or PWS signals corrupted with Impulsive and/or White noise by using edge-preserving regularization methods with non-smooth fidelity terms such as $\mathcal{L}^1$, Student's t-distribution, and M-estimators.

Summarizing, the main contributions of this paper are highlighted below:

- A class of functionals for robust signal denoising that combine two general terms for fitting and smoothing purposes.

6

- A theoretical proof of the existence of a minimum (existence Theorem) for any functional that belongs to the general class, provided that the original noisy signal satisfies the ϕ-compatible condition stated in Definition 1.

  It is shown that two problem domains based on widespread signals (PWC, PWS) corrupted by common noises (Impulsive, White Gaussian) satisfy the ϕ-compatible condition. Therefore, the existence theorem guarantees a fair solution to these denoising tasks. In addition, as a corollary of the existence Theorem, the existence of minimizers for a class of Mumford-Shah functionals is demonstrated, in order to theoretically accommodate discontinuous solutions.

- A novel proof methodology to establish the existence of a minimum in optimization functionals within the signal processing field, which has been adopted from a number of theoretical works developed in different application domains such as material science, physics of phase transitions, and fracture mechanics [27, 28, 29, 30, 31].

Due to the huge amount of empirical evidence on the usefulness of non-convex functionals on image and signal denoising and reconstruction, the scope of this work does not include practical experiments.

This work has been organized in four main parts. Section 2 characterizes the class of functionals $\mathcal{F}()$, introduced in Eq. (2), and formally states the main result of this paper, the existence theorem of a minimum, although no proof is here provided. Section 3 contains a number of handy definitions and theorems that will be required to prove the main theorem. In Section 4, the proof of the existence Theorem is rigorously written. Finally, conclusions and promising directions for future research are given in Section 5.

2. Formal statement of the existence theorem of a minimum for each functional of the class $\mathcal{F}(\cdot)$

This section is intended to introduce the main result of this paper. It consists in a theorem that establishes the existence of a minimum for each functional of the class $\mathcal{F}(\cdot)$, introduced in Eq. (2), for edge-preserving signal restoration. The scope of this section includes a formal characterization of two families of functions, $h(\cdot)$ and $\phi(\cdot)$, required to build the general fitting and potential terms, respectively, which can be combined to define the class of

functionals $\mathcal{F}(\cdot)$. The analysis presented hereafter is carried out in arbitrary dimensions.

To ease the comprehension of this section, the class of functionals $\mathcal{F}(\cdot)$ stated in (2) is reintroduced below:

$$\mathcal{F}(u) = \int_\Omega h(|u - g|)dx + \int_\Omega \phi(\nabla u)dx,$$

where the domain $\Omega \subset \mathbb{R}^n$, $n \geq 1$, is an open and bounded set, the function $u : \Omega \to \mathbb{R}$ belongs to the Sobolev space

$$W^{1,p}(\Omega) = \{f \in L^p(\Omega) \ : \ \nabla f \in L^p(\Omega)\}$$

with $p > 1$, and the noisy signal g belongs to $L^\infty(\Omega)$.

In order to state the existence theorem, several properties of the fidelity and potential functions h and ϕ need to be formulated. First, five properties (F1, F2, F3, F4, F5) of h are listed:

Non-negativity: The data-fidelity function $h : [0, +\infty) \to \mathbb{R}$ satisfies,

$$h(x) \geq 0, \ \forall x \in [0, +\infty). \tag{F1}$$

Zero-value at $x = 0$: The function h maps the origin into the origin:

$$h(0) = 0. \tag{F2}$$

Locally Lipschitz: The function h is piecewise Lipschitz, i.e., there exists a collection of open sets B_i, $\cup B_i = \mathbb{R}$, and a set of corresponding constants $K_i > 0$ such that,

$$|h(x) - h(y)| \leq K_i|x - y|, \ \forall x, y \in B_i \tag{F3}$$

Strictly increasing function: The function h satisfies,

$$x < y \Rightarrow h(x) < h(y), \ \forall x, y \in [0, +\infty). \tag{F4}$$

This condition implies the existence of the derivative h' everywhere except at a Lebesgue zero-measure set S, as shown in [32].

Lower boundedness of the derivative: Given any interval $[a, b] \subset (0, +\infty)$, h' satisfies

$$\inf_{t \in [a,b] \backslash S} h'(t) = w_{a,b} > 0. \tag{F5}$$

In the case of the potential function ϕ, four other properties (P1, P2, P3, P4) are discussed below. For the purpose of establishing (P4), and by a common practice to minimize a non-convex functional, the convex envelope of ϕ is built and is denoted as $\phi^{**} : \mathbb{R}^n \to [0, +\infty]$. In addition, to state (P3) and (P4), let us assume that the *detachment set*

$$\mathcal{A}^\phi = \{ z \in \mathbb{R}^n : \phi^{**}(z) \neq \phi(z) \},$$

is a countable union of disjoint open sets $\mathcal{A}_i^\phi$, that is $\mathcal{A}^\phi = \bigcup_{i=1}^\infty \mathcal{A}_i^\phi$. These properties are listed below:

Global minimum: The potential reaches a global minimum at 0,

$$\phi(0) = \min_{z \in \mathbb{R}^n} \phi(z). \tag{P1}$$

Grow order: There exist $L_2 \geq L_1 > 0$, $\mu_2 \geq \mu_1 \geq 0$ and $p > 1$ such that

$$L_1(\mu_1 + |z|^2)^{p/2} \leq \phi(z) \leq L_2(\mu_2 + |z|^2)^{p/2}. \tag{P2}$$

Local convexity at zero: The detachment set $\mathcal{A}^\phi$ does not contain zero:

$$0 \notin \mathcal{A}^\phi. \tag{P3}$$

This implies that ϕ is locally convex at zero, and the converse is also true.

Affinity: The restriction of ϕ^{**} to each $\mathcal{A}_i^\phi$ of the detachment set is an affine function
$$\phi^{**}(z) = n_i \cdot z + l_i, \ \forall z \in \mathcal{A}_i^\phi, \tag{P4}$$

where $n_i \in \mathbb{R}^n$ and $l_i \in \mathbb{R}$. This is a standard condition used to prove the existence of minimum of non-convex functionals (see [33, 34, 27, 35, 30, 31]).

Moreover, the original noisy signal g has to satisfy the condition stated in the Definition 1, which is similar to the one proposed by March et al. [7] and it also holds in the Chipot's approach. Following the notation introduced in [7], this is called here ϕ-compatible, because its formulation depends on ϕ. This condition is expressed in terms of i) the Lebesgue measure, denoted by $\mathcal{L}^n$, ii) the approximate derivative of g, denoted by $ap\, Dg(x)$, and iii) the set of points where g is approximately derivable, denoted by D_g.

Definition 1 (ϕ-compatible condition): We say that a signal $g : \Omega \to \mathbb{R}$ satisfies the ϕ-compatible condition if

$$\mathcal{L}^n \left\{ x \in D_g : ap\, Dg(x) \in \mathcal{A}^\phi \right\} = 0. \tag{3}$$

As commented in the Introduction section, two popular families of signals with practical interest (PWC, PWS), when affected by certain types of noises, fulfill the ϕ-compatible condition. Next, a brief explanation is provided for each case:

- PWC signals both noiseless and contaminated with Impulsive noise, White Gaussian noise, and mixtures of them satisfy (3) when $0 \notin \mathcal{A}^\phi$. Since PWC signals corrupted with Impulsive noise keep the PWC structure, they can be mathematically represented as simple measurable functions $g(x) = \sum_{k=1}^\infty a_k \chi_{F_k}(x)$. Then, it is easy to see that $ap\, Dg(x) = 0$ and thus,

$$\left\{ x \in D_g : ap\, Dg(x) \in \mathcal{A}^\phi \right\} = \emptyset. \tag{4}$$

  Therefore, (3) holds.

- As shown next, PWS signals contaminated with a Gaussian Process (GP) or a mixture of Impulsive noise and GP also satisfy (3). Given any PWS signal g (with finite approximate derivative at almost every point), let us assume that g is contaminated with a realization of a stochastic process X, producing a corrupted signal denoted by $g + X$. If X has infinite approximate derivative at almost every point, then $ap\, D(g + X)(t) = ap\, Dg(t) + ap\, DX(t) = \infty$ for almost every $t \in [a, b]$. Therefore, $g + X$ satisfies (4) and, consequently, (3).

  Using a GP X as a particular case of a stochastic process, it is shown that, under some conditions, $ap\, DX(t) = \infty$ at almost every point $t \in [a, b]$, as first deduced in Berman [36, Lemma 3.1, p. 1264]. As it is demonstrated in that work, a GP with stationary increments and with jointly continuous local time $\psi(x, t)$ has almost surely realizations $X(t)$ with infinite approximate derivative for almost every t. To this end, Berman stated several independent theorems to prove that a GP with stationary increments has jointly continuous local time. One of them (Theorem 4.1 [36, p. 1265]), with a remarkable practical interest, states that a GP with mean 0, stationary increments, and continuous

$\sigma(t) = E[(X(t) - X(a))^2]$ has jointly continuous local time almost surely, if the following condition

$$\int_a^b \int_a^b \sigma(t-s)^{-(2+\epsilon)} \, dt \, ds < +\infty \tag{5}$$

holds for some $\epsilon > 0$.

In particular, the classical White Gaussian noise corresponds to a family $X(t), t \in [a, b]$, of i.i.d random variables with zero mean and constant variance σ^2. Thus, this type of noise satisfies $(5)^1$.

The existence of minimum is demonstrated through a theorem that is formulated on the basis of the conditions (F1) - (F5) and (P1) - (P4) and having a signal g satisfying the ϕ-compatible condition.

Theorem 2.1. *Let Ω be an open and bounded subset of $\mathbb{R}^n$, $n \geq 1$. Let $h : [0, +\infty) \to [0, +\infty)$ and $\phi : \mathbb{R}^n \to \mathbb{R}$ be the data-fidelity and potential functions, respectively, such that they satisfy the conditions (F1) - (F5) and (P1) - (P4), respectively. Let us assume that $g \in L^\infty(\Omega)$ satisfies the ϕ-compatibility condition (Definition 1). Then, any functional of the class $\mathcal{F}(\cdot)$ has a minimizer on $W^{1,p}(\Omega)$, $p > 1$.*

The existence result stated by the above theorem is more general than that provided by the Chipot et al.'s theorem [6]. More formally, the particular potential ϕ of the functional (1) proposed by Chipot et al.

$$\phi(t) = \frac{t^2}{\alpha^2 + t^2} + \gamma \, t^2,$$

satisfies the affinity condition (P4) with a detachment set $\mathcal{A}^\phi = (-b, -a) \cup (a, b)$, where a and b are positive numbers which depend on ϕ. Consequently, $0 \notin \mathcal{A}^\phi$. Besides, the theorem given by Chipot et al. is formulated for simple measurable functions which, as we note above, satisfy the ϕ-compatible condition. Therefore, the existence of minimum for the functional (1), proposed by Chipot et al., can be also deduced from the Theorem 2.1.

At this point two main issues should be highlighted:

---

[1]Note that $\sigma(t-s) = 2\sigma^2$, $\forall t, s \in [a, b]$.

- Although this result is stated for arbitrary dimensions n, the most appealing applications can be found for $n = 1$ which correspond to functionals of the class $\mathcal{F}(\cdot)$ used in signal denoising.

- The previous theorem does not guarantee the existence of a minimizer for the functional of Chipot et al. on arbitrary dimensions. More formally, when this functional is formulated for $n = 2$ (image denoising), the potential term is defined by the following radial function,

$$\phi(z) = \frac{\|z\|_2^2}{\delta^2 + \|z\|_2^2} + \gamma\|z\|_2^2, \ z \in \mathbb{R}^2,$$

  which does not meet the condition (P4). In fact, some of the most widely used potentials in image processing ($n = 2$) adopt a radial function pattern, but none of these functions satisfy (P4).

Recently Harjulehto et al. [11] have demonstrated the existence of a minimum for the functional (1) of Chipot et al. in arbitrary dimensions, on the basis of the convexity of the quadratic fidelity term. In fact, this proof also holds for any convex fidelity term, but it is not valid for non-convex ones. Therefore, this result is not extensible to all functionals that fit the template $\mathcal{F}(\cdot)$ introduced in (2), which admit either convex or non-convex fidelity terms. The latter ones have shown to be quite robust at fitting signals in the presence of outliers [13, 14].

Summarizing, we extend the functional introduced in [6] to a class defined by general fidelity and potential functions h and ϕ, respectively. In particular, h can be chosen from a wide range of functions to fit signals under different types of additive noises. With regard to ϕ, it is intended to preserve high-quality edges, and can be any quasi-robust function (see [37]) that needs not be differentiable. Finally, a theorem was formulated to establish the existence of a minimum for every functional in the class $\mathcal{F}(\cdot)$, provided that the signal satisfies the ϕ-compatible condition introduced in Definition 1.

3. Preliminary definitions

This section is intended to provide handy definitions and a lemma that will be required to prove the main result of the paper. In brief, the following topics are introduced:

Definition of Convex Envelope. It is the basis of a strategy to deal with a non-convex functional, which consists in building and minimizing the closest convex functional. Under some conditions, the solution of this convex functional is also the solution of a related non-convex functional.

Definition of Approximate Limit. It is a preliminary concept to define the *approximate differentiability.*

Definition of Approximate Differentiability. This definition is a more general concept of differentiability that supports the next point.

Equality of Approximate Derivative Lemma. It provides a condition to ensure that two measurable functions have equal approximate derivatives. It partially supports that (1) has a minimizer when the signal satisfies (3).

Space of Special Functions of Bounded Variation. It is denoted by $SVB(\Omega)$, with Ω being a bounded open set. It plays a key role in proving the existence of minimum for the functional of Mumford-Shah [23].

The above points are accurately stated hereafter.

Definition 2 (Convex envelope): Given a semicontinuous function $\psi : \mathbb{R}^n \to \mathbb{R}$, its convex envelope $\psi^{**} : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\psi^{**}(x) = \sup\left\{l(x) : l \text{ is convex, } l(y) \leq \psi(y) \ \forall y \in \mathbb{R}^n\right\}.$$

It is easy to note that if ψ meets the conditions (P1) - (P4), then ψ^{**} also satisfies them.

A usual strategy to prove the existence of a minimizer for any functional of the type $\mathcal{F}(\cdot)$ is to work with an associated auxiliary functional that fits the model $\hat{\mathcal{F}} : W^{1,p}(\Omega) \to \mathbb{R}$,

$$\hat{\mathcal{F}}(u) = \int_\Omega h(|u - g|)\, dx + \int_\Omega \phi^{**}(\nabla u)\, dx = \int_\Omega \hat{f}(x, u, \nabla u)\, dx, \qquad (6)$$

where $\hat{f}(x, u, \nabla u) = h(|u-g|) + \phi^{**}(\nabla u)$ is the convex envelop of $f(x, u, \nabla u) = h(|u - g|) + \phi(\nabla u)$ with respect to the third variable ∇u. From the conditions imposed on h and ϕ in Section 2 it follows that (6) is semicontinuous

with respect to the weak topology of the Sobolev space, which guarantees the existence of a minimizer $\hat{u}$ for a generic functional of the class $\hat{\mathcal{F}}(\cdot)$. Besides, if the following equality

$$\hat{f}(x, \hat{u}, \nabla \hat{u}) = f(x, \hat{u}, \nabla \hat{u})$$

holds almost everywhere on Ω, then from the Ekeland-Teman's theorem in [38, Theorem 3.8, Chapter 10] $\hat{u}$ is also a minimizer of the associated original functional of the type $\mathcal{F}(\cdot)$. Since the conventional concepts of limit and derivative are not suitable in general for measurable functions, a density concept is introduced to define the *approximate limit* and the *approximate differentiability*. They extend the classical differentiability analysis to the class of measurable functions.

Given any measurable set $D \subset \mathbb{R}^n$, the upper and lower densities of D at x are defined respectively as

$$\Theta^*(D, x) = \limsup_{r \to 0} \frac{|B(x, r) \cap D|}{|B(x, r)|} \qquad \Theta_*(D, x) = \liminf_{r \to 0} \frac{|B(x, r) \cap D|}{|B(x, r)|}.$$

Whenever $\Theta^*(D, x) = \Theta_*(D, x)$ holds, we will simply write $\Theta(D, x)$. For more details about densities, the reader should refer to Ambrosio-Fusco-Pallara [9].

Definition 3 (Approximate limit): Given a measurable function $u : \Omega \to \mathbb{R}$, it is said that l is the *approximate limit* of u at x if the set

$$E_\epsilon = \{y \in \Omega : |u(y) - l| > \epsilon\}$$

has zero upper density at x for all $\epsilon > 0$, i.e. $\Theta^*(E_\epsilon, x) = 0$, $\forall \epsilon > 0$. We will denote l by ap $\lim_{y \to x} u(y)$. The approximate discontinuity set S_u of u is defined as

$$S_u = \left\{ x \in \Omega : \text{ap} \lim_{y \to x} u(y) \text{ does not exist} \right\}.$$

Definition 4 (Approximate differentiability): Let $D \subset \mathbb{R}^n$ be a measurable set, and $v : D \to \mathbb{R}$ be a measurable function. It is said that v is *approximately differentiable* at $x \in D$ if there exist $a_x \in \mathbb{R}^n$ such that

$$\text{ap} \lim_{y \to x} \frac{|v(y) - v(x) - a_x(y - x)|}{|y - x|} = 0.$$

$D_v \subset D$ will denote the set of points where v is approximately differentiable. The vector a_x will be denoted by ap $Dv(x)$, and it will be called *approximate derivative* of v at x.

The following lemma establish when two functions have equal approximate derivatives, and it can be found at Giaquinta-Modica-Saucĕk [39].

Lemma 3.1. *Let D be a measurable set. Let us assume that $u : \Omega \to \mathbb{R}$ is almost everywhere approximately differentiable in D, and $v : \Omega \to \mathbb{R}$ satisfies that $v = u$ a.e. in D. Then, $v : \Omega \to \mathbb{R}$ is a.e. approximately differentiable in D and ap $Dv = $ ap Du a.e. in D.*

It should be said that any function in the space $W^{1,p}(\Omega)$, $p \geq 1$, or in the space of bounded variation functions $BV(\Omega)$, is almost everywhere approximately differentiable (see [9, 39]).

Definition 5 (Space of Special Functions of Bounded Variation):

$$SBV(\Omega) = \{u \in BV(\Omega) \; : \; \text{supp}\, D^s u = S_u\}, \tag{7}$$

where $BV(\Omega)$ is the space of bounded variation functions, and $D^s u$ is the singular part of Du with respect to $\mathcal{L}^n$, that is, $Du = \nabla u \mathcal{L}^n + D^s u$.

A useful subspace of $SBV(\Omega)$ is $W^{1,p}(\Omega \setminus K)$, with K being a closed subset of Ω with finite $(n-1)$-Hausdorff measure. This subspace is a constituent part of the Mumford-Shah functional domain.

4. Existence theorem of a minimum for each functional of the class $\mathcal{F}(\cdot)$

This section presents the proof of the existence of a minimizer for a generic functional of the type $\mathcal{F}(\cdot)$ when the signal g satisfies the ϕ-compatible condition.

The proof consists of two parts. The first one is a regularity result that proves the differentiability (almost everywhere in Ω) of any minimizer of an associated auxiliary functional that fits $\hat{\mathcal{F}}(\cdot)$ (Theorem 4.1). The second part exploits the regularity result to demonstrate the existence of a minimum of the generic functional of $\mathcal{F}(\cdot)$, which is the full version of Theorem 2.1 (the reader will find it below in this section). This proof also relies on an intermediate result stated in the Theorem 4.3.

Theorem 4.1 (Regularity). *Let Ω be an open and bounded subset of $\mathbb{R}^n$, $n \geq 1$. Let $\hat{u}$ be a minimizer of an auxiliary functional of the type $\hat{\mathcal{F}}(\cdot)$ on the space $W^{1,p}(\Omega)$, $p \geq 1$ with the functions $h : [0, +\infty] \to [0, +\infty)$ and $\phi^{**} : \mathbb{R}^n \to \mathbb{R}$, satisfying conditions (F1) - (F5) and (P1) - (P4), respectively. Then $\hat{u}$ is almost everywhere differentiable.*

In the proof of this Theorem, we will distinguish three cases depending on the value of p. However, since all arguments are well established in the related literature, only useful guidelines are provided. For the case of $n = 1$, the proof is based on the well-known fact that any function $u \in W^{1,p}(\Omega)$, with $p \geq 1$ and Ω being a bounded open subset of $\mathbb{R}$, is almost everywhere differentiable in the classical sense [40, Section 2.2, Theorem 2.14]. For the case of $1 < n < p$, the conclusion follows from the Morrey's inequality [41]. Finally, when $1 < p \leq n$, the core ideas of the proof come from the works of Giaquinta and Giusti [42], Giusti [43, Theorem 7.1 and 7.6] and Celada and Perrota [44, Section 3, Theorem 3.1], except for a few minor modifications to suitably manage the functional of the class $\mathcal{F}(\cdot)$.

Theorem 4.1 guarantees that any minimizers $\hat{u}$ of the auxiliary functional of $\hat{\mathcal{F}}(\cdot)$ is almost everywhere differentiable in Ω. This result allows the definition of the following sets

$$E_i = \left\{ x \in \Omega : \nabla \hat{u}(x) \in \mathcal{A}_i^{\phi} \right\}, \, i = 1, \ldots, \infty. \tag{8}$$

Moreover, the differentiability of $\hat{u}$ permits the first order Taylor expansion of $\hat{u}$ almost everywhere in Ω. The proof that $\hat{u}$ is also a minimizer of the target generic functional of the type $\mathcal{F}(\cdot)$ requires to demonstrate that all E_i have Lebesgue measure equal to zero, i.e. $|E_i| = 0$. However, the proof of $|E_i| = 0$ requires first to demonstrate $\|\hat{u}\|_{L^{\infty}(\Omega)} \leq \|g\|_{L^{\infty}(\Omega)}$, which is addressed in the next Lemma.

Lemma 4.2. *Let Ω be an open and bounded subset of $\mathbb{R}^n$, $n \geq 1$. Let us assume that $g \in L^{\infty}(\Omega)$. There exists a minimizer $\hat{u}$ of an auxiliary functional of the type $\hat{\mathcal{F}}(\cdot)$ on $W^{1,p}(\Omega)$, $p > 1$ with function $h : [0, +\infty] \to [0, +\infty)$ and $\phi^{**} : \mathbb{R}^n \to \mathbb{R}$ being the data-fidelity and potential functions, respectively, such that they satisfy the conditions (F1) - (F5) and (P1) - (P4), respectively. Then $\hat{u}$ satisfies $\|\hat{u}\|_{L^{\infty}(\Omega)} \leq \|g\|_{L^{\infty}(\Omega)}$.*

PROOF. The proof of this Lemma follows from the same arguments used in Chipot et al. [6] (the truncation argument) and the property (F4). If $\bar{u}$ is a

minimizer of $\hat{\mathcal{F}}(\cdot)$, then the function

$$\hat{u}(x) = \begin{cases} \|g\|_{L^\infty(\Omega)}, & \text{if } |\bar{u}(x)| \geq \|g\|_{L^\infty(\Omega)} \\ \bar{u}(x), & \text{otherwise.} \end{cases} \tag{9}$$

satisfies $h(|\hat{u}(x) - g(x)|) \leq h(|\bar{u}(x) - g(x)|)$ for a.e., $x \in \Omega$ which entails

$$\int_\Omega h(|\hat{u}(x) - g(x)|)\, dx \leq \int_\Omega h(|\bar{u}(x) - g(x)|)\, dx. \tag{10}$$

From (9) follows that

$$\nabla\hat{u}(x) = \begin{cases} 0, & \text{if } |\tilde{u}(x)| \geq \|g\|_{L^\infty(\Omega)} \\ \nabla\bar{u}(x), & \text{otherwise.} \end{cases} \tag{11}$$

which satisfies $\phi^{**}(\nabla\hat{u}(x)) \leq \phi^{**}(\nabla\bar{u}(x))$ for a.e., $x \in \Omega$ due to property (P1). Therefore, the following inequality holds

$$\int_\Omega \phi^{**}(\nabla\hat{u})\, dx \leq \int_\Omega \phi^{**}(\nabla\bar{u})\, dx. \tag{12}$$

From (10) and (12) we obtain that $\hat{u}$ is a minimizer.

The next theorem takes a first step towards the proof of the result $|E_i| = 0$, by using some arguments presented by Dacorogna and Marcellini [35], Fonseca, Fusco and Marcellini [29] and Celada, Perrota and Guidorzi [45].

Theorem 4.3. *Let Ω be an open and bounded subset of $\mathbb{R}^n$, $n \geq 1$. Let $\hat{u}$ be a minimizer of an auxiliary functional of the type $\hat{\mathcal{F}}(\cdot)$ on $W^{1,p}(\Omega)$, $p > 1$ with function $h : [0, +\infty] \to [0, +\infty)$ and $\phi^{**} : \mathbb{R}^n \to \mathbb{R}$ being the data-fidelity and potential functions, respectively, such that they satisfy the conditions (F1) - (F5) and (P1) - (P4), respectively. Let us assume that $g \in L^\infty(\Omega)$ satisfies the ϕ-compatibility condition (Definition 1). Then we have either $|E_i| = 0$ or $\hat{u}(x) = g(x)$ for a.e. $x \in E_i$.*

PROOF. The strategy used to demonstrate the theorem is "proof by contradiction". Suppose that,

$$\mathcal{L}^n \{E_i \cap \{x \in \Omega : g(x) \neq \hat{u}(x)\}\} > 0.$$

17

Under this assumption, the proof focuses on building $\tilde{u}$ as a modification of the minimizer $\hat{u}$, in order to induce a decreasing of the functional $\hat{\mathcal{F}}(\cdot)$, i.e. $\hat{\mathcal{F}}(\tilde{u}) < \hat{\mathcal{F}}(\hat{u})$. But, this is a contradiction with the fact that $\hat{\mathcal{F}}(\hat{u})$ is the minimum value of $\hat{\mathcal{F}}(\cdot)$.

This modification will be accomplished only on a particular subset of Ω through a linear function, that will be denoted by $G_r(x_0)$ and $v^r_{x_0}(\cdot)$, respectively. Both entities are defined below. Formally,

$$\tilde{u}(x) = \begin{cases} v^r_{x_0}(x), & \text{if } x \in G_r(x_0) \\ \hat{u}(x), & \text{if } x \in \Omega \setminus G_r(x_0). \end{cases} \tag{13}$$

The proof method is based on finding $v^r_{x_0}(x)$ and $G_r(x_0)$ such that (14) and (15) are satisfied.

$$\int_\Omega \phi^{**}(\nabla \tilde{u}) \, dx \le \int_\Omega \phi^{**}(\nabla \hat{u}) \, dx \tag{14}$$

$$\int_\Omega h(|\tilde{u} - g|) \, dx < \int_\Omega h(|\hat{u} - g|) \, dx \tag{15}$$

Note that by adding (14) and (15), the contradiction $\hat{\mathcal{F}}(\tilde{u}) < \hat{\mathcal{F}}(\hat{u})$ comes to light.

This process is carried out in two parts. The first part is aimed at proving (14), which means that the potential value does not increase. The second part is devoted to demonstrate (15), which means that the fitting term decreases.

Part 1: Proof of (14).
Let us assume that $x_0 \in E_i \cap \{x \in \Omega : g(x) \ne \hat{u}(x)\}$ such that there exists $\nabla \hat{u}(x_0)$ by the Theorem 4.1. By the differentiability of $\hat{u}$ at x_0, using Taylor expansion in x_0, we have

$$\hat{u}(x) = \hat{u}(x_0) + \nabla \hat{u}(x_0)(x - x_0) + o(\|x - x_0\|_2)$$

with

$$\frac{|o(\|x - x_0\|_2)|}{\|x - x_0\|_2} < \gamma \le 1, \forall x \in B_\delta(x_0) \tag{16}$$

for some δ. Let us take $r < \delta$ to define $v^r_{x_0}(\cdot)$:

$$v^r_{x_0}(x) = \begin{cases} s_{x_0}(x) - w_{x_0}(x), & \text{if } g(x_0) < \hat{u}(x_0) \\ s_{x_0}(x) + w_{x_0}(x), & \text{if } g(x_0) > \hat{u}(x_0) \end{cases}$$

18

where,

$$
\begin{aligned}
s_{x_0}(x) &= \hat{u}(x_0) + \nabla \hat{u}(x_0)(x - x_0), \\
w_{x_0}(x) &= \gamma(r - 2\|x - x_0\|_2)
\end{aligned}
$$

with $x \in B_r(x_0) \subset \Omega$, s_{x_0} is a local linear approximation to the function $\hat{u}$ at x_0, and w_{x_0} is a correction term.

Let us now define $G_r(x_0)$ as,

$$
G_r(x_0) = \begin{cases} G_r^-(x_0), & \text{when } g(x_0) < \hat{u}(x_0) \\ G_r^+(x_0), & \text{when } g(x_0) > \hat{u}(x_0) \end{cases}
$$

where $G_r^-(x_0)$ and $G_r^+(x_0)$ are the closed sets around x_0, which are formulated as follows:

$$
G_r^-(x_0) = \{x \in B_\delta(x_0) : v_{x_0}^r(x) \leq \hat{u}(x)\}, \, G_r^+(x_0) = \{x \in B_\delta(x_0) : v_{x_0}^r(x) \geq \hat{u}(x)\}.
$$

The set $G_r(x_0)$ can be considered the cornerstone of the subsequent deduction. The study of $G_r(x_0)$ involves the analyses of $G_r^-(x_0)$ and $G_r^+(x_0)$. Due to their similarity, without loss of generality, hereinafter it will be assumed that $G_r(x_0) = G_r^+(x_0)$.

Next, some useful properties of $G_r(x_0)$ and $v_{x_0}^r(\cdot)$ are given, which are going to be used in both the part 1 and the part 2:

1. $B_{r/4}(x_0) \subset G_r(x_0) \subset B_r(x_0) \subset B_\delta(x_0)$.

   From right to left, the first inclusion is obvious since $r < \delta$. To prove the second one, we will proceed by contradiction. Suppose there exists $x \in G_r(x_0) \backslash B_r(x_0)$ which implies $r < \|x - x_0\|_2$, and from the definition of $G_r(x_0)$ we get $\|x - x_0\|_2 < \delta$. Consequently,

$$
\begin{aligned}
v_{x_0}^r(x) - \hat{u}(x) &= \gamma(r - 2\|x - x_0\|_2) - o(\|x - x_0\|_2) \\
&= \gamma(r - \|x - x_0\|_2) - \gamma\|x - x_0\|_2 - o(\|x - x_0\|_2).
\end{aligned}
$$

   Since $r < \|x - x_0\|_2$, we have $\gamma(r - \|x - x_0\|_2) < 0$. Then, we arrive to the next inequality:

$$
\begin{aligned}
v_{x_0}^r(x) - \hat{u}(x) &\leq -\gamma\|x - x_0\|_2 - o(\|x - x_0\|_2) \\
&= \|x - x_0\|_2 \left(-\gamma - \frac{o(\|x - x_0\|_2)}{\|x - x_0\|_2}\right) < 0
\end{aligned}
$$

19

where the last inequality follows from (16). This result is a contradiction because it matches with the definition of $G_r^-(x_0)$, but we assumed $G_r(x_0) = G_r^+(x_0)$.

The leftmost inclusion $B_{r/4}(x_0) \subset G_r(x_0)$ can be deduced in a similar way, which implies that $G_r(x_0)$ is not empty.

2. $\tilde{u} \in W^{1,p}(\Omega)$.

From the definition of $G_r(x_0)$, the function $\tilde{u}$, firstly defined in Eq. (13), can be rewritten as

$$\tilde{u} = \begin{cases} \max\{\hat{u}(x), s_{x_0}(x) + w_{x_0}(x)\}, & \text{if } x \in B_r(x_0) \\ \hat{u}(x), & \text{if } x \in \Omega \setminus B_r(x_0). \end{cases} \tag{17}$$

Since $\hat{u}(x) \geq s_{x_0}(x) + w_{x_0}(x)$ when $x \in \partial(B_r(x_0))$, then the two function pieces of (17) agree on $\partial(B_r(x_0))$.

In addition, from [32, Theorem 4, Section 4.2], $\max\{\hat{u}(x), s_{x_0}(x) + w_{x_0}(x)\} \in W^{1,p}(B_r(x_0))$ and, by the very definition of $\hat{u}(x)$, $\hat{u} \in W^{1,p}(\Omega \setminus B_r(x_0))$.

From the continuity of $\tilde{u}$ at $\partial(B_r(x_0))$ and the belonging of their two pieces to their respective Sobolev spaces, we conclude that $\tilde{u} \in W^{1,p}(\Omega)$.

3. Given $n \in \mathbb{R}^n$, $\int_{G_r(x_0)} n \cdot (\nabla v_{x_0}^r - \nabla \hat{u})\, dx = 0$.

This property can be proved as follows:

$$\int_{G_r(x_0)} n \cdot (\nabla v_{x_0}^r - \nabla \hat{u})\, dx = \int_{G_r(x_0)} n \cdot (\nabla \tilde{u} - \nabla \hat{u})\, dx = \int_{B_r(x_0)} n \cdot (\nabla \tilde{u} - \nabla \hat{u})\, dx$$

$$= -\int_{B_r(x_0)} \mathbf{div}(n)(\tilde{u} - \hat{u})\, dx = 0.$$

The first two equalities can be respectively deduced from the definition (13) of $\tilde{u}$ and the fact that $\tilde{u} - \hat{u}$ has support in $G_r(x_0) \subset B_r(x_0) \subset \Omega$. The third equality can be inferred from the property (2) which implies that $\tilde{u} - \hat{u} \in W^{1,p}(\Omega)$, the fact that $\tilde{u} - \hat{u}$ is zero in $\partial B_r(x_0)$ and the formula for integration by parts. The fourth equality follows from $\mathbf{div}(n) = 0$.

20

4. $|v_{x_0}^r(x) - \hat{u}(x)| \leq 4r$, $x \in G_r(x_0)$.

It can be easily proved:

$$|v_{x_0}^r(x) - \hat{u}(x)| \leq \gamma(r + 2\|x - x_0\|_2) + |o(\|x - x_0\|_2)| \leq 4\gamma r \leq 4r \quad (18)$$

5. $|v_{x_0}^r(x) - \hat{u}(x)| \geq \frac{\gamma r}{4}$, $x \in B_{r/4}(x_0)$.

It can be inferred as follows:

$$\begin{aligned} v_{x_0}^r(x) - \hat{u}(x) &= \gamma(r - 2\|x - x_0\|_2) - o(\|x - x_0\|_2) \\ &\geq \gamma r - 2\gamma\|x - x_0\|_2 - \gamma\|x - x_0\|_2 \\ &= \gamma r - \frac{3}{4}\gamma r = \frac{\gamma r}{4}. \end{aligned}$$

6. $\|\nabla v_{x_0}^r(x) - \nabla\hat{u}(x_0)\|_2 = 2\gamma$, $x \in G_r(x_0)$. It follows from the computation of the derivative of $v_{x_0}^r(x)$.

7. $\nabla v_{x_0}^r(x) \in \mathcal{A}_i^\phi$, $\forall x \in G_r(x_0)$.

It is deduced from $\nabla\hat{u}(x_0) \in \mathcal{A}_i^\phi$, with $\mathcal{A}_i^\phi$ being an open set, and by using the property (6) with values of γ small enough.

The property (7) together with (P4) imply that $\phi^{**}(\nabla v_{x_0}^r(x)) = n_i \cdot \nabla v_{x_0}^r(x) + l_i$, $\forall x \in G_r(x_0)$. From this fact, the inequality $\phi^{**}(z) \geq n_i \cdot z + l_i$ $\forall z \in \Omega$, the property (3), and the definition (13) of $\tilde{u}$, we have the following set of inequalities:

$$\begin{aligned} \int_\Omega \phi^{**}(\nabla\hat{u})\,dx &= \int_{\Omega\backslash G_r(x_0)} \phi^{**}(\nabla\hat{u})\,dx + \int_{G_r(x_0)} \phi^{**}(\nabla\hat{u})\,dx \\ &\geq \int_{\Omega\backslash G_r(x_0)} \phi^{**}(\nabla\hat{u})\,dx + \int_{G_r(x_0)} n_i \cdot \nabla\hat{u} + l_i\,dx \\ &= \int_{\Omega\backslash G_r(x_0)} \phi^{**}(\nabla\hat{u})\,dx + \int_{G_r(x_0)} n_i \cdot \nabla v_{x_0}^r + l_i\,dx \\ &= \int_{\Omega\backslash G_r(x_0)} \phi^{**}(\nabla\hat{u})\,dx + \int_{G_r(x_0)} \phi^{**}(\nabla v_{x_0}^r)\,dx \\ &= \int_{\Omega\backslash G_r(x_0)} \phi^{**}(\nabla\tilde{u})\,dx + \int_{G_r(x_0)} \phi^{**}(\nabla\tilde{u})\,dx \\ &= \int_\Omega \phi^{**}(\nabla\tilde{u})\,dx. \end{aligned}$$

21

Finally, the next inequality is obtained:

$$\int_\Omega \phi^{**}(\nabla \tilde{u}) \, dx \le \int_\Omega \phi^{**}(\nabla \hat{u}) \, dx. \tag{19}$$

Therefore, (14) is proved.

Part 2: Proof of (15).

Now we will prove that the fitting term of (6) decreases (see (15)). Without loss of generality, we assume that $g(x_0) > \hat{u}(x_0)$.

From Federer [46], the fact that g is measurable implies that for almost every point in Ω the function g is approximately continuous. Let us assume that g is approximately continuous at x_0. Let us define D^ϵ and $G_r^\epsilon(x_0)$ as shown below

$$D^\epsilon = \{x \in \Omega : |g(x) - g(x_0)| \le \epsilon\},$$

$$G_r^\epsilon(x_0) = G_r(x_0) \cap D^\epsilon = \{x \in G_r(x_0) : |g(x) - g(x_0)| \le \epsilon\} \; \forall \epsilon > 0.$$

Then, the set D^ϵ has density value equal to one at x_0 for all $\epsilon > 0$ with respect to the sequence $\{G_r(x_0)\}_{r \in I}$ (due to g is approximately continuous at x_0), where $I = (0, s)$ for some positive number s, i.e.

$$\lim_{r \to 0, r \in I} \frac{|G_r^\epsilon(x_0)|}{|G_r(x_0)|} = 1. \tag{20}$$

We denote by C_r and D_r the two sides of (15), respectively, constrained to the region $G_r(x_0)$ where these two terms are different (due to $\hat{u} \ne \tilde{u}$, $\tilde{u} = v_{x_0}^r$ in $G_r(x_0)$):

$$C_r = \int_{G_r(x_0)} h(|\hat{u}(x) - g(x)|) \, dx, \quad D_r = \int_{G_r(x_0)} h(|v_{x_0}^r(x) - g(x)|) \, dx.$$

They can be decomposed as

$$C_r = C_r' + C_r'', \quad D_r = D_r' + D_r'',$$

where

$$C_r' = \int_{G_r^\epsilon(x_0)} h(|\hat{u}(x) - g(x)|) \, dx, \quad C_r'' = \int_{G_r(x_0) \backslash G_r^\epsilon(x_0)} h(|\hat{u}(x) - g(x)|) \, dx,$$

and

$$D_r' = \int_{G_r^\epsilon(x_0)} h(|v_{x_0}^r(x) - g(x)|) \, dx, \quad D_r'' = \int_{G_r(x_0) \backslash G_r^\epsilon(x_0)} h(|v_{x_0}^r(x) - g(x)|) \, dx.$$

22

The proof of (15) is based on showing that the difference between C_r and D_r is strictly positive when r is small enough. To achieve this goal, this difference is conveniently rearranged as follows:

$$C_r - D_r = (C'_r - D'_r) + (C''_r - D''_r).$$

The idea is to show that the first term $(C'_r - D'_r)$ is positive and decreases slower than the absolute value of the second term $(C''_r - D''_r)$ when r approaches to zero. Then, their sum will be positive for r small enough, and (15) will be satisfied.

Firstly, a lower bound for $C'_r - D'_r$ will be found. To this end, a number of facts are going to be introduced. Let us take an ϵ, r and some fixed r_0 small enough such that,

$$0 < \epsilon < g(x_0) - \hat{u}(x_0), \tag{21}$$

$$0 < r \leq r_0, \tag{22}$$

$$\sup_{x \in G_r^\epsilon(x_0)} v_{x_0}^r(x) < g(x_0) - \epsilon, \tag{23}$$

$$g(x_0) - \frac{\epsilon}{2} < g(x), \forall x \in G_{r_0}^\epsilon(x_0). \tag{24}$$

Then, we have

$$\hat{u}(x) \leq v_{x_0}^r(x) < g(x_0) - \epsilon < g(x), \ \forall x \in G_r^\epsilon(x_0). \tag{25}$$

Consequently,

$$|\hat{u}(x) - g(x)| \geq |v_{x_0}^r(x) - g(x)|, \ \forall x \in G_r^\epsilon(x_0). \tag{26}$$

In addition, given the set $G_{r_0}^\epsilon(x_0)$, the following numbers are defined:

$$b_0 = \max_{x \in G_{r_0}^\epsilon(x_0)} |\hat{u}(x) - g(x)|, \quad a_0 = \min_{x \in G_{r_0}^\epsilon(x_0)} |v_{x_0}^r(x) - g(x)|.$$

These two numbers can be bounded as $b_0 < +\infty$ and $a_0 > \epsilon/2$. The first bound can be straightforwardly inferred from the boundedness of $\hat{u}(x)$ and $g(x)$. The second one follows from (24) and (25).

Let us define now two other numbers:

$$b = |\hat{u}(x) - g(x)|, \quad a = |v_{x_0}^r(x) - g(x)|.$$

23

From the definition of a, b, a_0 and b_0, the following interval inclusion is obtained:

$$[a, b] \subset [a_0, b_0].$$

Then, taking into account (F5), the inequality

$$w_{a,b} \geq w_{a_0,b_0} > 0 \tag{27}$$

is established.

Besides, since h is continuous and increasing (see (F4)), the following relation holds [9]:

$$h(d) - h(c) \geq \int_c^d h'(t) \, dt, \quad d \geq c. \tag{28}$$

Using (27) and (28) with $d = b$ and $c = a$, a lower estimate is obtained:

$$h(b) - h(a) \geq \int_a^b h'(t) \, dt \geq w_{a,b}(b - a) \geq w_{a_0,b_0}(b - a). \tag{29}$$

Then, by substituting a and b in (29) by their expressions, we get the following inequality:

$$h(|\hat{u}(x) - g(x)|) - h(|v_{x_0}^r(x) - g(x)|) \geq w_{a_0,b_0}(|\hat{u}(x) - g(x)| - |v_{x_0}^r(x) - g(x)|). \tag{30}$$

Given the equality

$$C_r' - D_r' = \int_{G_r^\epsilon(x_0)} h(|\hat{u}(x) - g(x)|) - h(|v_{x_0}^r(x) - g(x)|),$$

and using the inequality (30) and the property (5), the following sequence of inequalities is obtained:

$$
\begin{aligned}
C_r' - D_r' &\geq w_{a_0,b_0} \int_{G_r^\epsilon(x_0)} |\hat{u}(x) - g(x)| - |v_{x_0}^r(x) - g(x)| \, dx \\
&= w_{a_0,b_0} \int_{G_r^\epsilon(x_0)} v_{x_0}^r(x) - \hat{u}(x) \, dx \\
&\geq w_{a_0,b_0} \int_{G_r^\epsilon(x_0) \cap B_{r/4}(x_0)} v_{x_0}^r(x) - \hat{u}(x) \, dx \\
&\geq w_{a_0,b_0} \gamma \frac{r}{4} |G_r^\epsilon(x_0) \cap B_{r/4}(x_0)| \\
&\geq \frac{w_{a_0,b_0} \gamma r |B_{r/4}(x_0)|}{4} \frac{|G_r^\epsilon(x_0) \cap B_{r/4}(x_0)|}{|B_{r/4}(x_0)|} > 0
\end{aligned}
$$

where

$$G_r^\epsilon(x_0) \cap B_{r/4}(x_0) = \{x \in B_{r/4}(x_0) : |g(x) - g(x_0)| \le \epsilon\}. \tag{31}$$

From the limit stated in (20) and the property (1), the following limit expression is obtained:

$$\lim_{r \to 0} \frac{|\{x \in B_{r/4}(x_0) : |g(x) - g(x_0)| \le \epsilon\}|}{|B_{r/4}(x_0)|} = 1. \tag{32}$$

Finally, the targeted lower bound is deduced by using (32):

$$C_r' - D_r' \ge O(r^{n+1}). \tag{33}$$

Therefore, we conclude that $C_r' - D_r'$ converges to zero no faster than $O(r^{n+1})$.

Next, we will find an upper bound for $|C_r'' - D_r''|$. From the definitions of C_r'' and D_r'', we have

$$C_r'' - D_r'' = \int_{G_r(x_0) \backslash G_r^\epsilon(x_0)} h(|\hat{u}(x) - g(x)|)\, dx - \int_{G_r(x_0) \backslash G_r^\epsilon(x_0)} h(|v_{x_0}^r(x) - g(x)|)\, dx. \tag{34}$$

In order to bound (34), an additional property of h needs to be considered. The fitting function h is Lipschitz on the interval $I = [0, 2\|g\|_{L^\infty(\Omega)} + 4]$ with a Lipschitz constant M. This can be inferred from these two points:

- According to property (F3), h is locally Lipschitz.

- $|g(x) - v_{x_0}^r(x)|, |g(x) - \hat{u}(x)| \in [0, 2\|g\|_\infty + 4]$ for a.e. $x \in G_r(x_0)$, which is derived from $\|\hat{u}\|_{L^\infty(\Omega)} \le \|g\|_{L^\infty(\Omega)}$ (Lemma 4.2). The number 4 comes from adding and subtracting $\hat{u}(x)$ in $|g(x) - v_{x_0}^r(x)|$, from applying the subadditivity property of the modulus function, and finally from using the property (4).

The fact that the function h is M-Lipschitz on the interval $I = [0, 2\|g\|_{L^\infty(\Omega)} + 4]$ and the property (4) imply the following inequalities, where B denotes

$G_r(x_0) \setminus G_r^\epsilon(x_0)$,

$$
\begin{aligned}
|C_r'' - D_r''| &\leq \int_B M \big| |\hat{u}(x) - g(x)| - |v_{x_0}^r(x) - g(x)| \big| \, dx \\
&\leq \int_B M \big| \hat{u}(x) - v_{x_0}^r(x) \big| \, dx \\
&\leq M \, 4r \, |G_r(x_0) \setminus G_r^\epsilon(x_0)| \\
&\leq M \, 4r \, |B_{r/4}(x_0)| \, \frac{|G_r(x_0) \setminus G_r^\epsilon(x_0)|}{|B_{r/4}(x_0)|}.
\end{aligned}
\tag{35}
$$

By similar arguments as those exposed to deduce (32), we conclude that

$$
\lim_{r \to 0} \frac{|G_r(x_0) \setminus G_r^\epsilon(x_0)|}{|B_{r/4}(x_0)|} = 0.
\tag{36}
$$

Then, from (35) and (36) it is inferred that

$$
|C_r'' - D_r''| \leq o(r^{n+1}).
\tag{37}
$$

We conclude that $|C_r'' - D_r''|$ converges to zero faster than $o(r^{n+1})$. Now, by combining the lower and upper bounds (33) and (37) it is obtained, for r small enough,

$$
C_r - D_r = (C_r' - D_r') + (C_r'' - D_r'') > 0.
\tag{38}
$$

Consequently,

$$
\int_{G_r(x_0)} h(|\hat{u}(x) - g(x)|) \, dx > \int_{G_r(x_0)} h(|v_{x_0}^r(x) - g(x)|) \, dx.
\tag{39}
$$

Hence, integrating over the entire domain Ω, and using that $\tilde{u} = \hat{u}$ in $\Omega \setminus G_r(x_0)$, we get

$$
\int_\Omega h(|\hat{u}(x) - g(x)|) \, dx > \int_\Omega h(|\tilde{u}(x) - g(x)|) \, dx,
\tag{40}
$$

which corresponds to (15).

The above theorem can be considered the foundation of the result previously stated in Theorem 2.1, Section 2. It is reproduced below for ease of clarity, and its proof is provided here.

Theorem 2.1. *Let Ω be an open and bounded subset of $\mathbb{R}^n$, $n \geq 1$. Let $h : [0, +\infty) \to [0, +\infty)$ and $\phi : \mathbb{R}^n \to \mathbb{R}$ be the data-fidelity and potential functions, respectively, such that they satisfy the conditions (F1) - (F5) and (P1) - (P4), respectively. Let us assume that $g \in L^\infty(\Omega)$ satisfies the ϕ-compatibility condition (Definition 1). Then, any functional of the class $\mathcal{F}(\cdot)$ has a minimizer on $W^{1,p}(\Omega)$, $p > 1$.*

PROOF. Let us define

$$D = \{x \in \Omega : \nabla \hat{u}(x) \in \mathcal{A}^\phi\} = \bigcup_{i=1}^{\infty} E_i.$$

As regards the value of $\mathcal{L}^n\{D\}$, two possible alternatives need to be explored:

- The first alternative is a trivial case. If $\mathcal{L}^n\{D\} = 0$, then $\hat{u}$ is a minimizer of $(\mathcal{F})$ and the proof is completed.

- The second alternative corresponds to $\mathcal{L}^n\{D\} > 0$, which is a non-trivial case. By applying Theorem 4.3 we conclude that

$$\hat{u}(x) = g(x), \text{ for a.e. } x \in D.$$

  By using Lemma 3.1 on the functions $\hat{u} : \Omega \to \mathbb{R}$ and $g : \Omega \to \mathbb{R}$ we obtain that g is a.e. approximately differentiable on D and

$$\nabla \hat{u}(x) = \text{ap}\, Dg(x), \text{ for a.e. } x \in D.$$

  The previous equality and the fact that $\nabla \hat{u}(x) \in \mathcal{A}^\phi$ allow us to infer,

$$\text{ap}\, Dg(x) \in \mathcal{A}^\phi, \text{ a.e. } x \in D,$$

  which is a contradiction with the ϕ-compatible condition (Definition 1). Hence, the only possible alternative is the first one, which guarantees that $\hat{u}$ is a minimizer of $\mathcal{F}(\cdot)$.

In order to deal with functions having discontinuities such as PWC and PWS signals, the class of functionals $\mathcal{F}(\cdot)$ is extended to a more general class which is also a generalization of the Mumford and Shah functional, whose fidelity and regularization terms retain the properties (F1)-(F5) and (P1)-(P4), respectively.

5. Dealing with discontinuities via the Mumford-Shah functional

The proof of Theorem 4.3 requires the solutions to belong to $W^{1,p}(\Omega)$, $\Omega \subset \mathbb{R}^n$, $n \geq 1$, and $p > 1$, a space which does not include functions with discontinuities. However, numerical methods could yield satisfactory discontinuous solutions to approximate truly discontinuous target signals, or as discrete approximations to continuous sharp signals [47]. These solutions can be obtained by computationally efficient algorithms, and are generally perceived as good visual quality results.

Although these are valid solutions, it would be appropriate to extend the theoretical framework to accommodate them. To this end, the Mumford-Shah (MS) functional [23] can be generalized as follows:

$$\mathcal{M}(u, K) = \int_{\Omega \setminus K} h(|u - g|) \, dx + \int_{\Omega \setminus K} \phi(\nabla u) \, dx + \mathcal{H}^{n-1}(K \cap \Omega), \quad (41)$$

where K stands for a jump set that estimates the discontinuities of the target signal g, $\phi : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ satisfies the properties (P1) - (P4), $h : [0, +\infty) \to [0, \infty]$ satisfies the properties (F1) - (F5), and $\mathcal{H}^{n-1}$ denotes the $(n-1)$-dimensional Hausdorff measure. The goal is to determine a pair $(\tilde{u}, \tilde{K})$ that satisfies

$$(\tilde{u}, \tilde{K}) = \arg\min \mathcal{M}(u, K), \quad (42)$$

where $K \subset \overline{\Omega}$ is a closed set with $\mathcal{H}^{n-1}(K) < +\infty$ and $u \in W^{1,p}(\Omega \setminus K) \subset SBV(\Omega)$. Next, some functionals and a theorem introduced in related papers are revisited because they are necessary to demonstrate the existence of minimizers for the functional $\mathcal{M}$.

The MS functional is a particular case of (41), with $h(\cdot) = \phi(\cdot) = \|\cdot\|_2^2$. The existence of minimizer for this functional was proved by Mumford et al. [23], De Giorgi et al. [48], and Dal Maso et al. [49]. In particular, the proof given in [48] relies on minimizing a simpler functional, presented below, which was previously introduced by Ambrosio et al. [9].

$$\mathcal{G}(u) = \int_{\Omega} |u - g|^2 \, dx + \int_{\Omega} \|\nabla u\|_2^2 \, dx + \mathcal{H}^{n-1}(S_u), \quad (43)$$

where $u \in SBV(\Omega)$, $\nabla u \in L^2(\Omega)$, and S_u is the discontinuity set of u. The existence of minimizers of $\mathcal{G}$ is proved in [9] (by the Ambrosio's compactness

Theorem). Let $\tilde{u}$ be a minimizer of $\mathcal{G}$; then, in [48] is proved that $(\tilde{u}, \overline{S}_{\tilde{u}})$ is a minimizer of the MS functional.

Furthermore, Fonseca et al. [50] and Fusco et al. [51] defined the following generalization of the MS functional:

$$\mathcal{MC}(u, K) = \int_{\Omega \setminus K} |u - g|^q \, dx + \int_{\Omega \setminus K} \phi(\nabla u) \, dx + \mathcal{H}^{n-1}(K \cap \Omega), \, q \geq 1 \quad (44)$$

with $\phi(\cdot)$ being a convex function satisfying (P1), (P2), and the following two conditions:

- *Convexity at infinity*

$$\int_Q \phi(z + \nabla \eta) \, dx \geq \int_Q [\phi(z) + \nu(\mu^2 + |z|^2 + |\nabla \eta|^2)^{\frac{p-2}{2}} |\nabla \eta|^2] \, dx, \quad (M1)$$

where $v > 1$, $\mu \geq 0$, $\forall \eta \in C_0^1(Q)$, and $Q \subset \mathbb{R}^N$ is the unit cube.

- Given $0 < t_0 < t$, $0 < m < p$, and $\forall z \in S^{n-1}$, the p-recession function of ϕ, $\phi_p(z) = \limsup_{t \to +\infty} \frac{\phi(tz)}{t^p}$, satisfies

$$\left| \phi_p(z) - \frac{\phi(tz)}{t^p} \right| \leq \frac{c_o}{t^m}, \quad (M2)$$

The proof of existence of minimizer for the function $\mathcal{MC}$ relies on its equivalence with the following functional

$$\mathcal{GC}(u) = \int_\Omega |u - g|^q \, dx + \int_\Omega \phi(\nabla u) \, dx + \mathcal{H}^{n-1}(S_u), \, q \geq 1 \quad (45)$$

which is well known that has minimizers according to the Ambrosio's compactness Theorem as shown in Fonseca et al. [50].

With regards to the minimization of the nonconvex functionals of the class $\mathcal{M}$, the corresponding convexified functional $\mathcal{M}^{**}$ is introduced below:

$$\mathcal{M}^{**}(u, K) = \int_{\Omega \setminus K} h(|u - g|) \, dx + \int_{\Omega \setminus K} \phi^{**}(\nabla u) \, dx + \mathcal{H}^{n-1}(K \cap \Omega), \quad (46)$$

where ϕ^{**} is the convex envelope of ϕ, and $h : \mathbb{R} \to [0, \infty]$ is as in (41). The functional $\mathcal{M}^{**}$, which is also a generalization of the MS functional, is intended to facilitate the existence of minimum of $\mathcal{M}$.

Following the strategy used to prove the equivalence of (44) and (45) in the Theorem 3.5 in [51], the next functional is proposed to establish an equivalence with $\mathcal{M}^{**}$:

$$\mathcal{G}^{**}(u) = \int_\Omega h(|u - g|) \, dx + \int_\Omega \phi^{**}(\nabla u) \, dx + \mathcal{H}^{n-1}(S_u). \qquad (47)$$

The equivalence of (46) and (47) is formulated as in the previous equivalences: if $\tilde{u}$ is a minimizer of $\mathcal{G}^{**}$, then $(\tilde{u}, \overline{S}_u)$ is a minimizer of $\mathcal{M}^{**}$. The proof is similar to the Theorem 3.5 in [51], considering the two facts below:

1. If $\phi(\cdot)$ satisfies the conditions (P1), (P2), (M1), and (M2), required by the Theorem 3.5 in [51], then so does $\phi^{**}(\cdot)$.
2. The proof of the Theorem 3.5 in [51] does not depend on the nature of the fidelity function common to both functionals.

The existence of minimizers for functionals in $\mathcal{M}^{**}$ together with Theorem 2.1 stated in Sec. 4, are the key premises for proving the existence of minimizer for functionals in the class $\mathcal{M}$ in the next section.

5.1. Existence of Solution for the class of functionals $\mathcal{M}$

Theorem 5.1, which is formulated below, establishes the existence of minimizer for functionals in the class $\mathcal{M}$. This result follows from the Theorem 2.1 stated in Sec. 4, the Ambrosio's compactness Theorem, and the equivalence between functionals (46) and (47), provided that ϕ satisfies (F1)-(F5) and h satisfies the conditions (P1)-(P4) and (M1)-(M2).

Theorem 5.1. *Let Ω be an open and bounded subset of $\mathbb{R}^n$, $n \geq 1$. Let $h : [0, +\infty) \to [0, +\infty)$ be the data-fidelity function such that it satisfies the conditions (F1)-(F5), and let $\phi : \mathbb{R}^n \to \mathbb{R}$ be the potential function such that it satisfies the conditions (P1)-(P4) and (M1)-(M2). Let us assume that $g \in L^\infty(\Omega)$ satisfies the ϕ-compatibility condition (Definition 1). Then, any functional of the class $\mathcal{M}$ has a minimizer $(\tilde{u}, \tilde{K})$ where $u \in W^{1,p}(\Omega \setminus \tilde{K})$ and $\tilde{K}$ is a compact set with $\mathcal{H}^{n-1}(\tilde{K}) < \infty$.*

PROOF. Let $(\tilde{u}, \tilde{K})$ be a minimizer of some functional of the form $\mathcal{M}^{**}$. Then, $\mathcal{H}^{n-1}(\tilde{K}) < +\infty$ and $\tilde{u}$ is a minimizer of the corresponding functional

$$\mathcal{M}_{\tilde{K}}^{**}(u) = \int_{\Omega \setminus \tilde{K}} h(|u - g|) \, dx + \int_{\Omega \setminus \tilde{K}} \phi^{**}(\nabla u) \, dx,$$

in the space $W^{1,p}(\Omega \setminus \tilde{K})$. Since ap $Dg(x) \notin \mathcal{A}^\phi$ for almost everywhere $x \in D_g$ (ϕ-compatible condition), then by the Theorem 2.1 we obtain

$$\mathcal{M}_{\tilde{K}}^{**}(\tilde{u}) = \mathcal{M}_{\tilde{K}}(\tilde{u}) = \int_{\Omega \setminus \tilde{K}} h(|\tilde{u} - g|)\, dx + \int_{\Omega \setminus \tilde{K}} \phi(\nabla \tilde{u})\, dx.$$

Consequently,

$$\mathcal{M}^{**}(\tilde{u}, \tilde{K}) = \mathcal{M}(\tilde{u}, \tilde{K}). \tag{48}$$

In what follows, the proof focuses on proving that $(\tilde{u}, \tilde{K})$ is also a minimizer of $\mathcal{M}$. Since $\phi^{**} \leq \phi$, the next inequality holds

$$\mathcal{M}^{**}(u, K) \leq \mathcal{M}(u, K), \tag{49}$$

for all $u \in W^{1,p}(\Omega \setminus K)$ and $K \subset \Omega$ compact. Then, given (48) and (49), it is easy to see that

$$\mathcal{M}(\tilde{u}, \tilde{K}) = \mathcal{M}^{**}(\tilde{u}, \tilde{K}) \leq \mathcal{M}^{**}(u, K) \leq \mathcal{M}(u, K),$$

for all $u \in W^{1,p}(\Omega \setminus K)$ and $K \subset \Omega$ a closed set with $\mathcal{H}^{n-1}(\Omega) < +\infty$. Thus, the Theorem holds.

6. Conclusions

This paper defines a general class of denoising functionals, and presents a theoretical proof of the existence of minimum provided that the noisy signal satisfies a property denoted by ϕ-compatible condition. There is a wide range of signals that satisfy this condition, some of them with clear practical interest such as the family of PWC and PWS signals which can arise in a number of application domains. These functionals integrate two general terms based on M-estimators from robust statistics for fitting and smoothing purposes, respectively. Although this class includes both convex and non-convex functionals, the main contribution of this work is to have established this theoretical result, for the first time, on a broad subclass of non-convex models. There is much empirical evidence supporting that this type of functionals produces sharper results in signal reconstruction as compared to convex ones.

The above contributions suggest a substantial potential for solving complex denoising problems, when considering unexplored functionals that fit the class defined in this paper and the problem complexity .

Promising directions for future research can be outlined as follows:

- The functionals of this class are able to restore signals affected by additive noises. Thus, a natural extension of this work is to deal with other kinds of noises such as those which are multiplicative in nature (Poisson, speckle, etc.).

- The guarantee of existence of a minimum for this class of non-convex problems enables the use of numerical methods to solve them. However, a theoretical proof on the convergence of numerical methods to the minimum is still open.

- The possibility of proving the existence of minimum for less constrained functionals would allow to denoise, with theoretical guarantees, signals of arbitrary dimension with practical interest.

- Besides the PWC and PWS signals, it would be interesting to find out what other classes of signals with practical applications satisfy the ϕ-compatible condition.

Acknowledgements

References

[1] G. Aubert, P. Komprost, Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations (Applied Mathematical Sciences), 2nd Edition, Springer Verlag, 2006.

[2] T. Chan, J. Shen, Image Processing and Analysis - Variational, PDE, wavelet, and Stochastic Methods, SIAM publisher, 2005.

[3] S. Geman, E. D. McClure, Bayesian image analysis. an application to single photon emission tomography, Proceedings of the American Statistical Association 19 (1985) 12–18.

[4] D. Geman, G. Reynolds, Constraint restoration and recovery of discontinuities, IEEE Trans. Pattern Anal. Mach. Intell. 14 (3) (1992) 367–383.

[5] M. Nikolova, M. Ng, C. Tam, Fast nonconvex nonsmooth minimization methods for image restoration and reconstruction, IEEE Trans. on Image Processing 19 (12) (2010) 3073–3088.

[6] M. Chipot, R. March, M. Rosati, G. V. Caffarelli, Analysis of a nonconvex problem related to selective signal smoothing, Math. Models Methods Appl. Sci. 7 (3) (1997) 313–328.

[7] R. March, M. Rosati, A. Schiaffino, BV solutions of the one-dimensional perona malik equation, http://www.iac.rm.cnr.it/ march/perona-malik-13nov.pdf (2007).

[8] P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, IEEE Trans. on Pattern Analysis and Machine Intelligence 12 (7) (1990) 629–639.

[9] L. Ambrosio, N. Fusco, D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems, Oxford Science Publication, 2000.

[10] M. Nikolova, Analysis of the recovery of edges in images and signals by minimizing nonconvex regularized least-squares, SIAM Journal of Multiscale Modeling and Simulation 4 (3) (2005) 960–991.

[11] P. Harjulehto, P. Lavala, O. Toivanen, A variant of the geman-mcclure model for image restoration, J. Math. Anal. Appl. 399 (2013) 676–681.

[12] P. J. Huber, E. M. Ronchetti, Robust Statistics, 2nd Edition, Wiley, 2009.

[13] T. Rabie, Image deblurring in the presence of impulsive noise, IEEE Transaction on Image Processing 14 (11) (2005) 1755–1765.

[14] A. Hamza, H. Krim, Image denoising: A nonlinear robust statistical approach, IEEE Trans. Signal Process 49 (12) (2001) 3045–3054.

[15] L. Bar, N. Sochen, N. Kiryati, Image deblurring in the presence of impulsive noise, International Journal of Computer Vision 70 (3) (2006) 279–298.

[16] M. Black, G. Sapiro, D. Marimont, D. Heeger, Robust anisotropic diffusion, IEEE Trans. on Image Processing 7 (3) (1998) 421–432.

[17] M. Nikolova, M. Ng, C. W.-K. Zhang, S., Efficient reconstruction of piecewise constant images using nonsmooth nonconvex minimization, SIAM Journal of Imaging Science 1 (1) (2008) 2–25.

[18] A. Selesnick, A. Parekh, I. Bayram, Convex 1-d total variation denoising with non-convex regularization, IEEE Signal Processing Letter 22 (2) (2015) 141–144.

[19] L. Vese, Problèmes variationnels et EDP pour l'analyse d'images et l'évolution de courbes, Ph.D. thesis, Université de Nice Sophia-Antipolos (1996).

[20] P. Charbonnier, L. Blanc-Feraud, G. Aubert, M. Barlaud, Deterministic edge-preserving regularization in computed imaging, IEEE Trans. on Image Processing 6 (2) (1997) 298–311.

[21] G. Aubert, L. Blanc-Feraud, R. March, γ-convergence of discrete functionals with nonconvex perturbation for image, SIAM Journal of Numerical Analysis 42 (3) (2004) 1128–1145.

[22] L. Vese, T. Chan, Reduced non-convex functional approximations for image restoration and segmentation, Tech. rep., UCLA, computational and Applied Mathematics Reports (1997).

[23] D. Mumford, J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems, Comm. Pure Appl. Math XLII (1989) 577–685.

[24] A. Drobyshev, C. Machka, M. Horsch, M. Seltmann, V. Liebscher, H. de Angelis, J. Beckers, Specificity assessment from fractionation experiments (safe): a novel method to evaluate microarray probe specificity based on hybridisation stringencies, Nucleic acids research 31 (2) (2003) E1–1.

[25] Y. Sowa, D. Rowe, C. Leake, T. Yakushi, M. Homma, A. Ishijima, M. Berry, Direct observation of steps in rotation of the bacterial flagellar motor, Nature 437 (6) (2005) 916–919.

[26] G. Winkler, O. Wittich, V. Liebscher, A. Kempe, Dont shed tears over breaks, Jahresbericht DMV 107 (2005) 5787.

[27] F. D. Blasi, G. Panigiani, On the dirichlet problem for first order partial differential equations. a baire category approach, Nonlinear Differential Equation and Applications NoDEA 6 (1999) 13–34.

[28] B. Dacorogna, P. Marcellini, Implicit Partial Differential Equations, Birkhauser, 1999.

[29] I. Fonseca, N. Fusco, P. Marcellini, An existence result for nonconvex variational problem via regularity, ESAIM: Control, Optimization and Calculus of Variations 7 (2002) 69–95.

[30] M. Sychev, Characterization of homogeneous scalar variational problems solvable for all boundary data, Proceeding of the Royal Society of Edinburgh: Section A Mathematics 130 (3) (2000) 611–631.

[31] S. Zagatti, Minimization of functional of the gradient by baire's theorem, SIAM Journal on Control and Optimization 38 (2) (2000) 384–399.

[32] L. Evans, R. Gariepy, Measure Theory and Fine Properties of Functions, CRC Press, 1992.

[33] A. Cellina, On minima of a functional of the gradient: necessary condition, Nonlinear Analysis: Theory, Methods and Applications 20 (4) (1993) 337–341.

[34] A. Cellina, On minima of a functional of the gradient: sufficient condition, Nonlinear Analysis: Theory, Methods and Applications 20 (4) (1993) 343–347.

[35] B. Dacorogna, P. Marcellini, General existence theorems for hamilton-jacobi equations in the scalar and vectorial case, Acta Math. 178 (1997) 1–37.

[36] Berman, Gaussian processes with stationary increments: Local times and sample function properties, Annals of Mathematical Statistics 41 (1970) 1173–1396.

[37] M. Rivera, J. Marroquin, Efficient half-quadric regularization with granularity control, Journal of Image and Vision Computing 21 (4) (2003) 345–357.

[38] I. Ekeland, R. Temam, Convex Analysis and Variational Problems, SIAM, 1999.

[39] M. Giaquinta, G. Modica, J. Soucek, Cartesian Currents in the Calculus of Variations I, Springer Verlag, 1998.

[40] G. Buttazzo, M. Guiaquinta, S. Hildebrandt, One-dimensional Variational Problems An Introduction, Oxford University Press, 1998.

[41] A. Bressan, Lecture Notes on Sobolev Spaces, Unpublished Manuscript, 2012.

[42] M. Giaquinta, E. Giusti, On the regularity of the minima of variational integrals, Acta Math. 148 (1) (1982) 31–46.

[43] E. Giusti, Direct Methods in the Calculus of Variations, World Scientific, 1985.

[44] P. Celada, S. Perrotta, Minimizing non-convex multiple integrals: a density result, Proceeding of the Royal Society of Edinburgh 130A (2000) 721–741.

[45] P. Celada, G. Cupini, M. Guidorzi, Existence and regularity of minimizers of nonconvex integrals with p-q growth, ESAIM: Control, Optimization and Calculus of Variations 13 (2) (2007) 343–358.

[46] H. Federer, Geometric Measure Theory, Springer Verlag, 1969.

[47] M. Nikolova, Minimizers of cost-functions involving nonsmooth data-fidelity terms application to the processing of outliers, SIAM Journal of Numerical Analysis 40 (3) (2002) 965–994.

[48] E. D. Giorgi, Minimization of functional of the gradient by baire's theorem, SIAM Journal on Control and Optimization 38 (2) (2000) 384–399.

[49] G. D. Maso, J. M. Morel, S. Solomini, A variational method in image segmentation: Existence and approximation results, Acta Mathematica 168 (1) (1992) 384–399.

[50] I. Fonseca, N. Fusco, Regularity results for anistropic image segmentation model, Analli de la squolla Normale Superiore di Pissa 23 (4) (1997) 463–499.

[51] N. Fusco, G. Mingione, C. Trombetti, Regularity of minimizers for a class of anisotropic free discontinuity problems, Journal of Convex Analysis 8 (2) (2001) 349–367.