A Few Remarks on Variational Models for Denoising

Rustum Choksi∗, Irene Fonseca†, and Barbara Zwicknagl‡

November 13, 2012

Abstract

Variational models for image and signal denoising are based on the minimization of energy functionals consisting of a fidelity term together with higher-order regularization. In addition to the choices of function spaces to measure fidelity and impose regularization, different scaling exponents appear. In this note we present a few simple, yet novel, remarks on (i) the stability with respect to deterministic noise perturbations, captured via oscillatory sequences converging weakly to zero, and (ii) exact reconstruction.

Key words: denoising, higher-order regularization, stability, weak convergence, Brezis-Lieb condition.

AMS subject classifications: 49N45, 94A08.

1 Introduction

In recent years, a wealth of research has been dedicated to the study of higher-order regularized variational methods for image and signal denoising. The idea of using a higher-order regularization to overcome ill-posedness in variational methods associated with ill-posed inverse problems is well-known, and it dates back to Andrey Tikhonov (cf. [27]). Since the pioneering work of Rudin, Osher and Fatemi [25], much attention has been devoted to $TV$-regularization in the context of image denoising. Analytical work has mainly targeted the study of fast and efficient numerical implementations and comparison with other models (see [9, 12, 14, 20, 21, 23] and the references therein), the understanding of the structure of possible minimizers (see [1, 2, 3, 4, 6, 10, 11, 13, 15, 16, 18, 22, 24, 28]), and the analysis of the effects of $TV$-regularization in terms of edge, block, texture, and scale preservation (see [17, 21, 22, 26]).

Here we will consider the following class of higher-order regularized variational models for denoising. Given a noisy measured signal $f : \Omega \to \mathbb{R}$ in a domain $\Omega \subset \mathbb{R}^d$, with

$$f = f_0 + h$$

where $f_0$ is the undisturbed signal and $h$ is a noise sample, we seek to reconstruct the original signal $f_0$, or at least a good approximation of it, from the measured signal $f$. A standard technique for such problems is to minimize functionals of the form

$$J(u) := |u|^\alpha_X + \lambda \|u - f\|_Y^\beta, \quad u : \Omega \to \mathbb{R},$$

where $\alpha$, $\beta$, and $\lambda$ are constants. The choice of $\alpha$ and $\beta$ depends on the type of regularity one wishes to impose on the solution. For example, $\alpha = 1$ and $\beta = 2$ correspond to the standard $TV$-regularization, while $\alpha = 2$ and $\beta = 2$ correspond to the total second-order variation (TSO)-regularization. The parameter $\lambda$ controls the trade-off between the fidelity term and the regularization term.

In this note, we will present a few simple, yet novel, remarks on (i) the stability with respect to deterministic noise perturbations, captured via oscillatory sequences converging weakly to zero, and (ii) exact reconstruction.
with $X$ and $Y$ Banach function spaces such that, typically, $X \subset Y$, and $\alpha, \beta \geq 1$. The space $Y$ is often a Lebesgue space $L^p(\Omega)$ for some $1 \leq p < \infty$, but it may also be a normed space with a weaker norm such as the $H^{-1}$ or the $G$-norm of Meyer (see [22]). These weaker norms have proven to be successful when one wishes to retain textures in images (see [21, 22, 23]).

The space $Y$ is often a Lebesgue space $L^p(\Omega)$ for some $1 \leq p < \infty$, but it may also be a normed space with a weaker norm such as the $H^{-1}$ or the $G$-norm of Meyer (see [22]). These weaker norms have proven to be successful when one wishes to retain textures in images (see [21, 22, 23]). The most popular choices for the function space $X$ include the space $BV(\Omega)$ of functions of bounded variation, and Sobolev spaces $W^{2,k}(\Omega)$ for some $k \geq 1$. Having these examples in mind, we assume that $X$ is equipped with a norm $\| \cdot \|_X$ and a semi-norm $| \cdot |_X$, where the latter is sequentially lower semicontinuous with respect to an appropriate weak convergence in $X$, and such that for some $C > 0$

$$\|u\|_X \leq C (|u|_X + \|u\|_Y) \quad \text{for all } u \in X.$$  

(3)

The second term in (2) is called the fidelity term and it encourages minimizers to stay close to the observed signal. The first term is intended to be a regularization which should filter out (oscillatory) noisy parts of the observed image. The two terms are weighted by a regularization parameter $\lambda > 0$. The exponents $\alpha$ and $\beta$ are usually natural numbers. Typical examples of such functionals include the total variation based regularization in the Rudin-Osher-Fatemi (ROF) model [25] on two-dimensional bounded Lipschitz domains $\Omega \subset \mathbb{R}^2$, namely

$$J(u) := |u|_{BV} + \lambda \|u - f\|^2_{L^2} ,$$  

(4)

for $u \in L^1_{loc}(\Omega)$, with the $BV$-semi-norm$^1$ defined by (8), or variants with slightly different fidelity terms, such as

$$J(u) := |u|_{BV} + \lambda \|u - f\|_{L^2} , \text{ or } J(u) := |u|_{BV} + \lambda \|u - f\|_{L^1} .$$  

(5)

These functionals share a common feature: If we input for $f$ a highly oscillatory perturbation of a clean image $f_0$, minimization returns a signal which is both close to $f$ (the influence of the second term) in a way as to minimize oscillations (the influence of the first term). Not surprisingly, numerical simulations of minimizers for a noisy perturbation of clean images result in states closer to the actual image than the measured signal. In this way, we say that the original measured signal $f$ has been denoised. We refer to the book of Chan and Shen [14], and the references therein, for various examples and types of noise perturbations.

It is therefore natural to assess two properties of these variational problems. The first is their stability with respect to weakly vanishing noise. Precisely, if

$$J_n(u) := |u|_{X}^{\alpha} + \lambda \|u - f - h_n\|_{Y}^{\beta} \quad \text{and} \quad J(u) := |u|_{X}^{\alpha} + \lambda \|u - f\|_{Y}^{\beta} ,$$  

(6)

where the sequence of noise $\{h_n\}$ converges to zero in some weak sense, then the minimizers of the disturbed functionals $J_n$ should converge to minimizers of the undisturbed functional $J$. In fact, one would expect that $\{J_n\}$ Gamma converges to $J$ (cf. Definition 2.2). The fact that functionals of type (2) are stable with respect to small perturbations in the strong norm is easy to establish, and has been noted in [1]. However, the paradigm of these variational problems is that minimization of the regularization reduces highly oscillatory perturbations of the clean signal which one often attributes to noise. Thus it is natural to consider perturbations which do not go to zero in a strong sense but rather in a weak sense, for example increasing oscillations between two fixed values. Surprisingly, this notion of stability with respect to weak perturbations has rarely been address. In fact, we are aware of only [16].

$^1$Sometimes the regularization semi-norm $| \cdot |_{BV}$ is replaced by the full norm.
A second property pertains to the existence of a class of clean simple signals which remain invariant under energy minimization. As noted by Chan and Esedoğlu in [13] “It is in general desirable for image denoising algorithms to have a large class of “noise-free” images that they leave invariant.” Precisely, for \( f_0 \) in some class of signals, for \( \lambda \) sufficiently large, the model allows for the reconstruction of \( f_0 \), i.e., the unique minimizer of \( J \) with \( f = f_0 \) is \( u = f_0 \).

In this note, we present a few simple, yet novel, remarks on these two properties. The first set of remarks shows that a certain condition (which we refer to as a Brezis-Lieb-type condition) on the fidelity norm is a natural condition for insuring this notion of weak stability. In the second set, we remark that an elementary comparison principle yields some well-known results concerning loss of contrast for the original ROF and sufficiency relations based upon the existence of regularization-realizing vector fields.

2 Preliminaries

In the sequel we will use \( C \) to denote a constant that may vary from line to line, and from expression to expression. We let \( \Omega \subset \mathbb{R}^N \) be an open Lipschitz domain, and use the space of functions with bounded variation on \( \Omega \subset \mathbb{R}^N \) defined to be (see, for example, [5])

\[
BV(\Omega) := \left\{ u \in L^1(\Omega) : |u|_{BV(\Omega)} < \infty \right\}
\]

(7)

equipped with the norm \( \|u\|_{BV(\Omega)} := \|u\|_{L^1(\Omega)} + |u|_{BV(\Omega)} \), where

\[
|u|_{BV(\Omega)} := \sup \left\{ \int_{\Omega} u(x) \Div \phi(x) \, dx : \phi \in C_c^1(\Omega; \mathbb{R}^N), \|\phi\|_{L^\infty(\Omega)} \leq 1 \right\}.
\]

(8)

We call \( \psi \in C_c^1(\Omega; \mathbb{R}^N) \) a TV-realizing vector-field for \( u \in BV(\Omega) \) if \( \|\psi\|_{L^\infty(\Omega)} \leq 1 \) and

\[
|u|_{BV(\Omega)} = -\int_{\Omega} u(x) \Div \psi(x) \, dx.
\]

A sequence \( \{u_n\} \) in \( BV(\Omega) \) is said to converge weakly* to \( u \in BV(\Omega) \) if \( u_n \to u \) in \( L^1(\Omega) \) and \( Du_n \rightharpoonup Du \) in \( \Omega \). Here \( Du_n \) and \( Du \) denote the distributional derivatives which are Radon measures, and \( \rightharpoonup \) stands for weak convergence of measures. It can be shown that every sequence \( \{u_n\} \) with \( \sup_{n \in \mathbb{N}} \|u_n\|_{BV(\Omega)} < \infty \) admits a subsequence \( \{u_{n_k}\} \) that weakly* converges to a function \( u \in BV(\Omega) \).

As in [22], we denote by \( G \) the dual space of the closure of \( BV \) in the Schwartz class (see [22, Section 1.14] for a detailed discussion of this space and the associated \( G \)-norm). It turns out that \( L^2(\Omega) \subset G \), and if \( f \in BV(\Omega) \) and \( g \in L^2(\Omega) \) then (see [22, Lemma 3])

\[
\left| \int_{\Omega} f(x) g(x) \, dx \right| \leq \|f\|_{BV} \|g\|_G.
\]

(9)

We will use the following result of Brezis and Lieb on weakly converging sequences (see Theorem 1 in [8]).

**Lemma 2.1** Let \( \{f_n\} \) be a sequence of measurable functions which are uniformly bounded in \( L^p(\Omega) \) for some \( 0 < p < \infty \). If \( f_n \to f \) pointwise almost everywhere, then

\[
\lim_{n \to \infty} \left\{ \|f_n\|_{L^p(\Omega)}^p - \|f_n - f\|_{L^p(\Omega)}^p \right\} = \|f\|_{L^p(\Omega)}^p.
\]
We recall the definition and fundamental property of $\Gamma$-convergence (see, for example, [7]).

**Definition 2.2** A sequence of functionals $F_n : X \to [−\infty, +\infty]$ defined on a metric space $X \Gamma$-converges to $F : X \to [−\infty, +\infty]$, and we write $F_n \Gamma \to F$, if 

(i) (Lower bound) for every converging sequence $\{x_n\}$, $x_n \to x$ in $X$, one has $F(x) \leq \lim \inf_{n \to \infty} F_n(x_n)$, and

(ii) (Upper bound) for every $x \in X$, there exists a sequence $\{x_n\}$ such that $x_n \to x$ and $\lim \sup_{n \to \infty} F_n(x_n) = F(x)$.

A consequence of Gamma convergence (cf. [7]) is: If $F_n \Gamma \to F$ and if $x_n$ is a minimizer of $F_n$ with $x_n \to x$, then $x$ is a minimizer of $F$.

(10)

## 3 Remarks on the Stability with respect to Weak Perturbations

We address the stability of functionals of type (2) with respect to small weak² perturbations, i.e., we consider a sequence of signals $\{f_n\}$ such that $f_n = f_0 + h_n$. We assume that $h_n \to 0$ weakly in $\mathcal{Y}$. As a consequence of the Principle of Uniform Boundedness, $\sup_n \|h_n\|_\mathcal{Y} \leq C$, and we may restrict ourselves to (sub)sequences for which $\lim_{n \to \infty} \|h_n\|_\mathcal{Y}$ exists. Our analysis with a weak norm is similar in spirit to results presented in [16] for a second-order problem (see [16, Theorem 4.5]).

### 3.1 The Brezis-Lieb Condition and Convergence of Minimizers

In this subsection we use the notation "weak convergence" and "$\mathcal{X}$" for the appropriate type of weak (respectively weak*) convergence in the Banach space $\mathcal{X}$ (see (H1) and its discussion below). Here we make the following assumptions:

(H1) For every sequence $\{u_n\} \subset \mathcal{X}$ with $\sup_n \|u_n\|_\mathcal{X} \leq C$ there exists a subsequence $\{u_{n_k}\}$ and $u \in \mathcal{X}$ such that $u_{n_k} \mathcal{X} \to u$ in $\mathcal{X}$. The semi-norm $\|\cdot\|_\mathcal{X}$ is sequentially weakly lower semicontinuous.

(H2) Rellich-type property: Whenever $u_n \mathcal{X} \to u$ in $\mathcal{X}$, then $u_n \to u$ strongly in $\mathcal{Y}$.

(H3) Brezis-Lieb type condition: The sequence of noise samples $\{h_n\}$ is such that $h_n \to 0$ in $\mathcal{Y}$, and for every $f \in \mathcal{Y}$

$$
\|f\|_\mathcal{Y}^\beta = \lim_{n \to \infty} \left( \|f - h_n\|_\mathcal{Y}^\beta - \|h_n\|_\mathcal{Y}^\beta \right).
$$

(11)

The hypothesis (H1) is satisfied for the standard weak convergence in $\mathcal{X}$ if $\mathcal{X}$ is a reflexive Banach space. The Rellich-type hypothesis (H2) requires a compact embedding of the regularization space into the fidelity space. If $\mathcal{X} = BV (\Omega)$, then we denote by $\mathcal{X}$ the convergence with respect to the weak* topology (see Section 2). We first give two examples in which the Brezis-Lieb-type condition is satisfied.

²The fact that functionals (2) are stable with respect to small perturbations in the strong norm has been noted in [1]. If a weak fidelity measure, such as Meyer’s G-norm or negative Besov-norms (see [21, 22, 23]), is used, then the strong condition $\|h_n\|_\mathcal{Y} \to 0$ still allows for oscillatory sequences which do not converge (even up to a subsequence) in any classical sense (see [22, Lemma 14]). See [19] for an interesting error estimate with respect to the $L^2$ norm in the context of bar code reconstruction.
Example 3.1  

(i) Let \((\mathcal{Y}, (\cdot, \cdot)_\mathcal{Y})\) be a Hilbert space, and let \(\beta = 2\). Then the Brezis-Lieb type condition (H3) is satisfied for all sequences \(h_n \to 0\) in \(\mathcal{Y}\), since

\[
\lim_{n \to \infty} \left(\|f - h_n\|_Y^2 - \|h_n\|_Y^2\right) = \lim_{n \to \infty} \left(\|f\|_Y^2 + \|h_n\|_Y^2 - 2\langle f, h_n \rangle_\mathcal{Y} - \|h_n\|_Y^2\right) = \|f\|_Y^2 .
\]

(ii) Clearly, the Brezis-Lieb type condition is satisfied if \(\{h_n\}\) converges strongly. It allows, however, for some more general sequences, in particular for concentration phenomena. For \(1 \leq p < \infty\) let \(\{h_n\} \subset L^p(\Omega)\) converge weakly and pointwise almost everywhere to 0. Then for all \(f \in L^p(\Omega)\), the sequence \(\{f - h_n\}\) converges weakly in \(L^p(\Omega)\) and pointwise almost everywhere to \(f\). Applying the Brezis-Lieb Lemma (see Lemma 2.1, [8]) to \(f_n := f - h_n\), we observe that

\[
\|f\|_{L^p(\Omega)}^p = \lim_{n \to \infty} \left\{\|f_n\|_{L^p(\Omega)}^p - \|f - f_n\|_{L^p(\Omega)}^p\right\} = \lim_{n \to \infty} \left\{\|f - h_n\|_{L^p(\Omega)}^p - \|h_n\|_{L^p(\Omega)}^p\right\},
\]

and hence (H3) is satisfied. A typical example of such a sequence \(\{h_n\}\) is

\[
h_n(x) := \begin{cases} 
  n - n^2 x & 0 \leq x < \frac{1}{n} , \\
  0 & \frac{1}{n} \leq x \leq 1 .
\end{cases}
\]

In this subsection, define for \(n \in \mathbb{N}\)

\[
J_n(u) := |u|_X^n + \lambda \|u - f - h_n\|_Y^n \quad \text{and} \quad J(u) := |u|_X^n + \lambda \|u - f\|_Y^n .
\]

We note that these hypotheses immediately lead to the following \(\Gamma\)-convergence result.

Theorem 3.2  

(i) Compactness: If (H1) holds, then for every sequence \(\{u_n\} \subset X\) with 

\[
\sup_{n \in \mathbb{N}} J_n(u_n) < \infty ,
\]

there exists a subsequence (not relabeled) such that \(u_n \rightharpoonup u\) in \(X\).

(ii) \(\Gamma\)-limit: If (H1), (H2), (H3) hold, then setting 

\[
\tilde{J}(u) := |u|_X^n + \lambda \|u - f\|_Y^n + \lambda \lim_{n \to \infty} \|h_n\|_Y^n ,
\]

we have \(J_n \rightharpoonup \tilde{J}\) with respect to the weak topology in \(X\).

Proof: (i) By (3) we deduce that

\[
\|u_n\|_X \leq C \left(\|u_n\|_X + \|u_n\|_\mathcal{Y}\right) \leq C \left(\|u_n\|_X + \|u_n - f - h_n\|_\mathcal{Y} + \|f + h_n\|_\mathcal{Y}\right) \leq C \left(1 + (J_n(u_n))^{1/\alpha} + ((J_n(u_n))^{1/\beta} + \|f + h_n\|_\mathcal{Y}\right) \leq C ,
\]

where the constants \(C\) depend on \(\lambda, f\) and \(\{h_n\}\). Thus, \(\sup_{n \in \mathbb{N}} \|u_n\|_X < \infty\) and compactness follows from (H1).

(ii) To establish the \(\Gamma\)-liminf of \(\{J_n\}\), we observe that by (H2) and lower semicontinuity, we have for every sequence \(\{u_n\}\) with \(u_n \rightharpoonup u\) in \(X\)

\[
\liminf_{n \to \infty} J_n(u_n) = \liminf_{n \to \infty} \left(\|u_n\|_X^n + \lambda \|u_n - f - h_n\|_Y^n \right) \geq |u|_X^n + \lambda \|u - f\|_Y^n + \lambda \lim_{n \to \infty} \|h_n\|_Y^n ,
\]

5
where we used (H3).

Concerning the limsup inequality, if \( u \in \mathcal{X} \), then setting \( u_n := u \) for all \( n \in \mathbb{N} \), it follows again from (H3) that

\[
\lim_{n \to \infty} J_n (u) = |u|^2 + \lambda \lim_{n \to \infty} \| u - f - h_n \|_{\mathcal{Y}}^2 = |u|^2 + \lambda \left( \| u - f \|_{\mathcal{Y}}^2 + \lim_{n \to \infty} \| h_n \|_{\mathcal{Y}}^2 \right),
\]

and thus the constant sequence \( \{ u \} \) is a recovery sequence.

By the fundamental property (10) of Gamma Convergence, Theorem 3.2 implies

**Corollary 3.3** If (H1), (H2) and (H3) hold, and if \( u_n \) is a minimizer of \( J_n \), then (up to a sub-sequence) \( u_n \xrightarrow{\mathcal{X}} u \) in \( \mathcal{X} \) and \( u \) is a minimizer of \( J \). Additionally, the same holds true if, in place of (H2), we assume that the sequence of minimizers \( \{ u_n \} \) has a subsequence that converges strongly in \( \mathcal{Y} \).

If we consider \( L^p \)-penalization in the non-Hilbert space case \( p \neq 2 \), the Brezis-Lieb type condition does not hold true for general weakly converging sequences. In this case, we cannot expect convergence of minimizers in general. We give an example for the widely used \( L^1 \)-penalization.

**Example 3.4** Let \( f = 0 \), and let a sequence of noise samples \( \{ h_n \} \subset L^1 (0, 1) \) be given by

\[
h_n (x) := h (nx) \quad \text{with} \quad h (x) := \begin{cases} 
2 & 0 \leq x \leq \frac{1}{3}, \\
-1 & \frac{1}{3} < x < 1,
\end{cases}
\]

where \( h \) is extended periodically to \( \mathbb{R} \) with period 1. Consider the associated functionals

\[
J (u) := |u|_{BV} + \lambda \| u \|_{L^1 (0, 1)} \quad \text{and} \quad J_n (u) := |u|_{BV} + \lambda \| u - h_n \|_{L^1 (0, 1)}.
\]

The Brezis-Lieb condition (H3) is not satisfied since, e.g., for the constant function \( 1 \in L^1 (0, 1) \)

\[
\| 1 \|_{L^1 (0, 1)} = 1 \neq \frac{1}{3} = \lim \left( \| 1 - h_n \|_{L^1 (0, 1)} - \| h_n \|_{L^1 (0, 1)} \right).
\]

Clearly \( u = 0 \) is the unique minimizer of \( J \). However, for any \( n \in \mathbb{N} \),

\[
J_n (0) = \lambda \| h_n \|_{L^1 (0, 1)} = \lambda \left( \frac{1}{3} \cdot 2 + \frac{2}{3} \cdot 1 \right) = \frac{4}{3} \lambda,
\]

and

\[
J_n (-1) = \lambda \| 1 + h_n \|_{L^1 (0, 1)} = \lambda \left( \frac{1}{3} \cdot 3 + \frac{2}{3} \cdot 0 \right) = \lambda < \frac{4}{3} \lambda = J_n (0),
\]

so that \( u = 0 \) is not a minimizer for any \( n \). Also, if \( u_n \) is minimizer of \( J_n \), then there is no subsequence \( \{ u_{n_k} \} \) with \( u_{n_k} \to 0 \) strongly in \( L^1 (0, 1) \), or else

\[
J_{n_k} (u_{n_k}) = |u_{n_k}|_{BV (0, 1)} + \lambda \| u_{n_k} - h_{n_k} \|_{L^1 (0, 1)} \\
\geq \lambda \left( \| h_{n_k} \|_{L^1 (0, 1)} - \| u_{n_k} \|_{L^1 (0, 1)} \right) \\
\geq \lambda \left( \frac{4}{3} - \| u_{n_k} - u \|_{L^1 (0, 1)} \right) \\
> \lambda
\]

for large enough \( n_k \in \mathbb{N} \), contradicting the minimality of \( u_{n_k} \).
3.2 Convergence of minimizers in the ROF model

Here we consider the classical ROF model, i.e.,

\[ J(u) := |u|_{BV} + \lambda \|u - f\|_{L^2(\Omega)}^2 \quad \text{and} \quad J_n(u) := |u|_{BV} + \lambda \|u - h_n\|_{L^2(\Omega)}^2. \]

Since \( J \) is strictly convex, there exists a unique minimizer of \( J \) in \( BV(\Omega) \). The results in Subsection 3.1 do not directly apply to the classical ROF since (H2) is not satisfied. Under natural additional assumptions\(^3\), however, we obtain convergence of minimizers also for ROF. We outline two natural conditions which ensure that sequences \( \{u_n^*\} \) of minimizers \( u_n^* \) of \( J_n \) admit subsequences that converge strongly in \( L^2(\Omega) \) to a minimizer of \( J \): In Proposition 3.6 we note that by restricting our attention to bounded perturbations, the analogous result to Theorem 3.2 carries over. Proposition 3.7 is based upon a more general condition involving the \( G \)-norm, and its proof follows directly from arguments of Meyer (see [22, Section 1.14]).

For the first condition we will need the following auxiliary truncation lemma.

**Lemma 3.5** Let \( \Omega \subset \mathbb{R}^2 \) and let \( f \in L^\infty(\Omega) \). Then the minimizer \( u^* \) of \( J \) satisfies \( u^* \in L^\infty(\Omega) \) and \( \|u^*\|_{L^\infty(\Omega)} \leq \|f\|_{L^\infty(\Omega)} \).

**Proof:** Given \( u \in BV(\Omega) \) with \( \|u\|_{L^\infty(\Omega)} > \|f\|_{L^\infty(\Omega)} \) (possibly \( \|u\|_{L^\infty(\Omega)} = \infty \)), define \( \tilde{u} \in L^\infty(\Omega) \) pointwise almost everywhere by

\[ \tilde{u}(x) := \begin{cases} -\|f\|_{L^\infty(\Omega)} & \text{if } u(x) \leq -\|f\|_{L^\infty(\Omega)}, \\ u(x) & \text{if } -\|f\|_{L^\infty(\Omega)} \leq u(x) \leq \|f\|_{L^\infty(\Omega)}, \\ \|f\|_{L^\infty(\Omega)} & \text{if } \|f\|_{L^\infty(\Omega)} \leq u(x). \end{cases} \]

Clearly \( \|\tilde{u}\|_{L^\infty(\Omega)} \leq \|f\|_{L^\infty(\Omega)} \) and \( \|\tilde{u} - f\|_{L^2(\Omega)} < \|u - f\|_{L^2(\Omega)} \). We claim that further \( |\tilde{u}|_{BV} \leq |u|_{BV} \), and hence \( J(\tilde{u}) < J(u) \). Indeed, let \( \phi_\epsilon \) be mollifiers, and for \( a > 0 \) consider

\[ \tau_a(t) := \begin{cases} -a, & \text{if } t \leq -a \\ t, & \text{if } -a \leq t \leq a \\ a, & \text{if } t \geq a. \end{cases} \]

Let \( U \subset \Omega \) be such that \( |Du| (\partial U) = 0 \). Then by lower semicontinuity of the total variation,

\[ |Du| (\Omega) \geq |Du| (U) = \lim_{\epsilon \to 0} \int_U |\nabla (\phi_\epsilon \ast u)| \geq \limsup_{\epsilon \to 0} \int_U |\nabla (\tau_a \ast (\phi_\epsilon \ast u))| \geq |D (\tau_a \circ u)| (U). \]

Let \( U \neq \Omega \) to obtain \( |Du| (\Omega) \geq |D (\tau_a \circ u)| (\Omega) \).

It follows that, if \( u^* \) is a minimizer of \( J \), then \( u^* \in L^\infty(\Omega) \) and \( \|u^*\|_{L^\infty(\Omega)} \leq \|f\|_{L^\infty(\Omega)}. \)

---

\(^3\)In general, we do not expect to find a strongly convergent subsequence. Consider \( f \equiv 0 \), \( \Omega = \mathbb{R}^2 \), and \( h_n = n \chi_{B_{\frac{1}{n}}(0)} \), where \( \chi_{B_{\frac{1}{n}}(0)} \) denotes the characteristic function of the ball with radius \( \frac{1}{n} \) and center \( 0 \in \mathbb{R}^2 \). One can readily check that for large \( \lambda \) the minimizer \( u_n^* \) of \( J_n \) is \( c \chi_{B_1} \), where \( c = O(1) \) as \( n \to \infty \). Clearly, the sequence \( \{u_n^*\} \) converges weakly in \( L^2(\mathbb{R}^2) \) to \( u \equiv 0 \) as \( n \) tends to infinity, but there is no strongly converging subsequence.
Proposition 3.6  If \( f \in L^\infty(\Omega) \) and if the sequence of noise samples \( \{h_n\} \) is uniformly bounded, i.e.,

\[
\sup_{n \in \mathbb{N}} \|h_n\|_{L^\infty(\Omega)} < +\infty,
\]

and converges weakly to 0 in \( L^2(\Omega) \), then minimizers of \( J_n \) converge to minimizers of \( J \) weakly* in \( BV(\Omega) \).

Proof: In view of Corollary 3.3 and Example 3.1, it remains to show that for a sequence of minimizers \( \{u_n\} \) of \( J_n \), there exists a subsequence that converges strongly to \( u \) in \( L^2(\Omega) \). By the compact embedding \( BV(\Omega) \subset L^1(\Omega) \), there exists a subsequence (not relabeled) such that \( u_n \rightharpoonup u \) strongly in \( L^1(\Omega) \) for some \( u \in BV(\Omega) \).

By Lemma 3.5, \( \sup_n \|u_n\|_{L^\infty(\Omega)} \leq \sup_n \|f + h_n\|_{L^\infty(\Omega)} < +\infty \), and thus \( u \in L^\infty(\Omega) \) and

\[
\|u_n - u\|_{L^2(\Omega)} \leq \|u_n - u\|_{L^\infty(\Omega)} \|u - u_n\|_{L^1(\Omega)} \to 0.
\]

Hence \( u_n \rightharpoonup u \) strongly in \( L^2(\Omega) \).

\[\Box\]

Proposition 3.7  Let \( f \in L^2(\Omega) \), and let \( \{h_n\} \) be a sequence of noise samples such that \( h_n \rightharpoonup 0 \) in \( L^2(\Omega) \) and \( \|h_n\|_G \to 0 \). Let \( u_n \in BV(\Omega) \) be a minimizer of \( J_n \). Then there exists a subsequence (not relabeled) \( \{u_n\} \) and \( u \in BV(\Omega) \) such that \( u_n \rightharpoonup u \) in \( BV(\Omega) \), and \( u \) is minimizer of \( J \).

Proof: By Theorem 3.2 (i), there exists a subsequence (not relabeled) such that \( u_n \rightharpoonup u \) for some \( u \in BV(\Omega) \). For all \( v \in BV(\Omega) \), we have \( J_n(u_n) \leq J_n(v) \), i.e.,

\[
|u_n|_{BV(\Omega)} + \lambda \|u_n - f\|^2_{L^2} - 2\lambda (u_n, h_n)_{L^2} \leq |v|_{BV(\Omega)} + \lambda \|v - f\|^2_{L^2} - 2\lambda (v, h_n)_{L^2}.
\]

By (9), we have \( (u_n, h_n)_{L^2(\Omega)} \leq \|u_n\|_{BV(\Omega)} \|h_n\|_G \), hence

\[
|u_n|_{BV(\Omega)} (1 - 2\lambda \|h_n\|_G) - 2\lambda \|u_n\|_{L^1} \|h_n\|_G + \lambda \|u_n - f\|^2_{L^2} \leq |v|_{BV(\Omega)} + \lambda \|v - f\|^2_{L^2} - 2\lambda (v, h_n)_{L^2}.
\]

Since \( h_n \rightharpoonup 0 \) in \( L^2(\Omega) \), \( \|h_n\|_G \to 0 \) and \( \|u_n\|_{L^1} \to \|u\|_{L^1} \), using a lower semicontinuity argument, we obtain

\[
|u|_{BV(\Omega)} + \lambda \|u - f\|^2_{L^2(\Omega)} \leq |v|_{BV(\Omega)} + \lambda \|v - f\|^2_{L^2(\Omega)},
\]

and we conclude that \( u \) is a minimizer of \( J \).

We note that the stability properties presented in this section do not depend on the scaling exponent \( \alpha > 1 \) of the regularization term. We proved convergence of minimizers of the distorted functionals \( J_n \) to minimizers of \( J \), although in general a minimizer of \( J \) is not a minimizer of \( J_n \) for some \( n \in \mathbb{N} \). However, in the ROF model the trivial minimizer has the following higher stability property which follows directly from [22, Lemma 4].

Proposition 3.8  If \( \|f\|_G < \frac{1}{\sqrt{\lambda}} \), then \( u \equiv 0 \) is the minimizer of \( J \), and for any sequence \( \{h_n\} \) with \( h_n \rightharpoonup 0 \) weakly in \( L^2(\Omega) \) and \( \|h_n\|_G \to 0 \), there exists a \( n_0 \in \mathbb{N} \) such that, for all \( n \geq n_0 \), the function \( u \equiv 0 \) is a minimizer of \( J_n \), where

\[
J(u) := \|u\|_{BV(\Omega)} + \lambda \|u - f\|^2_{L^2(\Omega)} \quad \text{and} \quad J_n(u) := \|u\|_{BV(\Omega)} + \lambda \|u - f - h_n\|^2_{L^2(\Omega)}.
\]

8
The higher stability of the trivial minimizer depends heavily on $\alpha$, and fails in general if $\alpha > 1$. Indeed, set

$$h_n(x) := h(nx), \quad \text{where} \quad h(x) := \begin{cases} -1 & 0 \leq x \leq \frac{1}{2} \\ +1 & \frac{1}{2} \leq x \leq 1, \end{cases}$$

and $h$ is extended periodically to $\mathbb{R}$ with period 1, and consider

$$J_n(u) := |u|_{BV(0,1)}^2 + \lambda \|u - h_n\|_{L^2(0,1)}^2.$$

Then $J_n(0) = \lambda \|h_n\|_{L^2(0,1)}^2 = \lambda$. Now consider the competitor functions $u_n(x) := c_n h_n$ with $c_n := \frac{\lambda}{n^2 + \lambda}$. We have

$$J_n(u_n) = \frac{\lambda^2 n^2}{(n^2 + \lambda)^2} + \lambda \frac{n^4}{(n^2 + \lambda)^2} = \frac{\lambda n^2 + n^4}{(n^2 + \lambda)^2} \lambda < \lambda = J(0),$$

so that $u \equiv 0$ is not a minimizer of $J_n$ for any $n \in \mathbb{N}$.

### 4 Remarks on the Invariance for Clean, Simple Signals - Exact Reconstruction

We consider functionals $J$ of the general form (2). We say that $f$ is exactly reconstructable for some $\lambda > 0$, if the unique minimizer of $J$ is $f$. Fix $f \in X$ with $|f|_X \neq 0$. We first note the following elementary comparison principle: $f$ is a minimizer of $J$, i.e. $J(f) \leq J(u)$ for all $u \neq f$, if and only if

$$|f|_X^\alpha \leq |u|_X^\alpha + \lambda \|u - f\|_Y^\beta \quad \text{for all} \quad u \neq f,$$

or, equivalently,

$$\lambda \geq \sup_{u \neq f} \frac{|f|_X^\alpha - |u|_X^\alpha}{\|u - f\|_Y^\beta}. \quad (12)$$

The inequality in (12) is strict if and only if $f$ is the unique minimizer of $J$.

**Proposition 4.1 (Scaling and the Failure of Exact Reconstructibility)** If $\beta > 1$ and $|f|_X \neq 0$, then $f$ is not exactly reconstructable for any finite $\lambda > 0$.

**Proof:** For $0 < \epsilon < 1$ define the test functions $u_\epsilon := (1 - \epsilon) f$. Then

$$\sup_{u \neq f} \frac{|f|_X^\alpha - |u|_X^\alpha}{\|f - u\|_Y^\beta} \geq \sup_{0 < \epsilon < 1} \frac{(1 - (1 - \epsilon)^\alpha) |f|_X^\alpha}{\epsilon^\beta \|f\|_Y^\beta} = \infty,$$

since $1 - (1 - \epsilon)^n = \mathcal{O}(\epsilon)$ as $\epsilon \to 0$. \qed

Note that, as a consequence of this result, regardless of the choice of regularization, only the trivial images with $|f|_X = 0$ are invariant if one uses as fidelity norm the $L^2$-norm squared. We recall that in contrast, the $L^2$-norm squared satisfies the Brezis-Lieb condition, and hence yields stability of minimizers with respect to
weak perturbation. It is well-known that minimizers of the ROF functional (4) possess a loss of contrast, and indeed the only clean image which is left invariant is \( f \equiv 0 \) (see [13, 26]).

The next corollary addresses the borderline case \( \beta = 1 \). In particular, for \( L^p \) fidelity terms it has been noted that a signal \( f \) is left invariant by energy minimization (for \( \lambda \) sufficiently large) if and only if there exists a suitably regular, TV-realizing vector field (see [22, 13, 24, 15]). Let us first note here that the sufficiency of such a vector field immediately follows from our comparison estimate (12). Let \( 1 \leq p < \infty \), let \( Y := L^p(\Omega) \), and consider

\[
J(\lambda)(u) := |u|_X + \lambda \| u - f \|_{L^p(\Omega)}.
\]

Here we use the notation \( J(\lambda) \) to highlight the dependence on the regularization parameter \( \lambda \). We set \( p^* := \frac{p}{p-1} \) if \( p \neq 1 \), and \( p^* = \infty \) if \( p = 1 \).

**Corollary 4.2 (Exact Reconstruction and Regularization-realizing Vector Fields)** Suppose there is a function space \( \mathcal{F} \) such that

\[
|u|_X = \sup_{\phi \in \mathcal{F}} \int_{\Omega} u(x) \phi(x) \, dx \quad \text{for all } u \in \mathcal{X}.
\]

If there exists a \( \phi_f \in \mathcal{F} \cap L^{p^*}(\Omega) \) such that

\[
|f|_X = \int_{\Omega} f(x) \phi_f(x) \, dx,
\]

then for \( \lambda \geq \| \phi_f \|_{L^{p^*}(\Omega)} \) the signal \( f \) is a minimizer of \( J(\lambda) \) among all \( u \in \mathcal{X} \).

**Proof:** For \( u \in \mathcal{X} \), using Hölder’s inequality we estimate

\[
\frac{|f|_X - |u|_X}{\| f - u \|_{L^p(\Omega)}} \leq \frac{\int_{\Omega} |(f(x) - u(x)) \phi_f(x)| \, dx}{\| f - u \|_{L^p(\Omega)}} \leq \frac{\| f - u \|_{L^p(\Omega)} \| \phi_f \|_{L^{p^*}(\Omega)}}{\| f - u \|_{L^p(\Omega)}} = \| \phi_f \|_{L^{p^*}(\Omega)}.
\]

It now suffices to apply (12). □

A well-studied special case of Corollary 4.2 is TV-regularization (see [13, 15, 22, 24]), where \( \mathcal{X} := BV \), \( \mathcal{F} := \left\{ \text{Div} \, v : v \in C^1_c(\Omega; \mathbb{R}^N) , \| v \|_{L^\infty(\Omega)} \leq 1 \right\} \), and \( \mathcal{Y} := L^1(\Omega) \). Then \( f \) is reconstructable if there exists a vector field \( v \in L^1(\Omega) \) with \( \text{Div} \, v \in \mathcal{F} \) such that

\[
|f|_{BV} = \int_{\Omega} f \, \text{Div} \, v \, dx.
\]

The existence of the associated vector field is often also a necessary condition, (see for example [15, 24] for a proof using convex analysis).

Lastly, we show how the comparison estimate (12) can be used, in the case in which \( \mathcal{X} := BV \), to gain insight into when a smooth signal is reconstructable or not. Suppose that we are given a signal \( f \in W^{1,\infty}(\Omega) \), i.e.,
Since we may assume that $f$ has a strict local minimum at $x_0 := 0$. Without loss of generality, we may assume that $f(0) := 0$, and $f(x) \geq r(|x|)$ for all $x \in (-\epsilon, \epsilon)$. Since $r$ is strictly monotone, there exists $r^{-1}$ with $r^{-1}(x) \to 0$ as $x \to 0$. In particular, $\frac{1}{n} f(y_n) \geq r(|y_n|) \geq r^{-1}\left(\frac{n}{1}\right) \to 0$ as $n \to \infty$. Hence by (15) and (14) we deduce that

$$\sup_{n \geq N} \frac{|f|_{BV} - |u_n|_{BV}}{\|f - u_n\|_{L^p(\mathbb{R})}} \geq C \sup_{n \geq N} \frac{1/n}{1/n (y_n - x_n)^{1/p}} \geq \sup_{n \geq N} \frac{1}{y_n^{1/p}} = \infty ,$$

and the claim follows in view of (12).

Suppose now that $x_0 \in \partial \Omega$. Without loss of generality, we may assume that $x_0 := 0, (0, \epsilon) \subset \Omega$, $\mathbb{R} \cap \Omega = \emptyset$. Further, since $r$ is strictly monotone, there exists $r^{-1}$ with $r^{-1}(x) \to 0$ as $x \to 0$. In particular, $\frac{1}{n} = f(y_n) \geq r(|y_n|) \leq r^{-1}\left(\frac{n}{1}\right) \to 0$ as $n \to \infty$. Hence by (15) and (14) we deduce that

$$\sup_{n \geq N} \frac{|f|_{BV} - |u_n|_{BV}}{\|f - u_n\|_{L^p(\mathbb{R})}} \geq C \sup_{n \geq N} \frac{1/n}{1/n (y_n - x_n)^{1/p}} \geq \sup_{n \geq N} \frac{1}{y_n^{1/p}} = \infty ,$$

and the claim follows in view of (12).
and \( f(x) \geq r(|x|) \) for all \( x \in (0, \epsilon) \). Then there is a sequence \( \{y_n\} \) such that, for \( n \geq N \), we have \( 0 < y_n < \epsilon \), \( f(y_n) \leq \frac{1}{n} \) and \( f(x) \leq \frac{1}{n} \) for all \( x \in (0, y_n) \). We proceed as above. Define the test functions

\[
u_n(x) := \begin{cases} f(x) & x \in \Omega \setminus (0, y_n), \\ \frac{1}{n} & x \in (0, y_n). \end{cases}
\]

Then \( |f|_{BV} \geq |\nu_n|_{BV} + \frac{1}{n} \) and \( \|f - \nu_n\|_{L^p(\Omega)} \leq \frac{1}{n}y_n^{1/p} \), and we conclude as in (16).

We conclude by noting that in a series of papers (see [2, 3, 4]), Allard gives a detailed analysis of properties of minimizers for total variation based functionals. In particular, he proves a regularity result for the level sets and gives very restrictive properties of the minimizers in the case in which the input signal is a characteristic function. We note that Theorem 1.1 in [4] implies, as an easy corollary, that for \( L^2 \)-squared fidelity no indicator function of a simple set can be reconstructed exactly. This fact is now also a consequence of Proposition 4.1.

**Acknowledgments:** The research of RC was partly supported by an NSERC (Canada) Discovery Grant. The research of IF was partially funded by the National Science Foundation under Grant No. DMS-0905778. The research of BZ was funded by a postdoctoral fellowship of the National Science Foundation under Grant No. DMS-0905778. IF and BZ warmly thank the Center for Nonlinear Analysis (NSF Grant No. DMS-0635983) for its support during the preparation of this paper.

**References**


