

Minimizers of one-dimensional regularization problems with linear and nonlinear total variation

Luca Lussardi*

Marco Morandotti†

Federico Stra‡

August 1, 2026

Abstract

A study of the minimizers of one-dimensional Rudin–Osher–Fatemi-type functionals with linear or nonlinear total variation and fidelity term is undertaken. Conditions on the input datum and the parameters of the functional are found that force the input itself to be the minimizer or not. In the linear setting the discriminating conditions on the parameters are complementary highlighting the sharpness of the results. The nonlinear setting is substantially different: while the non-minimality of the datum is treated in analogy with the linear case, for the minimality of the input only partial answers are found. The results in the nonlinear setting hinge on auxiliary constrained or penalized minimization problems investigating the behavior of optimal transitions with prescribed height. Additionally, they are complemented by some numerical examples.

MSC2020 49K05 (26A45, 49J05)

Keywords energy minimization, (nonlinear) total variation, regularity of minimizers, optimality conditions.

Contents

1	Introduction	2
1.1	Main results and open questions	3
2	Preliminaries	4
2.1	Definitions and basic properties of BV functions	4
2.2	Technical results	7
2.3	Perspective of a convex function	14
3	Linear TV	17
3.1	The input is not a minimizer	17
3.2	The input is the minimizer	20
4	Nonlinear $TV_{\varphi,\theta}$	28
4.1	Boundary value problem	28
4.2	The input is not a minimizer	34
4.3	Is the input a minimizer?	35
4.3.1	Numerical examples	37

*Politecnico di Torino, Torino, luca.lussardi@polito.it, ORCID: 0000-0001-9130-3573

†Politecnico di Torino, Torino, marco.morandotti@polito.it, ORCID: 0000-0003-3528-6152

‡Politecnico di Torino, Torino, federico.stra@polito.it, ORCID: 0000-0001-8151-1045

A	Miscellanea	41
A.1	Alternative proofs of Proposition 4.4(i)	41
B	First variation of the functional in subsection 4.3	42
B.1	Differentiable setting	43
B.1.1	Special case $p = \gamma = 1$	43
	References	48

1 Introduction

In the framework of image processing, the total variation energy functional plays a fundamental role. Since the pioneering work of Rudin, Osher, and Fatemi [ROF92], total variation regularization has been extensively studied both from the theoretical and applied viewpoints, leading to a wide literature on qualitative properties of minimizers and on nonlinear variants of the classical model (see the survey [CCN15] and the references therein; in particular, see also [CL97] and [Cha04]).

Given an interval $I \subseteq \mathbb{R}$, a datum $g : I \rightarrow \mathbb{R}$ describing the gray level of an input image, $\lambda, \gamma > 0$, and $p \in [1, \infty]$, we study one-dimensional minimization problems for functionals of the form

$$F(u) = \text{TV}(u; I) + \lambda \|u - g\|_p^\gamma$$

as well as their nonlinear counterpart obtained by replacing the classical total variation TV with a generalized functional $\text{TV}_{\varphi, \theta}$. Precisely, in Definition 2.1 we define $\text{TV}(u; I)$ as the *pointwise* version of the total variation of the function u over the set I . Up to a choice of good representatives, this corresponds to the quantity $|Du|(I)$ defined for $u \in \text{BV}(I)$ according to [AFP00]. For $u \in \text{BV}(I)$, we define also the nonlinear total variation as

$$\text{TV}_{\varphi, \theta}(u; I) = \int_I \varphi(|u'(x)|) dx + |D^c u|(I) + \sum_{x \in J_u} \theta(|[u](x)|),$$

where $\varphi, \theta : [0, \infty) \rightarrow [0, \infty)$; see Definition 2.9 for the precise definition, including the requirements on φ and θ .

Among the relevant questions concerning this optimization problem, we are motivated by understanding the regularity of the minimizers and the degeneracy of the problem. In particular, in this article we address these questions:

Q1: Can the minimizer exhibit a Cantor part $D^c u \neq 0$?

Q2: Is the minimizer $u = g$?

We will obtain an answer to the former question as a byproduct of our investigation of the latter. With regard to Q2, our analysis shows that the behavior of g near the boundary of the domain is crucial. More specifically, under the assumption of at least or at most α -growth at the endpoints (see Definition 2.10), we determine conditions on the parameters p, γ which cause the input g to be a minimizer for sufficiently large λ , or cause it not to be a minimizer for any λ . The threshold $1/\gamma \leq 1 + 1/(p\alpha)$ that we find for the two behaviors is sharp.

The linear case turns out to admit a more detailed analysis. We obtain several sharp results characterizing both the failure and the validity of minimality of the datum. On the one hand, we prove that functions exhibiting sufficiently fast growth near the boundary cannot be minimizers. On the other hand, monotone or piecewise monotone data with controlled growth at the endpoints are shown to be essentially unique minimizers for sufficiently large values of the fidelity parameter

λ. Moreover, the Cantor staircase function provides an interesting example showcasing that the minimizers may exhibit only the Cantor part of the derivative, thus answering Q1 positively.

By contrast, the nonlinear setting is considerably more difficult and remains only partially understood. Although we are able to establish some preliminary structural results and derive partial analogues of the linear theory, the nonlinear nature of $TV_{\varphi,\theta}$ leads to significant differences and obstacles with respect to the linear problem; as a consequence, most of the techniques developed in the linear case do not extend directly to the nonlinear case, mostly due to the fact that the functional exhibits a worse behavior with respect to local truncations. To work around this issue, we resort to the auxiliary boundary value problem (4.1), which allows us to characterize optimal transitions with prescribed height. Moreover, the failure of the analogues to the linear theorems is highlighted by the behavior of the minimizers of the penalized functional (4.7). Yet, our treatment of the nonlinear problem is only partial and we regard this part of the paper as mainly exploratory, with some examples suggesting that the nonlinear theory may exhibit genuinely different phenomena. This section is complemented by numerical examples to showcase and visualize some possible behaviors of the minimizers of (4.7).

Our proofs are mainly based on fine constructions of competitors for the minimization problems heavily exploiting the convexity of the functionals. In addition, we resort to a range of techniques from calculus of variations and convex analysis, including a detailed study of Euler–Lagrange equations (for the nonlinear problem), the notion of perspective of a convex function, Legendre transform, and other inequalities inspired by optimal transport on the real line.

The article is organized as follows. In [Section 2](#) we collect the necessary preliminaries on functions of bounded variation and introduce the nonlinear total variation functional. [Section 3](#) contains the main results for the linear model, distinguishing between situations where the datum fails to be a minimizer and situations where it is optimal. In [Section 4](#) we turn to the nonlinear setting, where we analyze the associated boundary value problem, obtain partial results which leave some open questions, and showcase the behavior of optimal profiles by means of numerical examples. The approach is based on insight about the minimization problem developed in the appendix [Subsection B.1.1](#).

1.1 Main results and open questions

We summarize here the main results contained in the paper.

- **Linear total variation.** After the simple and illustrative example [Theorem 3.1](#) with the Cantor staircase function, in [Theorem 3.4](#) we show that if g has sufficiently fast growth near the endpoints and γ is large enough then g is not a minimizer. In [Subsection 3.2](#) we provide progressively more general results of minimality of the input datum: we start with the simple situation in [Theorem 3.7](#) where g is constant near the endpoints; we move on to the case in [Theorem 3.8](#) where g is monotone and has controlled growth; finally, with the help of the fundamental [Proposition 3.9](#), in [Theorem 3.10](#) we extend the result to piecewise monotone functions. The complementarity of the ranges for the exponent γ in [Theorem 3.4](#) and [Theorem 3.8](#) demonstrates the sharpness of our results.
- **Nonlinear total variation.** In [Theorem 4.9](#) we prove the analogue of [Theorem 3.4](#) under the assumption of a lower bound on the boundary value problem (4.1). [Subsection 4.3](#) is more exploratory: as it turns out surprisingly, the desired nonlinear statement corresponding to [Proposition 3.9](#) is false, a fact which we demonstrate with a very simple counterexample in [Proposition 4.10](#). The culprit is the non minimality of constant functions for a functional penalizing jumps at the right endpoint of the interval, as stated in [Proposition 4.11](#). We

demonstrate the behavior of minimizers of this penalized problem with some numerical examples in [Subsection 4.3.1](#).

We stress the fact that our study of the linear case ([Section 3](#)) provides more precise statements than what could be derived from the nonlinear case ([Section 4](#)) by specialising to $\varphi = \theta = \text{Id}$. This is because the [Definition 2.1](#) of TV is a pointwise version of total variation which is sensitive to pointwise modifications of the function, in contrast with $\text{TV}_{\varphi, \theta}$ of [Definition 2.9](#), which is invariant in the L^1 equivalence classes.

Our partial treatment of the regularity issue leaves open some questions more advanced than Q1, for instance whether there exists a minimizer $u \in \text{SBV}$ when the input is $g \in \text{SBV}$. This could have implications in regularization of images.

Finally, in the nonlinear setting, future research may help identify sufficient conditions for the minimality of the input, potentially based on the behavior in 0 and at ∞ of the non-linearities φ and θ . While our results show that there are situations where the input performs better than certain specific competitors (see [Proposition 4.10](#)), the exposition in [Subsection 4.3](#) demonstrates that the situation is substantially more involved than the linear case and a more detailed analysis is required to fully resolve this matter.

2 Preliminaries

2.1 Definitions and basic properties of BV functions

Definition 2.1. Given a subset $I \subseteq \mathbb{R}$ and a function $u : I \rightarrow \mathbb{R}$, the *total variation* of u over I is

$$\text{TV}(u; I) = \sup \left\{ \sum_{i=1}^n |u(x_i) - u(x_{i-1})| : x_i \in I \ \forall i \in \{0, \dots, n\} \wedge x_{i-1} \leq x_i \ \forall i \in \{1, \dots, n\} \right\}.$$

The vector space of *functions with finite total variation* over I is

$$\text{TV}(I) = \{u : I \rightarrow \mathbb{R} : \text{TV}(u; I) < \infty\}.$$

In the sequel, we may use the abbreviated notation $\text{TV}(u)$ instead of $\text{TV}(u; I)$ when it is clear from the context.

Definition 2.2. A function $u : I \rightarrow \mathbb{R}$ is said to be *regulated* if the left limit $u^-(x)$ exists and is finite at any $x \in \mathbb{R}$ which is an accumulation point of I from the left, and the right limit $u^+(x)$ exists and is finite at any $x \in \mathbb{R}$ which is an accumulation point of I from the right.

Proposition 2.3. *Every function $u \in \text{TV}(I)$ is regulated.*

Definition 2.4. A regulated function $u : I \rightarrow \mathbb{R}$ is said to be *regular* if

$$u(x) \in [u^-(x) \wedge u^+(x), u^-(x) \vee u^+(x)]$$

for every $x \in I$ which is a bilateral accumulation point of I and $u(x) = u^\pm(x)$ at any point $x \in I$ which is a unilateral accumulation point of I . The set of *regular functions with finite total variations* over I is

$$\text{TV}_{\text{reg}}(I) = \{u \in \text{TV}(I) : u \text{ is regular}\}.$$

The total variation $\text{TV}(u; I)$ and the space of functions $\text{TV}(I)$ are the pointwise version of $|\text{Du}|(I)$ and $\text{BV}(I)$ as defined in [\[AFP00\]](#). More explicitly, $\text{TV}(u; I)$ is precisely the *pointwise variation* defined in [\[AFP00, Definition 3.26\]](#); [\[AFP00, Theorem 3.27\]](#) shows that the variation

$|Du|(I)$ is equal to the relaxation of $\text{TV}(u, I)$ in the $\mathcal{L}^1$ equivalence class (see also [AFP00, (3.23)]); [AFP00, Theorem 3.28] shows that any function $u \in \text{BV}(I)$ admits a *good representative* $v \in \text{TV}_{\text{reg}}(I)$ such that $|Du|(I) = \text{TV}(v; I)$.

The following properties follow directly from the definitions.

Proposition 2.5. *Let $I \subseteq \mathbb{R}$ be an interval.*

- (i) *If $u \in \text{TV}(I)$, then u has at most countably many discontinuities (and they are jumps).*
- (ii) *If $u \in \text{TV}(I)$, then u is regular if and only if $\liminf_{x' \rightarrow x} u(x') \leq u(x) \leq \limsup_{x' \rightarrow x} u(x')$ for every accumulation point $x \in I$.*
- (iii) *If $u, v \in \text{TV}_{\text{reg}}(I)$ coincide $\mathcal{L}^1$ -a.e., then they share the same set of jump points and they coincide elsewhere.*
- (iv) *Every $u \in \text{BV}(I)$ admits an equivalent function $v \in \text{TV}_{\text{reg}}(I)$ which coincides $\mathcal{L}^1$ -a.e. with u .*
- (v) *If $u \in \text{TV}(I)$, then a function $v : I \rightarrow \mathbb{R}$ is regular and coincides with u $\mathcal{L}^1$ -a.e. if and only if $\liminf_{x' \rightarrow x} u(x') \leq v(x) \leq \limsup_{x' \rightarrow x} u(x')$ for every $x \in I$.*
- (vi) *If u is regulated and x is a bilateral accumulation point of I , then $u(x) \in [u^-(x) \wedge u^+(x), u^-(x) \vee u^+(x)]$ if and only if the “pointwise variation” $\llbracket u \rrbracket = |u(x) - u^-(x)| + |u^+(x) - u(x)|$ is equal to the jump $|\llbracket u \rrbracket(x)| = |u^+(x) - u^-(x)|$.*

Regulated, regular and TV functions are well behaved with regard to composition with continuous or Lipschitz functions.

Proposition 2.6. *Let $I \subseteq \mathbb{R}$.*

- (i) *If $u : I \rightarrow \mathbb{R}$ is regulated and $f \in C(\overline{u(I)})$, then $f \circ u$ is regulated.*
- (ii) *If $u : I \rightarrow \mathbb{R}$ is regular and $f \in C(\overline{u(I)})$ is monotone, then $f \circ u$ is regular.*
- (iii) *If $u \in \text{TV}(I)$ and $f \in \text{Lip}(u(I))$, then $f \circ u \in \text{TV}(I)$ and $\text{TV}(f \circ u; I) \leq \text{Lip}(f) \text{TV}(u; I)$.*
- (iv) *If $u \in \text{TV}_{\text{reg}}(I)$ and $f \in \text{Lip}(u(I))$ is monotone, then $f \circ u \in \text{TV}_{\text{reg}}(I)$.*

Proof. Since the rest are straightforward, we prove only (iii), which will play a crucial role in what follows.

- (iii) For every choice of sorted points $x_0 < \dots < x_n$ in I we have

$$\sum_{i=1}^n |f(u(x_i)) - f(u(x_{i-1}))| \leq \text{Lip}(f) \sum_{i=1}^n |u(x_i) - u(x_{i-1})| \leq \text{Lip}(f) \text{TV}(u; I);$$

taking the supremum on the left hand side proves the thesis. $\square$

As a special application of the last two points of Proposition 2.6 we get the following useful property.

Remark 2.7. Given $m, M \in \mathbb{R}$, the functions $f_m, f_m^M, f_m^M \in \text{Lip}(\mathbb{R})$ defined as

$$f_m(t) = t \vee m, \quad f^M(t) = t \wedge M, \quad f_m^M(t) = t \vee m \wedge M$$

are 1-Lipschitz and monotone, therefore if $u \in \text{TV}(I)$ then the truncated functions $u \vee m$, $u \wedge M$ and $u \vee m \wedge M$ are also in $\text{TV}(I)$ and

$$\text{TV}(u \vee m; I), \text{TV}(u \wedge M; I), \text{TV}(u \vee m \wedge M; I) \leq \text{TV}(u; I).$$

If $u \in \text{TV}_{\text{reg}}(I)$, the truncated functions are also in $\text{TV}_{\text{reg}}(I)$.

We collect here some further elementary properties of the pointwise total variation.

Proposition 2.8. *Let $I \subseteq \mathbb{R}$ be an interval and $u \in \text{TV}(I)$. Then the following hold true.*

- (i) *u can be expressed as the difference of the two non-decreasing functions $x \mapsto \text{TV}(u; (-\infty, x])$ and $x \mapsto \text{TV}(u; (-\infty, x]) - u(x)$;*
- (ii) *$\text{TV}(u; [a, c]) = \text{TV}(u; [a, b]) + \text{TV}(u, [b, c])$ for every $a \leq b \leq c$ in I ;*
- (iii) *$\text{TV}(u; [a, c]) = |u(a) - u^+(a)| + \text{TV}(u, (a, c])$ for every $a < c$ in I ;*
- (iv) *if $\tilde{u} : I \rightarrow \mathbb{R}$ is such that*

$$\liminf_{y \rightarrow x} u(y) \leq \tilde{u}(x) \leq \limsup_{y \rightarrow x} u(y),$$

then $\text{TV}(\tilde{u}; I) \leq \text{TV}(u; I)$;

Proof. Property (i) is a well known fact; the next two splitting properties are elementary; therefore, we only prove (iv).

- (iv) Let $x_0 < x_1 < \dots < x_n$ be a sorted partition of I . Given $\varepsilon > 0$, for every $i \in \{0, \dots, n\}$ there exist distinct points $x_{i,\varepsilon}^\pm \in I$ such that

$$\begin{aligned} u(x_{i,\varepsilon}^-) - \varepsilon &\leq \tilde{u}(x_i) \leq u(x_{i,\varepsilon}^+) + \varepsilon & \forall i \in \{0, \dots, n\}, \\ x_{i-1,\varepsilon}^\pm &< x_{i,\varepsilon}^\pm & \forall i \in \{1, \dots, n\}. \end{aligned} \quad (2.1)$$

For every $i \in \{0, \dots, n\}$, let $x'_{i,\varepsilon} < x''_{i,\varepsilon}$ such that $\{x'_{i,\varepsilon}, x''_{i,\varepsilon}\} = \{x_{i,\varepsilon}^-, x_{i,\varepsilon}^+\}$. The sequence of points $x'_{0,\varepsilon} < x''_{0,\varepsilon} < x'_{1,\varepsilon} < \dots < x'_{n,\varepsilon} < x''_{n,\varepsilon}$ forms a partition of I . By the triangle inequality and (2.1) we get

$$\begin{aligned} |u(x'_{i,\varepsilon}) - \tilde{u}(x_i)| + |\tilde{u}(x_i) - u(x''_{i,\varepsilon})| &= |u(x_{i,\varepsilon}^+) - \tilde{u}(x_i)| + |\tilde{u}(x_i) - u(x_{i,\varepsilon}^-)| \\ &\leq \left| (u(x_{i,\varepsilon}^+) + \varepsilon) - \tilde{u}(x_i) \right| + \left| \tilde{u}(x_i) - (u(x_{i,\varepsilon}^-) - \varepsilon) \right| + 2\varepsilon \\ &= \left| (u(x_{i,\varepsilon}^+) + \varepsilon) - (u(x_{i,\varepsilon}^-) - \varepsilon) \right| + 2\varepsilon \\ &\leq |u(x_{i,\varepsilon}^+) - u(x_{i,\varepsilon}^-)| + 4\varepsilon = |u(x'_{i,\varepsilon}) - u(x''_{i,\varepsilon})| + 4\varepsilon, \end{aligned}$$

from which we obtain

$$\begin{aligned} \sum_{i=1}^n |\tilde{u}(x_{i-1}) - \tilde{u}(x_i)| &\leq \sum_{i=1}^n \left(|\tilde{u}(x_{i-1}) - u(x''_{i-1,\varepsilon})| + |u(x''_{i-1,\varepsilon}) - u(x'_{i,\varepsilon})| + |u(x'_{i,\varepsilon}) - \tilde{u}(x_i)| \right) \\ &\leq \sum_{i=0}^n \left(|u(x'_{i,\varepsilon}) - \tilde{u}(x_i)| + |\tilde{u}(x_i) - u(x''_{i,\varepsilon})| \right) + \sum_{i=1}^n |u(x''_{i-1,\varepsilon}) - u(x'_{i,\varepsilon})| \\ &\leq \sum_{i=0}^n |u(x'_{i,\varepsilon}) - u(x''_{i,\varepsilon})| + \sum_{i=1}^n |u(x''_{i-1,\varepsilon}) - u(x'_{i,\varepsilon})| + 4(n+1)\varepsilon. \end{aligned}$$

By definition of total variation, the right hand side is less than or equal to $\text{TV}(u; I) + 4(n+1)\varepsilon$. Sending $\varepsilon \rightarrow 0$ and taking the supremum of the left hand side over the partition $(x_i)_{0 \leq i \leq n}$ we deduce the thesis. $\square$

We recall the classical decomposition $Du = D^a u + D^c u + D^j u$ of the derivative of a function $u \in \text{BV}(I)$; see [AFP00, Corollary 3.33]. The (linear) total variation of a function $u \in \text{TV}_{\text{reg}}(I)$ can be represented as

$$\text{TV}(u; I) = \int_I |u'(x)| \, dx + |D^c u|(I) + \sum_{x \in J_u} |[u](x)|,$$

where J_u is the jump set of u and $[u](x) = g^+(x) - g^-(x)$ is the jump value. This leads us to introduce the nonlinear total variation $\text{TV}_{\varphi, \theta}(u; I)$ as follows.

Definition 2.9. Given two non-decreasing functions $\varphi, \theta : [0, \infty) \rightarrow [0, \infty)$ such that φ is convex, θ is concave, $\varphi(0) = \theta(0) = 0$ and

$$\lim_{t \rightarrow \infty} \frac{\varphi(t)}{t} = \lim_{t \rightarrow 0^+} \frac{\theta(t)}{t} = \theta'(0) = 1,$$

the *nonlinear total variation* of $u \in \text{TV}(I)$ is

$$\text{TV}_{\varphi, \theta}(u; I) = \int_I \varphi(|u'(x)|) \, dx + |D^c u|(I) + \sum_{x \in J_u} \theta(|[u](x)|).$$

It is simple to verify that the following conditions are equivalent:

$$\varphi(t) = t, \, \forall t \in [0, \infty) \quad \Leftrightarrow \quad \exists t \in (0, \infty) : \varphi(t) = t \quad \Leftrightarrow \quad \varphi'(0) = 1.$$

Our main theorems rely on the following property of the input datum near the endpoints. The precise assumption will be anyway recalled explicitly in their statements.

Definition 2.10. Given $\alpha > 0$, a function $u : (a, b) \rightarrow \mathbb{R}$ which admits the limit $g^+(a)$ is said to have *at most α -growth* at the endpoint a if there exists $C > 0$ such that $|g(x) - g^+(a)| \leq C(x - a)^\alpha$ for $x \in (a, b)$. It is said to have *at least α -growth* if the reverse inequality holds.

2.2 Technical results

Definition 2.11. Given a measurable function $u : \Omega \rightarrow [0, \infty]$, let $\mu_u : [0, \infty] \rightarrow [0, |\Omega|]$ be its *super-level set measure*

$$\mu_u(t) = |\{x \in \Omega : u(x) > t\}|.$$

This function is non-increasing and lower semi-continuous, therefore it admits a non-increasing lower semi-continuous pseudo-inverse $(\mu_t)^{-1} : [0, |\Omega|] \rightarrow [0, \infty]$ given by

$$(\mu_u)^{-1}(s) = \inf\{t : \mu_u(t) \leq s\}.$$

We define the *non-increasing* and *non-decreasing rearrangements* $u_\dagger, u^\dagger : [0, |\Omega|] \rightarrow [0, \infty]$ of u as

$$u_\dagger(s) = (\mu_u)^{-1}(s), \quad u^\dagger(s) = (\mu_u)^{-1}(|\Omega| - s), \quad \forall s \in [0, |\Omega|].$$

The functions μ_u and $(\mu_u)^{-1}$ are related by the crucial property

$$\mu_u(t) \leq s \iff (\mu_u)^{-1}(s) \leq t.$$

Moreover, one can show that $\mu_{u_\dagger} = \mu_{u^\dagger} = \mu_u$ (see Remark 2.13, where it is implicitly proven), therefore by the layer-cake representation we have

$$\int_\Omega f(u(x)) \, dx = \int_0^{|\Omega|} f(u_\dagger(s)) \, ds = \int_0^{|\Omega|} f(u^\dagger(s)) \, ds$$

for every $f : [0, \infty] \rightarrow [0, \infty]$ measurable.

The following proposition collects two contractivity properties of the monotone rearrangement, both with respect to the L^p norm and the (linear) total variation. The proof is based upon two technical lemmas, which we postpone. See also [Tal16] for a survey about rearrangements.

Proposition 2.12. *Let $p \in [1, \infty]$ and $u, v : (a, b) \rightarrow [0, \infty]$. Then*

$$\|u^\dagger - v^\dagger\|_p \leq \|u - v\|_p, \quad (2.2)$$

$$\text{TV}(u^\dagger; (0, b - a)) \leq \text{TV}(u; (a, b)). \quad (2.3)$$

For $p > 1$, equality in (2.2) holds if and only if

$$(u(x_2) - u(x_1))(v(x_2) - v(x_1)) \geq 0, \quad \text{for almost every } x_1, x_2 \in (a, b); \quad (2.4a)$$

for $p = 1$, equality in (2.2) holds if and only if

$$\max\{u(x_1), v(x_2)\} \geq \min\{u(x_2), v(x_1)\}, \quad \text{for almost every } x_1, x_2 \in (a, b). \quad (2.4b)$$

Equality in (2.3) holds if and only if u is monotone,

Proof. For the convenience of the reader, we provide a sketch of the proof of the classical result (2.2) for the case $(a, b) = (0, 1)$. See also [Bae19, Theorem 2.8(c)] with $\Psi(x, y) = |x - y|^p$ or [Rak25, Theorem 1.3.1, Corollary 1.3.2] with $\rho(x, y) = |x - y|^p$ for independent and more complete proofs. By approximation with simple functions, let

$$u_n = \sum_{i=1}^n u_{n,i} \mathbf{1}_{(\frac{i-1}{n}, \frac{i}{n}]}, \quad v_n = \sum_{i=1}^n v_{n,i} \mathbf{1}_{(\frac{i-1}{n}, \frac{i}{n}]},$$

Let $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ be a permutation that rearranges the values $(u_{n,i})_{i=1}^n$ in non-decreasing order, i.e. $u_{n,\sigma(i)} \leq u_{n,\sigma(i+1)}$ for every $i \in \{1, \dots, n-1\}$. Suppose that $i \mapsto v_{n,\sigma(i)}$ is not non-decreasing: then there must be $j \in \{1, \dots, n-1\}$ such that $v_{n,\sigma(j)} > v_{n,\sigma(j+1)}$. Let π be the transposition that swaps j and $j+1$. By applying Lemma 2.14 below with

$$a = u_{n,\sigma(j)}, \quad A = u_{n,\sigma(j+1)}, \quad b = v_{n,\sigma(j+1)}, \quad B = v_{n,\sigma(j)},$$

we get that

$$\|u_n - v_n\|_p^p = \sum_{i=1}^n |u_{n,i} - v_{n,i}|^p = \sum_{i=1}^n |u_{n,\sigma(i)} - v_{n,\sigma(i)}|^p \geq \sum_{i=1}^n |u_{n,\sigma(i)} - v_{n,\pi \circ \sigma(i)}|^p.$$

The number of inversions in the sequence $(v_{n,\pi \circ \sigma(i)})_{i=1}^n$ is one less than that in $(v_{n,\sigma(i)})_{i=1}^n$. By iterating this argument, we find a permutation τ such that $(v_{n,\tau(i)})_{i=1}^n$ is non-decreasing and

$$\|u_n - v_n\|_p^p \geq \sum_{i=1}^n |u_{n,\sigma(i)} - v_{n,\tau(i)}|^p = \|u_n^\dagger - v_n^\dagger\|_p^p.$$

The proof is concluded letting $u_n \rightarrow u$ and $v_n \rightarrow v$, for which we have also $u_n^\dagger \rightarrow u^\dagger$ and $v_n^\dagger \rightarrow v^\dagger$.

The proof of the equality condition for (2.2) is reminiscent of the proof of c -cyclical monotonicity of optimal plans; see for instance [ABS21, Theorem 3.17]. Suppose $p > 1$ and assume by contradiction that (2.4a) is violated. Without loss of generality, we may assume that $x_1 < x_2$, $u(x_1) < u(x_2)$ and $v(x_1) > v(x_2)$, with x_1, x_2 Lebesgue points of both u and v . For sufficiently small radius $r > 0$, define the intervals $I_i^r = (x_i - r, x_i + r)$ and the rearranged function $v_r : (a, b) \rightarrow \mathbb{R}$ given by

$$v_r(x) = \begin{cases} v(x + (x_2 - x_1)) & x \in I_1^r, \\ v(x - (x_2 - x_1)) & x \in I_2^r, \\ v(x) & \text{otherwise.} \end{cases}$$

By Lemma 2.14 with $a = u(x_1)$, $A = u(x_2)$, $b = v(x_2)$, $B = v(x_1)$ and $f(t) = |t|^p$ we have

$$\begin{aligned} |u(x_1) - v(x_1)|^p + |u(x_2) - v(x_2)|^p &> |u(x_1) - v(x_2)|^p + |u(x_2) - v(x_1)|^p \\ &= |u(x_1) - v_r(x_1)|^p + |u(x_2) - v_r(x_2)|^p. \end{aligned}$$

By the Lebesgue property, this implies that for r sufficiently small

$$J_r := \int_{-r}^r \sum_{i=1}^2 (|u(x_i + y) - v(x_i + y)|^p - |u(x_i + y) - v_r(x_i + y)|^p) dy > 0.$$

Then

$$\|u - v\|_p^p - \|u - v_r\|_p^p = \int_{I_1' \cup I_2'} (|u(x) - v(x)|^p - |u(x) - v_r(x)|^p) dx = 2rJ_r > 0.$$

This proves that

$$\|u - v\|_p > \|u - v_r\|_p \geq \|u^\dagger - v_r^\dagger\|_p = \|u^\dagger - v^\dagger\|_p,$$

hence (2.2) is a strict inequality. The proof of (2.4b) for the case $p = 1$ is analogous.

Let us now turn to the proof of (2.3). Given $\varepsilon > 0$, let $x'_\varepsilon, x''_\varepsilon \in (a, b)$ be two different points such that

$$u(x'_\varepsilon) \leq \inf_{(a,b)} u + \varepsilon \quad \text{and} \quad u(x''_\varepsilon) \geq \sup_{(a,b)} u - \varepsilon.$$

By Definition 2.1,

$$\text{TV}(u; (a, b)) \geq |u(x'_\varepsilon) - u(x''_\varepsilon)| \geq \sup_{(a,b)} u - \inf_{(a,b)} u - 2\varepsilon.$$

Letting $\varepsilon \rightarrow 0$ shows that

$$\text{TV}(u; (a, b)) \geq \sup_{(a,b)} u - \inf_{(a,b)} u. \quad (2.5)$$

On the other hand,

$$\text{TV}(u^\dagger; (0, b-a)) = \sup_{(0,b-a)} u^\dagger - \inf_{(0,b-a)} u^\dagger = \text{ess sup}_{(a,b)} u - \text{ess inf}_{(a,b)} u \leq \sup_{(a,b)} u - \inf_{(a,b)} u,$$

therefore we deduce the thesis. Let us now prove that if equality in (2.5) holds then u is monotone. We prove the contrapositive. Assume that u is not monotone: there exist $x_1 < x_2$ and $x_3 < x_4$ such that $u(x_1) < u(x_2)$ and $u(x_3) > u(x_4)$. For every $\varepsilon > 0$, let $x'_\varepsilon, x''_\varepsilon$ as before. Let $S_\varepsilon = \{x_1, x_2, x_3, x_4, x'_\varepsilon, x''_\varepsilon\}$. Rename the points as $(y_0^\varepsilon, \dots, y_5^\varepsilon) = (x_1, x_2, x_3, x_4, x'_\varepsilon, x''_\varepsilon)$. There exists a permutation σ_ε of $\{0, \dots, 5\}$ such that $z_i^\varepsilon = y_{\sigma_\varepsilon(i)}^\varepsilon$ is increasing in i . Since the set of permutations is finite, there exists $\varepsilon_n \rightarrow 0^+$ such that σ_{ε_n} is always the same permutation σ .

Let $a_i^{\varepsilon_n} = u(y_i^{\varepsilon_n})$, hence $u(z_i^{\varepsilon_n}) = a_{\sigma(i)}^{\varepsilon_n}$. When $n \rightarrow \infty$, the finite sequence $(a_i^{\varepsilon_n})_{i=0}^5$ converges to some $(a_i)_{i=0}^5$ with the property that $(a_{\sigma(i)})_{i=0}^5$ is not monotone, because of the presence of the points x_1, x_2, x_3, x_4 which cause a violation of monotonicity. From $\sup u - \varepsilon \leq u(x'_\varepsilon) \leq \max_i a_i^\varepsilon \leq \sup u$ we deduce $\lim_{n \rightarrow \infty} \max_i a_i^{\varepsilon_n} = \sup u$, and similarly for min and inf. By Lemma 2.15 below we therefore have

$$\begin{aligned} \text{TV}(u) &\geq \liminf_{n \rightarrow \infty} \text{TV}(u; S_{\varepsilon_n}) = \liminf_{n \rightarrow \infty} \sum_{i=1}^5 |u(z_i^{\varepsilon_n}) - u(z_{i-1}^{\varepsilon_n})| = \liminf_{n \rightarrow \infty} \sum_{i=1}^5 |u(y_{\sigma(i)}^{\varepsilon_n}) - u(y_{\sigma(i-1)}^{\varepsilon_n})| \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^5 |a_{\sigma(i)}^{\varepsilon_n} - a_{\sigma(i-1)}^{\varepsilon_n}| = \sum_{i=1}^5 |a_{\sigma(i)} - a_{\sigma(i-1)}| > \max_i a_i - \min_i a_i = \sup_{(a,b)} u - \inf_{(a,b)} u, \end{aligned}$$

therefore the inequality in (2.5) is strict. $\square$

Remark 2.13. Inequality (2.2) in the previous Proposition 2.12 is closely related to the theory of optimal transport on the real line with a convex cost. Indeed, introduce the probabilities $\mu = u_{\#}(\mathcal{L}^1 \llcorner (0, 1))$, $\nu = v_{\#}(\mathcal{L}^1 \llcorner (0, 1)) \in \mathcal{P}([0, \infty))$. We claim that $\mu = (u_{\dagger})_{\#}(\mathcal{L}^1 \llcorner (0, 1))$, and similarly for ν . It is sufficient to prove that the measures coincide on the set (t, ∞) for every $t > 0$. We have

$$u^{-1}((t, \infty)) = \{x \in (0, 1) : u(x) > t\}$$

and

$$\begin{aligned} (u_{\dagger})^{-1}((t, \infty)) &= \{s \geq 0 : u_{\dagger}(s) > t\} = \{s \geq 0 : (\mu_u)^{-1}(s) > t\} \\ &= \{s \geq 0 : \mu_u(t) > s\} = [0, \mu_u(t)), \end{aligned}$$

therefore

$$\mu((t, \infty)) = \mathcal{L}^1(u^{-1}(t, \infty)) = \mathcal{L}^1(\{x \in (0, 1) : u(x) > t\}) = \mu_u(t)$$

and

$$(u_{\dagger})_{\#}(\mathcal{L}^1 \llcorner (0, 1))((t, \infty)) = \mathcal{L}^1((u_{\dagger})^{-1}((t, \infty))) = \mathcal{L}^1([0, \mu_u(t))) = \mu_u(t).$$

Define now the transport plans $\pi, \pi_{\dagger} \in \Pi(\mu, \nu)$ given by

$$\pi = (u, v)_{\#}(\mathcal{L}^1 \llcorner (0, 1)), \quad \pi_{\dagger} = (u_{\dagger}, v_{\dagger})_{\#}(\mathcal{L}^1 \llcorner (0, 1)).$$

The plan $\pi_{\dagger}$ is supported on a monotone graph¹, hence it is c -cyclically monotone with respect to the cost $c(x, y) = |x - y|^p$, therefore it is optimal (see [ABS21, Proposition 4.5] where the converse is proven, although it is actually an equivalence; observe that their inequality (4.5) there is related to our Lemma 2.14; see also [ABS21, Theorem 1.11] for the optimality of the monotone rearrangement). This implies that

$$\begin{aligned} \|u_{\dagger} - v_{\dagger}\|_p^p &= \int_0^1 |u_{\dagger}(z) - v_{\dagger}(z)|^p dz = \int_{[0, \infty)^2} |x - y|^p d\pi_{\dagger}(x, y) \\ &\leq \int_{[0, \infty)^2} |x - y|^p d\pi(x, y) = \int_0^1 |u(z) - v(z)|^p dz = \|u - v\|_p^p. \end{aligned}$$

With regard to the equality condition, if the inequality (2.2) is saturated and $p > 1$, then the plan π must be supported on a monotone graph too, which translates to (2.4a).

Lemma 2.14. *If $f : \mathbb{R} \rightarrow \mathbb{R}$ is a convex function and $a \leq A$ and $b \leq B$, then*

$$f(b - a) + f(B - A) \leq f(B - a) + f(b - A).$$

Moreover, the inequality is strict if f is strictly convex, $a < A$ and $b < B$.

In particular, for every $p \geq 1$ we have $|b - a|^p + |B - A|^p \leq |B - a|^p + |b - A|^p$. If $p > 1$ the inequality is strict if and only if $a < A$, and $b < B$. For $p = 1$, the inequality is strict if and only if $a < A$, $b < B$, $a < B$ and $b < A$, which is equivalent to $\max\{a, b\} < \min\{A, B\}$.

Proof. Let $x := b - A \leq b - a =: X$ and $\Delta b = B - b \geq 0$. The convexity of f implies the monotonicity of forward differences, therefore we have

$$f(x + \Delta b) - f(x) \leq f(X + \Delta b) - f(X),$$

that is

$$f(B - A) - f(b - A) \leq f(B - a) - f(b - a),$$

¹A set $\Gamma \subset \mathbb{R}^2$ such that $(x_2 - x_1)(y_2 - y_1) \geq 0$ for every $(x_1, y_1), (x_2, y_2) \in \Gamma$.

which is equivalent to the thesis. The inequality is strict if f is strictly convex, $x < X$ and $\Delta b > 0$, which are the stated conditions.

The inequality for $p \geq 1$ and the characterization of strict inequality for $p > 1$ follow using $f(t) = |t|^p$.

Let us turn to the characterization of strict inequality for $p = 1$, $f(t) = |t|$. If $x = X$ or $\Delta b = 0$, then equality holds. Let $x < X, \Delta b > 0$. By the monotonicity of difference quotients,

$$\frac{f(x + \Delta b) - f(x)}{\Delta b} \leq \frac{f(X + \Delta b) - f(x)}{X + \Delta b - x} \leq \frac{f(X + \Delta b) - f(X)}{\Delta b},$$

where equality is verified if and only if f is affine in $(x, X + \Delta b)$. Therefore, the strict inequality holds if and only if $0 \in (x, X + \Delta)$, which amounts to the required property on a, A, b, B . $\square$

The next lemma is a classical rearrangement inequality, whose proof can be found in any textbook about optimal transport, in relation to the c -cyclical monotonicity with respect to the cost $|x - y|$, see e.g. [ABS21].

Lemma 2.15. *Let $n \in \mathbb{N}_+$ and let $a = (a_i)_{i=0}^n$ be a finite sequence of real numbers. For every permutation $\sigma \in \mathcal{P}_n$ we have*

$$\sum_{i=1}^n |a_{\sigma(i)} - a_{\sigma(i-1)}| \geq \max_i a_i - \min_i a_i,$$

with equality if and only if $(a_{\sigma(i)})_{i=0}^n$ is monotone.

Proposition 2.16. *Let $g \in \text{BV}((a, b))$ and let $\lambda, \gamma > 0$, $p \in (0, \infty]$. If $u \in \text{BV}((a, b))$ minimizes the functional*

$$F(u) = \text{TV}(u) + \lambda \|u - g\|_p^\gamma$$

and $m \leq g \leq M$ then also $m \leq u \leq M$.

Proof. Let $\tilde{u} = u \vee m \wedge M$. Then $\text{TV}(\tilde{u}) \leq \text{TV}(u)$ by Remark 2.7, and $\|\tilde{u} - g\|_p \leq \|u - g\|_p$, with this latter inequality being strict unless $m \leq u \leq M$. $\square$

Proposition 2.17. *Let $g \in \text{BV}((a, b))$ be a non-decreasing function and $\lambda, \gamma > 0$, $p \in [1, \infty]$. If $u \in \text{BV}((a, b))$ minimizes the functional*

$$F(u) = \text{TV}(u) + \lambda \|u - g\|_p^\gamma$$

then it is non-decreasing.

Proof. If g is constant, then the only minimizer is $u = g$, which achieves $F(u) = 0$. Indeed, if u is any other constant function, then $\|u - g\|_p > 0$; otherwise, if u is not constant, then $\text{TV}(u) > 0$.

We may now assume that g is not constant; therefore, there exist $x_1 < x_2$ such that $g(x_1) < g(x_2)$. Let $\bar{u}(x) = u^\dagger(x - a)$ be the non-decreasing rearrangement of u suitably translated to preserve the same domain of definition. Since g is already non-decreasing, we have $g(x) = g^\dagger(x - a)$. By (2.2) and (2.3) of Proposition 2.12 we have

$$\begin{aligned} \|\bar{u} - g\|_{L^p((a, b))} &= \|u^\dagger - g^\dagger\|_{L^p((0, b-a))} \leq \|u - g\|_{L^p((a, b))}, \\ \text{TV}(\bar{u}; (a, b)) &= \text{TV}(u^\dagger; (0, b-a)) \leq \text{TV}(u; (a, b)), \end{aligned}$$

and both inequalities must be saturated. By the equality conditions stated in Proposition 2.12, the latter forces u to be monotone. We wish to prove that u is actually non-decreasing. By contradiction, suppose there exist $x_3 < x_4$ such that $u(x_3) > u(x_4)$. Since u is monotone, this

forces it to be non-increasing. Let $x' = \min\{x_1, x_3\}$ and $x'' = \max\{x_2, x_4\}$. Then for any points $y_1 \leq x'$ and $y_2 \geq x''$ we have

$$\begin{aligned} g(y_1) &\leq g(x') \leq g(x_1) < g(x_2) \leq g(x'') \leq g(y_2), \\ u(y_1) &\geq u(x') \geq u(x_1) > u(x_2) \geq u(x'') \geq u(y_2), \end{aligned}$$

from which $(u(y_2) - u(y_1))(g(y_2) - g(y_1)) < 0$. In the case $p > 1$, this fact contradicts the equality condition (2.4a).

We now treat separately the case $p = 1$. By Proposition 2.16, we have

$$m := \inf_{(a,b)} g \leq u \leq \sup_{(a,b)} g =: M.$$

In the limit $y_1 \rightarrow a^+$ and $y_2 \rightarrow b^-$ we have $g(y_1) < u(y_1)$ and $u(y_2) < g(y_2)$ because

$$\begin{aligned} \lim_{y_1 \rightarrow a^+} (u(y_1) - g(y_1)) &= \lim_{y_1 \rightarrow a^+} [(u(y_2) - g(y_1)) + (u(y_1) - u(y_2))] \\ &\geq (u(x_1) - u(x_2)) + \lim_{y_1 \rightarrow a^+} (u(y_2) - g(y_1)) \\ &\geq (u(x_1) - u(x_2)) + (u(y_2) - m) \geq u(x_1) - u(x_2) > 0 \end{aligned}$$

and similarly

$$\lim_{y_2 \rightarrow b^-} (g(y_2) - u(y_2)) \geq u(x_1) - u(x_2) > 0.$$

Therefore there is a non-negligible set of points y_1, y_2 which violate the inequality (2.4b). $\square$

The next lemma contains the explicit calculation of an integral which will be useful to estimate the fidelity term from below in Lemma 2.19.

Lemma 2.18. *Let $a \in \mathbb{R}$ and $h, C, \alpha \in \mathbb{R}_+$. Then*

$$\int_a^{a+(h/C)^{1/\alpha}} (h - C(x - a)^\alpha)^p dx = K(p, \alpha) C^{-1/\alpha} h^{p+1/\alpha},$$

where

$$K(p, \alpha) = \frac{B(p+1, 1/\alpha)}{\alpha} = \frac{\Gamma(p+1)\Gamma(1/\alpha)/\alpha}{\Gamma(p+1/\alpha+1)} = \frac{\Gamma(p+1)\Gamma(1/\alpha+1)}{\Gamma(p+1/\alpha+1)} = \binom{p+1/\alpha}{p}^{-1}$$

and

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt = \Gamma(x)\Gamma(y)/\Gamma(x+y) \quad (2.6)$$

is the Beta function.

Proof. After a translation by a and with the substitution $y = \frac{C}{h}x^\alpha$, $dx = \frac{(h/C)^{1/\alpha}}{\alpha} y^{1/\alpha-1} dy$, we have

$$\begin{aligned} \int_a^{a+(h/C)^{1/\alpha}} (h - C(x - a)^\alpha)^p dx &= \int_0^{(h/C)^{1/\alpha}} (h - Cx^\alpha)^p dx = h^p \int_0^{(h/C)^{1/\alpha}} \left(1 - \frac{C}{h}x^\alpha\right)^p dx \\ &= \frac{h^p(h/C)^{1/\alpha}}{\alpha} \int_0^1 (1-y)^p y^{1/\alpha-1} dy \\ &= \frac{h^p(h/C)^{1/\alpha}}{\alpha} B(p+1, 1/\alpha). \end{aligned} \quad \square$$

Lemma 2.19 (Fidelity lower bound). *Let $g : (a, b) \rightarrow \mathbb{R}$ be a measurable function for which there exists $C > 0$ such that $g(x) - g^+(a) \leq C(x - a)^\alpha$ for every $x \in (a, b)$, and let $u : (a, b) \rightarrow \mathbb{R}$ be a measurable function such that*

$$h := \inf_{(a,b)} u - g^+(a) \geq 0.$$

Then

$$\|u - g\|_{L^p((a,b))}^p \geq K(p, \alpha) C^{-1/\alpha} h^{p+1/\alpha},$$

where

$$K(p, \alpha) = \frac{B(p+1, 1/\alpha)}{\alpha} = \frac{\Gamma(p+1)\Gamma(1/\alpha)/\alpha}{\Gamma(p+1/\alpha+1)} = \frac{\Gamma(p+1)\Gamma(1/\alpha+1)}{\Gamma(p+1/\alpha+1)} = \binom{p+1/\alpha}{p}^{-1}.$$

If g satisfies also $g(x) - g^-(b) \geq -C(b - x)^\alpha$ for every $x \in (a, b)$ and

$$k := g^-(b) - \sup_{(a,b)} u \geq 0,$$

then

$$\|u - g\|_{L^p((a,b))}^p \geq K(p, \alpha) C^{-1/\alpha} (h^{p+1/\alpha} + k^{p+1/\alpha}). \quad (2.7)$$

Proof. In the interval $(a, a + (h/C)^{1/\alpha})$ we have

$$g(x) \leq g^+(a) + C(x - a)^\alpha \leq g^+(a) + h \leq u(x),$$

therefore, by Lemma 2.18,

$$\begin{aligned} \|u - g\|_{L^p((a,b))}^p &\geq \int_a^{a+(h/C)^{1/\alpha}} (u(x) - g(x))^p dx \geq \int_a^{a+(h/C)^{1/\alpha}} (h - C(x - a)^\alpha)^p dx \\ &\geq K(p, \alpha) C^{-1/\alpha} h^{p+1/\alpha}. \end{aligned}$$

Suppose that the additional assumptions on g and u are satisfied. Letting $U(x) = g^+(a) + C(x - a)^\alpha$ and $L(x) = g^-(b) - C(b - x)^\alpha$, we have

$$U(a + (h/C)^\alpha) = \inf u \leq \sup u = L(b - (k/C)^{1/\alpha}) \leq g(b - (k/C)^{1/\alpha}) \leq U(b - (k/C)^{1/\alpha}),$$

hence $a + (h/C)^{1/\alpha} \leq b - (k/C)^{1/\alpha}$. The interval $(b - (k/C)^{1/\alpha}, b)$ is disjoint from the previous one and in it we have

$$u(x) \leq g^-(b) - k \leq g^-(b) - C(b - x)^\alpha \leq g(x),$$

therefore

$$\begin{aligned} \|u - g\|_{L^p((a,b))}^p &\geq \int_a^{a+(h/C)^{1/\alpha}} (u(x) - g(x))^p dx + \int_{b-(k/C)^{1/\alpha}}^b (g(x) - u(x))^p dx \\ &\geq \int_a^{a+(h/C)^{1/\alpha}} (h - C(x - a)^\alpha)^p dx + \int_{b-(k/C)^{1/\alpha}}^b (k - C(b - x)^\alpha)^p dx \\ &\geq K(p, \alpha) C^{-1/\alpha} (h^{p+1/\alpha} + k^{p+1/\alpha}). \quad \square \end{aligned}$$

Definition 2.20. Given $n \in \mathbb{N}_+$, an exponent $p > 0$ and non-negative real numbers $x_1, \dots, x_n \in [0, \infty)$, their p -mean is

$$M_p(x_1, \dots, x_n) = \left(\frac{1}{n} \sum_{i=1}^n x_i^p \right)^{1/p}.$$

If is well known that $M_p \leq M_q$ for $p < q$. Moreover, clearly $M_p(x_1, \dots, x_n) = 0$ if and only if $x_1 = \dots = x_n = 0$.

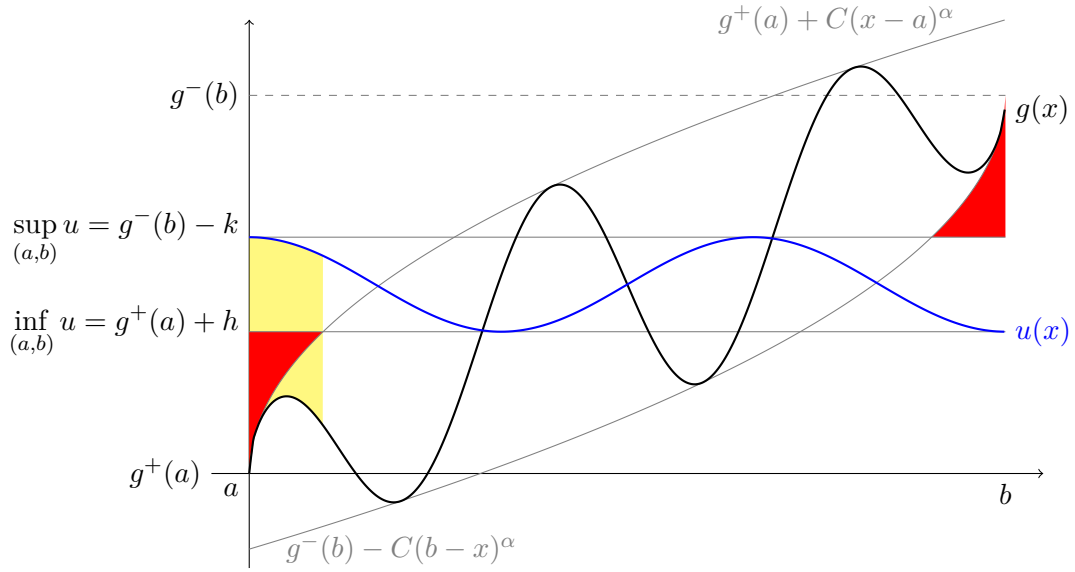


Figure 1: depiction of the situation described in Lemma 2.19.

2.3 Perspective of a convex function

Let $\varphi : [0, \infty) \rightarrow \mathbb{R}$ be a convex function which is continuous in 0.²

The *perspective* of φ is $\varphi : [0, \infty) \times (0, \infty) \rightarrow \mathbb{R}$ given by $\varphi(t, s) = s\varphi(t/s)$. It is clearly continuous, and it is also convex (see [BV23, Section 3.2.6]).

It is well known that the difference quotient function $t \mapsto \frac{\varphi(t) - \varphi(0)}{t}$ is non-decreasing. As a consequence, we can define the *asymptotic slope*

$$\varphi'(\infty) = \lim_{t \rightarrow \infty} \frac{\varphi(t)}{t} = \lim_{t \rightarrow \infty} \frac{\varphi(t) - \varphi(0)}{t} = \sup_{t > 0} \frac{\varphi(t) - \varphi(0)}{t} \in \mathbb{R} \cup \{\infty\}$$

and the *recession function* $\varphi^\infty : [0, \infty) \rightarrow (-\infty, \infty]$

$$\varphi^\infty(t) = \lim_{r \rightarrow \infty} \frac{\varphi(rt)}{r} = \lim_{r \rightarrow \infty} \frac{\varphi(rt) - \varphi(0)}{r} = \sup_{r > 0} \frac{\varphi(rt) - \varphi(0)}{r} \in \mathbb{R} \cup \{\infty\}.$$

It is immediate to verify that $\varphi^\infty(t) = \varphi'(\infty)t$ for every $t \in [0, \infty)$, with the usual convention $\infty \cdot 0 = 0$. Indeed for $t = 0$ we have $\varphi^\infty(0) = 0 = \varphi'(\infty)0$, and for $t > 0$ we have

$$\varphi^\infty(t) = \lim_{r \rightarrow \infty} \frac{\varphi(rt)}{r} = t \lim_{r \rightarrow \infty} \frac{\varphi(rt)}{rt} = t \lim_{r \rightarrow \infty} \frac{\varphi(r)}{r} = t\varphi'(\infty).$$

Moreover, if $\varphi(0) = 0$, then $\varphi(t, s)$ is non-increasing in s : indeed,

$$\frac{1}{t}\varphi(t, s) = \frac{s}{t}\varphi(t/s) = \frac{\varphi(t/s) - \varphi(0)}{t/s}$$

is non-decreasing in t/s , therefore non-increasing in s .

Lemma 2.21. *Using the notation $\varphi_s = \varphi(\cdot, s)$ to denote the perspective, we have the following rules of commutation with the Legendre transform: for every $s > 0$*

$$(\varphi_s)^* = s\varphi^*, \quad (\varphi^*)_s = (s\varphi)^*.$$

²One may extend it to $\varphi : \mathbb{R} \rightarrow \mathbb{R} \cup \{\infty\}$ by setting $\varphi|_{(-\infty, 0)} = \infty$, but of course this extension is no longer continuous in 0.

Proof. Straightforward computations show that, for every $\tau \in \mathbb{R}$,

$$\begin{aligned} (\varphi_s)^*(\tau) &= \sup_{t \geq 0} (\tau t - \varphi_s(t)) = \sup_{t \geq 0} (\tau t - s\varphi(t/s)) \\ &= s \cdot \sup_{t \geq 0} (\tau t/s - \varphi(t/s)) = s \cdot \sup_{u \geq 0} (\tau u - \varphi(u)) = s\varphi^*(\tau) \end{aligned}$$

and

$$(\varphi^*)_s(\tau) = s\varphi^*(\tau/s) = s \cdot \sup_{t \geq 0} ((\tau/s)t - \varphi(t)) = \sup_{t \geq 0} (\tau t - s\varphi(t)) = (s\varphi)^*(\tau). \quad \square$$

Proposition 2.22. *If $\varphi'(\infty) \in \mathbb{R}$ is finite, then the perspective $\varphi : [0, \infty) \times (0, \infty) \rightarrow \mathbb{R}$ can be extended continuously to $\varphi : [0, \infty) \times [0, \infty) \rightarrow \mathbb{R}$ setting $\varphi(t, 0) = \varphi^\infty(t) = \varphi'(\infty)t$ and this extension is convex.*

“Specific” proof. ³ Since for every $t \in [0, \infty)$ we have the directional limit

$$\lim_{s \rightarrow 0^+} \varphi(t, s) = \lim_{s \rightarrow 0^+} s\varphi(t/s) = \lim_{r \rightarrow \infty} \varphi(rt)/r = \varphi^\infty(t),$$

setting $\varphi(t, 0) = \varphi^\infty(t) = \varphi'(\infty)t$ is the only possible choice that may render the extension continuous. Let us verify that it is indeed the case.

From $\frac{\varphi(t) - \varphi(0)}{t} \leq \varphi'(\infty)$ we get the upper bound $\varphi(t) \leq \varphi'(\infty)t + \varphi(0)$, valid for every $t \in [0, \infty)$.

We claim that the Legendre transform $\varphi^*(\sigma) \in \mathbb{R}$ is finite for every $\sigma \in (-\infty, \varphi'(\infty))$.⁴ Indeed, fix $T \in [0, \infty)$ such that $\frac{\varphi(T) - \varphi(0)}{T} \geq \sigma$. Then for every $t \in [T, \infty)$ we have $\frac{\varphi(t) - \varphi(0)}{t} \geq \frac{\varphi(T) - \varphi(0)}{T} \geq \sigma$, hence $\varphi(t) \geq \sigma t + \varphi(0)$. This implies that

$$\begin{aligned} \varphi^*(\sigma) &= \sup_{t \in [0, \infty)} (\sigma t - \varphi(t)) = \max \left\{ \sup_{t \in [0, T]} (\sigma t - \varphi(t)), \sup_{t \in [T, \infty)} (\sigma t - \varphi(t)) \right\} \\ &= \max \left\{ \max_{t \in [0, T]} (\sigma t - \varphi(t)), \sup_{t \in [T, \infty)} (-\varphi(0)) \right\} < \infty \end{aligned}$$

is finite. This fact provides us with the lower bound $\varphi(t) \geq \sigma t - \varphi^*(\sigma)$, valid for every $t \in [0, \infty)$.

The lower bound and upper bound can be combined, evaluated at t/s and multiplied by s , yielding

$$\sigma t - \varphi^*(\sigma)s \leq s\varphi(t/s) \leq \varphi'(\infty)t + \varphi(0)s, \quad \forall (t, s) \in [0, \infty) \times (0, \infty).$$

Notice that both the lower and upper bounds are continuous functions of t and s . For a fixed $t_0 \in [0, \infty)$, we get

$$\limsup_{\substack{t \rightarrow t_0 \\ s \rightarrow 0^+}} \varphi(t, s) = \limsup_{\substack{t \rightarrow t_0 \\ s \rightarrow 0^+}} s\varphi(t/s) \leq \limsup_{\substack{t \rightarrow t_0 \\ s \rightarrow 0^+}} (\varphi'(\infty)t + \varphi(0)s) = \varphi'(\infty)t_0$$

and

$$\liminf_{\substack{t \rightarrow t_0 \\ s \rightarrow 0^+}} \varphi(t, s) = \liminf_{\substack{t \rightarrow t_0 \\ s \rightarrow 0^+}} s\varphi(t/s) \geq \liminf_{\substack{t \rightarrow t_0 \\ s \rightarrow 0^+}} (\sigma t - \varphi^*(\sigma)s) = \sigma t_0.$$

³This proof uses more specific properties of the perspective.

⁴Notice that in general it may be $\varphi^*(\varphi'(\infty)) = \infty$. This happens when $\varphi'(\infty)t - \varphi(t)$ does not go to zero as $t \rightarrow \infty$, for instance $\varphi(t) = o(t)$ but $\varphi(t) \not\rightarrow 0$. An explicit example is $\varphi(t) = -\sqrt{t}$.

By the arbitrariness of $\sigma < \varphi'(\infty)$ we deduce that the $\liminf$ is actually greater than or equal to $\varphi'(\infty)t_0$, which implies

$$\lim_{\substack{t \rightarrow t_0 \\ s \rightarrow 0^+}} \varphi(t, s) = \varphi'(\infty)t_0 = \varphi(t_0, 0),$$

hence the extension is continuous.

It is a general fact that a continuous extension is necessarily convex. □

“General” proof. ⁵ Since for every $t \in [0, \infty)$ we have the directional limit

$$\lim_{s \rightarrow 0^+} \varphi(t, s) = \lim_{s \rightarrow 0^+} s\varphi(t/s) = \lim_{r \rightarrow \infty} \varphi(rt)/r = \varphi^\infty(t),$$

setting $\varphi(t, 0) = \varphi^\infty(t) = \varphi'(\infty)t$ is the only possible choice that may render the extension continuous.

First of all, observe that the extension is convex. To this end, let $(t_0, s_0), (t_1, s_1) \in [0, \infty) \times [0, \infty)$, $\tau \in [0, 1]$ and $(t_\tau, s_\tau) = (1 - \tau)(t_0, s_0) + \tau(t_1, s_1)$. For every $\varepsilon > 0$ we have $(t_0, s_0 + \varepsilon), (t_1, s_1 + \varepsilon) \in [0, \infty) \times (0, \infty)$ and $(t_\tau, s_\tau + \varepsilon) = (1 - \tau)(t_0, s_0 + \varepsilon) + \tau(t_1, s_1 + \varepsilon)$, therefore

$$\varphi(t_\tau, s_\tau + \varepsilon) \leq (1 - \tau)\varphi(t_0, s_0 + \varepsilon) + \tau\varphi(t_1, s_1 + \varepsilon).$$

Sending $\varepsilon \rightarrow 0^+$ and using the directional continuity in s yields

$$\varphi(t_\tau, s_\tau) \leq (1 - \tau)\varphi(t_0, s_0) + \tau\varphi(t_1, s_1),$$

hence the extension is convex.

Let us verify that it is continuous. Notice that the extension $\varphi(t, 0) = \varphi'(\infty)t$ is a continuous function of t . Fix $t_0 \in (0, \infty)$ and $\varepsilon > 0$. There exists $\delta > 0$ such that $t_0 - \delta \geq 0$ and

$$|\varphi(t_0 \pm \delta, 0) - \varphi(t_0, 0)| \leq \varepsilon/2.$$

By the continuity in s , there exists $\eta > 0$ such that for every $s \in [0, \eta]$ we have

$$|\varphi(t_0, s) - \varphi(t_0, 0)| \leq \varepsilon/2, \quad |\varphi(t_0 \pm \delta, s) - \varphi(t_0 \pm \delta, 0)| \leq \varepsilon/2.$$

By the triangle inequality we deduce also $|\varphi(t_0 \pm \delta, s) - \varphi(t_0, 0)| \leq \varepsilon$. Let $(t, s) \in [t_0 - \delta, t_0 + \delta] \times [0, \eta]$ be an arbitrary point in this neighborhood of $(t_0, 0)$. By convexity we have the upper bound

$$\varphi(t, s) \leq \max\{\varphi(t_0 \pm \delta, 0), \varphi(t_0 \pm \delta, \eta)\} \leq \varphi(t_0, 0) + \varepsilon.$$

To find a lower bound, assume that $t_0 \leq t \leq t_0 + \delta$; the symmetric case $t_0 - \delta \leq t \leq t_0$ is analogous. We can write $t_0 = (1 - \tau)(t_0 - \delta) + \tau t$ with $\tau = \delta/(t - t_0 + \delta) \in [1/2, 1]$, therefore

$$\varphi(t_0, s) \leq (1 - \tau)\varphi(t_0 - \delta, s) + \tau\varphi(t, s),$$

from which

$$\begin{aligned} \varphi(t, s) &\geq \frac{1}{\tau}\varphi(t_0, s) - \frac{1 - \tau}{\tau}\varphi(t_0 - \delta, s) \geq \frac{1}{\tau}(\varphi(t_0, 0) - \varepsilon) - \frac{1 - \tau}{\tau}(\varphi(t_0, 0) + \varepsilon) \\ &= \varphi(t_0, 0) - \left(1 - \frac{2}{\tau}\right)\varepsilon \geq \varphi(t_0, 0) - 3\varepsilon. \end{aligned}$$

Combined with the upper bound, this shows that $|\varphi(t, s) - \varphi(t_0, 0)| \leq 3\varepsilon$ for every $(t, s) \in [t_0 - \delta, t_0 + \delta] \times [0, \eta]$, proving the continuity in $(t_0, 0)$, for $t_0 \in (0, \infty)$.

⁵This proof uses only the existence of the directional limit.

The continuity in $(0, 0)$ is proved similarly. Given $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|\varphi(\delta, 0) - \varphi(0, 0)| \leq \varepsilon/2, \quad |\varphi(2\delta, 0) - \varphi(0, 0)| \leq \varepsilon/2.$$

By the continuity in s , there exists $\eta > 0$ such that for every $s \in [0, \eta]$ we have

$$|\varphi(0, s) - \varphi(0, 0)| \leq \varepsilon/2, \quad |\varphi(\delta, s) - \varphi(\delta, 0)| \leq \varepsilon/2, \quad |\varphi(2\delta, s) - \varphi(2\delta, 0)| \leq \varepsilon/2.$$

By the triangle inequality we infer $|\varphi(i\delta, s) - \varphi(0, 0)| \leq \varepsilon$ for $i = 0, 1, 2$. If $(t, s) \in [0, 2\delta] \times [0, \eta]$ is an arbitrary point in this neighborhood of the origin, with a similar argument as before we get $|\varphi(t, s) - \varphi(0, 0)| \leq 3\varepsilon$, hence proving the continuity. $\square$

3 Linear TV

In this section we study the minimization of the functional

$$F(u) = \text{TV}(u) + \lambda \|u - g\|_p^\gamma$$

with linear total variation and fidelity with respect to an input datum g . We classify situations where the minimizer u is not the input g itself (Subsection 3.1) and situations where the minimizer u is indeed the input g (Subsection 3.2).

Our results are based on some conditions on g (α -growth in Definition 2.10, monotonicity, piecewise monotonicity) and on specific ranges of the parameters p, γ, λ . The complementarity of these ranges provides a full characterization and demonstrates the sharpness of our results (see Remark 3.11).

3.1 The input is not a minimizer

We start with a model situation where the input g is taken to be the Cantor–Lebesgue–Vitali function $\chi \in C([0, 1]; [0, 1])$. Subsequently, this example will be generalized in Theorem 3.4 to functions with at least α -growth at one endpoint (see Definition 2.10).

Theorem 3.1. *Let $\chi \in C([0, 1]; [0, 1])$ be the Cantor–Lebesgue–Vitali function and $\lambda, \gamma > 0$, $p \in [1, \infty]$. If*

$$\gamma > \left(\frac{1}{p} \cdot \frac{\log 3}{\log 2} + 1 \right)^{-1}$$

then the function χ is not a minimizer of the functional $F(u) = \text{TV}(u) + \lambda \|u - \chi\|_p^\gamma$.

Proof. Given $n \in \mathbb{N}_+$ let $\chi_n = 2^{-n} \vee \chi \wedge (1 - 2^{-n}) \in C([0, 1]; [2^{-n}, 1 - 2^{-n}])$. We have $F(\chi) = \text{TV}(\chi) = 1$ and $\text{TV}(\chi_n) = 1 - 2^{-n+1}$. Moreover, for $p = \infty$ we have $\|\chi_n - \chi\|_p = 2^{-n}$, whereas for $p \in [1, \infty)$ we have

$$\begin{aligned} \|\chi_n - \chi\|_p^p &= 2 \int_0^{3^{-n}} \chi(x)^p dx \leq 2 \int_0^{3^{-n}} \chi(x)^{p-1} \chi(x) dx \leq 2 \|\chi\|_{L^\infty([0, 3^{-n}])}^{p-1} \int_0^{3^{-n}} \chi(x) dx \\ &= 2 \cdot 2^{-n(p-1)} \frac{1}{2} 2^{-n} 3^{-n} = 2^{-np} 3^{-n}, \end{aligned}$$

hence $\|\chi_n - \chi\|_p \leq 2^{-n} 3^{-n/p}$ in both cases (interpreting $-n/p = 0$ for $p = \infty$). From this we get that

$$F(\chi_n) \leq 1 - 2^{-n+1} + \lambda 2^{-n\gamma} 3^{-n\gamma/p},$$

therefore $F(\chi_n) < F(\chi) \Leftrightarrow \lambda 2^{-n\gamma} 3^{-n\gamma/p} < 2^{-n+1}$, i.e. if and only if

$$\frac{\lambda}{2} < 2^{(\gamma-1)n} 3^{\frac{2}{p}n} = \exp\left(\left((\gamma-1)\log 2 + \frac{\gamma}{p}\log 3\right)n\right).$$

For n large enough this is satisfied if

$$(\gamma-1)\log 2 + \frac{\gamma}{p}\log 3 > 0 \Leftrightarrow (1-\gamma)\log 2 < \frac{\gamma}{p}\log 3 \Leftrightarrow \frac{1-\gamma}{\gamma} < \frac{1}{p} \cdot \frac{\log 3}{\log 2} \Leftrightarrow \frac{1}{\gamma} < \frac{1}{p} \cdot \frac{\log 3}{\log 2} + 1. \quad \square$$

Remark 3.2. The previous statement holds true for the functional

$$F(u) = \text{TV}_{\varphi,\theta}(u) + \lambda \|u - \chi\|_p^\gamma$$

where

$$\text{TV}_{\varphi,\theta}(u) = \int_0^1 \varphi(|u'|) + |D^c u|((0,1)) + \sum_{J_u} \theta(|u^+ - u^-|)$$

is the non-linear total variation.⁶ This is because $\text{TV}_{\varphi,\theta}$ coincides with TV for both χ and the competitor χ_n .

The next lemma is used in the proof of [Theorem 3.4](#). We will need it only with $\ell = \sup_{(a,c]} g$, but we prefer to give a more general statement, which might be of broader interest.

Lemma 3.3. *Let $g \in \text{BV}((a,b))$, $c \in (a,b)$ and*

$$\inf_{(a,c]} g \leq \ell \leq \sup_{(a,c]} g.$$

Define $u : (a,b) \rightarrow \mathbb{R}$ such that

$$u(x) = \begin{cases} \ell & x \in (a,c), \\ g(x) & x \in (c,b), \end{cases}$$

and u is regular in c . Then $u \in \text{BV}((a,b))$ and

$$\text{TV}(g; (a,b)) \geq \text{TV}(u; (a,b)) + |\ell - g^+(a)|.$$

Proof. Since u is constant in (a,c) , is regular in c , and $u = g$ in (c,b) , we have

$$\text{TV}(u; (a,b)) = \text{TV}(u; (a,c)) + |u^-(c) - u^+(c)| + \text{TV}(u; (c,b)) = |\ell - g^+(c)| + \text{TV}(g; (c,b)),$$

hence $u \in \text{BV}((a,b))$. Let $m = \inf_{(a,c]} g$ and $M = \sup_{(a,c]} g$. For every $\varepsilon > 0$ there exists $x'_\varepsilon, x''_\varepsilon \in (a,c]$ with $x'_\varepsilon \leq x''_\varepsilon$ such that

$$|g(x'_\varepsilon) - m| \leq \varepsilon \quad \wedge \quad |g(x''_\varepsilon) - M| \leq \varepsilon \quad \text{or} \quad |g(x'_\varepsilon) - M| \leq \varepsilon \quad \wedge \quad |g(x''_\varepsilon) - m| \leq \varepsilon.$$

Let us treat the first alternative; the second is analogous. Exploiting $|m - M| = |m - \ell| + |\ell - M|$, we can estimate

$$\begin{aligned} & \text{TV}(g; (a,b)) \\ &= \text{TV}(g; (a, x'_\varepsilon]) + \text{TV}(g; [x'_\varepsilon, x''_\varepsilon]) + \text{TV}(g; [x''_\varepsilon, c]) + \text{TV}(g; (c, b)) \\ &\geq |g^+(a) - g(x'_\varepsilon)| + |g(x'_\varepsilon) - g(x''_\varepsilon)| + |g(x''_\varepsilon) - g(c)| + |g(c) - g^+(c)| + \text{TV}(g; (c, b)) \\ &\geq |g^+(a) - m| + |m - M| + |M - g(c)| + |g(c) - g^+(c)| + \text{TV}(g; (c, b)) - 4\varepsilon \\ &= |g^+(a) - m| + |m - \ell| + |\ell - M| + |M - g(c)| + |g(c) - g^+(c)| + \text{TV}(g; (c, b)) - 4\varepsilon \\ &\geq |g^+(a) - \ell| + |\ell - g^+(c)| + \text{TV}(g; (c, b)) - 4\varepsilon \\ &= |g^+(a) - \ell| + \text{TV}(u; (a,b)) - 4\varepsilon. \end{aligned}$$

The thesis follows letting $\varepsilon \rightarrow 0^+$. $\square$

⁶The condition $\lim_{t \rightarrow \infty} \varphi(t)/t = \lim_{t \rightarrow 0^+} \theta(t)/t = 1$ is not required.

Theorem 3.4. Let $g \in \text{BV}((a, b))$ be a function with the property that there exist $\alpha, c > 0$ and $0 < \delta < b - a$ such that $g(x) - g^+(a) \geq c(x - a)^\alpha$ for all $x \in (a, a + \delta]$. If

$$\gamma > \left(1 + \frac{1}{p\alpha}\right)^{-1} = \frac{p\alpha}{p\alpha + 1} \quad (3.1)$$

then g is not a minimizer of the functional

$$F(u) = \text{TV}(u) + \lambda \|u - g\|_p^\gamma$$

for every $p \in [1, \infty]$ and $\lambda > 0$.

If instead equality holds in (3.1), the same conclusion is true provided that

$$\lambda < K(p, \alpha)^{-\gamma/p} c^{\gamma/(p\alpha)} = K(p, \alpha)^{-\alpha/(1+p\alpha)} c^{1/(1+p\alpha)}. \quad (3.2)$$

Proof. Define the non-decreasing function $\omega : (a, a + \delta] \rightarrow \mathbb{R}$

$$\omega(x) = \sup_{(a, x]} g$$

and observe that we have $\omega(x) \geq g(x) \geq g^+(a) + c(x - a)^\alpha$ and $\omega^+(a) = g^+(a)$. Given $y \in (a, a + \delta]$, introduce the auxiliary parameter $h = h(y) = \omega(y) - g^+(a) \geq c(y - a)^\alpha > 0$ and define the truncated function $u_y \in \text{BV}((a, b))$

$$u_y(x) = \begin{cases} \omega(y) & x \in (a, y], \\ g(x) & x \in (y, b), \end{cases}$$

which, thanks to Lemma 3.3 applied with $c = y$ and $\ell = \omega(y)$, satisfies

$$\text{TV}(g, (a, b)) \geq \text{TV}(u_y, (a, b)) + |\omega(y) - g^+(a)| = \text{TV}(u_y, (a, b)) + h.$$

For every $x \in (a, y]$ we have $g^+(a) + c(x - a)^\alpha \leq g(x) \leq \omega(y) = g^+(a) + h = u_y(x)$, therefore $|u_y(x) - g(x)| = u_y(x) - g(x) \leq h - c(x - a)^\alpha$. Moreover from $c(y - a)^\alpha \leq h$ we deduce $y \leq a + (h/c)^{1/\alpha}$, hence by Lemma 2.18 we get

$$\|u_y - g\|_{L^p((a, b))}^p \leq \int_a^y (h - c(x - a)^\alpha)^p dx \leq \int_a^{a + (h/c)^{1/\alpha}} (h - c(x - a)^\alpha)^p dx = K(p, \alpha) c^{-1/\alpha} h^{p+1/\alpha}.$$

Combining the estimates for the total variation and fidelity term we obtain

$$F(u_y) - F(g) \leq -h + \lambda K(p, \alpha)^{\gamma/p} c^{-\gamma/(p\alpha)} h^{\gamma(1 + \frac{1}{p\alpha})},$$

which is negative for y sufficiently small because $h(y) \rightarrow 0^+$ for $y \rightarrow a^+$ and the exponent of h in the last term is greater than 1 by (3.1), hence the negative linear term dominates.

If instead γ satisfies the equality in (3.1), the exponent of h is 1 and the negativity follows from (3.2). $\square$

The following remark highlights a strong structural property of the minimizers of F when the input is monotone.

Remark 3.5. Assume that the input g is non-decreasing and let u be a minimizer of $F(u) = \text{TV}(u) + \lambda \|u - g\|_p^\gamma$; u must be non-decreasing. Define

$$\begin{aligned} x' &= \sup\{x \in (a, b) : g \leq u^+(a) \text{ in } (a, x)\}, \\ x'' &= \inf\{x \in (a, b) : g \geq u^-(b) \text{ in } (x, b)\}, \end{aligned}$$

and the non-decreasing regular function

$$\bar{u}(x) = \begin{cases} u^+(a) & x \in (a, x'], \\ g(x) & x \in (x', x''), \\ u^-(b) & x \in [x'', b]. \end{cases}$$

We have $\text{TV}(\bar{u}) = u^-(b) - u^+(a) = \text{TV}(u)$. In (a, x') we have $g \leq \bar{u} \leq u$, whereas in (x'', b) the reverse inequalities hold, therefore

$$\|\bar{u} - g\|_{L^p((a, x'))} \leq \|u - g\|_{L^p((a, x'))}, \quad \|\bar{u} - g\|_{L^p((x'', b))} \leq \|u - g\|_{L^p((x'', b))}.$$

In (x', x'') we have $\bar{u} = g$, therefore

$$\|\bar{u} - g\|_{L^p((x', x''))} = 0 \leq \|u - g\|_{L^p((x', x''))}.$$

In all three cases, equality holds if and only if $u = \bar{u}$ in the respective intervals. This shows that $F(\bar{u}) \leq F(u)$, with equality if and only if $u = \bar{u}$ in (a, b) .

This proves that all minimizers must be of the form $\bar{u}$, although we do not show uniqueness because the values at the endpoints $u^+(a)$ and $u^-(b)$ might not be uniquely determined.

The main content of [Theorem 3.4](#) is showing that $u^+(a) > g^+(a)$, hence a minimizer cannot coincide with g .

Theorem 3.6. *Let $g \in \text{BV}((a, b))$ be a non-decreasing function and let $p \in [1, \infty]$. If $0 < \lambda < (b - a)^{-1/p}$ then every minimizer $u \in \text{BV}((0, 1))$ of the functional*

$$F(u) = \text{TV}(u) + \lambda \|u - g\|_p$$

is a constant function. Moreover, the minimizer is unique if $p > 1$.

Proof. We treat the case $p < \infty$; the case $p = \infty$ follows from a straightforward adaptation. Let $u \in \text{BV}((a, b))$ be a minimizer. By [Proposition 2.17](#) u is non-decreasing. Moreover, by [Proposition 2.16](#) we have $g^+(a) \leq u \leq g^-(b)$.

Let $c \in [u^+(a), u^-(b)]$ and define $h = c - u^+(a) \geq 0$ and $k = u^-(b) - c \geq 0$. Using the fact that

$$\|u - c\|_p \leq (b - a)^{1/p} \|u - c\|_\infty = (b - a)^{1/p} \max\{h, k\} \leq (b - a)^{1/p} (h + k)$$

we can estimate

$$\begin{aligned} F(u) - F(c) &= \text{TV}(u) - \text{TV}(c) + \lambda (\|u - g\|_p - \|c - g\|_p) \geq u^-(b) - u^+(a) - \lambda \|u - c\|_p \\ &\geq h + k - \lambda (b - a)^{1/p} (h + k) = (1 - \lambda (b - a)^{1/p}) (h + k). \end{aligned}$$

Under the assumption on λ , the inequality $F(u) \geq F(c)$ is strict unless $h + k = 0$, i.e. $h = k = 0$, that is $u \equiv c$. This proves that every minimizer is constant.

If $p > 1$, then $F(c) = \lambda \|c - g\|_p$ is strictly convex, therefore the minimizer is unique. $\square$

3.2 The input is the minimizer

In this section we find conditions that ensure that the input g is the minimizer. This has the interesting consequence that, under the assumptions of either [Theorem 3.7](#), [Theorem 3.8](#) or [Theorem 3.10](#), if $D^c g \neq 0$, then also $D^c u \neq 0$ where u is the minimizer of $F(u) = \text{TV}(u) + \lambda \|u - g\|_p^\gamma$, showing that in general minimizers can have the Cantor part of the derivative.

Theorem 3.7. *Let $g \in \text{BV}((0, 1); [0, 1])$ be a non-decreasing function with $g \equiv 0$ on $(0, \delta)$ and $g \equiv 1$ on $(1 - \delta, 1)$ for some $\delta \in (0, 1/2)$, and let $\lambda, \gamma > 0$, $p \in [1, \infty]$. If $\gamma \leq 1$ and $\lambda\delta^{\gamma/p} > 2$ then $u = g$ is the unique minimizer of the functional $F : \text{BV}((0, 1)) \rightarrow \mathbb{R}$ given by*

$$F(u) = \text{TV}(u) + \lambda\|u - g\|_p^\gamma.$$

Proof. Let $u \in \text{BV}((0, 1))$ be a minimizer. By [Proposition 2.17](#) u is non-decreasing. Moreover we have $0 \leq u \leq 1$ because $F(0 \vee u \wedge 1) \leq F(u)$ with strict inequality otherwise.

We claim that $u \equiv u(0^+)$ on $(0, \delta)$ and $u \equiv u(1^-)$ on $(1 - \delta, 1)$. In fact if this is not the case then the function

$$\tilde{u}(x) = \begin{cases} u(0^+) & x \in (0, \delta) \\ u(x) & x \in (\delta, 1 - \delta) \\ u(1^-) & x \in (1 - \delta, 1) \end{cases}$$

has the property $\text{TV}(\tilde{u}) \leq \text{TV}(u)$ and $\|\tilde{u} - g\|_p < \|u - g\|_p$.

We have the estimates $\|u - g\|_\infty \geq \max\{u(0^+), 1 - u(1^-)\}$ and

$$\begin{aligned} \|u - g\|_p^p &\geq \int_{(0, \delta) \cup (1 - \delta, 1)} |u(x) - g(x)|^p dx \\ &= \int_0^\delta u(0^+)^p dx + \int_{1 - \delta}^1 (1 - u(1^-))^p dx = \delta[u(0^+)^p + (1 - u(1^-))^p]. \end{aligned}$$

Letting $h = u(0^+)$ and $k = 1 - u(1^-)$, these can be summarized as $\|u - g\|_p \geq \delta^{1/p} M_p(h, k)$ where M_p is the p -mean.⁷ Notice that $M_p(h, k) \leq 1$, therefore $M_p(h, k)^\gamma \geq M_p(h, k) \geq M_1(h, k)$ for $\gamma \in (0, 1]$. Using these estimates and $F(g) = 1$ we can bound from below

$$F(u) = u(1^-) - u(0^+) + \lambda\|u - g\|_p^\gamma \geq 1 - (h + k) + \lambda\delta^{\gamma/p} M_p(h, k)^\gamma \geq F(g) + (\lambda\delta^{\gamma/p} - 2)M_1(h, k)$$

and this inequality is strict unless $M_1(h, k) = 0$, i.e. $h = k = 0$, that is $u(0^+) = 0$ and $u(1^-) = 1$.

In this case $\text{TV}(u) = u(1^-) - u(0^+) = 1$, hence $F(u) = 1 + \lambda\|u - g\|_p^\gamma = F(g) + \lambda\|u - g\|_p^\gamma$ and for u to be a minimizer we must have $u = g$. $\square$

The next theorem treats the case of a monotone input. This will be later generalized to piecewise monotone functions in [Theorem 3.10](#), with the additional difficulty of having to deal with the jumps between the subintervals of monotonicity.

Theorem 3.8. *Let $g \in \text{BV}_{\text{reg}}((a, b))$ be a non-decreasing function with at most α -growth at the endpoints a and b , i.e. there exist $\alpha > 0$ and $C > 0$ such that $g(x) - g^+(a) \leq C(x - a)^\alpha$ and $g^-(b) - g(x) \leq C(b - x)^\alpha$ for all $x \in (a, b)$. If the conditions*

$$\gamma \leq \left(1 + \frac{1}{p\alpha}\right)^{-1} = \frac{p\alpha}{p\alpha + 1} \quad \text{and} \quad \lambda > 2^{\frac{(p-1)\alpha+1}{p\alpha+1}} K(p, \alpha)^{-\gamma/p} C^{\gamma/(p\alpha)} H^{1-\gamma(1+1/(p\alpha))}$$

are satisfied, then g is essentially the unique minimizer of the functional

$$\begin{aligned} F : \text{BV}((x_0, x_n)) &\rightarrow [0, \infty) \\ F(u) &= \text{TV}(u) + \lambda\|u - g\|_p^\gamma, \end{aligned}$$

in the sense that any other minimizer coincides with g outside of the jump set of g .

⁷See [Definition 2.20](#).

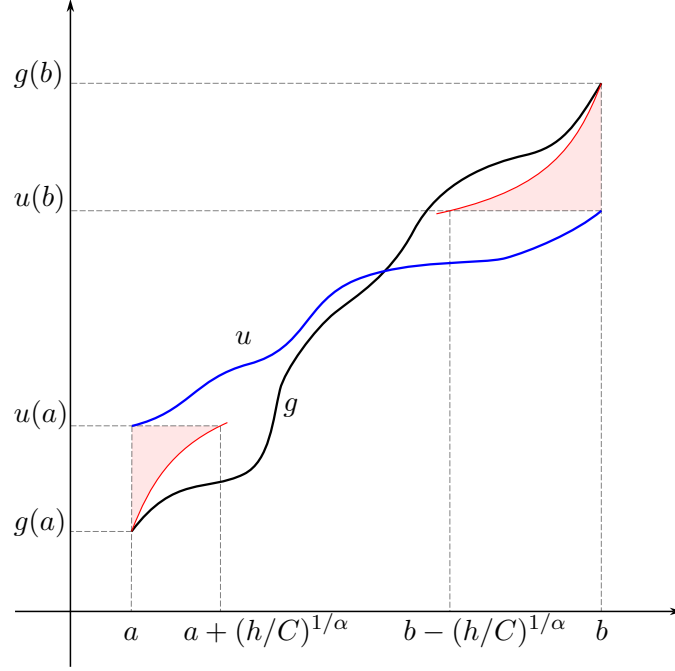


Figure 2: Visual aid for the proof of [Theorem 3.8](#). The two red arcs are $g(a) + C(x-a)^\alpha$ and $g(b) - C(b-x)^\alpha$, which bound g from above and below respectively. The shaded regions are the two contributions considered in the estimate of $\|u - g\|_p^p$.

Proof. Let u be a minimizer and let $H = g^-(b) - g^+(a)$. From [Proposition 2.16](#) and [Proposition 2.17](#) we know that u is non-decreasing and $g^+(a) \leq u(x) \leq g^-(b)$ for every $x \in (a, b)$. Let $u^+(a) = g^+(a) + h$ and $u^-(b) = g^-(b) - k$ with $0 \leq h, k \leq H$. Our first goal is to prove that $h = k = 0$. With regard to the total variation, we have $\text{TV}(u) = u^-(b) - u^+(a) = g^-(b) - g^+(a) - (h + k) = \text{TV}(g) - (h + k)$. With regard to the fidelity term, from [Lemma 2.19](#) we have that the estimate (2.7) holds, that is

$$\|u - g\|_p^p \geq K(p, \alpha) C^{-1/\alpha} (h^{p+1/\alpha} + k^{p+1/\alpha}).$$

For convenience, introduce the parameter $\beta = p + 1/\alpha > 1$ appearing as exponent in the previous estimate. Recalling that $h, k \leq H$, if $\gamma \leq p/\beta = \frac{p\alpha}{p\alpha+1}$ by the inequalities between means we have

$$h + k \leq 2^{1-1/\beta} (h^\beta + k^\beta)^{1/\beta}, \quad (h^\beta + k^\beta)^{\gamma/p} \geq H^{\beta\gamma/p-1} (h^\beta + k^\beta)^{1/\beta}.$$

Combining these estimates leads to

$$\begin{aligned} 0 &\geq F(u) - F(g) \geq -(h + k) + \lambda K(p, \alpha)^{\gamma/p} C^{-\gamma/(p\alpha)} (h^\beta + k^\beta)^{\gamma/p} \\ &\geq \left[\lambda K(p, \alpha)^{\gamma/p} C^{-\gamma/(p\alpha)} H^{\beta\gamma/p-1} - 2^{1-1/\beta} \right] (h^\beta + k^\beta)^{1/\beta}. \end{aligned}$$

The assumption on γ implies that the coefficient in the square brackets is positive, therefore we must have $h = k = 0$.

Now that we have $u^+(a) = g^+(a)$ and $u^-(b) = g^-(b)$, we conclude by noticing that $\text{TV}(u) = \text{TV}(g)$, hence $F(u) - F(g) = \lambda \|u - g\|^\gamma \geq 0$, with equality if and only if $u = g$. $\square$

Proposition 3.9. *Let $x_0 < \dots < x_n$ and let $g \in \text{BV}((x_0, x_n))$ be a regular function which is monotone in every interval $I_i = (x_{i-1}, x_i)$. If the conditions*

$$p \geq 1, \quad \gamma \leq 1, \quad \lambda > 2/G(I_i) \quad \forall i \in \{1, \dots, n\},$$

are satisfied, where for every $J \subseteq I := (x_0, x_n)$

$$G(J) = \frac{\gamma}{(H\mathcal{L}^1(I)^{1/p})^{1-\gamma}} \begin{cases} \mathcal{L}^1(J)^{1/p} & p = 1 \text{ or } g|_J \text{ constant,} \\ p \left(\frac{\|g - \inf_J g\|_{L^{p-1}(J)}}{\|g - \inf_J g\|_{L^p(I)}} \wedge \frac{\|g - \sup_J g\|_{L^{p-1}(J)}}{\|g - \sup_J g\|_{L^p(I)}} \right)^{p-1} & \text{otherwise,} \end{cases}$$

then any minimizer $u \in \text{BV}((x_0, x_n))$ of the functional

$$\begin{aligned} F : \text{BV}((x_0, x_n)) &\rightarrow [0, \infty) \\ F(u) &= \text{TV}(u) + \lambda \|u - g\|_p^\gamma. \end{aligned}$$

satisfies

$$\inf_{I_i} g \leq \inf_{I_i} u \leq \sup_{I_i} u \leq \sup_{I_i} g \quad \forall i \in \{1, \dots, n\}.$$

Proof. Let $H = \sup_I g - \inf_I g$.

We may assume that u is regular, otherwise regularizing it decreases strictly the total variation without affecting the fidelity term (see [Proposition 2.5\(v\)](#) and [Proposition 2.8\(iv\)](#)). We may also assume that $\inf_I g \leq \inf_I u \leq \sup_I u \leq \sup_I g$, thanks to [Remark 2.7](#).

Exploiting the symmetry, assume by contradiction that the inequality $\inf_{I_i} g \leq \inf_{I_i} u$ is violated for some index $i \in \{1, \dots, n\}$ and that g is increasing in I_i . Letting $m = \inf_{I_i} g$, our goal is to show that the function $\tilde{u}$ such that $\tilde{u} = u$ in $(x_0, x_n) \setminus I_i$ and $\tilde{u} = u \vee m$ in I_i performs strictly better than u , i.e. $F(\tilde{u}) < F(u)$, therefore contradicting the absurd assumption. To this end we let

$$\delta = \inf_{I_i} g - \sup_{I_i} u = m - \sup_{I_i} u$$

and we distinguish two cases: either $\delta > 0$, or $\delta \leq 0$.

Case $\delta > 0$. We start by showing that

$$\text{TV}(\tilde{u}; (x_0, x_n)) \leq \text{TV}(u; (x_0, x_n)) + 2\delta \quad (3.3)$$

We split

$$\text{TV}(u; (x_0, x_n)) = \text{TV}(u; (x_0, x_{i-1}]) + \text{TV}(u; [x_{i-1}, x_i]) + \text{TV}(u; [x_i, x_n))$$

and similarly for $\tilde{u}$. Since $\tilde{u}$ coincides with u in $(x_0, x_{i-1}] \cup [x_i, x_n)$, we have

$$\text{TV}(u; (x_0, x_{i-1}]) = \text{TV}(\tilde{u}; (x_0, x_{i-1}]), \quad \text{TV}(u; [x_i, x_n)) = \text{TV}(\tilde{u}; [x_i, x_n)).$$

Let us focus on the total variations of u and $\tilde{u}$ over the interval $[x_{i-1}, x_i]$. Since $\sup_{I_i} u = m - \delta$, for every $\varepsilon > 0$ there exists $x_\varepsilon^* \in I_i$ such that $m - \delta - \varepsilon \leq u(x_\varepsilon^*) \leq m - \delta$; moreover $\tilde{u} = m$ in (x_{i-1}, x_i) .

$$\begin{aligned} \text{TV}(u, [x_{i-1}, x_i]) &\geq |u(x_\varepsilon^*) - u(x_{i-1})| + |u(x_i) - u(x_\varepsilon^*)|, \\ \text{TV}(\tilde{u}, [x_{i-1}, x_i]) &= |m - \tilde{u}(x_{i-1})| + |\tilde{u}(x_i) - m|. \end{aligned}$$

From these and $|u(x_\varepsilon^*) - m| \leq \delta + \varepsilon$ we deduce $\text{TV}(\tilde{u}; [x_{i-1}, x_i]) \leq \text{TV}(u; [x_{i-1}, x_i]) + 2\delta + 2\varepsilon$, which leads to (3.3) by letting $\varepsilon \rightarrow 0^+$.

Let us now study the fidelity terms associated to u and $\tilde{u}$. In $I \setminus I_i$ the functions u and $\tilde{u}$ coincide, whereas in the interval I_i we have $u \leq m - \delta < \tilde{u} = m \leq g$, hence $|u - g| = g - u \geq g - m + \delta$ and $|\tilde{u} - g| = g - \tilde{u} = g - m$. Letting $f = |\tilde{u} - g|$ we have

$$\|u - g\|_p - \|\tilde{u} - g\|_p \geq \|f + \delta \mathbf{1}_{I_i}\|_p - \|f\|_p =: \Phi(\delta).$$

The function $\Phi : [0, \infty) \rightarrow \mathbb{R}$ is non-negative, increasing, convex, and $\Phi(0) = 0$. If $p = 1$ then $\Phi(\delta) = \mathcal{L}^1(I_i)\delta$; for $p > 1$ instead, if $f \equiv 0$ then $\Phi(\delta) = \mathcal{L}^1(I_i)^{1/p}\delta$, otherwise

$$\begin{aligned}\Phi'(\delta) &= \frac{d}{d\delta} \left(\int_I (f + \delta \mathbf{1}_{I_i})^p \right)^{\frac{1}{p}} = \frac{1}{p} \left(\int_I (f + \delta \mathbf{1}_{I_i})^p \right)^{\frac{1}{p}-1} \frac{d}{d\delta} \int_I (f + \delta \mathbf{1}_{I_i})^p \\ &= \left(\int_I (f + \delta \mathbf{1}_{I_i})^p \right)^{\frac{1-p}{p}} p \int_{I_i} (f + \delta)^{p-1} = p \left(\frac{\|f + \delta\|_{L^{p-1}(I_i)}}{\|f + \delta \mathbf{1}_{I_i}\|_{L^p(I)}} \right)^{p-1},\end{aligned}$$

hence

$$\Phi(\delta) \geq \Phi'(0)\delta = p \left(\frac{\|f\|_{L^{p-1}(I_i)}}{\|f\|_{L^p(I)}} \right)^{p-1} \delta.$$

In conclusion,

$$\|u - g\|_p - \|\tilde{u} - g\|_p \geq C_i \delta$$

where

$$C_i = \begin{cases} \mathcal{L}^1(I_i)^{1/p} & p = 1 \vee g \equiv m \text{ in } I_i, \\ p \left(\frac{\|f\|_{L^{p-1}(I_i)}}{\|f\|_{L^p(I)}} \right)^{p-1} & \text{otherwise.} \end{cases}$$

Since $0 \leq \|\tilde{u} - g\|_p \leq \|u - g\|_p \leq H \mathcal{L}^1(I)^{1/p}$, from the previous inequality and Lagrange's Mean Value Theorem we deduce

$$\|u - g\|_p^\gamma - \|\tilde{u} - g\|_p^\gamma \geq \frac{\gamma C_i \delta}{(H \mathcal{L}^1(I)^{1/p})^{1-\gamma}} \geq G(I_i) \delta. \quad (3.4)$$

Combining the estimates (3.3) for the total variation and (3.4) for the fidelity term, we get

$$0 \geq F(u) - F(\tilde{u}) = (\text{TV}(u) - \text{TV}(\tilde{u})) + \lambda \left(\|u - g\|_{L^p(I)}^\gamma - \|\tilde{u} - g\|_{L^p(I)}^\gamma \right) \geq (\lambda G(I_i) - 2) \delta,$$

which is a contradiction because both factors in the right hand side are strictly positive.

Case $\delta \leq 0$. We treat the total variation and the fidelity term separately. With regard to the total variation, our goal is to show that

$$\text{TV}(\tilde{u}; (x_0, x_n)) \leq \text{TV}(u; (x_0, x_n)). \quad (3.5)$$

By hypothesis $\sup_{I_i} u \geq \inf_{I_i} g$, so for every $\varepsilon > 0$ there is $x_\varepsilon^* \in I_i$ such that $u(x_\varepsilon^*) > \inf_{I_i} g - \varepsilon =: m_\varepsilon$. We split

$$\text{TV}(u; (x_0, x_n)) = \text{TV}(u; (x_0, x_{i-1}]) + \text{TV}(u; [x_{i-1}, x_\varepsilon^*]) + \text{TV}(u; [x_\varepsilon^*, x_i]) + \text{TV}(u; [x_i, x_n])$$

and similarly for $\tilde{u}$. Since $\tilde{u}$ coincides with u in $(x_0, x_{i-1}] \cup [x_i, x_n)$, we have

$$\text{TV}(u; (x_0, x_{i-1}]) = \text{TV}(\tilde{u}; (x_0, x_{i-1}]), \quad \text{TV}(u; [x_i, x_n]) = \text{TV}(\tilde{u}; [x_i, x_n]).$$

Let us focus on the total variations of u and $\tilde{u}$ over the interval $[x_{i-1}, x_\varepsilon^*]$, since those over $[x_\varepsilon^*, x_i]$ are treated analogously. In all following cases we will be proving that

$$\text{TV}(\tilde{u}; [x_{i-1}, x_\varepsilon^*]) \leq \text{TV}(u; [x_{i-1}, x_\varepsilon^*]) + \varepsilon, \quad (3.6)$$

so that passing to the limit $\varepsilon \rightarrow 0^+$ we finally get (3.5).

- (i) If $u(x_{i-1}) \geq m$, then $\tilde{u} = u \vee m$ on the whole interval $[x_{i-1}, x_\varepsilon^*]$, hence we have immediately $\text{TV}(\tilde{u}; [x_{i-1}, x_\varepsilon^*]) \leq \text{TV}(u; [x_{i-1}, x_\varepsilon^*])$.

(ii) Consider now $u(x_{i-1}) < m$.

(a) If $u^+(x_{i-1}) \geq m_\varepsilon$ then $u^+(x_{i-1}) \leq \tilde{u}^+(x_{i-1}) \leq u^+(x_{i-1}) + \varepsilon$, hence (3.6) follows as

$$\begin{aligned} \text{TV}(u; [x_{i-1}, x_\varepsilon^*]) &= |u^+(x_{i-1}) - u(x_{i-1})| + \text{TV}(u; (x_{i-1}, x_\varepsilon^*]) \\ &\geq |\tilde{u}^+(x_{i-1}) - \tilde{u}(x_{i-1})| - \varepsilon + \text{TV}(\tilde{u}; (x_{i-1}, x_\varepsilon^*]) \\ &= \text{TV}(\tilde{u}; [x_{i-1}, x_\varepsilon^*]) - \varepsilon. \end{aligned}$$

(b) Otherwise, if $u^+(x_{i-1}) < m_\varepsilon$, define

$$x'_\varepsilon = \max\{x \in (x_{i-1}, x_\varepsilon^*] : u \leq m_\varepsilon \text{ in } (x_{i-1}, x)\}.$$

We have that $u^-(x'_\varepsilon) \leq m_\varepsilon$.

i. If $x'_\varepsilon < x_\varepsilon^*$ then $u^+(x'_\varepsilon) \geq m_\varepsilon$, so we may assume without loss of generality that $u(x'_\varepsilon) = m_\varepsilon$; moreover $\tilde{u} = m$ in $(x_{i-1}, x'_\varepsilon]$.

$$\begin{aligned} \text{TV}(u; [x_{i-1}, x_\varepsilon^*]) &= \text{TV}(u; [x_{i-1}, x'_\varepsilon]) + \text{TV}(u; [x'_\varepsilon, x_\varepsilon^*]) \\ &\geq |u(x'_\varepsilon) - u(x_{i-1})| + \text{TV}(u; [x'_\varepsilon, x_\varepsilon^*]) \end{aligned}$$

and

$$\begin{aligned} \text{TV}(\tilde{u}; [x_{i-1}, x_\varepsilon^*]) &= \text{TV}(\tilde{u}; [x_{i-1}, x'_\varepsilon]) + \text{TV}(\tilde{u}; [x'_\varepsilon, x_\varepsilon^*]) \\ &= |m - u(x_{i-1})| + \text{TV}(\tilde{u}; [x'_\varepsilon, x_\varepsilon^*]) \end{aligned}$$

Observing that $|u(x'_\varepsilon) - m| = |m_\varepsilon - m| = \varepsilon$ and $\text{TV}(u; [x'_\varepsilon, x_\varepsilon^*]) \geq \text{TV}(\tilde{u}; [x'_\varepsilon, x_\varepsilon^*])$, we deduce (3.6).

ii. If instead $x'_\varepsilon = x_\varepsilon^*$, then $u \leq m_\varepsilon$ and $\tilde{u} = m$ in $(x_{i-1}, x_\varepsilon^*)$, hence

$$\text{TV}(u; [x_{i-1}, x_\varepsilon^*]) \geq |u(x_\varepsilon^*) - u(x_{i-1})|$$

and, since $u(x_{i-1}) < m \leq \tilde{u}(x_\varepsilon^*)$,

$$\text{TV}(\tilde{u}; [x_{i-1}, x_\varepsilon^*]) = |m - u(x_{i-1})| + |\tilde{u}(x_\varepsilon^*) - m| = |\tilde{u}(x_\varepsilon^*) - u(x_{i-1})|;$$

noticing that $u(x_\varepsilon^*) \leq \tilde{u}(x_\varepsilon^*) \leq u(x_\varepsilon^*) + \varepsilon$ we conclude again that (3.6) holds.

This concludes the proof of (3.6), and hence that of (3.5).

On the other hand, since $\inf_{I_i} u < \inf_{I_i} g$, there is a point $x^* \in I_i$ such that $u(x^*) < \inf_{I_i} g$. By regularity of u , there is an open interval (one endpoint of which is x^*) where $u < \inf_{I_i} g$. This implies that $\|\tilde{u} - g\|_p < \|u - g\|_p$, which together with (3.5) leads to $F(\tilde{u}) < F(u)$, which contradicts the minimality of u . $\square$

Theorem 3.10. *Let $x_0 < \dots < x_n$ and let $g \in \text{BV}_{\text{reg}}((x_0, x_n))$ be a function which is monotone in every interval $I_i = (x_{i-1}, x_i)$ and with at most α -growth at their extremes, i.e. there exist $\alpha, C > 0$ such that*

$$|g(x) - g^+(x_{i-1})| \leq C(x - x_{i-1})^\alpha \quad |g(x) - g^-(x_i)| \leq C(x_i - x)^\alpha \quad \forall x \in (x_{i-1}, x_i).$$

Given some points $\xi_i \in I_i$, let $I_i^- = (x_{i-1}, \xi_i)$ and $I_i^+ = (\xi_i, x_i)$. If the conditions of Proposition 3.9 hold for the intervals I_i , I_i^- and I_i^+ , and in addition the inequalities

$$\gamma \leq \left(1 + \frac{1}{p\alpha}\right)^{-1} = \frac{p\alpha}{p\alpha + 1} \quad \text{and} \quad \lambda > 2^{2+1/\beta} K(p, \alpha)^{-\gamma/p} C^{\gamma/(p\alpha)} n^{1-\gamma/p} H^{1-\beta\gamma/p}$$

are satisfied, then g is essentially the unique minimizer of the functional

$$F : \text{BV}((x_0, x_n)) \rightarrow [0, \infty)$$

$$F(u) = \text{TV}(u) + \lambda \|u - g\|_p^\gamma,$$

in the sense that any other minimizer coincides with g outside of the jump set of g .

In the course of the proof it would appear natural to rearrange monotonically $u|_{I_i}$ via [Proposition 2.17](#). Unfortunately, this transformation does not keep track of jumps at the endpoints of these intervals, which may cause an increase in energy. This obstruction is overcome by the additional requests on the intervals $I_i^\pm$ and a more involved ad hoc proof strategy.

Proof. Let $I = (x_0, x_n)$, $H = \sup_I g - \inf_I g$ and $H_i = \sup_{I_i} g - \inf_{I_i} g$. By [Proposition 3.9](#), for every $i \in \{1, \dots, n\}$ we have

$$\inf_{I_i} g \leq \inf_{I_i} u \leq \sup_{I_i} u \leq \sup_{I_i} g, \quad \inf_{I_i^\pm} g \leq \inf_{I_i^\pm} u \leq \sup_{I_i^\pm} u \leq \sup_{I_i^\pm} g.$$

Fix $i \in \{1, \dots, n\}$. Assume that g is non-decreasing in I_i . Define

$$h_i = \inf_{I_i} u - g^+(x_{i-1}) = \inf_{I_i} u - \inf_{I_i} g \geq 0, \quad k_i = g^-(x_i) - \sup_{I_i} u = \sup_{I_i} g - \sup_{I_i} u \geq 0.$$

Notice that $h_i + k_i \leq H_i \leq H$. Moreover

$$\inf_{I_i^-} u \leq \sup_{I_i^-} g = g^-(\xi_i) \leq g^+(\xi_i) = \inf_{I_i^+} g \leq \inf_{I_i^+} u$$

and

$$u(\xi_i) \geq \min\{u^-(\xi_i), u^+(\xi_i)\} \geq \min\left\{\inf_{I_i^-} u, \inf_{I_i^+} u\right\} \geq \inf_{I_i^-} u$$

therefore

$$\inf_{I_i^-} u = \inf_{I_i} u \quad \text{and} \quad \sup_{I_i^+} u = \sup_{I_i} u.$$

Given $\varepsilon > 0$, let $x_{i,\varepsilon}^\pm \in I_i^\pm$ such that

$$u(x_{i,\varepsilon}^-) \leq \inf_{I_i} u + \varepsilon \quad \text{and} \quad u(x_{i,\varepsilon}^+) \geq \sup_{I_i} u - \varepsilon.$$

For every $i \in \{1, \dots, n\}$ there exists $x_{i-1} < x'_{i,\varepsilon} < x''_{i,\varepsilon} < x_i$ such that

$$|u(x'_{i,\varepsilon}) - g^+(x_{i-1})| \leq h_i + \varepsilon, \quad |u(x''_{i,\varepsilon}) - g^-(x_i)| \leq k_i + \varepsilon. \quad (3.7)$$

By the monotonicity of $g|_{I_i}$ we have

$$\begin{aligned} \text{TV}(g; [x_0, x_n]) &= \sum_{i=1}^n \text{TV}(g; (x_{i-1}, x_i)) + \sum_{i=1}^{n-1} |[g](x_i)| \\ &= \sum_{i=1}^n |g^-(x_i) - g^+(x_{i-1})| + \sum_{i=1}^{n-1} |g^+(x_i) - g^-(x_i)|. \end{aligned} \quad (3.8)$$

On the other hand $\text{TV}(u; [x_0, x_n])$ can be estimated from below along the sorted list of points

$$x'_{1,\varepsilon} < x''_{1,\varepsilon} < \dots < x'_{i,\varepsilon} < x''_{i,\varepsilon} < \dots < x'_{n,\varepsilon} < x''_{n,\varepsilon}$$

and using (3.7) and (3.8), obtaining

$$\begin{aligned}
& \text{TV}(u; [x_0, x_n]) \\
& \geq \sum_{i=1}^n |u(x''_{i,\varepsilon}) - u(x'_{i,\varepsilon})| + \sum_{i=1}^{n-1} |u(x'_{i+1,\varepsilon}) - u(x''_{i,\varepsilon})| \\
& \stackrel{(3.7)}{\geq} \sum_{i=1}^n (|g^-(x_i) - g^+(x_{i-1})| - h_i - k_i - 2\varepsilon) + \sum_{i=1}^{n-1} (|g^+(x_i) - g^-(x_i)| - h_{i+1} - k_i - 2\varepsilon) \\
& \stackrel{(3.8)}{\geq} \text{TV}(g; [x_0, x_n]) - 2 \sum_{i=1}^n (h_i + k_i) - (4n - 2)\varepsilon.
\end{aligned}$$

Letting $\varepsilon \rightarrow 0$ we deduce $\text{TV}(u; [x_0, x_n]) \geq \text{TV}(g; [x_0, x_n]) - 2 \sum_{i=1}^n (h_i + k_i)$.

Applying Lemma 2.19 to every interval I_i and summing from 1 to n we get

$$\|u - g\|_{L^p(I)}^p \geq \sum_{i=1}^n \|u - g\|_{L^p(I_i)}^p \geq K(p, \alpha) C^{-1/\alpha} \sum_{i=1}^n (h_i^{p+1/\alpha} + k_i^{p+1/\alpha}).$$

Let us call $\beta = p + 1/\alpha > 1$ the exponent appearing in the fidelity term estimate. Using for convenience the shorter notation $M_q(h; k) = M_q(h_1, k_1, \dots, h_n, k_n)$ to denote the q -mean of the $2n$ parameters (see Definition 2.20), we have that $M_1(h; k) \leq M_\beta(h; k)$. Since $t \mapsto t^\beta$ is super-additive, we have also $h_i^\beta + k_i^\beta \leq (h_i + k_i)^\beta \leq H_i^\beta \leq H^\beta$, hence

$$M_\beta(h; k) = \left(\frac{1}{2n} \sum_{i=1}^n (h_i^\beta + k_i^\beta) \right)^{1/\beta} \leq \left(\frac{1}{2n} \sum_{i=1}^n H_i^\beta \right)^{1/\beta} \leq \left(\frac{1}{2n} n H^\beta \right)^{1/\beta} = 2^{-1/\beta} H.$$

The assumption $\gamma \leq p/\beta$ ensures that $\beta\gamma/p \leq 1$, therefore

$$\left(\frac{M_\beta(h; k)}{2^{-1/\beta} H} \right)^{\beta\gamma/p} \geq \frac{M_\beta(h; k)}{2^{-1/\beta} H}.$$

This implies

$$\begin{aligned}
\|u - g\|_{L^p(I)}^\gamma & \geq K(p, \alpha)^{\gamma/p} C^{-\gamma/(p\alpha)} \left(\sum_{i=1}^n (h_i^{p+1/\alpha} + k_i^{p+1/\alpha}) \right)^{\gamma/p} \\
& = K(p, \alpha)^{\gamma/p} C^{-\gamma/(p\alpha)} \left(2n M_\beta(h; k)^\beta \right)^{\gamma/p} \\
& = K(p, \alpha)^{\gamma/p} C^{-\gamma/(p\alpha)} (2n)^{\gamma/p} M_\beta(h; k)^{\beta\gamma/p} \\
& \geq K(p, \alpha)^{\gamma/p} C^{-\gamma/(p\alpha)} (2n)^{\gamma/p} 2^{-\gamma/(p+1/\beta)} H^{\beta\gamma/p-1} M_\beta(h; k) \\
& = 2^{-1/\beta} K(p, \alpha)^{\gamma/p} C^{-\gamma/(p\alpha)} n^{\gamma/p} H^{\beta\gamma/p-1} M_\beta(h; k).
\end{aligned}$$

Combining the resulting lower bounds for the total variation and the fidelity terms leads to

$$\begin{aligned}
0 & \geq F(u) - F(g) = \text{TV}(u; [x_0, x_n]) - \text{TV}(g; [x_0, x_n]) + \lambda \|u - g\|_{L^p([x_0, x_n])}^\gamma \\
& \geq -4n M_1(h; k) + \lambda 2^{-1/\beta} K(p, \alpha)^{\gamma/p} C^{-\gamma/(p\alpha)} n^{\gamma/p} H^{\beta\gamma/p-1} M_\beta(h; k) \\
& \geq [\lambda 2^{-1/\beta} K(p, \alpha)^{\gamma/p} C^{-\gamma/(p\alpha)} n^{\gamma/p} H^{\beta\gamma/p-1} - 4n] M_\beta(h; k).
\end{aligned}$$

If λ is sufficiently large (precisely as stated in the assumptions of the theorem), the square bracket is positive, hence we must have

$$h_1 = k_1 = \dots = h_n = k_n = 0.$$

This implies that $\text{TV}(u; [x_0, x_n]) \geq \text{TV}(g; [x_0, x_n])$, therefore $F(u) \geq F(g)$ and g is a minimizer.

Let us show that $u = g$ on $I \setminus J_g$. If it is not the case, there is $x \in I \setminus J_g$ such that $u(x) \neq g(x)$. Since u is regular, we have either $u^-(x) \neq g(x)$ or $u^+(x) \neq g(x)$. Moreover, x is a continuity point for g , hence there is a small interval of the form $(x - \varepsilon, x)$ or $(x, x + \varepsilon)$ where u and g differ, leading to $\|u - g\|_{L^p(I)} > 0$ and contradicting the optimality of u . $\square$

Remark 3.11. Notice that the assumption on γ in [Theorem 3.4](#)⁸ is complementary to that in [Theorem 3.8](#) and [Theorem 3.10](#), demonstrating the sharpness of our result.

4 Nonlinear $\text{TV}_{\varphi, \theta}$

In this section, we tackle the problem with the nonlinear total variation $\text{TV}_{\varphi, \theta}$. As we will see in the ensuing sections, this case is remarkably more difficult than the linear case. In particular, the functional does not behave well with respect to local truncations. To overcome this difficulty, we study a suitable boundary value problem in [Subsection 4.1](#), characterize its minimizers, and use them to construct competitors for the problem with the nonlinear total variation. [Subsection 4.2](#) is the nonlinear counterpart to [Subsection 3.1](#), whereas the results in [Subsection 3.2](#) cannot be reproduced in this setting: in [Subsection 4.3](#) we prove [Proposition 4.10](#) where we show that different behaviors of φ at infinity can disprove both [Theorem 3.8](#) and [Theorem 3.10](#). As a consequence, we propose to study the penalized functional in [\(4.7\)](#), for which numerical results illustrate the behavior of minimizers.

4.1 Boundary value problem

We want to minimize the nonlinear total variation on an interval of a function with prescribed limits at the endpoints, which can be formulated as the optimization problem

$$E(L, H) = \inf \left\{ \text{TV}_{\varphi, \theta}(u; (0, L)) : u \in \text{BV}((0, L)), u^+(0) = 0, u^-(L) = H \right\}, \quad (4.1)$$

where $L > 0$ and $H \geq 0$ are two prescribed constants.

Theorem 4.1. *The problem [\(4.1\)](#) admits at least a minimizer $u \in \text{BV}_{\text{reg}}((0, L))$ of the form*⁹

$$u(x) = ax + b\Theta(x - \bar{x}),$$

for some $a, b \geq 0$ and $\bar{x} \in (0, L)$, hence

$$\begin{aligned} E(L, H) &= \min \{ L\varphi(a) + \theta(b) : a, b \geq 0, La + b = H \} \\ &= \min \{ L\varphi(a) + \theta(H - La) : 0 \leq a \leq H/L \} \\ &= \min \{ L\varphi((H - b)/L) + \theta(b) : 0 \leq b \leq H \} \\ &= \min \{ \varphi_L(H\alpha) + \theta(H(1 - \alpha)) : 0 \leq \alpha \leq 1 \}. \end{aligned} \quad (4.2)$$

In particular, there is at least one minimizer which has no Cantor part. If φ is strictly convex and θ is strictly concave then any minimizer is of this form.

Proof. Let $I = (0, L)$, let u be a minimizer, let $Du = u' \mathcal{L}^1 + D^c u + [u] \mathcal{H}^0|_{J_u}$ be the decomposition of its derivative, and let

$$A = |D^a u|(I), \quad C = |D^c u|(I), \quad J = |D^j u|(I).$$

⁸As a particular case, [Theorem 3.1](#) falls in this setting too, with $\alpha = \log 2 / \log 3$.

⁹ $\Theta(\cdot)$ is the Heaviside step function.

Observe that $H = Du(I) = D^a u(I) + D^c u(I) + D^j u(I) \leq A + C + J$, with equality if and only if all three measures are non-negative, i.e. u is non-decreasing.

- *Monotonicity.* The function $\tilde{u}$ such that $D\tilde{u} = \frac{H}{A+C+J} (|D^a u| + |D^c u| + |D^j u|)$ is non-decreasing and satisfies

$$\begin{aligned} \text{TV}_{\varphi, \theta}(\tilde{u}) &= \int_0^L \varphi\left(\frac{H}{A+C+J}|u'(x)|\right) dx + \frac{H}{A+C+J}C + \sum_{x \in J_u} \theta\left(\frac{H}{A+C+J}|[u](x)|\right) \\ &\leq \int_0^L \varphi(|u'(x)|) dx + C + \sum_{x \in J_u} \theta(|[u](x)|) = \text{TV}_{\varphi, \theta}(u) \end{aligned}$$

because $\frac{H}{A+C+J} \leq 1$ and φ and θ are non-decreasing. This implies that there is always a non-decreasing minimizer.

Let us investigate the equality conditions. If $C > 0$ we must have $H = A + C + J$, i.e. u non-decreasing. If $A > 0$ we must have $\varphi\left(\frac{H}{A+C+J}|u'(x)|\right) = \varphi(|u'|)$, which implies $H = A + C + J$, i.e. u non-decreasing, if φ is strictly increasing.¹⁰ If $J > 0$ we must have $\theta\left(\frac{H}{A+C+J}|[u](x)|\right) = \theta(|[u](x)|)$, which implies $H = A + C + J$, i.e. u non-decreasing, if θ is strictly increasing.¹¹

This means that a minimizer must be non-decreasing if φ and θ are strictly increasing.

- *Transfer Cantor part to absolutely continuous part.* Assume that u is non-decreasing. Observe that $\varphi(\beta) - \varphi(\alpha) \leq \beta - \alpha$ for every $0 \leq \alpha < \beta$; moreover, if φ is strictly convex, then this inequality is strict for every $\alpha < \beta$.

– Assume $A > 0$. Then the function $\tilde{u}$ such that $D\tilde{u} = \left(1 + \frac{C}{A}\right) D^a u + D^j u$ satisfies

$$\begin{aligned} \text{TV}_{\varphi, \theta}(\tilde{u}) &= \int_0^L \varphi(\tilde{u}'(x)) dx + \sum_{x \in J_u} \theta(|[\tilde{u}](x)|) \\ &= \int_0^L \varphi\left(\left(1 + \frac{C}{A}\right) u'(x)\right) dx + \sum_{x \in J_u} \theta(|[u](x)|) \\ &\leq \int_0^L \varphi(u'(x)) dx + \int_0^L \frac{C}{A} u'(x) dx + \sum_{x \in J_u} \theta(|[u](x)|) \\ &= \int_0^L \varphi(u'(x)) dx + C + \sum_{x \in J_u} \theta(|[u](x)|) = \text{TV}_{\varphi, \theta}(u). \end{aligned}$$

In this case, if φ is strictly convex, equality holds if and only if $\frac{C}{A} u'(x) = 0$ for almost every x , hence $C = 0$.

– If $A = 0$, then the function $\tilde{u}$ such that $D\tilde{u} = \frac{C}{L} + D^j u$ satisfies

$$\begin{aligned} \text{TV}_{\varphi, \theta}(\tilde{u}) &= \int_0^L \varphi\left(\frac{C}{L}\right) dx + \sum_{x \in J_u} \theta(|[\tilde{u}](x)|) \\ &\leq \int_0^L \frac{C}{L} dx + \sum_{x \in J_u} \theta(|[u](x)|) \\ &= C + \sum_{x \in J_u} \theta(|[u](x)|) = \text{TV}_{\varphi, \theta}(u). \end{aligned}$$

¹⁰Notice that φ strictly convex implies φ strictly increasing.

¹¹Notice that θ strictly concave implies θ strictly increasing.

If $\varphi \neq \text{Id}$, then $\varphi(t) < t$ for every $t > 0$, therefore equality holds if and only if $C/L = 0$, hence $C = 0$.

- *Transfer Cantor part to jump part.* Assume that u is non-decreasing. Observe that the function θ is subadditive, i.e. $\theta(\alpha + \beta) \leq \theta(\alpha) + \theta(\beta)$ for every $\alpha, \beta \geq 0$; moreover, if θ is strictly concave, then this inequality is strict for every $\alpha, \beta > 0$.

– Assume $J > 0$. Then the function $\tilde{u}$ such that $D\tilde{u} = D^a u + \left(1 + \frac{C}{J}\right) D^j u$ satisfies

$$\begin{aligned} \text{TV}_{\varphi, \theta}(\tilde{u}) &= \int_0^L \varphi(|u'(x)|) dx + \sum_{x \in J_u} \theta\left(\left(1 + \frac{C}{J}\right) |[u](x)|\right) \\ &\leq \int_0^L \varphi(|u'(x)|) dx + \sum_{x \in J_u} \theta(|[u](x)|) + \sum_{x \in J_u} \theta\left(\frac{C}{J} |[u](x)|\right) \\ &\leq \int_0^L \varphi(|u'(x)|) dx + \sum_{x \in J_u} \theta(|[u](x)|) + \sum_{x \in J_u} \frac{C}{J} |[u](x)| \\ &= \int_0^L \varphi(|u'(x)|) dx + \sum_{x \in J_u} \theta(|[u](x)|) + C = \text{TV}_{\varphi, \theta}(u). \end{aligned}$$

If $\theta(t) < t$ for every $t > 0$, then equality holds if and only if $\frac{C}{J} |[u](x)| = 0$ for every $x \in J_u$, hence $C = 0$.

– If $J = 0$, then the function $\tilde{u}$ such that $D\tilde{u} = D^a u + C\delta_{\bar{x}}$, with $\bar{x} \in I$, satisfies

$$\text{TV}_{\varphi, \theta}(\tilde{u}) = \int_0^L \varphi(|u'(x)|) dx + \theta(C) \leq \int_0^L \varphi(|u'(x)|) dx + C = \text{TV}_{\varphi, \theta}(u).$$

If $\theta(t) < t$ for every $t > 0$, then equality holds if and only if $C = 0$.

- *Transfer Cantor part to jump part, additively (optional).* This step is presented as an alternative way to transfer the Cantor part to the jump part. Let $\tilde{u}$ be a function such that $D\tilde{u} = D^a u + D^j u + \rho + \sigma$ where ρ, σ are atomic measures satisfying $\rho(I) + \sigma(I) = D^c u(I)$, $|\rho|(I) + |\sigma|(I) = |D^c u|(I)$, $\text{supp}(\rho) \subseteq J_u$, $\sigma \perp D^j u$, $S = \text{supp}(\sigma)$. Then

$$\begin{aligned} \text{TV}_{\varphi, \theta}(\tilde{u}) &= \int_0^L \varphi(|u'(x)|) dx + \sum_{x \in J_u} \theta(|[u](x) + \rho(x)|) + \sum_{x \in S} \theta(|\sigma(x)|) \\ &\leq \int_0^L \varphi(|u'(x)|) dx + \sum_{x \in J_u} \theta(|[u](x)|) + \sum_{x \in J_u} \theta(|\rho(x)|) + \sum_{x \in S} \theta(|\sigma(x)|) \\ &\leq \int_0^L \varphi(|u'(x)|) dx + \sum_{x \in J_u} \theta(|[u](x)|) + \sum_{x \in J_u} |\rho(x)| + \sum_{x \in S} |\sigma(x)| \\ &\leq \int_0^L \varphi(|u'(x)|) dx + \sum_{x \in J_u} \theta(|[u](x)|) + |D^c u|(I) \leq \text{TV}_{\varphi, \theta}(u). \end{aligned}$$

This works in particular by taking either ρ or σ to be $D^c u(I)\delta_{\bar{x}}$ and the other equal to zero. Since this step is optional, we do not discuss here the equality conditions.

- *Averaging the absolutely continuous part.* By Jensen inequality we have

$$\varphi\left(\left|\frac{D^a u(I)}{L}\right|\right) = \varphi\left(\left|\int_0^L u'(x) dx\right|\right) \leq \int_0^L \varphi(|u'(x)|) dx = \frac{1}{L} \int_0^L \varphi(|u'(x)|) dx;$$

moreover, if φ is strictly convex, then equality holds if and only if u' is constantly equal to $D^a u(I)/L$.¹²

As a consequence, the function $\tilde{u}$ such that $D\tilde{u} = \frac{D^a u(I)}{L} \mathcal{L}^1 + D^c u + D^j u$ satisfies

$$\begin{aligned} \mathrm{TV}_{\varphi, \theta}(\tilde{u}) &= \int_0^L \varphi\left(\left|\frac{D^a u(I)}{L}\right|\right) dx + C + \sum_{x \in J_u} \theta(|[u](x)|) \\ &\leq \int_0^L \varphi(|u'(x)|) dx + C + \sum_{x \in J_u} \theta(|[u](x)|) = \mathrm{TV}_{\varphi, \theta}(u), \end{aligned}$$

with equality if and only if u' is constantly equal to $D^a u(I)/L$, under the strict convexity assumption of φ .

- *Coalescing jumps into a single one.* The function $\tilde{u}$ such that $D\tilde{u} = D^a u + D^c u + D^j u(I)\delta_{\bar{x}}$ for some $\bar{x} \in I$ satisfies

$$\begin{aligned} \mathrm{TV}_{\varphi, \theta}(\tilde{u}) &= \int_0^L \varphi(|u'(x)|) dx + |D^c u|(I) + \theta(|D^j u(I)|) \\ &= \int_0^L \varphi(|u'(x)|) dx + |D^c u|(I) + \theta\left(\left|\sum_{x \in J_u} [u](x)\right|\right) \\ &\leq \int_0^L \varphi(|u'(x)|) dx + |D^c u|(I) + \theta\left(\sum_{x \in J_u} |[u](x)|\right) \\ &\leq \int_0^L \varphi(|u'(x)|) dx + |D^c u|(I) + \sum_{x \in J_u} \theta(|[u](x)|) = \mathrm{TV}_{\varphi, \theta}(u). \end{aligned}$$

Let us investigate the equality conditions. In the first inequality, equality holds if and only if $|D^j u(I)| = |D^j u|(I)$ if θ is strictly increasing; in the second inequality, equality holds if and only if u has at most one jump if θ is strictly concave.

If φ is strictly convex and θ is strictly concave, then a minimizer u must necessarily be non-decreasing, without any Cantor part ($D^c u = 0$), with u' constant and at most one jump. Such functions can be written as $u(x) = ax + b\Theta(x - \bar{x})$, where Θ is the Heaviside step function, with $\bar{x} \in (0, L)$ and $a, b \in [0, \infty)$, in which case the functional to minimize is $\mathrm{TV}_{\varphi, \theta}(u) = L\varphi(a) + \theta(b)$.

The admissibility condition for the boundary value problem is $aL + b = H$, so we need to find $a \in [0, H/L]$ that minimizes $L\varphi(a) + \theta(H - La)$, or equivalently $b \in [0, H]$ that minimizes $L\varphi((H - b)/L) + \theta(b)$. The last formula in (4.2) follows letting $\alpha = \frac{L}{H}a$: then $a \in [0, H/L]$ if and only if $\alpha \in [0, 1]$ and we have

$$L\varphi(a) + \theta(H - La) = L\varphi\left(\frac{H}{L}\alpha\right) + \theta(H(1 - \alpha)) = \varphi_L(H\alpha) + \theta(H(1 - \alpha)). \quad \square$$

Remark 4.2. If φ and θ are differentiable, then a pair (a, b) optimal for (4.2) and such that $a, b > 0$ must necessarily satisfy $\varphi'(a) = \theta'(b)$. Notice that in general there can be a finite, countable or uncountable number of minimizers of (4.2) and pairs satisfying the previous condition on the derivative.

¹²Notice that the function $\varphi \circ |\cdot| : \mathbb{R} \rightarrow \mathbb{R}$ is strictly convex if and only if φ is strictly convex.

Remark 4.3. If we extend φ, θ to $\bar{\varphi}, \bar{\theta} : \mathbb{R} \rightarrow [0, \infty]$ by setting $\bar{\varphi}(t) = \bar{\theta}(t) = +\infty$ for $t < 0$,¹³ then (4.2) can be rewritten as

$$\begin{aligned} E(L, H) &= \min_{0 \leq b \leq H} \left(L\varphi\left(\frac{H-b}{L}\right) + \theta(b) \right) = \min_{b \geq 0} \left(L\bar{\varphi}\left(\frac{H-b}{L}\right) + \bar{\theta}(b) \right) \\ &= \min_{b \in \mathbb{R}} \left(L\bar{\varphi}\left(\frac{H-b}{L}\right) + \bar{\theta}(b) \right) = \min_{b \in \mathbb{R}} (\bar{\varphi}_L(H-b) + \bar{\theta}(b)) = (\bar{\varphi}_L \square \bar{\theta})(H), \end{aligned}$$

where $\bar{\varphi}_L(t) = L\bar{\varphi}(t/L)$ is the perspective of $\bar{\varphi}$ and $\square$ denotes the inf-convolution.

In the following propositions we study several properties of $E(L, H)$: monotonicity and subadditivity (Proposition 4.4), upper bounds (Proposition 4.5), lower bounds (Proposition 4.6), and continuous extension (Proposition 4.8).

Proposition 4.4. *The following properties are true.*

- (i) $E(L, H)$ is non-increasing in L and non-decreasing in H .
- (ii) E is jointly subadditive: $E(L_1 + L_2, H_1 + H_2) \leq E(L_1, H_1) + E(L_2, H_2)$.

Proof. We prove the two properties separately.

- (i) In the last representation formula in (4.2), observe that for every $\alpha \in [0, 1]$ the function $\varphi_L(H\alpha) + \theta(H(1-\alpha))$ to be minimized enjoys the claimed monotonicity (the monotonicity in L follows from the monotonicity of the perspective φ_L mentioned in Subsection 2.3).

In Subsection A.1 we provide two alternative proofs of this property.

- (ii) Let $u_1 \in \text{BV}((0, L_1))$ and $u_2 \in \text{BV}((0, L_2))$ be optimal for $E(L_1, H_1)$ and $E(L_2, H_2)$ respectively. Define $u \in \text{BV}((0, L_1 + L_2))$ by

$$u(x) = \begin{cases} u_1(x) & x \in (0, L_1), \\ H_1 & x = L_1, \\ H_1 + u_2(x - L_1) & x \in (L_1, L_1 + L_2). \end{cases}$$

Observe that u is continuous in L_1 and has the correct limits at 0 and $L_1 + L_2$ to be admissible for $E(L_1 + L_2, H_1 + H_2)$, therefore

$$\begin{aligned} E(L_1 + L_2, H_1 + H_2) &\leq \text{TV}_{\varphi, \theta}(u, (0, L_1 + L_2)) \\ &= \text{TV}_{\varphi, \theta}(u, (0, L_1)) + \text{TV}_{\varphi, \theta}(u, (L_1, L_1 + L_2)) \\ &= \text{TV}_{\varphi, \theta}(u_1, (0, L_1)) + \text{TV}_{\varphi, \theta}(u_2, (0, L_2)) \\ &= E(L_1, H_1) + E(L_2, H_2). \end{aligned} \quad \square$$

Proposition 4.5. *We have the upper bounds*

$$E(L, H) \leq \varphi_L(H) = L\varphi(H/L), \quad E(L, H) \leq \theta(H). \quad (4.3)$$

Proof. These upper bounds follow directly from (4.2), taking $a = H/L, b = 0$ for the former and $a = 0, b = H$ for the latter. $\square$

¹³Notice that $\bar{\varphi}$ is still convex, whereas $\bar{\theta}$ is neither convex nor concave globally.

Proposition 4.6. *We have the lower bounds*

$$E(L, H) \geq \varphi_L(H) + \theta(H) - H, \quad (4.4)$$

and

$$\begin{aligned} E(L, H) &\geq \varphi_L(H) - (\varphi_L'(H)H - \theta(H))_+ \\ &= \begin{cases} \varphi_L(H) & \varphi_L'(H) \leq \theta(H)/H, \\ \varphi_L(H) - \varphi_L'(H)H + \theta(H) & \varphi_L'(H) \geq \theta(H)/H, \end{cases} \end{aligned} \quad (4.5)$$

where $\varphi_L'(H) = \min \partial \varphi_L(H)$ is the smallest element of the subdifferential.¹⁴

Proof. Let us prove the first lower bound. Define the complementary functions $\psi, \rho : [0, \infty) \rightarrow [0, \infty)$ by $\psi(t) = t - \varphi(t)$ and $\rho(t) = t - \theta(t)$. Observe that $\psi(0) = \rho(0) = 0$, ψ is concave, increasing and sublinear at $+\infty$,¹⁵ whereas ρ is convex and $\rho'(0) = 0$, hence increasing too. If we define also $\psi_L(t) = L\psi(t/L)$, then we have $\psi_L(t) = t - \varphi_L(t)$. Exploiting the monotonicity of ψ_L and ρ , for every $\alpha \in [0, 1]$ we have

$$\begin{aligned} \varphi_L(H\alpha) + \theta(H(1 - \alpha)) &= H\alpha - \psi_L(H\alpha) + H(1 - \alpha) - \rho(H(1 - \alpha)) \\ &= H - \psi_L(H\alpha) - \rho(H(1 - \alpha)) \\ &\geq H - \psi_L(H) - \rho(H) = \varphi_L(H) + \theta(H) - H. \end{aligned}$$

Let us now prove the second lower bound. Since φ_L is convex and θ is concave, for every $t \in [0, H]$ we have $\varphi_L(t) \geq \varphi_L(H) + \varphi_L'(H)(t - H)$ and $\theta(t) \geq \theta(H)/Ht$, therefore

$$\begin{aligned} \varphi_L(H\alpha) + \theta(H(1 - \alpha)) &\geq \varphi_L(H) + \varphi_L'(H)(H\alpha - H) + \theta(H)(1 - \alpha) \\ &\geq \varphi_L(H) - (\varphi_L'(H)H - \theta(H))(1 - \alpha). \end{aligned}$$

If $\varphi_L'(H)H \leq \theta(H)$, the minimum is achieved for $1 - \alpha = 0$ and the lower bound is just $\varphi_L(H)$; otherwise it is achieved for $1 - \alpha = 1$, and the lower bound is $\varphi_L(H) - \varphi_L'(H)H + \theta(H)(1 - \alpha)$. $\square$

Remark 4.7. If $\varphi'(0) = 1$, then necessarily $\varphi(t) = t$ for every $t \in [0, \infty)$. In (4.5) we always have $\theta(H)/H \leq 1 = \varphi_L'(H)$, therefore $E(L, H) \geq \theta(H)$. Combined with (4.3) we get $E(L, H) = \theta(H)$.

If on the other hand $\varphi'(0) < 1$, since $\lim_{H \rightarrow 0^+} \theta(H)/H = 1$ we have that for H sufficiently small $\theta(H)/H \geq \varphi_L'(H)$, hence from (4.5) we get $E(L, H) \geq \varphi_L(H)$. Combined with (4.3) we get $E(L, H) = \varphi_L(H)$ for H sufficiently small.

More in general, one can rewrite the right hand side of the lower bound (4.4) as $\varphi_L(H) - \rho(H) = \theta(H) - \psi_L(H)$. Combined with the upper bounds (4.3), this leads to

$$\varphi_L(H) - \rho(H) \leq E(L, H) \leq \varphi_L(H), \quad \theta(H) - \psi_L(H) \leq E(L, H) \leq \theta(H).$$

Since $\rho(H) = o(H)$ for $H \rightarrow 0$, the former estimate is more useful for H small, whereas the latter is more useful for large values of H , especially when θ is sublinear.

Proposition 4.8. *$E(L, H)$ is continuous in $(0, \infty) \times [0, \infty)$ and it admits a continuous extension to $[0, \infty) \times [0, \infty)$, with $E(0, H) = \theta(H)$.*

¹⁴The estimate is available for every element of the subdifferential, but taking the smallest leads to the best one.

¹⁵In the sense that $\lim_{t \rightarrow \infty} \psi(t)/t = 0$.

Proof. Thanks to [Proposition 2.22](#), the function

$$g(L, H, \alpha) = \varphi_L(H\alpha) + \theta(H(1 - \alpha)) = \varphi(H\alpha, L) + \theta(H(1 - \alpha))$$

is continuous in $[0, \infty) \times [0, \infty) \times [0, 1]$, hence it is uniformly continuous in $[0, \bar{L}] \times [0, \bar{H}] \times [0, 1]$ for every $\bar{L}, \bar{H} > 0$. From this we deduce that in $(0, \bar{L}) \times [0, \bar{H})$ the function

$$E(L, H) = \min_{\alpha \in [0, 1]} g(L, H, \alpha)$$

is continuous. Recalling that $\varphi'(\infty) = 1$ and therefore $\varphi(H, 0) = H$, the continuous extension to $L = 0$ is provided by the minimization problem in the right hand side, i.e.

$$\begin{aligned} \min_{\alpha \in [0, 1]} g(0, H, \alpha) &= \min_{\alpha \in [0, 1]} [\varphi(H\alpha, 0) + \theta(H(1 - \alpha))] = \min_{\alpha \in [0, 1]} [H\alpha + \theta(H(1 - \alpha))] \\ &= \min_{\alpha \in [0, 1]} [H - \rho(H(1 - \alpha))] = H - \rho(H) = \theta(H). \end{aligned} \quad \square$$

4.2 The input is not a minimizer

Recall from [Remark 3.2](#) that the Cantor function provides a simple example where the input is not a minimizer even in the nonlinear setting. This is generalized in the following theorem to an input datum g with at least α -growth at one endpoint, providing the nonlinear analogue of [Theorem 3.4](#).

Theorem 4.9. *Let $g \in \text{BV}_{\text{reg}}((a, b))$ be a function with the property that there exist $\alpha, c > 0$ and $0 < \delta < b - a$ such that $g(x) - g^+(a) \geq c(x - a)^\alpha$ for all $x \in (a, a + \delta]$. Assume that there exist $C, \eta > 0$ such that $E(\delta, H) \geq CH^\eta$ for every $H \in [0, \bar{H}]$, where $\bar{H} = \text{ess sup}_{(a, a + \delta]} g - g^+(a)$. If*

$$\gamma > \eta \left(1 + \frac{1}{p\alpha}\right)^{-1} = \frac{p\alpha\eta}{p\alpha + 1}$$

then g is not a minimizer of the functional

$$F(u) = \text{TV}_{\varphi, \theta}(u) + \lambda \|u - g\|_p^\gamma$$

for any $p \in [1, \infty]$ and $\lambda > 0$.

Proof. Define the non-decreasing function $\omega : (a, a + \delta] \rightarrow \mathbb{R}$

$$\omega(x) = \text{ess sup}_{(a, x]} g$$

and observe that we have $\omega(x) \geq g^+(a) + c(x - a)^\alpha$ and $\omega^+(a) = g^+(a)$. Given $y \in (a, a + \delta]$, introduce the auxiliary parameter $h = h(y) = \omega(y) - g^+(a) \geq c(y - a)^\alpha > 0$.

For every $\varepsilon > 0$, the set $\{x \in (a, y] : g(x) \geq \omega(y) - \varepsilon\}$ has positive $\mathcal{L}^1$ -measure, and $\mathcal{L}^1$ -a.e. $x \in (a, y]$ is a continuity point of g , therefore there exists $x_{y, \varepsilon} \in (a, y]$ such that $g(x_{y, \varepsilon}) \geq \omega(y) - \varepsilon$ and g is continuous in $x_{y, \varepsilon}$. From this continuity, we deduce $g(x_{y, \varepsilon}) \leq \omega(y)$, hence $|\omega(y) - g(x_{y, \varepsilon})| \leq \varepsilon$. Define the truncated function $u_{y, \varepsilon} \in \text{BV}((a, b))$

$$u_{y, \varepsilon}(x) = \begin{cases} \omega(y) & x \in (a, x_{y, \varepsilon}), \\ g(x) & x \in [x_{y, \varepsilon}, b). \end{cases}$$

We have

$$\begin{aligned} \text{TV}_{\varphi, \theta}(u_{y, \varepsilon}; (a, b)) &= \text{TV}_{\varphi, \theta}(u_{y, \varepsilon}; (a, x_{y, \varepsilon})) + \theta(|u_{y, \varepsilon}^-(x_{y, \varepsilon}) - u_{y, \varepsilon}^+(x_{y, \varepsilon})|) + \text{TV}_{\varphi, \theta}(u_{y, \varepsilon}; (x_{y, \varepsilon}, b)) \\ &= \theta(|\omega(y) - g(x_{y, \varepsilon})|) + \text{TV}_{\varphi, \theta}(g; (x_{y, \varepsilon}, b)) \\ &\leq \theta(\varepsilon) + \text{TV}_{\varphi, \theta}(g; (x_{y, \varepsilon}, b)). \end{aligned}$$

On the other hand, recalling the continuity of g in $x_{y,\varepsilon}$, $x_{y,\varepsilon} - a \leq y - a \leq \delta$, the monotonicity stated in [Proposition 4.4](#), and the lower bound for $E(\delta, \cdot)$, we have

$$\begin{aligned} \text{TV}_{\varphi,\theta}(g; (a, b)) &= \text{TV}_{\varphi,\theta}(g; (a, x_{y,\varepsilon})) + \text{TV}_{\varphi,\theta}(g; (x_{y,\varepsilon}, b)) \\ &\geq E(x_{y,\varepsilon} - a, |g(x_{y,\varepsilon}) - g^+(a)|) + \text{TV}_{\varphi,\theta}(g; (x_{y,\varepsilon}, b)) \\ &\geq E(\delta, |\omega(y) - g^+(a)| - \varepsilon) + \text{TV}_{\varphi,\theta}(g; (x_{y,\varepsilon}, b)) \\ &= E(\delta, h - \varepsilon) + \text{TV}_{\varphi,\theta}(g; (x_{y,\varepsilon}, b)) \\ &\geq C(h - \varepsilon)^\eta + \text{TV}_{\varphi,\theta}(g; (x_{y,\varepsilon}, b)). \end{aligned}$$

From these two total variation estimates we get

$$\text{TV}_{\varphi,\theta}(g; (a, b)) \geq \text{TV}_{\varphi,\theta}(u_{y,\varepsilon}; (a, b)) + C(h - \varepsilon)^\eta - \theta(\varepsilon).$$

Since for $\mathcal{L}^1$ -a.e. $x \in (a, a + \delta]$ we have $u_{y,\varepsilon}(x) \geq g(x)$, with a similar computation as in [Theorem 3.4](#), we have

$$\|u_{y,\varepsilon} - g\|_{L^p((a,b))}^p \leq K(p, \alpha) C^{-1/\alpha} h^{p+1/\alpha}.$$

Combining the estimates for the total variation and the fidelity term we arrive at

$$F(u_{y,\varepsilon}) - F(g) \leq \theta(\varepsilon) - C(h - \varepsilon)^\eta + \lambda K(p, \alpha)^{\gamma/p} c^{-\gamma/(p\alpha)} h^{\gamma(1+\frac{1}{p\alpha})},$$

which can be made strictly negative under the assumption on the exponent γ sending $\varepsilon \rightarrow 0$ and $y \rightarrow 0$, which implies $h(y) \rightarrow 0$. $\square$

Observe that [Proposition 4.6](#) provides a convenient way to verify the assumption on $E(\delta, \cdot)$. For instance, if $\varphi(t) = ct^\eta + o(t^\eta)$ for $t \rightarrow 0$ with $\eta > 1$, then by [Remark 4.7](#) we have $E(L, H) = \varphi_L(H) = cL^{1-\eta}H^\eta + o(H^\eta)$ for $H \rightarrow 0$, hence the assumption is satisfied.

4.3 Is the input a minimizer?

The intent of this section is to try and replicate the results contained in [Subsection 3.2](#). Unfortunately, as it turns out, the expected statements are not true. Indeed, the culprit lies in the failure of the nonlinear analogue of [Proposition 3.9](#): more specifically, while it would be possible to replicate the majority of its proof, the sub-case (ii)(b) proves to be an insurmountable obstruction, due to the fact that not all monotone functions are minimizers of the boundary value problem for $\text{TV}_{\varphi,\theta}$, a simple fact that was exploited in (ii)(b) to perform the local truncation without degrading the linear total variation.

Case (i) of [Proposition 4.10](#) demonstrates a simple situation where the nonlinear analogues of [Proposition 3.9](#) and [Theorem 3.10](#) both fail, exploiting the behavior at infinity of the nonlinearity φ , see (4.6a). The sharpness of the assumption is highlighted by the opposite behavior under the complementary conditions in case (ii) of the same proposition. In the particular case $p = \gamma$, conditions (4.6) amount to the existence or non-existence of an oblique asymptote for φ ; in the case $\gamma > p$, they are related to a sufficiently fast convergence rate to the oblique asymptote.

A more detailed analysis of (ii)(b) leads us to consider the penalized problem in [Proposition 4.11](#), which reveals the underlying reason for the failure of the desired nonlinear statements: constant functions are not optimal profiles for transitioning between two levels, not even in the presence of a fidelity term with respect to a constant datum.

Proposition 4.10. *Let $L, H > 0$. Let $g \in \text{BV}((0, L))$ given by*

$$g(x) = H\Theta(x - x_0) = \begin{cases} 0 & x < x_0, \\ H & x > x_0, \end{cases}$$

with $x_0 \in (0, L)$. For $\alpha \geq \alpha^* := \frac{H}{2 \min\{x_0, L-x_0\}}$, define the competitor $u_\alpha \in \text{BV}((0, L))$ as

$$u_\alpha(x) = \left(\frac{H}{2} + \alpha(x - x_0) \right) \vee 0 \wedge H = \begin{cases} 0 & x < x_0 - \frac{H}{2\alpha}, \\ \frac{H}{2} + \alpha(x - x_0) & x_0 - \frac{H}{2\alpha} < x < x_0 + \frac{H}{2\alpha}, \\ H & x > x_0 + \frac{H}{2\alpha}. \end{cases}$$

(i) If $\theta(H) = H$ and

$$\inf_{\alpha \geq \alpha^*} \frac{\varphi(\alpha) - \alpha}{\alpha^{1-\gamma/p}} = -\infty, \quad (4.6a)$$

then for every $\lambda \geq 0$ there exists $\alpha \geq \alpha^*$ such that $F(u_\alpha) < F(g)$; hence the input g is not optimal.

(ii) If instead either $\theta(H) < H$, or $\theta(H) = H$ and

$$\inf_{\alpha \geq \alpha^*} \frac{\varphi(\alpha) - \alpha}{\alpha^{1-\gamma/p}} > -\infty, \quad (4.6b)$$

then for every $\lambda \geq 0$ sufficiently large we have $F(u_\alpha) \geq F(g)$ for all $\alpha \geq \alpha^*$; hence the input g performs better than these particular competitors.

Proof. We have $F(g) = \theta(H)$ and

$$F(u_\alpha) = H \frac{\varphi(\alpha)}{\alpha} + \frac{\lambda H^{\gamma(1+1/p)}}{2^\gamma (p+1)^{\gamma/p} \alpha^{\gamma/p}}.$$

Assume $\theta(H) < H$. The difference in energy can be written as

$$F(u_\alpha) - F(g) = \alpha^{-\gamma/p} \left[\alpha^{\gamma/p} \left(H \frac{\varphi(\alpha)}{\alpha} - \theta(H) \right) + \frac{\lambda H^{\gamma(1+1/p)}}{2^\gamma (p+1)^{\gamma/p}} \right].$$

Since $H \frac{\varphi(\alpha)}{\alpha} - \theta(H)$ is a bounded function of α and $\gamma/p > 0$, we have

$$\lim_{\alpha \rightarrow 0^+} \left[\alpha^{\gamma/p} \left(H \frac{\varphi(\alpha)}{\alpha} - \theta(H) \right) \right] = 0.$$

On the other hand, $\lim_{\alpha \rightarrow \infty} \left(H \frac{\varphi(\alpha)}{\alpha} - \theta(H) \right) = H - \theta(H) > 0$, therefore

$$\lim_{\alpha \rightarrow \infty} \left[\alpha^{\gamma/p} \left(H \frac{\varphi(\alpha)}{\alpha} - \theta(H) \right) \right] = \infty.$$

By continuity in α , it follows that

$$m := \inf_{\alpha \geq \alpha^*} \left[\alpha^{\gamma/p} \left(H \frac{\varphi(\alpha)}{\alpha} - \theta(H) \right) \right] \geq \inf_{\alpha > 0} \left[\alpha^{\gamma/p} \left(H \frac{\varphi(\alpha)}{\alpha} - \theta(H) \right) \right] > -\infty.$$

This implies that for every λ sufficiently large, namely $\lambda > -\frac{2^\gamma (p+1)^{\gamma/p}}{H^{\gamma(1+1/p)}} m$, we have $F(u_\alpha) - F(g) > 0$ for every α .

Assume $\theta(H) = H$. In this case the difference in energy can be written as

$$F(u_\alpha) - F(g) = \alpha^{-\gamma/p} \left[H \frac{\varphi(\alpha) - \alpha}{\alpha^{1-\gamma/p}} + \frac{\lambda H^{\gamma(1+1/p)}}{2^\gamma (p+1)^{\gamma/p}} \right].$$

In case (i) we have that for every $\lambda \geq 0$ the infimum of the square bracket is $-\infty$, therefore there exists $\alpha \geq \alpha^*$ for which it is negative, hence $F(u_\alpha) - F(g) < 0$. In case (ii) we have that for λ sufficiently large, namely

$$\lambda \geq -2^\gamma(p+1)^{\gamma/p} H^{1-\gamma(1+1/p)} \inf_{\alpha \geq \alpha^*} \frac{\varphi(\alpha) - \alpha}{\alpha^{1-\gamma/p}},$$

the infimum of the square bracket is non-negative, hence $F(u_\alpha) - F(g) \geq 0$ for all $\alpha \geq \alpha^*$. $\square$

To replicate the truncation argument used in (ii)(b), we would need the zero function to be a global minimizer of the functional

$$F(u) = \text{TV}_{\varphi, \theta}(u; (0, L)) + \theta(|H - u^-(L)|) + \lambda \|u\|_p^\gamma \quad (4.7)$$

among all functions $u \in BV((0, L))$ with $u^+(0) = 0$. In Proposition 4.11 below we show that this is not the case.

By a line of reasoning similar to that in the proof of Theorem 4.1, we can show that a minimizer u must be increasing (see the proof of monotonicity at the beginning of the proof) and with values in $[0, H]$, its Cantor and jump part can be moved forward to $x = L$ lowering both the total variation and the fidelity term, hence we may assume $u \in W^{1,1}((0, L))$, and finally u must be convex since the increasing rearrangement of u' does not change the total variation but lowers the fidelity.

Proposition 4.11. *Let $L, H > 0$, let $p, \gamma \geq 1$, and assume $\theta'_-(H) > \varphi'_+(0)$. Then the function $g = 0$ is not a local minimizer in $BV((0, L))$ of the functional*

$$F(u) = \text{TV}_{\varphi, \theta}(u; (0, L)) + \theta(|H - u^-(L)|) + \lambda \|u\|_p^\gamma$$

with boundary condition $u^+(0) = 0$ for any $\lambda > 0$.

Proof. Considering non-negative variations $v \in W^{1,1}((0, L); [0, \infty))$ with $v^+(0) = 0$, the one-sided directional derivative of the functional F at $u = g$ is (see Appendix B)

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0^+} \frac{F(g + \varepsilon v) - F(g)}{\varepsilon} &= \int_0^L \varphi'_+(0) v'(x) dx - \theta'_-(H) v^-(L) + \delta_{\gamma,1} \lambda \|v\|_p \\ &= (\varphi'_+(0) - \theta'_-(H)) v^-(L) + \delta_{\gamma,1} \lambda \|v\|_p. \end{aligned}$$

For any λ , this can be made strictly negative by taking $v(x) = (x/L)^q$ with $q > 0$ large enough. This disproves the local minimality of $g \in W^{1,1}((0, L))$, hence in $BV((0, L))$. $\square$

4.3.1 Numerical examples

We conclude this section by investigating numerically the nature of the global minimizer in a simple situation. Let us take $p = \gamma = 1$ and φ, θ differentiable. Recall that the minimizer is a convex function $u \in W^{1,1}((0, L); [0, H])$ with $u^+(0) = 0$. Let $x_0 = \max\{x \in [0, L] : u|_{(0,x]} = 0\}$. From Subsection B.1.1 we know that in the interval (x_0, L) the optimal profile u is given by (B.7):

$$u(x) = \begin{cases} 0 & x \in (0, x_0], \\ \frac{1}{\lambda} \varphi^*(\varphi'(0) + \lambda(x - x_0)) & x \in (x_0, L). \end{cases}$$

In this sub section, we consider two explicit examples:

$$\begin{aligned} \varphi_1(t) &= t - \log(1+t), & \varphi_2(t) &= \sqrt{1+t^2} - 1, \\ \theta_1(t) &= \log(1+t), & \theta_2(t) &= \sqrt{1+2t} - 1. \end{aligned}$$

For the non-linearity φ_1 we have $\varphi_1'(0) = 0$ and $\varphi_1^*(\tau) = -\tau - \log(1 - \tau)$, therefore in the interval (x_0, L)

$$u(x) = -(x - x_0) - \frac{\log(1 - \lambda(x - x_0))}{\lambda}.$$

On the other hand, for φ_2 we have $\varphi_2'(0) = 0$ and $\varphi_2^*(\tau) = 1 - \sqrt{1 - \tau^2}$, hence in the interval (x_0, L)

$$u(x) = \frac{1 - \sqrt{1 - \lambda^2(x - x_0)^2}}{\lambda}.$$

These two families of profiles are depicted in Figure 3.

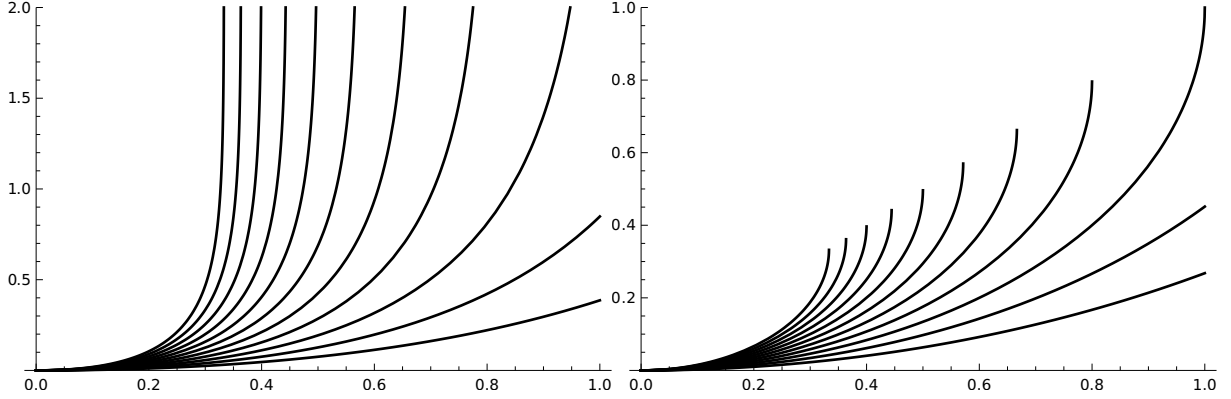


Figure 3: profiles u shifted at $x_0 = 0$ for the non-linearities φ_1 (left) and φ_2 (right), with the parameter λ ranging from $1/2$ (lowest curves) to 3 (highest curves) in steps of $1/4$. Notice that in both cases the interval of definition of the profiles is $[0, 1/\lambda]$.

In Figure 4 we illustrate some minimizers u of the functional $F(u)$ for the two sets of non-linearities and various choices of the parameter λ . These are computed numerically through a search of local minimizers and it is possible to verify that they follow the shape of the analytical profiles introduced above.

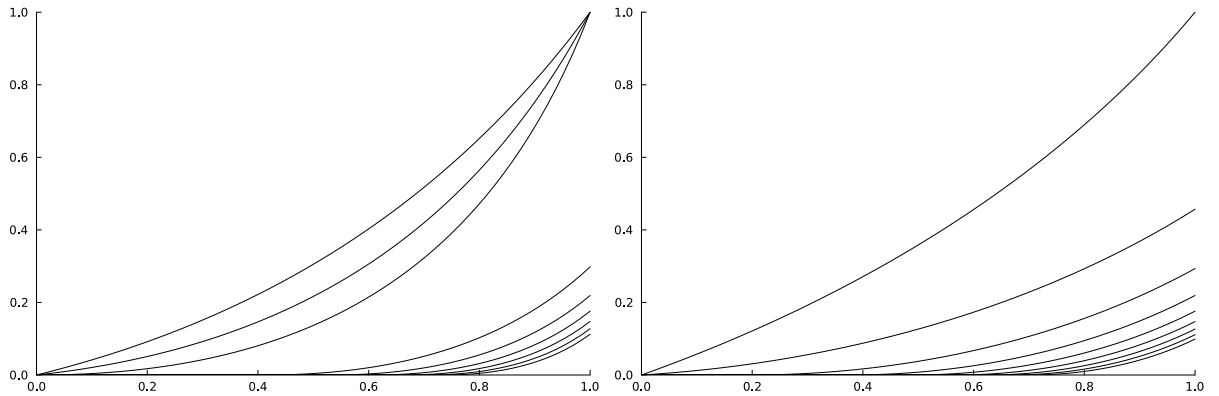


Figure 4: minimizers of F with non-linearities φ_1, θ_1 (left) and φ_2, θ_2 (right), $L = H = 1$, and with the parameter λ ranging from 0.4 (highest curves) to 2 (lowest curves) in steps of 0.2 .

In Figure 5 we plot the functions u , u' and $\varphi'(u')$, where u is the minimizer of the problem with non-linearities φ_1, θ_1 and parameters $L = H = 1$, $\lambda = 1.1$. As predicted by the ODE (B.4), the function $\varphi'(u')$ appears to be affine in the interval where $u > 0$.

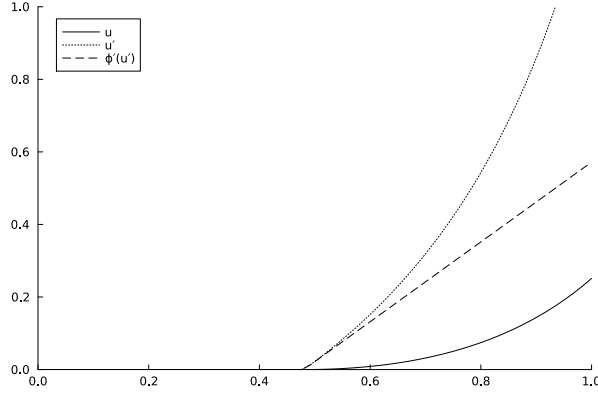


Figure 5: plot of u , u' and $\varphi'(u')$, where u is the minimizer of the problem with non-linearities φ_1, θ_1 and parameters $L = H = 1$, $\lambda = 1.1$; notice that in the interval where $u > 0$ the function $\varphi'(u')$ is affine, in accordance with (B.4). The value of x_0 , computed numerically with both Julia and Mathematica, is approximately $x_0 \approx 0.4801305$.

Numerical implementation The numerical approximations of the optimal profiles are obtained as follows. Knowing that the minimizer of the functional F should be a non-negative continuous function¹⁶, for a fixed $n \in \mathbb{N}_+$ we introduce the uniform mesh $x_i = i\Delta x$, where $\Delta x = L/n$, and approximate the unknown function u with a continuous piecewise affine function

$$\hat{u}(x) = \sum_{i=1}^n \left(\frac{x_i - x}{x_i - x_{i-1}} u_{i-1} + \frac{x - x_{i-1}}{x_i - x_{i-1}} u_i \right) \mathbf{1}_{(x_{i-1}, x_i]}(x)$$

parametrized by the vector $(u_1, \dots, u_n) \in [0, \infty)^n$, with the convention $u_0 = 0$. The values $u_i = \hat{u}(x_i)$ correspond to a sampling of u over the uniform mesh. This leads to introducing the function $F_n : [0, \infty)^n \rightarrow [0, \infty)$ given by

$$\begin{aligned} F_n(u_1, \dots, u_n) &= F(\hat{u}) = \sum_{i=1}^n \varphi \left(\left| \frac{u_i - u_{i-1}}{\Delta x} \right| \right) \Delta x + \theta(|H - u_n|) + \lambda \left(\sum_{i=1}^{n-1} u_i \Delta x + \frac{u_n}{2} \Delta x \right) \\ &= \left(\sum_{i=1}^n \varphi \left(\left| \frac{u_i - u_{i-1}}{\Delta x} \right| \right) + \lambda \left(\sum_{i=1}^{n-1} u_i + \frac{u_n}{2} \right) \right) \Delta x + \theta(|H - u_n|), \end{aligned}$$

which we minimize among the vectors $(u_1, \dots, u_n) \in [0, H]^n$. The function F_n amounts to approximating the L^1 norm in the fidelity with the trapezoidal rule, the derivative u' with finite differences, and the integral of $\varphi(|u'|)$ with the rectangle rule.

Alternatively, the problem can be reformulated in terms of the discrete slopes $(a_1, \dots, a_n) \in \mathbb{R}^n$, which are related to the values $(u_1, \dots, u_n) \in \mathbb{R}^n$ by the transformation

$$a_i = \frac{u_i - u_{i-1}}{\Delta x}, \quad u_i = \sum_{j=1}^i a_j \Delta x.$$

Notice that the mapping $\mathbb{R}^n \rightarrow \mathbb{R}^n$ between $(u_i)_{i=1}^n$ and $(a_i)_{i=1}^n$ is a bijection. Recalling the monotonicity of the optimal profile u , the slopes $(a_i)_{i=1}^n$ can be assumed to be non-negative.

¹⁶ Additionally, $u \in W^{1,1}((0, L); [0, H])$ non-decreasing and convex.

The function $G_n : [0, \infty)^n \rightarrow [0, \infty)$ which corresponds to F_n in this set of coordinates, i.e. $G_n(a_1, \dots, a_n) = F_n(u_1, \dots, u_n)$, is given by

$$\begin{aligned} G_n(a_1, \dots, a_n) &= \Delta x \left(\sum_{i=1}^n \varphi(a_i) + \lambda \Delta x \left(\sum_{i=1}^{n-1} \sum_{j=1}^i a_j + \frac{1}{2} \sum_{j=1}^n a_j \right) \right) + \theta \left(\left| H - \Delta x \sum_{j=1}^n a_j \right| \right) \\ &= \Delta x \left(\sum_{i=1}^n \varphi(a_i) + \lambda \Delta x \left(\sum_{i=1}^n a_i (n - i + 1/2) \right) \right) + \theta \left(\left| H - \Delta x \sum_{j=1}^n a_j \right| \right). \end{aligned}$$

Not only the formulation in terms of the slopes avoids the bad conditioning arising from the finite difference operator, but it also allows to easily restrict the search to monotone functions instead of merely non-negative ones. These benefits manifest empirically in a faster convergence of the optimization algorithm.

The actual numerical minimization has been implemented in the Julia programming language [Bez+17] using the box-constrained limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS-B) algorithm provided by the package `Optim.jl` [MR18]. In addition, a multistart strategy from some hand-picked and some random initial points has been employed in order to reduce the risk of converging to a local minimum.

The approximate profiles obtained via minimization of G_n have been validated by comparison with the semi-analytical approach described in Subsection B.1.1, which consists in minimizing the scalar function defined in (B.9) subjected to the constraint (B.10).

As an example, in Figure 6 we plot the function f defined in (B.9) associated with non-linearities φ_1, θ_1 and parameters $L = H = 1$, $\lambda = 1.1$; the minimum is found at $x_0 \approx 0.4801305$, as already mentioned in Figure 5. In Figure 7 we illustrate the similar situation for φ_2, θ_2 , $L = 1$, $H = 0.55$ and $\lambda = 1.5$.

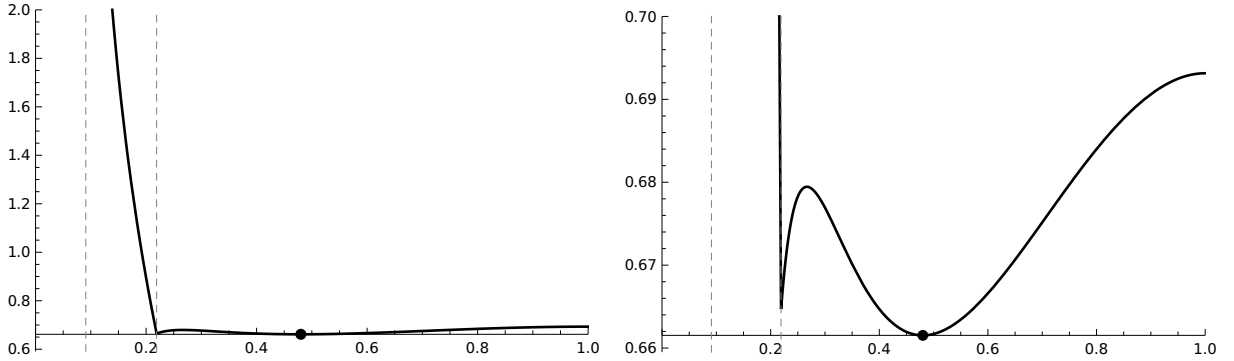


Figure 6: plot of the function f defined in (B.9) associated with non-linearities φ_1, θ_1 and parameters $L = H = 1$, $\lambda = 1.1$. The same plot is presented at two different vertical zoom levels. The vertical dashed lines, from left to right, are positioned at $L - \frac{1-\varphi'(0)}{\lambda} \approx 0.0909091$ and $L - \frac{(\varphi^*)^{-1}(\lambda H) - \varphi'(0)}{\lambda} \approx 0.21909$. The minimum, highlighted with a black dot, is found at $x_0 \approx 0.4801305$, as already mentioned in Figure 5.

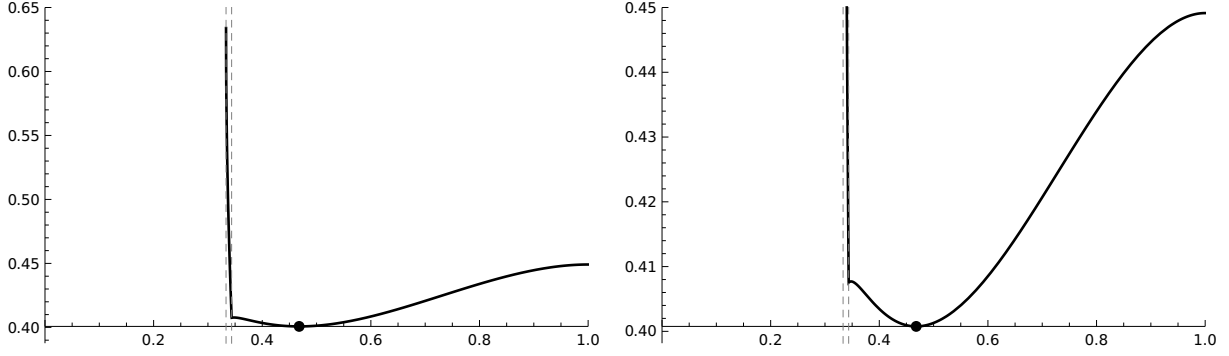


Figure 7: plot of the function f defined in (B.9) associated with non-linearities φ_2, θ_2 and parameters $L = 1$, $H = 0.55$, $\lambda = 1.5$. The same plot is presented at two different vertical zoom levels. The vertical dashed lines, from left to right, are positioned at $L - \frac{1-\varphi'(0)}{\lambda} \approx 0.333333$ and $L - \frac{(\varphi^*)^{-1}(\lambda H) - \varphi'(0)}{\lambda} \approx 0.343621$. The minimum, highlighted with a black dot, is found at $x_0 \approx 0.467959$.

A Miscellanea

A.1 Alternative proofs of Proposition 4.4(i)

Alternative proof of Proposition 4.4(i), from definition (4.1). Monotonicity in L . Let $L_1 \leq L_2$ and $u \in \text{BV}((0, L_1))$ be optimal for $E(L_1, H)$ and increasing. Define $v \in \text{BV}((0, L_2))$ by $v(x) = u(L_1/L_2 x)$. The limits at the endpoints are preserved, so v is admissible for $E(L_2, H)$. Setting $T(x) = L_2/L_1 x$, we have $Dv = T_\# Du$ and also $D^{a,c,j}v = T_\# D^{a,c,j}u$. Owing to the bijectivity of T , the contributions of D^c and D^j to the functional $\text{TV}_{\varphi,\theta}$ are preserved, because

$$|D^c v|((0, L_2)) = |T_\# D^c u|((0, L_2)) = T_\# |D^c u|((0, L_2)) = |D^c u|(T^{-1}((0, L_2))) = |D^c u|((0, L_1))$$

and $|D^j v|$ and $|D^j u|$ have the same atom values. On the other hand, since $L_1/L_2 \leq 1$, we have $\varphi((L_1/L_2)t) \leq (L_1/L_2)\varphi(t)$ for every $t \geq 0$, hence

$$\begin{aligned} \int_0^{L_2} \varphi(v'(x)) \, dx &= \int_0^{L_2} \varphi\left(\frac{L_1}{L_2} u'\left(\frac{L_1}{L_2} x\right)\right) \, dx \\ &\leq \int_0^{L_2} \frac{L_1}{L_2} \varphi\left(u'\left(\frac{L_1}{L_2} x\right)\right) \, dx = \int_0^{L_1} \varphi(u'(x)) \, dx \end{aligned}$$

Therefore $E(L_2, H) \leq \text{TV}_{\varphi,\theta}(v; (0, L_2)) = \text{TV}_{\varphi,\theta}(u; (0, L_1)) \leq E(L_1, H)$.

Monotonicity in H . Let $H_1 \leq H_2$ and $u \in \text{BV}((0, L))$ optimal for $E(L, H_2)$. Define $v \in \text{BV}((0, L))$ by $v(x) = (H_1/H_2)u(x)$. We have $v^+(0) = 0$ and $v^-(L) = H_1$, so v is admissible for $E(L, H_1)$. Since φ and θ are monotone and $|Dv| = (H_1/H_2)|Du| \leq |Du|$, we get $E(L, H_1) \leq \text{TV}_{\varphi,\theta}(v; (0, L)) \leq \text{TV}_{\varphi,\theta}(u; (0, L)) = E(L, H_2)$. $\square$

Alternative proof of Proposition 4.4(i), from (4.2). Monotonicity in L . Recall the formula

$$E(L, H) = \min \left\{ L\varphi\left(\frac{H-b}{L}\right) + \theta(b) : 0 \leq b \leq H \right\}$$

in (4.2). Observe that for every $t \geq 0$ the function $L \mapsto L\varphi(t/L)$ is decreasing. Indeed, if $L_1 \leq L_2$, then $\varphi(t/L_2) = \varphi((L_1/L_2)(t/L_1)) \leq (L_1/L_2)\varphi(t/L_1)$. This means that the objective function to minimize is decreasing in L , and therefore so is $E(L, H)$.

Monotonicity in H . Let $L_1 \leq L_2$ and $a, b \geq 0$ with $La + b = H_2$ be optimal in (4.2) for $E(L, H_2)$. Setting $\tilde{a} = (H_1/H_2)a \leq a$ and $\tilde{b} = (H_1/H_2)b \leq b$, we have $L\tilde{a} + \tilde{b} = H_1$. Since φ and θ are increasing, we deduce $E(L, H_1) \leq L\varphi(\tilde{a}) + \theta(\tilde{b}) \leq L\varphi(a) + \theta(b) = E(L, H_2)$. $\square$

B First variation of the functional in Subsection 4.3

Let $\sigma : [0, \infty) \rightarrow \mathbb{R}$ be a function which admits left and right derivatives at every point. We introduce the following notation for the one-sided directional derivative of the composition $\sigma \circ |\cdot|$:

$$\mathcal{D}_\sigma(a)[b] := (\sigma \circ |\cdot|)'_{\text{sign } b}(a)b = \lim_{\varepsilon \rightarrow 0^+} \frac{\sigma(|a + \varepsilon b|) - \sigma(|a|)}{\varepsilon} = \begin{cases} \sigma'_+(0)|b| & a = 0, \\ \sigma'_{\text{sign}(ab)}(|a|) \text{sign}(a)b & a \neq 0, \end{cases}$$

where σ'_-, σ'_+ denote the left and right derivatives respectively. Notice that, for every fixed $a \in \mathbb{R}$, the function $\mathcal{D}_\sigma(a)[b]$ is positively 1-homogeneous in b . Moreover, if σ is differentiable everywhere, then $\mathcal{D}_\sigma(a)[b] = \sigma'(|a|) \text{sign}(a)b$ is linear in b for $a \neq 0$, and it is linear for $a = 0$ if and only if $\sigma'_+(0) = 0$.

Let us compute the one-sided Gâteaux derivative of the functional F . Starting with the total variation, for $u, v \in W^{1,1}((0, L))$ we have

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0^+} \frac{\text{TV}_{\varphi, \theta}(u + \varepsilon v) - \text{TV}_{\varphi, \theta}(u)}{\varepsilon} &= \int_0^L \lim_{\varepsilon \rightarrow 0^+} \frac{\varphi(|u'(x) + \varepsilon v'(x)|) - \varphi(|u'(x)|)}{\varepsilon} dx \\ &= \int_0^L \mathcal{D}_\varphi(u'(x))[v'(x)] dx; \end{aligned}$$

on the other hand, the derivative of the penalization at the right endpoint is

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\theta(|u^-(L) + \varepsilon v^-(L) - H|) - \theta(|u^-(L) - H|)}{\varepsilon} = \mathcal{D}_\theta(u^-(L) - H)[v^-(L)].$$

Let us now compute the one-sided Gâteaux derivative of $\|u\|_p^\gamma$. When $u = 0$ is the identically zero function, we have

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\|u + \varepsilon v\|_p^\gamma - \|u\|_p^\gamma}{\varepsilon} = \lim_{\varepsilon \rightarrow 0^+} \frac{\varepsilon^\gamma}{\varepsilon} \|v\|_p^\gamma = \begin{cases} 0 & \gamma > 1, \\ \|v\|_p & \gamma = 1, \\ \infty & \gamma < 1. \end{cases}$$

Assume now $u \neq 0$, hence $\|u\|_p > 0$. For $p = 1$ we have

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0^+} \frac{\|u + \varepsilon v\|_1^p - \|u\|_1^p}{\varepsilon} &= \lim_{\varepsilon \rightarrow 0^+} \int_0^L \frac{|u(x) + \varepsilon v(x)| - |u(x)|}{\varepsilon} dx \\ &= \int_{\{u \neq 0\}} \text{sign } u(x) v(x) dx + \int_{\{u=0\}} |v(x)| dx, \end{aligned}$$

whereas for $p > 1$ we have

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0^+} \frac{\|u + \varepsilon v\|_p^p - \|u\|_p^p}{\varepsilon} &= \lim_{\varepsilon \rightarrow 0^+} \int_0^L \frac{|u(x) + \varepsilon v(x)|^p - |u(x)|^p}{\varepsilon} dx \\ &= \int_0^L p|u(x)|^{p-1} \text{sign } u(x) v(x) dx \\ &= \int_{\{u \neq 0\}} p|u(x)|^{p-1} \text{sign } u(x) v(x) dx. \end{aligned}$$

These last two formulas can be combined succinctly as

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\|u + \varepsilon v\|_p^p - \|u\|_p^p}{\varepsilon} = \int_{\{u \neq 0\}} |u(x)|^{p-1} \text{sign } u(x) v(x) dx + \delta_{p,1} \int_{\{u=0\}} |v(x)| dx, \quad (\text{B.1})$$

valid for all $p \geq 1$, where $\delta_{p,1}$ denotes the Kronecker delta. Writing $\|u\|_p^\gamma = (\|u\|_p^p)^{\gamma/p}$, by the chain rule we finally get

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\|u + \varepsilon v\|_p^\gamma - \|u\|_p^\gamma}{\varepsilon} = \frac{\gamma}{p} \|u\|_p^{\gamma-p} \lim_{\varepsilon \rightarrow 0^+} \frac{\|u + \varepsilon v\|_p^p - \|u\|_p^p}{\varepsilon},$$

where this latter right derivative is given by (B.1).

In conclusion, the first variation of the functional F at the identically zero function $u = 0$ is

$$\lim_{\varepsilon \rightarrow 0^+} \frac{F(0 + \varepsilon v) - F(0)}{\varepsilon} = \int_0^L \varphi'_+(0) |v'(x)| dx + \theta'_{\text{sign } v^-(L)}(H) v^-(L) + \begin{cases} 0 & \gamma > 1, \\ \lambda \|v\|_p & \gamma = 1, \\ \infty & \gamma < 1; \end{cases}$$

whereas for $u \not\equiv 0$ instead we have

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0^+} \frac{F(u + \varepsilon v) - F(u)}{\varepsilon} \\ &= \int_0^L \mathcal{D}_\varphi(u'(x)) [v'(x)] dx + \mathcal{D}_\theta(u^-(L) - H) [v^-(L)] \\ & \quad + \lambda \gamma \|u\|_p^{\gamma-p} \left(\int_{\{u \neq 0\}} |u(x)|^{p-1} \text{sign } u(x) v(x) dx + \delta_{p,1} \int_{\{u=0\}} |v(x)| dx \right). \end{aligned}$$

B.1 Differentiable setting

If φ and θ are differentiable everywhere, we can expand out the definition of $\mathcal{D}_\varphi, \mathcal{D}_\theta$ and obtain the simplified expression

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0^+} \frac{F(u + \varepsilon v) - F(u)}{\varepsilon} \\ &= \int_{\{u' \neq 0\}} \varphi'(u'(x)) \text{sign}(u'(x)) v'(x) dx + \int_{\{u'=0\}} \varphi'(0) |v'(x)| dx \\ & \quad + \begin{cases} |v^-(L)| & u^-(L) = H \\ \theta'(|u^-(L) - H|) \text{sign}(u^-(L) - H) v^-(L) & u^-(L) \neq H \end{cases} \\ & \quad + \lambda \gamma \|u\|_p^{\gamma-p} \left(\int_{\{u \neq 0\}} |u(x)|^{p-1} \text{sign } u(x) v(x) dx + \delta_{p,1} \int_{\{u=0\}} |v(x)| dx \right). \end{aligned}$$

B.1.1 Special case $p = \gamma = 1$

Since $u(x) = \int_0^x u'(y) dy$, for non-decreasing functions the functional can be rewritten as

$$\begin{aligned} F(u) &= \int_0^L \varphi(u'(x)) dx + \theta(H - u^-(L)) + \lambda \int_0^L \int_0^x u'(y) dy dx \\ &= \int_0^L [\varphi(u'(x)) + \lambda(L - x)u'(x)] dx + \theta(H - u^-(L)). \end{aligned}$$

The first integral is convex in u' , whereas the second addend is concave.

Optimality conditions Let us derive simplified optimality conditions obtained from vertical variations.

- If $u^-(L) = H$, the first variation with respect to $v \in C^1((0, L))$ with $v|_{(0, x_0]} = 0$ is

$$\lim_{\varepsilon \rightarrow 0^+} \frac{F(u + \varepsilon v) - F(u)}{\varepsilon} = \int_{x_0}^L [\varphi'(u'(x))v'(x) + \lambda(L - x)v'(x)] dx + |v^-(L)|.$$

When $v^-(L) \geq 0$ the one-sided variation is

$$\int_{x_0}^L [\varphi'(u'(x)) + \lambda(L - x) + 1]v'(x) dx.$$

If u is a local minimizer, this must be positive for every v such that $\int_{x_0}^L v'(x) dx \geq 0$, hence $\varphi'(u'(x)) + \lambda(L - x) + 1$ must be a non-negative constant.

When $v^-(L) \leq 0$ the one-sided variation is

$$\int_{x_0}^L [\varphi'(u'(x)) + \lambda(L - x) - 1]v'(x) dx.$$

If u is a local minimizer, this must be positive for every v such that $\int_{x_0}^L v'(x) dx \leq 0$, hence $\varphi'(u'(x)) + \lambda(L - x) - 1$ must be a non-positive constant.

Putting these two conditions together we get that $\varphi'(u'(x)) + \lambda(L - x)$ must be a constant in $[0, 1]$.

- If instead $u^-(L) < H$, then for every variation $v \in C^1((0, L))$ with $v|_{(0, x_0]} = 0$ we have

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \frac{F(u + \varepsilon v) - F(u)}{\varepsilon} &= \int_{x_0}^L [\varphi'(u'(x))v'(x) + \lambda(L - x)v'(x)] dx - \theta'(H - u^-(L))v^-(L) \\ &= \int_{x_0}^L [\varphi'(u'(x)) + \lambda(L - x) - \theta'(H - u^-(L))]v'(x) dx. \end{aligned}$$

If u is a local minimizer, then this variation must be zero, hence by the arbitrariness of v we deduce

$$\varphi'(u'(x)) + \lambda(L - x) = \theta'(H - u^-(L)) \quad \text{for a.e. } x \in (x_0, L). \quad (\text{B.2})$$

Under continuity assumptions, evaluating at $x = L$ we get

$$\varphi'(u'(L)) = \theta'(H - u^-(L)).$$

We may also consider horizontal variations arising from translations of u . Given $\alpha \in \mathbb{R}$, define the extended function

$$u_\alpha(x) = \begin{cases} 0 & x \leq 0, \\ u(x) & x \in (0, L) \\ u^-(L) + \alpha(x - L), & x \geq L, \end{cases}$$

and observe that $F(u) = F(u_\alpha)$.

- *Right translation.* Since $u'(L) \geq 0$, we have

$$\lim_{\tau \rightarrow 0^+} \frac{F(u_\alpha(\cdot - \tau)) - F(u_\alpha)}{\tau} = -\varphi(u'(L)) + \theta'(H - u^-(L))u'(L) - \lambda u^-(L) \geq 0.$$

- *Left translation.* The left translation is admissible if and only if u is identically zero in a right neighborhood of 0. If $u^-(L) < H$, we have

$$\lim_{\tau \rightarrow 0^+} \frac{F(u_\alpha(\cdot + \tau)) - F(u_\alpha)}{\tau} = \varphi(\alpha) - \theta'(H - u^-(L))\alpha + \lambda u^-(L) \geq 0,$$

from which we deduce the necessary condition

$$\lambda u^-(L) \geq \sup_{\alpha \in \mathbb{R}} (\theta'(H - u^-(L))\alpha - \varphi(\alpha)) = \varphi^*(\theta'(H - u^-(L))) = \varphi^*(\varphi'(u'(L))).$$

If instead $u^-(L) = H$, then

$$\lim_{\tau \rightarrow 0^+} \frac{F(u_\alpha(\cdot + \tau)) - F(u_\alpha)}{\tau} = \varphi(\alpha) + \theta'(H - u^-(L))|\alpha| + \lambda u^-(L) \geq 0,$$

which is a condition always satisfied since the three summands are individually non-negative.

Recall that the minimizer is a convex function $u \in W^{1,1}((0, L); [0, H])$ with $u^+(0) = 0$. Let $x_0 = \max\{x \in [0, L] : u|_{(0, x]} = 0\}$. From (B.2) we deduce that the optimal profile u must solve the ODE

$$\varphi'(u'(x)) + \lambda(L - x) = \theta'(H - u^-(L)) =: C, \quad (\text{B.3})$$

where C is a constant. From the inequalities $\varphi'(0) \leq \varphi' \leq \varphi'(\infty) = 1$ and $\theta'(\infty) \leq \theta' \leq \theta'(0) = 1$, evaluating (B.3) at $x = L$ we deduce $\max\{\varphi'(0), \theta'(\infty)\} \leq C \leq 1$; evaluating (B.3) at $x = x_0$ we deduce $\lambda(L - x_0) \leq 1 - \varphi'(0)$, that is

$$x_0 \geq L - \frac{1 - \varphi'(0)}{\lambda}.$$

If $\lambda L > 1 - \varphi'(0)$, this ensures $x_0 > 0$.¹⁷

We now wish to find a condition relating x_0 and C . To do so, we perform yet another variation in the following manner. Assuming $x_0 \in (0, L)$, fix an arbitrary point $\bar{x} \in (x_0, L)$ and, extending $u = 0$ before 0 as usual, define the function

$$u_\varepsilon(x) = \begin{cases} u((1 + \varepsilon)(x - \bar{x}) + \bar{x}) & x \in (0, \bar{x}), \\ u(x) & x \in [\bar{x}, L]. \end{cases}$$

A straightforward computation shows that

$$0 = \frac{d}{d\varepsilon} \Big|_{\varepsilon=0} F(u_\varepsilon) = \int_{x_0}^{\bar{x}} \{ \varphi'(u'(x)) [u'(x) + u''(x)(x - \bar{x})] + \lambda u(x) \} dx,$$

from which, dividing by $\bar{x} - x_0$ and sending $\bar{x} \rightarrow x_0^+$, we infer

$$0 = \varphi'(u'(x_0))u'(x_0) + \lambda u(x_0) = \varphi'(u'(x_0))u'(x_0).$$

If $\varphi' > 0$, then necessarily $u'(x_0) = 0$, hence $\varphi'(u'(x_0)) = \varphi'(0)$. Otherwise, let $[0, \delta] = \{t \geq 0 : \varphi'(t) = 0\}$. From the above condition we deduce either $u'(x_0) = 0$ or $u'(x_0) \in [0, \delta]$, therefore again $\varphi'(u'(x_0)) = \varphi'(0)$. Evaluating (B.3) at $x = x_0$ yields $C = \varphi'(0) + \lambda(L - x_0)$, and consequently (B.3) becomes

$$\varphi'(u'(x)) = \varphi'(0) + \lambda(x - x_0). \quad (\text{B.4})$$

¹⁷Incidentally, observe that in the linear case $\varphi(t) = t$ we have $\lambda(L - x_0) \leq 1 - 1 = 0$, therefore $x_0 = L$ and the optimal profile cannot detach from 0. This is consistent with the results in Subsection 3.2.

Explicit formula for the optimal profile Recall the Fenchel–Young inequality $\varphi(t) + \varphi^*(\tau) \geq \tau t$ with its equality conditions

$$\varphi(t) + \varphi^*(\tau) = t\tau \quad \Longleftrightarrow \quad \tau \in \partial\varphi(t) \quad \Longleftrightarrow \quad t = \partial\varphi^*(\tau) \quad (\text{B.5})$$

expressed in terms of the sub-differential. Even when φ is not strictly convex, the ODE (B.4) implies $u'(x) \in \partial\varphi^*(\varphi'(0) + \lambda(x - x_0))$. Since φ^* is locally Lipschitz in its domain, by Rademacher's theorem it is differentiable almost everywhere,¹⁸ therefore

$$u'(x) = (\varphi^*)'(\varphi'(0) + \lambda(x - x_0)) \quad \text{for almost every } x \in (x_0, L). \quad (\text{B.6})$$

This is sufficient to infer that

$$\begin{aligned} u(x) &= \int_{x_0}^x (\varphi^*)'(\varphi'(0) + \lambda(y - x_0)) \, dy = \frac{1}{\lambda} \int_{\varphi'(0)}^{\varphi'(0) + \lambda(x - x_0)} (\varphi^*)'(z) \, dz \\ &= \frac{1}{\lambda} [\varphi^*(\varphi'(0) + \lambda(x - x_0)) - \varphi^*(\varphi'(0))] = \frac{1}{\lambda} \varphi^*(\varphi'(0) + \lambda(x - x_0)). \end{aligned} \quad (\text{B.7})$$

Formula for the resulting functional Setting $\tau = \varphi'(t)$ in (B.5) shows that

$$\varphi(t) = t\varphi'(t) - \varphi^*(\varphi'(t)).$$

Evaluating it at $t = u'(x)$ and using (B.4) and (B.7) leads to

$$\begin{aligned} \varphi(u'(x)) + \lambda u(x) &= u'(x)\varphi'(u'(x)) - \varphi^*(\varphi'(u'(x))) + \varphi^*(\varphi'(0) + \lambda(x - x_0)) \\ &= u'(x)(\varphi'(0) + \lambda(x - x_0)). \end{aligned}$$

Using (B.6), we get

$$\begin{aligned} \int_{x_0}^L \{\varphi(u'(x)) + \lambda u(x)\} \, dx &= \int_{x_0}^L (\varphi'(0) + \lambda(x - x_0)) (\varphi^*)'(\varphi'(0) + \lambda(x - x_0)) \, dx \\ &= \frac{1}{\lambda} \int_{\varphi'(0)}^{\varphi'(0) + \lambda(L - x_0)} y (\varphi^*)'(y) \, dy, \end{aligned}$$

hence the functional becomes

$$F(u) = \frac{1}{\lambda} \int_{\varphi'(0)}^{\varphi'(0) + \lambda(L - x_0)} y (\varphi^*)'(y) \, dy + \theta(|H - u(L)|), \quad (\text{B.8})$$

which has to be minimized over $x_0 \in \left[L - \frac{1 - \varphi'(0)}{\lambda}, L\right]$. In addition, we know that the minimizer satisfies $u(L) \leq H$; therefore, let us find a more explicit condition on x_0 which ensures this property and allows to remove the absolute value in the argument of θ .

We claim that $\varphi^* : [\varphi'(0), 1] \rightarrow [0, \varphi^*(1)]$ is strictly increasing, continuous and bijective. Suppose that $\varphi^*(\tau) = 0$ for some $\tau \in [\varphi'(0), 1]$; then

$$\varphi(t) - \varphi(0) = \varphi(t) = \varphi(t) + \varphi^*(\tau) \geq \tau t,$$

from which, dividing by t and letting $t \rightarrow 0^+$, we get $\varphi'(0) \geq \tau$, hence $\tau = \varphi'(0)$. This proves that $\tau > \varphi'(0)$ implies $\varphi^*(\tau) > 0$. Since φ^* is convex, for every $\sigma > \tau$ we have

$$\varphi^*(\sigma) \geq \varphi^*(0) + \frac{\sigma}{\tau} (\varphi^*(\tau) - \varphi^*(0)) = \frac{\sigma}{\tau} \varphi^*(\tau) > \varphi^*(\tau),$$

¹⁸Actually, being one-dimensional, it is differentiable outside of a countable set.

hence the claimed monotonicity. The continuity follows from the basic properties of the Legendre transform. The bijectivity follows from continuity and strict monotonicity.

Using (B.7), the condition $u(L) \leq H$ is therefore trivially satisfied if $\lambda H \geq \varphi^*(1)$, otherwise it is equivalent to $\varphi'(0) + \lambda(L - x_0) \leq (\varphi^*)^{-1}(\lambda H)$, that is

$$x_0 \geq L - \frac{(\varphi^*)^{-1}(\lambda H) - \varphi'(0)}{\lambda},$$

where by $(\varphi^*)^{-1}$ we denote the inverse of the strictly increasing restriction $\varphi^*|_{[\varphi'(0), 1]}$. This latter inequality is correct in both cases if we interpret $(\varphi^*)^{-1}(\lambda H) = 1$ when $\lambda H \geq \varphi^*(1)$. Notice that, since $(\varphi^*)^{-1}(\lambda H) \leq 1$, this condition on x_0 is more restrictive than the previously found lower bound $x_0 \geq L - \frac{1 - \varphi'(0)}{\lambda}$.

Reduction to a one-dimensional minimization problem Finding the optimal profile u has been reduced to minimizing the function

$$f(x_0) := \frac{1}{\lambda} \int_{\varphi'(0)}^{\varphi'(0) + \lambda(L - x_0)} y(\varphi^*)'(y) dy + \theta \left(H - \frac{1}{\lambda} \varphi^*(\varphi'(0) + \lambda(L - x_0)) \right) \quad (\text{B.9})$$

under the constraint

$$L - \frac{(\varphi^*)^{-1}(\lambda H) - \varphi'(0)}{\lambda} \leq x_0 \leq L. \quad (\text{B.10})$$

If φ is strictly convex then φ^* is differentiable everywhere, therefore we can compute

$$\begin{aligned} \frac{dF(u)}{dx_0} &= f'(x_0) = -y(\varphi^*)'(y)|_{y=\varphi'(0) + \lambda(L - x_0)} + \theta' \left(H - \frac{1}{\lambda} \varphi^*(y) \right) (\varphi^*)'(y)|_{y=\varphi'(0) + \lambda(L - x_0)} \\ &= (\varphi^*)'(y) [\theta' (H - \varphi^*(y)/\lambda) - y]|_{y=\varphi'(0) + \lambda(L - x_0)}. \end{aligned}$$

If the minimizer x_0 is an interior point, it must satisfy $f'(x_0) = 0$, which is equivalent to

$$\theta' \left(H - \frac{1}{\lambda} \varphi^*(\varphi'(0) + \lambda(L - x_0)) \right) = \varphi'(0) + \lambda(L - x_0),$$

because

$$(\varphi^*)'(\varphi'(0) + \lambda(L - x_0)) \geq \frac{\varphi^*(\varphi'(0) + \lambda(L - x_0)) - \varphi^*(\varphi'(0))}{\lambda(L - x_0)} > 0$$

cannot vanish by the strict monotonicity of φ^* .

Acknowledgements The authors are members of GNAMPA, INdAM. This work was partially supported by the European Union - Next Generation EU and Research Projects PRIN2022 PNRR “Geometric-Analytic Methods for PDEs and Applications” (2022SLTHCE, cup E53D23005880006), awarded with D.D. 104 - 02/02/2022 of the Italian Ministry for University and Research. This manuscript reflects only the authors’ views and opinions and the Ministry cannot be considered responsible for them.

Conflict of interest The authors declare no conflict of interest.

Data availability statement We do not analyse or generate any datasets, because our work proceeds within a theoretical and mathematical approach.

References

- [ABS21] L. Ambrosio, E. Brué, D. Semola. *Lectures on Optimal Transport*. Vol. 130. UNITEXT. Cham: Springer International Publishing, 2021. ISBN: 978-3-030-72161-9, 978-3-030-72162-6. DOI: [10.1007/978-3-030-72162-6](https://doi.org/10.1007/978-3-030-72162-6). URL: <https://link.springer.com/10.1007/978-3-030-72162-6>.
- [AFP00] L. Ambrosio, N. Fusco, D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Science Publications. Clarendon Press, 2000. ISBN: 9780198502456.
- [Bae19] A. Baernstein II. *Symmetrization in Analysis*. New Mathematical Monographs. Cambridge University Press, 2019.
- [Bez+17] J. Bezanson, A. Edelman, S. Karpinski, V. B. Shah. “Julia: A Fresh Approach to Numerical Computing”. In: *SIAM Review* 59.1 (2017), pp. 65–98. DOI: [10.1137/141000671](https://doi.org/10.1137/141000671). URL: <https://epubs.siam.org/doi/10.1137/141000671>.
- [BV23] S. P. Boyd, L. Vandenberghe. *Convex Optimization*. Version 29. Cambridge New York Melbourne New Delhi Singapore: Cambridge University Press, 2023. 716 pp. ISBN: 978-0-521-83378-3. URL: https://web.stanford.edu/%7Eboyd/cvxbook/bv_cvxbook.pdf.
- [CCN15] V. Caselles, A. Chambolle, M. Novaga. “Total Variation in Imaging”. In: *Handbook of Mathematical Methods in Imaging*. Ed. by O. Scherzer. New York, NY: Springer New York, 2015, pp. 1455–1499. ISBN: 978-1-4939-0790-8. DOI: [10.1007/978-1-4939-0790-8_23](https://doi.org/10.1007/978-1-4939-0790-8_23). URL: https://doi.org/10.1007/978-1-4939-0790-8_23.
- [Cha04] A. Chambolle. “An algorithm for total variation minimization and applications”. In: *Journal of Mathematical Imaging and Vision* 20.1 (Jan. 2004), pp. 89–97. ISSN: 1573-7683. DOI: [10.1023/B:JMIV.0000011325.36760.1e](https://doi.org/10.1023/B:JMIV.0000011325.36760.1e). URL: <https://doi.org/10.1023/B:JMIV.0000011325.36760.1e>.
- [CL97] A. Chambolle, P.-L. Lions. “Image recovery via total variation minimization and related problems”. In: *Numerische Mathematik* 76.2 (Apr. 1997), pp. 167–188. ISSN: 0945-3245. DOI: [10.1007/s002110050258](https://doi.org/10.1007/s002110050258). URL: <https://doi.org/10.1007/s002110050258>.
- [MR18] P. K. Mogensen, A. N. Riseth. “Optim: A mathematical optimization package for Julia”. In: *Journal of Open Source Software* 3.24 (2018), p. 615. DOI: [10.21105/joss.00615](https://doi.org/10.21105/joss.00615). URL: <https://joss.theoj.org/papers/10.21105/joss.00615>.
- [Rak25] J. M. Rakotoson. *Relative Rearrangement*. Vol. 2376. Lecture Notes in Mathematics. Cham: Springer Nature Switzerland, 2025. ISBN: 978-3-032-02227-1, 978-3-032-02228-8. DOI: [10.1007/978-3-032-02228-8](https://doi.org/10.1007/978-3-032-02228-8). URL: <https://link.springer.com/10.1007/978-3-032-02228-8>.
- [ROF92] L. I. Rudin, S. Osher, E. Fatemi. “Nonlinear total variation based noise removal algorithms”. In: *Physica D: Nonlinear Phenomena* 60.1 (1992), pp. 259–268. ISSN: 0167-2789. DOI: [https://doi.org/10.1016/0167-2789\(92\)90242-F](https://doi.org/10.1016/0167-2789(92)90242-F). URL: <https://www.sciencedirect.com/science/article/pii/016727899290242F>.
- [Tal16] G. Talenti. “The Art of Rearranging”. In: *Milan Journal of Mathematics* 84.1 (June 2016), pp. 105–157. ISSN: 1424-9294. DOI: [10.1007/s00032-016-0253-6](https://doi.org/10.1007/s00032-016-0253-6). URL: <https://link.springer.com/article/10.1007/s00032-016-0253-6>.