

ONE DIMENSIONAL FRACTIONAL ORDER TGV : GAMMA-CONVERGENCE AND BILEVEL TRAINING SCHEME

ELISA DAVOLI

*Faculty of Mathematics, University of Vienna
Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria*

PAN LIU

*Cambridge Image Analysis, Department of Applied Mathematics and Theoretical Physics,
University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK*

ABSTRACT. New fractional r -order seminorms, TGV^r , $r \in \mathbb{R}$, $r \geq 1$, are proposed in the one-dimensional (1D) setting, as a generalization of the integer order TGV^k -seminorms, $k \in \mathbb{N}$. The fractional r -order TGV^r -seminorms are shown to be intermediate between the integer order TGV^k -seminorms. A bilevel training scheme is proposed, where under a box constraint a simultaneous optimization with respect to parameters and order of derivation is performed. Existence of solutions to the bilevel training scheme is proved by Γ -convergence. Finally, the numerical landscape of the cost function associated to the bilevel training scheme is discussed for two numerical examples.

1. INTRODUCTION

In the last decades, Calculus of Variations and Partial Differential Equations (PDE) methods have proven to be very efficient in signal (1D) and image (2D) denoising problems. Signal (image) denoising consists, roughly speaking, in recovering a noise-free clean signal u_c starting from a corrupted signal $u_\eta = u_c + \eta$, by filtering out the noise encoded by η . One of the most successful variational approach to signal (image) denoising (see, for example [35, 36, 37]) relies on the ROF total-variational functional

$$ROF(u) := \|u - u_\eta\|_{L^2(I)}^2 + \alpha TV(u), \quad (1.1)$$

introduced in [35]. Here $I = (0, 1)$ represents the domain of a one-dimensional image (a signal), $\alpha \in \mathbb{R}^+$, and $TV(u) = |u'|_{\mathcal{M}_b(I)}$, stands for the total mass of the measure u' on I (see [2, Definition 1.4]).

An important role in determining the reconstruction properties of the ROF functional is played by the parameter α . Indeed, if α is too large, then the total variation of u is too penalized and the image turns out to be over-smoothed, with a resulting loss of information

E-mail addresses: elisa.davoli@univie.ac.at, panliu.0923@maths.cam.ac.uk.

2010 Mathematics Subject Classification. 26B30, 94A08, 47J20.

Key words and phrases. total generalized variation, fractional derivatives, optimization and control, computer vision and pattern recognition.

on the internal edges of the picture. Conversely, if α is too small then the noise remains un-removed. The choice of the “best” parameter α then becomes an important task.

In [24] the authors proposed a training scheme (\mathcal{B}) relying on a bilevel learning optimization defined in machine learning, namely on a semi-supervised training scheme that optimally adapts itself to the given “perfect data” (see [13, 14, 26, 27, 39, 40]). This training scheme searches for the optimal α so that the recovered image u_α , obtained as a minimizer of (1.1), optimizes the L^2 -distance from the clean image u_c . An implementation of (\mathcal{B}) equipped with total variation is the following:

$$\begin{aligned} \text{Level 1.} \quad & \alpha_m \in \arg \min \left\{ \|u_\alpha - u_c\|_{L^2(I)}^2 : \alpha > 0 \right\} \\ \text{Level 2.} \quad & u_\alpha := \arg \min \left\{ \|u - u_\eta\|_{L^2(I)}^2 + \alpha TV(u) : u \in SBV(I) \right\}, \end{aligned} \quad (1.2)$$

where $SBV(I)$ denotes the set of special functions of bounded variation in I (see, e.g. [2, Chapter 4]).

It is well known that the ROF model in (1.1) suffers drawbacks like the staircasing effect, and the training scheme (\mathcal{B}) inherits that feature, namely the optimized reconstruction function u_{α_m} also exhibits the staircasing effect. One approach to counteract this problem is to insert higher-order derivatives in the regularizer (see [9, 11, 16, 33]). Two of the most successful image-reconstruction functionals among those involving mixed first and higher order terms are the infimal-convolution total variation ($ICTV$) [9] and the total generalized variation (TGV) models [33]. Note that they coincide with each other in the one-dimensional setting.

For $I := (0, 1) \subset \mathbb{R}$, $u \in BV(I)$, $k \in \mathbb{N}$, and $\alpha = (\alpha_0, \dots, \alpha_k) \in \mathbb{R}_+^{k+1}$, the TGV_α^k regularizer (see [43]) is defined as

$$|u|_{TGV_\alpha^1(I)} := \alpha_0 TV(u),$$

and

$$\begin{aligned} |u|_{TGV_\alpha^{k+1}(I)} := \inf \left\{ \alpha_0 |u' - v_0|_{\mathcal{M}_b(I)} + \alpha_1 |v_0' - v_1|_{\mathcal{M}_b(I)} + \right. \\ \left. \dots + \alpha_{k-1} |v_{k-2}' - v_{k-1}|_{\mathcal{M}_b(I)} + \alpha_k |v_{k-1}'|_{\mathcal{M}_b(I)} : \right. \\ \left. v_i \in BV(I) \text{ for } 0 \leq i \leq k-1 \right\}. \end{aligned}$$

For instance, for $k = 1$, the $TGV_{\alpha_0, \alpha_1}^2$ regularizer reads as

$$|u|_{TGV_{\alpha_0, \alpha_1}^2(I)} := \inf \left\{ \alpha_0 |u' - v_0|_{\mathcal{M}_b(I)} + \alpha_1 |v_0'|_{\mathcal{M}_b(I)}, v_0 \in BV(I) \right\}.$$

Substituting $TGV_{\alpha_0, \alpha_1}^2$ into (1.2) provides a bilevel training scheme with TGV image-reconstruction model. We recall that large values of α_1 will yield regularized solutions that are close to TV -regularized reconstructions, and large values of α_0 will result in TV^2 -type solutions (see, e.g., [34]). The best choice of parameters α_0 and α_1 is determined by an adaptation of the training scheme (\mathcal{B}) above (see [22] for a detailed study).

In the existing literature a regularizer is fixed a priori, and the biggest effort is concentrated on studying how to identify the best parameters. In the case of the TGV^k -model, this amounts to set manually the value of k first, and then determine the optimal α_m in (1.2). However, there is no evidence suggesting that TGV^2 will always perform better than TV . In addition, the higher order seminorms TGV^k , $k \geq 2$, have rarely been analyzed,

and hence their performance is largely unknown. Numerical simulations show that for different images (signals in 1D), different orders of TGV^k might give different results. The main focus of this paper is exactly to investigate how to optimally tune both the weight α and the order k of the TGV_α^k -seminorm, in order to achieve the best reconstructed image.

Our result is threefold. First, we develop a bilevel training scheme, not only for parameter training, but also for determining the optimal order k of the regularizer TGV^k for image reconstruction. A straightforward modification of (\mathcal{B}) would be to just insert the order of the regularizer inside the learning level 2 in (1.2). Namely,

$$\begin{aligned} \text{Level 1.} \quad & (\tilde{\alpha}, \tilde{k}) := \arg \min \left\{ \|u_{\alpha,k} - u_c\|_{L^2(I)}^2 : \alpha > 0, k \in \mathbb{N} \right\} \\ \text{Level 2.} \quad & u_{\alpha,k} := \arg \min \left\{ \|u - u_\eta\|_{L^2(I)}^2 + |u|_{TGV_\alpha^k(I)} : u \in SBV(I) \right\}. \end{aligned}$$

Often, in order to show the existence of a solution of the training scheme and also for the numerical realization of the model, a *box constraint*

$$(\alpha, k) \in [P, 1/P]^{k+1} \times [1, 1/P], \quad (1.3)$$

where $P \in (0, 1)$ small is a fixed real number, needs to be imposed (see, e.g. [3, 20]). However, such constraint makes the above training scheme less interesting. To be precise, restricting the analysis to the case in which $k \in \mathbb{N}$ is an integer, the box constraint (1.3) would only allow k to take finitely many values, and hence the optimal order \tilde{k} of the regularizer would simply be determined by performing scheme (\mathcal{B}) finitely many times, at each time with different values of k . In addition, finer texture effects, for which an ‘‘intermediate’’ reconstruction between the one provided by TGV^k and TGV^{k+1} for some $k \in \mathbb{N}$ would be needed, might be neglected in the optimization procedure.

Therefore, a main challenge in the setup of such a training scheme is to give a meaningful interpolation between the spaces TGV^k and TGV^{k+1} , and hence to guarantee that the collection of such spaces itself exhibits certain compactness and lower semicontinuity properties. To this purpose, we modify the definition of the TGV^k functionals by incorporating the theory of fractional Sobolev spaces, and we introduce the notion of fractional order TGV^{k+s} spaces (see Definition 3.1), where $k \in \mathbb{N}$, and $0 < s < 1$. For $k = 1$, our definition reads as follows.

$$\begin{aligned} |u|_{TGV_\alpha^{1+s}(I)} := \inf \left\{ \alpha_0 |u' - sv_0|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) |v_0|_{W^{s,1+s(1-s)}(I)} \right. \\ \left. + \alpha_0 s(1-s) \left| \int_I v_0(x) dx \right| : v_0 \in W^{s,1+s(1-s)}(I) \right\}. \end{aligned}$$

In the expression above, $W^{s,1+s(1-s)}(I)$ is the *fractional Sobolev space* of order s and integrability $1 + s(1 - s)$. For every $k \in \mathbb{N}$ and $s \in [0, 1]$ we additionally introduce the sets

$$BGV_\alpha^{k+s}(I) := \left\{ u \in L^1(I) : |u|_{TGV_\alpha^{k+s}(I)} < +\infty \right\},$$

namely the classes of functions with bounded generalized total-variation seminorm.

In our first main result (see Theorem 3.2) we show that the TGV^{1+s} seminorm is indeed intermediate between TGV^1 and TGV^2 , i.e., we prove that, up to subsequences,

$$\lim_{s \nearrow 1} |u|_{TGV_\alpha^{1+s}(I)} = |u|_{TGV_\alpha^2(I)} \quad \text{and} \quad \lim_{s \searrow 0} |u|_{TGV_\alpha^{1+s}(I)} = \alpha_0 |u'|_{\mathcal{M}_b(I)}. \quad (1.4)$$

Equation (1.4) shows that, for $s \nearrow 1$, the behavior of the TGV^{1+s} -seminorm is close to the one of the standard TGV^2 -seminorm, whereas for $s \searrow 0$ it approaches the TV functional. We additionally prove (see Corollary 3.5) that analogous results hold for higher order TGV^{k+s} -seminorms. We point out that working with such interpolation spaces has many advantages. Indeed, TGV^{k+s} is expected to inherit the properties of fractional order derivatives, which have shown to be able to reduce the staircasing and contrast effects in noise-removal problems (see, e.g. [12]).

Our second and third main results (Theorems 4.2 and 5.2) concern the following improved training scheme (\mathcal{R}), which, under the box constraint (1.3), simultaneously optimizes both the parameter α and the order r of derivation:

$$\begin{aligned} \text{Level 1.} \quad & (\tilde{\alpha}, \tilde{r}) := \arg \min \left\{ \|u_{\alpha,r} - u_c\|_{L^2(I)}^2, (\alpha, r) \in [P, 1/P]^{\lfloor r \rfloor + 1} \times [1, 1/P] \right\}, \quad (1.5) \\ \text{Level 2.} \quad & u_{\alpha,r} := \arg \min \left\{ \|u - u_\eta\|_{L^2(I)}^2 + TGV_\alpha^r(u) : u \in BGV_\alpha^r(I) \right\}. \end{aligned}$$

In the definition above, $\lfloor r \rfloor$ denotes the largest integer strictly smaller than or equal to r . We first show in Theorem 4.2 that the fractional order TGV_α^r functionals

$$\mathcal{F}_\alpha^r(u) := \|u - u_\eta\|_{L^2(I)}^2 + TGV_\alpha^r(u) \quad \text{for every } u \in BGV_\alpha^r(I)$$

are continuous, in the sense of Γ -convergence in the weak* topology of $BV(I)$ (see [7] and [15]), with respect to the parameters α and the order r . Secondly, in Theorem 5.2 we exploit this Γ -convergence result, to prove existence of solutions to our training scheme (\mathcal{R}). Note that, according to the given noisy image u_η and noise-free image u_c , the Level 1 in our training scheme (\mathcal{R}) provides simultaneously an optimal regularizer $TGV^{\tilde{r}}$ and a corresponding optimal parameter $\tilde{\alpha} \in [P, 1/P]^{\lfloor \tilde{r} \rfloor + 1}$. We point out that, in general, the optimal order of derivation \tilde{r} might, or might not, be an integer. In other words, the fractional TGV^r seminorms are not intended as an improvement but rather as an extension of the integer order TGV^k seminorms, which for some classes of signals might provide optimal reconstruction and be selected by the bilevel training scheme.

Although this paper mainly focuses on a theoretical analysis of TGV^r and on showing the existence of optimal results for the training scheme (\mathcal{R}), in Section 6 some preliminary numerical examples are discussed (see Figures 1–2). We stress that a complete description of the optimality conditions and a reliable numerical scheme for identifying the optimal solution of the training scheme (1.5) are beyond the scope of this work, and are still a challenging open problem. We refer to [21, 23] for some preliminary results in this direction. The two-dimensional setting of fractional order TGV^r and $ICTV^r$ seminorms, as well as more extensive numerical analysis and examples for different type of images (with large flat areas, fine details, etc.), will be the subject of the follow-up work [18].

Our paper is organized as follows. In Section 2 we review the definitions and some basic properties of fractional Sobolev spaces. In Section 3 we introduce the fractional order TGV^r seminorms, we study their main properties, and prove that they are intermediate between integer-order seminorms (see Theorem 3.2). In Section 4 we characterize the asymptotic behavior of the functionals \mathcal{F}_α^r with respect to parameters and order of derivations (see Theorem 4.2). In Section 5 we introduce our training scheme (\mathcal{R}). In particular, in Theorem 5.2 we show that (\mathcal{R}) admits a solution under the box constraint (1.3). Lastly, in Section 6 some examples and insights are provided.

2. THE THEORY OF FRACTIONAL SOBOLEV SPACES

In what follows we will assume that $I = (0, 1)$. We first recall a few results from the theory of fractional Sobolev spaces. We refer to [25] for an introduction to the main results, and to [1, 29, 30, 32] and the references therein for a comprehensive treatment of the topic.

Definition 2.1 (Fractional Sobolev spaces). For $0 < s < 1$, $1 \leq p < +\infty$, and $u \in L^p(I)$, we define the *Gagliardo seminorm* of u by

$$|u|_{W^{s,p}(I)} := \left(\int_I \int_I \frac{|u(x) - u(y)|^p}{|x - y|^{1+sp}} dx dy \right)^{\frac{1}{p}}. \quad (2.1)$$

We say that $u \in W^{s,p}(I)$ if

$$\|u\|_{W^{s,p}(I)} := \|u\|_{L^p(I)} + |u|_{W^{s,p}(I)} < +\infty.$$

The following embedding results hold true ([25, Theorems 6.7, 6.10, and 8.2, and Corollary 7.2]).

Theorem 2.2 (Sobolev Embeddings - 1). *Let $s \in (0, 1)$ be given.*

1. *Let $p < \frac{1}{s}$. Then there exists a positive constant $C = C(p, s)$ such that for every $u \in W^{s,p}(I)$ there holds*

$$\|u\|_{L^q(I)} \leq C \|u\|_{W^{s,p}(I)} \quad (2.2)$$

for every $q \in [1, \frac{p}{1-sp}]$. If $q < \frac{p}{1-sp}$, then the embedding of $W^{s,p}(I)$ into $L^q(I)$ is also compact.

2. *Let $p = \frac{1}{s}$. Then the embedding in (2.2) holds for every $q \in [1, +\infty)$.*
3. *Let $p > \frac{1}{s}$. Then there exists a positive constant $C = C(p, s)$ such that for every $u \in W^{s,p}(I)$ we have*

$$\|u\|_{C^{0,\alpha}(I)} \leq C \|u\|_{W^{s,p}(I)},$$

with $\alpha := \frac{sp-1}{p}$, where

$$\|u\|_{C^{0,\alpha}(I)} := \|u\|_{L^\infty(I)} + \sup_{x \neq y \in I} \frac{|u(x) - u(y)|}{|x - y|^\alpha}.$$

The additional embedding result below is proved in [38, Corollary 19].

Theorem 2.3 (Sobolev Embeddings - 2). *Let $s \geq r$, $p \leq q$ and $s - 1/p \geq r - 1/q$, with $0 < r \leq s < 1$, and $1 \leq p \leq q \leq +\infty$. Then*

$$W^{s,p}(I) \subset W^{r,q}(I),$$

and

$$|u|_{W^{r,q}(I)} \leq \frac{36}{rs} |u|_{W^{s,p}(I)}.$$

The next inequality is a special case of [4, Theorem 1] and [31, Theorem 1].

Theorem 2.4 (Poincaré Inequality). *Let $p \geq 1$, and let $sp < 1$. There exists a constant $C > 0$ such that*

$$\left\| u - \int_I u(x) dx \right\|_{L^{\frac{p}{1-sp}}(I)}^p \leq \frac{Cs(1-s)}{(1-sp)^{p-1}} |u|_{W^{s,p}(I)}^p.$$

It is possible to construct a continuous extension operator from $W^{s,1}(I)$ to $W^{s,1}(\mathbb{R})$ (see, e.g., [25, Theorem 5.4]).

Theorem 2.5 (Extension Operator). *Let $s \in (0, 1)$, and let $1 \leq p < +\infty$. Then $W^{s,p}(I)$ is continuously embedded in $W^{s,p}(\mathbb{R})$, namely there exists a constant $C = C(p, s)$ such that for every $u \in W^{s,p}(I)$ there exists $\tilde{u} \in W^{s,p}(\mathbb{R})$ satisfying $\tilde{u}|_I = u$ and*

$$\|\tilde{u}\|_{W^{s,p}(\mathbb{R})} \leq C \|u\|_{W^{s,p}(I)}.$$

The next two theorems ([41, Section 2.2.2, Remark 3, and Section 2.11.2]) yield an identification between fractional Sobolev spaces and Besov spaces in \mathbb{R} , and guarantee the reflexivity of Besov spaces $B_{p,q}^s$ for p, q finite.

Theorem 2.6 (Identification with Besov spaces). *If $1 \leq p < +\infty$ and $s \in \mathbb{R}^+ \setminus \mathbb{N}$, then*

$$W^{s,p}(\mathbb{R}) = B_{p,p}^s(\mathbb{R})$$

Theorem 2.7 (Reflexivity of Besov spaces). *Let $-\infty < s < +\infty$, $1 \leq p < +\infty$ and $0 < q < +\infty$. Then*

$$(B_{p,q}^s(\mathbb{R}))' = B_{p',q'}^{-s}(\mathbb{R}),$$

where $(B_{p,q}^s(\mathbb{R}))'$ is the dual of the Besov space $B_{p,q}^s(\mathbb{R})$, and where p' and q' are the conjugate exponent of p and q , respectively.

In view of Theorems 2.6 and 2.7 the following characterization holds true.

Corollary 2.8 (Reflexivity of fractional Sobolev spaces). *Let $1 < p < +\infty$ and $s \in \mathbb{R}^+ \setminus \mathbb{N}$. Then the fractional Sobolev space $W^{s,p}(\mathbb{R})$ is reflexive.*

We conclude this section by recalling two theorems describing the limit behavior of the Gagliardo seminorm as $s \nearrow 1$ and $s \searrow 0$, respectively. The first result has been proved in [5, Theorem 3 and Remark 1], and [17, Theorem 1].

Theorem 2.9 (Asymptotic behavior as $s \nearrow 1$). *Let $u \in BV(I)$. Then*

$$\lim_{s \nearrow 1} (1-s) |u|_{W^{s,1}(I)} = |u'|_{\mathcal{M}_b(I)}.$$

Similarly, the asymptotic behavior of the Gagliardo seminorm has been characterized as $s \searrow 0$ in [31, Theorem 3].

Theorem 2.10 (Asymptotic behavior as $s \searrow 0$). *Let $u \in \cup_{0 < s < 1} W^{s,1}(\mathbb{R})$. Then,*

$$\lim_{s \searrow 0} s |u|_{W^{s,1}(\mathbb{R})} = 4 \|u\|_{L^1(\mathbb{R})}.$$

3. THE FRACTIONAL ORDER TGV SEMINORMS

Let $r \in (1, +\infty) \setminus \mathbb{N}$ be given. In this section we define the *fractional r -order total generalized variation* (TGV^r) seminorms, and we prove some first properties.

Definition 3.1 (The TGV^r seminorms). *Let $0 < s < 1$, $k \in \mathbb{N}$ be such that $k+s = r$, and let $\alpha = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_k) \in \mathbb{R}_+^{k+1}$. For every $u \in L^1(I)$, we define its fractional TGV^{k+s} seminorm as follows.*

Case 1. for $k = 1$

$$|u|_{TGV_\alpha^{1+s}(I)} := \inf \left\{ \alpha_0 |u' - sv_0|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) |v_0|_{W^{s,1+s(1-s)}(I)} \right. \\ \left. + \alpha_0 s(1-s) \left| \int_I v_0(x) dx \right| : v_0 \in W^{s,1+s(1-s)}(I) \right\}.$$

Case 2. for $k > 1$

$$\begin{aligned} |u|_{TGV_\alpha^{k+s}(I)} := \inf \left\{ \alpha_0 |u' - v_0|_{\mathcal{M}_b(I)} + \alpha_1 |v'_0 - v_1|_{\mathcal{M}_b(I)} + \right. \\ \left. \cdots + \alpha_{k-1} |v'_{k-2} - sv_{k-1}|_{\mathcal{M}_b(I)} + \alpha_k s(1-s) |v_{k-1}|_{W^{s,1+s(1-s)}(I)} \right. \\ \left. + \alpha_{k-1} s(1-s) \left| \int_I v_{k-1}(x) dx \right| : \right. \\ \left. v_i \in BV(I) \text{ for } 0 \leq i \leq k-2, v_{k-1} \in W^{s,1+s(1-s)}(I) \right\} \end{aligned}$$

Moreover, we say that u belongs to the space of *functions with bounded total generalized variation*, and we write $u \in BGV_\alpha^{k+s}(I)$ if

$$\|u\|_{BGV_\alpha^{k+s}(I)} := \|u\|_{L^1(I)} + |u|_{TGV_\alpha^{k+s}(I)} < +\infty,$$

where $0 \leq s < 1$, $k \in \mathbb{N}$, $\alpha = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_k) \in \mathbb{R}_+^{k+1}$. Additionally, we write $u \in BGV^{k+s}(I)$ if there exists $\alpha \in \mathbb{R}_+^{k+1}$ such that $u \in BGV_\alpha^{k+s}(I)$. Note that if $u \in BGV_\alpha^{k+s}(I)$ for some $\alpha \in \mathbb{R}_+^{k+1}$, then $u \in BGV_\beta^{k+s}(I)$ for every $\beta \in \mathbb{R}_+^{k+1}$.

We observe that the TGV^{k+s} seminorm is actually “intermediate” between the TGV^k seminorm and the TGV^{k+1} seminorm. To be precise, we have the following identification.

Theorem 3.2 (Asymptotic behavior of the fractional TGV seminorm-1). *For every $u \in BV(I)$, up to the extraction of a (non-relabelled) subsequence there holds*

$$\lim_{s \nearrow 1} |u|_{TGV_\alpha^{1+s}(I)} = |u|_{TGV_\alpha^2(I)} \quad \text{and} \quad \lim_{s \searrow 0} |u|_{TGV_\alpha^{1+s}(I)} = \alpha_0 |u'|_{\mathcal{M}_b(I)}.$$

Before proving Theorem 3.2 we state and prove an intermediate result that will be crucial in determining the asymptotic behavior of the TGV^{1+s} seminorm as $s \nearrow 1$.

Proposition 3.3. *Let $u \in W^{1,\infty}(I)$. Then*

$$\limsup_{s \nearrow 1} (1-s) |u|_{W^{s,1+s(1-s)}(I)} \leq |u'|_{\mathcal{M}_b(I)}.$$

Proof. Let $u \in W^{1,\infty}(I)$. Then there exists a constant $L > 0$ such that

$$|u(x) - u(y)| \leq L|x - y|^s$$

for every $x, y \in I$ and every $s \in (0, 1)$. Thus

$$\begin{aligned} |u|_{W^{s,1+s(1-s)}(I)}^{1+s(1-s)} &= \int_I \int_I \frac{|u(x) - u(y)|^{1+s(1-s)}}{|x - y|^{1+s(1+s(1-s))}} dx dy \\ &\leq L^{s(1-s)} \int_I \int_I \frac{|x - y|^{s^2(1-s)} |u(x) - u(y)|}{|x - y|^{1+s(1+s(1-s))}} dx dy \\ &= L^{s(1-s)} \int_I \int_I \frac{|u(x) - u(y)|}{|x - y|^{1+s}} dx dy = L^{s(1-s)} |u|_{W^{s,1}(I)}. \end{aligned}$$

This implies that

$$\begin{aligned} (1-s) |u|_{W^{s,1+s(1-s)}(I)} &\leq (1-s) L^{\frac{s(1-s)}{1+s(1-s)}} |u|_{W^{s,1}(I)}^{\frac{1}{1+s(1-s)}} \\ &= L^{\frac{s(1-s)}{1+s(1-s)}} \left[(1-s) |u|_{W^{s,1}(I)} \right]^{\frac{1}{1+s(1-s)}} e^{\frac{s(1-s) \log(1-s)}{1+s(1-s)}}. \end{aligned}$$

Therefore, by Theorem 2.9 we conclude that

$$\limsup_{s \nearrow 1} (1-s) |u|_{W^{s,1+s(1-s)}(I)} \leq \limsup_{s \nearrow 1} \left[(1-s) |u|_{W^{s,1}(I)} \right]^{\frac{1}{1+s(1-s)}} \leq |u'|_{\mathcal{M}_b(I)}. \quad \square$$

A crucial ingredient in the proof of Theorem 3.2 is a compactness and lower-semicontinuity result for maps with uniformly weighted averages and $W^{s,1+s(1-s)}$ -seminorms.

Proposition 3.4. *Let $\{s_n\} \subset (0, 1)$ be such that $s_n \rightarrow \bar{s}$, with $\bar{s} \in (0, 1]$. For every $n \in \mathbb{N}$ let $v_n \in W^{s_n,1+s_n(1-s_n)}(I)$ be such that*

$$\sup_{n \geq 1} s_n(1-s_n) \left\{ |v_n|_{W^{s_n,1+s_n(1-s_n)}(I)} + \left| \int_I v_n(x) dx \right| \right\} < +\infty. \quad (3.1)$$

Then, for $\bar{s} \in (0, 1)$, there exists $\bar{v} \in W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)$ such that, up to the extraction of a (non-relabeled) subsequence,

$$v_n \rightarrow \bar{v} \quad \text{strongly in } L^1(I), \quad (3.2)$$

and

$$\liminf_{n \rightarrow \infty} s_n(1-s_n) |v_n|_{W^{s_n,1+s_n(1-s_n)}(I)} \geq \bar{s}(1-\bar{s}) |\bar{v}|_{W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)}. \quad (3.3)$$

For $\bar{s} = 1$, there exists $\bar{v} \in BV(I)$ such that, up to the extraction of a (non-relabeled) subsequence,

$$v_n - \int_I v_n(x) dx \rightarrow \bar{v} \quad \text{strongly in } L^1(I), \quad (3.4)$$

and

$$\liminf_{n \rightarrow \infty} s_n(1-s_n) |v_n|_{W^{s_n,1+s_n(1-s_n)}(I)} \geq |\bar{v}'|_{\mathcal{M}_b(I)}. \quad (3.5)$$

Proof. We first observe that for $x, y \in I$, $1 \leq p < +\infty$, and $s < t$, we have

$$|x - y|^{1+sp} > |x - y|^{1+tp}.$$

Hence, in view of (2.1) there holds

$$|u|_{W^{s,p}(I)} \leq |u|_{W^{t,p}(I)}$$

for every $u \in W^{t,p}(I)$.

Without loss of generality (and up to the extraction of a non-relabeled subsequence) we can assume that the sequences $\{s_n\}$ and $\{s_n(1-s_n)\}$ converge monotonically to \bar{s} and $\bar{s}(1-\bar{s})$, respectively. According to the value of \bar{s} only 4 situations can arise:

Case 1: $\frac{1}{2} \leq \bar{s} < 1$: $s_n \searrow \bar{s}$ and $s_n(1-s_n) \nearrow \bar{s}(1-\bar{s})$;

Case 2: $0 < \bar{s} < \frac{1}{2}$: $s_n \searrow \bar{s}$ and $s_n(1-s_n) \searrow \bar{s}(1-\bar{s})$;

Case 3: $\frac{1}{2} < \bar{s} \leq 1$: $s_n \nearrow \bar{s}$ and $s_n(1-s_n) \searrow \bar{s}(1-\bar{s})$;

Case 4: $0 < \bar{s} \leq \frac{1}{2}$: $s_n \nearrow \bar{s}$ and $s_n(1-s_n) \nearrow \bar{s}(1-\bar{s})$.

For convenience of the reader we subdivide the proof into three steps.

Step 1: We first consider Case 1. By (3.1) there exists a constant C such that

$$\sup_{n \geq 1} \left\{ |v_n|_{W^{s_n,1+s_n(1-s_n)}(I)} + \left| \int_I v_n(x) dx \right| \right\} \leq C. \quad (3.6)$$

We point out that the function $f : (0, 1) \rightarrow \mathbb{R}$, defined as

$$f(x) := x - \frac{1}{1+x(1-x)} \quad \text{for every } x \in [0, 1], \quad (3.7)$$

is strictly increasing on $[0, 1]$. In particular, since $s_n \geq \bar{s}$, there holds $f(s_n) \geq f(\bar{s})$, namely

$$s_n - \frac{1}{1+s_n(1-s_n)} \geq \bar{s} - \frac{1}{1+\bar{s}(1-\bar{s})}.$$

By applying Theorem 2.3 with $s = s_n$, $r = \bar{s}$, $p = 1 + s_n(1 - s_n)$, and $q = 1 + \bar{s}(1 - \bar{s})$, we obtain that there exists a constant C such that

$$|v_n|_{W^{\bar{s}, 1 + \bar{s}(1 - \bar{s})}(I)} \leq C |v_n|_{W^{s_n, 1 + s_n(1 - s_n)}(I)} \quad (3.8)$$

for every $n \in \mathbb{N}$. The uniform bound (3.6) yields then that there exists a constant C such that

$$\sup_{n \geq 1} |v_n|_{W^{\bar{s}, 1 + \bar{s}(1 - \bar{s})}(I)} \leq C. \quad (3.9)$$

In view of Theorem 2.4, Corollary 2.8, and estimates (3.6) and (3.9) there exists $\bar{v} \in W^{\bar{s}, 1 + \bar{s}(1 - \bar{s})}(I)$ such that, up to the extraction of a (non-reabeled) subsequence, we have

$$v_n \rightharpoonup \bar{v} \quad \text{weakly in } W^{\bar{s}, 1 + \bar{s}(1 - \bar{s})}(I). \quad (3.10)$$

Since $\bar{s}(1 + \bar{s}(1 - \bar{s})) < 1$, and $1 < \frac{1 + \bar{s}(1 - \bar{s})}{1 - \bar{s}(1 + \bar{s}(1 - \bar{s}))}$, by Theorem 2.2 (1.), the embedding of $W^{\bar{s}, 1 + \bar{s}(1 - \bar{s})}(I)$ into $L^1(I)$ is compact. Property (3.2) follows then by (3.10).

By the lower semicontinuity of the $W^{\bar{s}, 1 + \bar{s}(1 - \bar{s})}(I)$ norm with respect to the weak convergence, and by (3.8) we deduce the inequality

$$\begin{aligned} \bar{s}(1 - \bar{s}) |\bar{v}|_{W^{\bar{s}, 1 + \bar{s}(1 - \bar{s})}(I)} &\leq \liminf_{n \rightarrow +\infty} \bar{s}(1 - \bar{s}) |v_n|_{W^{s_n, 1 + s_n(1 - s_n)}(I)} \\ &= \liminf_{n \rightarrow +\infty} s_n(1 - s_n) |v_n|_{W^{s_n, 1 + s_n(1 - s_n)}(I)}, \end{aligned}$$

which in turn yields (3.3).

Step 2: Consider now Case 2. The function $g : (0, 1) \rightarrow \mathbb{R}$, defined as

$$g(x) := \frac{1}{1 + x(1 - x)} \quad \text{for every } x \in (0, 1),$$

is strictly decreasing in $(0, \frac{1}{2}]$. By the definition of the maps f (defined in (3.7)) and g there holds

$$g(\bar{s}) + f(\bar{s}) = \bar{s}.$$

Therefore, by the monotonicity of g in $(0, \frac{1}{2}]$ there exist $\frac{1}{2} > \hat{s} > \bar{s}$, and $\lambda \in (0, 1)$, such that $g(\hat{s}) + f(\bar{s}) > \lambda \bar{s}$, namely

$$\bar{s} - \frac{1}{1 + \bar{s}(1 - \bar{s})} > \lambda \bar{s} - \frac{1}{1 + \hat{s}(1 - \hat{s})}.$$

By the monotonicity of f on $[0, 1]$, and the fact that $s_n \geq \bar{s}$ for every $n \in \mathbb{N}$, we have

$$s_n - \frac{1}{1 + s_n(1 - s_n)} = f(s_n) > f(\bar{s}) = \bar{s} - \frac{1}{1 + \bar{s}(1 - \bar{s})} > \lambda \bar{s} - \frac{1}{1 + \hat{s}(1 - \hat{s})}.$$

By (3.1) there exists a constant C such that

$$\sup_{n \geq 1} \left\{ |v_n|_{W^{s_n, 1 + s_n(1 - s_n)}(I)} + \left| \int_I v_n(x) dx \right| \right\} \leq C. \quad (3.11)$$

Since $\frac{1}{2} > \hat{s} > \bar{s}$, and $s_n(1 - s_n) \searrow \bar{s}(1 - \bar{s})$, there exists $n_0 \in \mathbb{N}$ such that

$$1 + s_n(1 - s_n) < 1 + \hat{s}(1 - \hat{s}) \quad \text{for every } n \geq n_0.$$

Hence, choosing $s = s_n$, $r = \lambda \bar{s}$, $p = 1 + s_n(1 - s_n)$, and $q = 1 + \hat{s}(1 - \hat{s})$ in Theorem 2.3, we deduce that there exists a constant C such that

$$|v_n|_{W^{\lambda \bar{s}, 1 + \hat{s}(1 - \hat{s})}(I)} \leq C |v_n|_{W^{s_n, 1 + s_n(1 - s_n)}(I)}$$

for every $n \geq n_0$. In particular, (3.1) yields the uniform bound

$$\sup_{n \geq 1} |v_n|_{W^{\lambda\bar{s}, 1+\hat{s}(1-\hat{s})}(I)} \leq C.$$

In view of Theorem 2.4, Corollary 2.8, and estimate (3.11) we deduce the existence of a map \bar{v} such that, up to the extraction of a (non-relabelled) subsequence,

$$v_n \rightharpoonup \bar{v} \quad \text{weakly in } W^{\lambda\bar{s}, 1+\hat{s}(1-\hat{s})}(I). \quad (3.12)$$

Since $\lambda\bar{s}(1 + \hat{s}(1 - \hat{s})) < \hat{s}(1 + \hat{s}(1 - \hat{s})) < 1$, and $1 < \frac{1+\hat{s}(1-\hat{s})}{1-\lambda\bar{s}(1+\hat{s}(1-\hat{s}))}$, by Theorem 2.2 (1) the space $W^{\lambda\bar{s}, 1+\hat{s}(1-\hat{s})}(I)$ embeds compactly into $L^1(I)$. Hence, the convergence in (3.12) holds also strongly in $L^1(I)$, and (3.2) follows. In particular, Fatou's Lemma yields

$$|v|_{W^{\bar{s}, 1+\bar{s}(1-\bar{s})}(I)}^{1+\bar{s}(1-\bar{s})} \leq \liminf_{n_k \rightarrow +\infty} |v_{n_k}|_{W^{s_{n_k}, 1+s_{n_k}(1-s_{n_k})}(I)}^{1+s_{n_k}(1-s_{n_k})},$$

which in turn implies (3.3).

Step 3: We omit the proof of the result in Case 4, and in Case 3 for $\bar{s} < 1$, as they follow from analogous arguments. Regarding Case 3 for $\bar{s} = 1$, by Hölder inequality we have

$$\begin{aligned} & \int_I \int_I \frac{|v_n(x) - v_n(y)|}{|x - y|^{1+\frac{s_n}{2-s_n}}} dx dy \\ & \leq \left[\int_I \int_I \left(\frac{|v_n(x) - v_n(y)|}{|x - y|^{1+\frac{s_n}{2-s_n}}} \right)^{1+s_n(1-s_n)} dx dy \right]^{\frac{1}{1+s_n(1-s_n)}} \\ & = \left[\int_I \int_I \frac{|v_n(x) - v_n(y)|^{1+s_n(1-s_n)}}{|x - y|^{1+\frac{s_n}{2-s_n}+s_n(1-s_n)+\frac{s_n^2}{2-s_n}(1-s_n)}} dx dy \right]^{\frac{1}{1+s_n(1-s_n)}}. \end{aligned} \quad (3.13)$$

Now,

$$1 + \frac{s_n}{2-s_n} + s_n(1-s_n) + \frac{s_n^2}{2-s_n}(1-s_n) < 1 + s_n + s_n^2(1-s_n)$$

for n big enough (because $s_n \nearrow 1$). Thus

$$\frac{1}{|x - y|^{1+\frac{s_n}{2-s_n}+s_n(1-s_n)+\frac{s_n^2}{2-s_n}(1-s_n)}} < \frac{1}{|x - y|^{1+s_n+s_n^2(1-s_n)}}$$

for every $x, y \in I$, $x \neq y$, and by (3.13) we obtain

$$|v_n|_{W^{\frac{s_n}{2-s_n}, 1}(I)} \leq |v_n|_{W^{s_n, 1+s_n(1-s_n)}(I)}$$

for every $n \in \mathbb{N}$. Property (3.1) yields the existence of a constant C such that

$$\sup_{n \in \mathbb{N}} (1 - s_n) \left(|v_n|_{W^{\frac{s_n}{2-s_n}, 1}(I)} + \left| \int_I v_n(x) dx \right| \right) \leq C. \quad (3.14)$$

Setting $t_n := \frac{s_n}{2-s_n}$, there holds $t_n \rightarrow 1$ as $n \rightarrow +\infty$, and (3.14) implies

$$\sup_{n \in \mathbb{N}} (1 - t_n) \left\{ |v_n|_{W^{t_n, 1}(I)} + \left| \int_I v_n(x) dx \right| \right\} \leq C.$$

Properties (3.4) and (3.5) are then a consequence of [5, Theorem 4]. \square

We now prove Theorem 3.2.

Proof of Theorem 3.2. Fix $\varepsilon > 0$. Let $v_0 \in BV(I)$ be such that

$$|u|_{TGV_\alpha^2(I)} \geq \alpha_0 |u' - v_0|_{\mathcal{M}_b(I)} + \alpha_1 |v_0'|_{\mathcal{M}_b(I)} - \varepsilon.$$

Let $v_0^k \in W^{1,\infty}(I)$ satisfy

$$|(v_0^k)'|_{\mathcal{M}_b(I)} - |v_0'|_{\mathcal{M}_b(I)}| < \varepsilon,$$

and

$$\|v_0 - v_0^k\|_{L^1(I)} < \varepsilon.$$

In view of Proposition 3.3 there holds

$$\begin{aligned} \limsup_{s \nearrow 1} |u|_{TGV_\alpha^{1+s}(I)} &\leq \limsup_{s \nearrow 1} \left\{ \alpha_0 |u' - sv_0^k|_{\mathcal{M}_b(I)} + \alpha_0 s(1-s) \left| \int_I v_0^k(x) dx \right| \right. \\ &\quad \left. + \alpha_1 s(1-s) |v_0^k|_{W^{s,1+s(1-s)}(I)} \right\} \\ &\leq \alpha_0 |u' - v_0^k|_{\mathcal{M}_b(I)} + \alpha_1 |(v_0^k)'|_{\mathcal{M}_b(I)} \\ &\leq \alpha_0 |u' - v_0|_{\mathcal{M}_b(I)} + \alpha_1 |v_0'|_{\mathcal{M}_b(I)} + (\alpha_0 + \alpha_1)\varepsilon \\ &\leq |u|_{TGV_\alpha^2(I)} + (\alpha_0 + \alpha_1 + 1)\varepsilon. \end{aligned}$$

The arbitrariness of ε yields

$$\limsup_{s \nearrow 1} |u|_{TGV_\alpha^{1+s}(I)} \leq |u|_{TGV_\alpha^2(I)}. \quad (3.15)$$

To prove the opposite inequality, for every $s \in (0, 1)$ let $v_0^s \in W^{s,1+s(1-s)}(I)$ be such that

$$\begin{aligned} \alpha_0 |u' - sv_0^s|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) |v_0^s|_{W^{s,1+s(1-s)}(I)} + \alpha_0 s(1-s) \left| \int_I v_0^s(x) dx \right| \\ \leq |u|_{TGV_\alpha^{1+s}(I)} + s. \end{aligned} \quad (3.16)$$

In view of (3.15) and Proposition 3.4, there exists $\tilde{v} \in BV(I)$ such that, up to the extraction of a (non-relabelled) subsequence,

$$v_0^s - \int_I v_0^s(x) dx \rightarrow \tilde{v} \quad \text{strongly in } L^1(I), \quad (3.17)$$

as $s \rightarrow 1$, and

$$\lim_{s \nearrow 1} s(1-s) |v_0^s|_{W^{s,1+s(1-s)}(I)} \geq |\tilde{v}'|_{\mathcal{M}_b(I)}. \quad (3.18)$$

Additionally, by (3.15) and (3.16) there holds

$$s \|v_0^s\|_{L^1(I)} \leq |u'|_{\mathcal{M}_b(I)} + \|u' - sv_0^s\|_{\mathcal{M}_b(I)} \leq |u'|_{\mathcal{M}_b(I)} + |u|_{TGV_\alpha^{1+s}(I)} + s \leq C$$

for every $s \in (0, 1)$. Thus, there exists a constant C such that

$$\lim_{s \rightarrow 1} \int_I v_0^s(x) dx = C.$$

In particular, setting $v := \tilde{v} + C$, by (3.17) and (3.18) there holds

$$v_0^s \rightarrow v \quad \text{strongly in } L^1(I),$$

and

$$\lim_{s \rightarrow 1} s(1-s) |v_0^s|_{W^{s,1+s(1-s)}(I)} \geq |v'|_{\mathcal{M}_b(I)}.$$

Passing to the limit in (3.16) we deduce the inequality

$$|u|_{TGV_\alpha^2(I)} \leq \alpha_0 |u' - v|_{\mathcal{M}_b(I)} + \alpha_1 |v'|_{\mathcal{M}_b(I)} \leq \liminf_{s \nearrow 1} |u|_{TGV_\alpha^{1+s}(I)},$$

which in turn implies the thesis.

To study the case $s \searrow 0$, we first observe that

$$\sup_{s \in (0,1)} |u|_{TGV_\alpha^{1+s}(I)} \leq \alpha_0 |u'|_{\mathcal{M}_b(I)}. \quad (3.19)$$

Thus we only need to prove the opposite inequality. To this aim, for every $s \in (0,1)$ let $v_0^s \in W^{s,1+s(1-s)}(I)$ be such that

$$\begin{aligned} & \alpha_0 |u' - sv_0^s|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) |v_0^s|_{W^{s,1+s(1-s)}(I)} + \alpha_0 s(1-s) \left| \int_I v_0^s(x) dx \right| \\ & \leq |u|_{TGV_\alpha^{1+s}(I)} + s. \end{aligned} \quad (3.20)$$

Since $s(1+s(1-s)) < 1$ for $s \in (0,1)$, by (3.19) and (3.20), and in view of Theorem 2.4, there exists a constant C such that

$$s \int_I v_0^s(x) dx \rightarrow C, \quad \text{and} \quad sv_0^s \rightarrow C \quad \text{strongly in } L^1(I).$$

Passing to the limit in (3.20) we deduce the inequality

$$\alpha_0 |u'|_{\mathcal{M}_b(I)} \leq \alpha_0 |u' - C|_{\mathcal{M}_b(I)} + \alpha_0 C \leq \liminf_{s \searrow 0} |u|_{TGV_\alpha^{1+s}(I)}.$$

The thesis follows owing to (3.19). \square

Corollary 3.5 (Asymptotic behavior of the fractional TGV seminorm-2). *Let $k \geq 2$. For every $u \in BV(I)$, up to the extraction of a (non-relabelled) subsequence there holds*

$$\lim_{s \nearrow 1} |u|_{TGV_\alpha^{k+s}(I)} = |u|_{TGV_\alpha^{k+1}(I)} \quad \text{and} \quad \lim_{s \searrow 0} |u|_{TGV_\alpha^{k+s}(I)} = |u|_{TGV_\alpha^k(I)},$$

where $\hat{\alpha} := (\alpha_0, \dots, \alpha_{k-1}) \in \mathbb{R}_+^k$.

Proof. The result follows by straightforward adaptations of the arguments in the proof of Theorem 3.2. \square

We proceed by showing that the minimization problem in Definition 3.1 has a solution.

Proposition 3.6. *If the infimum in Definition 3.1 is finite, then it is attained.*

Proof. Let $k = 1$. Let $\alpha \in \mathbb{R}_+^2$, and let $u \in BGV_\alpha^{1+s}(I)$. We need to show that

$$\begin{aligned} |u|_{TGV_\alpha^{1+s}(I)} = \min \left\{ \alpha_0 |u' - sv|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) |v|_{W^{s,1+s(1-s)}(I)} \right. \\ \left. + \alpha_0 s(1-s) \left| \int_I v(x) dx \right| : v \in W^{s,1+s(1-s)}(I) \right\}. \end{aligned} \quad (3.21)$$

We first observe that $u \in BV(I)$.

Indeed, let $\delta > 0$, and let $v \in W^{s,1+s(1-s)}(I)$ be such that

$$\alpha_0 |u' - sv|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) |v|_{W^{s,1+s(1-s)}(I)} + \alpha_0 s(1-s) \left| \int_I v(x) dx \right| \leq |u|_{TGV_\alpha^{1+s}(I)} + \delta.$$

By Hölder inequality there holds

$$\begin{aligned} \alpha_0 |u'|_{\mathcal{M}_b(I)} & \leq \alpha_0 |u' - sv|_{\mathcal{M}_b(I)} + \alpha_0 s \|v\|_{L^1(I)} \\ & \leq \alpha_0 |u' - sv|_{\mathcal{M}_b(I)} + \alpha_1 s |v|_{W^{s,1+s(1-s)}(I)} + \alpha_0 s \|v\|_{L^{1+s(1-s)}(I)} + \alpha_0 s(1-s) \left| \int_I v(x) dx \right| \\ & \leq |u|_{TGV_\alpha^{1+s}(I)} + \delta + \alpha_1 s^2 |v|_{W^{s,1+s(1-s)}(I)} + \alpha_0 s \|v\|_{L^{1+s(1-s)}(I)}, \end{aligned}$$

which implies the claim.

Let now $\{v_n\} \subset W^{s,1+s(1-s)}(I)$ be a minimizing sequence for (3.21). Since $s(1+s(1-s)) < 1$ for $s \in (0, 1)$, by Theorem 2.2 (1.) there exists a constant C such that

$$\sup_{n \in \mathbb{N}} \|v_n\|_{W^{s,1+s(1-s)}(I)} \leq C.$$

Thus, by Corollary 2.8 there exists $\bar{v} \in W^{s,1+s(1-s)}(I)$ such that, up to the extraction of a (non-relabelled) subsequence, there holds

$$v_n \rightharpoonup \bar{v} \quad \text{weakly in } W^{s,1+s(1-s)}(I),$$

and hence by Theorem 2.2 (1.),

$$v_n \rightarrow \bar{v} \quad \text{strongly in } L^1(I).$$

The thesis follows now by the lower semicontinuity of the total variation and the $W^{s,1+s(1-s)}$ -norm with respect to the L^1 convergence and the weak convergence in $W^{s,1+s(1-s)}(I)$, respectively.

For $k = 2$, let $\{v_0^n\} \subset BV(I)$ and $\{v_1^n\} \subset W^{s,1+s(1-s)}(I)$ be such that

$$\lim_{n \rightarrow +\infty} \left\{ \alpha_0 |u' - v_0^n|_{\mathcal{M}_b(I)} + \alpha_1 |(v_0^n)' - sv_1^n|_{\mathcal{M}_b(I)} + \alpha_2 s(1-s) |v_1^n|_{W^{s,1+s(1-s)}(I)} + \alpha_0 s(1-s) \left| \int_I v_1^n(x) dx \right| \right\} = TGV_\alpha^{2+s}(I).$$

Since $s(1+s(1-s)) < 1$ for $s \in (0, 1)$, by Theorem 2.2 (1.) we obtain that $\{v_1^n\}$ is uniformly bounded in $W^{s,1+s(1-s)}(I)$. Therefore, $\{v_0^n\}$ is uniformly bounded in $BV(I)$, and there exist $v_0 \in BV(I)$ and $v_1 \in W^{s,1+s(1-s)}(I)$ such that, up to the extraction of a (non-relabelled) subsequence,

$$v_0^n \xrightarrow{*} v_0 \quad \text{weakly* in } BV(I),$$

and

$$v_1^n \rightharpoonup v_1 \quad \text{weakly in } W^{s,1+s(1-s)}(I).$$

In particular, by Theorem 2.2 (1.),

$$v_1^n \rightarrow v_1 \quad \text{strongly in } L^1(I).$$

The minimality of v_0 and v_1 is a consequence of lower semicontinuity. The thesis for $k > 2$ follows by analogous arguments. \square

We observe that the TGV^{k+s} seminorms are all topologically equivalent to the total variation seminorm.

Proposition 3.7. *For every $k \geq 1$ and $0 < s < 1$, we have*

$$BV(I) \sim BGV^k(I) \sim BGV^{k+s}(I),$$

namely the three function spaces are topologically equivalent.

Proof. We only show that

$$BV(I) \sim BGV^{1+s}(I) \sim BGV^2(I). \quad (3.22)$$

The proof of the inequality for $k > 1$ is analogous. In view of (3.19), to prove the first equivalence relation in (3.22) we only need to show that there exist a constant C and a multi-index $\alpha \in \mathbb{R}_+^2$ such that

$$|u'|_{\mathcal{M}_b(I)} \leq C |u|_{TGV_\alpha^{1+s}(I)}.$$

By Theorem 2.2 we have

$$\begin{aligned} |u'|_{\mathcal{M}_b(I)} &\leq |u' - sv_0|_{\mathcal{M}_b(I)} + s \|v_0\|_{L^1(I)} \\ &\leq |u' - sv_0|_{\mathcal{M}_b(I)} + Cs |v_0|_{W^{s,1+s(1-s)}(I)} + s(1-s) \left| \int_I v_0(x) dx \right| \\ &= |u' - sv_0|_{\mathcal{M}_b(I)} + \frac{C}{(1-s)} s(1-s) |v_0|_{W^{s,1+s(1-s)}(I)} + s(1-s) \left| \int_I v_0(x) dx \right| \end{aligned}$$

for every $v_0 \in W^{s,1+s(1-s)}(I)$. Thus

$$|u'|_{\mathcal{M}_b(I)} \leq C |u|_{TGV_{1, \frac{C}{1-s}}^{1+s}}(I)$$

for every $s \in (0, 1)$. This completes the proof of the first equivalence in (3.22). Property (3.22) follows now by [8, Theorem 3.3]. \square

4. THE FRACTIONAL r -ORDER TGV^r FUNCTIONAL

In this section we introduce the fractional r -order TGV^r functional and prove a Γ -convergence result with respect to the parameters α and s .

Definition 4.1. Let $r \in [1, +\infty)$, $\alpha \in \mathbb{R}_+^{\lfloor r \rfloor + 1}$, and $u_\eta \in L^2(I)$. We define the functional $\mathcal{F}_\alpha^r : BV(I) \rightarrow [0, +\infty)$ as

$$\mathcal{F}_\alpha^r(u) := \int_I |u - u_\eta|^2 dx + |u|_{TGV_\alpha^r(I)}$$

for every $u \in BV(I)$. Note that the definition is well-posed due to Proposition 3.7.

The main result of this section reads as follows.

Theorem 4.2 (Γ -convergence of the fractional order TGV functional). *Let $k \in \mathbb{N}$ and $u_\eta \in L^2(I)$. Let $\{s_n\} \subset [0, 1]$ and $\{\alpha_n\} \subset \mathbb{R}_+^{k+1}$ be such that $s_n \rightarrow s$, and $\alpha_n \rightarrow \alpha$. Then, if $s \in (0, 1]$, the functional $\mathcal{F}_{\alpha_n}^{k+s_n}$ Γ -converges to \mathcal{F}_α^{k+s} in the weak* topology of $BV(I)$, namely for every $u \in BV(I)$ the following two conditions hold:*

(LI) If

$$u_n \xrightarrow{*} \text{weakly}^* \text{ in } BV(I),$$

then

$$\mathcal{F}_\alpha^{k+s}(u) \leq \liminf_{n \rightarrow +\infty} \mathcal{F}_{\alpha_n}^{k+s_n}(u_n).$$

(RS) There exists $\{u_n\} \subset BV(I)$ such that

$$u_n \xrightarrow{*} u \quad \text{weakly}^* \text{ in } BV(I),$$

and

$$\limsup_{n \rightarrow +\infty} \mathcal{F}_{\alpha_n}^{k+s_n}(u_n) \leq \mathcal{F}_\alpha^{k+s}(u).$$

The same result holds for $s = 0$, by replacing $\alpha = (\alpha_0, \dots, \alpha_k)$ with $\hat{\alpha} := (\alpha_0, \dots, \alpha_{k-1})$ in (LI) and (RS).

Remark 4.3. We recall that Γ -convergence is a variational convergence, originally introduced by E. De Giorgi and T. Franzoni in the seminar paper [19], which guarantees, roughly speaking, convergence of minimizers of the sequence of functionals to minimizers of the Γ -limit. The first condition in Theorem 4.2, known as *liminf inequality* ensures that the Γ -limit provides a lower bound for the asymptotic behavior of the $TGV_{\alpha_n}^{k+s_n}$ functionals, whereas the second condition, namely the *existence of a recovery sequence* guarantees that this lower bound is attained. We refer to [7] and [15] for a thorough discussion of the topic.

We subdivide the proof of Theorem 4.2 into two propositions. The next result will be crucial for establishing the liminf inequality.

Proposition 4.4. *Let $k \in \mathbb{N}$ and $u_\eta \in L^2(I)$. Let $\{s_n\} \subset [0, 1]$ and $\{\alpha_n\} \subset \mathbb{R}_+^{k+1}$ be such that $s_n \rightarrow s$, and $\alpha_n \rightarrow \alpha$. Let $\{u_n\} \in BV(I)$ be such that*

$$\sup_{n \in \mathbb{N}} \mathcal{F}_{\alpha_n}^{k+s_n}(u_n) < +\infty. \quad (4.1)$$

Then, there exists $u \in BV(I)$ such that, up to the extraction of a (non-relabeled) subsequence, there holds

$$u_n \xrightarrow{*} u \text{ in } BV(I). \quad (4.2)$$

In addition, if $s \in (0, 1]$ there holds

$$|u|_{TGV_\alpha^{k+s}(I)} \leq \liminf_{n \rightarrow +\infty} |u_n|_{TGV_{\alpha_n}^{k+s_n}(I)}. \quad (4.3)$$

If $s = 0$ we have

$$|u|_{TGV_{\hat{\alpha}}^k(I)} \leq \liminf_{n \rightarrow +\infty} |u_n|_{TGV_{\alpha_n}^{k+s_n}(I)}, \quad (4.4)$$

where $\hat{\alpha} \in \mathbb{R}_+^k$ is the multi-index $\hat{\alpha} := (\alpha_0, \dots, \alpha_{k-1})$.

Proof. We prove the statement for $k = 1$. The proof of the result for $k > 1$ follows via straightforward modifications.

For $k = 1$, we have $\{\alpha^n\} \subset \mathbb{R}^2$, and

$$(\alpha_0^n, \alpha_1^n) \rightarrow (\alpha_0, \alpha_1). \quad (4.5)$$

By Proposition 3.6 we deduce that there exists $v_0^n \in W^{s_n, 1+s_n(1-s_n)}(I)$ such that

$$\begin{aligned} |u_n|_{TGV_{\alpha^n}^{1+s_n}(I)} &= \alpha_0^n |u_n' - s_n v_0^n|_{\mathcal{M}_b(I)} + \alpha_1^n s_n (1-s_n) |v_0^n|_{W^{s_n, 1+s_n(1-s_n)}(I)} \\ &\quad + \alpha_0^n s_n (1-s_n) \left| \int_I v_0^n(x) dx \right|. \end{aligned} \quad (4.6)$$

We preliminary observe that (4.1), (4.5), and (4.6) yield the existence of a constant C such that

$$s_n(1-s_n) \left(|v_0^n|_{W^{s_n, 1+s_n(1-s_n)}(I)} + \left| \int_I v_0^n(x) dx \right| \right) + |u_n' - s_n v_0^n|_{\mathcal{M}_b(I)} \leq C |u_n|_{TGV_{\alpha^n}^{k+s_n}(I)} \leq C \quad (4.7)$$

for every $n \in \mathbb{N}$. For convenience of the reader we subdivide the proof into three steps.

Step 1: Assume first that $s \in (0, 1)$. By Proposition 3.4 there exists $v_0 \in W^{s, 1+s(1-s)}(I)$ such that

$$v_0^n \rightarrow v_0 \text{ strongly in } L^1(I), \quad (4.8)$$

and

$$\liminf_{n \rightarrow \infty} s_n(1-s_n) |v_0^n|_{W^{s_n, 1+s_n(1-s_n)}(I)} \geq s(1-s) |v_0|_{W^{s, 1+s(1-s)}(I)}. \quad (4.9)$$

By (4.1), (4.5), (4.6), and (4.8) there exists a constant C such that

$$|u_n'|_{\mathcal{M}_b(I)} \leq C,$$

and hence, again by (4.1),

$$\sup_{n \in \mathbb{N}} \|u_n\|_{BV(I)} \leq C \sup_{n \in \mathbb{N}} \|u_n\|_{BGV_{\alpha^n}^{1+s_n}(I)} < +\infty$$

which implies (4.2).

By (4.2), (4.6), (4.8), (4.9), and since $0 < s < 1$, there holds

$$\begin{aligned} \liminf_{n \rightarrow \infty} |u_n|_{TGV_\alpha^{1+s_n}(I)} &\geq \liminf_{n \rightarrow \infty} \alpha_0^n |u'_n - s_n v_0^n|_{\mathcal{M}_b(I)} \\ &\quad + \liminf_{n \rightarrow \infty} \alpha_1^n s_n (1 - s_n) |v_0^n|_{W^{s_n, 1+s_n(1-s_n)}(I)} + \liminf_{n \rightarrow \infty} \alpha_0^n s_n (1 - s_n) \left| \int_I v_0^n(x) dx \right| \\ &\geq \alpha_0 |u' - s v_0|_{\mathcal{M}_b(I)} + \alpha_1 s (1 - s) |v_0|_{W^{s, 1+s(1-s)}(I)} \\ &\quad + \alpha_0 s (1 - s) \left| \int_I v_0(x) dx \right| \geq |u|_{TGV_\alpha^{1+s}(I)}, \end{aligned}$$

where in the last inequality we used the definition of the TGV_α^{1+s} -seminorm. In particular, we deduce (4.3).

Step 2: Consider now the case in which $s = 1$. In view of Proposition 3.4, estimate (4.7) yields the existence of a map $\tilde{v}_0 \in BV(I)$ such that

$$v_0^n - \int_I v_0^n(x) dx \rightarrow \tilde{v}_0 \quad \text{strongly in } L^1(I) \quad (4.10)$$

and

$$|\tilde{v}'_0|_{\mathcal{M}_b(I)} \leq \liminf_{n \rightarrow +\infty} s_n (1 - s_n) |v_0^n|_{W^{s_n, 1+s_n(1-s_n)}(I)}. \quad (4.11)$$

On the other hand, by (4.7) there holds

$$\left| u'_n - s_n \int_I v_0^n(x) dx \right|_{\mathcal{M}_b(I)} \leq |u'_n - s_n v_0^n|_{\mathcal{M}_b(I)} + s_n \left\| v_0^n - \int_I v_0^n(x) dx \right\|_{L^1(I)} \leq C$$

for every $n \in \mathbb{N}$. Thus, there exists $\tilde{u} \in BV(I)$ such that

$$u_n - u_n(0) - s_n \left(\int_I v_0^n(x) dx \right) x \xrightarrow{*} \tilde{u} \quad \text{weakly* in } BV(I). \quad (4.12)$$

By (4.1) and by Definition 3.1, we have

$$\sup_{n \geq 1} \|u_n\|_{L^1(I)} \leq C. \quad (4.13)$$

In view of (4.12), testing the map $u_n - u_n(0) - s_n \left(\int_I v_0^n(x) dx \right) x$ against the function $x - \frac{1}{2}$, we obtain that

$$\lim_{n \rightarrow +\infty} \int_I u_n(x) \left(x - \frac{1}{2} \right) dx + \frac{s_n}{6} \int_I v_0^n(x) dx = \int_I \tilde{u}(x) \left(x - \frac{1}{2} \right) dx.$$

Thus, by (4.13),

$$\frac{s_n}{6} \left| \int_I v_0^n(x) dx \right| \leq \left| \int_I u_n(x) \left(x - \frac{1}{2} \right) dx + \frac{s_n}{6} \int_I v_0^n(x) dx \right| + \frac{1}{2} \|u_n\|_{L^1(I)} \leq C \quad (4.14)$$

for every $n \in \mathbb{N}$. By combining (4.10), (4.11), (4.12), and (4.14) we deduce that there exists $v_0 \in BV(I)$ such that

$$v_0^n \rightarrow v_0 \quad \text{strongly in } L^1(I), \quad (4.15)$$

and

$$|v'_0|_{\mathcal{M}_b(I)} \leq \liminf_{n \rightarrow +\infty} s_n (1 - s_n) |v_0^n|_{W^{s_n, 1+s_n(1-s_n)}(I)}. \quad (4.16)$$

In particular, by combining (4.7), (4.15), and (4.16) we have that

$$|u'_n|_{\mathcal{M}_b(I)} \leq C$$

for every $n \in \mathbb{N}$, which by (4.1) yields (4.2).

Step 3: Consider finally the case in which $s = 0$. In view of (4.7), and by Theorem 2.4 there holds

$$\left\| v_0^n - \int_I v_0^n(x) dx \right\|_{L^1(I)} \leq C \quad (4.17)$$

for every $n \in \mathbb{N}$. On the other hand, (4.7) yields

$$s_n \left| \int_I v_0^n(x) dx \right| \leq C \quad (4.18)$$

for every $n \in \mathbb{N}$. Combining (4.17) and (4.18) we conclude that there exists a constant $\lambda \in \mathbb{R}$ such that, up to the extraction of a (non-relabeled) subsequence there holds

$$s_n v_0^n \rightarrow \lambda \quad \text{strongly in } L^1(I). \quad (4.19)$$

As a result, in view of (4.1), Definition 3.1, and (4.7) we deduce the existence of a map $u \in BV(I)$ such that, up to the extraction of a (non-relabeled) subsequence,

$$u_n \xrightarrow{*} u \quad \text{weakly* in } BV(I). \quad (4.20)$$

Hence, by (4.6), (4.19), and (4.20) we have

$$\begin{aligned} \liminf_{n \rightarrow \infty} |u_n|_{TGV_{\alpha_n}^{1+s_n}(I)} &\geq \liminf_{n \rightarrow \infty} \alpha_0^n |u'_n - s_n v_0^n|_{\mathcal{M}_b(I)} \\ &+ \liminf_{n \rightarrow \infty} \alpha_1^n s_n (1 - s_n) |v_0^n|_{W^{s_n, 1+s_n(1-s_n)}(I)} + \liminf_{n \rightarrow \infty} \alpha_0^n s_n (1 - s_n) \left| \int_I v_0^n(x) dx \right| \\ &\geq \alpha_0 |u' - \lambda|_{\mathcal{M}_b(I)} + \alpha_0 |\lambda| \geq \alpha_0 |u'|_{\mathcal{M}_b(I)}, \end{aligned}$$

which in turn implies (4.4). This concludes the proof of the proposition. \square

The following result is instrumental for the construction of a recovery sequence.

Proposition 4.5. *Let $k \in \mathbb{N}$ and $u_\eta \in L^2(I)$. Let $\{s_n\} \subset [0, 1]$ and $\{\alpha_n\} \subset \mathbb{R}_+^{k+1}$ be such that $s_n \rightarrow s$, and $\alpha_n \rightarrow \alpha$. Let $u \in BV(I)$. Then, if $s \in (0, 1]$ there holds*

$$\lim_{n \rightarrow \infty} |u|_{TGV_{\alpha_n}^{k+s_n}(I)} = |u|_{TGV_\alpha^{k+s}(I)}, \quad (4.21)$$

and if $s = 0$,

$$\lim_{n \rightarrow \infty} |u|_{TGV_{\alpha_n}^{k+s_n}(I)} = |u|_{TGV_{\hat{\alpha}}^{k+s}(I)}, \quad (4.22)$$

where $\hat{\alpha} \in \mathbb{R}_+^k$ is the multi-index $\hat{\alpha} := (\alpha_0, \dots, \alpha_{k-1})$.

Proof. We prove the proposition for $k = 1$. The thesis for $k > 1$ can be proven by analogous arguments. The special cases $s = 0$ and $s = 1$ have already been analyzed in Theorem 3.2, in the situation in which $\alpha_n = \alpha$ for all $n \in \mathbb{N}$. The thesis for general sequences $\{\alpha_n\}$ follows by straightforward adaptations. Therefore, without loss of generality, we can assume that $s \in (0, 1)$. As a consequence of Proposition 4.4 we immediately have that

$$\liminf_{n \rightarrow \infty} |u|_{TGV_{\alpha}^{1+s_n}(I)} \geq |u|_{TGV_\alpha^{1+s}(I)}. \quad (4.23)$$

To prove the opposite inequality, fix $\varepsilon > 0$. By the definition of the $TGV_\alpha^{1+s}(I)$ -seminorm, and in view of [25, Theorems 2.4 and 5.4], there exists $v_0 \in C^\infty(I) \cap W^{s, 1+s(1-s)}(I)$ such that

$$|u|_{TGV_\alpha^{1+s}(I)} \geq \alpha_0 |u' - s v_0|_{\mathcal{M}_b(I)} + \alpha_1 s (1 - s) |v_0|_{W^{s, 1+s(1-s)}(I)} + \alpha_0 s (1 - s) \left| \int_I v_0(x) dx \right| - \varepsilon.$$

In particular, there holds $v_0 \in C^\infty(I) \cap W^{s_n, 1+s_n(1-s_n)}(I)$ for every $n \in \mathbb{N}$. Hence

$$|u|_{TGV_\alpha^{1+s_n}(I)} \leq \alpha_0 |u' - s_n v_0|_{\mathcal{M}_b(I)}$$

$$+ \alpha_1 s_n (1 - s_n) |v_0|_{W^{s_n, 1+s_n(1-s_n)}(I)} + \alpha_0 s_n (1 - s_n) \left| \int_I v_0(x) dx \right|,$$

for every $n \in \mathbb{N}$, and

$$\limsup_{n \rightarrow \infty} |u|_{TGV_\alpha^{1+s_n}(I)} \leq |u|_{TGV_\alpha^{1+s}(I)} + \varepsilon. \quad (4.24)$$

The thesis follows by the arbitrariness of ε , and by combining (4.23) and (4.24). \square

We conclude this section by proving Theorem 4.2.

Proof of Theorem 4.2. Property (LI) is a direct consequence of Proposition 4.4. Property (RS) follows by Proposition 4.5, choosing $u_n = u$ for every $n \in \mathbb{N}$. \square

5. THE BILEVEL TRAINING SCHEME EQUIPPED WITH TGV^r REGULARIZER

Let $r \in [1, +\infty)$ be given and recall $\lfloor r \rfloor$ denote the largest integer smaller than or equal to r . We propose the following training scheme (\mathcal{R}) which takes into account the order of derivation r of the regularizer and the parameter $\alpha \in \mathbb{R}_+^{\lfloor r \rfloor + 1}$ simultaneously. We restrict our analysis to the case in which α and r satisfy the *box constraint*

$$(\alpha, r) \in [P, 1/P]^{\lfloor r \rfloor + 1} \times [1, 1/P] \quad (5.1)$$

where $P \in (0, 1)$ small is a fixed real number.

Our new training scheme (\mathcal{R}) is defined as follows:

$$\text{Level 1.} \quad (\tilde{\alpha}, \tilde{r}) := \arg \min \left\{ \mathcal{I}(\alpha, r), (\alpha, r) \in [P, 1/P]^{\lfloor r \rfloor + 1} \times [1, 1/P] \right\}, \quad (5.2)$$

$$\text{Level 2.} \quad u_{\alpha, r} := \arg \min \{ \mathcal{F}_\alpha^r(u) : u \in BV(I) \} \quad (5.3)$$

where $\mathcal{I}(\alpha, r)$, defined as

$$\mathcal{I}(\alpha, r) := \|u_{\alpha, r} - u_c\|_{L^2(I)}^2 \quad (5.4)$$

is the *cost function* associated to the training scheme (\mathcal{R}) , \mathcal{F}_α^r is introduced in Definition 4.1, the map $u_c \in L^2(I)$ represents a noise-free test signal, and $u_\eta \in L^2(I)$ is the noise (corrupted) signal.

Note that we only allow the parameters α and the order r of regularizers to lie within a prescribed finite range. This is needed to force the optimal reconstructed signal $u_{\tilde{\alpha}, \tilde{r}}$ to remain inside our proposed space $BV(I)$ (see Proposition 4.4). In particular, if some of the components of $\tilde{\alpha}$ blow up to ∞ , we might end up in the space $W^{r,1}(I)$, which is outside the purview of this paper. We point out that P can be chosen as small as the user wants. Thus, despite the box constraint, our analysis still incorporates a large class of regularizers, such as TV and TGV^2 (see, e.g., [22]).

Before we state the main theorem of this section, we prove a technical lemma and show that (5.3) has a unique minimizer for each given (α, r) .

Lemma 5.1. *For every $r \in [1, 1/P]$ and $\alpha \in \mathbb{R}_+^{\lfloor r \rfloor + 1}$, there exists a unique $u_{\alpha, r} \in BV(I)$ solving the minimization problem (5.3).*

Proof. Let $\{u_n\} \subset BV(I)$ be a minimizing sequence for (5.3). By Proposition 3.7, $\{u_n\}$ is uniformly bounded in $BV(I)$. Thus there exists $u_{\alpha, r} \in BV(I)$ such that

$$u_n \xrightarrow{*} u_{\alpha, r} \quad \text{weakly* in } BV(I),$$

and hence also strongly in $L^2(I)$. The minimality of $u_{\alpha,r}$ follows then by Proposition 4.4, whereas the uniqueness of the minimum is a consequence of the strict convexity of the functional. \square

The next result guarantees existence of solutions to our training scheme.

Theorem 5.2. *Let $u_\eta, u_c \in BV(I)$ be given. Under the box constraint (5.1), the training scheme (\mathcal{R}) admits at least one solution $(\tilde{\alpha}, \tilde{r}) \in [P, 1/P]^{[\tilde{r}]+1} \times [1, 1/P]$ and provides an associated optimally reconstructed signal $u_{\tilde{\alpha}, \tilde{r}} \in BV(I)$.*

Proof. The result follows by Theorem 4.2 and by standard properties of Γ -convergence. We highlight the main steps for convenience of the reader.

Let $\{(\alpha_n, r_n)\}$ be a minimizing sequence for (5.2) with $\{r_n\} \subset [1, 1/P]$. Then, up to the extraction of a (non-reabeled) subsequence, there exists $\tilde{r} \in [1, 1/P]$ such that $r_n \rightarrow \tilde{r}$, and $\tilde{N} \in \mathbb{N}$ large enough such that for all $n \geq \tilde{N}$, one of the following two cases arises.

1. $r_n \in [[\tilde{r}], [\tilde{r}] + 1]$, or
2. $r_n \in [[\tilde{r}] - 1, [\tilde{r}]]$.

Suppose statement 1 holds (statement 2 can be handled by an analogous argument), then for $n \geq \tilde{N}$, we have $(\alpha_n, r_n) \in [P, 1/P]^{[\tilde{r}]+1} \times [[\tilde{r}], [\tilde{r}] + 1]$ for all $n \geq \tilde{N}$, and, up to extracting a further subsequence, $\alpha_n \rightarrow \tilde{\alpha}$ with $\tilde{\alpha} \in [P, 1/P]^{[\tilde{r}]+1}$. Let u_{α_n, r_n} be the unique solution to (5.3) provided by Lemma 5.1. By (5.3) and Proposition 3.7, there holds

$$\|u_{\alpha_n, r_n}\|_{BGV_{\alpha_n}^{r_n}(I)} \leq |u_\eta|_{TGV_{\alpha_n}^{r_n}(I)} \leq C |u'_\eta|_{\mathcal{M}_b(I)}$$

for all $n \geq \tilde{N}$. Next, in view of Proposition 4.4, there exists a $\tilde{u} \in BGV_{\tilde{\alpha}}^{\tilde{r}}(I)$ such that

$$u_{\alpha_n, r_n} \xrightarrow{*} \tilde{u} \quad \text{weakly* in } BV(I).$$

Thus, in particular, strongly in $L^2(I)$.

We claim that

$$\tilde{u} = u_{\tilde{\alpha}, \tilde{r}} = \arg \min \{ \mathcal{F}_{\tilde{\alpha}}^{\tilde{r}}(u) : u \in BV(I) \}.$$

Indeed, by Propositions 4.4 and 4.5 there holds

$$\begin{aligned} \int_I |\tilde{u} - u_\eta|^2 dx + |\tilde{u}|_{TGV_{\tilde{\alpha}}^{\tilde{r}}(I)} &\leq \liminf_{n \rightarrow +\infty} \int_I |u_{\alpha_n, r_n} - u_\eta|^2 dx + |u_{\alpha_n, r_n}|_{TGV_{\alpha_n}^{r_n}(I)} \\ &\leq \lim_{n \rightarrow +\infty} \left[\int_I |v - u_\eta|^2 dx + |v|_{TGV_{\alpha_n}^{r_n}(I)} \right] \leq \int_I |v - u_\eta|^2 dx + \limsup_{n \rightarrow \infty} |v|_{TGV_{\alpha_n}^{r_n}(I)} \\ &= \int_I |v - u_\eta|^2 dx + |v|_{TGV_{\tilde{\alpha}}^{\tilde{r}}(I)} \end{aligned}$$

for every $v \in BV(I)$, where at the last equality we used (4.21) or (4.22). This completes the proof of the claim and of the theorem. \square

Remark 5.3. The box constraint (5.1) is only used to guarantee that a minimizing sequence $\{(\alpha_n, r_n)\}$ has a convergent subsequence whose limit is bounded away from 0. Alternatively, different box-constraints for each parameter α and r might be enforced, such as

$$(\alpha, r) \in [P_1, Q_1] \times [P_2, Q_2] \times \cdots \times [1, Q_{[r]+2}]$$

where $0 < P_i < Q_i < +\infty$, $i = 1, \dots, [r] + 1$ and $1 < Q_{[r]+2} < +\infty$.

6. EXAMPLES AND INSIGHT

In order to gain further insight into the cost function $\mathcal{I}(\alpha, r)$, defined in (5.4), we compute it for a grid of values of α and r . We perform this analysis for two signals presenting different features, namely for a signal exhibiting corners (see Figure 1A) and for a signal with flat areas (see Figure 1B). In both cases, for simplicity, we assume $\alpha_0 = \alpha_1 = \alpha$ and we consider the discrete box-constraint

$$(\alpha, r) \in \{0, 0.005, 0.01, 0.015, \dots, 2.5\} \times \{1, 1.0025, 1.005, \dots, 2\} \quad (6.1)$$

(see Remark 5.3). The reconstructed signal $u_{\alpha, r}$ in (5.3) is computed by using the primal-dual algorithm presented in [10] and [28].

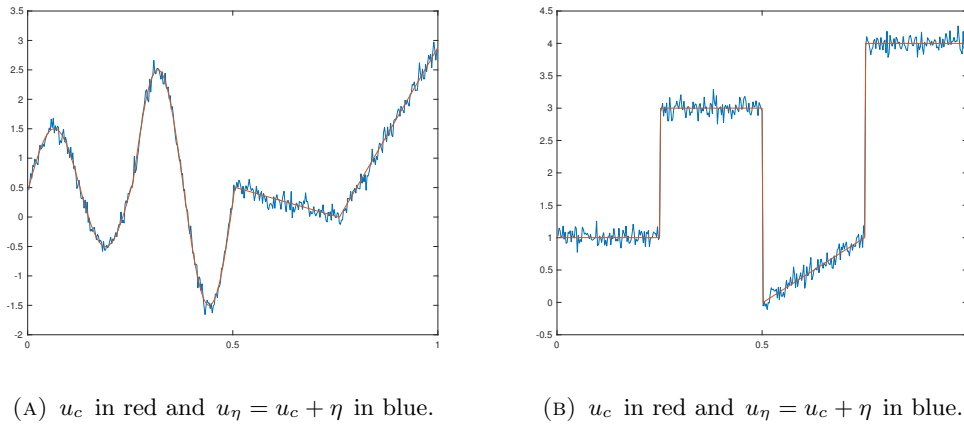


FIGURE 1. The artificial noise η is generated by using a Gaussian noise distribution.

The numerical landscapes of the cost function $\mathcal{I}(\alpha, r)$ are visualized in Figure 2A and Figure 2B, respectively.

As we can see, in Figure 2B the optimal \tilde{r} seems to lie away from the boundary of the discrete box-constraint (6.1), which are the integer values $r = 1$ and $r = 2$, showing an example in which the optimal signal reconstruction with respect to the L^2 -distance can be achieved by the fractional order TGV^r . We can also see in Figure 3 that the optimal denoising results are quite satisfactory. However, it is also possible that the optimal result \tilde{r} is an integer (in Figure 2A \tilde{r} is indeed very close to an integer). For example, for a complete flat signal, i.e., $u_c \equiv 1$, then $u_{\alpha, 1} = u_c$ for α large enough, provided that the noise η satisfies $\int_I \eta(x) dx = 0$ (zero-average assumption on noise is a reasonable assumption, see [42]). We point out once more that the introduction of fractional r -order TGV^r only meant to expand the training choices for the bilevel scheme, but not to provide a superior seminorm to integer order TGV^k . The optimal solution $\tilde{r} \in [1, 1/P]$, fractional order or integer order, is completely up to the given data u_c and u_η .

Although from the numerical landscapes the cost function $\mathcal{I}(\alpha, r)$ (Figure 2) appears to be almost quasiconvex (see, [6, Section 3.4]) in the variable α , this is not the case. In the forthcoming paper [21] some explicit counterexamples showing that at least for certain piecewise constant signals $\mathcal{I}(\alpha, 1)$ is not quasiconvex will be presented. Additionally, Figure 2A

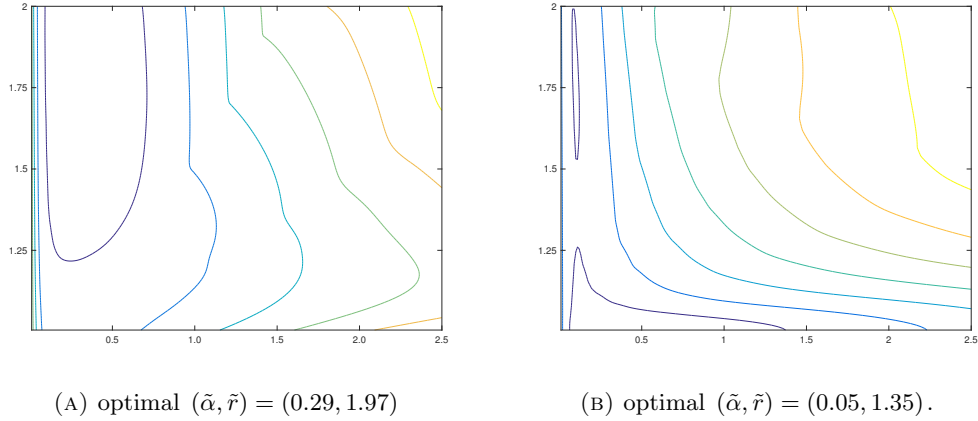


FIGURE 2. Numerical landscape (contour map) for the cost function $\mathcal{I}(\alpha, r)$. Darker colors correspond to smaller values of $\mathcal{I}(\alpha, r)$

and 2B both show that $\mathcal{I}(\alpha, r)$ is not quasiconvex in the r variable. The non-quasiconvexity of the cost function $\mathcal{I}(\alpha, r)$, implies that the training scheme (\mathcal{R}) may not have a unique solution, i.e., the global minimizer of $\mathcal{I}(\alpha, r)$ might be not unique. In particular, the non-quasiconvexity of $\mathcal{I}(\alpha, r)$ prevents us from using standard gradient descent methods to find a global minimizer (which is the optimal solution we are looking for in (5.2)). Therefore, the identification of a reliable numerical scheme for solving the bilevel problem (upper level problem) remains an open question.

As a final remark, we point out that the development of a numerical scheme to identify the global minimizers of $\mathcal{I}(\alpha, 1)$, $\alpha \in \mathbb{R}^+$ has been undertaken in [21], where the *Bouligand* differentiability and the finite discretization of $\mathcal{I}(\alpha, 1)$ will be analyzed.

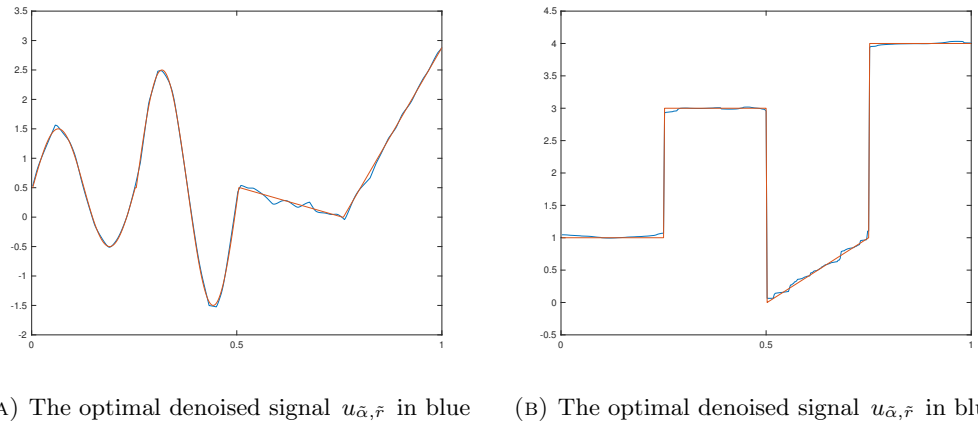


FIGURE 3. The optimal denoised signal $u_{\tilde{\alpha}, \tilde{r}}$ in both cases are very close to the given clean signal.

ACKNOWLEDGEMENTS

P. Liu is partially funded by the National Science Foundation under Grant No. DMS - 1411646. E. Davoli is supported by the Austrian Science Fund (FWF) projects P27052 and F65. The authors wish to thank Irene Fonseca for suggesting this project and for many helpful discussions and comments. The authors also thank Giovanni Leoni for comments on the subject of Section 2.

REFERENCES

- [1] R.A. Adams, J.J.F. Fournier. *Sobolev spaces*. Elsevier/Academic Press, Amsterdam, second edition, 2003.
- [2] L. Ambrosio, N. Fusco, D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.
- [3] M. Bergounioux. Optimal control of problems governed by abstract elliptic variational inequalities with state constraints. *SIAM J. Control Optim.* **36**(1) (1998), 273–289.
- [4] J. Bourgain, H. Brezis, P. Mironescu. Limiting embedding theorems for $W^{s,p}$ when $s \uparrow 1$ and applications. *J. Anal. Math.* **87** (2002), 77–101.
- [5] J. Bourgain, H. Brezis, P. Mironescu. Another look at sobolev spaces. *Optimal Control and Partial Differential Equations*, J. L. Menaldi, E. Rofman, A. Sulem (Eds.), a volume in honor of A. Bensoussan's 60th birthday, IOS Press, Amsterdam, pages 439–455, 2001.
- [6] S. Boyd, L. Vandenberghe. *Convex optimization*. Cambridge Univ. Press, Cambridge, 2004.
- [7] A. Braides. Γ -convergence for beginners. Oxford Lecture Series in Mathematics and its Applications. Oxford University Press, Oxford, 2002.
- [8] K. Bredies, T. Valkonen. Inverse problems with second-order total generalized variation constraints. *Proceedings of SampTA 2011 - 9th International Conference on Sampling Theory and Applications, Singapore*, 2011.
- [9] M. Burger, K. Papafitsoros, E. Papoutsellis, C.-B. Schönlieb. Infimal convolution regularisation functionals of BV and L^p spaces. Part I: the finite p case. *J. Math. Imaging Vision*, **55**(3) (2016), 343–369.
- [10] A. Chambolle, T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vision*, **40** (1) (2011), 120–145.
- [11] T. Chan, A. Marquina, P. Mulet. High-order total variation-based image restoration. *SIAM J. Sci. Comput.* **22**(2) (2000), 503–516.
- [12] D. Chen, S. Sun, C. Zhang, Y. Chen, D. Xue. Fractional-order $TV-L^2$ model for image denoising. *Central European Journal of Physics*, **11**(10) (2013), 1414–1422.
- [13] Y. Chen, T. Pock, R. Ranftl, H. Bischof. Revisiting loss-specific training of filter-based mrfs for image restoration. In J. Weickert, M. Hein, and B. Schiele, editors, *Pattern Recognition: 35th German Conference, GCPR 2013, Saarbrücken, Germany, September 3-6, 2013. Proceedings*, pages 271–281. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
- [14] Y. Chen, R. Ranftl, T. Pock. Insights into analysis operator learning: From patch-based sparse models to higher order mrfs. *IEEE Transactions on Image Processing*, **23**(3) (2014), 1060–1072.
- [15] G. Dal Maso. *An introduction to Γ -convergence*. Birkhäuser Boston, Inc., Boston, MA, 1993.
- [16] G. Dal Maso, I. Fonseca, G. Leoni, M. Morini. A higher order model for image restoration: the one-dimensional case. *SIAM J. Math. Anal.* **40**(6) (2009), 2351–2391.
- [17] J. Dávila. On an open question about functions of bounded variation. *Calc. Var. Partial Differential Equations*, **15**(4) (2002), 519–527.
- [18] E. Davoli, P. Liu. 2D Fractional order TGV: Gamma-convergence and bilevel training scheme. *In preparation*.
- [19] E. De Giorgi, T. Franzoni. Su un tipo di convergenza variazionale. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* **58** (1975), 842–850.
- [20] J. C. De los Reyes, C.-B. Schönlieb. Image denoising: learning the noise model via nonsmooth PDE-constrained optimization. *Inverse Probl. Imaging*, **7**(4) (2013), 1183–1214.
- [21] J. C. De los Reyes, P. Liu, C.-B. Schönlieb. Numerical analysis of bilevel parameter training scheme in total variation image reconstruction. *In preparation*.
- [22] J. C. De los Reyes, C.-B. Schönlieb, T. Valkonen. Bilevel parameter learning for higher-order total variation regularisation models. *J. Math. Imaging Vision*, (2016), 1–25.
- [23] M. Benning, C.-B. Schönlieb, T. Valkonen, V. Vlačić. Explorations on anisotropic regularisation of dynamic inverse problems by bilevel optimisation. *ArXiv:1602.01278*, (2017).

- [24] J.C. De los Reyes, C.-B. Schönlieb, T. Valkonen. The structure of optimal parameters for image restoration problems. *J. Math. Anal. Appl.* **434**(1) (2016), 464–500.
- [25] E. Di Nezza, G. Palatucci, E. Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.* **136**(5) (2012), 521–573.
- [26] J. Domke. Generic methods for optimization-based modeling. In N.D. Lawrence and M.A. Girolami, editors, *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics (AISTATS-12)*, volume 22, pages 318–326, 2012.
- [27] J. Domke. Learning graphical model parameters with approximate marginal inference. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **35**(10) (2013), 2454–2467.
- [28] J.B. Garnett, T.M. Le, Y. Meyer, L.A. Vese. Image decompositions using bounded variation and generalized homogeneous Besov spaces. *Appl. Comput. Harmon. Anal.* **23**(1) (2007), 25–56.
- [29] N.S. Landkof. *Foundations of modern potential theory*. Springer-Verlag, New York-Heidelberg, 1972.
- [30] G. Leoni. *A first course in Sobolev spaces*. American Mathematical Society, Providence, RI, 2009.
- [31] V.G. Maz'ya, T.O. Shaposhnikova. On the Bourgain, Brezis, and Mironescu theorem concerning limiting embeddings of fractional Sobolev spaces. *J. Funct. Anal.* **195**(2) (2002), 230–238.
- [32] V.G. Maz'ya, T.O. Shaposhnikova. *Theory of Sobolev multipliers*. Springer-Verlag, Berlin, 2009.
- [33] K. Papafitsoros, K. Bredies. A study of the one dimensional total generalised variation regularisation problem. *Inverse Probl. Imaging*, **9**(2) (2015), 511–550.
- [34] K. Papafitsoros, C.B. Schönlieb. A combined first and second order variational approach for image reconstruction. *J. Math. Imaging Vision*, **48**(2) (2014), 308–338.
- [35] L.I. Rudin. Segmentation and restoration using local constraints. *Cognitech Report* **36** (1993).
- [36] L.I. Rudin, S. Osher. Total variation based image restoration with free local constraints. In *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, volume 1, pages 31–35 vol.1, Nov 1994.
- [37] L.I. Rudin, S. Osher, E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, **60** (1992), 259–268.
- [38] J. Simon. Sobolev, Besov and Nikol'skiĭ fractional spaces: embeddings and comparisons for vector valued spaces on an interval. *Ann. Mat. Pura Appl.* **157** (1990), 117–148 .
- [39] M.F. Tappen. Utilizing variational optimization to learn markov random fields. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, June 2007.
- [40] M.F. Tappen, C. Liu, E.H. Adelson, W.T. Freeman. Learning gaussian conditional random fields for low-level vision. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, June 2007.
- [41] H. Triebel. *Theory of function spaces*. Modern Birkhäuser Classics. Birkhäuser/Springer Basel AG, Basel, 2010.
- [42] Chambolle, Antonin and Lions, Pierre-Louis. *Image recovery via total variation minimization and related problems*. Numer. Math., 1997
- [43] T. Valkonen, K. Bredies, F. Knoll. Total generalized variation in diffusion tensor imaging. *SIAM J. Imaging Sci.* **6**(1) (2013), 487–525.