Deep Variational Models

Part III Essay

A thesis submitted in partial fulfilment of the requirements for the degree of

Master of Advanced Study in Mathematics

by

Timo Sprekeler

May 2017

University of Cambridge, Trinity College

Supervisor: Dr. Carola-Bibiane Schönlieb

Abstract

The purpose of this essay is to analyse a recently proposed bilevel learning scheme for image reconstruction in one dimension which optimises under a box constraint with respect to both the regularisation parameter and the order of the regulariser. The learning scheme is based on $ICTV$ regularisation and is distinguished by the fact that it involves fractional order $ICTV$ seminorms, which happens to reduce the staircasing effect. Along the analysis of the existence and uniqueness for the bilevel learning scheme, the fractional Sobolev spaces are studied, leading to a result concerning the asymptotic behaviour of the Gagliardo seminorm due to Bourgain, Brezis and Mironescu. The fractional $ICTV$ seminorms are introduced and the asymptotic behaviour is investigated. In particular they are shown to lie intermediate between the surrounding integer $ICTV$ seminorms. Further, the existence of extremal functions for these new seminorms is established and the connection with the total variation is analysed. Finally, a theorem guaranteeing the existence and uniqueness of a solution to the learning scheme is proven.

Contents

1 Introduction

In Image Analysis and Signal Processing, variational noise removal techniques and PDE methods have gained lots of popularity over the last decades. In a paper from 1992, the mathematicians L. I. Rudin, S. Osher and E. Fatemi introduced a total variation based noise removal method which can be considered as the starting point of this development, see [FOR92]. This method relies on minimising the so-called ROF-functional

$$ROF_\alpha(u) := \frac{1}{2}\|u - u_0\|^2_{L^2(I)} + \alpha TV(u),$$

where $I := (0,1)$ is the domain of a 1D-image/signal, $u_0 \in L^2(I)$ represents the corrupted image, and $TV(u)$ denotes the total variation of u, i.e.

$$TV(u) = \sup_{\substack{\phi \in C_c^1(I) \\ \|\phi\|_\infty \leq 1}} \int_0^1 u(x)\phi'(x)\, dx = \sup_{\substack{P = \{x_0, \ldots, x_{n_P}\} \\ P \text{ partition of } I}} \sum_{k=0}^{n_P - 1} |u(x_{k+1}) - u(x_k)|.$$

The corrupted image u_0 is of the form

$$u_0 = u_c + \eta,$$

where u_c represents the clean picture and η is the noise that we want to remove.

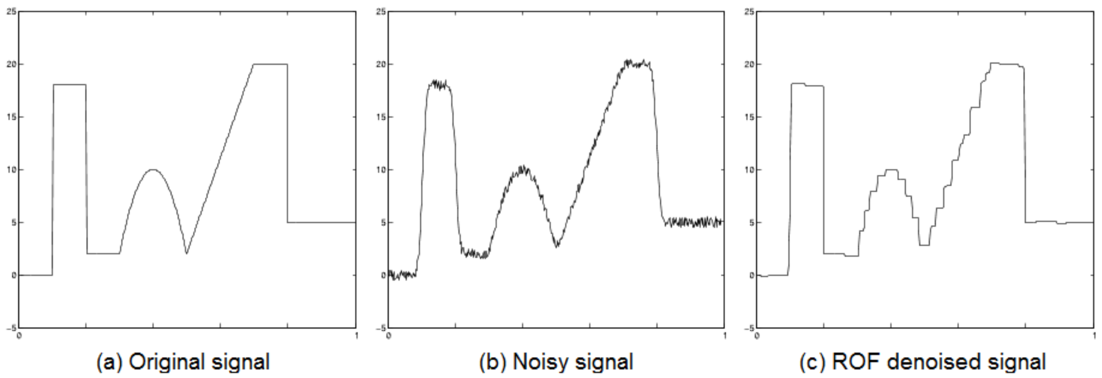
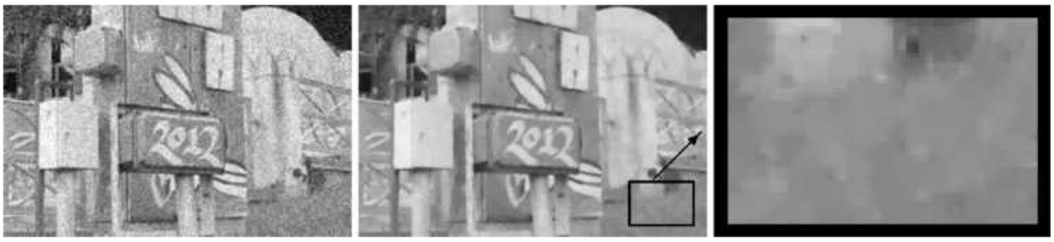


Figure 1: (taken from [WX09]) *ROF denoising of a signal including piecewise constant, piecewise linear and piecewise parabolic parts.*

As figure 1 shows, the ROF method has the property that it preserves discontinuities in noisy step functions. However, there is a problem with this scheme. When using total variation based image reconstruction schemes, the solutions are usually piecewise constant which results in an effect called staircasing. In two dimensions, the staircasing effect is responsible for the formation of blocks in the reconstructed image, see figure 2.



(a) Noisy image (b) TV reconstructed image (c) Staircasing (block-formation)

Figure 2: (taken from [PS14]) *TV image reconstruction and Staircasing.*

The outcome of minimising the ROF-functional depends on the choice of the so-called regularisation parameter $\alpha \in \mathbb{R}_+$. On the one hand, if we choose α to be very small, the functional basically minimises the L^2-error to the corrupted image, which means that we get a reconstructed image that is very close to u_0, but where the noise is still not removed. On the other hand, if we choose α to be very large, the functional basically just minimises the total variation, which results in an over-smoothed image that is not necessarily very close to u_0. Therefore, one is interested in finding "good" or even optimal regularisation parameters.
In order to find optimal regularisation parameters, one uses so-called bilevel learning schemes, which adapt themselves to given perfect data. Bilevel learning schemes are constrained optimisation problems where the constraint consists of solving an optimisation problem itself. The latter one is called the lower level problem or the second level problem. The idea of the learning scheme is now as follows. Suppose our variational method is based on minimising the functional $\mathcal{J} = \mathcal{J}_\alpha$, where α is the vector consisting of all regularisation parameters. It would be a very difficult task to tune the regularisation parameters by hand in order to find a suitable constellation (particularly when optimising for more than one parameter as in the subsequently discussed $ICTV$ regularisation, cf. (1.1)). Instead, we choose a quality measure Q such that $Q(u_\alpha)$ evaluates the quality of a minimiser $u_\alpha \in \arg\min_u J_\alpha(u)$. Then the bilevel learning scheme minimises the quality measure with respect to the parameter vector α, i.e. it is of the form

$$\begin{cases} \min\limits_{\alpha} & Q(u_\alpha) \\ \text{s.t.} & u_\alpha \in \arg\min\limits_{u} \mathcal{J}_\alpha(u) \end{cases}.$$

An example for a bilevel learning scheme which uses the ROF-functional as variational approach and the squared L^2-error to some clean test picture $u_c \in L^2(I)$ as quality measure is the following:

$$(\mathcal{B}_1) \begin{cases} \min\limits_{\alpha>0} & \|u_\alpha - u_c\|^2_{L^2(I)} \\ \text{s.t.} & u_\alpha = \arg\min\limits_{u \in BV(I)} ROF_\alpha(u) \end{cases}.$$

The scheme $(\mathcal{B}_1)$ then looks for the optimal regularisation parameter α^* such that the ROF-reconstruction u^*_α minimises the L^2-error to u_c. The disadvantage of this scheme is again the staircasing effect, since it is based on minimising the ROF-functional.
There are two popular regularisation methods that are known for reducing the staircasing effect, the $ICTV$ (infimal-convolution total variation) regulariser and the TGV (total generalised variation) regulariser, which coincide in one dimension.
The integer $ICTV$ seminorm/regulariser of order $k+1 \in \mathbb{N}_{>1}$ and weight $\alpha = (\alpha_0, \dots, \alpha_k) \in \mathbb{R}^{k+1}_+$ on I is defined by

$$|u|_{ICTV^{k+1}_\alpha(I)} := \inf_{\substack{v_l \in BV(I) \\ \forall 0 \le l \le k-1}} \left\{ \alpha_0 |u' - v_0|_{\mathcal{M}_b(I)} + \sum_{i=0}^{k-2} \alpha_{i+1} |v_i' - v_{i+1}|_{\mathcal{M}_b(I)} + \alpha_k |v_{k-1}'|_{\mathcal{M}_b(I)} \right\}, \tag{1.1}$$

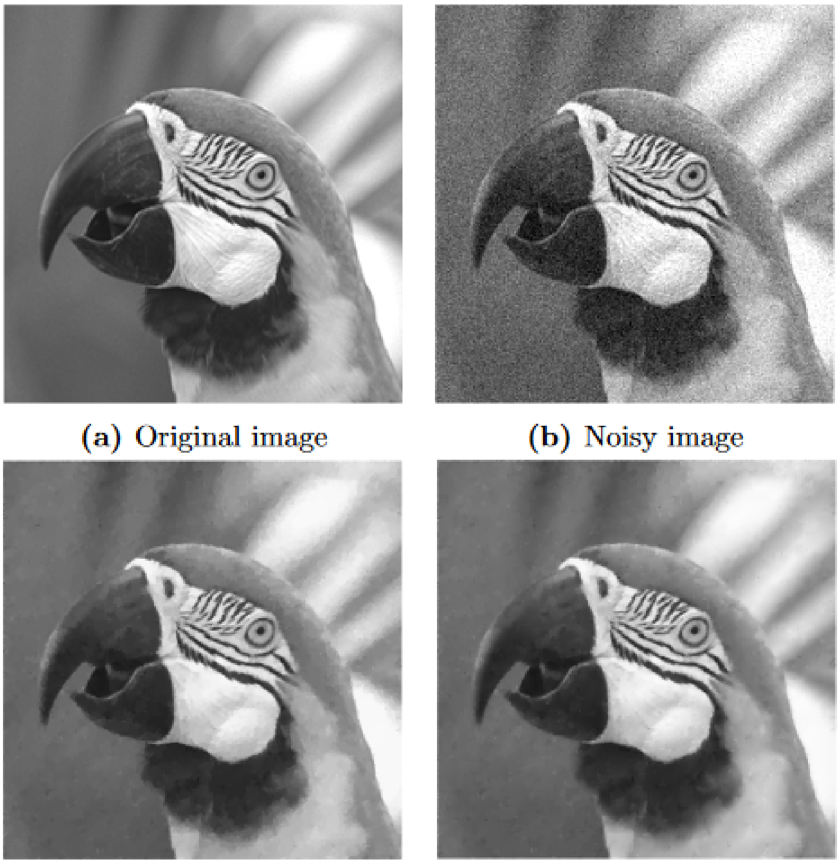
where we are using the notation $|u'|_{\mathcal{M}_b} \equiv TV(u)$.
An often used regulariser in this class is the $ICTV$ regulariser of order two, which reads as

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

The definition of the regularisers can easily be adapted to a two dimensional setting. A result of image reconstruction by using $ICTV^2$ (with squared L^2 fidelity/data term) is shown in figure 3.
We see that the staircasing effect in the $ICTV^2$ reconstructed image is reduced compared to the TV reconstructed image. To see this mathematically, we take a function $v_0 \in BV(I)$ and look at the expression in the $ICTV^2$ regulariser. If we think of v_0 as the distributional derivative of some other function $w_0 \in BV(I)$, then the first term of the regulariser can be seen as first order TV term involving the total variation of w_0 and the second term can be seen as second order TV term involving the total variation of w_0', which is small if w_0 is piecewise linear. Therefore, this second order TV term can counteract possible staircasing created by the first order TV term.
In this essay, we will investigate the following bilevel learning scheme that optimises both the regularisation parameter and the order of the $ICTV$ regulariser (the order of derivation) proposed by E. Davoli and P. Liu in [DL16]:

$$(\mathcal{B}) \begin{cases} (\alpha^*, r^*) & := \underset{(\alpha, r)}{\arg\min} \left\{ \|u_{\alpha,r} - u_c\|^2_{L^2(I)} : (\alpha, r) \in [a, A]^{\lfloor r \rfloor + 1} \times [1, R] \right\} \\ u_{\alpha,r} & := \underset{u \in BCV^r_\alpha(I)}{\arg\min} \left(\|u - u_0\|^2_{L^2(I)} + |u|_{ICTV^r_\alpha(I)} \right) \end{cases}.$$

In this scheme, $a, A > 0$ and $R > 1$ are fixed real numbers and BCV^r_α denotes the space of functions with corresponding finite $ICTV^r_\alpha$ seminorm (definition 10). To give sense to the scheme, we need to define $ICTV$ seminorms for non-integer orders.



(a) Original image (b) Noisy image

(c) TV denoising (d) ICTV^2 denoising

Figure 3: (taken from [DSV17]) *TV and $ICTV^2$ image reconsruction.*

The advantage in working with seminorms of fractional order is that fractional order derivatives reduce contrast and staircasing effects.
We will define the fractional order $ICTV$ seminorms $|\cdot|_{ICTV_\alpha^{k+s}}$ for $k \in \mathbb{N}$ and $s \in (0,1)$ in such a way that they are in some sense intermediate between the surrounding integer $ICTV$-spaces. For instance (see definition 9), the fractional $ICTV$ regulariser of order $1+s$ for $s \in (0,1)$ is defined by

$$|u|_{ICTV_\alpha^{1+s}(I)} := \inf_{\substack{v_0 \in W^{s,1+s(1-s)}(I),\\ (v_0)_I = 0}} \left\{ \alpha_0 |u' - sv_0|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s)|v_0|_{W^{s,1+s(1-s)}(I)} \right\}$$

and the corresponding $ICTV$ space is defined by

$$BCV_\alpha^{1+s}(I) := \{u \in L^1(I) : |u|_{ICTV_\alpha^{1+s}(I)} < \infty\}.$$

This definition already shows that we need to study fractional Sobolev spaces before we can understand the fractional $ICTV$ spaces. The structure of this essay is as follows.

Structure. The goal of this essay is to analyse the existence and uniqueness of a solution (α^*, r^*) to $(\mathcal{B})$ (which corresponds to the reconstructed image u_{α^*,r^*}).
The main reference for this essay is [DL16].
Section 2 is based on [Alt06], [EG92] and [Sch16] and provides a quick review of the cornerstones in the theory of functions of bounded variation. The main results of this section will be approximation and compactness in the space BV.
In section 3, we are going to study fractional Sobolev spaces. The first part of the section serves as an introduction to the main results in the theory of fractional Sobolev spaces and provides an important toolbox including embedding theorems and a Poincaré-type inequality, which will be used frequently throughout this essay. The mentioned results concerning classical Sobolev space theory can be found in [Eva10]. In the second part of the section, we are going to study the asymptotic behaviour of the so-called Gagliardo seminorm on which section 4 crucially relies. The results of this subsection are based on [Ma14] and [BBM02], where the latter one is the main reference.
In section 4 we define both the integer $ICTV$ regulariser as well as the fractional $ICTV$ regulariser and study the connection between them by looking at the limit behaviour of the fractional $ICTV$ seminorms. The section is concluded by investigating the existence of extremal functions for the fractional $ICTV$ seminorms and a theorem relating the $ICTV$ seminorms to the total variation. The results in this section provide the foundation for the study of the bilevel learning scheme $(\mathcal{B})$ in section 5.
In section 5, we can finally use our results from the sections 3 and 4 to prove an existence and uniqueness theorem for the learning scheme $(\mathcal{B})$. The results in the sections 4 and 5 are based on [DL16].

Contribution. This essay is not intended to provide new research results. Rather, the aim of this essay is to come up with an improved version of the original paper [DL16]. This involves firstly the correction of any mistakes made in [DL16] which I was able to find, and secondly an attempt to improve the clarity of the argumentation in the proofs.
In particular, this version of the proof of theorem 11 corrects and simplifies the original one, and arose from collaboration with Pan Liu (especially lemma 2).
Further, this essay is aimed to be as self-contained as possible and hence it might be accessible to a wider audience.

Acknowledgements. First of all I would like to thank my advisor Dr Carola-Bibiane Schönlieb for suggesting this interesting topic and for her continuous support. Throughout the essay-writing process she provided plenty of good ideas and always had an open ear for my questions.
Secondly I would like to thank Pan Liu for a very fruitful discussion of his paper [DL16] and in particular for working together with me on the proof of theorem 11.
It was a great pleasure for me to discuss mathematics with both Carola Schönlieb and Pan Liu on this modern topic of bilevel learning.

2 Functions of Bounded Variation

2.1 Recap on signed Radon-measures

In this section, we will recall some facts about signed Radon-measures. The results will not be proven in this essay, but we refer to [Rud87] for a more extensive discussion of the topic. We start by defining signed Radon-measures.

Definition 1 (Signed Radon-measure)**.** *Let $\mathcal{U}$ be a set and let $\mathcal{A}$ be a σ-algebra on $\mathcal{U}$. A map $\mu : \mathcal{A} \to (-\infty, \infty)$ is called a signed Radon-measure, if the following two properties are satisfied:*

(i) μ is a signed measure, i.e. for any pairwise disjoint family of sets $(A_n)_n \subset \mathcal{A}$ there holds

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

(ii) The variation of μ,

$$|\mu| : \mathcal{A} \longrightarrow [0,\infty]$$

$$|\mu|(A) = \sup\left\{\sum_{n\in\mathbb{N}} |\mu(A_n)| : \bigcup_{n\in\mathbb{N}} A_n \subset A,\ A_m \cap A_r = \emptyset\ \forall m \neq r\right\},$$

defines a Radon-measure, i.e. a Borel-measure that is finite on compact sets and that satisfies the exterior regularity condition

$$\forall A \in \mathcal{U} :\ |\mu|(A) = \inf_{\substack{B\supset A \\ B\subset\mathcal{U}\ open}} |\mu|(B).$$

We denote by $\mathcal{M}(\mathcal{U})$ the space of all signed Radon-measures.

Remark 1. *The space $\mathcal{M}(\mathcal{U})$ endowed with the total variation norm*

$$\|\mu\|_{\mathcal{M}(\mathcal{U})} := |\mu|(\mathcal{U})$$

is a Banach space.

The main theorem in the study of the space $\mathcal{M}$ is Riesz's representation theorem. Before we state the theorem, we recall the definition of two important function spaces.

Definition 2. *Let $\mathcal{U} \subset \mathbb{R}^N$. We define the space of compactly supported continuous functions*

$$C_c(\mathcal{U}) := \{f \in C(\mathcal{U};\mathbb{R}) : \text{supp}(u) \text{ compact}\}$$

and the space of continuous functions that vanish at infinity by

$$\begin{aligned} C_0(\mathcal{U}) &:= \overline{C_c(\mathcal{U})}^{\|\cdot\|_\infty} \\ &= \{f \in C(\mathcal{U};\mathbb{R}) : \exists (f_n)_n \subset C_c(\mathcal{U})\ s.t.\ f_n \longrightarrow f \text{ uniformly}\} \\ &= \{f \in C(\mathcal{U};\mathbb{R}) : \forall \varepsilon > 0\ \exists K \subset \mathcal{U} \text{ compact } s.t.\ |f(x)| < \varepsilon\ \forall x \in \mathcal{U}\backslash K\} \\ &= \left\{f \in C(\mathcal{U};\mathbb{R}) : \lim_{x\to\pm\infty} f(x) = 0\right\}. \end{aligned}$$

Remark 2. *The space $C_0(\mathcal{U})$ endowed with the norm $\|\cdot\|_\infty$ is a Banach space.*

Riesz's representation theorem provides a connection of the spaces $C_0(\mathcal{U})$ and $\mathcal{M}(\mathcal{U})$. To be precise, it says that $\mathcal{M}(\mathcal{U})$ is the dual of the space $C_0(\mathcal{U})$.

Theorem 1 (Riesz's Representation Theorem, $C_0' = \mathcal{M}$)**.** *Let $\mathcal{U} \subset \mathbb{R}^N$ and let $\lambda : C_0(\mathcal{U}) \to \mathbb{R}$ be linear and continuous. Then there exists an unique signed Radon-measure $\mu \in \mathcal{M}(\mathcal{U})$ such that*

$$\lambda(f) = \int_{\mathcal{U}} f \; d\mu.$$

There holds

$$\|\mu\|_{\mathcal{M}(\mathcal{U})} = \|\lambda\|_{\mathcal{L}(C_0(\mathcal{U});\mathbb{R})} := \sup_{\substack{f \in C_0(\mathcal{U}) \\ \|f\|_\infty \leq 1}} |\lambda(f)|.$$

We note that this theorem implies that for any bounded sequence $(u_n)_n$ in $L^1(\mathcal{U}) \subset \mathcal{M}(\mathcal{U}) = C_0'(\mathcal{U})$ there exists a signed Radon-measure $\mu \in \mathcal{M}(\mathcal{U})$ and a subsequence such that $u_n \overset{*}{\rightharpoonup} \mu$ in $\mathcal{M}(\mathcal{U})$ along this subsequence, i.e.

$$\int_{\mathcal{U}} u_n \cdot f \; d\mathcal{L}^N \longrightarrow \int_{\mathcal{U}} f \; d\mu$$

for any $f \in C_0(\mathcal{U})$.
Another important application of the theorem is the following. Consider the space $X := C_0(\mathbb{R}^N)$. Since X is separable, there holds weak-$*$ compactness in the dual $X' = \mathcal{M}(\mathbb{R}^N)$ which means that for any bounded sequence $(\mu_n)_n \subset \mathcal{M}(\mathbb{R}^N)$ there exists some $\mu \in \mathcal{M}(\mathbb{R}^N)$ and a subsequence such that $\mu_n \overset{*}{\rightharpoonup} \mu$ in $\mathcal{M}(\mathbb{R}^N)$, i.e.

$$\int_{\mathbb{R}^N} f \; d\mu_n \longrightarrow \int_{\mathbb{R}^N} f \; d\mu$$

for any $f \in C_0(\mathbb{R}^N)$.
Besides the weak-$*$ convergence, there is also a notion of weak convergence in $\mathcal{M}(\mathbb{R}^N)$.

Definition 3 (Weak Convergence in $\mathcal{M}(\mathbb{R}^N)$)**.** *Let $(\mu_n)_n, \mu$ be positive Radon-measures on $\mathbb{R}^N$. Then we say that the sequence $(\mu_n)_n$ converges weakly to μ, denoted by $\mu_n \overset{\mathcal{M}}{\rightharpoonup} \mu$, if*

$$\int_{\mathbb{R}^N} f \; d\mu_n \longrightarrow \int_{\mathbb{R}^N} f \; d\mu$$

for any $f \in C_c(\mathbb{R}^N)$.

To round the section off, we note the following connection between weak and weak-$*$ convergence:

$$\mu_n \overset{*}{\rightharpoonup} \mu \iff \mu_n \overset{\mathcal{M}}{\rightharpoonup} \mu \text{ and } \|\mu_n\|_{\mathcal{M}(\mathbb{R}^N)} \leq C.$$

2.2 Definition and first properties

In some tasks in image analysis, in the analysis of partial differential equations and in the calculus of variations, the given problem turns out to naturally provide no more than an estimate on the L^1-norm, which often results in difficulties. The problem is that since L^1 is not reflexive, we cannot extract a convergent subsequence for a L^1 bounded sequence. In particular, we do not get easily a limit function (in the PDE context the limit would be a candidate for the solution of the equation). To overcome this problem, one passes into the space of signed Radon-measures $\mathcal{M}$ instead of working in L^1.
The motivation for defining the space of functions of bounded variation is the problem described above for the non-reflexive Sobolev space $W^{1,1}$. We would like to have a Banach space that contains $W^{1,1}$ and where a bounded sequence has a convergent subsequence. We will see that the function space BV possesses these properties.

Notation. *Let $\mathcal{U} \subset \mathbb{R}^N$ be open. For a function u of integrability class $L^1(\mathcal{U})$, we can identify u with the distribution*

$$\langle u \rangle : \ C_c^\infty(\mathcal{U}) \longrightarrow \mathbb{R}, \ \phi \longmapsto \langle u, \phi \rangle := \int_\mathcal{U} u\phi \ d\mathcal{L}^N,$$

and we write $\nabla u \in \mathcal{M}(\mathcal{U}; \mathbb{R}^N)$ if the distributional gradient

$$\nabla \langle u \rangle : \ C_c^\infty(\mathcal{U}; \mathbb{R}^N) \longrightarrow \mathbb{R}, \ \phi \longmapsto \langle \nabla u, \phi \rangle := -\langle u, \nabla \cdot \phi \rangle = -\int_\mathcal{U} u \left(\nabla \cdot \phi\right) \ d\mathcal{L}^N$$

can be extended to an element of $C_0(\mathcal{U}; \mathbb{R}^N)' = \mathcal{M}(\mathcal{U}; \mathbb{R}^N)$, i.e. to a continuous linear functional on $C_0(\mathcal{U}; \mathbb{R}^N)$.

Definition 4 (The space BV)**.** *Let $\mathcal{U} \subset \mathbb{R}^N$ be open. We define*

$$BV(\mathcal{U}) := \{u \in L^1(\mathcal{U}) : \nabla u \in \mathcal{M}(\mathcal{U}; \mathbb{R}^N)\}$$

to be the space of functions of bounded variation over $\mathcal{U}$.

We note that the so-defined function space, endowed with the right norm, is actually a Banach space which contains $W^{1,1}(\mathcal{U})$. This means that the condition $\nabla u \in \mathcal{M}(\mathcal{U}; \mathbb{R}^N)$ is weaker than the condition $\nabla u \in L^1(\mathcal{U}; \mathbb{R}^N)$, in the sense that the latter one is more restrictive. Moreover, we could have defined $BV(\mathcal{U})$ equivalently in a different way as the following theorem shows.

Theorem 2. *Let $\mathcal{U} \subset \mathbb{R}^N$ be open. Then the following holds:*

(i) *The space $BV(\mathcal{U})$ can be equivalently written as*

$$BV(\mathcal{U}) = \{u \in L^1(\mathcal{U}) : TV(u) < \infty\},$$

where the total variation of the function u is defined by

$$TV(u) = \sup_{\substack{\phi \in C_c^1(\mathcal{U}; \mathbb{R}^N) \\ \|\phi\|_\infty \leq 1}} \int_\mathcal{U} u \left(\nabla \cdot \phi\right) \ d\mathcal{L}^N.$$

(ii) $W^{1,1}(\mathcal{U})$ *is contained in* $BV(\mathcal{U})$.

(iii) $BV(\mathcal{U})$ *endowed with the norm*

$$\|u\|_{BV(\mathcal{U})} := \|u\|_{L^1(\mathcal{U})} + TV(u) \tag{2.1}$$

is a Banach space.

Proof. (*i*) Firstly, let $u \in BV(\mathcal{U})$. Then ∇u is understood as the extension of $\nabla\langle u\rangle$ to $C_0(\mathcal{U};\mathbb{R}^N)$. By definition of $\|\cdot\|_{\mathcal{M}(\mathcal{U};\mathbb{R}^N)} = \|\cdot\|_{\mathcal{L}(C_0(\mathcal{U};\mathbb{R}^N);\mathbb{R})}$, we have for all $\phi \in C_c^1(\mathcal{U};\mathbb{R}^N) \subset C_0(\mathcal{U};\mathbb{R}^N)$ that

$$\int_{\mathcal{U}} u\,(\nabla\cdot\phi)\;d\mathcal{L}^N = -\langle\nabla u,\phi\rangle \le |\langle\nabla u,\phi\rangle| \le \|\nabla u\|_{\mathcal{M}(\mathcal{U};\mathbb{R}^N)}\|\phi\|_\infty.$$

So $TV(u) \le \|\nabla u\|_{\mathcal{M}(\mathcal{U};\mathbb{R}^N)} < \infty$.
For the other direction, let $u \in L^1(\mathcal{U})$ with $TV(u) < \infty$. By definition of the distributional gradient $K = \nabla\langle u\rangle$, the finite total variation of u implies

$$\sup_{\substack{\phi\in C_c^\infty(\mathcal{U};\mathbb{R}^N)\\ \|\phi\|_\infty\le 1}} |K(\phi)| \le C < \infty. \tag{2.2}$$

This estimate implies that we can extend K to a continuous linear functional on $C_0(\mathcal{U};\mathbb{R}^N)$ as follows: Let $\phi^* \in C_0(\mathcal{U};\mathbb{R}^N)$. By approximation, there exists a sequence $(\phi_n)_n \subset C_c^\infty(\mathcal{U};\mathbb{R}^N)$ with $\phi_n \to \phi^*$ uniformly. Using linearity of K and our estimate (2.2), we obtain

$$|K(\phi_m) - K(\phi_r)| = |K(\phi_m - \phi_r)| \le C\|\phi_m - \phi_r\|_{L^\infty(\mathcal{U})} \underset{m,r\to\infty}{\longrightarrow} 0.$$

Hence, $(K(\phi_n))_n$ is as Cauchy sequence in $\mathbb{R}$ convergent and we can extend K by $K(\phi^*) := \lim_{n\to\infty} K(\phi_n)$ to $K \in \mathcal{L}(C_0(\mathcal{U};\mathbb{R}^N);\mathbb{R})$.

(*ii*) Let $u \in W^{1,1}(\mathcal{U})$. By definition, $u \in L^1(\mathcal{U})$ and the distributional gradient can be represented by an element in $L^1(\mathcal{U})$, i.e. there exists a function $g \in L^1(\mathcal{U})$ such that $K(\phi) := \langle\nabla u,\phi\rangle = \langle g,\phi\rangle$ for all test functions $\phi \in C_c^\infty(\mathcal{U};\mathbb{R}^N)$. The Cauchy-Schwarz inequality yields

$$\sup_{\substack{\phi\in C_c^\infty(\mathcal{U};\mathbb{R}^N)\\ \|\phi\|_\infty\le 1}} |K(\phi)| \le \|g\|_{L^1(\mathcal{U})} < \infty.$$

As we have seen in the proof of (*i*), this estimate implies that we can extend K to a continuous linear functional on $C_0(\mathcal{U};\mathbb{R}^N)$. Hence, $u \in BV(\mathcal{U})$.

(*iii*) It is easy to see that $BV(\mathcal{U})$ is a vector subspace of $L^1(\mathcal{U})$ and that (2.1) defines a norm on $BV(\mathcal{U})$. It remains to show that the space is complete. This follows directly from the completeness of $L^1(\mathcal{U})$ and the lower semi-continuity of the total variation with respect to L^1-convergence. We will prove the latter fact at the beginning of the next subsection.

□

The strong convergence in BV provided by the norm defined in (2.1) is just rarely used. We will define a more useful type of convergence in BV, the weak-star convergence. The advantage of the weak-star convergence is that one can show that given a bounded sequence in BV, we can extract a weakly-star convergent subsequence in BV.

Definition 5 (Weak-$*$ convergence in BV)**.** *Let $\mathcal{U} \subset \mathbb{R}^N$ be open and let $(u_n)_n$ be a sequence in $BV(\mathcal{U})$. We say that $(u_n)_n$ converges weakly-$*$ in $BV(\mathcal{U})$ to some limit function $u \in BV(\mathcal{U})$ if*

$$u_n \cdot \mathcal{L}^N \overset{*}{\rightharpoonup} u \cdot \mathcal{L}^N, \quad \nabla u_n \overset{*}{\rightharpoonup} \nabla u \quad \text{in } \mathcal{M}(\mathcal{U}).$$

2.3 Approximation and Compactness

The goal of this section is to prove that we can approximate BV-functions by smooth functions and that bounded sequences in BV have L^1-convergent subsequences. As a first step, we have to show that L^1-limits are actually in BV.

Lemma 1 (Lower semi-continuity)**.** *Let $\mathcal{U} \subset \mathbb{R}^N$ be open. Further, let $(u_n)_n \subset BV(\mathcal{U})$ with $u_n \to u$ in $L^1(\mathcal{U})$. Then there holds the inequality*

$$\|u\|_{BV(\mathcal{U})} \leq \liminf_{n\to\infty} \|u_n\|_{BV(\mathcal{U})}. \tag{2.3}$$

In particular, if the right hand side is finite, then $u \in BV(\mathcal{U})$.

Proof. Since $u_n \to u$ strongly in $L^1(\mathcal{U})$, we have norm-convergence, i.e.

$$\lim_{n\to\infty} \|u_n\|_{L^1(\mathcal{U})} = \|u\|_{L^1(\mathcal{U})}. \tag{2.4}$$

Let now $\phi \in C_c^1(\mathcal{U};\mathbb{R}^N)$ with $\|\phi\|_\infty \leq 1$. Using the strong L^1-convergence we obtain

$$\int_{\mathcal{U}} u\,(\nabla\cdot\phi)\ d\mathcal{L}^N = \lim_{n\to\infty} \int_{\mathcal{U}} u_n\,(\nabla\cdot\phi)\ d\mathcal{L}^N \leq \liminf_{n\to\infty} TV(u_n).$$

Hence, there holds

$$TV(u) \leq \liminf_{n\to\infty} TV(u_n),$$

which implies together with (2.4),

$$\begin{aligned}
\|u\|_{BV(\mathcal{U})} = \|u\|_{L^1(\mathcal{U})} + TV(u) &\leq \lim_{n\to\infty} \|u_n\|_{L^1(\mathcal{U})} + \liminf_{n\to\infty} TV(u_n) \\
&\leq \liminf_{n\to\infty} \big(\|u_n\|_{L^1(\mathcal{U})} + TV(u_n)\big) \\
&= \liminf_{n\to\infty} \|u_n\|_{BV(\mathcal{U})},
\end{aligned}$$

the desired inequality (2.3). □

One of the most important and useful theorems in the theory of functions of bounded variation is the result that functions of bounded variation can be approximated by smooth functions. This result is useful since it allows to prove things first for smooth functions, which is often a lot easier than proving it directly, and once this is done, one uses an approximation argument based on the following approximation theorem.

Theorem 3 (Approximation Theorem)**.** *Let $\mathcal{U} \subset \mathbb{R}^N$ be open. Then given a function $u \in BV(\mathcal{U})$, there exists a sequence $(u_n)_n \subset BV(\mathcal{U}) \cap C^\infty(\mathcal{U})$ of smooth functions with*

$$u_n \longrightarrow u \text{ in } L^1(\mathcal{U}), \text{ and}$$
$$TV(u_n) \longrightarrow TV(u).$$

The proof in the general case is rather technical due to boundary difficulties. To see the idea behind the proof, we will prove the theorem for the special case $\mathcal{U} = \mathbb{R}^N$ and refer to [EG92] for the proof of the general case.

Proof. ($\mathcal{U} = \mathbb{R}^N$) Let $u \in BV(\mathbb{R}^N)$ be arbitrary. Then by definition of the space BV, $u \in L^1(\mathbb{R}^N)$ and $\nabla u \in \mathcal{M}(\mathbb{R}^N; \mathbb{R}^N)$. We choose $\chi \in C_c^\infty(\mathbb{R}^N)$ with $0 \leq \chi \leq 1$, $\int\limits_{\mathbb{R}^N} \chi \, d\mathcal{L}^N = 1$ and define the mollifiers

$$\chi_\varepsilon(x) := \frac{1}{\varepsilon^N} \chi\left(\frac{x}{\varepsilon}\right), \quad \varepsilon > 0.$$

By standard Lebesgue-theory, the functions $u_n := u * \chi_{\frac{1}{n}}$ are of regularity class $C^\infty(\mathbb{R}^N)$ and there holds $u_n \to u$ in $L^1(\mathbb{R}^N)$.
In particular, this implies $\|u_n\|_{L^1(\mathbb{R}^N)} \to \|u\|_{L^1(\mathbb{R}^N)}$, thus it remains to show that $TV(u_n) \to TV(u)$. To this end, let $\phi \in C_c^1(\mathbb{R}^N; \mathbb{R}^N)$ with $\|\phi\|_\infty \leq 1$ and compute, using the properties of the convolution,

$$\begin{aligned}
\int\limits_{\mathbb{R}^N} u_n \left(\nabla \cdot \phi\right) \, d\mathcal{L}^N &= \int\limits_{\mathbb{R}^N} \left(u * \chi_{\frac{1}{n}}\right) \left(\nabla \cdot \phi\right) \, d\mathcal{L}^N \\
&= \int\limits_{\mathbb{R}^N} u \left(\left(\nabla \cdot \phi\right) * \chi_{\frac{1}{n}}\right) \, d\mathcal{L}^N \\
&= \int\limits_{\mathbb{R}^N} u \; \nabla \cdot \left(\phi * \chi_{\frac{1}{n}}\right) \, d\mathcal{L}^N \leq TV(u),
\end{aligned}$$

where we used in the last step that $\phi * \chi_{\frac{1}{n}} \in C_c^1(\mathbb{R}^N; \mathbb{R}^N)$ and $\|\phi * \chi_{\frac{1}{n}}\|_\infty \leq 1$. This calculation shows $(u_n)_n \subset BV(\mathbb{R}^N)$ and

$$\limsup_{n\to\infty} TV(u_n) \leq TV(u).$$

By using lemma 1, we obtain

$$TV(u) \leq \liminf_{n\to\infty} TV(u_n) \leq \limsup_{n\to\infty} TV(u_n) \leq TV(u).$$

Hence $TV(u_n) \to TV(u)$ and the theorem is proven for the case $\mathcal{U} = \mathbb{R}^N$. □

When we think back to our motivation of defining the space of functions with bounded variations, the key point was that we wanted to have a space where bounded sequences possess convergent subsequences. With the approximation theorem at hand, we can state and prove the desired compactness result for the space BV.

Theorem 4 (Compactness Theorem)**.** *Let $\mathcal{U} \subset \mathbb{R}^N$ be open and bounded. Then given a sequence of functions $(u_n)_n \subset BV(\mathcal{U})$ with $\|u_n\|_{BV(\mathcal{U})} \leq C$ for some constant $C > 0$, there exists a function $u \in BV(\mathcal{U})$ such that along a subsequence there holds*

$$u_n \longrightarrow u \text{ in } L^1(\mathcal{U}).$$

Proof. Step 1: We show that for a smooth function $f \in BV(\mathcal{U}) \cap C^\infty(\mathcal{U})$ there holds

$$TV(f) = \int_{\mathcal{U}} |\nabla f| \; d\mathcal{L}^N. \tag{2.5}$$

For all $\phi \in C_c^1(\mathcal{U}; \mathbb{R}^N)$ with $\|\phi\|_\infty \leq 1$ we find by using Gauß's theorem

$$\int_{\mathcal{U}} f\,(\nabla \cdot \phi) \; d\mathcal{L}^N = -\int_{\mathcal{U}} \nabla f \cdot \phi \; d\mathcal{L}^N \leq \|\phi\|_\infty \int_{\mathcal{U}} |\nabla f| \; d\mathcal{L}^N \leq \int_{\mathcal{U}} |\nabla f| \; d\mathcal{L}^N,$$

which shows

$$TV(f) \leq \int_{\mathcal{U}} |\nabla f| \; d\mathcal{L}^N. \tag{2.6}$$

To show the other direction, we consider a sequence of functions $(\phi_n)_n \subset C_c^1(\mathcal{U}; \mathbb{R}^N)$ with $\|\phi_n\|_\infty \leq 1$ and

$$\phi_n \underset{n\to\infty}{\longrightarrow} \phi_f := -\frac{\nabla f}{|\nabla f|} \quad \text{in } L^1(\mathcal{U}).$$

This is possible since $\phi_f \in L^1(\mathcal{U}; \mathbb{R}^N)$ with $\|\phi_f\|_\infty = 1$ due to density of $C_c^1(\mathcal{U}; \mathbb{R}^N)$ in $L^1(\mathcal{U}; \mathbb{R}^N)$. We use again Gauß's theorem to obtain

$$\int_{\mathcal{U}} f\,(\nabla \cdot \phi_n) \; d\mathcal{L}^N = -\int_{\mathcal{U}} \nabla f \cdot \phi_n \; d\mathcal{L}^N \underset{n\to\infty}{\longrightarrow} -\int_{\mathcal{U}} \nabla f \cdot \phi_f \; d\mathcal{L}^N = \int_{\mathcal{U}} |\nabla f| \; d\mathcal{L}^N.$$

Together with (2.6), we deduce (2.5).
Step 2: Let now $(u_n)_n$ be a bounded sequence in $BV(\mathcal{U})$. We apply the approximation theorem to each u_n and obtain a sequence $(f_n)_n \subset BV(\mathcal{U}) \cap C^\infty(\mathcal{U})$ with

$$\|f_n - u_n\|_{L^1(\mathcal{U})} \leq \frac{1}{n} \quad \text{and} \quad TV(f_n) \leq c \quad \forall n \in \mathbb{N}$$

for some constant $c > 0$. Using $\|u_n\|_{BV(\mathcal{U})} \leq C$ and "Step 1", this implies the boundedness of $(f_n)_n$ in $W^{1,1}(\mathcal{U})$. Since the embedding $W^{1,1}(\mathcal{U}) \hookrightarrow L^1(\mathcal{U})$ is compact, there exists some $u \in L^1(\mathcal{U})$ such that $f_n \to u$ in $L^1(\mathcal{U})$ for a subsequence. Then there also holds $u_n \to u$ in $L^1(\mathcal{U})$ along a subsequence since

$$\|u - u_n\|_{L^1(\mathcal{U})} \leq \|u - f_n\|_{L^1(\mathcal{U})} + \|f_n - u_n\|_{L^1(\mathcal{U})} \leq \|u - f_n\|_{L^1(\mathcal{U})} + \frac{1}{n} \longrightarrow 0.$$

It remains to show that $u \in BV(\mathcal{U})$, but this follows from lemma 1. □

To conclude this section, we remark that by adaptations (and using Rellich-Kondrachov), one can show the following stronger compactness result:

Remark 3. *If $\mathcal{U} \subset \mathbb{R}^N$ is a bounded domain with Lipschitz regularity and $(u_n)_n$ a bounded sequence in $BV(\mathcal{U})$, then there exists a subsequence that converges strongly in $L^q(\mathcal{U})$ for all $q \in [1, 1^*)$, where $1^* := \frac{N}{N-1}$ ($:= \infty$ if $N = 1$).*

3 Fractional Sobolev Spaces

3.1 Definition and main results

As motivation let us firstly recall that the Sobolev space $W^{s,p}(I)$ for $s \in \mathbb{N}$ and $p \in [1, \infty]$ is defined as the space of $L^p(I)$ functions whose distributional derivatives up to order s can be represented by a $L^p(I)$ function, i.e.

$$W^{s,p}(I) := \{u \in L^p(I) : D^\alpha u \in L^p(I)\ \forall \alpha \leq s\}.$$

The space $W^{s,p}(I)$ endowed with the norm

$$\|u\|_{W^{s,p}(I)} = \sum_{\alpha \leq s} \|D^\alpha u\|_{L^p(I)},$$

is a Banach space. We would like to allow the derivative-index s also to lie in $\mathbb{R}_+ \backslash \mathbb{N}$. So we want to define for $s \in \mathbb{R}_+ \backslash \mathbb{N}$ a Banach space $W^{s,p}$ which is intermediate between the integer derivative-index Sobolev spaces $W^{\lfloor s \rfloor, p}$ and $W^{\lfloor s \rfloor + 1, p}$ whose norm also measures fractional-order derivatives. The way this is done is by a modified Hölder condition.

Definition 6 (Gagliardo seminorm)**.** *Let $s \in (0,1)$ and $p \in [1, \infty)$. We define for $u \in L^p(I)$ the Gagliardo seminorm by*

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

Definition 7 (Fractional Sobolev spaces)**.** *Let $s \in \mathbb{R}_+ \backslash \mathbb{N}$ and $p \in [1, \infty)$. We define the Fractional Sobolev space with derivative-index s and integrability-index p to be*

$$W^{s,p}(I) := \left\{ u \in W^{\lfloor s \rfloor, p}(I) : |D^{\lfloor s \rfloor} u|_{W^{s - \lfloor s \rfloor, p}(I)} < \infty \right\}$$

which becomes a Banach space endowed with the norm

$$\|u\|_{W^{s,p}(I)} := \|u\|_{W^{\lfloor s \rfloor, p}(I)} + |D^{\lfloor s \rfloor} u|_{W^{s - \lfloor s \rfloor, p}(I)}.$$

Remark 4. *In this essay we will mostly consider Fractional Sobolev spaces with derivative-index $s \in (0,1)$, for which the definition reads as follows:*

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

One of the most important tools in Sobolev space theory are the Sobolev embeddings, so it is just natural to ask whether we also have them for fractional Sobolev spaces. Fortunately the answer to this question is positive. We will just state embeddings into L^q spaces and into Sobolev spaces, but it should be mentioned that there hold also embeddings into C^α spaces.

Before we state the theorem, we want to think about what we would expect by simply adapting the conditions from the Sobolev embedding theorems for integer derivative-order Sobolev spaces.
We denote the Sobolev number corresponding to the Sobolev space $W^{s,p}$ by $\gamma_{s,p}$. Recall that in one dimension

$$\gamma_{s,p} = s - \frac{1}{p}.$$

For a non-positive Sobolev number, i.e. in the case $sp \leq 1$, we would expect

$$W^{s,p} \hookrightarrow L^q$$

for some $q \geq 1$ if

$$\gamma_{s,p} \geq \gamma_{0,q},$$

which is true without further restriction if $sp = 1$, and for $q \leq \frac{p}{1-sp}$ if $sp < 1$.
The following theorem confirms that our expectations are correct. It includes also an important compactness statement for q strictly less than the "fractional critical exponent" $p^* = \frac{p}{1-sp}$.

Theorem 5 (Sobolev Embedding into L^q spaces). *Let $s \in (0,1)$ and $p \in [1,\infty)$ such that $sp \leq 1$.*

(*i*) *If $sp < 1$, there exists a constant $C = C(s,p) > 0$ such that there holds*

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

for all $q \in [1, p^*]$, where

$$p^* = \frac{p}{1-sp}$$

is the fractional critical exponent.
Moreover, the embedding $W^{s,p}(I) \hookrightarrow L^q(I)$ is compact for all $1 \leq q < p^*$.

(*ii*) *If $sp = 1$, then there holds the continuous embedding*

$$W^{s,p}(I) \hookrightarrow L^q(I)$$

for all $1 \leq q < \infty$.

Proof. We refer to [DPV12]. □

Another powerful Sobolev embedding theorem is about embeddings into other Sobolev spaces

$$W^{s,p} \hookrightarrow W^{r,q}$$

with higher integrability index $q \geq p$. By adapting the classical theory we would expect the condition

$$\gamma_{s,p} \geq \gamma_{r,q}.$$

Since $q \geq p$, this implies $r \leq s$, which means that we cannot gain both higher integrability and higher differentiability. Our expectations are again correct as the following theorem shows:

Theorem 6 (Sobolev Embedding into $W^{r,q}$ spaces). *Let $s, r \in (0,1)$ and $p, q \in [1, \infty)$ such that $q \geq p$, $r \leq s$ and*

$$s - \frac{1}{p} \geq r - \frac{1}{q}.$$

Then there holds $W^{s,p}(I) \subset W^{r,q}(I)$ and

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

Proof. We refer to [Sim90]. □

Apart from the embedding theorems, there are also some nice other properties. By identification of fractional Sobolev spaces with Besov spaces, one can show that the fractional $W^{s,p}$ spaces are reflexive for $p > 1$, see e.g. [Tri10].

Theorem 7 (Reflexivity). *The fractional Sobolev spaces $W^{s,p}$ are reflexive for $s \in \mathbb{R}_+ \backslash \mathbb{N}$, $p \in (1, \infty)$.*

In many cases, we will have a bound on the Gagliardo seminorm $|\cdot|_{W^{s,p}(I)}$ for some sequence, but we would like to bound the full norm $\|\cdot\|_{W^{s,p}(I)}$ such that we can make use of the reflexivity to deduce the existence of a weakly convergent subsequence. This can be done by using the following Poincaré inequality.

Theorem 8 (Poincaré inequality). *Let $s \in (0,1)$ and $p \in [1, \infty)$ such that $sp < 1$. Then there exists a constant $C > 0$ such that*

$$\|u - (u)_I\|^p_{L^{p^*}(I)} \leq C \frac{s(1-s)}{(1-sp)^{p-1}} |u|^p_{W^{s,p}(I)},$$

where $(u)_I$ denotes the mean value of u over I, i.e. $(u)_I = \frac{1}{|I|} \int u(x)\, dx$.

Proof. We refer to [BBM02] where the result is proven for $s \geq \frac{1}{2}$ and to [MS02] for the generalisation to $s \in (0,1)$. □

Note that since I is bounded and since the fractional critical exponent satisfies $p^* \geq p$ whenever $sp < 1$, this theorem provides in particular an estimate for the $L^p(I)$-norm.

3.2 Limit behaviour of the Gagliardo seminorm as $s \nearrow 1$

In this section we will study the limit behaviour of the Gagliardo seminorm

$$|u|_{W^{s,1}(I)} = \int_I \int_I \frac{|u(x) - u(y)|}{|x - y|^{1+s}} \, dx \, dy \qquad (0 < s < 1) \tag{3.1}$$

as $s \nearrow 1$. Clearly, for non-constant smooth functions $u \in C^\infty(I)$, there holds

$$\lim_{s \nearrow 1} |u|_{W^{s,1}(I)} = \infty.$$

In particular the Gagliardo seminorm (3.1) does not converge to

$$|u|_{W^{1,1}(I)} = \|u'\|_{L^1(I)} = \int_I |u'(x)| \, dx$$

as $s \nearrow 1$. However, as we shall see, there holds the following result, which goes back to Brezis, Bourgain and Mironescu, and can be found in [BBM02].

Theorem 9. *Let $u \in BV(I)$. Then there holds*

$$\lim_{s \nearrow 1} (1 - s)|u|_{W^{s,1}(I)} = TV(u). \tag{3.2}$$

As a first step, we take an arbitrary sequence $(\rho_\varepsilon)_{\varepsilon>0}$ of radial mollifiers, i.e. a family of nonnegative radial functions $\rho_\varepsilon : \mathbb{R} \to \mathbb{R}^+$ which are absolutely integrable and satisfy

- $\int_0^\infty \rho_\varepsilon(x) dx = 1,$
- $\forall \delta > 0 : \lim_{\varepsilon \to 0} \int_\delta^\infty \rho_\varepsilon(x) dx = 0.$

We observe that it is enough to show that for all $u \in BV(I)$ there holds

$$\int_I \int_I \frac{|u(x) - u(y)|}{|x - y|} \rho_\varepsilon(x - y) \, dx \, dy \longrightarrow TV(u) \quad \text{ as } \varepsilon \to 0. \tag{3.3}$$

Indeed, to see that (3.3) implies (3.2), let $\varepsilon = 1 - s$ and choose the specific sequence $(\rho_\varepsilon)_{\varepsilon>0}$ defined by

$$\rho_\varepsilon(x) = \frac{\varepsilon}{|x|^{1-\varepsilon}} \mathbb{1}_{[0,1]}(|x|).$$

An important role in the proof of (3.3) plays the limit behaviour of the integral over the set $\{(x, y) \in I \times I : x \geq y\}$.

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

$$\int_0^1 \int_y^1 \frac{|u(x) - u(y)|}{|x-y|} \rho_\varepsilon(x-y)\ dx\ dy \longrightarrow \int_0^1 |u'(y)| dy \quad \text{as } \varepsilon \to 0. \tag{3.4}$$

Proof. Step 1: We show that for any $u \in W^{1,1}(I)$ and any nonnegative $\rho \in L^1(\mathbb{R})$:

$$\int_0^1 \int_0^1 \frac{|u(x) - u(y)|}{|x-y|} \rho(x-y)\ dx\ dy \lesssim |u|_{W^{1,1}(I)} \|\rho\|_{L^1(\mathbb{R})}. \tag{3.5}$$

By the Sobolev extension theorems, there exists an extension $\bar{u} \in W^{1,1}(\mathbb{R})$ such that $\bar{u} = u$ in I, $\bar{u}$ has compact support, and $|\bar{u}|_{W^{1,1}(\mathbb{R})} \lesssim |u|_{W^{1,1}(I)}$. Note that for any $h \in \mathbb{R} \backslash \{0\}$ there holds

$$\frac{1}{|h|} \int_{\mathbb{R}} |\bar{u}(x+h) - \bar{u}(x)|\ dx \leq |\bar{u}|_{W^{1,1}(\mathbb{R})} \lesssim |u|_{W^{1,1}(I)}.$$

Hence, we find

$$\begin{aligned}
\int_0^1 \int_0^1 \frac{|u(x) - u(y)|}{|x-y|} \rho(x-y)\ dx\ dy &\leq \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{|u(x) - u(y)|}{|x-y|} \rho(x-y)\ dx\ dy \\
&= \int_{\mathbb{R}} \left(\frac{1}{|h|} \int_{\mathbb{R}} |\bar{u}(x+h) - \bar{u}(x)|\ dx \right) \rho(h)\ dh \\
&\lesssim |u|_{W^{1,1}(I)} \|\rho\|_{L^1(\mathbb{R})},
\end{aligned}$$

the desired estimate (3.5).
Step 2: We claim that it suffices to show (3.4) for a dense subset of $W^{1,1}(I)$, for instance $C^2(\bar{I})$.
Using the estimate (3.5), we obtain for any $u, v \in W^{1,1}(I)$

$$\begin{aligned}
&\left| \int_0^1 \int_y^1 \frac{|u(x) - u(y)|}{|x-y|} \rho_\varepsilon(x-y)\ dx\ dy - \int_0^1 \int_y^1 \frac{|v(x) - v(y)|}{|x-y|} \rho_\varepsilon(x-y)\ dx\ dy \right| \\
&\leq \int_0^1 \int_y^1 \frac{||u(x) - u(y)| - |v(x) - v(y)||}{|x-y|} \rho_\varepsilon(x-y)\ dx\ dy \\
&\leq \int_0^1 \int_y^1 \frac{|(u-v)(x) - (u-v)(y)|}{|x-y|} \rho_\varepsilon(x-y)\ dx\ dy \\
&\lesssim |u-v|_{W^{1,1}(I)},
\end{aligned}$$

which implies the claim.
Step 3: We show (3.4) for $u \in C^2(\bar{I})$. Let $y \in I$ be arbitrary, but fixed.
By Taylor's theorem we have

$$\left| \frac{|u(x)-u(y)|}{|x-y|} - |u'(y)| \right| \lesssim |x-y| \quad \forall x \in (y, y+\delta)$$

for some $0 < \delta \leq 1 - y$. Hence, we find

$$\begin{aligned} \left| \int_y^{y+\delta} \left(\frac{|u(x)-u(y)|}{|x-y|} - |u'(y)| \right) \rho_\varepsilon(x-y)\, dx \right| &\lesssim \int_y^{y+\delta} (x-y)\rho_\varepsilon(x-y)\, dx \\ &\lesssim \int_0^\delta h\rho_\varepsilon(h)\, dh \\ &\underset{\varepsilon\to 0}{\longrightarrow} 0, \end{aligned}$$

which implies by the properties of the radial mollifiers:

$$\begin{aligned} \lim_{\varepsilon\to 0} \int_y^1 \frac{|u(x)-u(y)|}{|x-y|} \rho_\varepsilon(x-y)\, dx &= \lim_{\varepsilon\to 0} \int_y^{y+\delta} \frac{|u(x)-u(y)|}{|x-y|} \rho_\varepsilon(x-y)\, dx \\ &= |u'(y)| \lim_{\varepsilon\to 0} \int_y^{y+\delta} \rho_\varepsilon(x-y)\, dx \\ &= |u'(y)| \lim_{\varepsilon\to 0} \int_0^\delta \rho_\varepsilon(h)\, dh \\ &= |u'(y)|. \end{aligned}$$

Since y was arbitrary, we have shown the pointwise convergence

$$\int_y^1 \frac{|u(x)-u(y)|}{|x-y|} \rho_\varepsilon(x-y)\, dx \underset{\varepsilon\to 0}{\longrightarrow} |u'(y)| \quad \forall y \in I. \tag{3.6}$$

We note that $u \in C^2(\bar{I})$ is certainly Lipschitz. Therefore the convergence in $L^1(I)$ follows from (3.6) and

$$\int_y^1 \frac{|u(x)-u(y)|}{|x-y|} \rho_\varepsilon(x-y)\, dx \leq \mathrm{Lip}(u)$$

by the dominated convergence theorem. It follows (3.4), which is nothing else than the convergence of the $L^1(I)$-norms. $\square$

Remark 5. • *Using the same procedure, we see that any $\varphi \in C_c^1(I)$ satisfies*

$$\lim_{\varepsilon \to 0} \int_y^1 \frac{\varphi(x) - \varphi(y)}{x - y} \rho_\varepsilon(x - y)\ dx = \varphi'(y) \quad \forall y \in I, \tag{3.7}$$

which will be crucial in proving (3.3).

• *A similar computation as in "Step 1" for $u \in BV(I)$ – using the estimate $\frac{1}{|h|} \int |u(x+h) - u(x)| dx \leq TV(u)$ – yields :*

$$\limsup_{\varepsilon \to 0} \int_0^1 \int_0^1 \frac{|u(x) - u(y)|}{|x - y|} \rho_\varepsilon(x - y)\ dx\ dy \leq TV(u). \tag{3.8}$$

Now we can prove theorem 9.

Proof of Theorem 9. As we have already seen, it suffices to show (3.3).
Step 1: We claim that for any $u \in L^1(\mathbb{R})$, $\varphi \in C_c^1(\mathbb{R})$ and any nonnegative radial function $\rho \in L^1(\mathbb{R})$:

$$\left| \int_{\mathbb{R}} \int_y^\infty u(y) \frac{\varphi(x) - \varphi(y)}{x - y} \rho(x - y) dx dy \right| \leq \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{|u(x) - u(y)|}{|x - y|} |\varphi(x)| \rho(x - y) dx dy. \tag{3.9}$$

Without loss of generality we can assume that $\rho \equiv 0$ in some ball around the origin. (Otherwise replace ρ by $\rho_n = \mathbb{1}_{\{|x| > \frac{1}{n}\}} \rho$ and take $n \to \infty$)
Observe that since ρ is radial, there holds for all $x \in \mathbb{R}$:

$$\int_x^\infty \frac{\rho(x - y)}{y - x}\ dy = \int_{-\infty}^x \frac{\rho(x - y)}{x - y}\ dy. \tag{3.10}$$

Therefore, by renaming integration variables, (3.10) and Fubini's theorem, we find

$$\begin{aligned}
\int_{\mathbb{R}} \int_y^\infty u(y)\varphi(y) \frac{\rho(x - y)}{x - y}\ dx\ dy &= \int_{\mathbb{R}} \int_x^\infty u(x)\varphi(x) \frac{\rho(x - y)}{y - x}\ dy\ dx \\
&= \int_{\mathbb{R}} \int_{-\infty}^x u(x)\varphi(x) \frac{\rho(x - y)}{x - y}\ dy\ dx \\
&= \int_{\mathbb{R}} \int_y^\infty u(x)\varphi(x) \frac{\rho(x - y)}{x - y}\ dx\ dy.
\end{aligned}$$

Hence, we have

$$\begin{aligned}\left|\int_{\mathbb{R}}\int_{y}^{\infty} u(y)\frac{\varphi(x)-\varphi(y)}{x-y}\rho(x-y)dxdy\right| &= \left|\int_{\mathbb{R}}\int_{y}^{\infty} \frac{u(y)-u(x)}{x-y}\varphi(x)\rho(x-y)dxdy\right| \\ &\leq \int_{\mathbb{R}}\int_{y}^{\infty}\left|\frac{u(y)-u(x)}{x-y}\varphi(x)\rho(x-y)\right| dxdy \\ &\leq \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{|u(x)-u(y)|}{|x-y|}|\varphi(x)|\rho(x-y)dxdy.\end{aligned}$$

<u>Step 2:</u> Let now $u \in BV(I)$ and let $\varphi \in C_c^1(I)$ with $\|\varphi\|_\infty \leq 1$.
By extending u and φ with zero outside of I, we are in the situation of "Step 1". We write $K := \text{supp}(\varphi)$ and $r := \text{dist}(I^c, K)$. By using (3.9) we then obtain

$$\begin{aligned}\left|\int_0^1\int_{y}^{\infty} u(y)\frac{\varphi(x)-\varphi(y)}{x-y}\rho_\varepsilon(x-y)dxdy\right| &\leq \int_{\mathbb{R}}\int_{K} \frac{|u(x)-u(y)|}{|x-y|}|\varphi(x)|\rho_\varepsilon(x-y)dxdy \\ \leq \int_0^1\int_K \frac{|u(x)-u(y)|}{|x-y|}\rho_\varepsilon(x-y)dxdy &+ \int_{I^c}\int_K \frac{|u(x)-u(y)|}{|x-y|}\rho_\varepsilon(x-y)dxdy \\ \leq \int_0^1\int_0^1 \frac{|u(x)-u(y)|}{|x-y|}\rho_\varepsilon(x-y)dxdy &+ \frac{\|u\|_{L^1(I)}}{r}\int_r^\infty \rho_\varepsilon(h)\, dh \\ \leq \int_0^1\int_0^1 \frac{|u(x)-u(y)|}{|x-y|}\rho_\varepsilon(x-y)dxdy &+ o(1). \qquad (\varepsilon \to 0)\end{aligned}$$

We note that for fixed $y \in I$, using (3.7), there holds

$$\begin{aligned}\int_y^\infty u(y)\frac{\varphi(x)-\varphi(y)}{x-y}\rho_\varepsilon(x-y)dx &= u(y)\int_y^1 \frac{\varphi(x)-\varphi(y)}{x-y}\rho_\varepsilon(x-y)dx + o(1) \\ &\xrightarrow[\varepsilon\to 0]{} u(y)\varphi'(y).\end{aligned}$$

Since $\varphi \in C_c^1(I)$, it is certainly Lipschitz and hence for ε sufficiently small:

$$\begin{aligned}\left|\int_y^\infty u(y)\frac{\varphi(x)-\varphi(y)}{x-y}\rho_\varepsilon(x-y)dx\right| &\leq \left|u(y)\int_y^1 \frac{\varphi(x)-\varphi(y)}{x-y}\rho_\varepsilon(x-y)dx\right| + C \\ &\leq \text{Lip}(\varphi)|u(y)| + C.\end{aligned}$$

Since $u \in BV(I)$, the right-hand side (as a function of $y \in I$) is absolutely integrable over I and we can apply the dominated convergence theorem, which yields

$$\lim_{\varepsilon \to 0} \int_0^1 \int_y^\infty u(y) \frac{\varphi(x) - \varphi(y)}{x - y} \rho_\varepsilon(x - y)\ dx\ dy = \int_0^1 u(y)\varphi'(y)\ dy.$$

By taking the limit inferior in the previous calculation we find

$$\left| \int_0^1 u(y)\varphi'(y) dy \right| \le \liminf_{\varepsilon \to 0} \int_0^1 \int_0^1 \frac{|u(x) - u(y)|}{|x - y|} \rho_\varepsilon(x - y)\ dx\ dy. \tag{3.11}$$

The relation (3.11) holds for all $\varphi \in C_c^1(I)$ with $\|\varphi\|_\infty \le 1$, hence

$$TV(u) \le \liminf_{\varepsilon \to 0} \int_0^1 \int_0^1 \frac{|u(x) - u(y)|}{|x - y|} \rho_\varepsilon(x - y)\ dx\ dy.$$

Together with (3.8), we obtain

$$\lim_{\varepsilon \to 0} \int_0^1 \int_0^1 \frac{|u(x) - u(y)|}{|x - y|} \rho_\varepsilon(x - y)\ dx\ dy = TV(u),$$

which is exactly (3.3), what we wanted to show. □

Instead of working with some fixed u, one could also consider the case of a sequence. Bourgain, Brezis and Mironescu showed a strong compactness result by using a variant of the Riesz-Fréchet-Kolmogorov theorem, see [BBM02, Theorem 4]. For our purposes, the following special case is sufficient for this essay.

Theorem 10 (Case of a sequence, Compactness)**.** *Let $(s_n)_n \subset (0, 1)$ be a sequence with $s_n \nearrow 1$ as $n \to \infty$. Further, let $(u_n)_n \subset L^1(I)$ satisfy*

(i) $(u_n)_I = 0$ *for all* $n \in \mathbb{N}$, *and*

(ii) $(1 - s_n)|u_n|_{W^{s_n,1}(I)} \le C$ *uniformly in* n.

Then $(u_n)_n$ is relatively compact in $L^1(I)$, and there exists a function $u \in BV(I)$ such that

$$u_n \underset{n \to \infty}{\longrightarrow} u \quad \text{in } L^1(I)$$

up to a subsequence.

Proof. We refer to [BBM02, Theorem 4]. □

4 Fractional $ICTV$ Spaces

4.1 Definition of the space

Firstly, let us recall the definition of the $ICTV_\alpha^{k+1}$ seminorm for $k \in \mathbb{N}$.
Throughout this section we will make use of the notation $|u'|_{\mathcal{M}_b(I)} \equiv TV(u)$. This becomes particularly handy when for instance some $v \in L^1(I)$ is the distributional derivative of a function $\tilde{v} \in BV(I)$ such that we can write $|v|_{\mathcal{M}_b(I)} \equiv TV(\tilde{v})$.

Definition 8 (Integer-order $ICTV$ seminorm)**.** *Let $u \in BV(I)$. Then we define for $k \in \mathbb{N}$ and $\alpha = (\alpha_0, \dots, \alpha_k) \in \mathbb{R}_+^{k+1}$ the integer ICTV seminorm/ ICTV regulariser of order $k+1$ and weight α on I by*

$$|u|_{ICTV_\alpha^{k+1}(I)} := \inf_{\substack{v_l \in BV(I) \\ \forall 0 \le l \le k-1}} \left\{ \alpha_0 |u' - v_0|_{\mathcal{M}_b(I)} + \sum_{i=0}^{k-2} \alpha_{i+1} |v_i' - v_{i+1}|_{\mathcal{M}_b(I)} + \alpha_k |v_{k-1}'|_{\mathcal{M}_b(I)} \right\}.$$

For $k = 1$ the above definition reads

$$|u|_{ICTV_\alpha^2(I)} := \inf_{v_0 \in BV(I)} \left\{ \alpha_0 |u' - v_0|_{\mathcal{M}_b(I)} + \alpha_1 |v_0'|_{\mathcal{M}_b(I)} \right\}. \tag{4.1}$$

Note that in case of the $ICTV_\alpha^2$ regularisation, the choice of a very large α_1 in (4.1) will result in a TV-similar regularisation.
We will now define the fractional-order $ICTV$ space via defining a suitable seminorm. We want to define the seminorm in such a way that it is intermediate between the surrounding integer-order $ICTV$ seminorms. As we shall see, the following definition works.

Definition 9 (Fractional-order $ICTV$ seminorm)**.** *Let $u \in L^1(I)$ and let $s \in (0,1)$.*

(i) *For $k = 1$ and $\alpha = (\alpha_0, \alpha_1) \in \mathbb{R}_+^2$, the fractional ICTV seminorm/ ICTV regulariser of order $1+s$ and weight α on I is defined by*

$$|u|_{ICTV_\alpha^{1+s}(I)} := \inf_{\substack{v_0 \in W^{s,1+s(1-s)}(I), \\ (v_0)_I = 0}} \left\{ \alpha_0 |u' - s v_0|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) |v_0|_{W^{s,1+s(1-s)}(I)} \right\}.$$

(ii) *For $k \in \mathbb{N}_{>1}$ and $\alpha = (\alpha_0, \dots, \alpha_k) \in \mathbb{R}_+^{k+1}$, the fractional ICTV seminorm/ ICTV regulariser of order $k+s$ and weight α on I is defined by*

$$|u|_{ICTV_\alpha^{k+s}(I)} :=$$
$$:= \inf_{\substack{v_0, \dots, v_{k-2} \in BV(I), \\ v_{k-1} \in W^{s,1+s(1-s)}(I), \\ (v_{k-1})_I = 0}} \{ \alpha_0 |u' - v_0|_{\mathcal{M}_b(I)} + \sum_{i=0}^{k-3} \alpha_{i+1} |v_i' - v_{i+1}|_{\mathcal{M}_b(I)} +$$
$$+ \alpha_{k-1} |v_{k-2}' - s v_{k-1}|_{\mathcal{M}_b(I)} + \alpha_k s(1-s) |v_{k-1}|_{W^{s,1+s(1-s)}(I)} \}.$$

Definition 10 (Fractional $ICTV$ spaces). *Let $k \in \mathbb{N}$, $s \in (0,1)$ and $\alpha \in \mathbb{R}_+^{k+1}$. We denote by $BCV_\alpha^{k+s}(I)$ the space of functions with finite fractional $ICTV_\alpha^{k+s}(I)$ seminorm,*

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

and we define the norm on this space to be

$$\|u\|_{BCV_\alpha^{k+s}(I)} := \|u\|_{L^1(I)} + |u|_{ICTV_\alpha^{k+s}(I)}.$$

Moreover, we define

$$BCV^{k+s}(I) := \{u \in L^1(I) \mid \exists\, \alpha \in \mathbb{R}_+^{k+1} : u \in BCV_\alpha^{k+s}(I)\}.$$

4.2 Limit behaviour of the fractional $ICTV$ seminorm

The first thing we have to check is whether our definition of the fractional $ICTV$ seminorm is compatible with the definition of the integer $ICTV$ seminorm. For simplicity we will just consider the case $k = 1$. We will state a general result in the end which follows by simple adaptations. Our goal is to prove the following theorem:

Theorem 11 (Limit behaviour of $ICTV$ seminorm). *Let $u \in BV(I)$. Then there holds, up to a subsequence:*

(i) $\lim_{s\to 0} |u|_{ICTV_\alpha^{1+s}(I)} = \alpha_0 |u'|_{\mathcal{M}_b(I)}$

(ii) $\liminf_{s\to 1} |u|_{ICTV_\alpha^{1+s}(I)} \geq |u|_{ICTV_\alpha^2(I)}$.

As a first step we need to investigate the limit behaviour of the $W^{s,1+s(1-s)}$ seminorm as $s \to 1$. Fortunately, a part of the work of this was already done in section 3.2 in deriving the limit behaviour of the $W^{s,1}$ seminorm. We will use theorem 9 to show:

Lemma 2 (Limit behaviour of $W^{s,1+s(1-s)}$ seminorm). *Let $u \in BV(I) \cap C^\infty(I)$. Then there holds*

$$\limsup_{s\to 1}\, (1-s)|u|_{W^{s,1+s(1-s)}(I)} \leq |u'|_{\mathcal{M}_b(I)}. \tag{4.2}$$

Proof. Since $u \in BV(I) \cap C^\infty(I)$ and $s \in (0,1)$, there holds the pointwise estimate

$$|u(x) - u(y)| \leq |u'|_{\mathcal{M}_b(I)}|x-y| \leq |u'|_{\mathcal{M}_b(I)}|x-y|^s$$

for all $x, y \in I$. Hence, we have

$$\begin{aligned} |u|^{1+s(1-s)}_{W^{s,1+s(1-s)}(I)} &= \int_0^1\int_0^1 \frac{|u(x)-u(y)|^{1+s(1-s)}}{|x-y|^{1+s+s^2(1-s)}}\,dx\,dy \\ &\leq |u'|^{s(1-s)}_{\mathcal{M}_b(I)} \int_0^1 \frac{|u(x)-u(y)|}{|x-y|^{1+s}}\,dx\,dy \\ &= |u'|^{s(1-s)}_{\mathcal{M}_b(I)}\,|u|_{W^{s,1}(I)}, \end{aligned}$$

which implies the estimate

$$|u|_{W^{s,1+s(1-s)}(I)} \leq |u'|_{\mathcal{M}_b(I)}^{\frac{s(1-s)}{1+s(1-s)}} \, |u|_{W^{s,1}(I)}^{\frac{1}{1+s(1-s)}}. \tag{4.3}$$

Without loss of generality, we may assume that $|u'|_{\mathcal{M}_b(I)} \neq 0$ (otherwise (4.2) is trivial). Then there holds

$$|u'|_{\mathcal{M}_b(I)}^{\frac{s(1-s)}{1+s(1-s)}} \underset{s\to 1}{\longrightarrow} 1$$

and we deduce from (4.3) and theorem 9 that

$$\begin{aligned}
\limsup_{s\to 1}\,(1-s)|u|_{W^{s,1+s(1-s)}(I)} &\leq \limsup_{s\to 1}\,(1-s)|u|_{W^{s,1}(I)}^{\frac{1}{1+s(1-s)}} \\
&= \limsup_{s\to 1}\left[\Big((1-s)|u|_{W^{s,1}(I)}\Big)^{\frac{1}{1+s(1-s)}} \cdot (1-s)^{\frac{s(1-s)}{1+s(1-s)}}\right] \\
&= \limsup_{s\to 1}\left[e^{\frac{1}{1+s(1-s)}\log\left((1-s)|u|_{W^{s,1}(I)}\right)} \cdot e^{\frac{s(1-s)}{1+s(1-s)}\log(1-s)}\right] \\
&= |u'|_{\mathcal{M}_b(I)} \cdot 1 = |u'|_{\mathcal{M}_b(I)}.
\end{aligned}$$

□

Besides the asymptotic behaviour of the $W^{s,1+s(1-s)}$ seminorm, we need some compactness and lower semicontinuity result for $W^{s,1+s(1-s)}$ functions with mean zero. This becomes very important once we consider minimal sequences for the infimum in the definition of the fractional $ICTV$ seminorm.

Lemma 3 (Compactness and Lower Semicontinuity). *Let $(s_n)_n \subset (0,1)$ be a sequence with $s_n \to \bar{s}$ for some $0 < \bar{s} \leq 1$. Further, let $(v_n)_n \subset W^{s_n,1+s_n(1-s_n)}(I)$ satisfy*

(1) $(v_n)_I = 0$ *for all* $n \in \mathbb{N}$, *and*

(2) $\sup_{n\in\mathbb{N}} s_n(1-s_n)|v_n|_{W^{s_n,1+s_n(1-s_n)}(I)} < \infty.$

Then there holds the following:

(*i*) *If $\bar{s} < 1$, there exists a function $v \in W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)$ such that, up to a subsequence,*

$$v_n \underset{n\to\infty}{\longrightarrow} v \text{ in } L^1(I)$$

and

$$\bar{s}(1-\bar{s})|v|_{W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)} \leq \liminf_{n\to\infty} s_n(1-s_n)|v_n|_{W^{s_n,1+s_n(1-s_n)}(I)}. \tag{4.4}$$

(ii) *If* $\bar{s} = 1$*, there exists a function* $v \in BV(I)$ *such that, up to a subsequence,*

$$v_n \underset{n\to\infty}{\longrightarrow} v \text{ in } L^1(I)$$

and

$$|v'|_{\mathcal{M}_b(I)} \leq \liminf_{n\to\infty} s_n(1-s_n)|v_n|_{W^{s_n,1+s_n(1-s_n)}(I)}. \tag{4.5}$$

Proof. For $n \in \mathbb{N}$, we write $p_n := 1+s_n(1-s_n)$ and we set $\bar{p} := 1+\bar{s}(1-\bar{s})$. Without loss of generality – up to a subsequence – we may assume that the sequences $(s_n)_n$ and $(p_n)_n$ converge monotonically. We obtain the following five cases:

- *Case 1:* $\bar{s} \in [\frac{1}{2}, 1)$, $s_n \searrow \bar{s}$ and $p_n \nearrow \bar{p}$,
- *Case 2:* $\bar{s} \in (0, \frac{1}{2})$, $s_n \searrow \bar{s}$ and $p_n \searrow \bar{p}$,
- *Case 3:* $\bar{s} \in (0, \frac{1}{2}]$, $s_n \nearrow \bar{s}$ and $p_n \nearrow \bar{p}$,
- *Case 4:* $\bar{s} \in (\frac{1}{2}, 1)$, $s_n \nearrow \bar{s}$ and $p_n \searrow \bar{p}$,
- *Case 5:* $\bar{s} = 1$, $s_n \nearrow 1$ and $p_n \searrow 1$.

Case 1: In this case, there holds for all n that $\bar{s} \leq s_n$ and $\bar{p} \geq p_n$. Since the function

$$f : (0,1) \to \mathbb{R}, \quad f(x) := x - \frac{1}{1+x(1-x)}$$

is strictly increasing, we also have

$$s_n - \frac{1}{p_n} = f(s_n) \geq f(\bar{s}) = \bar{s} - \frac{1}{\bar{p}}.$$

Hence, we can apply theorem 6 and obtain

$$|v_n|_{W^{\bar{s},\bar{p}}(I)} \leq \frac{36}{s_n\bar{s}}|v_n|_{W^{s_n,p_n}(I)} \leq 144|v_n|_{W^{s_n,p_n}(I)} \leq C \tag{4.6}$$

for all n, where we used $s_n \geq \bar{s} \geq \frac{1}{2}$ and assumption (2) in the last two steps. Note that the constant on the right-hand side does not depend on n, i.e. we have an uniform bound on the $W^{\bar{s},\bar{p}}$ seminorm. In order to use the reflexivity of $W^{\bar{s},\bar{p}}$ (theorem 7), we need a uniform bound on the full norm. This is done via the Poincaré inequality (theorem 8):
We have $\bar{s}\bar{p} < 1$, hence by theorem 8 and (4.6) there holds

$$\|v_n\|_{L^{\frac{\bar{p}}{1-\bar{s}\bar{p}}}(I)} \lesssim |v_n|_{W^{\bar{s},\bar{p}}(I)} \leq C, \tag{4.7}$$

where we used $(v_n)_I = 0$. Since $\frac{\bar{p}}{1-\bar{s}\bar{p}} \geq \bar{p}$ and I is bounded, (4.7) implies in particular a control of the $L^{\bar{p}}$-norm and hence

$$\|v_n\|_{W^{\bar{s},\bar{p}}(I)} = \|v_n\|_{L^{\bar{p}}(I)} + |v_n|_{W^{\bar{s},\bar{p}}(I)} \lesssim \|v_n\|_{L^{\frac{\bar{p}}{1-\bar{s}\bar{p}}}(I)} + |v_n|_{W^{\bar{s},\bar{p}}(I)} \lesssim |v_n|_{W^{\bar{s},\bar{p}}(I)} \leq C.$$

By theorem 7, $W^{\bar{s},\bar{p}}(I)$ is reflexive. Thus, there exists a function $v \in W^{\bar{s},\bar{p}}(I)$ such that, up to a subsequence,

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

Since $\bar{s}\bar{p} < 1$ and $\bar{p}^* \geq \bar{p} > 1$, the embedding $W^{\bar{s},\bar{p}}(I) \hookrightarrow L^1(I)$ is compact by theorem 5. Hence, up to a subsequence,

$$v_n \longrightarrow v \quad \text{strongly in } L^1(I). \tag{4.8}$$

The first part of the claim is proven and it remains to show (4.4). To this end, we note that (4.8) implies, up to a subsequence, convergence almost everywhere such that we find by Fatou's lemma

$$|v|^{\bar{p}}_{W^{\bar{s},\bar{p}}(I)} \leq \liminf_{n\to\infty} |v_n|^{p_n}_{W^{s_n,p_n}(I)},$$

which implies the desired estimate (4.4).

<u>*Case 2:*</u> Our goal is again to apply theorem 6 in order to get boundedness of our sequence $(v_n)_n$ in some fractional Sobolev space $W^{s,p}$. We show that there exists some $s \leq \bar{s}$ and some $p \geq \bar{p}$ such that, for n sufficiently large, there holds

$$s \leq s_n, \; p \geq p_n \tag{4.9}$$

and

$$s_n - \frac{1}{p_n} \geq s - \frac{1}{p}. \tag{4.10}$$

We simply set $s := \bar{s} - \varepsilon_s$ and $p := \bar{p} + \varepsilon_p$ with some $\varepsilon_s \in (0, \frac{\bar{s}}{2})$ and $\varepsilon_p > 0$ to be determined later. We observe that the condition (4.9) is satisfied, for n sufficiently large, for any choice of ε_s and ε_p. Note that for all n, we have

$$s_n - \frac{1}{p_n} \geq \bar{s} - \frac{1}{p_n} \geq \bar{s} - \frac{1}{\bar{p}} = s + \varepsilon_s - \frac{1}{p - \varepsilon_p},$$

hence

$$\left(s_n - \frac{1}{p_n}\right) - \left(s - \frac{1}{p}\right) \geq \varepsilon_s + \frac{1}{p} - \frac{1}{p - \varepsilon_p} = \varepsilon_s - \frac{\varepsilon_p}{p(p - \varepsilon_p)} = \varepsilon_s - \frac{\varepsilon_p}{(\bar{p} + \varepsilon_p)\bar{p}}.$$

For any choice of $\varepsilon_s \in (0, \frac{\bar{s}}{2})$ we can find some $\varepsilon_p > 0$ such that the right-hand side is nonnegative, i.e. such that (4.10) holds.
Now we can conclude exactly as in the first case

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

and hence by Poincaré, reflexivity of $W^{s,p}(I)$ and compactness of the embedding $W^{s,p}(I) \hookrightarrow L^1(I)$, there exists a function $v_{s,p} \in W^{s,p}(I)$ such that, up to subsequences,

$$\begin{aligned} v_n &\rightharpoonup v_{s,p} \quad \text{weakly in } W^{s,p}(I), \text{ and} \\ v_n &\to v_{s,p} \quad \text{strongly in } L^1(I). \end{aligned}$$

If we now let $\varepsilon_s, \varepsilon_p \to 0$ and use a diagonal argument, we obtain the desired $v \in W^{\bar{s},\bar{p}}(I)$. As in the first case, by using Fatou's lemma, we are then able to deduce the estimate (4.4).

Case 3: We show that there exists some $s \leq \bar{s}$ such that, for n sufficiently large, there holds $s \leq s_n$ and

$$s_n - \frac{1}{p_n} \geq s - \frac{1}{\bar{p}}.$$

We set $s := s_1 - \varepsilon$ for some $\varepsilon \in (0, \frac{s_1}{2})$ fixed. Then clearly $s \leq s_1 \leq s_n \leq \bar{s}$. Since $p_n \to \bar{p}$, we have for n sufficiently large

$$\frac{1}{p_n} - \frac{1}{\bar{p}} \leq \varepsilon,$$

hence

$$s_n - \frac{1}{p_n} \geq s_1 - \frac{1}{p_n} \geq s_1 - \left(\frac{1}{\bar{p}} + \varepsilon\right) = s - \frac{1}{\bar{p}}.$$

Thus, we can apply theorem 6 and analogously to the previous cases, we find a function $v \in W^{s,\bar{p}}(I)$ such that, up to subsequences,

$$\begin{aligned} v_n &\rightharpoonup v \quad \text{weakly in } W^{s,p}(I), \text{ and} \\ v_n &\to v \quad \text{strongly in } L^1(I). \end{aligned}$$

By using Fatou's lemma, we obtain (4.4), which also shows that $v \in W^{\bar{s},\bar{p}}(I)$ since the right hand side of (4.4) is finite by assumption (2).

Case 4: In this case, there holds for all n that $s_1 \leq s_n$ and $p_1 \geq p_n$. Further, we observe

$$s_n - \frac{1}{p_n} = f(s_n) \geq f(s_1) = s_1 - \frac{1}{p_1},$$

where we used that f is strictly increasing on $(0, 1)$. So we can apply theorem 6 and proceed as previously.

Case 5: By assumption (2), there holds

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

uniformly in n. Observe, that we are in the situation of theorem 10, which shows the first part of the claim. Regarding the estimate (4.5), we refer to [BBM02]. □

Now we will use both the asymptotic behaviour of the $W^{s,1+s(1-s)}$ norm from lemma 2, and the compactness and lower semicontinuity result from lemma 3, in order to prove theorem 11.

Proof of Theorem 11. (i) We note that by definition of $|u|_{ICTV_\alpha^{1+s}(I)}$: (take $v_0 = 0$)

$$|u|_{ICTV_\alpha^{1+s}(I)} \leq \alpha_0 |u'|_{\mathcal{M}_b(I)} \tag{4.11}$$

for all $s \in (0,1)$. Hence, there holds

$$\limsup_{s\to 0} |u|_{ICTV_\alpha^{1+s}(I)} \leq \alpha_0 |u'|_{\mathcal{M}_b(I)}. \tag{4.12}$$

By definition of the fractional $ICTV$ seminorm, we find for each $0 < s < 1$ some function $v_s \in W^{s,1+s(1-s)}(I)$ with $(v_s)_I = 0$ and

$$\alpha_0 |u' - s v_s|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) |v_s|_{W^{s,1+s(1-s)}(I)} \leq |u|_{ICTV_\alpha^{1+s}(I)} + s. \tag{4.13}$$

We write $p := 1 + s(1-s)$ and observe that $sp < 1$. Thus, we can use theorem 8, the Poincaré inequality, to obtain

$$\|v_s\|_{L^1(I)} \lesssim \|v_s\|_{L^{\frac{p}{1-sp}}(I)} \lesssim \left(\frac{s(1-s)}{(1-sp)^{p-1}}\right)^{\frac{1}{p}} |v_s|_{W^{s,p}(I)}. \tag{4.14}$$

By (4.13) and (4.11), there holds the uniform estimate (in s)

$$s(1-s)|v_s|_{W^{s,p}(I)} \leq C$$

for some $C > 0$, such that we deduce from (4.14),

$$\|s v_s\|_{L^1(I)} \lesssim \frac{1}{1-s} \left(\frac{s(1-s)}{(1-sp)^{p-1}}\right)^{\frac{1}{p}} \underset{s\to 0}{\longrightarrow} 0,$$

i.e.

$$s v_s \underset{s\to 0}{\longrightarrow} 0 \ \text{ in } L^1(I).$$

We note that we have by (4.13),

$$\alpha_0 |u' - s v_s|_{\mathcal{M}_b(I)} \leq |u|_{ICTV_\alpha^{1+s}(I)} + s,$$

such that taking the limit inferior yields

$$\alpha_0 |u'|_{\mathcal{M}_b(I)} \leq \liminf_{s\to 0} |u|_{ICTV_\alpha^{1+s}(I)}. \tag{4.15}$$

Combining (4.12) and (4.15) results in

$$\limsup_{s\to 0} |u|_{ICTV_\alpha^{1+s}(I)} \leq \alpha_0 |u'|_{\mathcal{M}_b(I)} \leq \liminf_{s\to 0} |u|_{ICTV_\alpha^{1+s}(I)},$$

which implies

$$\lim_{s\to 0} |u|_{ICTV_\alpha^{1+s}(I)} = \alpha_0 |u'|_{\mathcal{M}_b(I)}.$$

(ii) Let $v \in BV(I)$ be an extremal function for $|u|_{ICTV_\alpha^2(I)}$, i.e. such that

$$\alpha_0 |u' - v|_{\mathcal{M}_b(I)} + \alpha_1 |v'|_{\mathcal{M}_b(I)} = |u|_{ICTV_\alpha^2(I)}.$$

By the approximation theorem (theorem 3), there exists a sequence of smooth functions $(v_n) \subset BV(I) \cap C^\infty(I)$ such that

$$\begin{aligned} v_n &\longrightarrow v \quad \text{in } L^1 \text{ , and} \\ |v_n'|_{\mathcal{M}_b(I)} &\longrightarrow |v'|_{\mathcal{M}_b(I)}. \end{aligned}$$

By definition of the fractional $ICTV$ seminorm, we have

$$\begin{aligned} |u|_{ICTV_\alpha^{1+s}(I)} &\leq \alpha_0 |u' - s\,(v_n - (v_n)_I)|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) |v_n - (v_n)_I|_{W^{s,1+s(1-s)}(I)} \\ &= \alpha_0 |u' - s v_n + s(v_n)_I|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) |v_n|_{W^{s,1+s(1-s)}(I)}. \end{aligned}$$

Using lemma 2, we obtain

$$\limsup_{s\to 1} |u|_{ICTV_\alpha^{1+s}(I)} \leq \alpha_0 |u' - v_n|_{\mathcal{M}_b(I)} + \alpha_0 |(v_n)_I| + \alpha_1 |v_n'|_{\mathcal{M}_b(I)}$$

and passing to the limit $n \to \infty$, we find

$$\begin{aligned} \limsup_{s\to 1} |u|_{ICTV_\alpha^{1+s}(I)} &\leq \alpha_0 |u' - v|_{\mathcal{M}_b(I)} + \alpha_0 |(v)_I| + \alpha_1 |v'|_{\mathcal{M}_b(I)} \\ &= |u|_{ICTV_\alpha^2(I)} + \alpha_0 |(v)_I|. \end{aligned} \tag{4.16}$$

By definition of the fractional $ICTV$ seminorm, we find for each $0 < s < 1$ some function $v_s \in W^{s,1+s(1-s)}(I)$ with $(v_s)_I = 0$ and

$$\alpha_0 |u' - s v_s|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) |v_s|_{W^{s,1+s(1-s)}(I)} \leq |u|_{ICTV_\alpha^{1+s}(I)} + (1-s). \tag{4.17}$$

With regard to (4.16), the conditions (1) and (2) of lemma 3 are satisfied. By lemma 3, we find a function $v \in BV(I)$ such that, up to a subsequence,

$$v_s \underset{s\to 1}{\longrightarrow} v \ \text{ in } L^1(I),$$

and

$$|v'|_{\mathcal{M}_b(I)} \leq \liminf_{s\to 1} s(1-s) |v_s|_{W^{s,1+s(1-s)}(I)}.$$

Taking the limit inferior in (4.17), we obtain

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

which implies

$$|u|_{ICTV_\alpha^2(I)} \leq \liminf_{s\to 1} |u|_{ICTV_\alpha^{1+s}(I)}$$

by definition of the $ICTV_\alpha^2$ seminorm. □

By adaptations of the arguments used in the proof, we obtain for $k \in \mathbb{N}_{>1}$ the following result:

Remark 6. *Let $k \in \mathbb{N}_{>1}$ and let $u \in BV(I)$. Then there holds, up to a subsequence:*

(i) $\lim_{s \to 0} |u|_{ICTV^{k+s}_{(\alpha_0,\dots,\alpha_k)}(I)} = |u|_{ICTV^{k}_{(\alpha_0,\dots,\alpha_{k-1})}(I)}$

(ii) $\liminf_{s \to 1} |u|_{ICTV^{k+s}_{\alpha}(I)} \geq |u|_{ICTV^{k+1}_{\alpha}(I)}$.

4.3 Extremal functions and seminorm equivalences

4.3.1 Extremal functions for the fractional $ICTV$ seminorm

In the first part of this subsection, we deal with the existence of extremal functions for the fractional $ICTV$ seminorms from definition 9. For simplicity, we just consider the case $k = 1$. The general case follows then by simple adaptations.
In the case $k = 1$, we want to solve for given $u \in BCV^{1+s}_{\alpha}(I)$ the following constrained minimisation problem:

$$(P_u) \begin{cases} \min \ \alpha_0 |u' - sv|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s)|v|_{W^{s,1+s(1-s)}(I)} \\ \text{s.t.} \ \ v \in W^{s,1+s(1-s)}(I), \ (v)_I = 0 \end{cases} .$$

Before we prove the existence of a solution to the optimisation problem (P_u), we first note that every $BCV^{1+s}_{\alpha}(I)$-function is of bounded variation.

Remark 7. *Let $u \in BCV^{1+s}_{\alpha}(I)$ for some $\alpha \in \mathbb{R}^2_+$ and $s \in (0,1)$. Then $u \in BV(I)$.*

Proof. By definition of the fractional $ICTV$ seminorm, we can find a function $v \in W^{s,1+s(1-s)}(I)$ with $(v)_I = 0$ such that

$$\alpha_0 |u' - sv|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s)|v|_{W^{s,1+s(1-s)}(I)} \leq |u|_{ICTV^{1+s}_{\alpha}(I)} + 1.$$

We note that $s(1 + s(1-s)) < 1$, hence by theorem 5

$$\|v\|_{L^1(I)} \leq C \|v\|_{W^{s,1+s(1-s)}(I)}$$

for some $C > 0$, and we obtain

$$\begin{aligned} |u'|_{\mathcal{M}_b(I)} &\leq |u' - sv|_{\mathcal{M}_b(I)} + s\|v\|_{L^1(I)} \\ &\leq |u' - sv|_{\mathcal{M}_b(I)} + Cs\|v\|_{W^{s,1+s(1-s)}(I)} \\ &= \frac{1}{\alpha_0}\left(\alpha_0 |u' - sv|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s)|v|_{W^{s,1+s(1-s)}(I)}\right) \\ &\qquad + \left(Cs - \frac{\alpha_1}{\alpha_0} s(1-s)\right)|v|_{W^{s,1+s(1-s)}(I)} + Cs\|v\|_{L^{1+s(1-s)}(I)} \\ &\leq \frac{|u|_{ICTV^{1+s}_{\alpha}(I)} + 1}{\alpha_0} + s\left(C - \frac{\alpha_1}{\alpha_0}(1-s)\right)|v|_{W^{s,1+s(1-s)}(I)} + Cs\|v\|_{L^{1+s(1-s)}(I)} \\ &< \infty. \end{aligned}$$

□

Now we are going to prove, that for any function $u \in BCV_\alpha^{1+s}(I)$, we can find an extremal function for the fractional $ICTV$ seminorm. In other words, we prove the existence of a solution to (P_u).

Theorem 12 (Existence of extremal functions)**.** *Let $u \in BCV_\alpha^{1+s}(I)$ for some $\alpha \in \mathbb{R}_+^2$ and $s \in (0,1)$. Then there exists a solution v^* to the optimisation problem (P_u), i.e. there exists $v^* \in W^{s,1+s(1-s)}(I)$ with $(v^*)_I = 0$, such that*

$$P_u^{v^*} := \alpha_0 |u' - sv^*|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s)|v^*|_{W^{s,1+s(1-s)}(I)} = |u|_{ICTV_\alpha^{1+s}(I)}.$$

Proof. By definition of the expression $|u|_{ICTV_\alpha^{1+s}(I)}$, there exists a minimising sequence $(v_n)_n \subset W^{s,1+s(1-s)}(I)$ with $(v_n)_I = 0$ for all $n \in \mathbb{N}$ and

$$\alpha_0 |u' - sv_n|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s)|v_n|_{W^{s,1+s(1-s)}(I)} \underset{n\to\infty}{\longrightarrow} |u|_{ICTV_\alpha^{1+s}(I)}. \tag{4.18}$$

In particular, since $|u|_{ICTV_\alpha^{1+s}(I)} < \infty$ and the left-hand side is convergent, we have

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

for some $c > 0$. We note that $s(1 + s(1-s)) < 1$, hence theorem 8 yields

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

where we used the fact that $(v_n)_I = 0$ for all $n \in \mathbb{N}$. By reflexivity of $W^{s,1+s(1-s)}(I)$ (theorem 7), we find a function $v^* \in W^{s,1+s(1-s)}(I)$ such that, up to a subsequence,

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

By theorem 5, the embedding $W^{s,1+s(1-s)}(I) \hookrightarrow L^1(I)$ is compact and hence, up to a subsequence,

$$v_n \longrightarrow v^* \quad \text{strongly in } L^1(I). \tag{4.19}$$

By lemma 1, the total variation is lower semicontinuous with respect to the L^1-convergence, and as we have seen earlier, the $W^{s,1+s(1-s)}$ seminorm is weakly lower semicontinuous. Thus, we conclude that

$$\begin{aligned} P_u^{v^*} &\leq \liminf_{n\to\infty} \alpha_0 |u' - sv_n|_{\mathcal{M}_b(I)} + \liminf_{n\to\infty} \alpha_1 s(1-s)|v_n|_{W^{s,1+s(1-s)}(I)} \\ &\leq \liminf_{n\to\infty} \left[\alpha_0 |u' - sv_n|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s)|v_n|_{W^{s,1+s(1-s)}(I)}\right] \\ &= |u|_{ICTV_\alpha^{1+s}(I)}, \end{aligned}$$

where we used (4.18) in the last step. We note, that there also holds

$$P_u^{v^*} \geq |u|_{ICTV_\alpha^{1+s}(I)},$$

which follows from the definition of $|u|_{ICTV_\alpha^{1+s}(I)}$ and the fact that $v^* \in W^{s,1+s(1-s)}(I)$ satisfies

$$(v^*)_I = \lim_{n\to\infty} (v_n)_I = 0$$

by (4.19) and since $(v_n)_I = 0$ for all $n \in \mathbb{N}$. $\square$

By adaptations of the arguments used in the proof, we obtain for $k \in \mathbb{N}_{>1}$ the following result:

Remark 8. *Let $u \in BCV_\alpha^{k+s}(I)$ with some $k \in \mathbb{N}_{>1}$, $\alpha \in \mathbb{R}_+^{k+1}$ and $s \in (0,1)$. Then the infimum in definition 9 is attained.*

4.3.2 Seminorm equivalences

We want to investigate the relationship between the fractional $ICTV$ seminorms and the total variation. As we shall see, they are equivalent. For simplicity, we will again consider the case $k = 1$.

Proposition 2. *Let $s \in (0,1)$. Then there holds*

$$BV(I) \sim BCV^{1+s}(I).$$

Proof. Regarding (4.11), it is enough to show that there exists an $\alpha \in \mathbb{R}_+^2$ such that

$$|u'|_{\mathcal{M}_b(I)} \lesssim |u|_{ICTV_\alpha^{1+s}(I)}. \tag{4.20}$$

Let $v \in W^{s,1+s(1-s)}(I)$ with $(v)_I = 0$ be arbitrary. We use theorem 5 and theorem 8 to deduce

$$\begin{aligned} |u'|_{\mathcal{M}_b(I)} &\leq |u' - sv|_{\mathcal{M}_b(I)} + s\|v\|_{L^1(I)} \\ &\leq |u' - sv|_{\mathcal{M}_b(I)} + Cs|v|_{W^{s,1+s(1-s)}(I)} \end{aligned}$$

for some constant $C > 0$. Hence, by arbitrariness of v, there holds (4.20) with $\alpha := (1, \frac{C}{1-s})$. □

By adaptations of the proof, we obtain for $k \in \mathbb{N}$ the following result:

Remark 9. *Let $k \in \mathbb{N}$ and let $s \in (0,1)$. Then there holds*

$$BV(I) \sim BCV^k(I) \sim BCV^{k+s}(I) \sim BCV^{k+1}(I). \tag{4.21}$$

The last equivalence relation in (4.21) can be shown by [BV11, Theorem 3.3].

5 Analysis of the Bilevel Learning Scheme

5.1 The Bilevel Learning Scheme

We want to analyse the following bilevel learning scheme, which is looking for the optimal parameter vector $\alpha \in \mathbb{R}_+^{k+1}$ and the optimal order $r \geq 1$ of the $ICTV$ regulariser at the same time:

$$(\mathcal{B}) \begin{cases} (\alpha^*, r^*) & := \underset{(\alpha,r)}{\arg\min} \left\{ \|u_{\alpha,r} - u_c\|_{L^2(I)}^2 : (\alpha, r) \in [a, A]^{\lfloor r \rfloor + 1} \times [1, R] \right\} \\ u_{\alpha,r} & := \underset{u \in BCV_\alpha^r(I)}{\arg\min} \left(\|u - u_0\|_{L^2(I)}^2 + |u|_{ICTV_\alpha^r(I)} \right) \end{cases}.$$

In this scheme, $u_0 \in BV(I)$ represents the noisy image and $u_c \in BV(I)$ a clean (free of noise) test picture. Moreover, the numbers $a, A > 0$ and $R > 1$ are fixed. We note that the integer $ICTV$ seminorm of order $r = 1$ is not defined yet, however it is natural to set

$$|u|_{ICTV_{\alpha_0}^1(I)} := \alpha_0 |u'|_{\mathcal{M}_b(I)}.$$

The bilevel learning scheme $(\mathcal{B})$ then results in an optimal reconstructed image u_{α^*,r^*}. Our aim is to show that $(\mathcal{B})$ admits a unique solution and hence our reconstruction is well-defined.
As we shall see, the restriction of the range of possible α and r to lie within the box

$$(\alpha, r) \in [a, A]^{\lfloor r \rfloor + 1} \times [1, R], \tag{5.1}$$

ensures that u_{α^*,r^*} belongs to the class $BCV_{\alpha^*}^{r^*}(I)$. Further, (5.1) is necessary in order to realise the bilevel scheme $(\mathcal{B})$ numerically. For a numerical algorithm, we refer to the appendix and [DL16], where the authors propose a first order primal-dual algorithm which relies on non-smooth convex optimisation and Besov space techniques.

5.2 Existence and Uniqueness of Solutions

The following lemma will be essential in proving the existence of a solution to $(\mathcal{B})$, since it provides a compactness and lower semicontinuity result.

Lemma 4. *Let $k \in \mathbb{N}$, $(s_n)_n \subset (0,1)$ and let $(\alpha^n)_n \subset \mathbb{R}_+^{k+1}$ be a sequence of vectors $\alpha^n = (\alpha_0^n, \dots \alpha_k^n)$ that satisfy*

$$a < \inf_{\substack{n \in \mathbb{N} \\ i \in \{0,\dots,k\}}} \alpha_i^n \leq \sup_{\substack{n \in \mathbb{N} \\ i \in \{0,\dots,k\}}} \alpha_i^n < A \tag{5.2}$$

for some $a, A > 0$. Further, let $(u_n)_n \subset L^1(I)$ with $u_n \in BCV_{\alpha^n}^{k+s}(I)$ for all $n \in \mathbb{N}$ and

$$\sup_{n \in \mathbb{N}} \|u_n\|_{BCV_{\alpha^n}^{k+s_n}(I)} < \infty. \tag{5.3}$$

Then there exists $0 \leq \bar{s} \leq 1$, $\alpha = (\alpha_0, \ldots, \alpha_k) \in \mathbb{R}^{k+1}$ *and a function* $u \in BCV_\alpha^{k+\bar{s}}(I)$ ($u \in BV(I)$ *respectively if* $k = 1$ *and* $\bar{s} = 0$) *such that there holds, up to a subsequence,* $s_n \to \bar{s}$, $\alpha_n \to \alpha$ *and*

$$u_n \overset{*}{\rightharpoonup} u \text{ in } BV(I).$$

Moreover, there holds

$$\liminf_{n\to\infty} |u_n|_{ICTV_{\alpha^n}^{k+s_n}(I)} \geq \begin{cases} |u|_{ICTV_\alpha^{k+\bar{s}}(I)} & , \text{ if } \bar{s} > 0. \\ \alpha_0 |u'|_{\mathcal{M}_b(I)} & , \text{ if } \bar{s} = 0,\ k = 1. \\ |u|_{ICTV^k_{(\alpha_0,\ldots,\alpha_{k-1})}(I)} & , \text{ if } \bar{s} = 0,\ k > 1. \end{cases}$$

Proof. (Just for $k = 1$)
By the boundedness assumption (5.2) and the Bolzano-Weierstraß theorem, there holds, up to a subsequence,

$$\alpha^n = (\alpha_0^n, \alpha_1^n) \longrightarrow (\alpha_0, \alpha_1),$$
$$s_n \longrightarrow \bar{s},$$

for some $\alpha := (\alpha_0, \alpha_1) \in (a, A) \times (a, A)$ and $\bar{s} \in [0, 1]$.
Since $u_n \in BCV_{\alpha^n}^{1+s}(I)$ for all n, by using theorem 12, we can find functions $v_n \in W^{s_n,1+s_n(1-s_n)}(I)$ with $(v_n)_I = 0$ that satisfy

$$\alpha_0^n |u_n' - s_n v_n|_{\mathcal{M}_b(I)} + \alpha_1^n s_n(1 - s_n)|v_n|_{W^{s_n,1+s_n(1-s_n)}(I)} = |u_n|_{ICTV_{\alpha^n}^{1+s_n}(I)}. \tag{5.4}$$

We write $p_n := 1 + s_n(1 - s_n)$ and observe that $s_n p_n < 1$ for all n. Hence, by theorem 8, we have

$$\begin{aligned} \|v_n\|_{L^1(I)} &\lesssim \|v_n\|_{L^{\frac{p_n}{1-s_np_n}}(I)} \\ &\lesssim \left(\frac{s_n(1 - s_n)}{(1 - s_n p_n)^{p_n - 1}} \right)^{\frac{1}{p_n}} |v_n|_{W^{s_n,p_n}(I)} \\ &= [s_n(1 - s_n)(1 - s_n p_n)]^{\frac{1-p_n}{p_n}} \cdot s_n(1 - s_n)|v_n|_{W^{s_n,p_n}(I)}. \end{aligned}$$

It is easy to see that the first term in the last line is uniformly bounded in n, i.e. we obtain

$$\|v_n\|_{L^1(I)} \lesssim \|v_n\|_{L^{\frac{p_n}{1-s_np_n}}(I)} \lesssim s_n(1 - s_n)|v_n|_{W^{s_n,p_n}(I)}, \tag{5.5}$$

and we find (using $s_n < 1$)

$$\begin{aligned} |u_n'|_{\mathcal{M}_b(I)} &\leq |u_n' - s_n v_n|_{\mathcal{M}_b(I)} + s_n \|v_n\|_{L^1(I)} \\ &\lesssim |u_n' - s_n v_n|_{\mathcal{M}_b(I)} + s_n(1 - s_n)|v_n|_{W^{s_n,p_n}(I)}. \end{aligned} \tag{5.6}$$

By the boundedness assumption of the sequence $(\alpha^n)_n$ and (5.3), we deduce from (5.6) that

$$\sup_{n\in\mathbb{N}} \|u_n\|_{BV(I)} \lesssim \sup_{n\in\mathbb{N}} \|u_n\|_{BCV^{1+s_n}_{\alpha^n}(I)} < \infty,$$

i.e. the sequence $(u_n)_n$ is uniformly bounded in $BV(I)$. Hence, there exists a function $u \in BV(I)$ such that, up to a subsequence, there holds

$$u_n \overset{*}{\rightharpoonup} u \text{ in } BV(I).$$

We note that since $s_n p_n < 1$ and by (5.5), we have

$$\|v_n\|_{L^{p_n}(I)} \lesssim \|v_n\|_{L^{\frac{p_n}{1-s_np_n}}(I)} \lesssim s_n(1-s_n)|v_n|_{W^{s_n,p_n}(I)},$$

hence (note $s_n < 1$)

$$\begin{aligned} s_n(1-s_n)\|v_n\|_{W^{s_n,p_n}(I)} &\leq \|v_n\|_{L^{p_n}(I)} + s_n(1-s_n)|v_n|_{W^{s_n,p_n}(I)} \\ &\lesssim s_n(1-s_n)|v_n|_{W^{s_n,p_n}(I)} \\ &\lesssim \|u_n\|_{BCV^{1+s_n}_{\alpha^n}(I)}, \end{aligned}$$

where we used the boundedness assumption of the sequence $(\alpha^n)_n$ in the last step. In view of (5.3), this implies

$$\sup_{n\in\mathbb{N}} s_n(1-s_n)\|v_n\|_{W^{s_n,p_n}(I)} < \infty. \tag{5.7}$$

We distinguish the cases $\bar{s} = 0$, $\bar{s} \in (0,1)$ and $\bar{s} = 1$:
Case $\bar{s} = 0$: As in the proof of theorem 11, we obtain from (5.7) that, up to a subsequence,

$$s_n v_n \longrightarrow 0 \text{ in } L^1(I).$$

By using (5.4), we conclude

$$\begin{aligned} \liminf_{n\to\infty} |u_n|_{ICTV^{1+s_n}_{\alpha^n}(I)} &= \liminf_{n\to\infty} \left[\alpha_0^n|u_n' - s_nv_n|_{\mathcal{M}_b(I)} + \alpha_1^n s_n(1-s_n)|v_n|_{W^{s_n,p_n}(I)}\right] \\ &\geq \liminf_{n\to\infty} \ \alpha_0^n|u_n' - s_nv_n|_{\mathcal{M}_b(I)} \\ &\geq \alpha_0|u'|_{\mathcal{M}_b(I)}. \end{aligned}$$

Case $\bar{s} \in (0,1)$: We note that we are in the situation of lemma 3. Hence, there exists a function $v \in W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)$ such that, up to a subsequence,

$$v_n \underset{n\to\infty}{\longrightarrow} v \text{ in } L^1(I) \tag{5.8}$$

and

$$\bar{s}(1-\bar{s})|v|_{W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)} \leq \liminf_{n\to\infty} s_n(1-s_n)|v_n|_{W^{s_n,p_n}(I)}.$$

By using (5.4), we conclude

$$\begin{aligned}
\liminf_{n\to\infty} |u_n|_{ICTV^{1+s_n}_{\alpha^n}(I)} &= \liminf_{n\to\infty} \left[\alpha_0^n |u_n' - s_n v_n|_{\mathcal{M}_b(I)} + \alpha_1^n s_n(1-s_n)|v_n|_{W^{s_n,p_n}(I)}\right] \\
&\geq \liminf_{n\to\infty}\ \alpha_0^n |u_n' - s_n v_n|_{\mathcal{M}_b(I)} + \liminf_{n\to\infty}\ \alpha_1^n s_n(1-s_n)|v_n|_{W^{s_n,p_n}(I)} \\
&\geq \alpha_0 |u' - \bar{s} v|_{\mathcal{M}_b(I)} + \alpha_1 \bar{s}(1-\bar{s})|v|_{W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)} \\
&\geq |u|_{ICTV^{1+\bar{s}}_{\alpha}(I)},
\end{aligned}$$

where we used in the last step the definition of the fractional $ICTV$ seminorm together with the fact $(v)_I = 0$ (follows from (5.8) and $(v_n)_I = 0$).
Case $\bar{s} = 1$: In this case, by lemma 3, there exists a function $v \in BV(I)$ such that, up to a subsequence,

$$v_n \underset{n\to\infty}{\longrightarrow} v \text{ in } L^1(I)$$

and

$$|v'|_{\mathcal{M}_b(I)} \leq \liminf_{n\to\infty} s_n(1-s_n)|v_n|_{W^{s_n,1+s_n(1-s_n)}(I)}.$$

Similar to the previous case, we obtain

$$\liminf_{n\to\infty} |u_n|_{ICTV^{1+s_n}_{\alpha^n}(I)} \geq \alpha_0|u' - v|_{\mathcal{M}_b(I)} + \alpha_1 |v'|_{\mathcal{M}_b(I)} \geq |u|_{ICTV^2_\alpha(I)}.$$

□

Now we are able to prove the existence and uniqueness of solutions for our bilevel learning scheme ($\mathcal{B}$). The main work lies in proving the existence, whereas uniqueness will easily follow from strict convexity.

Theorem 13 (Existence and Uniqueness for ($\mathcal{B}$))**.** *Let $u_0, u_c \in BV(I)$ be given. Then there exists a unique solution*

$$(\alpha^*, r^*) \in [a, A]^{\lfloor r^* \rfloor + 1} \times [1, R]$$

to ($\mathcal{B}$) with the corresponding optimal reconstructed image

$$u_{\alpha^*, r^*} \in BCV^{r^*}_{\alpha^*}(I).$$

Proof. Step 1: We show that for each $(\alpha, r) \in [a, A]^{\lfloor r \rfloor + 1} \times [1, R]$, there exists a unique function $u_{\alpha,r} \in BCV^r_\alpha(I)$ satisfying

$$\|u_{\alpha,r} - u_0\|^2_{L^2(I)} + |u_{\alpha,r}|_{ICTV^r_\alpha(I)} = \min_{u \in BCV^r_\alpha(I)} \left[\|u - u_0\|^2_{L^2(I)} + |u|_{ICTV^r_\alpha(I)}\right]. \tag{5.9}$$

To this end, we first note that

$$\inf_{u \in BCV^r_\alpha(I)} \left[\|u - u_0\|^2_{L^2(I)} + |u|_{ICTV^r_\alpha(I)}\right] =: \gamma \in [0, \infty).$$

Hence, by definition of γ, there exists a minimising sequence $(u_n)_n \subset BCV_\alpha^r(I)$ with

$$\|u_n - u_0\|_{L^2(I)}^2 + |u_n|_{ICTV_\alpha^r(I)} \underset{n\to\infty}{\longrightarrow} \gamma. \tag{5.10}$$

In view of remark 9, the sequence $(u_n)_n$ is uniformly bounded in $BV(I)$, hence we find some $u_{\alpha,r} \in BV(I)$ such that (up to a subsequence)

$$u_n \overset{*}{\rightharpoonup} u_{\alpha,r} \text{ in } BV(I)$$

and by compactness

$$u_n \to u_{\alpha,r} \text{ in } L^2(I).$$

We use lemma 4 (and the fact that integer order $ICTV$ seminorms are lower semi-continuous with respect to weak-star convergence in BV) to conclude

$$\begin{aligned}
\|u_{\alpha,r} - u_0\|_{L^2(I)}^2 + |u_{\alpha,r}|_{ICTV_\alpha^r(I)} &\leq \liminf_{n\to\infty} \|u_n - u_0\|_{L^2(I)}^2 + \liminf_{n\to\infty} |u_n|_{ICTV_\alpha^r(I)} \\
&\leq \liminf_{n\to\infty} \left[\|u_n - u_0\|_{L^2(I)}^2 + |u_n|_{ICTV_\alpha^r(I)}\right] \\
&= \gamma,
\end{aligned}$$

where we used (5.10) in the last equality. Note that this shows in particular that $u \in BCV_\alpha^r(I)$ and hence (5.9) follows.
To show uniqueness, it is sufficient to show that the functional

$$J : BCV_\alpha^r(I) \to \mathbb{R}, \quad J(v) := \|v - u_0\|_{L^2(I)}^2 + |v|_{ICTV_\alpha^r(I)}$$

is strictly convex. This is clearly true, since the squared L^2-norm is strictly convex and the $ICTV$ seminorm is convex.
Step 2: We write

$$\Gamma := \inf\left\{\|u_{\alpha,r} - u_c\|_{L^2(I)}^2 : (\alpha, r) \in [a, A]^{\lfloor r\rfloor+1} \times [1, R]\right\}$$

and note that $\Gamma \in [0, \infty)$. Hence, there exists a minimising sequence $\{(\alpha^n, r_n)\}_n$ with

$$(\alpha^n, r_n) \in [a, A]^{\lfloor r_n\rfloor+1} \times [1, R] \tag{5.11}$$

for all $n \in \mathbb{N}$ and

$$\|u_{\alpha^n,r_n} - u_c\|_{L^2(I)}^2 \underset{n\to\infty}{\longrightarrow} \Gamma,$$

where $u_{\alpha^n,r_n} \in BCV_{\alpha^n}^{r_n}(I)$ is the unique function from step 1 satisfying

$$\|u_{\alpha^n,r_n} - u_0\|_{L^2(I)}^2 + |u_{\alpha^n,r_n}|_{ICTV_{\alpha^n}^{r_n}(I)} = \min_{u\in BCV_{\alpha^n}^{r_n}(I)} \left[\|u - u_0\|_{L^2(I)}^2 + |u|_{ICTV_{\alpha^n}^{r_n}(I)}\right]. \tag{5.12}$$

We observe that by (5.11) and the definition of the $ICTV$ seminorms, (take all $v_i = 0$)

$$|u_0|_{ICTV^{r_n}_{\alpha^n}(I)} \lesssim |u_0'|_{\mathcal{M}_b(I)},$$

which implies $u_0 \in BCV^{r_n}_{\alpha^n}(I)$. Then, in view of (5.12), there holds

$$\|u_{\alpha^n,r_n} - u_0\|^2_{L^2(I)} + |u_{\alpha^n,r_n}|_{ICTV^{r_n}_{\alpha^n}(I)} \leq |u_0|_{ICTV^{r_n}_{\alpha^n}(I)} \lesssim |u_0'|_{\mathcal{M}_b(I)}$$

and we deduce

$$\sup_{n\in\mathbb{N}} \|u_{\alpha^n,r_n}\|_{BCV^{r_n}_{\alpha^n}(I)} < \infty.$$

By (5.11) and Bolzano-Weierstraß, there exist some $r^*, k \in [1, R]$ such that, up to a subsequence,

$$r_n \longrightarrow r^*, \quad \lfloor r_n \rfloor \longrightarrow k.$$

Clearly, for n sufficiently large, there holds $\alpha^n \in [a, A]^{k+1}$ and

$$r_n = k + s_n,$$

where $s_n := r_n - k \to r^* - k =: \bar{s} \in [0, 1]$. Hence, we are in the situation of lemma 4 and we find $\alpha^* \in [a, A]^{\lfloor r^* \rfloor + 1}$ and a function $u_{\alpha^*,r^*} \in BCV^{r^*}_{\alpha^*}(I)$ such that, up to a subsequence,

$$u_{\alpha^n,r_n} \overset{*}{\rightharpoonup} u_{\alpha^*,r^*} \text{ in } BV(I)$$

and by compactness

$$u_{\alpha^n,r_n} \to u_{\alpha^*,r^*} \text{ in } L^2(I).$$

Thus, we can deduce

$$\|u_{\alpha^*,r^*} - u_c\|^2_{L^2(I)} \leq \lim_{n\to\infty} \|u_{\alpha^n,r_n} - u_c\|^2_{L^2(I)} = \Gamma.$$

It remains to show that the minimiser is unique, but this follows directly from the strict convexity of the squared $L^2(I)$-norm. □

6 Conclusion

Now that we have finished the mathematical analysis of the bilevel learning scheme ($\mathcal{B}$), we want to have a look back on what we have done in this essay. Further, we will have a look at a simulation of the scheme and discuss open problems in the end.

Review. Let us briefly remind ourselves of the cornerstones in this essay. After having recalled the main results from the theory of the space BV in section 2 (approximation and compactness), we started studying fractional Sobolev spaces in section 3. In the first part of section 3, we provided, apart from the definition of the spaces, a very useful toolbox consisting of embedding theorems (including compactness), a Poincaré inequality and a theorem stating reflexivity. We used these tools throughout our analysis to find convergent subsequences. In the second part of section 3, we studied the asymptotic behaviour of the Gagliardo seminorm and proved our first big theorem (theorem 9) which built the basis for section 4.
In section 4, we investigated the asymptotic behaviour of the fractional $ICTV$ seminorms (theorem 11, remark 6). On our way to this result, we proved a very useful compactness and lower semicontinuity result (lemma 3) which has been used in the proof of lemma 4, which was crucial for establishing the existence of a solution of ($\mathcal{B}$). In the last part of section 4, we saw firstly, that for an $ICTV$ function there always exists an extremal function for the infimum in the definition of the $ICTV$ seminorm (theorem 12, remark 8) and secondly, that the fractional $ICTV$ seminorms are actually equivalent to the total variation (proposition 2, remark 9).
Section 5 was then all about the proof of theorem 13, the existence and uniqueness of a solution to ($\mathcal{B}$). In the first step of the proof, we showed that minimising sequences for the minimisation problem in the second level of the bilevel learning scheme are bounded in BV such that we were able to use compactness in BV and lemma 4, the key lemma in the proof, to deduce the existence of a solution. The second step was then to show with a minimising sequence for the first level of the scheme ($\mathcal{B}$) at hand, that we are in the situation of lemma 4, which concludes the proof. This shows that the main work was in the proof of lemma 4, which made use of our results from sections 4.2 and 4.3.
Finally, deducing uniqueness of the solution was straightforward by using strict convexity.

Staircasing effect. One motivation in using our bilevel learning scheme was to reduce the staircasing effect which occurs in total variation based image reconstruction schemes. Figure 4 illustrates the signal denoising via ($\mathcal{B}$) where the clean signal includes a piecewise linear part. Apparently, the staircasing effect is attenuated and we get a very satisfactory result.

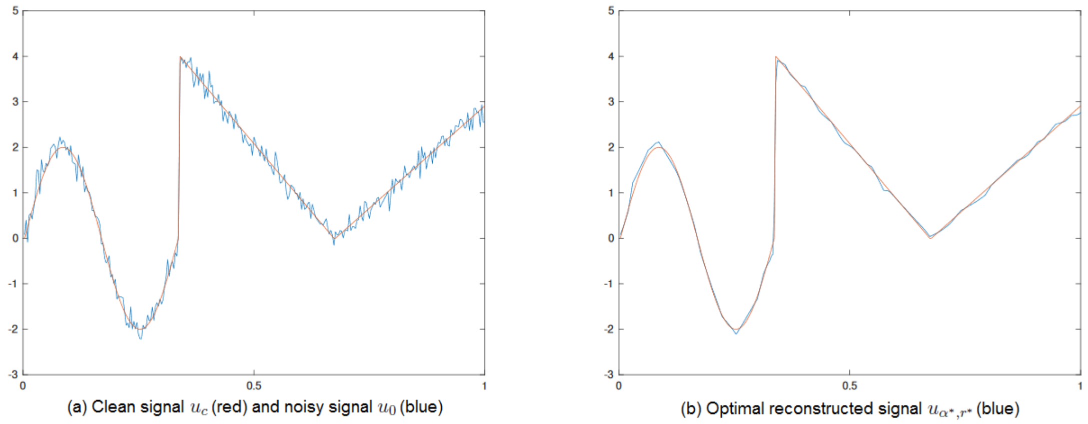


Figure 4: (taken from [DL16]) *Denoising via the bilevel learning scheme* ($\mathcal{B}$) *with* $a = 0.001$, $A = 2.5$ *and* $R = 2$.

Current and future works. The following problems are currently/can be taken up for further research:

- *Higher dimensions*: A generalisation of the bilevel learning scheme ($\mathcal{B}$) to the two dimensional setting (analysis and numerics).
  In the one dimensional setting we used the Gagliardo seminorm as a way to express fractional derivatives, which can be seen as a global definition (no pointwise information). Another approach would be to use fractional derivatives in the sense of Riemann-Liouville. The so-called left-sided Riemann-Liouville derivative provides a definition of the fractional derivative pointwise almost everywhere, see [CZ15].
  Using this pointwise approach, in [CZ15] the function space BV^s for $s \in (0,1)$ is introduced. This space consists of L^1 functions with s-bounded variation, that is
  $$TV^s(u) := \sup_{\substack{\phi \in C_c^\infty(\Omega) \\ \|\phi\|_\infty \leq 1}} \int_\Omega -u(x)\, \mathrm{div}^s \phi(x)\, dx,$$
  where $\Omega \subset \mathbb{R}^2$ is the domain of the image u, and div^s is the sum of the Riemann-Liouville fractional derivatives of order s over all coordinate directions.
  This approach could be used as starting point to construct a regulariser in two dimensions similar to the fractional $ICTV$ seminorms. The key challenge is the lack of knowledge of the space BV^s, i.e. whereas we were able to use all the interpolation space tools and Sobolev space properties in the Gagliardo setting, we basically start from scratch in two dimensions with this new starting point.

- *New regularisers*: A further extension to fractional order $ICTV$ seminorms. The authors of [DL16] are currently investigating this point with the help of quasiconvexity theory.

- *Denoising locally*: Improving $(\mathcal{B})$ by a spatially dependent learning method. A trilevel learning scheme which improves $(\mathcal{B})$ is suggested by P. Liu in [Liu16]. The idea of the scheme is as follows. Given a partition of the domain in cubes, we find for each cube the optimal regularisation parameter separately and then glue the corresponding reconstructed images together. The learning scheme then finds the best such partition and results in the image corresponding to this optimal partition.

7 Appendix

7.1 Notation

- $I := (0,1) \subset \mathbb{R}$.
- $\mathbb{R}_+ := (0,\infty) \subset \mathbb{R}$.
- X': Dual space of X.
- $\mathcal{L}^N$: Lebesgue-measure of dimension N.
- $|u'|_{\mathcal{M}_b(I)}$: Total variation of u on I.
- $\mathbb{1}_\Omega(x) := \begin{cases} 1 \ , x \in \Omega \\ 0 \ , x \notin \Omega \end{cases}$: Characteristic function on the set $\Omega \subset \mathbb{R}^N$.
- $(u)_\Omega := \frac{1}{|\Omega|}\int_\Omega u$: Mean of u on Ω, where $|\Omega|$ is the Lebesgue-measure of Ω.
- $(u_n), (u_n)_n, (u_n)_{n\in\mathbb{N}}$: Sequences with running index n.
- $\operatorname{supp}(u) := \overline{\{x \in \Omega : u(x) \neq 0\}}$: Support of $u : \Omega \to \mathbb{R}$.
- $\operatorname{dist}(\Omega_1, \Omega_2)$: Distance of the sets $\Omega_1, \Omega_2 \subset \mathbb{R}^N$.
- $\operatorname{Lip}(u)$: Lipschitz constant of u.
- $D^\alpha u$: The α-th derivative of u.
- $o, \mathcal{O}$: Landau-symbols.
- $\lfloor x \rfloor := \max\{m \in \mathbb{Z} : m \leq x\}$: Floor-function.
- $\lesssim$: Write $A \lesssim B$ if there exists $C > 0$ s.t. $A \leq CB$.

7.2 Numerics

The following numerical method for solving the lower level problem in ($\mathcal{B}$) is presented in [DL16]. It relies on the fact that we can identify fractional Sobolev spaces with Besov spaces, see [Tri10]. Let us consider the case $k = 1$.
By definition of the fractional $ICTV$ seminorms and using the minimax formulation of Besov spaces presented in [GLM07], we find

$$\begin{aligned}
&\min_u \left[\frac{1}{2}\|u-u_0\|^2_{L^2(I)} + |u|_{ICTV^{1+s}_\alpha(I)}\right] \\
&= \min_{u,v_0} \left[\frac{1}{2}\|u-u_0\|^2_{L^2(I)} + \alpha_0|u'-sv_0|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s)|v_0|_{W^{s,1+s(1-s)}(I)}\right] \\
&= \min_{u,v_0}\max_{\varphi,t} \left[\frac{1}{2}\|u-u_0\|^2_{L^2(I)} + \langle u'-sv_0,\varphi\rangle - \chi_{\varphi,\alpha_0} + \alpha_1 s(1-s)\left(\int_I |K^s_t * v_0|^{s'}\right)^{\frac{1}{s'}}\right],
\end{aligned}$$

where K^s_t denotes the operator from [GLM07, Section 3.4], $s' := 1 + s(1-s)$, and

$$\chi_{\varphi,\alpha_0}(x) := \begin{cases} 0 & , \text{ if } |\varphi(x)| \le \alpha_0 \\ \infty & , \text{ else} \end{cases}$$

for $x \in I$. We observe that the reformulated minimax problem is actually a saddle-point problem which fits into the framework of [CP11] and can be solved by a first-order primal-dual algorithm.
The numerical solution of the bilevel problem for $ICTV^k$ is a matter of future research. The case $k = 2$ has been investigated in [DSV17].

References

[Alt06] Alt, H. W., *Lineare Funktionalanalysis*, 5. überarb. Auflage, Springer-Verlag Berlin Heidelberg, 2006

[BBM02] Bourgain, J., Brezis, H., and Mironescu, P., *Limiting embedding theorems for $W^{s,p}$ when $s \nearrow 1$ and applications*, J. Anal. Math., No. 87, 2002

[BV11] Bredies, K., and Valkonen, T., *Inverse problems with second-order total generalized variation constraints*, Proceedings of SampTA 2011 - 9th Intern. Conf. on Samp. Th. and Appl., 2011

[CDS16] Chung, C., De Los Reyes, J. C., and Schönlieb, C.-B., *Learning optimal spatially-dependent regularization parameters in total variation image restoration*, 2016

[CP11] Chambolle, A., and Pock, T., *A first-order primal-dual algorithm for convex problems with applications to imaging*, J. Math. Imag. Vision, Vol. 40, pp. 120–145, 2011

[CZ15] Chen, K., and Zhang, J., *A total fractional-order variation model for image restoration with nonhomogeneous boundary conditions and its numerical solution*, SIAM J. Imaging Sci. Vol. 8, No. 4, pp. 2487–2518

[DL16] Davoli, E., and Liu, P., *Optimal Parameters and regularizers for image reconstruction problems - one dimensional setting*, 2016

[DPV12] Di Nezza, E., Palatucci, G., and Valdinoci, E., *Hitchhiker's guide to the fractional Sobolev spaces*, Bull. Sci. math. 136, No. 5, pp. 521–573, 2012.

[DSV17] De Los Reyes, J. C., Schönlieb, C.-B., and Valkonen, T., *Bilevel parameter learning for higher-order total variation regularisation models*, J. Math. Imaging and Vision, No. 57, pp. 1–25, 2017

[Eva10] Evans, L. C., *Partial Differential Equations*, Amer. Math. Soc., Vol. 19, Second Edition, 2010

[EG92] Evans, L. C., and Gariepy, R. F., *Measure Theory and Fine Properties of Functions*, Studies in Advanced Mathematics, CRC Press, 1992

[FOR92] Fatemi, E., Osher, S., and Rudin, L. I., *Nonlinear total variation based noise removal algorithms*, Physica D, Vol. 60, pp. 259-268, 1992

[GLM07] Garnett, J. B., Le, T. M., Meyer, Y., and Vese, L. A., *Image decompositions using bounded variation and generalized homogeneous Besov spaces*, Appl. Comp. Harm. Anal., Vol. 23, pp. 25–56, 2007

[Liu16] Liu, P., *The spatially dependent bi-level learning scheme*, 2016

[Ma14] Ma, D., *Asymmetric anisotropic fractional Sobolev norms*, Arch. Math. 103, pp. 167–175, 2014

[MS02] Maz'ya, V., and Shaposhnikova, T., *On the Bourgain, Brezis, and Mironescu theorem concerning limiting embeddings of fractional Sobolev spaces*, J. Funct. Anal., No. 195, pp. 230–238, 2002

[PS14] Papafitsoros, K., and Schönlieb, C.-B., *A combined first and second order variational approach for image reconstruction*, J. Math. Imaging and Vision, No. 48(2), pp. 308–338, 2014

[Rud87] Rudin, W., *Real and complex analysis*, 3rd ed. New York: McGraw-Hill Book Co., 1987

[Sch16] Schweizer, B., *Maßtheorie für Partielle Differentialgleichungen*, lecture notes TU Dortmund, Sommersemester 2016

[Sim90] Simon, J., *Sobolev, Besov and Nikolskii Fractional Spaces: Imbeddings and Comparisons for Vector Valued Spaces on an Interval*, Ann. di Mat. pura ed appl., No. 157, pp. 117–148, 1990.

[Tri10] Triebel, H., *Theory of Function Spaces*, Modern Birkhäuser Classics, Basel, 2010

[WX09] Wei, S., and Xu, H., *Staircasing reduction model applied to total variation based image reconstruction*, 17th Europ. Signal Processing Conf., 2009