NFN - Nationales Forschungsnetzwerk

Geometry + Simulation

http://www.gs.jku.at

Regularization with Metric Double Integrals of Functions with Values in a Set of High-Dimensional Vectors

René Ciak and Melanie Hirzmann and Otmar Scherzer

G+S Report No. 72

June 2018

Der Wissenschaftsfonds.

JOHANNES KEPLER UNIVERSITY LINZ

Regularization with Metric Double Integrals of Functions with Values in a Set of High-Dimensional Vectors

René Ciak[1] Melanie Hirzmann[1] Otmar Scherzer[1,2]
rene.ciak@univie.ac.at melanie.hirzmann@univie.ac.at otmar.scherzer@univie.ac.at

May 19, 2018

[1]Computational Science Center
University of Vienna
Oskar-Morgenstern-Platz 1
A-1090 Vienna, Austria

[2]Johann Radon Institute for Computational and Applied Mathematics (RICAM)
Altenbergerstraße 69
A-4040 Linz, Austria

Abstract

We present an approach for variational regularization of inverse and imaging problems for recovering functions with values in a set of vectors. We introduce regularization functionals, which are derivative-free double integrals of such functions. These regularization functionals are motivated from double integrals, which approximate Sobolev semi-norms of intensity functions. These were introduced in Bourgain, Brézis & Mironescu, *"Another Look at Sobolev Spaces". In: Optimal Control and Partial Differential Equations*-Innovations & Applications, IOS press, Amsterdam, 2001. For the proposed regularization functionals we prove existence of minimizers as well as a stability and convergence result for functions with values in a set of vectors.

1. Introduction

Functions with values in a set of vectors appear in several applications of imaging and in inverse problems, such as (we emphasize that in this paper vectors are synonym for vectors, matrices and tensors, respectively):

- *Interferometric Synthetic Aperture Radar (InSAR)* is a technique used in remote sensing and geodesy to generate for example digital elevation maps of the earth's surface. InSAR images represent phase differences of waves between two or more SAR images, cf. [42, 51]. Therefore InSAR data are functions $f : \Omega \to \mathbb{S}^1 \subseteq \mathbb{R}^2$. The pointwise function values are on the $\mathbb{S}^1$, which is considered embedded into $\mathbb{R}^2$.

- A *color image* can be represented as a function in *HSV*-space (hue, saturation, value) (see e.g. [46]). Color images are then described as functions $f : \Omega \to K \subseteq \mathbb{R}^3$. Here Ω is a plane in $\mathbb{R}^2$, the image domain, and K (representing the HSV-space) is a cone in 3-dimensional space $\mathbb{R}^3$.

- Estimation of the *foliage angle distribution* has been considered for instance in [37, 49]. Thereby the imaging function is from $\Omega \subset \mathbb{R}^2$, a part of the Earth's surface, into $\mathbb{S}^2 \subseteq \mathbb{R}^3$, representing foliage angle orientation.

- Estimation of functions with values in $SO(3) \subseteq \mathbb{R}^{3\times 3}$. Such problems appear in *Cryo-Electron Microscopy* (see for instance [36, 56, 59]).

We emphasize that we are analyzing *vector*, *matrix*, *tensor*-valued functions, where pointwise function evaluations are elements of some given set. This should not be confused with set-valued functions, where every function evaluation can be a set.

Inverse problems and imaging tasks, such as the ones mentioned above, might be unstable, or even worse, the solution could be ambiguous. Therefore, numerical algorithms for imaging need to be *regularizing* to obtain approximations of the desired solution in a stable manner. Consider the operator equation

$$\mathrm{F}(w) = v^0, \tag{1.1}$$

where we assume that only (noisy) measurement data v of v^0 become available. In this paper the method of choice is *variational regularization* which consists in calculating a minimizer of the variational regularization functional

$$\mathcal{F}(w) := \mathcal{D}(\mathrm{F}(w), v) + \alpha\mathcal{R}(w). \tag{1.2}$$

Here

w is an element of the *set* $\mathcal{W}$ of admissible functions.

F is an operator modeling the image formation process.

$\mathcal{D}$ is called the *data* or *fidelity term*, which is used to compare a pair of data in the image domain, that is to quantify the difference of the two data sets.

$\mathcal{R}$ is called *regularization functional*, which is used to impose certain properties onto a minimizer of the regularization functional $\mathcal{F}$.

$\alpha > 0$ is called *regularization parameter* and provides a trade off between stability and approximation properties of the minimizer of the regularization functional $\mathcal{F}$.

v denotes measurement data, which we consider noisy.

v^0 denotes the exact data, which we assume to be not necessarily available.

The main objective of this paper is to introduce a general class of regularization functionals for functions with values in a set of vectors. In order to motivate our proposed class of regularization functionals we review a class of regularization functionals appropriate for analyzing *intensity data*.

Variational regularization for reconstruction of intensity data. Opposite to what we consider in the present paper, most commonly, imaging data v and admissible functions w, respectively, are considered to be representable as *intensity functions*. That is, they are functions from some subset Ω of an Euclidean space with *real values*.

In such a situation the most widely used regularization functionals use regularization terms consisting of powers of Sobolev (see [12, 15, 16]) or total variation semi-norms [52]. It is common to speak about *Tikhonov regularization* (see for instance [57]) when the data term and the regularization functional are squared Hilbert space norms, respectively. For the *Rudin, Osher, Fatemi (ROF)* regularization [52], also known as total variation regularization, the data term is the squared L^2-norm and $\mathcal{R}(w) = |w|_{TV}$ is the total variation semi-norm. Other widely used regularization functionals are *sparsity promoting* [22, 39], *Besov space norms* [44, 40] and anisotropic regularization norms [45, 54]. Aside from various regularization terms there also have been proposed different fidelity terms other than quadratic norm fidelities, like the p-th powers of ℓ^p and L^p-norms of the differences of $F(w)$ and v , [53, 55], Maximum Entropy [25, 26] and Kullback-Leibler divergence [50] (see [48] for some reference work).

Our work utilizes results from the seminal paper of Bourgain, Brézis, and Mironescu[13], which provides an equivalent *derivative-free* characterization of Sobolev spaces and the space $BV(\Omega, \mathbb{R}^M)$, the space of functions of bounded total variation, which consequently, in this context, was analyzed in Dávila and Ponce [23, 47], respectively. It is shown in [13, Theorems 2 & 3] and [23, Theorem 1] that when $(\rho_\varepsilon)_{\varepsilon>0}$ is a suitable sequence of non-negative, radially symmetric, radially decreasing mollifiers, then

$$\begin{aligned}\lim_{\varepsilon\to 0+} \tilde{\mathcal{R}}_\varepsilon(w) &:= \lim_{\varepsilon\to 0+} \int\limits_{\Omega\times\Omega} \frac{\|w(x)-w(y)\|_{\mathbb{R}}^p}{\|x-y\|_{\mathbb{R}^N}^p}\rho_\varepsilon(x-y)\,\mathrm{d}(x,y) \\ &= \begin{cases} C_{p,N}|w|_{W^{1,p}}^p & \text{if } w \in W^{1,p}(\Omega,\mathbb{R}),\ 1<p<\infty, \\ C_{1,N}|w|_{TV} & \text{if } w \in BV(\Omega,\mathbb{R}),\ p=1, \\ \infty & \text{otherwise,}\end{cases}\end{aligned} \tag{1.3}$$

Hence $\tilde{\mathcal{R}}_\varepsilon$ approximates powers of Sobolev semi-norms and the total variation semi-norm, respectively. Variational imaging, consisting in minimization of $\mathcal{F}$ from Equation 1.2 with $\mathcal{R}$ replaced by $\tilde{\mathcal{R}}_\varepsilon$, has been considered in [3, 11].

Regularization of functions with values in a set of vectors. In this paper we generalize the derivative-free characterization of Sobolev spaces and functions of bounded variation to functions, $u : \Omega \to K$, where K is some set of high-dimensional vectors, and use these functionals for variational regularization. The applications we have in mind contain that K is a closed subset of $\mathbb{R}^M$ (for instance HSV-data) with non-zero measure, or that K is a sub-manifold (such as for instance InSAR-data).

The reconstruction of manifold–valued data with variational regularization methods has already been subject to intensive research (see for instance [38, 19, 18, 17, 4, 60]). The variational approaches mentioned above use regularization and fidelity functionals based on Sobolev and TV semi-norms: [18, 19] introduced a total variation regularizer for cyclic data on $\mathbb{S}^1$, see also [7, 9, 10]. In [4, 6] combined first and second order differences and derivatives were used for regularization to restore manifold valued data. The later mentioned papers, however, are formulated in a finite dimensional setting, opposed to ours, which is considered in an infinite dimensional setting. Algorithms for total variation minimization problems, including half-quadratic minimization and non-local patch based methods, are given for example in [4, 5, 8] as well as in [35, 41]. On the theoretical side the total variation of functions with values in a manifold was investigated by Giaquinta and Mucci using the theory of Cartesian currents in [32, 33], and earlier [31] if the manifold is a $\mathbb{S}^1$.

The contents and the particular achievements of the paper are as follows. The contribution of this paper is to introduce and analytically analyze double integral regularization functionals for reconstructing functions with values in a set of vectors, generalizing functionals of the form Equation 1.3. Moreover, we develop and analyze fidelity terms for comparing manifold valued data. Summing these two terms provides a new class of regularization functionals of the form Equation 1.2 for reconstructing manifold valued data.

When analyzing our functionals we encounter several differences to existing regularization theory (compare Section 2):

(i) The *set of admissible functions*, where we minimize the regularization functional on, *cannot* be associated with a *linear* space of functions. As a consequence, well–posedness of the variational method (that is, existence of a minimizer of the energy functional) cannot be proven by applying standard direct methods in the Calculus of Variations [21, 20].

(ii) The regularization functionals are defined via elementary metrics and not norms, see Section 3.

(iii) In general, the fidelity terms are *non-convex*. Stability and convergence results are proven in Section 4.

The model is validated in Section 6 where we present numerical results for InSAR data denoising and inpainting.

2. Setting

In the following we introduce the basic notation and the set of admissible set of functions which we are regularizing on.

Assumption 2.1 All along this paper we assume that

- $p_1, p_2 \in [1, +\infty)$, $s \in (0, 1]$,
- $\Omega_1, \Omega_2 \subseteq \mathbb{R}^N$ are nonempty, bounded, and connected open sets with Lipschitz boundary, respectively,
- $k \in [0, N]$,
- $K_1 \subseteq \mathbb{R}^{M_1}, K_2 \subseteq \mathbb{R}^{M_2}$ are nonempty and closed subsets of $\mathbb{R}^{M_1}$ and $\mathbb{R}^{M_2}$, respectively.

Moreover,

- $\|\cdot\|_{\mathbb{R}^N}$ and $\|\cdot\|_{\mathbb{R}^{M_i}}$, $i = 1, 2$, are the Euclidean norms on $\mathbb{R}^N$ and $\mathbb{R}^{M_i}$, respectively.
- $\mathrm{d}_{\mathbb{R}^{M_i}} : \mathbb{R}^{M_i} \times \mathbb{R}^{M_i} \to [0, +\infty)$ denotes the Euclidean distance on $\mathbb{R}^{M_i}$ for $i = 1, 2$ and
- $\mathrm{d}_i := \mathrm{d}_{K_i} : K_i \times K_i \to [0, +\infty)$ denote arbitrary metrics on K_i, which fulfill for $i = 1$ and $i = 2$
  - $\mathrm{d}_{\mathbb{R}^{M_i}}|_{K_i \times K_i} \leq d_i$,
  - d_i is continuous with respect to $\mathrm{d}_{\mathbb{R}^{M_i}}|_{K_i \times K_i}$, meaning that for a sequence $(a_n)_{n\in\mathbb{N}}$ in $K_i \subseteq \mathbb{R}^{M_i}$ converging to some $a \in K_i$ we have that $\mathrm{d}_i(a_n, a) \to 0$.

  In particular, this assumption is valid if the metric d_i is equivalent to $\mathrm{d}_{\mathbb{R}^{M_i}}|_{K_i \times K_i}$.
- For all $\varepsilon > 0$ ρ_ε is a non-negative, radially symmetric mollifier: That is
  (i) $\rho_\varepsilon \in \mathcal{C}_c^\infty(\mathbb{R}^N, \mathbb{R})$ is radially symmetric,
  (ii) $\rho_\varepsilon \geq 0$,
  (iii) $\int\limits_{\mathbb{R}^N} \rho_\varepsilon(x)\,\mathrm{d}x = 1$, and
  (iv) for all $\delta > 0$, $\lim_{\varepsilon\to 0+} \int_{\left\{\|y\|_{\mathbb{R}^N} > \delta\right\}} \rho_\varepsilon(y)\,\mathrm{d}y = 0$.
- When we write p, Ω, K, then we mean p_i, Ω_i, K_i, for either $i = 1, 2$. In the following we will often omit the subscript indices whenever possible.

Example 2.2 *Let $\hat{\rho} \in C_c^\infty(\mathbb{R}, \mathbb{R}_+)$, symmetric at 0 and satisfy*

$$\left|\mathbb{S}^{N-1}\right| \int_0^\infty \hat{t}^{N-1} \hat{\rho}\left(\hat{t}\right) d\hat{t} = 1 \,.$$

Then for every $\varepsilon > 0$

$$x \in \mathbb{R}^n \to \rho_\varepsilon(x) := \frac{1}{\varepsilon^N} \hat{\rho}\left(\frac{\|x\|_{\mathbb{R}^N}}{\varepsilon}\right)$$

is a mollifier. To see this, we note that

- *by substitution $x = t\theta$ with $t > 0, \theta \in \mathbb{S}^{N-1}$ and $\hat{t} = \frac{t}{\varepsilon}$,*

$$\begin{aligned} \int_{\mathbb{R}^N} \rho_\varepsilon(x)\,\mathrm{d}x &= \frac{1}{\varepsilon^N} \int_{\mathbb{R}^N} \hat{\rho}\left(\frac{\|x\|_{\mathbb{R}^N}}{\varepsilon}\right) \mathrm{d}x \\ &= \frac{1}{\varepsilon^N} \int_0^\infty t^{N-1} \hat{\rho}\left(\frac{t}{\varepsilon}\right) dt \int_{\mathbb{S}^{N-1}} d\theta \\ &= \left|\mathbb{S}^{N-1}\right| \int_0^\infty \hat{t}^{N-1} \hat{\rho}\left(\hat{t}\right) d\hat{t} = 1 \,. \end{aligned} \tag{2.1}$$

  Here, $d\theta$ denotes the canonical spherical measure.
- *Again by the same substitutions, taking into account that $\hat{\rho}$ has compact support, it follows for $\varepsilon > 0$ sufficiently small that*

$$\begin{aligned} \int_{\left\{\|y\|_{\mathbb{R}^N} > \delta\right\}} \rho_\varepsilon(x)\,\mathrm{d}x &= \frac{1}{\varepsilon^N} \int_{\left\{\|y\|_{\mathbb{R}^N} > \delta\right\}} \hat{\rho}\left(\frac{\|x\|_{\mathbb{R}^N}}{\varepsilon}\right) \mathrm{d}x \\ &= \frac{1}{\varepsilon^N} \int_\delta^\infty t^{N-1} \hat{\rho}\left(\frac{t}{\varepsilon}\right) dt \int_{\mathbb{S}^{N-1}} d\theta \\ &= \left|\mathbb{S}^{N-1}\right| \int_{\delta/\varepsilon}^\infty \hat{t}^{N-1} \hat{\rho}\left(\hat{t}\right) d\hat{t} = 0 \,. \end{aligned} \tag{2.2}$$

Thus $\{\rho_\varepsilon : \varepsilon > 0\}$ is a set of mollifiers.

In the following we write down the basic spaces and sets, which will be used in the course of the paper.

Definition 2.3 • The *Lebesgue–Bochner space* of $\mathbb{R}^M$–valued functions on Ω consists of the set

$$L^p(\Omega,\mathbb{R}^M) := \{w:\Omega\to\mathbb{R}^M : w \text{ is Lebesgue-Borel measurable and } \|w(\cdot)\|^p_{\mathbb{R}^M}:\Omega\to\mathbb{R} \text{ is Lebesgue–integrable on } \Omega\},$$

which is associated with the norm $\|\cdot\|_{L^p(\Omega,\mathbb{R}^M)}$, given by

$$\|w\|_{L^p(\Omega,\mathbb{R}^M)} := \Big(\int_\Omega \|w(x)\|^p_{\mathbb{R}^M}\Big)^{1/p}.$$

- Let $0<s<1$. Then the *fractional Sobolev space* of order s is defined (see [1]) as the set

$$W^{s,p}(\Omega,\mathbb{R}^M) := \left\{ w\in L^p(\Omega,\mathbb{R}^M) : \frac{\|w(x)-w(y)\|_{\mathbb{R}^M}}{\|x-y\|^{\frac{N}{p}+s}_{\mathbb{R}^N}} \in L^p(\Omega\times\Omega,\mathbb{R})\right\}$$
$$= \{w\in L^p(\Omega,\mathbb{R}^M) : |w|_{W^{s,p}(\Omega,\mathbb{R}^M)}<\infty\},$$

equipped with the norm

$$\|\cdot\|_{W^{s,p}(\Omega,\mathbb{R}^M)} := \left(\|\cdot\|^p_{L^p(\Omega,\mathbb{R}^M)} + |\cdot|^p_{W^{s,p}(\Omega,\mathbb{R}^M)}\right)^{1/p}, \tag{2.3}$$

where $|\cdot|_{W^{s,p}(\Omega,\mathbb{R}^M)}$ is the semi-norm for $W^{s,p}(\Omega,\mathbb{R}^M)$, given by

$$|w|_{W^{s,p}(\Omega,\mathbb{R}^M)} := \Big(\int_{\Omega\times\Omega} \frac{\|w(x)-w(y)\|^p_{\mathbb{R}^M}}{\|x-y\|^{N+ps}_{\mathbb{R}^N}}\,\mathrm{d}(x,y)\Big)^{1/p}, \quad w\in W^{s,p}(\Omega,\mathbb{R}^M). \tag{2.4}$$

- For $s=1$ the Sobolev space $W^{1,p}(\Omega,\mathbb{R}^M)$ consists of all functions in $L^1(\Omega,\mathbb{R}^M)$ for which

$$\|w\|_{W^{1,p}(\Omega,\mathbb{R}^M)} := \Big(\|w\|^p_{L^p(\Omega,\mathbb{R}^M)} + \int_\Omega \|\nabla w(x)\|^p_{\mathbb{R}^M}\,\mathrm{d}x\Big)^{1/p} < \infty\,.$$

- Moreover, we recall the definition of the space $BV(\Omega,\mathbb{R}^M)$ from [2], which consists of all measurable functions $w:\Omega\to\mathbb{R}^M$ for which

$$\|w\|_{BV(\Omega,\mathbb{R}^M)} := \|w\|_{L^1(\Omega,\mathbb{R}^M)} + |w|_{BV(\Omega,\mathbb{R}^M)} < \infty,$$

where

$$|w|_{BV(\Omega,\mathbb{R}^M)} := \sup\left\{\int_\Omega w(x)\cdot\mathrm{Div}\varphi(x)\,\mathrm{d}x : \ \varphi\in C^1_c(\Omega,\mathbb{R}^{M\times N}) \text{ s.t. } \|\varphi\|_\infty := \operatorname*{ess\,sup}_{x\in\Omega}\|\varphi(x)\|_F \le 1\right\},$$

where $\|\varphi(x)\|_F$ denotes the Frobenius-norm of the matrix $\varphi(x)$.

Lemma 2.4 *Let $0<s\le 1$ and $p\in[1,\infty)$, then $W^{s,p}(\Omega,K)\subseteq L^p(\Omega,K)$ and the embedding is compact. Moreover, the embedding $BV(\Omega,K)\subseteq L^p(\Omega,K)$ is compact for all* $1\le p<1^* := \begin{cases} +\infty & \text{if } N=1 \\ \frac{N}{N-1} & \text{otherwise}\end{cases}$.

Proof: The first result can be found in [1] for $0<s<1$ and in [27] for $s=1$. The second assertion is stated in [2]. □

Remark 2.5 *Let Assumption 2.1 hold. We recall some basic properties of weak convergence in $W^{s,p}(\Omega,\mathbb{R}^M)$, $W^{1,p}(\Omega,\mathbb{R}^M)$ and weak* convergence in $BV(\Omega,\mathbb{R}^M)$ (see for instance [1, 2, 28, 27]) :*

- *Let $p>1$ and assume that $(w_n)_{n\in\mathbb{N}}$ is bounded in $W^{s,p}(\Omega,\mathbb{R}^M)$. Then there exists a subsequence $(w_{n_k})_{k\in\mathbb{N}}$ which converges weakly in $W^{s,p}(\Omega,\mathbb{R}^M)$.*

- *Assume that $(w_n)_{n\in\mathbb{N}}$ is bounded in $BV(\Omega,\mathbb{R}^M)$. Then there exists a subsequence $(w_{n_k})_{k\in\mathbb{N}}$ which converges weakly* in $BV(\Omega,\mathbb{R}^M)$.*

Before introducing the regularization functional, which we investigate theoretically and numerically, we give the definition of four sets of admissible functions.

Definition 2.6 For $0 < s \le 1$ and $p \ge 1$ we define the sets

$$\begin{aligned} L^p(\Omega,K) &:= \{w \in L^p(\Omega,\mathbb{R}^M) : w(x) \in K \text{ for a.e. } x \in \Omega\};\\ W^{s,p}(\Omega,K) &:= \{w \in W^{s,p}(\Omega,\mathbb{R}^M) : w(x) \in K \text{ for a.e. } x \in \Omega\},\\ W^{1,p}(\Omega,K) &:= \{w \in W^{1,p}(\Omega,\mathbb{R}^M) : w(x) \in K \text{ for a.e. } x \in \Omega\},\\ BV(\Omega,K) &:= \{w \in BV(\Omega,\mathbb{R}^M) : w(x) \in K \text{ for a.e. } x \in \Omega\}. \end{aligned} \tag{2.5}$$

Let $K \subseteq \mathbb{R}^M$ be closed. We consider

- $L^p(\Omega,K) \subseteq L^p(\Omega,\mathbb{R}^M)$ and associate the set $L^p(\Omega,K)$ with the strong $L^p(\Omega,\mathbb{R}^M)$-topology,
- $W^{s,p}(\Omega,K) \subseteq W^{s,p}(\Omega,\mathbb{R}^M)$ with the weak $W^{s,p}(\Omega,\mathbb{R}^M)$-topology, and
- $BV(\Omega,K) \subseteq BV(\Omega,\mathbb{R}^M)$ with the weak* $BV(\Omega,\mathbb{R}^M)$-topology.

Moreover, we define

$$W(\Omega,K) := \begin{cases} W^{s,p}(\Omega,K) & \text{for } p \in (1,\infty) \text{ and } s \in (0,1],\\ BV(\Omega,K) & \text{for } p = 1 \text{ and } s = 1\,. \end{cases} \tag{2.6}$$

Consistently, $W(\Omega,K)$

- is associated with the weak $W^{s,p}(\Omega,\mathbb{R}^M)$-topology in the case $p \in (1,\infty)$ and $s \in (0,1]$ and
- with the weak* $BV(\Omega,\mathbb{R}^M)$-topology when $p = 1$ and $s = 1$.

When we speak about

$$\text{convergence on } W(\Omega,K) \text{ we write } \overset{W}{\to}$$

and mean weak convergence on $W(\Omega,K) = W^{s,p}(\Omega,K)$, $W^{1,p}(\Omega,K)$ and weak* convergence on $BV(\Omega,K)$, respectively.

Remark 2.7
- *In general $L^p(\Omega,K), W^{s,p}(\Omega,K), W^{1,p}(\Omega,K)$ and $BV(\Omega,K)$ are sets which do not form a linear space.*
- *From Lemma 2.4 it follows that $W(\Omega,K) \subseteq L^p(\Omega,K)$.*
- *If $K = \mathbb{S}^1$, then $W^{s,p}(\Omega,K) = W^{s,p}(\Omega,\mathbb{S}^1)$ as introduced in [14].*
- *For an embedded manifold K the dimension of the manifold is not necessarily identical with the space dimension of $\mathbb{R}^M$. For instance if $K = \mathbb{S}^1 \subseteq \mathbb{R}^2$, then the dimension of $\mathbb{S}^1$ is 1 and $M = 2$.*

The following lemma shows that $W(\Omega,K)$ is a closed subset of $W(\Omega,\mathbb{R}^M)$.

Lemma 2.8 (Sequential closedness of $\boldsymbol{W(\Omega,K)}$) *Let $w_* \in W(\Omega,\mathbb{R}^M)$ and $(w_n)_{n\in\mathbb{N}}$ be a sequence in $W(\Omega,K) \subseteq W(\Omega,\mathbb{R}^M)$ with $w_n \overset{W}{\to} w_*$ as $n \to \infty$. Then $w_* \in W(\Omega,K)$ and $w_n \to w_*$ in $L^p(\Omega,K)$. Moreover, there exists a subsequence $(w_{n_k})_{k\in\mathbb{N}}$ which converges to w_* pointwise almost everywhere, i.e. $w_{n_k}(x) \to w_*(x)$ for almost every $x \in \Omega$ as $k \to +\infty$.*

Proof: The proof follows from standard convergence arguments in $W^{s,p}(\Omega,\mathbb{R}^M)$, $BV(\Omega,\mathbb{R}^M)$ and $L^p(\Omega,\mathbb{R}^M)$, respectively. □

In the following we postulate the assumptions on the operator F which will be used throughout the paper:

Assumption 2.9 Let $W(\Omega_1,K_1)$ be as in Equation 2.6 and assume that $\mathrm{F} : W(\Omega_1,K_1) \to L^{p_2}(\Omega_2,K_2)$ is well-defined.

We continue with the definition of our regularization functionals:

Definition 2.10 Let Assumption 2.1 and Assumption 2.9 hold. Moreover, let $\varepsilon > 0$ be fixed and let $\rho := \rho_\varepsilon$ be a mollifier.

The regularization functional $\mathcal{F}_\alpha^v[\mathrm{d}_2, \mathrm{d}_1] : W(\Omega_1, K_1) \to [0, \infty]$ is defined as follows

$$\boxed{\mathcal{F}_\alpha^v[\mathrm{d}_2, \mathrm{d}_1](w) := \int_{\Omega_2} \mathrm{d}_2^{p_2}(\mathrm{F}(w)(x), v(x))\,\mathrm{d}x + \alpha \int_{\Omega_1 \times \Omega_1} \frac{\mathrm{d}_1^{p_1}(w(x), w(y))}{\|x-y\|_{\mathbb{R}^N}^{k+p_1 s}} \rho^l(x-y)\,\mathrm{d}(x,y),} \tag{2.7}$$

where

(i) $v \in L^{p_2}(\Omega_2, K_2)$,

(ii) $s \in [0, 1]$,

(iii) $\alpha \in (0, +\infty)$ is the regularization parameter,

(iv) $l \in \{0, 1\}$ is an indicator and

(v) $\begin{cases} k \le N & \text{if } W(\Omega_1, K_1) = W^{s,p_1}(\Omega_1, K_1),\ 0 < s < 1, \\ k = 0 & \text{if } W(\Omega_1, K_1) = W^{1,p}(\Omega_1, K_1) \text{ or if } W(\Omega_1, K_1) = BV(\Omega_1, K_1),\ \text{respectively.} \end{cases}$

Setting

$$\boxed{[\![\xi, \nu]\!]_{[\mathrm{d}_2]} := \left(\int_{\Omega_2} \mathrm{d}_2^{p_2}(\xi(x), \nu(x))\,\mathrm{d}x \right)^{\frac{1}{p_2}},} \tag{2.8}$$

and

$$\boxed{\mathcal{R}_{[\mathrm{d}_1]}(w) := \int_{\Omega_1 \times \Omega_1} \frac{\mathrm{d}_1^{p_1}(w(x), w(y))}{\|x-y\|_{\mathbb{R}^N}^{k+p_1 s}} \rho^l(x-y)\,\mathrm{d}(x,y),} \tag{2.9}$$

Equation 2.7 can be expressed in compact form

$$\boxed{\mathcal{F}_\alpha^v[\mathrm{d}_2, \mathrm{d}_1](w) = [\![\mathrm{F}(w), v]\!]_{[\mathrm{d}_2]}^{p_2} + \alpha \mathcal{R}_{[\mathrm{d}_1]}(w).} \tag{2.10}$$

For convenience we will often skip some of the super or subscript, and use compact notations like e.g.

$$\mathcal{F}^v, \mathcal{F}[\mathrm{d}_2, \mathrm{d}_1] \text{ or } \mathcal{F}(w) = [\![\mathrm{F}(w), v]\!]^{p_2} + \alpha \mathcal{R}(w).$$

Remark 2.11 *(i)* $l = \{0, 1\}$ *is an indicator which allows to consider approximations of Sobolev semi-norms and double integral representations of the type of Bourgain, Brézis, and Mironescu [13] in a uniform manner.*

- *when* $k = 0$, $s = 1$, $l = 1$ *and when* d_1 *is the Euclidean distance, we get the double integrals of the Bourgain, Brézis, and Mironescu-form [13]. Compare with Equation 1.3.*
- *When* d_1 *is the Euclidean distance,* $k = N$ *and* $l = 0$*, we get Sobolev semi-norms.*

We expect a relation between the two classes of functionals for $l = 0$ *and* $l = 1$ *as stated in Subsection 5.2.*

(ii) The second term in Equation 2.7 is similar to the one used in [3, 11] and [13, 47, 23].

In the following we analyze basic properties of $[\![\cdot, \cdot]\!]_{[\mathrm{d}_2]}$ and the functional $\mathcal{F}$.

Proposition 2.12 *Let Assumption 2.1 hold.*

(i) Then the mapping $[\![\cdot, \cdot]\!]_{[\mathrm{d}_2]} : L^{p_2}(\Omega_2, K_2) \times L^{p_2}(\Omega_2, K_2) \to [0, +\infty]$ *satisfies the metric axioms.*

(ii) *Let, in addition, Assumption 2.9 hold, assume that $v \in L^{p_2}(\Omega_2, K_2)$ and that both metrics d_i, $i = 1, 2$, are equivalent to $\mathrm{d}_{\mathbb{R}^{M_i}}|_{K_i \times K_i}$, respectively. Then the functional $\mathcal{F}_\alpha^v[\mathrm{d}_2, \mathrm{d}_1]$ defined on $W(\Omega_1, K_1) \neq \emptyset$ does not attain the value $+\infty$.*

Proof: *(i)* We show that $[\![\cdot,\cdot]\!]_{[\mathrm{d}_2]}$ satisfies the metric axioms. Non-negativity, positive-definiteness and symmetry follow since d_2 is a metric. To prove the triangle inequality let $u, \xi, w \in L^{p_2}(\Omega_2, K_2)$. Then by application of Hölder's inequality we get:

$$\begin{aligned}
[\![u, w]\!]_{[\mathrm{d}_2]}^{p_2} &= \int_{\Omega_2} \mathrm{d}_2^{p_2}(u(x), w(x))\,\mathrm{d}x \\
&\leq \int_{\Omega_2} \mathrm{d}_2(u(x), \xi(x))\mathrm{d}_2^{p_2-1}(u(x), w(x))\,\mathrm{d}x + \int_{\Omega_2} \mathrm{d}_2(\xi(x), w(x))\mathrm{d}_2^{p_2-1}(u(x), w(x))\,\mathrm{d}x \\
&\leq \left(\left(\int_{\Omega_2} \mathrm{d}_2^{p_2}(u(x), \xi(x))\,\mathrm{d}x \right)^{\frac{1}{p_2}} + \left(\int_{\Omega_2} \mathrm{d}_2^{p_2}(\xi(x), w(x))\,\mathrm{d}x \right)^{\frac{1}{p_2}} \right) \\
&\quad * \left(\left(\int_{\Omega_2} \mathrm{d}_2^{p_2}(u(x), w(x))\,\mathrm{d}x \right)^{\frac{p_2-1}{p_2}} \right) \\
&= \left([\![u, \xi]\!]_{[\mathrm{d}_2]} + [\![\xi, w]\!]_{[\mathrm{d}_2]} \right) [\![u, w]\!]_{[\mathrm{d}_2]}^{p_2-1}.
\end{aligned}$$

(ii) We emphasize that $W(\Omega_1, K_1) \neq \emptyset$ because the function $w = \text{const} \in K_1$ belongs to all spaces, $W^{s,p}(\Omega, K_1)$ for $p_1 \in (1, \infty)$ and $s \in (0, 1]$ and $BV(\Omega, K_1)$ for $p_1 = 1$ and $s = 1$.

Assume now that the metrics d_i are equivalent to $\mathrm{d}_{\mathbb{R}^{M_i}}|_{K_i \times K_i}$ for $i = 1$ and $i = 2$, respectively. Hence we especially have an upper bound $d_i \leq C\mathrm{d}_{\mathbb{R}^{M_i}}|_{K_i \times K_i}$. We need to prove that $\mathcal{F}_\alpha^v[\mathrm{d}_2, \mathrm{d}_1](w) < \infty$ for every $w \in W(\Omega_1, K_1)$.

- For $W(\Omega_1, K_1) = BV(\Omega_1, K_1)$ this is guaranteed by [47, Theorem 1.2].
- For $W(\Omega_1, K_1) = W^{1,p}(\Omega_1, K_1)$ by [13, Theorem 1].
- For $W(\Omega_1, K_1) = W^{s,p_1}(\Omega_1, K_1)$ we distinguish between two cases. If $\|x - y\|_{\mathbb{R}^N} < 1$ we have that $\frac{1}{\|x-y\|_{\mathbb{R}^N}^{k+p_1 s}} \leq \frac{1}{\|x-y\|_{\mathbb{R}^N}^{N+p_1 s}}$ for $k \leq N$ and hence $\mathcal{R}_{[\mathrm{d}_1]}(w) \leq \|\rho\|_\infty |w|_{W^{s,p_1}(\Omega_1, \mathbb{R}^M)}^{p_1} < \infty$. If $\|x - y\|_{\mathbb{R}^N} \geq 1$ we can estimate $\mathcal{R}_{[\mathrm{d}_1]}(w) \leq \|\rho\|_\infty 2^{p_1} |\Omega| \, \|w\|_{L^{p_1}(\Omega_1, \mathbb{R}^M)}^{p_1} < \infty$. □

3. EXISTENCE

In order to prove existence of a minimizer of the functional $\mathcal{F}$ we apply the Direct Method in the Calculus of Variations (see e.g. [21, 20]). To apply these results we verify auxiliary results such as the closedness of $W(\Omega_1, K_1)$ and continuity properties of $[\![\cdot,\cdot]\!]_{[\mathrm{d}_2]}$, $\mathcal{R}_{[\mathrm{d}_1]}$ and $\mathcal{F}[\mathrm{d}_2, \mathrm{d}_1]$.

We will use the following assumption:

Assumption 3.1 Let Assumption 2.1 hold and let $W(\Omega_1, K_1)$ and the associated topology be as defined in Equation 2.6.

In addition we assume

- that $\mathrm{F} : W(\Omega_1, K_1) \to L^{p_2}(\Omega_2, K_2)$ is well-defined and sequentially continuous with respect to the specified topology on $W(\Omega_1, K_1)$ and
- that for every $t > 0$ and $\alpha > 0$ the level sets

$$\mathrm{level}_t(\mathcal{F}_\alpha^v[\mathrm{d}_2, \mathrm{d}_1]) := \{w \in W(\Omega_1, K_1) \ : \ \mathcal{F}_\alpha^v[\mathrm{d}_2, \mathrm{d}_1] \leq t\} \tag{3.1}$$

are sequentially pre-compact subsets of $W(\Omega_1, \mathbb{R}^{M_1})$.

Lemma 3.2 *Let Assumption 3.1 hold. Then the mappings $[\![\cdot,\cdot]\!]_{[\mathrm{d}_2]}$, $\mathcal{R}_{[\mathrm{d}_1]}$ and $\mathcal{F}[\mathrm{d}_2, \mathrm{d}_1]$ have the following continuity properties:*

(i) *The mapping* $[\![\cdot,\cdot]\!]_{[\mathrm{d}_2]} : L^{p_2}(\Omega_2, K_2) \times L^{p_2}(\Omega_2, K_2)] \to [0, +\infty]$ *is sequentially lower semi-continuous in both arguments, i.e. whenever sequences* $(w_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$ *in* $L^{p_2}(\Omega_2, K_2)$ *converge to* $w_* \in L^{p_2}(\Omega_2, K_2)$ *and* $v_* \in L^{p_2}(\Omega_2, K_2)$*, respectively, we have* $[\![w_*, v_*]\!]_{[\mathrm{d}_2]} \le \liminf_{n\to+\infty} [\![w_n, v_n]\!]_{[\mathrm{d}_2]}$.

(ii) *The functional* $\mathcal{R}_{[\mathrm{d}_1]} : W(\Omega_1, K_1) \to [0, \infty]$ *is sequentially lower semi-continuous, i.e. whenever a sequence* $(w_n)_{n\in\mathbb{N}}$ *in* $W(\Omega_1, K_1)$ *converges to some* $w_* \in W(\Omega_1, K_1)$ *we have*

$$\mathcal{R}_{[\mathrm{d}_1]}(w_*) \le \liminf_{n\to+\infty} \mathcal{R}_{[\mathrm{d}_1]}(w_n).$$

(iii) *The functional* $\mathcal{F}[\mathrm{d}_2, \mathrm{d}_1] : W(\Omega_1, K_1) \to [0, \infty]$ *is sequentially lower semi-continuous.*

Proof: *(i)* It is sufficient to show that for every pair of sequences $(w_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$ in $L^{p_2}(\Omega_2, K_2)$ which converge to $w_* \in L^{p_2}(\Omega_2, K_2)$ and $v_* \in L^{p_2}(\Omega_2, K_2)$, respectively, we can extract subsequences $(w_{n_j})_{j\in\mathbb{N}}$ and $(v_{n_j})_{j\in\mathbb{N}}$, respectively, with

$$[\![w_*, v_*]\!]_{[\mathrm{d}_2]} \le \liminf_{j\to+\infty} [\![w_{n_j}, v_{n_j}]\!]_{[\mathrm{d}_2]}.$$

To this end let $(w_n)_{n\in\mathbb{N}}, (v_n)_{n\in\mathbb{N}}$ be some sequences in $L^{p_2}(\Omega_2, K_2)$ with $w_n \to w_*$ and $v_n \to v_*$ in $L^{p_2}(\Omega_2, K_2)$. Lemma 2.8 ensures that there exist subsequences $(w_{n_j})_{j\in\mathbb{N}}, (v_{n_j})_{j\in\mathbb{N}}$ converging to w_* and v_* pointwise almost everywhere, which in turn implies $\big(w_{n_j}(\cdot), v_{n_j}(\cdot)\big) \to \big(w_*(\cdot), v_*(\cdot)\big)$ pointwise almost everywhere. Thereform, together with the continuity of $\mathrm{d}_2 : K_2 \times K_2 \to [0, \infty)$ with respect to $\mathrm{d}_{\mathbb{R}^{M_2}}$, cf. Section 2, we obtain by using the quadrangle inequality that for all $x \in \Omega_2$

$$|\mathrm{d}_2(w_{n_j}(x), v_{n_j}(x)) - \mathrm{d}_2(w_*(x), v_*(x))| \le \mathrm{d}_2(w_{n_j}(x), w_*(x)) + \mathrm{d}_2(v_{n_j}(x), v_*(x)) \to 0,$$

and hence

$$\mathrm{d}_2^{p_2}\big(w_{n_j}(x), v_{n_j}(x)\big) \to \mathrm{d}_2^{p_2}\big(w_*(x), v_*(x)\big) \text{ for almost every } x \in \Omega_2.$$

Applying Fatou's lemma we obtain

$$[\![w_*, v_*]\!]_{[\mathrm{d}_2]} = \int_{\Omega_2} \mathrm{d}_2^{p_2}(w_*(x), v_*(x))\,\mathrm{d}x \le \liminf_{j\to+\infty} \int_{\Omega_2} \mathrm{d}_2^{p_2}(w_{n_j}(x), v_{n_j}(x))\,\mathrm{d}x = \liminf_{j\to+\infty} [\![w_{n_j}, v_{n_j}]\!]_{[\mathrm{d}_2]}.$$

(ii) Let $(w_n)_{n\in\mathbb{N}}$ be a sequence in $W(\Omega_1, K_1)$ with $w_n \overset{W}{\to} w_*$ as $n \to +\infty$. From Lemma 2.8 it then follows that there exists a subsequence $w_{n_j} \to w_*$ with respect to the L^{p_1}-norm and therefore this subsequence is also converging pointwise almost everywhere. This further implies that

$$\mathrm{d}_1^{p_1}\big(w_{n_j}(x), w_{n_j}(y)\big) \to \mathrm{d}_1^{p_1}\big(w_*(x), w_*(y)\big)$$

for almost every

$$(x, y) \in \Omega_1 \times \Omega_1 \supseteq \{(x, y) \in \Omega_1 \times \Omega_1 : x \ne y\} =: A. \tag{3.2}$$

Defining

$$f_j(x, y) := \begin{cases} \dfrac{\mathrm{d}_1^{p_1}(w_{n_j}(x), w_{n_j}(y))}{\|x-y\|_{\mathbb{R}^N}^{k+ps}} \rho^l(x-y) & \text{for} \quad (x, y) \in A, \\ 0 & \text{for} \quad (x, y) \in (\Omega_1 \times \Omega_1) \setminus A, \end{cases} \qquad \text{for all } j \in \mathbb{N}.$$

and

$$f_*(x, y) := \begin{cases} \dfrac{\mathrm{d}_1^{p_1}(w_*(x), w_*(y))}{\|x-y\|_{\mathbb{R}^N}^{k+ps}} \rho^l(x-y) & \text{for} \quad (x, y) \in A, \\ 0 & \text{for} \quad (x, y) \in (\Omega_1 \times \Omega_1) \setminus A \end{cases}$$

we thus have $f_*(x, y) = \lim_{j\to+\infty} f_j(x, y)$ for almost every $(x, y) \in \Omega_1 \times \Omega_1$. Applying Fatou's lemma to the functions f_j yields the assertion.

(iii) It is sufficient to prove that the components $\mathcal{G}(\cdot) = [\![F(\cdot), v]\!]_{[\mathrm{d}_2]}$ and $\mathcal{R} = \mathcal{R}_{[\mathrm{d}_1]}$ of $\mathcal{F}[\mathrm{d}_1, \mathrm{d}_2] = \mathcal{G} + \alpha\mathcal{R}$ are sequentially lower semi-continuous. To prove that $\mathcal{G}$ is sequentially lower semi-continuous in every $w_* \in W(\Omega_1, K_1)$ let $(w_n)_{n\in\mathbb{N}}$ be a sequence in $W(\Omega_1, K_1)$ with $w_n \overset{W}{\to} w_*$ as $n \to +\infty$. Assumption 3.1, ensuring the sequential continuity of $\mathrm{F} : W(\Omega_1, K_1) \to L^{p_2}(\Omega_2, K_2)$, implies hence $\mathrm{F}(w_n) \to \mathrm{F}(w_*)$ in $L^{p_2}(\Omega_2, K_2)$ as $n \to +\infty$. By item *(i)* we thus obtain $\mathcal{G}(w_*) = [\![\mathrm{F}(w_*), v]\!] \le \liminf_{n\to+\infty} [\![\mathrm{F}(w_n), v]\!] = \liminf_{n\to+\infty} \mathcal{G}(w_n)$.
$\mathcal{R}$ is sequentially lower semi-continuous by item *(ii)*. □

3.1. Existence of minimizers. The proof of well–posedness of a minimizer of $\mathcal{F}[\mathrm{d}_2, \mathrm{d}_1]$ is along the lines of the proof in [53]. However, it differs slightly because the underlying set W is not a linear space.

Theorem 3.3 *Let Assumption 3.1 hold. Then the functional* $\mathcal{F}[\mathrm{d}_2, \mathrm{d}_1] : W(\Omega_1, K_1) \to [0, \infty]$ *attains a minimizer.*

Proof: We prove the existence of a minimizer via the Direct Method. We shortly write $\mathcal{F}$ for $\mathcal{F}[\mathrm{d}_2, \mathrm{d}_1]$. Let $(w_n)_{n\in\mathbb{N}}$ be a sequence in $W(\Omega_1, K_1)$ with

$$\lim_{n\to+\infty} \mathcal{F}(w_n) = \inf_{w\in W(\Omega_1,K_1)} \mathcal{F}(w) < \infty. \tag{3.3}$$

In particular there is some $c \in \mathbb{R}$ such that $\mathcal{F}(w_n) \le c$ for every $n \in \mathbb{N}$. Since the level set $\mathrm{level}_c(\mathcal{F})$ is pre-compact with respect to the topology given to $W(\Omega_1, \mathbb{R}^{M_1})$ we get the existence of a subsequence $(w_{n_k})_{k\in\mathbb{N}}$ which converges to some $w_* \in W(\Omega_1, \mathbb{R}^{M_1})$, where actually $w_* \in W(\Omega_1, K_1)$ due to Lemma 2.8. Because $\mathcal{F}$ is sequentially lower semi-continuous, see Lemma 3.2, we have $\mathcal{F}(w_*) \le \liminf_{k\to+\infty} \mathcal{F}(w_{n_k})$. Combining this with Equation 3.3 we obtain

$$\inf_{w\in W(\Omega_1,K_1)} \mathcal{F}(w) \le \mathcal{F}(w_*) \le \liminf_{k\to+\infty} \mathcal{F}(w_{n_k}) = \lim_{n\to+\infty} \mathcal{F}(w_n) = \inf_{w\in W(\Omega_1,K_1)} \mathcal{F}(w).$$

In particular $\mathcal{F}(w_*) = \inf_{w\in W(\Omega_1,K_1)} \mathcal{F}(w)$, meaning that w_* is a minimizer of $\mathcal{F}$. □

In the following we investigate two examples, which are relevant for the numerical examples in Section 6.

Example 3.4 *We consider that* $W(\Omega_1, K_1) = W^{s,p_1}(\Omega_1, K_1)$ *with* $p_1 > 1,\ 0 < s < 1$ *and fix* $k = N$. *Here we assume that the operator* F *is norm-coercive in the sense that the implication*

$$\|w_n\|_{L^{p_1}(\Omega_1,\mathbb{R}^{M_1})} \to +\infty \Rightarrow \|\mathrm{F}(w_n)\|_{L^{p_2}(\Omega_2,\mathbb{R}^{M_2})} \to +\infty \tag{3.4}$$

holds true for every sequence $(w_n)_{n\in\mathbb{N}}$ *in* $W^{s,p_1}(\Omega_1, K_1) \subseteq W^{s,p_1}(\Omega_1, \mathbb{R}^{M_1})$, *then the functional*

$$\mathcal{F}[\mathrm{d}_2, \mathrm{d}_1] = [\![\mathrm{F}(w), v]\!]^{p_2}_{[\mathrm{d}_2]} + \alpha \mathcal{R}_{[\mathrm{d}_1]}(w) : W^{s,p_1}(\Omega_1, K_1) \to [0, \infty]$$

is coercive. This can be seen as follows:

The assumption of d_1 *and* $\mathrm{d}_{\mathbb{R}^{M_1}}|_{K_1\times K_1}$ *resp.* d_2 *and* $\mathrm{d}_{\mathbb{R}^{M_2}}|_{K_2\times K_2}$ *carries over to* $\mathcal{F}[\mathrm{d}_2, \mathrm{d}_1]$ *and* $\mathcal{F}[\mathrm{d}_{\mathbb{R}^{M_2}}|_{K_2\times K_2}, \mathrm{d}_{\mathbb{R}^{M_1}}|_{K_1\times K_1}]$: *in particular there exists a constant* $c > 0$ *such that*

$$\mathcal{F}[\mathrm{d}_2, \mathrm{d}_1](w) \ge c\mathcal{F}[\mathrm{d}_{\mathbb{R}^{M_2}}|_{K_2\times K_2}, \mathrm{d}_{\mathbb{R}^{M_1}}|_{K_1\times K_1}](w) \text{ for all } w \in W^{s,p_1}(\Omega_1, K_1).$$

Thus it is sufficient to show that $\mathcal{F}[\mathrm{d}_{\mathbb{R}^{M_2}}|_{K_2\times K_2}, \mathrm{d}_{\mathbb{R}^{M_1}}|_{K_1\times K_1}] : W^{s,p_1}(\Omega_1, K_1) \to [0, \infty]$ *is coercive: To prove this we write shortly* $\mathcal{F}$ *instead of* $\mathcal{F}[\mathrm{d}_{\mathbb{R}^{M_2}}|_{K_2\times K_2}, \mathrm{d}_{\mathbb{R}^{M_1}}|_{K_1\times K_1}]$ *and consider sequences* $(w_n)_{n\in\mathbb{N}}$ *in* $W^{s,p_1}(\Omega_1, K_1)$ *with* $\|w_n\|_{W^{s,p_1}(\Omega_1,\mathbb{R}^{M_1})} \to +\infty$ *as* $n \to +\infty$. *We show that* $\mathcal{F}(w_n) \to +\infty$, *as* $n \to +\infty$. *Since*

$$\|w_n\|_{W^{s,p_1}(\Omega_1,\mathbb{R}^{M_1})} = \left(\|w_n\|^{p_1}_{L^{p_1}(\Omega_1,\mathbb{R}^{M_1})} + |w_n|^{p_1}_{W^{s,p_1}(\Omega_1,\mathbb{R}^{M_1})} \right)^{\frac{1}{p_1}}$$

we can prove this by investigating two different cases $\|w_n\|_{L^{p_1}(\Omega_1,\mathbb{R}^{M_1})} \to +\infty$ *and* $|w_n|_{W^{s,p_1}(\Omega_1,\mathbb{R}^{M_1})} \to +\infty$.

Case 1 $\|w_n\|_{L^{p_1}(\Omega_1,\mathbb{R}^{M_1})} \to +\infty$.
The inverse triangle inequality and the norm-coercivity of F, *Equation 3.4, give* $\|\mathrm{F}(w_n) - v\|_{L^{p_2}(\Omega_2,\mathbb{R}^{M_2})} \ge \|\mathrm{F}(w_n)\|_{L^{p_2}(\Omega_2,\mathbb{R}^{M_2})} - \|v\|_{L^{p_2}(\Omega_2,\mathbb{R}^{M_2})} \to +\infty$. *Therefore also*

$$\mathcal{F}(w_n) = \|\mathrm{F}(w_n) - v\|^{p_2}_{L^{p_2}(\Omega_2,\mathbb{R}^{M_2})} + \alpha \int\limits_{\Omega_1\times\Omega_1} \frac{\|w_n(x) - w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}} \rho^l(x-y)\, \mathrm{d}(x,y) \to +\infty.$$

Case 2 $|w_n|_{W^{s,p_1}(\Omega_1,\mathbb{R}^{M_1})} \to +\infty$.
If $l = 0$, *then* $\mathcal{R}_{[\mathrm{d}_1]}$ *is exactly the* $W^{s,p_1}(\Omega_1, \mathbb{R}^{M_1})$*-semi-norm* $|w|_{W^{s,p_1}(\Omega_1,\mathbb{R}^{M_1})}$ *and we trivially get the desired result.*

Hence we assume from now on that $l = 1$*. The assumptions on* ρ *ensure that there exists a* $\tau > 0$ *and* $1 \geq \eta_\tau > 0$ *such that*

$$\begin{aligned}\mathcal{S}_\tau &:= \{(x,y) \in \Omega_1 \times \Omega_1 : \rho(x-y) \geq \tau\} \\ &= \{(x,y) \in \Omega_1 \times \Omega_1 : \|x-y\|_{\mathbb{R}^N} \leq \eta_\tau\},\end{aligned}$$

cf. Figure 1.

Splitting $\Omega_1 \times \Omega_1$ *into* $\mathcal{S}_\tau =: \mathcal{S}$ *and its complement* $(\Omega_1 \times \Omega_1) \setminus \mathcal{S}_\tau =: \mathcal{S}^c$ *we accordingly split the integrals* $|w_n|_{W^{s,p_1}(\Omega_1,\mathbb{R}^{M_1})} = \int\limits_{\Omega_1\times\Omega_1} \frac{\|w_n(x)-w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}} \,\mathrm{d}(x,y)$ *and consider again two cases* $\int\limits_{\mathcal{S}} \frac{\|w_n(x)-w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}} \,\mathrm{d}(x,y) \to +\infty$ *and* $\int\limits_{\mathcal{S}^c} \frac{\|w_n(x)-w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}} \,\mathrm{d}(x,y) \to +\infty$*, respectively.*

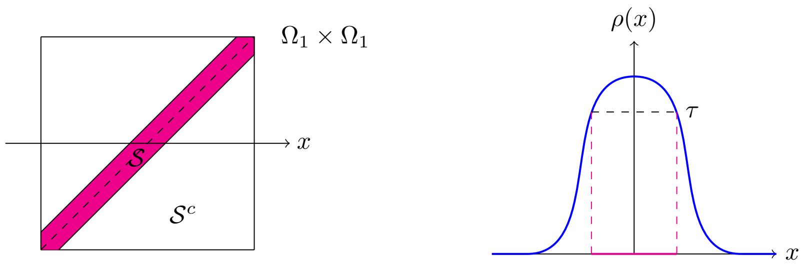


Figure 1. *The stripe* $\mathcal{S} = \mathcal{S}_\tau$ *if* Ω_1 *is an open interval and its connection to the radial mollifier* ρ *for fixed* $y \in \Omega_1$*.*

Case 2.1 $\int\limits_{\mathcal{S}} \frac{\|w_n(x)-w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}} \,\mathrm{d}(x,y) \to +\infty.$

By definition of $\mathcal{S}$ *we have* $\rho(x-y) \geq \tau > 0$ *for all* $(x,y) \in \mathcal{S}$*. Therefore*

$$\int\limits_{\mathcal{S}} \frac{\|w_n(x)-w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}} \rho(x-y) \,\mathrm{d}(x,y) \geq \tau \int\limits_{\mathcal{S}} \frac{\|w_n(x)-w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}} \,\mathrm{d}(x,y) \to +\infty.$$

Since $\alpha > 0$ *it therefore follows*

$$\begin{aligned}\mathcal{F}(w_n) = \|\mathrm{F}(w_n) - v\|^{p_2}_{L^{p_2}(\Omega_2,\mathbb{R}^{M_2})} + \underbrace{\alpha \int\limits_{\mathcal{S}} \frac{\|w_n(x)-w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}} \rho(x-y) \,\mathrm{d}(x,y)}_{\to+\infty} \\ + \underbrace{\alpha \int\limits_{\mathcal{S}^c} \frac{\|w_n(x)-w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}} \rho(x-y) \,\mathrm{d}(x,y)}_{\geq 0} \to +\infty.\end{aligned}$$

Case 2.2 $\int\limits_{\mathcal{S}^c} \frac{\|w_n(x)-w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}} \,\mathrm{d}(x,y) \to +\infty.$

For $(x,y) \in \mathcal{S}^c$ *it might happen that* $\rho(x-y) = 0$*, and thus instead of proving* $\mathcal{F}(w_n) \geq \int\limits_{\mathcal{S}^c} \frac{\|w_n(x)-w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}} \rho(x-y) \,\mathrm{d}(x,y) \to +\infty$ *,as in Case 2.1, we rather show that* $\mathcal{F}(w_n) \geq \|\mathrm{F}(w_n) - v\|^{p_2}_{L^{p_2}(\Omega_2,\mathbb{R}^{M_2})} \to +\infty$*. For this it is sufficient to show that for every* $c > 0$ *there is some* $C \in \mathbb{R}$ *such that the implication*

$$\|\mathrm{F}(w) - v\|^{p_2}_{L^{p_2}(\Omega_2,\mathbb{R}^{M_2})} \leq c \implies \int\limits_{\mathcal{S}^c} \frac{\|w(x)-w(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}} \,\mathrm{d}(x,y) \leq C,$$

holds true for all $w \in W^{s,p_1}(\Omega_1, K_1) \subseteq W^{s,p_1}(\Omega_1, \mathbb{R}^{M_1})$*. To this end let* $c > 0$ *be given and consider an arbitrarily chosen* $w \in W^{s,p_1}(\Omega_1, K_1)$ *fulfilling* $\|\mathrm{F}(w) - v\|^{p_2}_{L^{p_2}(\Omega_2,\mathbb{R}^{M_2})} \leq c$.

Then $\|\mathrm{F}(w)-v\|_{L^{p_2}(\Omega_2,\mathbb{R}^{M_2})} \leq \sqrt[p_2]{c}$*. Using the triangle inequality and the monotonicity of the function* $h : t \mapsto t^{p_2}$ *on* $[0,+\infty)$ *we get further*

$$\begin{aligned}\|\mathrm{F}(w)\|^{p_2}_{L^{p_2}(\Omega_2,\mathbb{R}^{M_2})} &= \|\mathrm{F}(w)-v+v\|^{p_2}_{L^{p_2}(\Omega_2,\mathbb{R}^{M_2})}\\ &\leq \Big(\|\mathrm{F}(w)-v\|_{L^{p_2}(\Omega_2,\mathbb{R}^{M_2})} + \|v\|_{L^{p_2}(\Omega_2,\mathbb{R}^{M_2})}\Big)^{p_2}\\ &\leq \big(\sqrt[p_2]{c} + \|v\|_{L^{p_2}(\Omega_2,\mathbb{R}^{M_2})}\big)^{p_2} =: \tilde{c}. \end{aligned} \tag{3.5}$$

Due to the norm-coercivity, it thus follows that $\|w\|_{L^{p_1}(\Omega_1,\mathbb{R}^{M_1})} \leq \bar{c}$, $\bar{c}$ *some constant. Using [53, Lemma 3.20] it then follows that*

$$\|w(x)-w(y)\|^{p_1}_{\mathbb{R}^{M_1}} \leq 2^{p_1-1}\|w(x)\|^{p_1}_{\mathbb{R}^{M_1}} + 2^{p_1-1}\|w(y)\|^{p_1}_{\mathbb{R}^{M_1}} \tag{3.6}$$

for all $(x,y) \in \Omega_1 \times \Omega_1$*. Using Equation 3.6, Fubini's Theorem and Equation 3.5 we obtain*

$$\begin{aligned}\int_{\Omega_1\times\Omega_1} \|w(x)-w(y)\|^{p_1}_{\mathbb{R}^{M_1}}\,\mathrm{d}(x,y) &\leq \int_{\Omega_1\times\Omega_1} 2^{p_1-1}\|w(x)\|^{p_1}_{\mathbb{R}^{M_1}} + 2^{p_1-1}\|w(y)\|^{p_1}_{\mathbb{R}^{M_1}}\,\mathrm{d}(x,y)\\ &= |\Omega_1|\int_{\Omega_1} 2^{p_1-1}\|w(x)\|^{p_1}_{\mathbb{R}^{M_1}}\,\mathrm{d}x + |\Omega_1|\int_{\Omega_1} 2^{p_1-1}\|w(y)\|^{p_1}_{\mathbb{R}^{M_1}}\,\mathrm{d}y\\ &= 2\,|\Omega_1|\int_{\Omega_1} 2^{p_1-1}\|w(x)\|^{p_1}_{\mathbb{R}^{M_1}}\,\mathrm{d}x\\ &= 2^{p_1}\,|\Omega_1|\;\|w\|^{p_1}_{L^{p_1}(\Omega_1,\mathbb{R}^{M_1})} \leq 2^{p_1}\,|\Omega_1|\,\bar{c}^{p_1}.\end{aligned}$$

Combining $\|x-y\|_{\mathbb{R}^N} \geq \eta_\tau > 0$ *for all* $(x,y) \in \mathcal{S}^c$ *with the previous inequality we obtain the needed estimate*

$$\begin{aligned}\int_{\mathcal{S}^c} \frac{\|w(x)-w(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}}\,\mathrm{d}(x,y) &\leq \frac{1}{\eta_\tau^{N+p_1 s}}\int_{\mathcal{S}^c} \|w(x)-w(y)\|^{p_1}_{\mathbb{R}^{M_1}}\,\mathrm{d}(x,y)\\ &\leq \frac{1}{\eta_\tau^{N+p_1 s}}\int_{\Omega_1\times\Omega_1} \|w(x)-w(y)\|^{p_1}_{\mathbb{R}^{M_1}}\,\mathrm{d}(x,y)\\ &\leq \frac{2^{p_1}\,|\Omega_1|\,\bar{c}^{p_1}}{\eta_\tau^{N+p_1 s}} =: C.\end{aligned}$$

The second example concerns the analysis coercivity of $\mathcal{F}[\mathrm{d}_2,\mathrm{d}_1]$, defined in Equation 2.10, when F denotes the *inpainting operator*. To prove this result we require the following auxiliary lemma:

Lemma 3.5 *There exists a constant* $C \in \mathbb{R}$ *such that for all* $w \in W^{s,p_1}(\Omega_1,\mathbb{R}^{M_1})$, $0 < s < 1$, $l \in \{0,1\}$, $1 < p_1 < \infty$ *and* $D \subseteq \Omega_1$

$$\|w\|_{L^{p_1}(D,\mathbb{R}^{M_1})} \leq C\left(\|w\|_{L^{p_1}(\Omega_1\setminus D,\mathbb{R}^{M_1})} + \int_{\Omega_1\times\Omega_1} \frac{\|w(x)-w(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}}\rho^l(x-y)\,\mathrm{d}(x,y)\right). \tag{3.7}$$

Proof: The proof is inspired by the proof of Poincaré's inequality in [27]. It is included here for the sake of completeness.

Assume first that $l = 1$. Let $\mathcal{S}$ be as above,

$$\begin{aligned}\mathcal{S} &:= \{(x,y) \in \Omega_1\times\Omega_1 : \rho(x-y) \geq \tau\}\\ &= \{(x,y) \in \Omega_1\times\Omega_1 : \|x-y\|_{\mathbb{R}^N} \leq \eta\}.\end{aligned}$$

If the stated inequality Equation 3.7 would be false, then for every $n \in \mathbb{N}$ there would exists a function $w_n \in W^{s,p_1}(\Omega_1,\mathbb{R}^{M_1})$ satisfying

$$\|w_n\|_{L^{p_1}(D,\mathbb{R}^{M_1})} \geq n\big(\|w_n\|_{L^{p_1}(\Omega_1\setminus D,\mathbb{R}^{M_1})} + \int_{\Omega_1\times\Omega_1} \frac{\|w_n(x)-w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}}\rho(x-y)\,\mathrm{d}(x,y)\big).$$

By normalizing we can assume without loss of generality *(i)* $\|w_n\|_{L^{p_1}(D,\mathbb{R}^{M_1})} = 1$,

(ii) $\|w_n\|_{L^{p_1}(\Omega_1\setminus D,\mathbb{R}^{M_1})} < \frac{1}{n}$,

(iii) $\int\limits_{\Omega_1\times\Omega_1} \frac{\|w_n(x)-w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}}\rho(x-y)\,\mathrm{d}(x,y) < \frac{1}{n}$.

By item *(i)* and item *(ii)* we get that $\|w_n\|_{L^{p_1}(\Omega_1,\mathbb{R}^{M_1})} = \|w_n\|_{L^{p_1}(D,\mathbb{R}^{M_1})} + \|w_n\|_{L^{p_1}(\Omega_1\setminus D,\mathbb{R}^{M_1})} < 1+\frac{1}{n} < 2$ is bounded. Moreover

$$
\begin{aligned}
|w_n|^{p_1}_{W^{s,p_1}(\Omega_1,\mathbb{R}^{M_1})} &= \int\limits_{\mathcal{S}} \frac{\|w_n(x)-w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}}\,\mathrm{d}(x,y) + \int\limits_{\mathcal{S}^c} \frac{\|w_n(x)-w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}}\,\mathrm{d}(x,y)\\
&\le \frac{1}{\tau}\int\limits_{\mathcal{S}} \frac{\|w_n(x)-w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}}\rho(x-y)\,\mathrm{d}(x,y) + \frac{2_1^p\,|\Omega_1|}{\eta^{N+p_1 s}}\,\|w_n\|^{p_1}_{L^{p_1}(\Omega_1,\mathbb{R}^{M_1})}\\
&< \frac{1}{\tau n} + \frac{2^{p_1+1}\,|\Omega_1|}{\eta^{N+p_1 s}} =: c < \infty,
\end{aligned}
$$

where c is independent of n. This yields that the sequence $(w_n)_{n\in\mathbb{N}}$ is bounded in $W^{s,p_1}(\Omega_1,\mathbb{R}^{M_1})$ by $(2^{p_1}+c)^{\frac{1}{p_1}}$. By Lemma 2.8 there exists a subsequence $(w_{n_k})_{k\in\mathbb{N}}$ of $(w_n)_{n\in\mathbb{N}}$ such that $w_{n_k}(x)\to w^*(x)$ converges pointwise almost everywhere to some $w^*\in W^{s,p_1}(\Omega_1,\mathbb{R}^{M_1})$.

Using the continuity of the norm and dominated convergence we obtain

(i) $\|w^*\|_{L^{p_1}(D,\mathbb{R}^{M_1})} = 1$, in particular w^* is not the null-function on D,

(ii) $\|w^*\|_{L^{p_1}(\Omega_1\setminus D,\mathbb{R}^{M_1})} = 0$ since $n\in\mathbb{N}$ is arbitrary and hence $w^*\equiv 0$ on $\Omega_1\setminus D$.

(iii)

$$
\frac{1}{n} > \int\limits_{\mathcal{S}} \frac{\|w_n(x)-w_n(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}}\rho(x-y)\,\mathrm{d}(x,y) \ge \frac{\tau}{\eta^{N+p_1 s}}\int\limits_{\mathcal{S}} \|w^*(x)-w^*(y)\|^{p_1}_{\mathbb{R}^{M_1}},
$$

i.e. $w^*(x)=w^*(y)$ for $\|x-y\|_{\mathbb{R}^N}\le\eta$ yielding that w^* is even constant since Ω_1 is connected,

which gives the contradiction.

In the case $l=0$ we use similar arguments, where the distance $\|x-y\|_{\mathbb{R}^N}$ can be estiamted by diam$|\Omega_1|$ since Ω_1 is bounded. □

Remark 3.6 *In case $l=1$ it follows that the sharper inequality holds true: There exists a constant $C\in\mathbb{R}$ such that for all $w\in W^{s,p_1}(\Omega_1,\mathbb{R}^{M_1}),\ 0<s<1,\ 1<p_1<\infty$ and $D\subseteq\Omega_1$*

$$
\|w\|_{L^{p_1}(D,\mathbb{R}^{M_1})} \le C\left(\|w\|_{L^{p_1}(\Omega_1\setminus D,\mathbb{R}^{M_1})} + \int\limits_{\mathcal{S}} \frac{\|w(x)-w(y)\|^{p_1}_{\mathbb{R}^{M_1}}}{\|x-y\|^{N+p_1 s}_{\mathbb{R}^N}}\rho^l(x-y)\,\mathrm{d}(x,y)\right). \tag{3.8}
$$

Example 3.7 *As in Example 3.4 we consider that $W(\Omega_1,K_1)=W^{s,p_1}(\Omega_1,K_1)$ with $p_1>1,\ 0<s<1$ and fix $k=N$.*

Assume that F *is the inpainting operator, i.e.*

$$
\mathrm{F}(w) = \chi_{\Omega_1\setminus D}(w),
$$

where $D\subseteq\Omega_1,\ w\in W^{s,p_1}(\Omega_1,K_1)$. Since the dimension of the data w and the image data $\mathrm{F}(w)$ have the same dimension at every point $x\in\Omega_1$, we write $M:=M_1=M_2$.
Then the functional

$$
\mathcal{F}[\mathrm{d}_2,\mathrm{d}_1] = [\![\mathrm{F}(w),v]\!]^{p_2}_{[\mathrm{d}_2]} + \alpha\mathcal{R}_{[\mathrm{d}_1]}(w) : W^{s,p_1}(\Omega_1,K_1)\to[0,\infty]
$$

is coercive for $p_2\ge p_1$:
The proof is done using the same arguments as in the proof of Example 3.4, where we additionally split ***Case 1*** *into the two sub-cases*

Case 1.1 $\|w_n\|_{L^{p_1}(D,\mathbb{R}^M)} \to +\infty$

Case 1.2 $\|w_n\|_{L^{p_1}(\Omega_1 \setminus D,\mathbb{R}^M)} \to +\infty$

*The fact that $p_2 \geq p_1$ and that Ω_1 is bounded ensures that $L^{p_2}(\Omega_1 \setminus D, \mathbb{R}^M) \subseteq L^{p_1}(\Omega_1, \mathbb{R}^M)$ which is needed in **Case 1.2** and **Case 2.2**. The sharper inequality (3.8) is needed in **Case 2.2**.*

4. Stability and Convergence

In this section we will first show a stability and afterwards a convergence result. We use the notation introduced in Section 2. In particular $W(\Omega_1, K_1)$ is as defined in Equation 2.6. We also stress that we use notationally simplified versions $\mathcal{F}$ and $\mathcal{F}^v$ of $\mathcal{F}[\mathrm{d}_2, \mathrm{d}_1]$ and $\mathcal{R}$ of $\mathcal{R}_{[\mathrm{d}_1]}$ whenever possible. See Equation 2.7, Equation 2.8 and Equation 2.9.

Lemma 4.1 *Let Assumption 3.1 be satisfied. Then*

$$\mathcal{F}^{v_\star}[\mathrm{d}_2, \mathrm{d}_1](w) \leq 2^{p_2-1}\mathcal{F}^{v_\diamond}[\mathrm{d}_2, \mathrm{d}_1](w) + 2^{p_2-1}[\![v_\diamond, v_\star]\!]_{[\mathrm{d}_2]}^{p_2}$$

for every $w \in W(\Omega_1, K_1)$ and $v_\star, v_\diamond \in L^{p_2}(\Omega_2, K_2)$.

Proof: Using the fact that for $p \geq 1$ we have that $|a+b|^p \leq 2^{p-1}(|a|^p + |b|^p),\ a, b \in \mathbb{R}$ and that $[\![\cdot,\cdot]\!]_{[\mathrm{d}_2]}$ fulfills the triangle inequality we obtain

$$\begin{aligned}\mathcal{F}^{v_\star}[\mathrm{d}_2, \mathrm{d}_1](w) &= [\![\mathrm{F}(w), v_\star]\!]_{[\mathrm{d}_2]}^{p_2} + \alpha\mathcal{R}_{[\mathrm{d}_1]}(w)\\ &\leq 2^{p_2-1}\big([\![\mathrm{F}(w), v_\diamond]\!]_{[\mathrm{d}_2]}^{p_2} + [\![v_\diamond, v_\star]\!]_{[\mathrm{d}_2]}^{p_2}\big) + \alpha\mathcal{R}_{[\mathrm{d}_1]}(w)\\ &\leq 2^{p_2-1}\big(\mathcal{F}^{v_\diamond}[\mathrm{d}_2, \mathrm{d}_1](w) + [\![v_\diamond, v_\star]\!]_{[\mathrm{d}_2]}^{p_2}\big).\end{aligned}$$ □

Theorem 4.2 *Let $v^\delta \in L^{p_2}(\Omega_2, K_2)$. Let $(v_n)_{n\in\mathbb{N}}$ be a sequence in $L^{p_2}(\Omega_2, K_2)$ such that $[\![v_n, v^\delta]\!] \to 0$. Then every sequence $(w_n)_{n\in\mathbb{N}}$ with*

$$w_n \in \arg\min\{\mathcal{F}^{v_n}[\mathrm{d}_2, \mathrm{d}_1](w) \ :\ w \in W(\Omega_1, K_1)\}$$

has a converging subsequence w.r.t. the topology of $W(\Omega_1, K_1)$. The limit $\tilde{w}$ of any such converging subsequence $(w_{n_k})_{k\in\mathbb{N}}$ is a minimizer of $\mathcal{F}^{v^\delta}[\mathrm{d}_2, \mathrm{d}_1]$. Moreover, $(\mathcal{R}(w_{n_k}))_{k\in\mathbb{N}}$ converges to $\mathcal{R}(\tilde{w})$.

The subsequent proof of Theorem 4.2 is similar to the proof of [53, Theorem 3.23].

Proof: For the ease of notation we simply write $\mathcal{F}^{v^\delta}$ instead of $\mathcal{F}^{v^\delta}[\mathrm{d}_2, \mathrm{d}_1]$ and $[\![v, \tilde{v}]\!] = [\![v, \tilde{v}]\!]_{[\mathrm{d}_2]}$

By assumption the sequence $([\![v_n, v^\delta]\!])_{n\in\mathbb{N}}$ converges to 0 and thus is bounded, i.e., there exists $B \in (0, +\infty)$ such that

$$[\![v_n, v^\delta]\!] \leq B \text{ for all } n \in \mathbb{N}. \tag{4.1}$$

Because $w_n \in \arg\min\{\mathcal{F}^{v_n}(w) \ :\ w \in W(\Omega_1, K_1)\}$ it follows that

$$\mathcal{F}^{v_n}(w_n) \leq \mathcal{F}^{v_n}(w) \text{ for all } w \in W(\Omega_1, K_1). \tag{4.2}$$

Now, take an arbitrary feasible $\overline{w} \in W(\Omega_1, K_1)$ and set $c := 2^{p_2-1}$. Applying Lemma 4.1, Equation 4.2 and Equation 4.1 implies that for all $n \in \mathbb{N}$

$$\begin{aligned}\mathcal{F}^{v^\delta}(w_n) &\leq c\mathcal{F}^{v_n}(w_n) + c[\![v_n, v^\delta]\!]^{p_2}\\ &\leq c\mathcal{F}^{v_n}(\overline{w}) + cB^{p_2}\\ &\leq c\big[c\mathcal{F}^{v^\delta}(\overline{w}) + c[\![v^\delta, v_n]\!]^{p_2}\big] + cB^{p_2}\\ &\leq c^2\mathcal{F}^{v^\delta}(\overline{w}) + (c^2 + c)B^{p_2} =: m.\end{aligned}$$

Hence, from Assumption 3.1, Equation 3.1 it follows that the sequence $(w_n)_{n\in\mathbb{N}}$ contains a converging subsequence.

Let now $(w_{n_k})_{k\in\mathbb{N}}$ be an arbitrary subsequence of $(w_n)_{n\in\mathbb{N}}$ which converges in $W(\Omega_1, K_1)$ to some $\tilde{w} \in W(\Omega_1, \mathbb{R}^{M_1})$. Then, from Lemma 2.8 and the continuity properties of F it follows that $\tilde{w} \in W(\Omega_1, K_1)$ and $(\mathrm{F}(w_{n_k}), v_{n_k}) \to (\mathrm{F}(\tilde{w}), v^\delta)$ in $L^{p_2}(\Omega_2, K_2)\times L^{p_2}(\Omega_2, K_2)$. Moreover, using Lemma 3.2 and Equation 4.2 it follows that for every $w \in W(\Omega_1, K_1)$ the following estimate holds true

$$\begin{aligned}
\mathcal{F}^{v^\delta}(\tilde{w}) &= [\![\mathrm{F}(\tilde{w}), v^\delta]\!]^{p_2} + \alpha\mathcal{R}(\tilde{w}) \leq [\![\mathrm{F}(\tilde{w}), v^\delta]\!]^{p_2} + \alpha \liminf_{k\to+\infty} \mathcal{R}(w_{n_k})\\
&\leq \liminf_{k\to+\infty} [\![\mathrm{F}(w_{n_k}), v_{n_k}]\!]^{p_2} + \alpha \liminf_{k\to+\infty} \mathcal{R}(w_{n_k}) = \limsup_{k\to+\infty} \mathcal{F}^{v_{n_k}}(w_{n_k}) \leq \limsup_{k\to+\infty} \mathcal{F}^{v_{n_k}}(w)\\
&= \limsup_{k\to+\infty} \left([\![\mathrm{F}(w), v_{n_k}]\!]^{p_2} + \alpha\mathcal{R}(w)\right) = [\![\mathrm{F}(w), v^\delta]\!]^{p_2} + \alpha\mathcal{R}(w)\\
&= \mathcal{F}^{v^\delta}(w).
\end{aligned}$$

This shows that $\tilde{w}$ is a minimizer of $\mathcal{F}^{v^\delta}$. Choosing $w = \tilde{w}$ in the previous estimate we obtain the equality

$$[\![\mathrm{F}(\tilde{w}), v^\delta]\!]^{p_2} + \alpha\mathcal{R}(\tilde{w}) = [\![\mathrm{F}(\tilde{w}), v^\delta]\!]^{p_2} + \alpha \liminf_{k\to+\infty} \mathcal{R}(w_{n_k}) = [\![\mathrm{F}(\tilde{w}), v^\delta]\!]^{p_2} + \alpha \limsup_{k\to+\infty} \mathcal{R}(w_{n_k}),$$

so that

$$\mathcal{R}(\tilde{w}) = \lim_{k\to+\infty} \mathcal{R}(w_{n_k}). \qquad \square$$

Before proving the next theorem we need the following definition, cf. [53].

Definition 4.3 An element $w^* \in W(\Omega_1, K_1)$ such that

$$\begin{aligned}
&\mathrm{F}(w^*) = v^0\\
&\mathcal{R}(w^*) = \min\{\mathcal{R}(w) \;:\; w \in W(\Omega_1, K_1),\ \mathrm{F}(w) = v^0\}.
\end{aligned} \tag{4.3}$$

is fulfilled is called an *$\mathcal{R}$-minimizing solution*.

The following theorem and its proof are inspired by [53, Theorem 3.26].

Theorem 4.4 *Let there exist an $\mathcal{R}$-minimizing solution $w^\dagger \in W(\Omega_1, K_1)$ and let $\alpha : (0,\infty) \to (0,\infty)$ be a function satisfying*

$$\alpha(\delta) \to 0 \text{ and } \frac{\delta^{p_2}}{\alpha(\delta)} \to 0 \text{ for } \delta \to 0. \tag{4.4}$$

Let $(\delta_n)_{n\in\mathbb{N}}$ be a sequence of positive real numbers converging to 0. Moreover, let $(v_n)_{n\in\mathbb{N}}$ be a sequence in $L^{p_2}(\Omega_2, K_2)$ with $[\![v^0, v_n]\!]_{[\mathrm{d}_2]} \leq \delta_n$ and set $\alpha_n := \alpha(\delta_n)$.

Then every sequence $(w_n)_{n\in\mathbb{N}}$ of minimizers

$$w_n \in \arg\min\{\mathcal{F}^{v_n}_{\alpha_n}[\mathrm{d}_2, \mathrm{d}_1](w) \;:\; w \in W(\Omega_1, K_1)\}$$

has a converging subsequence $w_{n_k} \overset{W}{\to} \tilde{w}$ as $k \to \infty$, where $\tilde{w}$ is an $\mathcal{R}$-minimizing solution. In addition, $\mathcal{R}(w_{n_k}) \to \mathcal{R}(\tilde{w})$.

Moreover, if $w^\dagger$ is unique it follows that $w_n \overset{W}{\to} w^\dagger$ and $\mathcal{R}(w_n) \to \mathcal{R}(w^\dagger)$.

Proof: We write shortly $[\![\cdot,\cdot]\!]$ for $[\![\cdot,\cdot]\!]_{[\mathrm{d}_2]}$. Taking into account that $w_n \in \operatorname{argmin}\{\mathcal{F}^{v_n}_{\alpha_n}[\mathrm{d}_2, \mathrm{d}_1](w) \;:\; w \in W(\Omega_1, K_1)\}$ it follows that

$$[\![\mathrm{F}(w_n), v_n]\!]^{p_2} \leq \mathcal{F}^{v_n}_{\alpha_n}(w_n) \leq \mathcal{F}^{v_n}_{\alpha_n}(w^\dagger) = [\![v^0, v_n]\!]^{p_2} + \alpha_n\mathcal{R}(w^\dagger) \leq \delta_n^{p_2} + \alpha_n\mathcal{R}(w^\dagger) \to 0,$$

yielding $[\![\mathrm{F}(w_n), v_n]\!] \to 0$ as $n \to +\infty$. From the triangle inequality it follows that $[\![\mathrm{F}(w_n), v^0]\!] \leq [\![\mathrm{F}(w_n), v_n]\!] + [\![v_n, v^0]\!] \to 0$ as $n \to +\infty$, so that, by taking into account that $\mathrm{d}_{\mathbb{R}^{M_2}}\big|_{K_2\times K_2} \leq \mathrm{d}_2$, it follows that

$$\mathrm{F}(w_n) \to v^0 \text{ in } L^{p_2}(\Omega_2, \mathbb{R}^{M_2}). \tag{4.5}$$

Since

$$\mathcal{R}(w_n) \leq \frac{1}{\alpha_n}\mathcal{F}^{v_n}_{\alpha_n}(w_n) \leq \frac{1}{\alpha_n}\mathcal{F}^{v_n}_{\alpha_n}(w^\dagger) = \frac{1}{\alpha_n}\left([\![v^0, v_n]\!]^{p_2} + \alpha_n\mathcal{R}(w^\dagger)\right) \leq \frac{\delta_n^{p_2}}{\alpha_n} + \mathcal{R}(w^\dagger),$$

we also get

$$\limsup_{n\to+\infty} \mathcal{R}(w_n) \le \mathcal{R}(w^\dagger). \tag{4.6}$$

Set $\alpha_{\max} := \max\{\alpha_n : n \in \mathbb{N}\}$. Since

$$\limsup_{n\to+\infty} \mathcal{F}_{\alpha_n}^{v^0}(w_n) \le \limsup_{n\to+\infty} \left([\![\mathrm{F}(w_n), v^0]\!]^{p_2} + \alpha_n \mathcal{R}(w_n) \right) \le \alpha_{\max} \mathcal{R}(w^\dagger)$$

the sequence $\mathcal{F}_{\alpha}^{v^0}(w_n)$ is bounded. From Assumption 3.1 Equation 3.1 it follows that there exists a converging subsequence $(w_{n_k})_{k\in\mathbb{N}}$ of $(w_n)_{n\in\mathbb{N}}$. The limit of $(w_{n_k})_{k\in\mathbb{N}}$ is denoted by $\tilde{w}$. Then, from Lemma 2.8 it follows that $\tilde{w} \in W(\Omega_1, K_1)$. Since the operator F is sequentially continuous it follows that $\mathrm{F}(w_{n_k}) \to \mathrm{F}(\tilde{w})$ in $L^{p_2}(\Omega_2, K_2)$. This shows that actually $\mathrm{F}(\tilde{w}) = v^0$ since Equation 4.5 is valid. Then, from Lemma 3.2 it follows that the functional $\mathcal{R} : W(\Omega_1, K_1) \to [0, +\infty]$ is sequentially lower semi-continuous, so that $\mathcal{R}(\tilde{w}) \le \liminf_{k\to+\infty} \mathcal{R}(w_{n_k})$. Combining this with Equation 4.6 we also obtain

$$\mathcal{R}(\tilde{w}) \le \liminf_{k\to+\infty} \mathcal{R}(w_{n_k}) \le \limsup_{k\to+\infty} \mathcal{R}(w_{n_k}) \le \mathcal{R}(w^\dagger) \le \mathcal{R}(\tilde{w}),$$

using the definition of $w^\dagger$. This, together with the fact that $\mathrm{F}(\tilde{w}) = v^0$ we see that $\tilde{w}$ is an $\mathcal{R}$-minimizing solution and that $\lim_{k\to+\infty} \mathcal{R}(w_{n_k}) = \mathcal{R}(\tilde{w})$.

Now assume that the solution fulfilling Equation 4.3 is unique; we call it $w^\dagger$. In order to prove that $w_n \overset{W}{\to} w^\dagger$ it is sufficient to show that any subsequence has a further subsequence converging to $w^\dagger$, cf. [53, Lemma 8.2]. Hence, denote by $(w_{n_k})_{k\in\mathbb{N}}$ an arbitrary subsequence of (w_n), the sequence of minimizers. Like before we can show that $\mathcal{F}_{\alpha}^{v^0}(w_{n_k})$ is bounded and we can extract a converging subsequence $(w_{n_{k_l}})_{l\in\mathbb{N}}$. The limit of this subsequence is $w^\dagger$ since it is the unique solution fulfilling Equation 4.3, showing that $w_n \overset{W}{\to} w^\dagger$. Moreover, $w^\dagger \in W(\Omega_1, K_1)$. Following the arguments above we obtain as well $\lim_{n\to+\infty} \mathcal{R}(w_n) = \mathcal{R}(w^\dagger)$.□

Remark 4.5 *Theorem 4.2 guarantees that the minimizers of $\mathcal{F}_{\alpha}^{v_n}[\mathrm{d}_2, \mathrm{d}_1]$ depend continuously on v^δ while Theorem 4.4 ensures that they converge to a solution of $\mathrm{F}(w) = v^0$, v^0 the exact data, while α tends to zero.*

5. Discussion of the Results and Conjectures

In this section we summarize some open problems related double integral representations of functions with values on manifolds.

5.1. Relation to single integral representations. In the following we show for one particular case of functions that have values in a manifold, the double integral formulation $\mathcal{R}_{[\mathrm{d}_1]}$, defined in Equation 2.9, approximates a single energy integral. The basic ingredient for this derivation is the exponential map related to the metric d_1 on the manifold. In the following we investigate manifold-valued functions $w \in W^{1,2}(\Omega, \mathcal{M})$, where we consider $\mathcal{M} \subseteq \mathbb{R}^{M\times 1}$ to be a connected, complete Riemannian manifold. In this case some of the regularization functionals $\mathcal{R}_{[\mathrm{d}_1]}$, defined in Equation 2.9, can be considered as approximations of *single* integrals. In particular we aim to generalize Equation 1.3 in the case $p = 2$.

We have that

$$\nabla w = \begin{bmatrix} \frac{\partial w_1}{\partial x_1} & \cdots & \frac{\partial w_1}{\partial x_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial w_M}{\partial x_1} & \cdots & \frac{\partial w_M}{\partial x_N} \end{bmatrix} \in \mathbb{R}^{M\times N}.$$

In the following we will write $\mathcal{R}_{[\mathrm{d}_1],\varepsilon}$ instead of $\mathcal{R}_{\mathrm{d}_1}$ to stress the dependence on ε in contrast to above. Moreover, let $\hat{\rho} : \mathbb{R}_+ \to \mathbb{R}_+$ be in $C_c^\infty(\mathbb{R}_+)$ and satisfy

$$\left|\mathbb{S}^{N-1}\right| \int_0^\infty \hat{t}^{N-1} \hat{\rho}\left(\hat{t}\right) d\hat{t} = 1\ .$$

Then for every $\varepsilon > 0$

$$x \in \mathbb{R}^n \to \rho_\varepsilon(x) := \frac{1}{\varepsilon^N} \hat{\rho}\left(\frac{\|x\|_{\mathbb{R}^N}}{\varepsilon} \right)$$

is a mollifier, cf. Example 2.2.
$\mathcal{R}_{[\mathrm{d}_1],\varepsilon}$ (with $p=2$) then reads as follows:

$$\lim_{\varepsilon\to 0+} \mathcal{R}_{[\mathrm{d}_1],\varepsilon}(w) := \lim_{\varepsilon\to 0+} \frac{1}{2} \int\limits_{\Omega\times\Omega} \frac{d_1^2(w(x),w(y))}{\|x-y\|_{\mathbb{R}^N}^2} \rho_\varepsilon(x-y)\,\mathrm{d}(x,y)\,, \tag{5.1}$$

where we added the factor $\frac{1}{2}$ due to reasons of calculation. Integration by parts with spherical coordinates $y = x - t\theta \in \mathbb{R}^{N\times 1}$ with $\theta \in \mathbb{S}^{N-1} \subseteq \mathbb{R}^{N\times 1}$, $t \geq 0$ gives

$$\lim_{\varepsilon\to 0+} \mathcal{R}_{[\mathrm{d}_1],\varepsilon}(w) = \lim_{\varepsilon\to 0+} \frac{1}{\varepsilon^N} \int\limits_{\Omega} \int\limits_{\mathbb{S}^{N-1}} \int\limits_0^\infty \frac{1}{2} d_1^2(w(x),w(x-t\theta)) t^{N-3} \hat{\rho}\left(\frac{t}{\varepsilon}\right) \mathrm{d}t\mathrm{d}\theta\,\mathrm{d}x\,. \tag{5.2}$$

Now, using that for $m_1 \in \mathcal{M}$ fixed and $m_2 \in \mathcal{M}$ such that m_1 and m_2 are joined by a unique minimizing geodesic (see for instance [29] where the concept of exponential mappings is explained)

$$\frac{1}{2}\partial_2 d_1^2(m_1,m_2) = -(\exp_{m_2})^{-1}(m_1) \in \mathbb{R}^{M\times 1}, \tag{5.3}$$

where ∂_2 denotes the derivative of d_1^2 with respect to the second component. By application of the chain rule we get

$$\frac{1}{2}\nabla_y d_1^2(w(x),w(y)) = -(\nabla w(y))^T (\exp_{w(y)})^{-1}(w(x)) \in \mathbb{R}^{N\times M} \times \mathbb{R}^{M\times 1} = \mathbb{R}^{N\times 1}\,,$$

where $w(x)$ and $w(y)$ are joined by a unique minimizing geodesic. This assumption is justifiable due to the fact that we consider the case $\varepsilon \to 0+$. Let $\cdot$ denote the scalar multiplication of two vectors in $\mathbb{R}^{N\times 1}$, then the last inequality shows that

$$\begin{aligned}\frac{1}{2}d_1^2(w(x),w(x-t\theta)) &= \frac{1}{2}d_1^2(w(x),w(x-t\theta)) - \frac{1}{2}d_1^2(w(x),w(x)) \\ &\approx -t\left((\nabla w(x-t\theta))^T (\exp_{w(x-t\theta)})^{-1}(w(x))\right)\cdot\theta\,.\end{aligned}$$

Thus from Equation 5.2 it follows that

$$\begin{aligned}&\lim_{\varepsilon\to 0+} \mathcal{R}_{[\mathrm{d}_1],\varepsilon}(w) \\ &\approx \lim_{\varepsilon\to 0+} -\frac{1}{\varepsilon^N} \int\limits_{\Omega} \int\limits_{\mathbb{S}^{N-1}} \int\limits_0^\infty \left((\nabla w(x-t\theta))^T (\exp_{w(x-t\theta)})^{-1}(w(x))\right)\cdot\theta \left(t^{N-2}\hat{\rho}\left(\frac{t}{\varepsilon}\right)\right) \mathrm{d}t\mathrm{d}\theta\,\mathrm{d}x\,.\end{aligned} \tag{5.4}$$

Now we will use a Taylor series of power 0 for $\nabla w(x-t\theta)$ and of power 1 for $(\exp_{w(x-t\theta)})^{-1}(w(x))$ to rewrite Equation 5.4. We write

$$F(w(x);t,\theta) := -(\exp_{w(x-t\theta)})^{-1}(w(x)) \in \mathbb{R}^{M\times 1} \tag{5.5}$$

and define

$$\dot{F}(w(x);\theta) := -\lim_{t\to 0+} \frac{1}{t\theta} \left((\exp_{w(x-t\theta)})^{-1}(w(x)) - \underbrace{(\exp_{w(x)})^{-1}(w(x))}_{=0} \right) \in \mathbb{R}^{M\times 1}. \tag{5.6}$$

Note that in the case that $\nabla w(x) \neq 0$ this is the leading order approximation of $\nabla w(x-t\theta)$. On the other hand because $(\exp_{w(x)})^{-1}(w(x))$ vanishes, $\dot{F}(w(x);\theta)$ is the leading order term of the expansion of $(\exp_{w(x-t\theta)})^{-1}(w(x))$ with respect to t. In summary we are calculating the leading order term of the expansion with respect to t.

Then from Equation 5.4 it follows that

$$\lim_{\varepsilon\to 0+} \mathcal{R}_{[\mathrm{d}_1],\varepsilon}(w) \approx \underbrace{\lim_{\varepsilon\to 0+} \frac{1}{\varepsilon^N} \int\limits_0^\infty t^{N-1}\hat{\rho}\left(\frac{t}{\varepsilon}\right)\mathrm{d}t}_{=|\mathbb{S}^{N-1}|^{-1}} \int\limits_{\Omega} \int\limits_{\mathbb{S}^{N-1}} \left((\nabla w(x))^T \dot{F}(w(x);\theta)\right) \mathrm{d}\theta\,\mathrm{d}x\,. \tag{5.7}$$

The previous calculations show that the double integral simplifies to a double integral where the inner integration domain has one dimension less than the original integral. Under certain assumption the integration domain can be further simplified:

Example 5.1 *If* $d_1(x,y) = \|x-y\|_{\mathbb{R}^M}$, $p = 2$, *then*

$$\dot{F}(w(x),\theta) = -\lim_{t\to 0+} \frac{1}{t\theta}\left(w(x) - w(x-t\theta)\right) = -\nabla w \in \mathbb{R}^{M\times 1}.$$

Thus from Equation 5.7 it follows that

$$\lim_{\varepsilon\to 0+} \mathcal{R}_{[\mathrm{d}_1],\varepsilon}(w) \approx \int_\Omega \underbrace{(\nabla w(x))^T \nabla w(x)}_{\|\nabla w(x)\|^2_{\mathbb{R}^M}} \,\mathrm{d}x\,. \tag{5.8}$$

This is exactly the identity derived in Bourgain, Brézis, and Mironescu [13].

From these considerations we can view $\lim_{\varepsilon\to 0+} \mathcal{R}_{[\mathrm{d}_1],\varepsilon}$ as functionals, which generalize Sobolev and BV semi-norms to functions with values on manifolds.

5.2. A conjecture on Sobolev semi-norms. Starting point for this conjecture is Equation 2.9. We will write Ω, M and p instead of Ω_1, M_1 and p_1.

- In the case $l = 0$, $k = N$, $0 < s < 1$ and $\mathrm{d}_1(w(x), w(y)) = \|w(x) - w(y)\|_{\mathbb{R}^M}$, $\mathcal{R}_{[\mathrm{d}_1]}$ from Equation 2.9 simplifies to the p-th power of the Sobolev semi-norm

$$\int_{\Omega\times\Omega} \frac{\|w(x)-w(y)\|^p_{\mathbb{R}^M}}{\|x-y\|^{N+ps}_{\mathbb{R}^N}} \,\mathrm{d}(x,y). \tag{5.9}$$

  For a recent survey on fractional Sobolev Spaces see [24].

- On the other hand, when we choose $k = 0$, $l = 1$ and $\mathrm{d}_1(w(x), w(y)) = \|w(x) - w(y)\|_{\mathbb{R}^M}$, then $\mathcal{R}_{[\mathrm{d}_1]}$ from Equation 2.9 (note $\rho = \rho_\varepsilon$ by simplification of notation):

$$\int_{\Omega\times\Omega} \frac{\|w(x)-w(y)\|^p_{\mathbb{R}^M}}{\|x-y\|^{ps}_{\mathbb{R}^N}} \rho_\varepsilon(x-y) \,\mathrm{d}(x,y). \tag{5.10}$$

- Therefore, in analogy what we know for $s = 1$ from [13], we conjecture that

$$\lim_{\varepsilon\to 0} \int_{\Omega\times\Omega} \frac{\|w(x)-w(y)\|^p_{\mathbb{R}^M}}{\|x-y\|^{ps}_{\mathbb{R}^N}} \rho_\varepsilon(x-y) \,\mathrm{d}(x,y) = C \int_{\Omega\times\Omega} \frac{\|w(x)-w(y)\|^p_{\mathbb{R}^M}}{\|x-y\|^{N+ps}_{\mathbb{R}^N}} \,\mathrm{d}(x,y). \tag{5.11}$$

  The form Equation 5.11 is numerically preferable to the standard Sobolev semi-norm Equation 5.9, because $\rho = \rho_\varepsilon$ and thus the integral kernel has compact support.

6. Numerical Examples

In this section we present some numerical examples for denoising and inpainting of functions with values on the circle $\mathbb{S}^1$. Functions with values on a sphere have already been investigated very diligently (see for instance [14] out of series of publications of these authors). Therefore we review some of their results first.

6.1. $\mathbb{S}^1$-Valued Data. Let $\emptyset \neq \Omega \subset \mathbb{R}$ or $\mathbb{R}^2$ be a bounded and simply connected open set. In [14] the question was considered when $w \in W^{s,p}(\Omega, \mathbb{S}^1)$ can be represented by some function $u \in W^{s,p}(\Omega, \mathbb{R})$ satisfying

$$\Phi(u) := \mathrm{e}^{\mathrm{i}u} = w. \tag{6.1}$$

That is, the function u can be *lifted* to w.

Lemma 6.1 ([14]) • *Let* $\Omega \subset \mathbb{R}$, $0 < s < \infty$, $1 < p < \infty$. *Then for all* $w \in W^{s,p}(\Omega, \mathbb{S}^1)$ *there exists* $u \in W^{s,p}(\Omega, \mathbb{R})$ *satisfying Equation 6.1.*

- *Let $\Omega \subset \mathbb{R}^N$, $N \geq 2$, $0 < s < 1$, $1 < p < \infty$. Moreover, let $sp < 1$ or $sp \geq N$, then for all $w \in W^{s,p}(\Omega, \mathbb{S}^1)$ there exists $u \in W^{s,p}(\Omega, \mathbb{R})$ satisfying Equation 6.1.*

  If $sp \in [1, N)$, then there exist functions $w \in W^{s,p}(\Omega, \mathbb{S}^1)$ such that Equation 6.1 does not hold with some function $u \in W^{s,p}(\Omega, \mathbb{R})$.

We consider the functional (note that by simplification of notation below $\rho = \rho_\varepsilon$ denotes a mollifier)

$$\mathcal{R}_{[\mathrm{d}_{\mathbb{S}^1}]}(w) = \int\limits_{\Omega \times \Omega} \frac{\mathrm{d}^p_{\mathbb{S}^1}(w(x), w(y))}{\|x - y\|^{k+ps}_{\mathbb{R}^N}} \rho^l(x - y) \,\mathrm{d}(x, y), \tag{6.2}$$

on $w \in W^{s,p}(\Omega, \mathbb{S}^1)$, which we defined in Equation 2.9.

Writing $w = \Phi(u)$ as in Equation 6.1 we get the lifted functional

$$\mathcal{R}^{\Phi}_{[\mathrm{d}_{\mathbb{S}^1}]}(u) := \int\limits_{\Omega \times \Omega} \frac{\mathrm{d}^p_{\mathbb{S}^1}(\Phi(u)(x), \Phi(u)(y))}{\|x - y\|^{k+ps}_{\mathbb{R}^N}} \rho^l(x - y) \,\mathrm{d}(x, y) \tag{6.3}$$

over the space $W^{s,p}(\Omega, \mathbb{R})$, where

$$\mathrm{d}_{\mathbb{S}^1}(w(x), w(y)) := \arccos(w(x)^T w(y)). \tag{6.4}$$

Remark 6.2
- *We note that in the case $k = 0$, $s = 1$ and $l = 1$ these integrals correspond with the ones considered in Bourgain, Brézis, and Mironescu [13] for functions with values on $\mathbb{S}^1$.*
- *If we choose $k = N$, $s = 1$ and $l = 0$, then this corresponds with Sobolev semi-norms on manifolds.*
- *Let $\varepsilon > 0$ fixed (that is, we consider neither a standard Sobolev regularization nor the limiting case $\varepsilon \to 0$ as in [13]). In this case we have proven coercivity of the functional $\mathcal{F} : W^{s,p}(\Omega, \mathbb{S}^1) \to [0, \infty)$, $0 < s < 1$, only with the following regularization functional, cf. Example 3.4 and Example 3.7:*

$$\int\limits_{\Omega \times \Omega} \frac{\mathrm{d}^p_{\mathbb{S}^1}(w(x), w(y))}{\|x - y\|^{N+ps}_{\mathbb{R}^N}} \rho_\varepsilon(x - y) \,\mathrm{d}(x, y).$$

We summarize a few results: The first lemma follows from elementary calculations:

Lemma 6.3 *$\mathrm{d}_{\mathbb{S}^1}$ and $\mathrm{d}_{\mathbb{R}^2}\big|_{\mathbb{S}^1 \times \mathbb{S}^1}$ are equivalent.*

Below we show that $\mathcal{R}^{\Phi}_{[\mathrm{d}_{\mathbb{S}^1}]}$ is finite on $W^{s,p}(\Omega, \mathbb{S}^1)$.

Lemma 6.4 *Let $\emptyset \neq \Omega \subset \mathbb{R}$ or $\mathbb{R}^2$ be a bounded and simply connected open set. Let $1 < p < \infty$ and $s \in (0, 1)$. If $N = 2$ assume that $sp < 1$ or $sp \geq 2$. Moreover, let Assumption 3.1 be satisfied. Then $\mathcal{R}^{\Phi}_{[\mathrm{d}_{\mathbb{S}^1}]}$ maps $W^{s,p}(\Omega, \mathbb{R})$ into $[0, \infty)$ (i.e. does not attain the value $+\infty$).*

Proof: From Lemma 6.3 and the assumptions of this lemma it follows that Assumption 2.1 holds. Therefore, it follows from Proposition 2.12 item *(ii)* that $\mathcal{R}_{[\mathrm{d}_{\mathbb{S}^1}]}(w) < \infty$ for all $w \in W(\Omega, \mathbb{S}^1)$. Then, from Lemma 6.1 and Lemma 6.3 it follows that $\mathcal{R}^{\Phi}_{[\mathrm{d}_{\mathbb{S}^1}]}(u) < \infty$ for all $u \in W(\Omega, \mathbb{R})$. which gives the assertion. □

6.2. Setting of numerical examples. In all numerical examples presented we use a simplified setting with

$$M_1 = M_2 =: M,\ \Omega_1 = \Omega_2 =: \Omega,\ K_1 = K_2 =: \mathbb{S}^1,\ p_1 = p_2 =: p,\ k = N,\ l = 1$$

and

$$W(\Omega, \mathbb{S}^1) = W^{s,p}(\Omega, \mathbb{S}^1).$$

As particular mollifier we use ρ_ε (see Example 2.2), which is defined via the one-dimensional normal-distribution $\hat{\rho}(x) = \frac{1}{\sqrt{\pi}} \mathrm{e}^{-x^2}$.

Regularization functionals. Let $\mathcal{R}_{[\mathrm{d}_{\mathbb{S}^1}]}$ and $\mathcal{R}^{\Phi}_{[\mathrm{d}_{\mathbb{S}^1}]}$ be as defined in Equation 6.2 and Equation 6.3, respectively. In what follows we consider the following regularization functional

$$\mathcal{F}_{\alpha}^{v^{\delta}}[\mathrm{d}_{\mathbb{S}^1}](w) := \int_{\Omega} \mathrm{d}_{\mathbb{S}^1}^{p}(\mathrm{F}(w)(x), v^{\delta}(x))\,\mathrm{d}x + \alpha \mathcal{R}_{[\mathrm{d}_{\mathbb{S}^1}]}(w), \tag{6.5}$$

on $W^{s,p}(\Omega, \mathbb{S}^1)$ and the lifted variant

$$\tilde{\mathcal{F}}_{\alpha}^{u^{\delta}}[\mathrm{d}_{\mathbb{S}^1}](u) := \int_{\Omega} \mathrm{d}_{\mathbb{S}^1}^{p}(\mathrm{F}(\Phi(u))(x), \Phi(u^{\delta}(x)))\,\mathrm{d}x + \alpha \mathcal{R}^{\Phi}_{[\mathrm{d}_{\mathbb{S}^1}]}(u) \tag{6.6}$$

over the space $W^{s,p}(\Omega, \mathbb{R})$ (as in Subsection 6.1), where Φ is defined as in (6.1). Note that $\tilde{\mathcal{F}} = \mathcal{F} \circ \Phi$ and $v^{\delta} = \Phi(u^{\delta})$.

Lemma 6.5 *Let $\emptyset \neq \Omega \subset \mathbb{R}$ or $\mathbb{R}^2$ be a bounded and simply connected open set. Let $1 < p < \infty$ and $s \in (0, 1)$. If $N = 2$ assume that $sp < 1$ or $sp \geq 2$. Moreover, let Assumption 3.1 be satisfied. Then the mapping $\tilde{\mathcal{F}}_{\alpha}^{u^{\delta}}[\mathrm{d}_{\mathbb{S}^1}] : W^{s,p_1}(\Omega_1, \mathbb{R}) \to [0, \infty)$ attains a minimizer.*

Proof: Due to the assumptions $\mathcal{F}_{\alpha}^{v^{\delta}}[\mathrm{d}_{\mathbb{S}^1}]$ attains a minimizer $w^* \in W^{s,p}(\Omega, \mathbb{S}^1)$. It follows from Lemma 6.1 that there exists a function $u^* \in W^{s,p_1}(\Omega_1, \mathbb{R})$ that can be lifted to w^*, i.e. $w^* = \Phi(u^*) := e^{\mathrm{i}u^*}$. Then u^* is a minimizer of (6.6) by definition of $\tilde{\mathcal{F}}$ and Φ. Moreover, $\tilde{\mathcal{F}}(u) < \infty$ by Lemma 6.4 and Lemma 6.3.□

6.3. Numerical minimization. In our concrete examples we will consider two different operators F. For numerical minimization we consider the functional from Equation 6.6 in a discretized setting. For this purpose we approximate the functions $u \in W^{s,p}(\Omega, \mathbb{R})$, $0 < s < 1, 1 < p < \infty$ by quadratic B-Spline functions and optimize with respect to the coefficients. We remark that this approximation, denoted by u_{approx}, is continuous and thus that sharp edges correspond to very steep slopes.
The noisy data u^{δ} is obtained by adding Gaussian white noise with variance σ^2 to the approximation or the discretized approximation of u.

We apply a simple Gradient Descent scheme with fixed step length implemented in MATLAB.

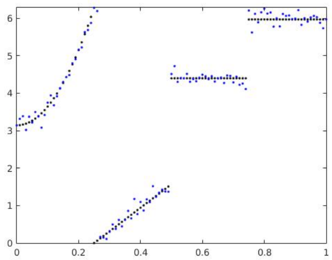
6.4. Denoising of $\mathbb{S}^1$-valued functions - The InSAR problem. In this case the operator $\mathrm{F} : W^{s,p}(\Omega, \mathbb{S}^1) \to L^p(\Omega, \mathbb{S}^1)$ is the inclusion operator. It is norm-coercive in the sense of Equation 3.4 and hence Assumption 3.1 is fulfilled. For $\emptyset \neq \Omega \subset \mathbb{R}$ or $\mathbb{R}^2$ a bounded and simply connected open set, $1 < p < \infty$ and $s \in (0, 1)$ such that additionally $sp < 1$ or $sp \geq 2$ if $N = 2$ we can apply Lemma 6.5 which ensures that the lifted functional $\tilde{\mathcal{F}}_{\alpha}^{u^{\delta}}[\mathrm{d}_{\mathbb{S}^1}] : W^{s,p}(\Omega, \mathbb{R}) \to [0, \infty)$ attains a minimizer $u \in W^{s,p}(\Omega, \mathbb{R})$.

In the examples we will just consider the continuous approximation u_{approx}.

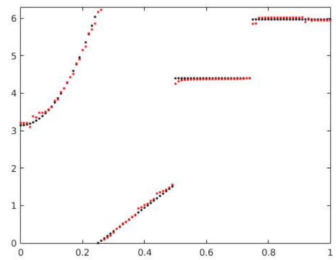
One dimensional test case. Let $\Omega = (0, 1)$ and consider the signal $u_{\text{approx}} : \Omega \to [0, 2\pi)$ representing the angle of a cyclic signal.
For the discrete approximation shown in Figure 2(A) the domain Ω is sampled equally at 100 points. u_{approx} is affected by an additive white Gaussian noise with $\sigma = 0.1$ to obtain the noisy signal which is colored in blue in Figure 2(A).

In this experiment we show the influence of the parameters s and p. In all cases the choice of the regularization parameter α is 0.19 and $\varepsilon = 0.01$.
The red signal in Figure 2(B) is obtained by choosing $s = 0.1$ and $p = 1.1$. We see that the periodicity of the signal is handled correctly and that there is nearly no staircasing. In Figure 2(C) the parameter s is changed from 0.1 to 0.6. The value of the parameter p stays fixed. Increasing of s leads the signal to be more smooth. We can observe an even stronger similar effect when increasing p (here from 1.1 to 2) and letting s fixed, see Figure 2(D). This fits the expectation since s only appears once in the denominator of the regularizer. At a jump increasing of s leads thus to an increasing of the regularization term. The parameter p appears twice in the regularizer. Huge jumps are hence weighted even more.

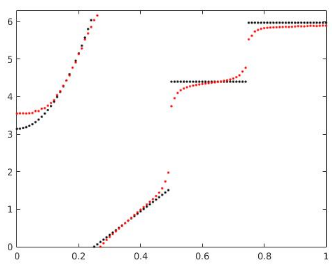
In Figure 3(A) we considered a simple signal with a single huge jump. Again it is described by the angular value. We proceeded as above to obtain the approximated discrete original data (black) and noisy signal with $\sigma = 0.1$ (blue). We chose again $\varepsilon = 0.01$.



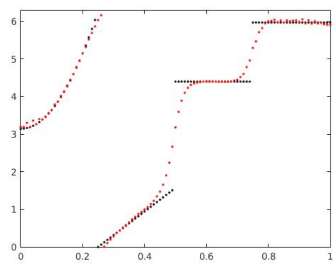
(A) *Original and noisy data*



(B) *Denoised data*



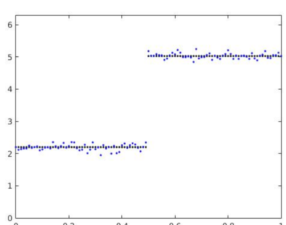
(C) *Increasing of s*



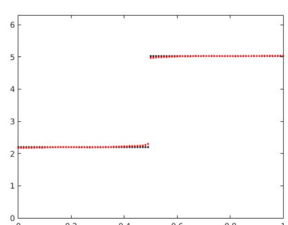
(D) *Increasing of p*

FIGURE 2. *Function on $\mathbb{S}^1$ represented in $[0, 2\pi)$: Left to right, top to bottom: Original data (black) and noisy data (blue) with 100 data points. Denoised data (red) where we chose $s = 0.1, p = 1.1, \alpha = 0.19$. Denoised data with $s = 0.6, p = 1.1, \alpha = 0.19$ resp. $s = 0.1, p = 2, \alpha = 0.19$.*

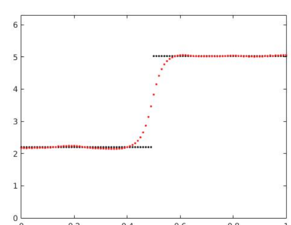
As we have seen above increasing of s leads to a more smooth signal. This effect can be compensated by choosing a rather small value of p, i.e. $p \approx 1$. In Figure 3(B) the value of s is 0.9. We see that it is still possible to reconstruct jumps by choosing e.g. $p = 1.01$.
Moreover, we have seen that increasing of p leads to an even more smooth signal. In Figure 3(C) we choose a quite large value of p, $p = 2$ and a rather small value of s, $s = 0.001$. Even for this very simple signal is was not possible to get sharp edges. This is due to the fact that the parameter p (but not s) additionally weights the height of jumps in the regularizing term.



(A) *Original and noisy data*



(B) $s = 0.9, \ p = 1.01$



(C) $s = 0.001, \ p = 2$

FIGURE 3. *Left to right: Original data (black) and noisy data (blue) sampled at 100 data points. Denoised data (red) where we chose $s = 0.9, p = 1.01, \alpha = 0.03$. Denoised data with $s = 0.001, p = 2, \alpha = 0.9$.*

Denoising of a $\mathbb{S}^1$-Valued Image. Our next example concerned a two-dimensional $\mathbb{S}^1$-valued image represented by the corresponding angular values. We remark that in this case where $N = 2$ the existence of such a representation is always guaranteed in the cases where $sp < 1$ or $sp \geq 2$, see Lemma 6.1.

The domain Ω is sampled into 60×60 data points and can be considered as discrete grid, $\{1, \dots, 60\} \times \{1, \dots, 60\}$. The B-Spline approximation evaluated at that grid is given by

$$u_{\text{approx}}(i, j) := 4\pi \frac{i}{60} \mod 2\pi, \quad i, j \in \{1, \dots, 60\}.$$

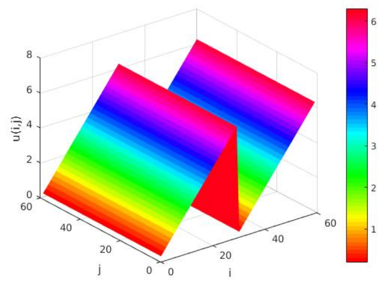
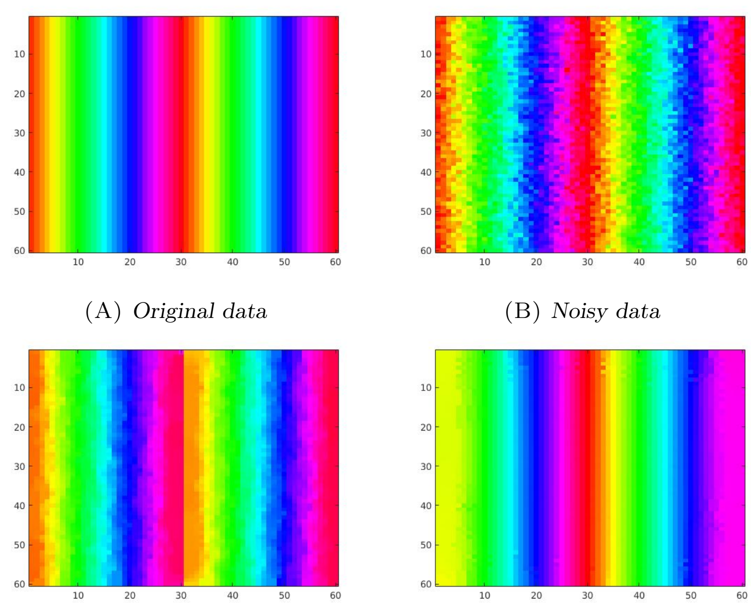


FIGURE 4. *The function u_{approx} evaluated on the discrete grid.*

The function u_{approx} is shown in Figure 4. We used the hsv colormap provided in MATLAB transferred to the interval $[0, 2\pi]$.

This experiment shows the difference of our regularizer respecting the periodicity of the data in contrast to the classical Total Variation regularizer. The classical TV-minimization is solved using a fixed point iteration ([43]); for the method see also [58].

In Figure 5(A) the function u_{approx} can be seen from the top, i.e. the axes correspond to the i resp. j axis in Figure 4. The noisy data is obtained by adding white Gaussian noise with $\sigma = \sqrt{0.001}$ using the built-in function `imnoise` in MATLAB. It is shown in Figure 5(B). We choose as parameters $s = 0.9,\ p = 1.1,\ \alpha = 1$, and $\varepsilon = 0.01$. We observe significant noise reduction in both cases. However, only in Figure 5(D) the color transitions are handled correctly. This is due to the fact, that our regularizer respects the periodicity, i.e. for the functional there is no jump in Figure 4 since 0 and 2π are identified. Using the classical TV regularizer the values 0 and 2π are not identified and have a distance of 2π. Hence, in the TV-denoised image there is a sharp edge in the middle of the image, see Figure 5(C).



(A) *Original data* (B) *Noisy data*

(C) *TV-denoised data* (D) *Denoised data*

FIGURE 5. *Left to right, top to bottom: Original and noisy data of an 60×60 image. TV-denoised data using a fixed point iteration method. Denoised data where we chose $s = 0.9, p = 1.1, \alpha = 1$, 400 steps.*

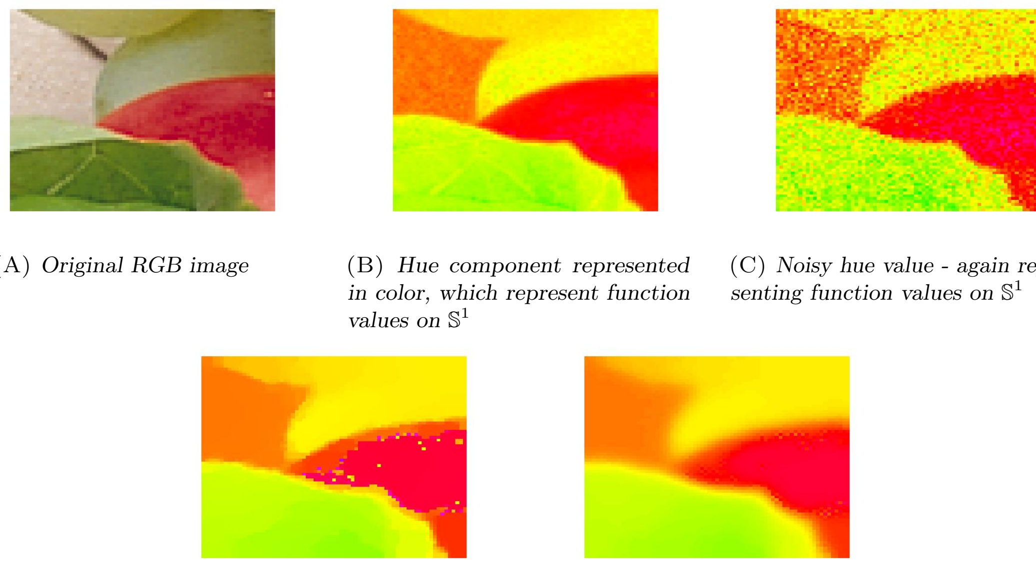
Hue Denoising. The HSV color space is shorthand for Hue, Saturation, Value (of brightness). The hue value of a color image is $\mathbb{S}^1$-valued, while saturation and value of brightness are real-valued. Representing colors in this space better match the human perception than representing colors in the RGB space.

In Figure 6(A) we see a part of size 70×70 of the RGB image “fruits” (https://homepages.cae.wisc.edu/~ece533/images/).

The corresponding hue data is shown in Figure 6(B), where we used again the colormap hsv, cf. Figure 4. Each pixel-value lies, after transformation, in the interval $[0, 2\pi)$ and represents the angular value. Gaussian white noise with $\sigma = \sqrt{0.001}$ is added to obtain a noisy image, see Figure 6(C).
To obtain the denoised image Figure 6(D) we again used the same fixed point iteration, cf. [43], as before.

We see that the denoised image suffers from artifacts due to the non-consideration of periodicity. The pixel-values in the middle of the apple (the red object in the original image) are close to 2π while those close to the border are nearly 0, meaning they have a distance of around 2π.
We use this TV-denoised image as starting image to perform the minimization of our energy functional. As parameters we choose $s = 0.49,\ p = 2,\ \alpha = 2,\ \varepsilon = 0.006$.

Since the cyclic structure is respected the disturbing artifacts in image Figure 6(D) are removed correctly. The edges are smoothed due to the high value of p, see Figure 6(E).



(A) *Original RGB image*

(B) *Hue component represented in color, which represent function values on $\mathbb{S}^1$*

(C) *Noisy hue value - again representing function values on $\mathbb{S}^1$*

(D) *TV-denoised data*

(E) *Denoised data*

Figure 6. *Left to right, top to bottom: Original RGB image and its Hue component. Noisy Hue data with $\sigma^2 = 0.001$. TV minimization is done using an iterative approach. It is serving as starting point for the GD minimization. Denoised data with $s = 0.49, p = 2, \alpha = 2$, 500 steps.*

6.5. $\mathbb{S}^1$-Valued Image Inpainting. In this case the operator $\mathrm{F} : W^{s,p}(\Omega, \mathbb{S}^1) \to L^p(\Omega, \mathbb{S}^1)$ is the inpainting operator, i.e.

$$\mathrm{F}(w) = \chi_{\Omega \setminus D}(w),$$

where $D \subseteq \Omega$ is the area to be inpainted.

We consider the functional

$$\mathcal{F}_\alpha^{v^\delta}[\mathrm{d}_{\mathbb{S}^1}](w) := \int\limits_{\Omega \setminus D} \mathrm{d}_{\mathbb{S}^1}^p(w(x), v^\delta(x))\,\mathrm{d}x + \alpha \int\limits_{\Omega \times \Omega} \frac{\mathrm{d}_{\mathbb{S}^1}^p(w(x), w(y))}{\|x - y\|_{\mathbb{R}^2}^{2+ps}} \rho_\varepsilon(x - y)\,\mathrm{d}(x, y),$$

on $W^{s,p}(\Omega, \mathbb{S}^1)$.

According to Example 3.7 the functional $\mathcal{F}$ is coercive and Assumption 3.1 is satisfied. For $\emptyset \neq \Omega \subset \mathbb{R}$ or $\mathbb{R}^2$ a bounded and simply connected open set, $1 < p < \infty$ and $s \in (0, 1)$ such that additionally $sp < 1$ or $sp \geq 2$ if $N = 2$ Lemma 6.5 applies which ensures that there exists a minimizer $u \in W^{s,p}(\Omega, \mathbb{R})$ of the lifted functional $\tilde{\mathcal{F}}_\alpha^{u^\delta}[\mathrm{d}_{\mathbb{S}^1}] : W^{s,p}(\Omega, \mathbb{R}) \to [0, \infty)$ $u \in W^{s,p}(\Omega, \mathbb{R})$

Inpainting of a $\mathbb{S}^1$-Valued Image. As a first inpainting test-example we consider two $\mathbb{S}^1$-valued images of size 28×28, see Figure 7, represented by its angular values. In both cases the ground truth can be seen in Figure 7(A) and Figure 7(F). We added Gaussian white noise with $\sigma = \sqrt{0.001}$ using the MATLAB build-in function `imnoise`. The noisy images can be seen in Figure 7(B) and Figure 7(G). The region D consists of the nine red squares in Figure 7(C) and Figure 7(H).

The reconstructed data are shown in Figure 7(D) and Figure 7(I).
For the two-colored image we used as parameters $\alpha = s = 0.3$, $p = 1.01$ and $\varepsilon = 0.05$. We see that the reconstructed edge appears sharp. The unknown squares, which are completely surrounded by one color are inpainted perfectly. The blue and green color changed slightly.

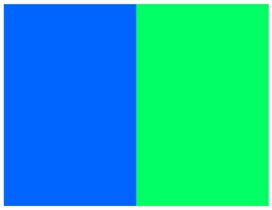
As parameters for the three-colored image we used $\alpha = s = 0.4$, $p = 1.01$ and $\varepsilon = 0.05$. Here again the unknown regions lying entirely in one color are inpainted perfectly. The edges are preserved. Just the corner in the middle of the image is slightly smoothed.
In Figure 7(E) and Figure 7(J) the TV-reconstructed data is shown. The underlying algorithm ([30]) uses the split Bregman method (see [34]).

In Figure 7(E) the edge is not completely sharp. There are some lighter parts on the blue side. This can be caused by the fact that the unknown domain in this area is not exactly symmetric with respect to the edge. This is also the case in Figure 7(J) where we observe the same effect. Unknown squares lying entirely in one color are perfectly inpainted.

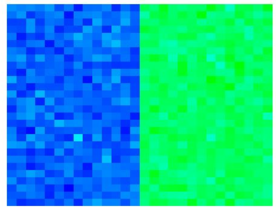
Hue Inpainting. As a last example we consider again the Hue-component of the image "fruits", see Figure 8(A). The unknown region D is the string *01.01* which is shown in Figure 8(B). As parameters we choose $p = 1.1$, $s = 0.1$, $\alpha = 2$ and $\varepsilon = 0.006$. We get the reconstructed image shown in Figure 8(C). The edges are preserved and the unknown area is restored quite well. This can be also observed in the TV reconstructed image, Figure 8(D), using again the split Bregman method as before, cf. [30].

6.6. Conclusion. In this paper we developed a functional for regularization of functions with values in a set of vectors. The regularization functional is a derivative-free, nonlocal term, which is based on a characterization of Sobolev spaces of *intensity data* derived by Bourgain, Brézis, Mironescu & Dávila. Our objective has been to extend their double integral functionals in a natural way to functions with values in a set of vectors, in particular functions with values on an embedded manifold. These new integral representations are used for regularization on a subset of the (fractional) Sobolev space $W^{s,p}(\Omega, \mathbb{R}^M)$ and the space $BV(\Omega, \mathbb{R}^M)$, respectively. We presented numerical results for denoising of artificial InSAR data as well as an example of inpainting. Moreover, several conjectures are at hand on relations between double metric integral regularization functionals and single integral representations.

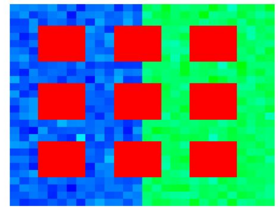
Acknowledgements. We thank Peter Elbau for very helpful discussions and comments. MH and OS acknowledge support from the Austrian Science Fund (FWF) within the national research network Geometry and Simulation, project S11704 (Variational Methods for Imaging on Manifolds). Moreover, OS is supported by the Austrian Science Fund (FWF), with SFB F68, project F6807-N36 (Tomography with Uncertainties) and I3661-N27 (Novel Error Measures and Source Conditions of Regularization Methods for Inverse Problems).



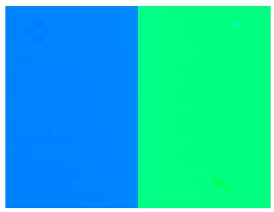
(A) *Original image*



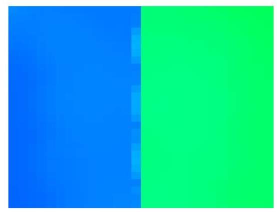
(B) *Noisy image*



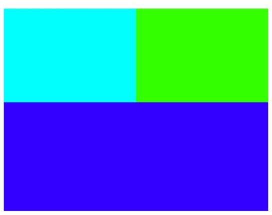
(C) *Noisy masked image*



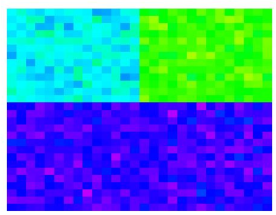
(D) *Reconstructed image*



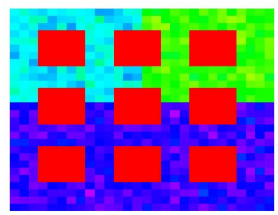
(E) *TV-reconstructed image*



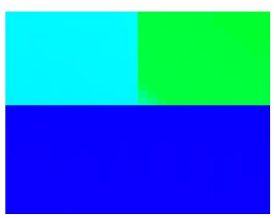
(F) *Original image*



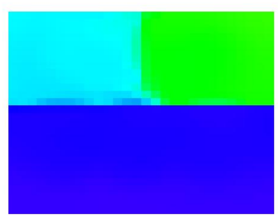
(G) *Noisy image*



(H) *Noisy masked image*

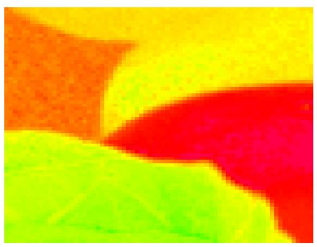


(I) *Reconstructed image*



(J) *TV-reconstructed image*

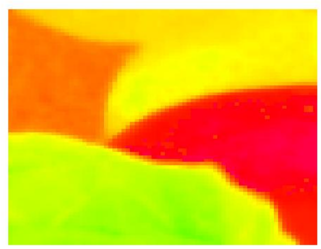
FIGURE 7. *Left to right. Top to bottom: Original image and the noisy data with* $\sigma^2 = 0.001$. *Noisy image with masking filter and denoised data with* $s = 0.3, p = 1.01, \alpha = 0.3$, *6000 steps. TV denoised data.*
Original image and the noisy data with $\sigma^2 = 0.001$. *Noisy image with masking filter and denoised data with* $s = 0.4, p = 1.01, \alpha = 0.4$, *10000 steps. TV denoised image.*



(A) *Hue component*

(B) *Image with masked region*

(C) *TV-reconstructed image*



(D) *Reconstructed image*

FIGURE 8. *Left to right, top to bottom: Original image and image with masked region. Reconstructed image with parameters* $p = 1.1,\ s = 0.1,\ \alpha = 2$ *and* $\varepsilon = 0.006$, *2000 steps. TV-reconstructed image.*

References

[1] R. A. Adams. "Sobolev Spaces". New York: Academic Press, 1975 (cited on page 5).

[2] L. Ambrosio, N. Fusco, and D. Pallara. "Functions of bounded variation and free discontinuity problems". Oxford Mathematical Monographs. New York: Oxford University Press, 2000. xviii+434. isbn: 0-19-850245-1 (cited on page 5).

[3] G. Aubert and P. Kornprobst. "Can the nonlocal characterization of Sobolev spaces by Bourgain et al. be useful for solving variational problems?" In: *SIAM J. Numer. Anal.* 47.2 (2009), pp. 844–860. doi: 10.1137/070696751 (cited on pages 3, 7).

[4] M. Bačák, R. Bergmann, G. Steidl, and A Weinmann. "A second order non-smooth variational model for restoring manifold-valued images". In: *SIAM J. Sci. Comput.* 38.1 (2016), A567–A597. doi: 10.1137/15M101988X (cited on page 3).

[5] R. Bergmann, R. H. Chan, R. Hielscher, J. Persch, and G. Steidl. "Restoration of manifold-valued images by half-quadratic minimization". In: *Inverse Probl. Imaging* 10.2 (2016), 281–304. doi: 10.3934/ipi.2016001 (cited on page 3).

[6] R. Bergmann, J. H. Fitschen, J. Persch, and G. Steidl. "Priors with coupled first and second order differences for manifold-valued image processing". In: (2017) (cited on page 3).

[7] R. Bergmann, F. Laus, G. Steidl, and A. Weinmann. "Second order differences of cyclic data and applications in variational denoising". In: *SIAM J. Imaging Sciences* 7.4 (2014), 2916–2953. doi: 10.1137/140969993 (cited on page 3).

[8] R. Bergmann, J. Persch, and G. Steidl. "A parallel Douglas–Rachford algorithm for restoring images with values in symmetric Hadamard manifolds". In: *SIAM J. Imaging Sciences* 9.3 (2016), pp. 901–937. doi: 10.1137/15M1052858 (cited on page 3).

[9] R. Bergmann and A. Weinmann. "A second order TV-type approach for inpainting and denoising higher dimensional combined cyclic and vector space data". In: *J. Math. Imaging Vision* 55.3 (2016), 401–427. doi: 10.1007/s10851-015-0627-3 (cited on page 3).

[10] R. Bergmann and A. Weinmann. "Inpainting of cyclic data using first and second order differences". In: *EMMCVPR 2015*. Ed. by X.-C. Tai, E. Bae, T. F. Chan, S. Y. Leung, and M. Lysaker. Springer, 2015, 155–168. doi: 10.1007/978-3-319-14612-6_12 (cited on page 3).

[11] J. Boulanger, P. Elbau, C. Pontow, and O. Scherzer. "Non-Local Functionals for Imaging". In: *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. Ed. by H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz. 1st. Vol. 49. Springer Optimization and Its Applications. New York: Springer, 2011, pp. 131–154. isbn: 978-1-4419-9568-1. doi: 10.1007/978-1-4419-9569-8 (cited on pages 3, 7).

[12] C. Bouman and K. Sauer. "A generalized Gaussian image model for edge-preserving MAP estimation". In: *IEEE Trans. Image Process.* 2.3 (1993), pp. 296–310 (cited on page 2).

[13] J. Bourgain, H. Brézis, and P. Mironescu. "Another Look at Sobolev Spaces". In: *Optimal Control and Partial Differential Equations-Innovations & Applications: In honor of Professor Alain Bensoussan's 60th anniversary*. Ed. by J.L. Menaldi, E. Rofman, and A. Sulem. Amsterdam: IOS press, 2001, pp. 439–455 (cited on pages 2, 7, 8, 18, 19).

[14] J. Bourgain, H. Brezis, and P. Mironescu. "Lifting in Sobolev spaces". In: *J. Anal. Math.* 80 (2000), pp. 37–86 (cited on pages 6, 18).

[15] A. Chambolle and P.-L. Lions. "Image recovery via total variation minimization and related problems". In: *Numer. Math.* 76.2 (1997), pp. 167–188 (cited on page 2).

[16] I. Cimrák and V. Melicher. "Mixed Tikhonov regularization in Banach spaces based on domain decomposition". In: *Appl. Math. Comput.* 218.23 (2012), pp. 11583–11596. doi: 10.1016/j.amc.2012.05.042 (cited on page 2).

[17] D. Cremers, S. Koetter, J. Lellmann, and E. Strekalovskiy. "Total Variation Regularization for Functions with Values in a Manifold". In: *IEEE International Conference on Computer Vision, ICCV 2013, Sydney, Australia, December 1-8, 2013*. 2013, pp. 2944–2951. doi: 10.1109/ICCV.2013.366 (cited on page 3).

[18] D. Cremers and E. Strekalovskiy. "Total Cyclic Variation and Generalizations". In: *J. Math. Imaging Vision* 47.3 (2013), pp. 258–277 (cited on page 3).

[19] D. Cremers and E. Strekalovskiy. "Total variation for cyclic structures: Convex relaxation and efficient minimization." In: *CVPR*. IEEE Computer Society, 2011, pp. 1905–1911. ISBN: 978-1-4577-0394-2 (cited on page 3).

[20] B. Dacorogna. "Direct Methods in the Calculus of Variations". Vol. 78. Applied Mathematical Sciences. Berlin: Springer Verlag, 1989 (cited on pages 3, 8).

[21] B. Dacorogna. "Weak Continuity and Weak Lower Semicontinuity of Non-Linear Functionals". Vol. 922. Lecture Notes in Mathematics. Berlin, Heidelberg, New York: Springer Verlag, 1982 (cited on pages 3, 8).

[22] I. Daubechies, M. Defrise, and C. De Mol. "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint". In: *Comm. Pure Appl. Math.* 57.11 (2004), pp. 1413–1457 (cited on page 2).

[23] J. Dávila. "On an open question about functions of bounded variation". In: *Calc. Var. Partial Differential Equations* 15.4 (2002), pp. 519–527 (cited on pages 2, 7).

[24] E. Di Nezza, G. Palatucci, and E. Valdinoci. "Hitchhiker's guide to the fractional Sobolev spaces". In: *Bull. Sci. Math.* 136.5 (2012), pp. 521–573. DOI: 10.1016/j.bulsci.2011.12.004 (cited on page 18).

[25] P. P. B. Eggermont. "Maximum entropy regularization for Fredholm integral equations of the first kind". In: *SIAM J. Math. Anal.* 24.6 (1993), pp. 1557–1576 (cited on page 2).

[26] H. W. Engl and G. Landl. "Convergence rates for maximum entropy regularization". In: *SIAM J. Numer. Anal.* 30.5 (1993), pp. 1509–1536 (cited on page 2).

[27] L. C. Evans. "Partial Differential Equations". Second. Vol. 19. Graduate Studies in Mathematics. Providence, RI: American Mathematical Society, 2010. ISBN: 978-0-8218-4974-3 (cited on pages 5, 12).

[28] L. C. Evans and R. F. Gariepy. "Measure theory and fine properties of functions". Studies in Advanced Mathematics. Boca Raton, FL: CRC Press, 1992. viii+268. ISBN: 0-8493-7157-0 (cited on page 5).

[29] A. Figalli and C. Villani. "Optimal transport and curvature". In: *Nonlinear PDE's and applications*. Vol. 2028. Lecture Notes in Math. Springer, Heidelberg, 2011, pp. 171–217. DOI: 10.1007/978-3-642-21861-3_4 (cited on page 17).

[30] P. Getruer. *tvreg*. From MathWorks–File Exchange (cited on page 24).

[31] M. Giaquinta, G. Modica, and J. Souček. "Variational problems for maps of bounded variation with values in S1". In: *Calc. Var. Partial Differential Equations* 1.1 (1993), pp. 87–121. DOI: 10.1007/BF02163266 (cited on page 3).

[32] M. Giaquinta and D. Mucci. "Maps of Bounded Variation with Values into a Manifold: Total Variation and Relaxed Energy". In: *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* 5.4 (2006), pp. 483–548 (cited on page 3).

[33] M. Giaquinta and D. Mucci. "The BV-energy of maps into a manifold: relaxation and density results". In: *Int. J. Pure Appl. Math.* 3.2 (2007), pp. 513–538 (cited on page 3).

[34] T. Goldstein and S. Osher. "The Split Bregman Method for L1-regularized Problems". In: *SIAM J. Imaging Sciences* 2 (2009), pp. 323–343 (cited on page 24).

[35] P. Grohs and M. Sprecher. "Total Variation Regularization by Iteratively Reweighted Least Squares on Hadamard Spaces and the Sphere". Tech. rep. 2014-39. Switzerland: Seminar for Applied Mathematics, ETH Zürich, 2014 (cited on page 3).

[36] R. Hadani and A. Singer. "Representation theoretic patterns in three dimensional Cryo-Electron Microscopy I: The intrinsic reconstitution algorithm". In: *Ann. Math.* 174.2 (2011), pp. 1219–1241. ISSN: 0003486X (cited on page 1).

[37] C.A. Helliwell, R.S. Anderssen, M. Robertson, and E.J. Finnegan. "How is FLC repression initiated by cold?" In: *Trends Plant Sci.* 20 (2015), pp. 76–82 (cited on page 1).

[38] R. Kimmel and N.A. Sochen. "Orientation Diffusion or How to Comb a Porcupine". In: *J. Vis. Commun. Image Represent.* 13.1-2 (2002), pp. 238–248. DOI: 10.1006/jvci.2001.0501 (cited on page 3).

[39] V. Kolehmainen, M. Lassas, K. Niinimäki, and S. Siltanen. "Sparsity-promoting Bayesian inversion". In: *Inverse Probl.* 28.2 (2012), pp. 025005, 28. DOI: 10.1088/0266-5611/28/2/025005 (cited on page 2).

[40] Matti Lassas, Eero Saksman, and Samuli Siltanen. "Discretization-invariant Bayesian inversion and Besov space priors". In: *Inverse Problems and Imaging* 3.1 (2009), pp. 87–122 (cited on page 2).

[41] F. Laus, M. Nikolova, J. Persch, and G. Steidl. "A Nonlocal Denoising Algorithm for Manifold-Valued Images Using Second Order Statistics". In: *SIAM J. Imaging Sciences* 10.1 (2017), pp. 416–448 (cited on page 3).

[42] J.G. Liu and J. Mason. "Image Processing and GIS for Remote Sensing". 2nd ed. London: John Wiley and Sons, 2016 (cited on page 1).

[43] C. Loendahl and P. Magiera. *ROF Denoising Algorithm*. From MathWorks–File Exchange (cited on pages 22, 23).

[44] D. Lorenz and D. Trede. "Optimal convergence rates for Tikhonov regularization in Besov scales". In: *Inverse Probl.* 24.5 (2008), 055010 (14pp) (cited on page 2).

[45] S. Osher and S. Esedoglu. "Decomposition of Images by the anisotropic Rudin–Osher–Fatemi model". In: *Comm. Pure Appl. Math.* 57.12 (2004), pp. 1609–1626 (cited on page 2).

[46] K. Plataniotis and A.N. Venetsanopoulos. "Color Image Processing and Applications". Berlin: Springer, 2000 (cited on page 1).

[47] A. Ponce. "A new approach to Sobolev spaces and connections to Γ-Convergence". In: *Calc. Var. Partial Differential Equations* 19 (2004), pp. 229–255 (cited on pages 2, 7, 8).

[48] C. Pöschl. "Tikhonov Regularization with General Residual Term". English. PhD thesis. Innsbruck: University of Innsbruck, Austria, Oct. 2008 (cited on page 2).

[49] E. Puttonen, C. Briese, G. Mandlburger, M. Wieser, M Pfennigbauer, A. Zlinszky, and N. Pfeifer. "Quantification of Overnight Movement of Birch (Betula pendula) Branches and Foliage with Short Interval Terrestrial Laser Scanning". In: *Front. Plant Sci.* (2016) (cited on page 1).

[50] E. Resmerita and R. S. Anderssen. "Joint additive Kullback–Leibler residual minimization and regularization for linear inverse problems". In: *Math. Methods Appl. Sci.* 30.13 (2007), pp. 1527–1544 (cited on page 2).

[51] F. Rocca, C. Prati, and A. Ferretti. "An Overview of SAR Interferometry". In: *3rd ERS Symposiumm, Florence 97 - Abstracts and Papers*. 1997 (cited on page 1).

[52] L. I. Rudin, S. Osher, and E. Fatemi. "Nonlinear total variation based noise removal algorithms". In: *Phys. D* 60.1–4 (1992), pp. 259–268 (cited on page 2).

[53] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. "Variational methods in imaging". Applied Mathematical Sciences 167. New York: Springer, 2009. ISBN: 978-0-387-30931-6. DOI: 10.1007/978-0-387-69277-7 (cited on pages 2, 10, 12, 14–16).

[54] O. Scherzer and J. Weickert. "Relations between regularization and diffusion filtering". In: *J. Math. Imaging Vision* 12.1 (2000), pp. 43–63. ISSN: 0924-9907. DOI: 10.1023/A:1008344608808 (cited on page 2).

[55] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski. "Regularization methods in Banach spaces". Radon Series on Computational and Applied Mathematics 10. Berlin, Boston: De Gruyter, 2012. xii+283. DOI: 10.1515/9783110255720 (cited on page 2).

[56] A. Singer and Y. Shkolnisky. "Viewing Direction Estimation in Cryo-EM Using Synchronization". In: *SIAM J. Imaging Sciences* 5.3 (2012), pp. 1088–1110. DOI: 10.1137/120863642 (cited on page 1).

[57] A. N. Tikhonov and V. Y. Arsenin. "Solutions of Ill-Posed Problems". Washington, D.C.: John Wiley & Sons, 1977 (cited on page 2).

[58] C. R. Vogel and M. E. Oman. "Iterative methods for total variation denoising". In: *SIAM J. Sci. Comput.* 17 (1996), pp. 227–238 (cited on page 22).

[59] L. Wang, A. Singer, and Z. Wen. "Orientation determination of cryo-EM images using least unsquared deviations". In: *SIAM J. Imaging Sciences* 6.4 (2013), pp. 2450–2483. DOI: 10.1137/130916436 (cited on page 1).

[60] A. Weinmann, L. Demaret, and M. Storath. "Total Variation Regularization for Manifold-Valued Data". In: *SIAM J. Imaging Sciences* 7.4 (2014), pp. 2226–2257 (cited on page 3).