



Integer Optimal Control with Fractional Perimeter Regularization

Harbir Antil¹ · Paul Manns² 

Accepted: 13 June 2024
© The Author(s) 2024

Abstract

Motivated by many applications, optimal control problems with integer controls have recently received a significant attention. Some state-of-the-art work uses perimeter-regularization to derive stationarity conditions and trust-region algorithms. However, the discretization is difficult in this case because the perimeter is concentrated on a set of dimension $d - 1$ for a domain of dimension d . This article proposes a potential way to overcome this challenge by using the fractional nonlocal perimeter with fractional exponent $0 < \alpha < 1$. In this way, the boundary integrals in the perimeter regularization are replaced by volume integrals. Besides establishing some non-trivial properties associated with this perimeter, a Γ -convergence result is derived. This result establishes convergence of minimizers of fractional perimeter-regularized problem, to the standard one, as the exponent α tends to 1. In addition, the stationarity results are derived and algorithmic convergence analysis is carried out for $\alpha \in (0.5, 1)$ under an additional assumption on the gradient of the reduced objective. The theoretical results are supplemented by a preliminary computational experiment. We observe that the isotropy of the total variation may be approximated by means of the fractional perimeter functional.

Keywords Integer optimal control · Fractional perimeter regularization · Trust-region algorithms

✉ Paul Manns
paul.manns@tu-dortmund.de

Harbir Antil
hantil@gmu.edu

¹ Department of Mathematical Sciences and Center for Mathematics and Artificial Intelligence, George Mason University, Fairfax, VA 22030, USA

² Faculty of Mathematics, TU Dortmund University, 44227 Dortmund, Germany

1 Introduction

Let $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$, be a bounded, polyhedral domain. Let $M \in \mathbb{N}$ and $\eta > 0$. This article is concerned with solving minimization problems of the form

$$\begin{aligned} & \underset{w \in L^1(\Omega)}{\text{minimize}} \quad J_\alpha(w) := F(w) + \eta R_\alpha(w) \\ & \text{subject to } w(x) \in W := \{w_1, \dots, w_M\} \subset \mathbb{Z} \text{ a.e. in } \Omega. \end{aligned} \quad (\mathbf{P}_\alpha)$$

Here, $F : L^1(\Omega) \rightarrow \mathbb{R}$, which we assume to be bounded below, is the principle part of the objective that is due to the application and R_α is a regularizer that will provide desirable features of the solution to (\mathbf{P}_α) . The scalar $\alpha \in (0, 1)$ parameterizes the regularizer R_α and in turn (\mathbf{P}_α) . The specific role of α will become clear soon. In order to clarify possible misunderstandings, since the term polyhedron may or may not imply convexity in the literature, the term polyhedron in this paper is used for sets whose boundaries are unions of convex polytopes.

Recent work [18, 22, 25, 29] motivates and analyzes the use of a total variation regularization of w , which corresponds to a penalization of the perimeters of the level sets because

$$\text{TV}(w) \leq \sum_{i=1}^M |w_i| P(w^{-1}(\{w_i\}); \mathbb{R}^d) \leq 2 \max_i |w_i| \text{TV}(w)$$

if $\text{TV}(w) < \infty$, where $P(A; B)$ denotes the perimeter of A in B and $\text{TV}(w)$ denotes the total variation of the function w , where we assume that w is extended by outside of Ω for feasible w in (\mathbf{P}_α) , see [22, Lemma 2.1].

We refer the reader to [19] for extensive information on sets of finite perimeter and their properties. The key property of this regularization that is exploited in the aforementioned publications is the compactness it induces on the control space, specifically bounded sequences of feasible points of (\mathbf{P}_α) , like sequences produced by descent algorithms, have subsequences that converge in $L^1(\Omega)$.

In the multi-dimensional case, $d \geq 2$, the discretization of the subproblems that are proposed in [18, 22] is challenging. The reason is that the arguments in the analysis of the finite difference or finite element discretizations, as are for example carried out in [4, 8, 10], use that w can take values in \mathbb{R} (or at least $\text{conv } W$, the convex hull of W) and that the superordinate minimization problem is convex. Both of these features are not available in our setting. Moreover, any piecewise constant ansatz for w on a fixed decomposition of the domain into say polytopes restricts the geometry of the level sets and therefore introduces a potential gap between (\mathbf{P}_α) and its discretization. This gap is due to the local structure of the perimeter regularization. Specifically, the information on the perimeter is concentrated on the (reduced) boundary of the level sets and hence the discretization, which has finite $d - 1$ -dimensional Hausdorff measure. The very recent work [26] provides a two-level discretization that can overcome this issue but is computationally expensive and efficient implementations that make use of the underlying structure as in [23] are not available so far.

This motivates us to study regularization terms R_α that are close to the perimeter regularization, provide compactness in $L^1(\Omega)$, but also have non-local properties so that they might give a fruitful computational vantage point because they allow to replace the difficult localized boundary integrals: specifically, R_α is given by means of a double volume integral so that its numerical approximation can be improved by improving the quadrature of the volume integrals. Let w be feasible for (P_α) and let $E_i := w^{-1}(\{w_i\})$ for $i \in \{1, \dots, M\}$. Specifically, we consider

$$R_\alpha(w) := (1 - \alpha) \sum_{i=1}^M |w_i| P_\alpha(E_i)$$

for $\alpha \in (0, 1)$, where $P_\alpha(E)$ is the so-called *fractional perimeter*

$$P_\alpha(E) := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{|\chi_E(x) - \chi_E(y)|}{|x - y|^{d+s}} dx dy, \quad (1)$$

introduced and analyzed in [7, 31] and χ_E denotes the $\{0, 1\}$ -valued indicator function of $E \subset \mathbb{R}^d$, see also [3, 12, 13]. The limit problem that we approximate with (P_α) is the perimeter-regularized integer optimal control problem

$$\begin{aligned} & \underset{w \in L^1(\Omega)}{\text{minimize}} \quad J(w) := F(w) + \eta R(w) \\ & \text{subject to } w(x) \in W := \{w_1, \dots, w_M\} \subset \mathbb{Z} \text{ a.e. in } \Omega, \end{aligned} \quad (P)$$

where $R(w) = \omega_{d-1} \sum_{i=1}^M |w_i| P(E_i; \mathbb{R}^d)$ and ω_{d-1} denotes the volume of the unit ball in \mathbb{R}^{d-1} . In particular, this means that our optimization variable w can be modified only on the domain Ω but the regularizer in (P_α) , (P) takes into account the boundary of Ω . In other words, w is implicitly extended by zero outside of Ω and the jumps across $\partial\Omega$ are counted. We note that we are not the first ones to make steps in the direction of computationally exploiting these properties of the fractional perimeter and particularly point to the works [3, 5, 15].

1.1 Contributions

We make the following contributions. The existence of solutions to (P_α) is—in our opinion—not immediate since we are not aware of a Banach–Alaoglu theorem for the Sobolev space $W^{\alpha,1}(\mathbb{R}^d)$ that induces the fractional perimeter as defined in (1). Thus we prove it by means of an argument that exploits the specific structure of our feasible set that only consists of W -valued functions, $|W| < \infty$.

We show compactness in $L^1(\Omega)$ and Γ -convergence for $\alpha \nearrow 1$. For all $\alpha \in (0.5, 1)$, we prove stationarity conditions and asymptotics of a trust-region algorithm parallel to [22]. In order to achieve this, we currently require the regularity assumption $\nabla F(w^n) \in C^2(\bar{\Omega})$ for the iterates w^n produced by the algorithm, which is quite strong compared to the regularity $\nabla F(w^n) \in C(\bar{\Omega})$ that is required for the perimeter-regularized case, see [22].

We also provide a preliminary and qualitative computational experiment, in which we apply the trust-region algorithm for the choices $\alpha = 0.5$ and $\alpha = 0.9$ as well as for $R(w) = \omega_{d-1} \sum_{i=1}^M |w_i| P(E_i; \mathbb{R}^d)$. The geometric restriction induced by the piecewise constant ansatz for our control functions is visible in the limiting case but this behavior is alleviated for $\alpha = 0.5$ and $\alpha = 0.9$. Unfortunately, the subproblem solves in the trust-region algorithm are extremely expensive even for a relatively coarse discretization. Moreover, since computationally tractable discretizations of the subproblems for the limit case are not available so far, it is difficult to interpret and compare the results. Therefore, we emphasize that much more work is needed from a computational point of view in order to provide more efficient discretization and solution algorithms.

1.2 Structure of the Remainder

After introducing some notation and the necessary concepts regarding modes of convergence and local variations of the elements of the feasible set in Sect. 2, we provide the existence of solutions in Sect. 3. Compactness and Γ -convergence are analyzed in Sect. 4. The analysis of local minimizers and the trust-region algorithm and its asymptotics are provided and analyzed in Sect. 5. We provide a computational experiment and discuss its implications in Sect. 6.

2 Notation and Auxiliary Results

If not indicated otherwise, we assume that $\alpha \in (0, 1)$ is fixed but arbitrary in the whole article without further mention. We denote the complement of a set $A \subset \mathbb{R}^d$ by $A^c := \mathbb{R}^d \setminus A$. We denote the symmetric difference between A and a further set $B \subset \mathbb{R}^d$ by $A \Delta B$. We will frequently use the following reformulation of (1). Let $E \subset \mathbb{R}^d$, then $P_\alpha(E)$ satisfies

$$P_\alpha(E) = 2 \int_E \int_{E^c} \frac{1}{|x - y|^{d+\alpha}} \, dx \, dy. \quad (2)$$

We immediately obtain that the function P_α is submodular, see also (2.1) in [9].

Lemma 2.1 *Let $\alpha \in (0, 1)$. Let E, F be measurable subsets of Ω . Then*

$$P_\alpha(E \cap F) + P_\alpha(E \cup F) \leq P_\alpha(E) + P_\alpha(F) \quad (3)$$

and

$$P_\alpha(E \setminus F) + P_\alpha(F \setminus E) \leq P_\alpha(E) + P_\alpha(F). \quad (4)$$

Proof We consider the formulation of $P_\alpha(E)$ from (2) and define $g(x, y) := |x - y|^{-(d+\alpha)}$ for $x, y \in \mathbb{R}^d$. Then inserting the definitions, elementary computations, and

the positivity of g yield

$$\begin{aligned}
 \frac{1}{2}P_\alpha(E \cap F) + \frac{1}{2}P_\alpha(E \cup F) &= \int_{E \cap F} \int_{(E \cap F)^c} g + \int_{E \cup F} \int_{(E \cup F)^c} g \\
 &= \int_{E \cap F} \int_{(E \setminus F) \cup (F \setminus E)} g + \int_E \int_{(E \cup F)^c} g + \int_F \int_{(E \cup F)^c} g \\
 &\leq \int_{E \cap F} \int_{E \setminus F} g + \int_E \int_{E^c} g + \int_F \int_{(E \cup F)^c} g \\
 &\leq \int_E \int_{E^c} g + \int_F \int_{F^c} g = \frac{1}{2}P_\alpha(E; \Omega) + \frac{1}{2}P_\alpha(F; \Omega),
 \end{aligned}$$

which proves (3). In order to see (4), we consider

$$\begin{aligned}
 \frac{1}{2}P_\alpha(E \setminus F) + \frac{1}{2}P_\alpha(F \setminus E) &= \int_{E \setminus F} \int_{(E \setminus F)^c} g + \int_{F \setminus E} \int_{(F \setminus E)^c} g \\
 &= \int_E \int_{E^c} g + \int_{E \setminus F} \int_{E \cap F} g - \int_{E \cap F} \int_{E^c} g \\
 &\quad + \int_F \int_{F^c} g + \int_{F \setminus E} \int_{E \cap F} g - \int_{E \cap F} \int_{F^c} g \\
 &\leq \frac{1}{2}P_\alpha(E) + \frac{1}{2}P_\alpha(F).
 \end{aligned}$$

The proof is complete. \square

We will sometimes switch between the view of a function w that is feasible for (P $_\alpha$) and the partition of Ω that is given by its level sets. If there is no ambiguity, we will denote the level sets by $E_i := w^{-1}(\{w_i\})$, $i \in \{1, \dots, M\}$, without further mention. Let λ denote the Lebesgue measure on \mathbb{R}^d .

Let $\alpha \in (0, 1)$. Then $\nabla^\alpha f$, defined as

$$\nabla^\alpha f(x) = \int_{\mathbb{R}^d} \frac{(y - x) \cdot (f(y) - f(x))}{|y - x|^{d+1+\alpha}} dy$$

for $x \in \mathbb{R}^d$ is the so-called fractional gradient [3, 12, 30] for all f , where the integrand of this integral is an integrable function.

We denote the feasible set by

$$\mathcal{F} := \{w \in L^1(\Omega) \mid w(x) \in W \text{ for a.e. } x \in \Omega\}.$$

2.1 Modes of Convergence

The Gagliardo seminorm of the Sobolev space $W^{\alpha,1}(\mathbb{R}^d)$, see [14], with fractional order of differentiability $\alpha \in (0, 1)$ corresponds to the fractional perimeter P_α as defined above. To the best of the authors' knowledge, this space does not admit a

predual so that, in contrast to $\text{BV}(\mathbb{R}^d)$, there is no weak-* topology that gives existence of limits for bounded subsequences. However, this property can be recovered when restricting to our feasible set \mathcal{F} of a.e. W -valued integrable functions. We therefore refer to this property as pseudo-weakly-* in this article. In Banach spaces that are not uniformly convex, that is they do not have a norm that satisfies a uniform midpoint convexity property, having weak-* convergence or weak convergence together with convergence of the values of the norm does not necessarily imply convergence in norm. Consequently, there is sometimes an important mode of convergence for this situation like so-called strict convergence in $\text{BV}(\Omega)$. In an analogy to this, we define strict convergence for our setting as convergence in $L^1(\Omega)$ in combination with convergence of the regularizer, which of course implies pseudo-weak-* convergence.

Definition 2.2 Let $\alpha \in (0, 1)$ be fixed. We say that $\{w^n\}_n \subset \mathcal{F}$ converges to $w \in \mathcal{F}$ pseudo-weakly-* in \mathcal{F} and write $w^n \xrightarrow{p*} w$ if

$$\begin{aligned} \sum_{i=1}^M w_i \chi_{E_i^n} &= w^n \rightarrow w = \sum_{i=1}^M w_i \chi_{E_i} \text{ in } L^1(\Omega) \text{ and} \\ \sup_{n \in \mathbb{N}} P_\alpha(E_i^n) &< \infty \text{ for all } i \in \{1, \dots, M\}. \end{aligned}$$

We say that $\{w^n\}_n \subset \mathcal{F}$ converges to $w \in \mathcal{F}$ strictly in \mathcal{F} if

$$\begin{aligned} \sum_{i=1}^M w_i \chi_{E_i^n} &= w^n \rightarrow w = \sum_{i=1}^M w_i \chi_{E_i} \text{ in } L^1(\Omega) \text{ and} \\ R_\alpha(w^n) &\rightarrow R_\alpha(w). \end{aligned}$$

We obtain lower semicontinuity with respect to convergence in $L^1(\Omega)$ and in turn also for our regularizer, which we briefly show below.

Lemma 2.3 Let $\alpha \in (0, 1)$. Let $\chi_{E^n} \rightarrow \chi_E$ in $L^1(\Omega)$ for measurable sets $E^n, E \subset \Omega$. Then

$$P_\alpha(E) \leq \liminf_{n \rightarrow \infty} P_\alpha(E^n).$$

In particular, for $w^n \rightarrow w$ in $L^1(\Omega)$ and $w^n, w \in \mathcal{F}$ we obtain

$$R_\alpha(w) \leq \liminf_{n \rightarrow \infty} R_\alpha(w^n).$$

Proof Clearly, the first claim holds if $\liminf_{n \rightarrow \infty} P_\alpha(E^n; \Omega) = \infty$. If this is not the case, we observe that for $s = 0.5\alpha$ and by means of $|\chi_E(x) - \chi_E(y)| = |\chi_E(x) - \chi_E(y)|^2$ it holds that

$$(2) \quad P_\alpha(E) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{|\chi_E(x) - \chi_E(y)|^2}{|x - y|^{d+2s}} dx dy,$$

where we have extended χ_E with the value zero outside of Ω and where the right hand side is the squared Gagliardo seminorm of the Hilbert space $W^{s,2}(\mathbb{R}^d)$, see, for example, [14]. Moreover, the boundedness of the characteristic functions gives $\chi_{E^n} \rightarrow \chi_E$ in $L^2(\Omega)$ too. We infer that all subsequences $\chi_{E^{n_k}}$ such that $P_\alpha(E^{n_k})$ is bounded converge weakly to χ_E . Then the weak lower semicontinuity of the seminorm of $W^{s,2}(\mathbb{R}^d)$ yields the first claim. The second claim follows directly from the first and the definition of R_α . \square

2.2 Local Variations

We follow the ideas presented in [22] in order to derive stationarity conditions for (P α) and obtain a corresponding sufficient decrease that in turn allows to prove convergence of a sufficient decrease condition. Both rely on the analysis of a perturbation of the partition E_1, \dots, E_M . We introduce such perturbations by means of so-called *local variations*, where we follow [22], which in turn is based on [19].

Definition 2.4 (Definition 3.1 in [22])

- (a) A *one-parameter family of diffeomorphisms of \mathbb{R}^d* is a smooth function $f : (-\varepsilon, \varepsilon) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ for some $\varepsilon > 0$ such that for all $t \in (-\varepsilon, \varepsilon)$, the function $f_t(\cdot) := f(t, \cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a diffeomorphism.
- (b) Let $A \subset \mathbb{R}^d$ be open. Then the family $(f_t)_{t \in (-\varepsilon, \varepsilon)}$ is a *local variation in A* if in addition to (a) we have $f_0(x) = x$ for all $x \in \mathbb{R}^d$ and there is a compact set $K \subset A$ such that $\{x \in \mathbb{R}^d \mid f_t(x) \neq x\} \subset K$ for all $t \in (-\varepsilon, \varepsilon)$.
- (c) For a local variation, we define its *initial velocity*: $\phi(x) := \frac{\partial f}{\partial t}(0, x)$ for $x \in \mathbb{R}^d$.

Proposition 2.5 (Proposition 3.2 in [22]) *Let $\phi \in C_c^\infty(\Omega, \mathbb{R}^d)$. Let $f_t := I + t\phi$ for $t \in \mathbb{R}$. Then $(f_t)_{t \in (-\varepsilon, \varepsilon)}$ is a local variation in Ω with initial velocity ϕ for some $\varepsilon > 0$.*

Proof We refer the reader to Proposition 3.2 in [22]. \square

Let $w = \sum_{i=1}^M w_i \chi_{E_i}$ and $(f_t)_{t \in (-\varepsilon, \varepsilon)}$ be a local variation in Ω . Then we define the functions

$$f_t^\# w := \sum_{i=1}^M w_i \chi_{f_t(E_i)}$$

for all $t \in (-\varepsilon, \varepsilon)$. For the results below, we consider an arbitrary but fixed local variation $(f_t)_{t \in (-\varepsilon, \varepsilon)}$ in Ω with initial velocity $\phi \in C_c^\infty(\Omega, \mathbb{R}^d)$.

Proposition 2.6 *Let $\{E_1, \dots, E_M\}$ be a partition of Ω . Then the transformed sets $\{f_t(E_1), \dots, f_t(E_M)\}$ are a partition of Ω for all $t \in (-\varepsilon, \varepsilon)$.*

Proof This follows because the f_t are diffeomorphisms, see also [22, §3]. \square

We mention that local variations also allow for a notion of stability, see Definition 1.6 in [11].

Proposition 2.7 (Variation of the fractional perimeter) *Let $\phi \in C_c^\infty(\Omega; \mathbb{R}^d)$ and let $E \subset \Omega$ satisfy $P_\alpha(E) < \infty$. Then,*

$$P_\alpha(f_t(E)) - P_\alpha(E) = t L_\alpha(E, \phi) + o(t),$$

holds with the definitions

$$L_\alpha(E, \phi) := \int_E \int_{E^c} \frac{C(x, y)}{|x - y|^{d+\alpha}} dx dy < \infty$$

and

$$C(x, y) := \operatorname{div} \phi(x) + \operatorname{div} \phi(y) - (d + \alpha) \frac{(x - y) \cdot (\phi(x) - \phi(y))}{|x - y|^2}.$$

Proof By following the arguments in the proof of Lemma 4.1.1 with the choice $s = \alpha/2$ on p. 182 in [15], we obtain the desired identity

$$P_\alpha(f_t(E)) - P_\alpha(E) = t \int_E \int_{E^c} \frac{C(x, y)}{|x - y|^{d+\alpha}} dx dy + o(t).$$

Before proving that $|C(x, y)|$ is uniformly bounded, which concludes the proof, we briefly extend an argument in the proof of Lemma 4.1.1 in [15]. We first note that by means of Fubini's theorem and the substitution formula, see for example Theorem 263D (v) in [16], we obtain that $P_\alpha(f_t(E))$ is finite if and only if

$$\int_E \int_{E^c} \frac{1}{|f_t(x) - f_t(y)|^{d+\alpha}} |\det Df_t(x)| |\det Df_t(y)| dx dy \quad (5)$$

is finite, where $Df_t = I + t \nabla \phi$.

We do this because we did not immediately see why the $o(\varepsilon)$ -term ($o(t)$ in our notation) in its proof that is due to the remainder term of the Taylor expansion of $|x - y|^{-d-\alpha}$ is an $o(t)$ -term even after integrating. Loosely speaking, why were the authors of [15] able to deduce $\int_E \int_{E^c} o(t) dx dy = o(t)$? Since we did not see the argument directly ourselves, we provide the arguments in more detail below for convenience. They show both that the arguments in [15] are correct and we just required some more steps and, moreover, that the arguments from [15] even allow to prove the claim for the assumed generality $P_\alpha(E) < \infty$.

Applying Lemma 17.4 from [19] and inspecting its proof (note that ϕ has compact support and is Lipschitz), we obtain that there exist $t_0 > 0$ and functions $c_1(t)$, $c_2(t)$ that are uniformly bounded and whose bounds depend only on t_0 and ϕ such that for all t with $|t| < t_0$ we obtain for all $x, y \in \mathbb{R}^d$

$$\begin{aligned} |\det Df_t(x)| |\det Df_t(y)| &= (1 + t \operatorname{div} \phi(x) + t^2 c_1(t))(1 + t \operatorname{div} \phi(y) + t^2 c_1(t)) \\ &= (1 + t \operatorname{div} \phi(x))(1 + t \operatorname{div} \phi(y)) + t^2 c_2(t). \end{aligned} \quad (6)$$

Moreover, we deduce for $x \neq y$

$$\begin{aligned} \frac{1}{|f_t(x) - f_t(y)|^{d+\alpha}} &= \frac{1}{|x - y + t(\phi(x) - \phi(y))|^{d+\alpha}} \\ &= \frac{1}{|x - y|^{d+\alpha}} - t(d + \alpha) \underbrace{\frac{(x - y)^T(\phi(x) - \phi(y))}{|x - y|^{d+\alpha+2}}}_{=: \ell(x, y)} + \tilde{r}(x, y), \end{aligned}$$

where the remainder $\tilde{r}(x, y)$ term of the Taylor expansion of $z \mapsto |z|^{-d-\alpha}$ at $x - y$ is given by

$$\tilde{r}(x, y) = t^2(\phi(x) - \phi(y))^T \left[\frac{(d + \alpha + 2)(d + \alpha)}{|\xi|^{d+\alpha+4}} \xi \xi^T - \frac{d + \alpha}{|\xi|^{d+\alpha+2}} I \right] (\phi(x) - \phi(y))$$

for some $\xi = \xi(x, y)$ in the line segment between $[x - y, x - y + t(\phi(x) - \phi(y))]$. Because ϕ is compactly supported and Lipschitz continuous with Lipschitz constant L , there exists $0 < t_1 \leq t_0$ so that for all $|t| \leq t_1$ it holds that $|\xi(x, y)| \geq \frac{1}{2}|x - y|$ uniformly for all $x, y \in \mathbb{R}^d$. Combining this with the Lipschitz continuity of ϕ , we deduce

$$\tilde{r}(x, y) \leq t^2 L^2 \left[\frac{(d + \alpha + 2)(d + \alpha)}{2^{d+\alpha+4}} + \frac{d + \alpha}{2^{d+\alpha+2}} \right] \frac{1}{|x - y|^{d+\alpha}}.$$

Computing the double integral over the remainder term of the Taylor expansion implies for $r(t) := \int_E \int_{E^c} \tilde{r}(x, y) \, dx \, dy$ that

$$|r(t)| = \left| \int_E \int_{E^c} \tilde{r}(x, y) \, dx \, dy \right| \leq t^2 c_3 P_\alpha(E)$$

for some $c_3 > 0$ and all $|t| \leq t_1$, which gives that $r(t) = o(t)$. With a similar argument, one can deduce that there exists $c_4 > 0$ such that

$$\left| \int_E \int_{E^c} \ell(x, y) \, dx \, dy \right| \leq c_4 P_\alpha(E).$$

Combining these considerations with (6) and using that ϕ is compactly supported therein, we obtain that (5) is finite and in turn also $P_\alpha(f_t(E))$ is finite for all t with $|t| \leq t_1$.

Due to these arguments, we can combine all $o(t)$ terms and obtain the claimed formula for $P_\alpha(f_t(E)) - P_\alpha(E)$. It remains to show that $|C(x, y)|$ is uniformly bounded on $E \times E^c$. Because $\phi \in C_c^\infty(\Omega; \mathbb{R}^d)$, $|\operatorname{div} \phi(x)|$ and $|\operatorname{div} \phi(y)|$ are uniformly bounded. Moreover, because $\phi \in C_c^\infty(\Omega; \mathbb{R}^d)$ is globally Lipschitz with Lipschitz constant $L \geq 0$, we obtain that the absolute value of the third term is bounded by $L(d + \alpha)$ in combination with the Cauchy–Schwarz inequality. \square

We note that an alternative proof of this claim follows from the arguments in Sect. 3 of [20]. We also note that the strategy above uses the same basic steps as the arguments in Sect. 17 of [19] for the limit case $\alpha = 1$.

Lemma 2.8 *Let $\phi \in C_c^\infty(\Omega, \mathbb{R}^d)$. Let $(f_t)_{t \in (-\varepsilon, \varepsilon)}$ be the local variation defined by $f_t := I + t\phi$ for $t \in (-\varepsilon, \varepsilon)$. Then there exist $0 < \varepsilon_0 < \varepsilon$ and $L > 0$ such that for all $s, t \in (-\varepsilon_0, \varepsilon_0)$, and measurable sets $E, F \subset \Omega$ with $P_\alpha(E) < \infty$ it holds that*

$$|\lambda(f_t(E) \cap F) - \lambda(E \cap F)| \leq L|t|^\alpha P_\alpha(E)$$

and in particular,

$$\lambda(f_t(E) \Delta E) \leq L|t|^\alpha P_\alpha(E).$$

Proof We follow the proof strategy of Lemma 17.9 in [19]. To this end, let $g_t := f_t^{-1}$ for all $t \in (-\varepsilon, \varepsilon)$ and let $s, t \in (-\varepsilon, \varepsilon)$. Let $u_\delta \rightarrow \chi_E$ in $W^{\alpha,1}(\mathbb{R}^d)$ with $u_\delta \in C_c^\infty(\mathbb{R}^d)$, which exists by Theorem A.1 in [12]. Then we obtain

$$|\lambda(f_t(E) \cap F) - \lambda(E \cap F)| \leq \int_F |\chi_E(g_t(x)) - \chi_E(x)| \, dx \quad (7)$$

$$= \lim_{\delta \rightarrow 0} \underbrace{\int_F |u_\delta(g_t(x)) - u_\delta(x)| \, dx}_{=: d_\delta}, \quad (8)$$

where we note that the right hand side of the inequality (7) is identical to $\lambda(f_t(E) \Delta E)$ for the choice $F = \Omega$ so that the second claim follows from the first in this case.

Because the function u_δ is smooth, we can follow the proof of Proposition 3.14 in [12] (which in turn relies on the fractional fundamental theorem of calculus as it is given in Theorem 3.12 in [12]) in order to obtain

$$d_\delta \leq \gamma_{d,\alpha} \sup_{x \in \mathbb{R}^d} \|\phi_t(x) - x\|^\alpha \|\nabla^\alpha u_\delta\|_{L^1(\mathbb{R}^d, \mathbb{R}^d)},$$

where y in the proof of Proposition 3.14 in [12] is replaced by $\phi_t(x) - x$, Hölder's inequality is applied to obtain the term $\sup_{x \in \mathbb{R}^d} \|\phi_t(x) - x\|^\alpha$ and the constant $\gamma_{d,\alpha}$ is from Proposition 3.14 in [12] too. Then, the identity

$$\sup_{x \in \mathbb{R}^d} \|\phi_t(x) - x\| = |t| \sup_{x \in \mathbb{R}^d} \|\phi(x)\|$$

and the fact that $\|\phi(x)\|$ is bounded because $\phi \in C_c^\infty(\Omega, \mathbb{R}^d)$ yield

$$d_\delta \leq c \gamma_{d,\alpha} |t|^\alpha \|\nabla^\alpha u_\delta\|_{L^1(\mathbb{R}^d, \mathbb{R}^d)} \leq c \gamma_{d,\alpha} \mu_{d,\alpha} |t|^\alpha \underbrace{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{|u_\delta(x) - u_\delta(y)|}{|x - y|^{d+\alpha}} \, dx \, dy}_{=: [u_\delta]_{W^{\alpha,1}(\mathbb{R}^d)}} \quad$$

for some $c > 0$ and the constant $\mu_{d,\alpha}$ from (1.2) in [12].

Since $[u_\delta]_{W^{\alpha,1}(\mathbb{R}^d)}$ is the seminorm of $W^{\alpha,1}(\mathbb{R}^d)$ and u_δ approximates χ_E in norm in $W^{\alpha,1}(\mathbb{R}^d)$, we obtain

$$[u_\delta]_{W^{\alpha,1}(\mathbb{R}^d)} \rightarrow [\chi_E]_{W^{\alpha,1}(\mathbb{R}^d)} = P_\alpha(E).$$

□

Proposition 2.9 (Variation of the linearized objective) *Let $g \in C^2(\bar{\Omega})$. Let $E \subset \Omega$ satisfy $P_\alpha(E) < \infty$. Let $\phi \in C_c^\infty(\Omega, \mathbb{R}^d)$. Let $(f_t)_{t \in (-\varepsilon, \varepsilon)}$ be the local variation defined by $f_t := I + t\phi$ for $t \in (-\varepsilon, \varepsilon)$. Then,*

$$(g, \chi_{f_t(E)} - \chi_E)_{L^2} = t \int_E \operatorname{div}(g(x)\phi(x)) \, dx + O(t^2).$$

Proof The (first half of the) arguments in the proof of Proposition 17.8 in [19] imply

$$(g, \chi_{f_t(E)} - \chi_E)_{L^2} = t \int_E \operatorname{div}(g(x)\phi(x)) \, dx + O(t^2),$$

where $\int_E \operatorname{div}(g(x)\phi(x)) \, dx$ exists and is finite because of the assumed regularity of g . □

Remark 2.10 We note that the assumed regularity on g for the Taylor expansion of the Lebesgue measure in Proposition 2.5 may be deemed to be unrealistically high. Thus improving the required regularity is an important question for further research, in particular, because the much weaker requirement $g \in C(\bar{\Omega})$ is sufficient for proving a Taylor expansion in the non-local case, see [19, Proposition 17.8]. We have, however, not found a means to do so until this point.

3 Standing Assumptions and Existence of Solutions

We provide two standing assumptions on our problem that allow us to deduce the existence of solutions as well as first-order optimality conditions for (P_α) and (P) .

Assumption 3.1

- A.1 Let $F : L^1(\Omega) \rightarrow \mathbb{R}$ be bounded below.
- A.2 Let $F : L^2(\Omega) \rightarrow \mathbb{R}$ be twice continuously Fréchet differentiable. For some $C > 0$ and all $\xi \in L^2(\Omega)$, let the bilinear form induced by the Hessian $\nabla^2 F(\xi) : L^2(\Omega) \times L^2(\Omega) \rightarrow \mathbb{R}$ satisfy $|\nabla^2 F(\xi)(u, w)| \leq C\|u\|_{L^1}\|w\|_{L^1}$.

The second part of Assumption 3.1 is rather restrictive. The estimate on the Hessian with respect to the L^1 -norms is, for example, not satisfied if $F(w) = \frac{1}{2}\|w\|_{L^2}^2$ and it implies that F involves some Lipschitz operation that maps $L^1(\Omega)$ to a smaller space. This may be a convolution operator or a solution operator of a PDE, see also the discussions in [18, 21, 22, 24].

The existence of minimizers for the limit problem (P) under Assumption 3.1 follows as in [18, 22]. The existence of minimizers for the problems (P_α) follows from the compactness and lower semicontinuity properties of nonlocal perimeters, see, for example, [13, Sect. 3.7]. We briefly recap how the existence is achieved below.

Proposition 3.2 *Let $\alpha \in (0, 1)$. Let Assumption 3.1 hold. Then (P_α) admits a minimizer $w = \sum_{i=1}^M w_i \chi_{E_i}$.*

Proof We apply the direct method of calculus of variations. There is a real infimum of (P_α) because all terms of the objective are bounded. We thus consider a minimizing sequence $w^n = \sum_{i=1}^M w_i \chi_{E_i^n}$ with a corresponding sequence of partitions $(\{E_1^n, \dots, E_M^n\})_n$ of Ω . Because all terms of the objective are bounded below, we obtain that the sequences $(P_\alpha(E_i^n))_n$ are bounded. Consequently, the sequences $(\chi_{E_i^n})_n$ are bounded in a space that, similar to $\text{BV}(\Omega)$, admits sequential weak-* compactness properties, see [12, Corollary 4.6 and Proposition 4.8], which yields a limit function $w = \sum_{i=1}^M w_i \chi_{E_i}$ such that $\chi_{E_i^n} \rightarrow \chi_{E_i}$ in $L^1(\Omega)$ for all $i \in \{1, \dots, M\}$. Because convergence in $L^1(\Omega)$ implies pointwise a.e. convergence for a subsequence, we obtain that the limit sets E_i are a partition of Ω except for a set of Lebesgue measure zero. In other words, $w \in \mathcal{F}$. Moreover, Lebesgue's dominated convergence theorem gives $w^n \rightarrow w$ in $L^2(\Omega)$.

Because R_α is lower semicontinuous with respect to convergence in $L^1(\Omega)$ on our feasible set, see Lemma 2.3, and F is continuous, see Assumption 3.1, we obtain that w minimizes (P_α) .

4 Compactness and Γ -Convergence for $\alpha \nearrow 1$

We deduce compactness in $L^1(\Omega)$ and a Γ -convergence-type result from the results in [2]. It is immediate that $P_\alpha(E)$ as defined in (1) equals $\mathcal{J}_s^1(E, \mathbb{R}^d)$ in [2] and $\mathcal{J}_s^2(E, \mathbb{R}^d)$ in [2] equals zero for all $E \subset \Omega$. We start with the compactness result, continue with the \liminf - and \limsup -inequalities for Γ -convergence, and conclude that global minimizers converge to global minimizers.

Proposition 4.1 (Corollary of Theorem 1 in [2]) *Let Assumption 3.1 hold. Let $\alpha \nearrow 1$. Let w^α be feasible for (P_α) . Let $\sup_{\alpha \nearrow 1} J_\alpha(w^\alpha) < \infty$. Then $\{w^\alpha\}_\alpha$ is relatively compact in $L^1(\Omega)$, admits a limit point, and all limit points are W -valued.*

Proof Because F and the summands in R_α are bounded below it follows that $\sup_{\alpha \nearrow 1} (1-\alpha) P_\alpha(E_i^\alpha) < \infty$ for all $i \in \{1, \dots, M\}$, where $E_i^\alpha = (w^\alpha)^{-1}(\{w_i\})$. Consequently, we can apply Theorem 1 in [2] in order to obtain that for all $i \in \{1, \dots, M\}$, the sequence $(\chi_{E_i^\alpha})_\alpha$ is relatively compact in $L_{loc}^1(\mathbb{R}^d)$ and in turn in $L^1(\Omega)$. Recalling $w^\alpha = \sum_{i=1}^M w_i \chi_{E_i^\alpha}$, we obtain that the sequence $(w^\alpha)_\alpha$ is relatively compact. Because convergence in $L^1(\Omega)$ implies pointwise convergence a.e. for a subsequence, we obtain that the limit is W -valued, see also [18]. \square

Proposition 4.2 (Corollary of Theorem 2 in [2]) *Let $w = \sum_{i=1}^M w_i \chi_{E_i}$ for a partition $\{E_1, \dots, E_M\}$ of Ω into measurable sets E_i . Then:*

1. if $w^\alpha = \sum_{i=1}^M w_i \chi_{E_i^\alpha} \rightarrow w$ in $L^1(\Omega)$ for $i \in \{1, \dots, M\}$ for partitions $\{E_1^\alpha, \dots, E_M^\alpha\}$ of Ω , it holds that $J(w) \leq \liminf_{\alpha \nearrow 1} J_\alpha(w)$,
2. there is a sequence $w^\alpha = \sum_{i=1}^M w_i \chi_{E_i^\alpha} \rightarrow w$ in $L^1(\Omega)$ for $i \in \{1, \dots, M\}$ for partitions $\{E_1^\alpha, \dots, E_M^\alpha\}$ of Ω such that $J(w) \geq \limsup_{\alpha \nearrow 1} J_\alpha(w^\alpha)$.

Proof The first claim (lim inf-inequality) follows directly from Theorem 2 in [2]. The second claim (lim sup-inequality) is immediate for $R(w) = \infty$. If $R(w) < \infty$, the partition $\{E_1, \dots, E_M\}$ is a so-called Caccioppoli partition.

Then $\{E_1, \dots, E_M\}$ is a Caccioppoli partition of Ω and [6] asserts that polyhedral partitions are dense in the Caccioppoli partitions, that is there exist sets $\{T_1^k, \dots, T_M^k\}_k \subset \mathbb{R}^d$ whose boundaries are composed of finitely many convex polytopes such that

$$\sum_{i=1}^M w_i \chi_{T_i^k} \rightarrow \sum_{i=1}^M w_i \chi_{E_i} \text{ in } L^1(\Omega) \text{ for } k \rightarrow \infty, \quad (9)$$

$$P(T_i^k \cap \Omega; \mathbb{R}^d) \rightarrow P(E_i; \mathbb{R}^d) \text{ for } k \rightarrow \infty \text{ and all } i \in \{1, \dots, M\}, \quad (10)$$

where the second convergence follows from Corollary 2.5 in [6] in combination with $P(T_i^k \cap \Omega; \mathbb{R}^d) = \mathcal{H}^{d-1}(\Omega \cap \partial^* T_i^k) + \mathcal{H}^{d-1}(\partial \Omega \cap \overline{T_i^k})$, where $\partial^* A$ of a set $A \subset \mathbb{R}^d$ denotes its reduced boundary.

Because of (9) and (10) we can assume that all T_i^k are contained in a bounded *hold-all domain*, e.g., a large ball. Because the sets $T_i^k \cap \Omega$ are polyhedral (recall that we assumed that Ω is polyhedral at the beginning of the article), we can apply Lemmas 8 and 9 in [2] to them and obtain

$$\lim_{\alpha \nearrow 1} (1 - \alpha) P_\alpha(T_i^k \cap \Omega) = \omega_{n-1} P(T_i^k \cap \Omega). \quad (11)$$

Combining (10) and (11), we choose a suitable diagonal sequence indexed by $(k_n)_n$ of the polyhedral partitions $\{T_1^{n_k} \cap \Omega, \dots, T_M^{n_k} \cap \Omega\}_k$ in order to obtain

$$\lim_{\alpha \nearrow 1} P_\alpha(T_i^{n_k} \cap \Omega; \Omega) = \omega_{n-1} P(E_i; \Omega),$$

which implies the assertion. \square

We note that the proof of Proposition 4.2 is the only point in this article, where we use the assumption that the domain Ω is polyhedral. As a corollary of the compactness established in Proposition 4.1 and the liminf- and limsup-inequalities established in Proposition 4.2, we obtain the convergence of global minimizers to global minimizers below.

Corollary 4.3 *Let Assumption 3.1 hold. Let $\alpha \nearrow 1$. Let w^α be a global minimizer of (P_α) for α . Then there exists a W -valued accumulation point $w \in L^1(\Omega)$ with $\tilde{w} \in \text{BV}(\mathbb{R}^d)$, where \tilde{w} is the extension of w by zero outside of Ω , $w(x) \in W$ a.e., such that $w^\alpha \rightarrow w$ in $L^1(\Omega)$ and $J_\alpha(w^\alpha) \rightarrow J(w)$. Moreover, for all accumulation points of $(w^\alpha)_\alpha$ are W -valued, global minimizers of (P) , and their extensions by zero $(\tilde{w}^\alpha)_\alpha$ outside of Ω are in $\text{BV}(\mathbb{R}^d)$.*

5 Local Minimizers and Trust-Region Algorithm

As is noted in [18, 22], it makes sense to consider local minimizers and stationary points in the settings of (P_α) and (P) because L^1 -neighborhoods of feasible points contain further feasible points. We briefly translate these concepts from the perimeter-regularized case, see [22], to our setting in Sect. 5.1. Then we introduce and analyze non-local variants of the trust-region subproblem from [22] in Sect. 5.2. We introduce and describe a trust-region algorithm that builds on these subproblems in Sect. 5.3 and prove its asymptotics in Sect. 5.4.

5.1 Local Minimizers and Stationary Points

We start by defining local minimizers and stationary points and then verify that local minimizers are stationary.

Definition 5.1 Let $\alpha \in (0, 1)$ and let $w = \sum_{i=1}^M w_i \chi_{E_i}$ be feasible for (P_α) .

- We say that w is *locally optimal* for (P_α) if there is $r > 0$ such that $J_\alpha(w) \leq J_\alpha(v)$ for all v that are feasible for (P_α) and satisfy $\|v - w\|_{L^1} \leq r$.
- Let $\alpha \in (0.5, 1)$. Let $\nabla F(w) \in C^2(\bar{\Omega})$. We say that w is *stationary* for (P_α) if

$$\sum_{i=1}^M w_i \left(\int_{E_i} \operatorname{div}(\nabla F(w)(x)\phi(x)) \, dx + L_\alpha(E_i, \phi) \right) = 0 \quad (12)$$

for all $\phi \in C_c^\infty(\Omega; \mathbb{R}^d)$.

Proposition 5.2 Let Assumption 3.1 hold. Let $\alpha \in (0.5, 1)$. $w = \sum_{i=1}^M w_i \chi_{E_i}$ be locally optimal for (P_α) . Then w is stationary for (P_α) .

Proof Let $\phi \in C_c^\infty(\Omega, \mathbb{R}^d)$. Let $(f_t)_{t \in (-\varepsilon, \varepsilon)}$ for some $\varepsilon > 0$ be the local variation defined by $f_t := I + t\phi$. the function $t \mapsto F(f_t^\# w) + \eta R_\alpha(f_t^\# w)$ is differentiable at $t = 0$, which can be seen as follows. Assumption 3.1 implies that

$$\frac{d}{dt} F(f_t^\# w) \Big|_{t=0} = \frac{d}{dt} (\nabla F(w), f_t^\# w)_{L^2} \Big|_{t=0} + \frac{d}{dt} \underbrace{\nabla^2 F(\xi^t)(f_t^\# w - w, f_t^\# w - w)}_{=: r_t} \Big|_{t=0}$$

for some ξ^t in the line segment between w and $f_t^\# w$. Assumption 3.1 and Lemma 2.8 imply

$$\left| \nabla^2 F(\xi^t)(f_t^\# w - w, f_t^\# w - w) \right| \leq C |t|^{2\alpha} \sum_{i=1}^M P_\alpha(E_i; \Omega)$$

for some large enough $C > 0$ and all $t \in (-\varepsilon, \varepsilon)$. Because $\alpha > 0.5$, we obtain that r_t is differentiable at $t = 0$ with value zero so that

$$\frac{d}{dt} F(f_t^\# w) \Big|_{t=0} = \frac{d}{dt} (\nabla F(w), f_t^\# w)_{L^2} \Big|_{t=0}.$$

Then (12) follows from Propositions 2.7 and 2.9. \square

5.2 Trust-Region Subproblems

We introduce trust-region and analyze subproblems by following the ideas from [18, 22], that is the principal part of the objective enters the trust-region subproblem by means of a linear model and the regularization term is considered exactly. We analyze Γ -convergence of the trust-region subproblems with respect to convergence of the linearization point, and provide optimality conditions for the trust-region subproblem.

Let $\Delta > 0$ and \bar{w} be feasible for (P_α) with $R_\alpha(\bar{w}) < \infty$. The trust-region subproblem reads

$$TR_\alpha(\bar{w}, g, \Delta) := \begin{cases} \min_{w \in L^1(\Omega)} (g, w - \bar{w})_{L^2} + \eta R_\alpha(w) - \eta R_\alpha(\bar{w}) \\ \text{s.t. } \|w - \bar{w}\|_{L^1} \leq \Delta, \\ w(x) \in W \text{ for a.e. } x \in \Omega, \end{cases} \quad (TR_\alpha)$$

where we recover the linearized principal part of the objective of (P_α) with the choice $g = \nabla F(\bar{w})$. The trust-region subproblem $TR_\alpha(\bar{w}, g, \Delta)$ admits a minimizer, which we briefly show below.

Proposition 5.3 *Let \bar{w} be feasible for (P_α) with $R_\alpha(\bar{w}) < \infty$, $g \in L^2(\Omega)$, and $\Delta \geq 0$. Then $TR_\alpha(\bar{w}, g, \Delta)$ admits a minimizer.*

Proof Because $g \in L^2(\Omega)$ and $w, \bar{w} \in L^\infty(\Omega)$, the first term of the objective of $TR_\alpha(\bar{w}, g, \Delta)$ is bounded below. Because of the $L^\infty(\Omega)$ -bounds (W is a finite set), convergence in $L^1(\Omega)$ of feasible points implies convergence in $L^2(\Omega)$ and we obtain continuity of the first term of the objective if a sequence of feasible points converges in $L^1(\Omega)$. Moreover, the term $\eta R_\alpha(\bar{w})$ is constant. Consequently, the assumptions of Proposition 3.2 are satisfied on the non-empty feasible set of $TR_\alpha(\bar{w}, g, \Delta)$. Thus the existence of solutions to (P_α) follows as a corollary from (the proof of) Proposition 3.2. \square

Moreover, if the linearization point minimizes $TR_\alpha(w, \nabla F(w), \Delta)$, it is stationary for (P_α) too.

Proposition 5.4 *Let $\alpha \in (0.5, 1)$. Let $\{E_1, \dots, E_M\}$ be a partition of Ω such that $w = \sum_{i=1}^M w_i \chi_{E_i}$ satisfies $R_\alpha(w) < \infty$. Let $\nabla F(w) \in C^2(\bar{\Omega})$. If w is locally optimal for $TR_\alpha(w, \nabla F(w), \Delta)$ for some $\Delta > 0$, then w is stationary for (P_α) .*

Proof Let $g := \nabla F(w)$. We choose $\tilde{F}(v) := (g, v)_{L^2}$ for $v \in L^2(\Omega)$ and obtain $\nabla \tilde{F}(v) = g$ and $\nabla^2 \tilde{F}(v) = 0$ so that \tilde{F} satisfies Assumption 3.1 on \mathcal{F} . We apply Proposition 5.2 with \tilde{F} for F and obtain that w is stationary and satisfies (12) with $\nabla \tilde{F}(w) = g = \nabla F(w)$, which means that w is also stationary for (P_α) . \square

Next, we analyze Γ -convergence of the trust-region subproblems with respect to strict and pseudo-weak-* convergence of feasible points of (P_α) . As in [22], this will be a key ingredient of our convergence analysis of our trust-region algorithm.

Theorem 5.5 Let $\alpha \in (0, 1)$. Let $v^n \rightarrow v$ strictly in \mathcal{F} . Let $g^n \rightarrow g$ in $L^2(\Omega)$. Let $\Delta > 0$. Then the functionals $T^n : (\mathcal{F}, \text{pseudo-weak-}^*) \rightarrow \mathbb{R}$, defined as

$$T^n(w) := (g^n, w - v^n)_{L^2} + \eta R_\alpha(w) - R_\alpha(v^n) + \delta_{[0, \Delta]}(\|w - v^n\|_{L^1})$$

for $w \in \mathcal{F}$, Γ -converge to $T : (\mathcal{F}, \text{pseudo-weak-}^*) \rightarrow \mathbb{R}$, defined as

$$T(w) := (g, w - v)_{L^2} + \eta R_\alpha(w) - R_\alpha(v) + \delta_{[0, \Delta]}(\|w - v\|_{L^1})$$

for $w \in \mathcal{F}$, where $\delta_{[0, \Delta]}$ is the $\{0, \infty\}$ -valued indicator function of $[0, \Delta]$.

Proof We follow the proof strategy of [22, Theorem 5.2].

Part 1: Lower bound inequality. $T(w) \leq \liminf_{n \rightarrow \infty} T^n(w^n)$ for $w^n \xrightarrow{p^*} w$ in \mathcal{F} . Because of the uniform $L^\infty(\Omega)$ -bounds on \mathcal{F} , we obtain $w^n \rightarrow w$ in $L^2(\Omega)$ and $v^n \rightarrow v$ in $L^2(\Omega)$. In combination with $g^n \rightarrow g$ in $L^2(\Omega)$, we obtain $(g^n, w^n - v^n)_{L^2} \rightarrow (g, w - v)_{L^2}$.

Because $w^n \xrightarrow{p^*} w$ in \mathcal{F} and $v^n \rightarrow v$ strictly in \mathcal{F} , we obtain with the help of Lemma 2.3 that $R_\alpha(w) - R_\alpha(v) \leq \liminf_{n \rightarrow \infty} R_\alpha(w^n) - R_\alpha(v^n)$. Moreover, if $\|w^{n_k} - v^{n_k}\|_{L^1} \leq \Delta$ holds for an infinite subsequence, then the convergence of $\{w^n\}_n$ and $\{v^n\}_n$ in $L^1(\Omega)$ and the triangle inequality imply $\|w - v\|_{L^1} \leq \Delta$ so that the last term of T is zero and the lower bound inequality is satisfied. If there is no such subsequence, then $T^n \equiv \infty$ and the lower bound inequality holds trivially.

Part 2: Upper bound inequality. For each $w \in \mathcal{F}$ with $R_\alpha(w) < \infty$, there exists a sequence $w^n \xrightarrow{p^*} w$ in \mathcal{F} such that $T(w) \geq \limsup_{n \rightarrow \infty} T^n(w^n)$. We make a case distinction on the possible values of the norm difference $\|w - v\|_{L^1}$.

Case 2a $\|w - v\|_{L^1} > \Delta$: Then $T(w) = \infty$ and we can choose $w^n := w$ for all $n \in \mathbb{N}$.

Case 2b $\|w - v\|_{L^1} < \Delta$: We choose again $w^n := w$ for all $n \in \mathbb{N}$. We obtain

$$\begin{aligned} (g^n, w^n - v^n)_{L^2} + \eta R_\alpha(w^n) - R_\alpha(v^n) &\rightarrow (g, w - v)_{L^2} \\ &+ R_\alpha(w) - R_\alpha(v) + \underbrace{\delta_{[0, \Delta]}(\|w - v\|_{L^1})}_{=0}. \end{aligned}$$

The convergence of $\{w^n\}_n$ and $\{v^n\}_n$ in $L^1(\Omega)$ and the triangle inequality imply that $\|w^n - v^n\|_{L^1} \leq \Delta$ holds for all large enough $n \in \mathbb{N}$. Consequently, $T^n(w^n) \rightarrow T(w)$.

Case 2c $\|w - v\|_{L^1} = \Delta$: Because $\Delta > 0$, there exists a set

$$D := \{x \in \Omega \mid v(x) = w_1 \neq w_2 = w(x)\}$$

with $\lambda(D) > 0$. We note that the specific values w_1 and w_2 are without loss of generality because we may reorder the indices of the elements of W as necessary. The set D satisfies

$$\begin{aligned}
P_\alpha(D) &= P_\alpha(v^{-1}(\{w_1\}) \cap w^{-1}(\{w_2\})) \\
&\leq P_\alpha(v^{-1}(\{w_1\})) + P_\alpha(w^{-1}(\{w_2\})) \\
&\leq R_\alpha(v) + R_\alpha(w) < \infty,
\end{aligned}$$

where the first inequality follows from (3). The second inequality follows from the fact that at most one of the w_i can be zero. Because D has strictly positive Lebesgue measure $\lambda(D) > 0$, it has a point of density 1, that is there exists $\bar{x} \in D$ such that

$$\lim_{r \searrow 0} \frac{\lambda(D \cap B_r(\bar{x}))}{\lambda(B_r(\bar{x}))} = 1. \quad (13)$$

We define $\kappa^n := \|v^n - v\|_{L^1}$ for $n \in \mathbb{N}$. Because of (13), there exist a sequence $(r^n)_n$ and $n_0 \in \mathbb{N}$ such that $r^n \searrow 0$ and for all $n \geq n_0$:

$$\lambda(D \cap B_{r^n}(\bar{x})) \geq \kappa^n \geq \frac{1}{2} \lambda(B_{r^n}(\bar{x}))$$

and $B_{r^n}(\bar{x}) \subset \Omega$. We now restrict to $n \geq n_0$ and define w^n by

$$w^n(x) := \begin{cases} v(x) & \text{if } x \in B_{r^n}(\bar{x}), \\ w(x) & \text{else} \end{cases}$$

for a.e. $x \in \Omega$. This gives

$$\begin{aligned}
\|w^n - v^n\|_{L^1} &\leq \|v - v^n\|_{L^1} + \|w^n - v\|_{L^1} \\
&\leq \kappa^n + \|w - v\|_{L^1(\Omega \setminus B_{r^n}(\bar{x}))} \\
&\leq \kappa^n + \Delta - \underbrace{|w_1 - w_2|}_{\geq 1} \underbrace{\lambda(D \cap B_{r^n}(\bar{x}))}_{\geq \kappa^n} \underbrace{- \|w - v\|_{L^1(B_{r^n}(\bar{x}) \setminus D)}}_{\leq 0} \leq \Delta.
\end{aligned}$$

The construction of the w^n implies $w^n \xrightarrow{p^*} w$ in \mathcal{F} . In order to obtain the lim sup-inequality, we show $R_\alpha(w^n) \rightarrow R_\alpha(w)$. To this end, let $E_i := w^{-1}(\{w_i\})$, $E_i^n := (w^n)^{-1}(\{w_i\})$ for $i \in \{1, \dots, M\}$ and $n \geq n_0$. Then $E_1^n = E_1 \cup B_{r^n}(\bar{x})$ and $E_i^n = E_i \setminus B_{r^n}(\bar{x})$ for $i \geq 2$. For $\{E_1^n\}_n$, we deduce by means of (3)

$$\begin{aligned}
P_\alpha(E_1^n) &\leq P_\alpha(E_1) + P_\alpha(B_{r^n}(\bar{x})) \\
&\leq P_\alpha(E_1) + (r^n)^{d-\alpha} P_\alpha(B_1(\bar{x})) \rightarrow P_\alpha(E_1)
\end{aligned}$$

and, analogously, by means of (4)

$$P_\alpha(E_i^n) \leq P_\alpha(E_i) + (r^n)^{d-\alpha} P_\alpha(B_1(\bar{x})) \rightarrow P_\alpha(E_i)$$

for all $i \in \{1, \dots, M\}$. Summing the terms, we obtain $R_\alpha(w^n) \rightarrow R_\alpha(w)$. \square

5.3 Trust-Region Algorithm

We propose to solve (P_α) for locally optimal or stationary points with a variant of the trust-region algorithm that is proposed and analyzed in [18, 22]. It is stated as Algorithm 1 below and consists of two loops. The outer loop is indexed by n and in each iteration of the outer loop a new feasible iterate $w^n \in \mathcal{F}$ with $R_\alpha(w^n) < \infty$ is computed that improves acceptably over the previous iterate w^{n-1} . An acceptable improvement is achieved if the new iterate w^n satisfies

$$\text{ared}(w^{n-1}, w^n) \geq \sigma \text{ pred}(w^{n-1}, \Delta^{n,k}) \quad (14)$$

for a fixed $\sigma \in (0, 1)$ and the trust-region radius $\Delta^{n,k}$ that is determined by the inner loop (see below). In (14), the left hand side is defined by

$$\text{ared}(w^{n-1}, w) := F(w^{n-1}) + \eta R_\alpha(w^{n-1}) - F(w) - \eta R_\alpha(w)$$

for $w \in \mathcal{F}$ and is the *actual reduction* of the objective that is achieved by w . The right hand side is the *predicted reduction* that is achieved by the solution $\tilde{w}^{n,k}$ of the trust-region subproblem $\text{TR}_\alpha(w^{n-1}, \nabla F(w^{n-1}), \Delta^{n,k})$

$$\text{pred}(w^{n-1}, \Delta^{n,k}) := (\nabla F(w^{n-1}), w^{n-1} - \tilde{w}^{n,k}) + \eta R_\alpha(w^{n-1}) - \eta R_\alpha(\tilde{w}^{n,k})$$

is the predicted reduction by the (negative objective of the) trust-region subproblem for the current trust-region radius and thus its solution $\tilde{v}^{n,k}$.

To this end, the inner loop, indexed by k , starts from the reset trust-region radius $\Delta^{n,0} = \Delta^0$ and solves the trust-region subproblems $\text{TR}_\alpha(w^{n-1}, \nabla F(w^{n-1}), \Delta^{n,k})$ with linearization (model) point w^{n-1} . If the solution of the trust-region subproblem $\tilde{w}^{n,k}$ satisfies (14), the new iterate w^n is set as $\tilde{w}^{n,k}$ and the inner loop terminates. Else, the trust-region radius is halved and the next iteration of the inner loop begins. If the $\text{pred}(w^{n-1}, \Delta^{n,k}) = 0$, w^{n-1} is a minimizer of the trust-region subproblem for a positive trust-region radius and thus stationary for (P_α) by virtue of Proposition 5.4 if $\alpha \in (0.5, 1)$. In this case, Algorithm 1 terminates.

5.4 Asymptotics of the Trust-Region Algorithm

With the results that have been established in the previous sections, the asymptotics of Algorithm 1 can be analyzed by following the strategy from [22], which in turn is an extension of the analysis and ideas in [18]. We therefore only provide the information, where the proofs of the corresponding results in [22] require modification to match the situation of this work. We begin with the proof of the asymptotics of the inner loop and continue with the asymptotics of the outer loop.

Proposition 5.6 (Corollary 6.3 in [22], Corollary 4.20 in [18]) *Let $\alpha \in (0.5, 1)$. Let Assumption 3.1 hold. Let w^{n-1} produced by Algorithm 1 satisfy $\nabla F(w^{n-1}) \in C^2(\bar{\Omega})$. Then iteration n satisfies one of the following outcomes.*

1. *The inner loop terminates after finitely many iterations and*

Algorithm 1 Trust-region Algorithm leaning on SLIP from [18, 22]

Input: $\alpha \in (0, 1)$, F sufficiently regular, $\Delta^0 > 0$, $w^0 \in \mathcal{F}$ with $R_\alpha(w^n) < \infty$, $\sigma \in (0, 1)$.

```

1: for  $n = 0, \dots$  do
2:    $k \leftarrow 0$ 
3:    $\Delta^{n,0} \leftarrow \Delta^0$ 
4:   while not sufficient decrease according to (14) do
5:      $\tilde{w}^{n,k} \leftarrow$  minimizer of  $\mathbf{TR}_\alpha(w^{n-1}, \nabla F(w^{n-1}), \Delta^{n,k})$ .
6:      $\text{pred}(w^{n-1}, \Delta^{n,k}) \leftarrow (\nabla F(w^{n-1}), w^{n-1} - \tilde{w}^{n,k})_{L^2} + \eta R_\alpha(w^{n-1}) - \eta R_\alpha(\tilde{w}^{n,k})$ 
7:      $\text{ared}(v^{n-1}, \tilde{v}^{n,k}) \leftarrow F(v^{n-1}) + \eta R_\alpha(w^{n-1}) - F(\tilde{v}^{n,k}) - \eta R_\alpha(\tilde{w}^{n,k})$ 
8:     if  $\text{pred}(w^{n-1}, \Delta^{n,k}) \leq 0$  then
9:       Terminate. The predicted reduction for  $w^{n-1}$  is zero.
10:      else if not sufficient decrease according to (14) then
11:         $k \leftarrow k + 1$ 
12:         $\Delta^{n,k} \leftarrow \Delta^{n,k-1}/2$ .
13:      else
14:         $w^n \leftarrow \tilde{w}^{n,k}$ 
15:      end if
16:    end while
17: end for

```

(a) the sufficient decrease condition (14) is satisfied or
 (b) the predicted reduction is zero (and the iterate w^{n-1} is stationary for (\mathbf{P}_α)).
 2. The inner loop does not terminate and the iterate w^{n-1} is stationary.

Proof The proof follows as in Corollary 6.3 with the major steps of the proof being Lemma 6.1 and Lemma 6.2 in [22], where the violation of L -stationarity in Lemma 6.2 is replaced by a violation of (12) and the TV-term is replaced by R_α . The roles of Lemma 3.3, Lemma 3.5, and Proposition 5.5 in [22] are taken by Proposition 2.7, 2.9 and 5.2. The role of Lemma 3.8 in [22] is taken by Lemma 2.8. It leads to the term $(\varepsilon^k)^{2\alpha}$ instead of $(\varepsilon^k)^2$ in the proof of Lemma 6.2, which is still dominated by $\varepsilon^k \eta$ for $\varepsilon^k \searrow 0$ if $\alpha \in (0.5, 1)$ as assumed. \square

Theorem 5.7 (Theorem 6.4 in [22], Theorem 4.23 in [18]) *Let $\alpha \in (0.5, 1)$. Let Assumption 3.1 hold. Let the iterates $(w^n)_n$ be produced by Algorithm 1. Let $\nabla F(w^n) \in C^2(\bar{\Omega})$ for all $n \in \mathbb{N}$. Then all iterates are feasible for (\mathbf{P}_α) and the sequence of objective values $(J_\alpha(w^n))_n$ is monotonically decreasing. Moreover, one of the following mutually exclusive outcomes holds:*

1. The sequence $(w^n)_n$ is finite. The final element w^N of $(w^n)_n$ solves the trust-region subproblem $\mathbf{TR}_\alpha(w^N, \nabla F(w^N), \Delta)$ for some $\Delta > 0$ and is stationary for (\mathbf{P}_α) .
2. The sequence $(w^n)_n$ is finite and the inner loop does not terminate for the final element v^N , which is stationary for (\mathbf{P}_α) .
3. The sequence $(w^n)_n$ has a pseudo-weak-* accumulation point in \mathcal{F} . Every pseudo-weak-* accumulation point of $(w^n)_n$ is feasible, and strict. If w is a pseudo-weak-* accumulation point of $(w^n)_n$ that satisfies $\nabla F(w) \in C^2(\bar{\Omega})$, then it is stationary for (\mathbf{P}_α) .

If the trust-region radii are bounded away from zero for a subsequence $(w^{n_\ell})_\ell$, that is, if $0 < \underline{\Delta} := \liminf_{n_\ell \rightarrow \infty} \min_k \Delta^{n_\ell+1,k}$ and \bar{w} is a pseudo-weak-* accumulation point of $(w^{n_\ell})_\ell$ with $\nabla F(\bar{w}) \in C(\bar{\Omega})$, then \bar{w} solves $\mathbf{TR}_\alpha(\bar{w}, \nabla F(\bar{w}), \underline{\Delta}/2)$.

Proof As in the proofs of Theorem 6.4 in [22] and Theorem 4.23 in [18] it follows that Algorithm 1 produces a sequence of feasible iterates $(w^n)_n$ with corresponding monotonically decreasing sequence of objective function values $(J_\alpha(w^n))_n$. Again, as in the proofs of Theorem 6.4 in [22] and Theorem 4.23, it suffices to prove Outcome 3 in case that Outcomes 1 and 2 do not hold. In this argument Proposition 5.6 takes the role of Lemma 6.2 in [22] and Lemma 4.19 in [18]. As in [22], we consider four steps of the proof that Outcome 3 holds into four parts.

Outcome 3. (1) existence and feasibility of pseudo-weak-* accumulation points: This follows with the same arguments as are carried out for the existence of minimizers in Proposition 3.2.

Outcome 3. (2) pseudo-weak-* accumulation points are strict: This follows with the same arguments as are carried out in the corresponding paragraph in the proof of Theorem 6.4 in [22]. The only change is that the TV-term is replaced by the R_α .

Outcome 3. (3) strict accumulation points are optimal for \mathbf{TR}_α if the trust-region is bounded away from zero: This follows with the same arguments as are carried out in the corresponding paragraph in the proof of Theorem 6.4 in [22] when the role of Theorem 5.2 in [22] is taken by Algorithm 5.5 and the role of Proposition 5.5 in [22] is taken by Proposition 5.2.

Outcome 3. (4) strict accumulation points are stationary if the trust-region radius vanishes: This follows with the same arguments as are carried out in the corresponding paragraph in the proof of Theorem 6.4 in [22] with the following adaptions. The assumed violation of *L-stationarity* in Theorem 6.4 is replaced by a violation of (12) and the TV-term is replaced by R_α . The roles of Lemmas 3.3 and 3.5 in [22] are taken by Proposition 2.7 and 2.9. For the conditions (a), (b), and (c) on the choice of Δ^* in Theorem 6.4, we replace ε_1 by ε_1^α in (a) and the two occurrences of $\Delta^* \kappa^{-1}$ by $(\Delta^* \kappa^{-1})^{\frac{1}{\alpha}}$ in order to account for the exponent α in the estimate Lemma 2.8. Note that the non-negativity in (b) can be achieved because of the assumption $\alpha > \frac{1}{2}$, which implies that $(\Delta^*)^{\frac{1}{\alpha}}$ dominates $(\Delta^*)^2$ for $\Delta^* \searrow 0$. Moreover, below (a), (b), and (c) we choose t as $\left(\frac{\Delta^*}{2\kappa}\right)^{\frac{1}{\alpha}}$ instead of $\frac{\Delta^*}{2\kappa}$. Then the remaining steps can be carried out as in the proof of Theorem 6.4 in [22]. \square

6 Computational Experiment

We provide a computational example to give a qualitative impression of the behavior of the resulting discretization. We consider the binary control of an elliptic boundary value problem by means of a source term that enters the right hand side of the PDE. The main objective is a tracking-type functional so that the problems (\mathbf{P}_α) become

$$\begin{aligned}
 & \underset{u, w}{\text{minimize}} \quad \frac{1}{2} \|u - u_d\|_{L^2}^2 + \eta R_\alpha(w) \\
 & \text{subject to} \quad -\nu \Delta u = w \text{ in } \Omega, \quad u|_{\partial\Omega} = 0, \\
 & \quad w(x) \in W := \{0, 1\} \subset \mathbb{Z} \text{ a.e. in } \Omega.
 \end{aligned} \tag{P}_\alpha$$

We choose $\Omega = (0, 1) \times (0, 1)$, $u_d \in L^2(\Omega)$ such that it cannot be realized, implying the first term in the objective is bounded below away from zero. Moreover, we choose $\nu = \frac{1}{25}$ and $\eta = 5 \cdot 10^{-5}$. We discretize the problem and run the trust-region algorithm for the choices $\alpha = 0.5$, $\alpha = 0.9$, and for the limit with $R(w) = \omega_{d-1} P(E_1)$. In all cases, we initialize the algorithm with $w^0 = 0$.

In order to discretize the problem, we choose a piecewise constant ansatz for the control input w on a uniform grid of squares of size $n \times n$ with $n = 48$. For the limit case, an isotropic discretization of the total variation seminorm in integer optimization is computationally difficult and recent approaches [26] are not computationally mature enough so far. Therefore, we compute the total variation as the length of the interfaces along the boundaries of the grid cells, which implies an anisotropic behavior in the limit. We then obtain integer linear programs for the trust-region subproblems as derived in Appendix B of [22]. We expect that rectangular shapes are preferred for this anisotropic discretization. For the choices $\alpha = 0.5$, $\alpha = 0.9$, we tabulate the possible contributions of pairs of different cells to the double integral (2) and formulate the resulting trust-region subproblems as integer linear programs. Due to the double integral, this amounts to a number of variables in the order of n^4 , which is too much for standard integer programming solvers to handle easily. To alleviate this issue, we limit the contributions that are taken into account in the inner integral so that only cells whose center is within an ℓ_2 -distance of at most $7\frac{1}{n}$ to the center of a cell in the outer integral are taken into account in the inner integral. This is justified by the decay of the integral kernel of the Gagliardo seminorm with distance to the current point. For the limit case, we compute the limiting regularizer as the sum of the interface lengths multiplied by their jump heights as in [22].

We discretize the PDE using the open source library FEniCSx¹ [1, 27, 28], where we use a much finer mesh than 48×48 cells to solve the PDE and solve the subproblems using the integer programming solver Gurobi² [17]. We run the problems in single CPU mode on one node of TU Dortmund's Linux HPC cluster LiDO3 (node: $2 \times$ AMD EPYC 7542 32-Core CPUs and 1024 GB RAM). Even with our relatively coarse discretization and the approximation of the inner integral, many of the subproblems for $\alpha = 0.5$ and $\alpha = 0.9$ are very expensive, require a lot of time to solve and the computations take almost 3 weeks (with several subproblems being solved inexactly because they did not solve to global optimality within 30 h). The limit case is solved in a couple of hours.

This high computational burden shows that more work is necessary to solve these structured integer linear programs efficiently and approximate the involved integrals sensibly. For $\alpha = 0.5$, the trust-region algorithm accepts 31 steps until the trust-region radius contracts, that is it drops below the volume of one cell in the control

¹ <https://fenicsproject.org/>

² <https://www.gurobi.com/>

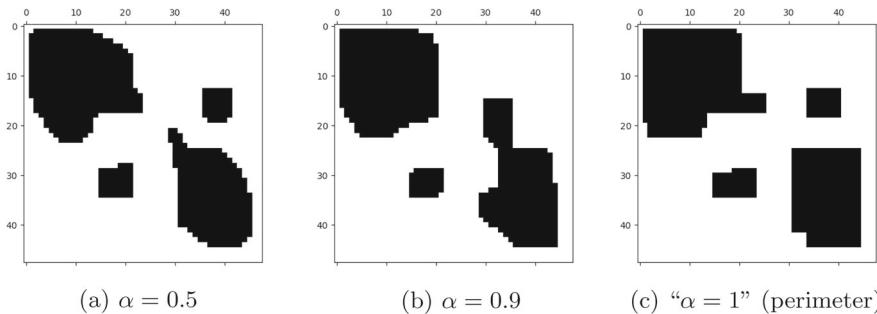


Fig. 1 Resulting shapes for the different values of α at the final iteration of the trust-region algorithm executed using discretized subproblems

grid. For $\alpha = 0.9$, the trust-region algorithm accepts 28 steps until the trust-region radius contracts. For the limiting case, the trust-region algorithm accepts 36 until the trust-region radius contracts.

The limiting case shows an anisotropic behavior. This is expected because the geometric restriction induced by the control function ansatz implies that all interface lengths (jump height multiplied by length) are computed along the boundaries of the discretization into squares, see also the comments at the end of Sect. 2 in [29]. This is alleviated for $\alpha = 0.5$ and $\alpha = 0.9$ and a less anisotropic behavior can be observed. The three resulting controls at the respective final iterations are shown in Fig. 1. We also note that when running the experiments again (where parallelization in Gurobi is turned on), slightly different results are returned because the integer problems are numerically challenging and we are at the limits of what Gurobi can handle so that several similar integer configurations are within the tolerances and the outcome is not deterministic. We have also observed that this leads our trust-region algorithm to follow slightly different paths and contract at different stationary points. This highlights even more that more work is necessary to solve the subproblems efficiently to global optimality (or a constant factor approximation).

Conclusion

Our theoretical analysis opens a sensible way of approaching the computationally difficult approximation of the anisotropic total variation in contexts with discreteness restrictions on the variables and discretizations with fixed geometries. Specifically, we have approximated the boundary integral by a double volume integral, where the approximation properties are carried out by means of the fractional nonlocal perimeter. Like other recent steps in this direction, our computational experiments show that we end up with a computationally very challenging problem, which needs to be understood and scaled to meaningful problem sizes in the future.

Acknowledgements The authors gratefully acknowledge computing time on the LiDO3 HPC cluster at TU Dortmund, partially funded in the Large-Scale Equipment Initiative by the Deutsche Forschungsgemeinschaft (DFG) as project 271512359. The authors like to thank an anonymous referee for valuable feedback on the manuscript.

Funding Harbir Antil H. Antil is partially supported by NSF grant DMS-2110263, Air Force Office of Scientific Research (AFOSR) under Award NO: FA9550-22-1-0248, and Office of Naval Research (ONR) under Award NO: N00014-24-1-2147. Paul Manns is partially supported by Deutsche Forschungsgemeinschaft (DFG) under Project No. 515118017.

Declarations

Competing Interests The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Alnaes, M.S., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., Ring, J., Rognes, M.E., Wells, G.N.: The FEniCS project version 1.5. *Arch. Numer. Softw.* **3** (2015)
2. Ambrosio, L., De Philippis, G., Martinazzi, L.: Gamma-convergence of nonlocal perimeter functionals. *Manuscr. Math.* **134**, 377–403 (2011)
3. Antil, H., Díaz, H., Jing, T., Schikorra, A.: Nonlocal bounded variations with applications. *SIAM J. Math. Anal.* **56**(2), 1903–1935 (2024)
4. Bartels, S.: Total variation minimization with finite elements: convergence and iterative solution. *SIAM J. Numer. Anal.* **50**(3), 1162–1180 (2012)
5. Borthagaray, J.P., Li, W., Nochetto, R.H.: Fractional elliptic problems on lipschitz domains: regularity and approximation. In: A^3N^2M : Approximation, Applications, and Analysis of Nonlocal, Nonlinear Models: Proceedings of the 50th John H. Barrett Memorial Lectures, pp. 27–99. Springer, New York (2023)
6. Braides, A., Conti, S., Garroni, A.: Density of polyhedral partitions. *Calc. Var. Partial Differ. Equ.* **56**, 1–10 (2017)
7. Caffarelli, L., Roquejoffre, J.-M., Savin, O.: Nonlocal minimal surfaces. *Commun. Pure Appl. Math.* **63**(9), 1111–1144 (2010)
8. Caillaud, C., Chambolle, A.: Error estimates for finite differences approximations of the total variation. *IMA J. Numer. Anal.* **43**(2), 692–736 (2023)
9. Chambolle, A., Morini, M., Ponsiglione, M.: Minimizing movements and level set approaches to nonlocal variational geometric flows. In: Geometric Partial Differential Equations proceedings, pp. 93–104. Springer, New York (2013)
10. Chambolle, A., Tan, P., Vaiter, S.: Accelerated alternating descent methods for Dykstra-like problems. *J. Math. Imaging Vis.* **59**, 481–497 (2017)
11. Cinti, E., Serra, J., Valdinoci, E.: Quantitative flatness results and bv -estimates for stable nonlocal minimal surfaces. *J. Differ. Geom.* **112**(3), 447–504 (2019)
12. Comi, G.E., Stefani, G.: A distributional approach to fractional Sobolev spaces and fractional variation: existence of blow-up. *J. Funct. Anal.* **277**(10), 3373–3435 (2019)
13. Cozzi, M., Figalli, A.: Regularity theory for local and nonlocal minimal surfaces: an overview. In: Nonlocal and Nonlinear Diffusions and Interactions: New Methods and Directions, Volume 2186 of Lecture Notes in Mathematics, pp. 117–158 (2017)

14. Di Nezza, E., Palatucci, G., Valdinoci, E.: Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.* **136**(5), 521–573 (2012)
15. Dipierro, S., Valdinoci, E.: Nonlocal minimal surfaces: interior regularity, quantitative estimates and boundary stickiness. In: *Recent Developments in Nonlocal Theory*, pp. 165–209 (2016)
16. Fremlin, D.H.: *Measure Theory*, vol. 2. Torres Fremlin, Colchester (2010)
17. Gurobi Optimization, LLC: *Gurobi Optimizer Reference Manual* (2023)
18. Leyffer, S., Manns, P.: Sequential linear integer programming for integer optimal control with total variation regularization. *ESAIM Control Optim. Calc. Variat.* **28**, 66 (2022)
19. Maggi, F.: *Sets of Finite Perimeter and Geometric Variational Problems: An Introduction to Geometric Measure Theory*, Number 135. Cambridge University Press, London (2012)
20. Maggi, F., Valdinoci, E.: Capillarity problems with nonlocal surface tension energies. *Commun. Partial Differ. Equ.* **42**(9), 1403–1446 (2017)
21. Manns, P., Nikolić, V.: Homotopy trust-region method for phase-field approximations in perimeter-regularized binary optimal control. *arXiv preprint arXiv:2310.12478* (2023)
22. Manns, P., Schiemann, A.: On integer optimal control with total variation regularization on multidimensional domains. *SIAM J. Control Optim.* **61**(6), 3415–3441 (2023)
23. Manns, P., Severitt, M.: On discrete subproblems in integer optimal control with total variation regularization in two dimensions. *arXiv preprint arXiv:2403.09213* (2024)
24. Manns, P., Hahn, M., Kirches, C., Leyffer, S., Sager, S.: On convergence of binary trust-region steepest descent. *J. Nonsmooth Anal. Optim.* **4**(3) (2023)
25. Marko, J., Wachsmuth, G.: Integer optimal control problems with total variation regularization: optimality conditions and fast solution of subproblems. *arXiv preprint arXiv:2207.05503* (2022)
26. Schiemann, A., Manns, P.: Discretization of total variation in optimization with integrality constraints. *arXiv preprint arXiv:2403.08346* (2024)
27. Scroggs, M.W., Baratta, I.A., Richardson, C.N., Wells, G.N.: Basix: a runtime finite element basis evaluation library. *J. Open Source Softw.* **7**(73), 3982 (2022)
28. Scroggs, M.W., Dokken, J.S., Richardson, C.N., Wells, G.N.: Construction of arbitrary order finite element degree-of-freedom maps on polygonal and polyhedral cell meshes. *ACM Trans. Math. Softw.* **48**(2), 18:1–18:23 (2022)
29. Severitt, M., Manns, P.: Efficient solution of discrete subproblems arising in integer optimal control with total variation regularization. *INFORMS J. Comput.* **35**(4), 869–885 (2023)
30. Šilhavý, M.: Fractional vector analysis based on invariance requirements (critique of coordinate approaches). *Contin. Mech. Thermodyn.* **32**(1), 207–228 (2020)
31. Visintin, A.: Generalized coarea formula and fractal sets. *Jpn. J. Ind. Appl. Math.* **8**, 175–201 (1991)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.