# A convergent overlapping domain decomposition method for total variation minimization

Fornasier, M., Langer, A. & Schönlieb, C-B.

Numerische
Mathematik

# A convergent overlapping domain decomposition method for total variation minimization

**Massimo Fornasier · Andreas Langer ·
Carola-Bibiane Schönlieb**

**Abstract**    In this paper we are concerned with the analysis of convergent sequential and parallel overlapping domain decomposition methods for the minimization of functionals formed by a discrepancy term with respect to the data and a total variation constraint. To our knowledge, this is the first successful attempt of addressing such a strategy for the nonlinear, nonadditive, and nonsmooth problem of total variation minimization. We provide several numerical experiments, showing the successful application of the algorithm for the restoration of 1D signals and 2D images in interpolation/inpainting problems, respectively, and in a compressed sensing problem, for recovering piecewise constant medical-type images from partial Fourier ensembles.

**Mathematics Subject Classification (2000)**    65K10 · 65N55 · 65N21 · 65Y05 · 90C25 · 52A41 · 49M30 · 49M27 · 68U10

M. Fornasier (✉) · A. Langer
Johann Radon Institute for Computational and Applied Mathematics (RICAM),
Austrian Academy of Sciences, Altenbergerstrasse 69, 4040 Linz, Austria
e-mail: massimo.fornasier@oeaw.ac.at

A. Langer
e-mail: andreas.langer@oeaw.ac.at

C.-B. Schönlieb
Institute for Numerical and Applied Mathematics, University of Göttingen,
Lotzestr. 16-18, 37083 Göttingen, Germany
e-mail: c.schoenlieb@math.uni-goettingen.de

## 1 Introduction

In concrete applications, e.g., for image processing, one might be interested in recovering at best a digital image provided only partial linear or nonlinear measurements, possibly corrupted by noise. Given the observation that natural and man-made images can be characterized by a relatively small number of edges and extensive relatively uniform parts, one may want to help the reconstruction by imposing that the interesting solution is the one which matches the given data and has also a few discontinuities localized on sets of lower dimension.

In the context of *compressed sensing* [6–8,21], it has been clarified that the minimization of $\ell_1$-norms occupies a fundamental role for the promotion of sparse solutions. This understanding furnishes an important interpretation of *total variation minimization*, i.e., the minimization of the $L^1$-norm of derivatives [33], as a regularization technique for image restoration. The problem can be modelled as follows; let $\Omega \subset \mathbb{R}^d$, for $d = 1, 2$ be a bounded open set with Lipschitz boundary, and $\mathcal{H} = L^2(\Omega)$. For $u \in L^1_{loc}(\Omega)$

$$V(u, \Omega) := \sup \left\{ \int_\Omega u \, \mathrm{div}\, \varphi \mathrm{d}x : \varphi \in \left[ C^1_c(\Omega) \right]^d, \ \|\varphi\|_\infty \le 1 \right\}$$

is the variation of $u$. Further, $u \in BV(\Omega)$, the space of bounded variation functions [1,24], if and only if $V(u, \Omega) < \infty$. In this case, we denote $|D(u)|(\Omega) = V(u, \Omega)$. If $u \in W^{1,1}(\Omega)$ (the Sobolev space of $L^1$-functions with $L^1$-distributional derivatives), then $|D(u)|(\Omega) = \int_\Omega |\nabla u| \mathrm{d}x$. We consider as in [12,37] the minimization in $BV(\Omega)$ of the functional

$$\mathcal{J}(u) := \|Tu - g\|^2_{L^2(\Omega)} + 2\alpha |D(u)|(\Omega), \tag{1}$$

where $T : L^2(\Omega) \to L^2(\Omega)$ is a bounded linear operator, $g \in L^2(\Omega)$ is a datum, and $\alpha > 0$ is a fixed *regularization parameter* [23]. Several numerical strategies to efficiently perform total variation minimization have been proposed in the literature. Without claiming of being exhaustive, we list a few of the relevant methods, ordered by their chronological appearance:

(1) the linearization approach of Dobson and Vogel [20] and of Chambolle and Lions [12] by iteratively re-weighted least squares, see also [18] for generalizations and refinements in the context of compressed sensing;
(2) the primal-dual approach of Chan et al. [13];
(3) variational approximation via locally quadratic functionals as in the work of Vese et al. [2,37];
(4) iterative thresholding algorithms based on projections onto convex sets as in the work of Chambolle [10] as well as in the work of Combettes and Wajs [15] and Daubechies et al. [19];
(5) iterative minimization of the Bregman distance as in the work of Osher et al. [32] (also notice the very recent Bregman split approach [27]);
(6) graph cuts [11,16] for the minimization of (1) with $T = I$ (the identity operator) and an anisotropic total variation;

(7) the approach proposed by Nesterov [31] and its modifications by Weiss et al. [39].

These approaches differ significantly, and they provide a convincing view of the interest this problem has been able to generate and of its applicative impact. However, because of their iterative-sequential formulation, none of the mentioned methods is able to address extremely large problems, such as 4D imaging (spatial plus temporal dimensions) from functional magnetic-resonance in nuclear medical imaging, astronomical imaging or global terrestrial seismic tomography, in real-time, or at least in an acceptable computational time. For such large scale simulations we need to address methods which allow us to reduce the problem to a finite sequence of sub-problems of a more manageable size, perhaps computable by one of the methods listed above. With this aim we introduced subspace correction and domain decomposition methods both for $\ell_1$-norm and total variation minimizations [25,26,34]. We address the interested reader to the broad literature included in [26] for an introduction to domain decomposition methods both for PDEs and convex minimization.

## 1.1 Difficulty of the problem

Due to the nonsmoothness and nonadditivity of the total variation with respect to a nonoverlapping domain decomposition (note that the total variation of a function on the whole domain equals the sum of the total variations on the subdomains plus the size of the jumps at the interfaces [26, formula (3.4)]), one encounters additional difficulties in showing convergence of such decomposition strategies to global minimizers. In particular, we stress very clearly that well-known approaches as in [9,14,35,36] are not directly applicable to this problem, because either they address additive problems or smooth convex minimizations, which is not the case of total variation minimization. We emphasize that the successful convergence of such alternating algorithms is far from being obvious for nonsmooth and nonadditive problems, as many counterexamples can be constructed, see for instance [38]. Moreover, for total variation minimization, the interesting solutions may be discontinuous, e.g., along curves in 2D. These discontinuities may cross the interfaces of the domain decomposition patches. Hence, the crucial difficulty is the correct numerical treatment of interfaces, with the preservation of crossing discontinuities and the correct matching where the solution is continuous instead, see [26, Section 7.1.1].

The work [26] was particularly addressed to *nonoverlapping* domain decompositions $\Omega_1 \cup \Omega_2 \subset \Omega \subset \bar{\Omega}_1 \cup \bar{\Omega}_2$ and $\Omega_1 \cap \Omega_2 = \emptyset$. Associated to the decomposition define $V_i = \{u \in L^2(\Omega) : \text{supp}(u) \subset \Omega_i\}$, for $i = 1, 2$ which constitute an orthogonal splitting $L^2(\Omega) = V_1 \oplus V_2$. With this splitting we wanted to minimize $\mathcal{J}$ by suitable instances of the following alternating algorithm: Pick an initial $V_1 \oplus V_2 \ni u_1^{(0)} + u_2^{(0)} := u^{(0)}$, for example $u^{(0)} = 0$, and iterate

$$\begin{cases} u_1^{(n+1)} \approx \arg\min_{v_1 \in V_1} \mathcal{J}\left(v_1 + u_2^{(n)}\right) \\ u_2^{(n+1)} \approx \arg\min_{v_2 \in V_2} \mathcal{J}\left(u_1^{(n+1)} + v_2\right) \\ u^{(n+1)} := u_1^{(n+1)} + u_2^{(n+1)}. \end{cases}$$

In [26] we proposed an implementation of this algorithm which is guaranteed to converge and to decrease the objective energy $\mathcal{J}$ monotonically. We could prove its convergence to minimizers of $\mathcal{J}$ only under technical conditions on the interfaces of the subdomains. However, in our numerical experiments, the algorithm seems to always converge robustly to the expected minimizer.

This discrepancy between theoretical analysis and numerical evidences motivated our investigation of *overlapping* domain decompositions. The hope was that the redundancy given by overlapping patches and the avoidance of boundary interfaces could allow for a technically easier theoretical analysis.

## 1.2 Our approach, results, and technical issues

In this paper we show how to adapt our previous algorithm [26] to the case of an *overlapping* domain decomposition. The setting of an overlapping domain decomposition eventually provides us with a framework in which we successfully prove in Theorem 5.7 and Theorem 6.2 its convergence to minimizers of $\mathcal{J}$, both in its sequential and parallel form, and these constitute the main results of this paper. Let us stress that to our knowledge this is the first method which addresses a domain decomposition strategy for total variation minimization with a formal, theoretical justification of convergence. We reiterate the relevance of our result, especially in light of the classical counterexamples mentioned above, which may hold for alternating algorithms on orthogonal decompositions of the space. Nevertheless, it is important to mention that there are also other very recent attempts of addressing domain decomposition methods for total variation minimization with successful numerical results [30,40], although lacking of a rigorous theoretical analysis.

Our analysis is performed for a discrete approximation of the continuous functional (1), for ease again denoted $\mathcal{J}$ in (3). Essentially we approximate functions $u$ by their sampling on a regular grid and their gradient $Du$ by finite differences $\nabla u$. It is well-known that such a discrete approximation   -converges to the continuous functional (see [4]). In particular, discrete minimizers of (3), interpolated by piecewise linear functions, converge in weak-*-topology of $BV$ to minimizers of the functional (1) in the continuous setting. Of course, when dealing with numerical solutions, only the discrete approach matters together with its approximation properties to the continuous problem. However, the need of working in the discrete setting is not only practical, it is also topological. In fact bounded sets in $BV$ are (only) weakly-*-compact, and this property is fundamental for showing that certain sequences have converging subsequences. Unfortunately, the weak-*-topology of $BV$ is "too weak" for our purpose of proving convergence of the domain decomposition algorithm; for instance, the trace on boundary sets is not a continuous operator with respect to this topology. This difficulty can be avoided, for example, by   -approximating the functional (1) by means of quadratic functionals (as in [2,12,37]) and working with the topology of $W^{1,2}(\Omega)$, the Sobolev space of $L^2$-functions with $L^2$-distributional first derivatives. However, this strategy changes the singular nature of the problem which makes it both interesting and difficult. Hence, the discrete approach has the virtues of being practical for numerical implementations, of correctly approximating the continuous setting, and of retaining

the main features which make the problem interesting. Note further that in the discrete setting where topological issues are not a concern anymore, also the dimension $d$ can be arbitrary, contrary to the continuous setting where the dimension $d$ has to be linked to boundedness properties of the operator $T$, see [37, property H2, pag. 134]. For ease of presentation, and in order to avoid unnecessary technicalities, we limit our analysis to splitting the problem into two subdomains $\Omega_1$ and $\Omega_2$. This is by no means a restriction. The generalization to multiple domains comes quite natural in our specific setting, see also [26, Remark 5.3]. When dealing with discrete subdomains $\Omega_i$, for technical reasons, we will require a certain splitting property for the total variation, i.e.,

$$|\nabla u|(\Omega) = |\nabla u|_{\Omega_1}(\Omega_1) + c_1\left(u|_{(\Omega_2 \setminus \Omega_1) \cup \Gamma_1}\right),$$
$$|\nabla u|(\Omega) = |\nabla u|_{\Omega_2}(\Omega_2) + c_2\left(u|_{(\Omega_1 \setminus \Omega_2) \cup \Gamma_2}\right), \tag{2}$$

where $c_1$ and $c_2$ are suitable functions which depend only on the restrictions $u|_{(\Omega_2 \setminus \Omega_1) \cup \Gamma_1}$ and $u|_{(\Omega_1 \setminus \Omega_2) \cup \Gamma_2}$, respectively, see (9) (symbols and notations are clarified once for all in the following section). Note that this formula is the discrete analogue of [26, formula (3.4)] in the continuous setting. The simplest examples of discrete domains with such a property are discrete $d$-dimensional rectangles (*d-orthotopes*). For instance, with our notations, it is easy to check that for $d = 1$ and for $\Omega$ being a discrete interval, one computes $c_1(u|_{(\Omega_2 \setminus \Omega_1) \cup \Gamma_1}) = |\nabla u|_{(\Omega_2 \setminus \Omega_1) \cup \Gamma_1}((\Omega_2 \setminus \Omega_1) \cup \Gamma_1)$, $c_2(u|_{(\Omega_1 \setminus \Omega_2) \cup \Gamma_2}) = |\nabla u|_{(\Omega_1 \setminus \Omega_2) \cup \Gamma_2}((\Omega_1 \setminus \Omega_2) \cup \Gamma_2)$; it is straightforward to generalize the computation to $d > 1$. Hence, for ease of presentation, we will assume to work with $d$-orthotope domains, also noting that such decompositions are already sufficient for any practical use in image processing, and stressing that the results can be generalized also to subdomains with different shapes as long as (2) is satisfied.

### 1.3 Organization of the work

The paper is organized as follows. In Sect. 2 we collect the relevant notations and symbols for the paper. Section 3 introduces the problem and the overlapping domain decomposition algorithm which we want to analyze. In Sect. 4 we address the approximate solution of the local problems defined on the subdomains $\Omega_i$, and we show how we interface them, by means of a suitable Lagrange multiplier. Sections 5 and 6 are concerned with the convergence of the sequential and parallel forms of the algorithm. In particular, in Sect. 5 we provide a characterization of minimizers by a discrete representation of the subdifferential of $\mathcal{J}$. This characterization is used in the convergence proofs in order to check the reached optimality. The final Sect. 7 provides a collection of applications and numerical examples.

## 2 Notations

Let us fix the main notations. Since we are interested in a discrete setting we define the *discrete d-orthotope* $\Omega = \{x_1^1 < \cdots < x_{N_1}^1\} \times \cdots \times \{x_1^d < \cdots < x_{N_d}^d\} \subset \mathbb{R}^d, d \in \mathbb{N}$ and the considered function spaces are $\mathcal{H} = \mathbb{R}^{N_1 \times N_2 \times \cdots \times N_d}$, where $N_i \in \mathbb{N}$ for $i = 1, \ldots, d$. For $u \in \mathcal{H}$ we write $u = u(x_\mathbf{i})_{\mathbf{i} \in \mathcal{I}}$ with

$$\mathcal{I} := \prod_{k=1}^{d} \{1, \ldots, N_k\}$$

and

$$u(x_{\mathbf{i}}) = u\left(x_{i_1}^1, \ldots, x_{i_d}^d\right)$$

where $i_k \in \{1, \ldots, N_k\}$ and $(x_{\mathbf{i}})_{\mathbf{i} \in \mathcal{I}} \in \mathbb{R}$. Then we endow $\mathcal{H}$ with the norm

$$\|u\|_{\mathcal{H}} = \|u\|_2 = \left(\sum_{\mathbf{i} \in \mathcal{I}} |u(x_{\mathbf{i}})|^2\right)^{1/2} = \left(\sum_{x \in \mathbb{R}} |u(x)|^2\right)^{1/2}.$$

We define the scalar product of $u, v \in \mathcal{H}$ as

$$\langle u, v \rangle_{\mathcal{H}} = \sum_{\mathbf{i} \in \mathcal{I}} u(x_{\mathbf{i}}) v(x_{\mathbf{i}}),$$

and the scalar product of $p, q \in \mathcal{H}^d$ as

$$\langle p, q \rangle_{\mathcal{H}^d} = \sum_{\mathbf{i} \in \mathcal{I}} \langle p(x_{\mathbf{i}}), q(x_{\mathbf{i}}) \rangle_{\mathbb{R}^d},$$

with $\langle y, z \rangle_{\mathbb{R}^d} = \sum_{j=1}^{d} y_j z_j$ for every $y = (y_1, \ldots, y_d) \in \mathbb{R}^d$ and $z = (z_1, \ldots, z_d) \in \mathbb{R}^d$. We will consider also other norms, in particular

$$\|u\|_p = \left(\sum_{\mathbf{i} \in \mathcal{I}} |u(x_{\mathbf{i}})|^p\right)^{1/p}, \quad 1 \le p < \infty,$$

and

$$\|u\|_\infty = \sup_{\mathbf{i} \in \mathcal{I}} |u(x_{\mathbf{i}})|.$$

We denote the discrete gradient $\nabla u$ by

$$(\nabla u)(x_{\mathbf{i}}) = \left((\nabla u)^1(x_{\mathbf{i}}), \ldots, (\nabla u)^d(x_{\mathbf{i}})\right),$$

with

$$(\nabla u)^j(x_{\mathbf{i}}) = \begin{cases} u\left(x_{i_1}^1, \ldots, x_{i_j+1}^j, \ldots, x_{i_d}^d\right) - u\left(x_{i_1}^1, \ldots, x_{i_j}^j, \ldots, x_{i_d}^d\right) & \text{if } i_j < N_j, \\ 0 & \text{if } i_j = N_j, \end{cases}$$

for all $j = 1, \ldots, d$ and for all $\mathbf{i} = (i_1, \ldots, i_d) \in \mathcal{I}$.

Let $\varphi : \mathbb{R} \to \mathbb{R}$, we define for $\omega \in \mathcal{H}^d$

$$\varphi(|\omega|)(\Omega) = \sum_{\mathbf{i} \in \mathcal{I}} \varphi(|\omega(x_{\mathbf{i}})|) = \sum_{x \in \mathbb{R}} \varphi(|\omega(x)|),$$

where $|y| = \sqrt{y_1^2 + \cdots + y_d^2}$. In particular we define the *total variation* of $u$ by setting $\varphi(s) = s$ and $\omega = \nabla u$, i.e.,

$$|\nabla u|(\Omega) := \sum_{\mathbf{i} \in \mathcal{I}} |\nabla u(x_{\mathbf{i}})| = \sum_{x \in} |\nabla u(x)|.$$

For an operator $T$ we denote by $T^*$ its adjoint. Further we introduce the *discrete divergence* div $: \mathcal{H}^d \to \mathcal{H}$ defined, in analogy with the continuous setting, by div $= -\nabla^*$ ($\nabla^*$ is the adjoint of the gradient $\nabla$). The discrete divergence operator is explicitly given by

$$(\text{div } p)(x_{\mathbf{i}}) = \begin{cases} p^1\left(x_{i_1}^1, \ldots, x_{i_d}^d\right) - p^1(x_{i_1-1}^1, \ldots, x_{i_d}^d) & \text{if } 1 < i_1 < N_1 \\ p^1\left(x_{i_1}^1, \ldots, x_{i_d}^d\right) & \text{if } i_1 = 1 \\ -p^1\left(x_{i_1-1}^1, \ldots, x_{i_d}^d\right) & \text{if } i_1 = N_1 \end{cases}$$

$$+ \cdots + \begin{cases} p^d\left(x_{i_1}^1, \ldots, x_{i_d}^d\right) - p^d\left(x_{i_1}^1, \ldots, x_{i_d-1}^d\right) & \text{if } 1 < i_d < N_d \\ p^d\left(x_{i_1}^1, \ldots, x_{i_d}^d\right) & \text{if } i_d = 1 \\ -p^d\left(x_{i_1}^1, \ldots, x_{i_d-1}^d\right) & \text{if } i_d = N_d, \end{cases}$$

for every $p = (p^1, \ldots, p^d) \in \mathcal{H}^d$ and for all $\mathbf{i} = (i_1, \ldots, i_d) \in \mathcal{I}$. (Note that if we considered discrete domains which are not discrete $d$-orthotopes, then the definitions of gradient and divergence operators should be adjusted accordingly.) With these notations, we define the closed convex set

$$K := \left\{ \text{div } p : p \in \mathcal{H}^d, |p(x)|_\infty \leq 1 \text{ for all } x \in \right\},$$

where $|p(x)|_\infty = \max\{|p^1(x)|, \ldots, |p^d(x)|\}$, and denote $P_K(u) = \arg\min_{v \in K} \|u - v\|_2$ the *orthogonal projection onto $K$*. We will often use the symbol 1 to indicate the constant vector with entry values 1, and $1_D$ to indicate the characteristic function of the domain $D \subset$ .

## 3 The overlapping domain decomposition algorithm

We are interested in the minimization of the functional

$$\mathcal{J}(u) := \|Tu - g\|_2^2 + 2\alpha |\nabla(u)|(\Omega), \tag{3}$$

where $T \in \mathcal{L}(\mathcal{H})$ is a linear operator, $g \in \mathcal{H}$ is a datum, and $\alpha > 0$ is a fixed constant. In order to guarantee the existence of minimizers for (3) we assume that:

(C) $\mathcal{J}$ is coercive in $\mathcal{H}$, i.e., there exists a constant $C > 0$ such that $\{\mathcal{J} \le C\} := \{u \in \mathcal{H} : \mathcal{J}(u) \le C\}$ is nonempty and bounded in $\mathcal{H}$.

It is well known that if $1 \notin \ker(T)$ then condition (C) is satisfied, see [37, Proposition 3.1].

Now, instead of minimizing (3) on the whole domain $\Omega$ we decompose $\Omega$ into two overlapping subdomains $\Omega_1$ and $\Omega_2$ such that $\Omega = \Omega_1 \cup \Omega_2$, $\Omega_1 \cap \Omega_2 = \emptyset$, and (2) is fulfilled. For consistency of the definitions of gradient and divergence, we assume that also the subdomains $\Omega_i$ are discrete $d$-orthotopes as well as $\Omega$, stressing that this is by no means a restriction, but only for ease of presentation. Due to this domain decomposition, $\mathcal{H}$ is split into two closed subspaces $V_j = \{u \in \mathcal{H} : \mathrm{supp}(u) \subset \Omega_j\}$, for $j = 1, 2$. Note that $\mathcal{H} = V_1 + V_2$ is not a direct sum of $V_1$ and $V_2$, but just a linear sum of subspaces. Thus any $u \in \mathcal{H}$ has a nonunique representation

$$u(x) = \begin{cases} u_1(x) & x \in \Omega_1 \setminus \Omega_2 \\ u_1(x) + u_2(x) & x \in \Omega_1 \cap \Omega_2, \\ u_2(x) & x \in \Omega_2 \setminus \Omega_1 \end{cases} \qquad u_i \in V_i, \quad i = 1, 2. \qquad (4)$$

We denote by $\Gamma_1$ the interface between $\Omega_1$ and $\Omega_2 \setminus \Omega_1$ and by $\Gamma_2$ the interface between $\Omega_2$ and $\Omega_1 \setminus \Omega_2$ (the interfaces are naturally defined in the discrete setting).

We introduce the trace operator of the restriction to a boundary $\Gamma_i$

$$\mathrm{Tr}\,|_{\Gamma_i} : V_i \to \mathbb{R}^{\Gamma_i}, \quad i = 1, 2,$$

with $\mathrm{Tr}\,|_{\Gamma_i} v_i = v_i \,|_{\Gamma_i}$, the restriction of $v_i$ on $\Gamma_i$. Note that $\mathbb{R}^{\Gamma_i}$ is as usual the set of maps from $\Gamma_i$ to $\mathbb{R}$. The trace operator is clearly a linear and continuous operator. We additionally fix a *bounded uniform partition of unity* (BUPU) $\{\chi_1, \chi_2\} \subset \mathcal{H}$ such that

(a) $\mathrm{Tr}\,|_{\Gamma_i} \chi_i = 0$ for $i = 1, 2$,
(b) $\chi_1 + \chi_2 = 1$,
(c) $\mathrm{supp}\,\chi_i \subset \Omega_i$ for $i = 1, 2$,
(d) $\max\{\|\chi_1\|_\infty, \|\chi_2\|_\infty\} = \kappa < \infty$.

We would like to solve

$$\arg\min_{u \in \mathcal{H}} \mathcal{J}(u),$$

by picking an initial $V_1 + V_2 \ni \tilde{u}_1^{(0)} + \tilde{u}_2^{(0)} := u^{(0)} \in \mathcal{H}$, e.g., $\tilde{u}_i^{(0)} = 0, i = 1, 2$, and iterating

$$\begin{cases} u_1^{(n+1)} \approx \arg\min_{\substack{v_1 \in V_1 \\ \mathrm{Tr}|_{\Gamma_1} v_1 = 0}} \mathcal{J}\left(v_1 + \tilde{u}_2^{(n)}\right) \\ u_2^{(n+1)} \approx \arg\min_{\substack{v_2 \in V_2 \\ \mathrm{Tr}|_{\Gamma_2} v_2 = 0}} \mathcal{J}\left(u_1^{(n+1)} + v_2\right) \\ u^{(n+1)} := u_1^{(n+1)} + u_2^{(n+1)} \\ \tilde{u}_1^{(n+1)} := \chi_1 \cdot u^{(n+1)} \\ \tilde{u}_2^{(n+1)} := \chi_2 \cdot u^{(n+1)}. \end{cases} \qquad (5)$$

Note that we are minimizing over functions $v_i \in V_i$ for $i = 1, 2$ which vanish on the interior boundaries, i.e., $\mathrm{Tr}\,|_{\Gamma_i}\, v_i = 0$. Moreover $u^{(n)}$ is the sum of the local minimizers $u_1^{(n)}$ and $u_2^{(n)}$, which are not uniquely determined on the overlapping part. Therefore we introduced a suitable correction by $\chi_1$ and $\chi_2$ in order to force the sub-minimizing sequences $(u_1^{(n)})_{n \in \mathbb{N}}$ and $(u_2^{(n)})_{n \in \mathbb{N}}$ to keep uniformly bounded. This issue will be explained in detail below, see Lemma 5.5. From the definition of $\chi_i$, $i = 1, 2$, it is clear that

$$u_1^{(n+1)} + u_2^{(n+1)} = u^{(n+1)} = (\chi_1 + \chi_2)u^{(n+1)} = \tilde{u}_1^{(n+1)} + \tilde{u}_2^{(n+1)}.$$

Note that in general $u_1^{(n)} = \tilde{u}_1^{(n)}$ and $u_2^{(n)} = \tilde{u}_2^{(n)}$. In (5) we use "$\approx$" (the approximation symbol) because in practice we never perform the exact minimization. In the following section we discuss how to realize the approximation to the individual subspace minimizations.

## 4 Local minimization by Lagrange multipliers

Let us consider, for example, the subspace minimization on $\Omega_1$

$$\arg\min_{\substack{v_1 \in V_1 \\ \mathrm{Tr}|_{\Gamma_1} v_1 = 0}} \mathcal{J}(v_1 + u_2) = \arg\min_{\substack{v_1 \in V_1 \\ \mathrm{Tr}|_{\Gamma_1} v_1 = 0}} \| Tv_1 - (g - Tu_2) \|_2^2$$
$$+ 2\alpha |\nabla(v_1 + u_2)|(\Omega). \qquad (6)$$

First of all, observe that $\{u \in \mathcal{H} : \mathrm{Tr}\,|_{\Gamma_1}\, u = \mathrm{Tr}\,|_{\Gamma_1}\, u_2, \mathcal{J}(u) \leq C\} \subset \{\mathcal{J} \leq C\}$, hence the former set is also bounded by assumption (C) and the minimization problem (6) has solutions.

It is useful to us to introduce an auxiliary functional $\mathcal{J}_1^s$, called the *surrogate functional* of $\mathcal{J}$ (cf. [26]): Assume $a, u_1 \in V_1, u_2 \in V_2$, and define

$$\mathcal{J}_1^s(u_1 + u_2, a) := \mathcal{J}(u_1 + u_2) + \| u_1 - a \|_2^2 - \| T(u_1 - a) \|_2^2. \qquad (7)$$

A straightforward computation shows that

$$\mathcal{J}_1^s(u_1 + u_2, a) = \| u_1 - a + T^*(g - Tu_2 - Ta) |_{\Omega_1} \|_2^2 + 2\alpha |\nabla(u_1 + u_2)|(\Omega)$$
$$+ \Phi(a, g, u_2),$$

where $\Phi$ is a function of $a, g, u_2$ only. Note that now the variable $u_1$ is not anymore affected by the action of $T$. Consequently, we want to realize an approximate solution to (6) by using the following algorithm: For $u_1^{(0)} = \tilde{u}_1^{(0)} \in V_1$,

$$u_1^{(\ell+1)} = \arg\min_{\substack{u_1 \in V_1 \\ \mathrm{Tr}|_{\Gamma_1} u_1 = 0}} \mathcal{J}_1^s \left( u_1 + u_2, u_1^{(\ell)} \right), \ell \geq 0. \qquad (8)$$

Additionally in (8) we can restrict the total variation on $\Gamma_1$ only, since we have

$$|\nabla(u_1 + u_2)|(\Omega) = |\nabla(u_1 + u_2)|_{\Gamma_1}(\Omega_1) + c_1 |u_2|_{(\Omega_2 \setminus \Gamma_1) \cup \Gamma_1}. \tag{9}$$

where we used (2) and the assumption that $u_1$ vanishes on the interior boundary $\Gamma_1$. Hence (8) is equivalent to

$$\arg \min_{\substack{u_1 \in V_1 \\ \mathrm{Tr}|_{\Gamma_1} u_1 = 0}} \mathcal{J}_1^s\left(u_1 + u_2, u_1^{(\ell)}\right) = \arg \min_{\substack{u_1 \in V_1 \\ \mathrm{Tr}|_{\Gamma_1} u_1 = 0}} \|u_1 - z_1\|_2^2$$
$$+ 2\alpha |\nabla(u_1 + u_2)|_{\Gamma_1}|(\Omega_1),$$

where $z_1 = u_1^{(\ell)} + (T^*(g - Tu_2 - Tu_1^{(\ell)}))|_{\Gamma_1}$.

Before proving the convergence of this algorithm, we first need to clarify how to practically compute $u_1^{(\ell+1)}$ for $u_1^{(\ell)}$ given. To this end we need to introduce further notions and to recall some useful results.

### 4.1 Generalized Lagrange multipliers for nonsmooth objective functions

Let us begin this section with the notion of a *subdifferential* in finite dimensions.

**Definition 4.1** For a finite dimensional and locally convex space $V$ and for a convex function $F : V \to \mathbb{R} \cup \{-\infty, +\infty\}$, we define the *subdifferential* of $F$ at $x \in V$, as the set valued function

$$\partial F(x) := \begin{cases} \emptyset & \text{if } F(x) = \infty \\ \{x^* \in V : \langle x^*, y - x \rangle + F(x) \le F(y) \quad \forall y \in V\} & \text{otherwise.} \end{cases}$$

It is obvious from this definition that $0 \in \partial F(x)$ if and only if $x$ is a minimizer of $F$. Since we deal with several spaces, namely, $\mathcal{H}$, $V_i$, it will turn out to be useful to sometimes distinguish in which space the subdifferential is defined by imposing a subscript $\partial_V F$ for the subdifferential considered on the space $V$.

We consider the following problem

$$\arg \min_{x \in V} \{F(x) : Gx = b\}, \tag{10}$$

where $G : V \to V$ is a linear operator on $V$. We have the following useful result.

**Theorem 4.2** *[28, Theorem 2.1.4, p. 305] Let $N = \{G^*\lambda : \lambda \in V\} = \mathrm{Range}(G^*)$. Then, $x_0 \in \{x \in V : Gx = b\}$ solves the constrained minimization problem (10) if and only if*

$$0 \in \partial F(x_0) + N.$$

## 4.2 Oblique thresholding

We want to exploit Theorem 4.2 in order to produce an algorithmic solution to each iteration step (8), which practically stems from the solution of a problem of this type

$$\arg\min_{\substack{u_1 \in V_1 \\ \mathrm{Tr}|_{\Omega_1} u_1 = 0}} \ \|u_1 - z_1\|_2^2 + 2\alpha \ |\nabla(u_1 + u_2 \mid_{\Omega_1})| (\Omega_1).$$

It is well-known how to solve this problem if $u_2 \equiv 0$ in $\Omega_1$ and the trace condition is not imposed. For the general case we propose the following solution strategy. In what follows all the involved quantities are restricted to $\Omega_1$, e.g., $u_2 = u_2 \mid_{\Omega_1}$.

**Theorem 4.3** (Oblique thresholding) *For $u_2 \in V_2$ and for $z_1 \in V_1$ the following statements are equivalent:*

(i)  $u_1^* = \arg\min_{\substack{u_1 \in V_1 \\ \mathrm{Tr}|_{\Omega_1} u_1 = 0}} \ \|u_1 - z_1\|_2^2 + 2\alpha |\nabla(u_1 + u_2)|(\Omega_1);$

(ii)  *there exists* $\eta \in \overline{\mathrm{Range}(\mathrm{Tr}\mid_{\Omega_1})^*} = \{\eta \in V_1 \ with \ \mathrm{supp}(\eta) = \overline{\Omega_1}\}$ *such that* $0 \in u_1^* - (z_1 - \eta) + \alpha \partial_{V_1} |\nabla(\cdot + u_2)|(\Omega_1)(u_1^*);$

(iii)  *there exists* $\eta \in V_1$ *with* $\mathrm{supp}(\eta) = \overline{\Omega_1}$ *such that* $u_1^* = (I - P_{\alpha K})(z_1 + u_2 - \eta) - u_2 \in V_1$ *and* $\mathrm{Tr}\mid_{\Omega_1} u_1^* = 0;$

(iv)  *there exists* $\eta \in V_1$ *with* $\mathrm{supp}(\eta) = \overline{\Omega_1}$ *such that* $\mathrm{Tr}\mid_{\Omega_1} \eta = \mathrm{Tr}\mid_{\Omega_1} z_1 + \mathrm{Tr}\mid_{\Omega_1} P_{\alpha K}(\eta - (z_1 + u_2))$ *or equivalently* $\eta = (\mathrm{Tr}\mid_{\Omega_1})^* \mathrm{Tr}\mid_{\Omega_1} (z_1 + P_{\alpha K}(\eta - (z_1 + u_2))).$

We call the solution operation provided by this theorem an *oblique thresholding*, in analogy to the terminology in [17], because it performs a thresholding of the derivatives, i.e., it sets to zero most of the derivatives of $u = u_1 + u_2 \approx z_1$ on $\Omega_1$, provided $u_2$, which is a fixed vector in $V_2$.

*Proof* Let us show the equivalence between (i) and (ii). The problem in (i) can be reformulated as

$$u_1^* = \arg\min_{u_1 \in V_1} \{F(u_1) := \ \|u_1 - z_1\|_2^2 + 2\alpha |\nabla(u_1 + u_2)|(\Omega_1), \mathrm{Tr}\mid_{\Omega_1} u_1 = 0\}. \tag{11}$$

Recall that $\mathrm{Tr}\mid_{\Omega_1} : V_1 \to \mathbb{R}^{\Omega_1}$ is a surjective map with closed range. This means that $(\mathrm{Tr}\mid_{\Omega_1})^*$ is injective and that $\mathrm{Range}(\mathrm{Tr}\mid_{\Omega_1})^* = \{\eta \in V_1 \ with \ \mathrm{supp}(\eta) = \overline{\Omega_1}\}$ is closed. Using Theorem 4.2 the optimality of $u_1^*$ is equivalent to the existence of $\eta \in \mathrm{Range}(\mathrm{Tr}\mid_{\Omega_1})^*$ such that

$$0 \in \partial_{V_1} F\left(u_1^*\right) + 2\eta. \tag{12}$$

Due to the continuity of $\|u_1 - z_1\|_2^2$ in $V_1$, we have, by [22, Proposition 5.6], that

$$\partial_{V_1} F\left(u_1^*\right) = 2\left(u_1^* - z_1\right) + 2\alpha \partial_{V_1} |\nabla(\cdot + u_2)|(\Omega_1)\left(u_1^*\right). \tag{13}$$

Thus, the optimality of $u_1^*$ is equivalent to

$$0 \in u_1^* - z_1 + \eta + \alpha \partial_{V_1} |\nabla(\cdot + u_2)|(\Omega_1)\ u_1^* \ . \tag{14}$$

This concludes the equivalence of (i) and (ii). Let us show now that (iii) is equivalent to (ii). The condition in (iii) can be rewritten as

$$\xi^* = (I - P_{\alpha K})(z_1 + u_2 - \eta), \quad \xi^* = u_1^* + u_2.$$

Since $|\nabla(\cdot)|(\Omega_1) \geq 0$ is 1-homogeneous and lower-semicontinuous, by [26, Example 4.2.2], the latter is equivalent to

$$0 \in \xi^* - (z_1 + u_2 - \eta) + \alpha \partial_{V_1} |\nabla(\cdot)|(\Omega_1)(\xi^*),$$

and equivalent to (ii). Note that in particular we have $\partial_{V_1}|\nabla(\cdot)|(\Omega_1)(\xi^*) = \partial_{V_1}|\nabla(\cdot + u_2)|(\Omega_1)(u_1^*)$, which is easily shown by a direct computation from the definition of subdifferential. We prove now the equivalence between (iii) and (iv). We have, for some $\eta \in V_1$ with $\mathrm{supp}(\eta) = \Omega_1$, and $\mathrm{Tr}\,|_{\Omega_1} u_1^* = 0$,

$$u_1^* = (I - P_{\alpha K})(z_1 + u_2 - \eta) - u_2 \in V_1,$$
$$= z_1 - \eta - P_{\alpha K}(z_1 + u_2 - \eta).$$

By applying $\mathrm{Tr}\,|_{\Omega_1}$ to both sides of the latter equality we get

$$0 = \mathrm{Tr}\,|_{\Omega_1} z_1 - \mathrm{Tr}\,|_{\Omega_1} \eta - \mathrm{Tr}\,|_{\Omega_1} P_{\alpha K}(z_1 + u_2 - \eta).$$

By observing that $-\mathrm{Tr}\,|_{\Omega_1} P_{\alpha K}(\xi) = \mathrm{Tr}\,|_{\Omega_1} P_{\alpha K}(-\xi)$, we obtain the fixed point equation

$$\mathrm{Tr}\,|_{\Omega_1} \eta = \mathrm{Tr}\,|_{\Omega_1} z_1 + \mathrm{Tr}\,|_{\Omega_1} P_{\alpha K}(\eta - (z_1 + u_2)). \tag{15}$$

Conversely, since all the considered quantities in

$$(I - P_{\alpha K})(z_1 + u_2 - \eta) - u_2$$

are in $V_1$, the whole expression is an element in $V_1$ and hence $u_1^*$ as defined in (iii) is an element in $V_1$ and $\mathrm{Tr}\,|_{\Omega_1} u_1^* = 0$. This shows the equivalence between (iii) and (iv) and therewith finishes the proof.

We wonder now whether any of the conditions in Theorem 4.3 is indeed practically satisfied. In particular, we want to show that $\eta \in V_1$ as in (iii) or (iv) of the previous theorem is provided as the limit of the following iterative algorithm:

$$\eta^{(0)} \in V_1, \quad \mathrm{supp}\,\eta^{(0)} = \Omega_1,$$
$$\eta^{(m+1)} = (\mathrm{Tr}\,|_{\Omega_1})^* \mathrm{Tr}\,|_{\Omega_1} \left( z_1 + P_{\alpha K}\left( \eta^{(m)} - (z_1 + u_2) \right) \right), \quad m \geq 0. \tag{16}$$

**Proposition 4.4** *The following statements are equivalent:*

(i) *there exists $\eta \in V_1$ such that $\eta = (\mathrm{Tr} \mid_1)^* \mathrm{Tr} \mid_1 (z_1 + P_{\alpha K}(\eta - (z_1 + u_2)))$ (which is in turn condition (iv) of Theorem 4.3);*

(ii) *the iteration (16) converges to some $\eta \in V_1$ which satisfies (15).*

For the proof of this proposition we need to recall some well-known notions and results.

**Definition 4.5** A nonexpansive map $\mathcal{T} : \mathcal{H} \rightarrow \mathcal{H}$ is strongly nonexpansive if for $(u_n - v_n)_{n \in \mathbb{N}}$ bounded and $\mathcal{T}(u_n) - \mathcal{T}(v_n)_2 - \| u_n - v_n_2 \rightarrow 0$ we have

$$u_n - v_n - (\mathcal{T}(u_n) - \mathcal{T}(v_n)) \rightarrow 0, \quad n \rightarrow \infty.$$

**Proposition 4.6** (Corollaries 1.3, 1.4, and 1.5 [5]) *Let $\mathcal{T} : \mathcal{H} \rightarrow \mathcal{H}$ be a strongly nonexpansive map. Then* fix $\mathcal{T} = \{u \in \mathcal{H} : \mathcal{T}(u) = u\} = \emptyset$ *if and only if $(\mathcal{T}^n u)_{n \in \mathbb{N}}$ converges to a fixed point $u_0 \in$ fix $\mathcal{T}$ for any choice of $u \in \mathcal{H}$.*

*Proof* (Proposition 4.4) Projections onto convex sets are strongly nonexpansive [3, Corollary 4.2.3]. Moreover, the composition of strongly nonexpansive maps is strongly nonexpansive [5, Lemma 2.1]. By an application of Proposition 4.6 we immediately have the result, since any map of the type $\mathcal{T}(\xi) = Q(\xi) + \xi_0$ is strongly nonexpansive whenever $Q$ is such a map (this is a simple observation from the definition of strongly nonexpansive maps). Indeed, we are looking for fixed points of $\eta = (\mathrm{Tr} \mid_1)^* \mathrm{Tr} \mid_1 (z_1 + P_{\alpha K}(\eta - (z_1 + u_2)))$ or, equivalently, of $\xi = \underbrace{(\mathrm{Tr} \mid_1)^* \mathrm{Tr} \mid_1 P_{\alpha K}(\xi)}_{:=Q} - \underbrace{((\mathrm{Tr} \mid_1)^* \mathrm{Tr} \mid_1 u_2)}_{:=\xi_0}$, where $\xi = (\mathrm{Tr} \mid_1)^* \mathrm{Tr} \mid_1 (\eta - (z_1 + u_2))$.

### 4.3 Convergence of the subspace minimization

From the results of the previous section it follows that the iteration (8) can be explicitly computed by

$$u_1^{(\ell+1)} = S_\alpha \left( u_1^{(\ell)} + T^* \left( g - Tu_2 - Tu_1^{(\ell)} \right) + u_2 - \eta^{(\ell)} \right) - u_2, \qquad (17)$$

where $S_\alpha := I - P_{\alpha K}$ and $\eta^{(\ell)} \in V_1$ is any solution of the fixed point equation

$$\eta = (\mathrm{Tr} \mid_1)^* \mathrm{Tr} \mid_1 \left( u_1^{(\ell)} + T^* \left( g - Tu_2 - Tu_1^{(\ell)} \right) - P_{\alpha K} \left( u_1^{(\ell)} \right. \right.$$
$$\left. \left. + T^* \left( g - Tu_2 - Tu_1^{(\ell)} \right) + u_2 - \eta \right) \right).$$

The computation of $\eta^{(\ell)}$ can be implemented by the algorithm (16).

**Proposition 4.7** *Assume $u_2 \in V_2$ and $T < 1$. Then the iteration (17) converges to a solution $u_1^* \in V_1$ of (6) for any initial choice of $u_1^{(0)} \in V_1$.*

The proof of this proposition is standard, see [15,17,26].

*Remark 4.8* Note that the condition $\|T\| < 1$ in Proposition 4.7, and assumed also later, is not automatically satisfied in general. However, in case the norm of the given operator $T$ exceeds 1, a proper rescaling of the problem re-establishes the desired setting. The rescaling can be easily obtained by multiplying the functional $\mathcal{J}$ by a positive constant $\gamma < \frac{1}{\|T\|^2}$, and minimizing the resulting functional

$$\mathcal{J}_\gamma(u) = \| \sqrt{\gamma} T u - \sqrt{\gamma} g \|_2^2 + 2\gamma\alpha|D(u)|(\Omega).$$

Note that the minimizers of the rescaled problem coincide with minimizers of $\mathcal{J}$.

Let us conclude this section by mentioning that all the results presented here hold symmetrically for the minimization on $V_2$, and that the notations should be just adjusted accordingly.

## 5 Convergence of the sequential alternating subspace minimization

In this section we want to prove the convergence of the algorithm (5) to minimizers of $\mathcal{J}$. In order to do that, we need a characterization of solutions of the minimization problem (3) as the one provided in [37, Proposition 4.1] for the continuous setting. We specify the arguments in [37, Proposition 4.1] for our discrete setting and we highlight the significant differences with respect to the continuous one.

### 5.1 Characterization of solutions

We make the following assumptions:

$(A_\varphi)$   $\varphi : \mathbb{R} \to \mathbb{R}$ is a convex function, nondecreasing in $\mathbb{R}^+$ such that
     (i)   $\varphi(0) = 0$.
     (ii)   There exist $c > 0$ and $b \geq 0$ such that $cz - b \leq \varphi(z) \leq cz + b$, for all $z \in \mathbb{R}^+$.

The particular example we have in mind is simply $\varphi(s) = s$, but we keep a more general notation for uniformity with respect to the continuous version in [37, Proposition 4.1]. In this section we are concerned with the following more general minimization problem

$$\arg\min_{u\in\mathcal{H}}\{\mathcal{J}_\varphi(u) := \|Tu - g\|_2^2 + 2\alpha\varphi(|\nabla u|)(\Omega)\}, \tag{18}$$

where $g \in \mathcal{H}$ is a datum, $\alpha > 0$ is a fixed constant (in particular for $\varphi(s) = s$).

To characterize the solution of the minimization problem (18) we use duality results from [22]. Therefore we recall the definition of the *conjugate (or Legendre transform)* of a function (see [22, Def. 4.1, p. 17]):

**Definition 5.1** Let $V$ and $V^*$ be two vector spaces placed in duality by a bilinear pairing denoted by $\langle \cdot, \cdot \rangle$ and $\phi : V \to \mathbb{R}$ a convex function. The *conjugate function (or Legendre transform)* $\phi^* : V^* \to \mathbb{R}$ is defined by

$$\phi^*(u^*) = \sup_{u \in V}\{\langle u, u^* \rangle - \phi(u)\}.$$

**Proposition 5.2** *Let $\zeta, u \in \mathcal{H}$. If the assumption $(A_\varphi)$ is fulfilled, then $\zeta \in \partial \mathcal{J}_\varphi(u)$ if and only if there exists $M = (M_0, \bar{M}) \in \mathcal{H} \times \mathcal{H}^d$, $\frac{|\bar{M}(x)|}{2\alpha} \leq c_1 \in [0, +\infty)$ for all $x \in$   such that*

$$\bar{M}(x), (\nabla u)(x) _{\mathbb{R}^d} + 2\alpha\varphi(|(\nabla u)(x)|) + 2\alpha\varphi_1^* \quad \frac{|\bar{M}(x)|}{2\alpha} \quad = 0 \ \text{ for all } x \in \Omega, \tag{19}$$

$$T^* M_0 - \text{div } \bar{M} + \zeta = 0, \tag{20}$$

$$-M_0 = 2(Tu - g), \tag{21}$$

*where $\varphi_1^*$ is the conjugate function of $\varphi_1$ defined by $\varphi_1(s) = \varphi(|s|)$, for $s \in \mathbb{R}$.*

*If additionally $\varphi$ is differentiable and $|(\nabla u)(x)| = 0$ for $x \in$   , then we can compute $\bar{M}$ as*

$$\bar{M}(x) = -2\alpha \frac{\varphi'(|(\nabla u)(x)|)}{|(\nabla u)(x)|}(\nabla u)(x). \tag{22}$$

The proof of this proposition specifies the one of [37, Proposition 4.1] to our discrete setting, it is technical, and it is deferred to the Appendix.

*Remark 5.3*  (i)  For $\varphi(s) = s$ the function $\varphi_1$ from Proposition 5.2 turns out to be $\varphi_1(s) = |s|$. Its conjugate function $\varphi_1^*$ is then given by

$$\varphi_1^*(s^*) = \sup_{s \in \mathbb{R}}\{\langle s^*, s \rangle - |s|\} = \begin{array}{ll} 0 & \text{for } |s^*| \leq 1 \\ \infty & \text{else} \end{array}.$$

Hence condition (19) specifies as follows

$$\bar{M}(x), (\nabla u)(x) _{\mathbb{R}^d} + 2\alpha|(\nabla u)(x)| = 0,$$

and, directly from the proof of Proposition 5.2 in the Appendix A, $|\bar{M}(x)| \leq 2\alpha$ for all $x \in$   .

(ii)  We want to highlight a few important differences with respect to the continuous case. Due to our definition of the gradient and its relationship with the divergence operator $-\text{div} = \nabla^*$ no boundary conditions are needed. Therefore condition (10) of [37, Proposition 4.1] has no discrete analogue in our setting. The continuous total variation of a function can be decomposed into an absolutely continuous part with respect to the Lebesgue measure and a singular part, whereas no singular part appears in the discrete setting. Therefore conditions (6) and (7) of [37, Proposition 4.1] have not a discrete analogue either.

(iii) An interesting consequence of Proposition 5.2 is that the map $S_\alpha = (I - P_{\alpha K})$ is bounded, i.e., $\|S_\alpha(z^k)\|_2 \to \infty$ if and only if $\|z^k\|_2 \to \infty$, for $k \to \infty$. In fact, since

$$S_\alpha(z) = \arg\min_{u \in \mathcal{H}} \|u - z\|_2^2 + 2\alpha|\nabla u|(\Omega),$$

from (20) and (21), we immediately obtain

$$S_\alpha(z) = z - \frac{1}{2}\operatorname{div}\bar{M},$$

and $\bar{M}$ is uniformly bounded.

### 5.2 Convergence properties

We return to the sequential algorithm (5). Let us explicitly express the algorithm as follows: Pick an initial $V_1 + V_2 \ni \tilde{u}_1^{(0)} + \tilde{u}_2^{(0)} := u^{(0)} \in \mathcal{H}$, for example, $\tilde{u}_i^{(0)} = 0, i = 1, 2$, and iterate

$$
\begin{cases}
\begin{cases}
u_1^{(n+1,0)} = \tilde{u}_1^{(n)} \\
u_1^{(n+1,\ell+1)} = \arg\min_{\substack{u_1 \in V_1 \\ \mathrm{Tr}|_{\Gamma_1} u_1 = 0}} \mathcal{J}_1^s\left(u_1 + \tilde{u}_2^{(n)}, u_1^{(n+1,\ell)}\right), \ell = 0, \ldots, L-1,
\end{cases} \\
\begin{cases}
u_2^{(n+1,0)} = \tilde{u}_2^{(n)} \\
u_2^{(n+1,m+1)} = \arg\min_{\substack{u_2 \in V_2 \\ \mathrm{Tr}|_{\Gamma_2} u_2 = 0}} \mathcal{J}_2^s\left(u_1^{(n+1,L)} + u_2, u_2^{(n+1,m)}\right), m = 0, \ldots, M-1,
\end{cases} \\
u^{(n+1)} := u_1^{(n+1,L)} + u_2^{(n+1,M)} \\
\tilde{u}_1^{(n+1)} := \chi_1 \cdot u^{(n+1)} \\
\tilde{u}_2^{(n+1)} := \chi_2 \cdot u^{(n+1)}.
\end{cases}
\tag{23}
$$

Algorithm (23) consists of two nested iterations. The inner iterations with indexes $\ell$ and $m$ constitute the iterative solution for the sequence of surrogate functionals on each subspace. Hence, these iterations approximatively compute minimizers for the functional $\mathcal{J}$ on the subspaces. The outer iteration with index $n$ stems from our domain decomposition approach and iteratively computes the minimizer of $\mathcal{J}$ on the whole space. Note that we do prescribe a finite number $L$ and $M$ of inner iterations for each subspace, respectively, and that $u^{(n+1)} = \tilde{u}_1^{(n+1)} + \tilde{u}_2^{(n+1)}$, with $u_i^{(n+1)} = \tilde{u}_i^{(n+1)}, i = 1, 2$, in general. In this section we want to prove the convergence of the algorithm in (23) for any choice of $L$ and $M$.

Observe that, for $a \in V_i$ and $\|T\| < 1$,

$$\|u_i - a\|_2^2 - \|Tu_i - Ta\|_2^2 \geq C\|u_i - a\|_2^2, \tag{24}$$

for $C = (1 - \|T\|^2) > 0$. Hence

$$\mathcal{J}(u) = \mathcal{J}_i^s(u, u_i) \leq \mathcal{J}_i^s(u, a), \tag{25}$$

and

$$\mathcal{J}_i^s(u, a) - \mathcal{J}_i^s(u, u_i) \geq C \ \|u_i - a\|_2^2. \tag{26}$$

**Proposition 5.4** (Convergence properties) *Let us assume that* $\|T\| < 1$. *The algorithm in* (23) *produces a sequence* $(u^{(n)})_{n \in \mathbb{N}}$ *in* $\mathcal{H}$ *with the following properties:*

(i)  $\mathcal{J}(u^{(n)}) > \mathcal{J}(u^{(n+1)})$ *for all* $n \in \mathbb{N}$ *(unless* $u^{(n)} = u^{(n+1)}$*);*
(ii)  $\lim_{n \to \infty} \|u^{(n+1)} - u^{(n)}\|_2 = 0$;
(iii)  *the sequence* $(u^{(n)})_{n \in \mathbb{N}}$ *has subsequences which converge in* $\mathcal{H}$.

*Proof* Let us first observe that

$$\mathcal{J}(u^{(n)}) = \mathcal{J}_1^s\left(\tilde{u}_1^{(n)} + \tilde{u}_2^{(n)}, \tilde{u}_1^{(n)}\right) = \mathcal{J}_1^s\left(\tilde{u}_1^{(n)} + \tilde{u}_2^{(n)}, u_1^{(n+1,0)}\right).$$

By definition of $u_1^{(n+1,1)}$ and the minimal properties of $u_1^{(n+1,1)}$ in (23) we have

$$\mathcal{J}_1^s\left(\tilde{u}_1^{(n)} + \tilde{u}_2^{(n)}, u_1^{(n+1,0)}\right) \geq \mathcal{J}_1^s\left(u_1^{(n+1,1)} + \tilde{u}_2^{(n)}, u_1^{(n+1,0)}\right).$$

From (25) we have

$$\mathcal{J}_1^s\left(u_1^{(n+1,1)} + \tilde{u}_2^{(n)}, u_1^{(n+1,0)}\right) \geq \mathcal{J}_1^s\left(u_1^{(n+1,1)} + \tilde{u}_2^{(n)}, u_1^{(n+1,1)}\right)$$
$$= \mathcal{J}\left(u_1^{(n+1,1)} + \tilde{u}_2^{(n)}\right).$$

Concatenating these inequalities we obtain

$$\mathcal{J}(u^{(n)}) \geq \mathcal{J}\left(u_1^{(n+1,1)} + \tilde{u}_2^{(n)}\right).$$

In particular, from (26) we have

$$\mathcal{J}(u^{(n)}) - \mathcal{J}\left(u_1^{(n+1,1)} + \tilde{u}_2^{(n)}\right) \geq C \ \left\|u_1^{(n+1,1)} - u_1^{(n+1,0)}\right\|_2^2.$$

After $L$ steps we conclude the estimate

$$\mathcal{J}(u^{(n)}) \geq \mathcal{J}\left(u_1^{(n+1,L)} + \tilde{u}_2^{(n)}\right),$$

and

$$\mathcal{J}(u^{(n)}) - \mathcal{J}\left(u_1^{(n+1,L)} + \tilde{u}_2^{(n)}\right) \geq C \sum_{\ell=0}^{L-1} \left\|u_1^{(n+1,\ell+1)} - u_1^{(n+1,\ell)}\right\|_2^2.$$

By definition of $u_2^{(n+1,1)}$ and its minimal properties we have

$$\mathcal{J}\left(u_1^{(n+1,L)} + \tilde{u}_2^{(n)}\right) \geq \mathcal{J}_2^s\left(u_1^{(n+1,L)} + u_2^{(n+1,1)}, u_2^{(n+1,0)}\right).$$

By similar arguments as above we finally find the decreasing estimate

$$\mathcal{J}(u^{(n)}) \geq \mathcal{J}\left(u_1^{(n+1,L)} + u_2^{(n+1,M)}\right) = \mathcal{J}(u^{(n+1)}) = \mathcal{J}\left(\tilde{u}_1^{(n+1)} + \tilde{u}_2^{(n+1)}\right), \quad (27)$$

and

$$\mathcal{J}(u^{(n)}) - \mathcal{J}(u^{(n+1)})$$
$$\geq C\left(\sum_{\ell=0}^{L-1} \left\|u_1^{(n+1,\ell+1)} - u_1^{(n+1,\ell)}\right\|_2^2 + \sum_{m=0}^{M-1} \left\|u_2^{(n+1,m+1)} - u_2^{(n+1,m)}\right\|_2^2\right), \quad (28)$$

which verifies $(i)$.

From (27) we have $\mathcal{J}(u^{(0)}) \geq \mathcal{J}(u^{(n)})$. By the coerciveness condition (C) $(u^{(n)})_{n \in \mathbb{N}}$ is uniformly bounded in $\mathcal{H}$, hence there exists a convergent subsequence $(u^{(n_k)})_{k \in \mathbb{N}}$ and hence $(iii)$ holds. For simplicity, we rename such a subsequence by $(u^{(n)})_{n \in \mathbb{N}}$. Moreover, since the sequence $(\mathcal{J}(u^{(n)}))_{n \in \mathbb{N}}$ is monotonically decreasing and bounded from below by 0, it is also convergent. From (28) and the latter convergence we deduce

$$\sum_{\ell=0}^{L-1} \left\|u_1^{(n+1,\ell+1)} - u_1^{(n+1,\ell)}\right\|_2^2 + \sum_{m=0}^{M-1} \left\|u_2^{(n+1,m+1)} - u_2^{(n+1,m)}\right\|_2^2 \to 0, \quad n \to \infty.$$
$$(29)$$

In particular, by the standard inequality $(a^2 + b^2) \geq \frac{1}{2}(a + b)^2$ for $a, b > 0$ and the triangle inequality, we have also

$$\left\|u^{(n)} - u^{(n+1)}\right\|_2 \to 0, \quad n \to \infty. \quad (30)$$

This gives $(ii)$ and completes the proof.

The use of the partition of unity $\{\chi_1, \chi_2\}$ allows us not only to guarantee the boundedness of $(u^{(n)})_{n \in \mathbb{N}}$, but also of the sequences $(\tilde{u}_1^{(n)})_{n \in \mathbb{N}}$ and $(\tilde{u}_2^{(n)})_{n \in \mathbb{N}}$.

**Lemma 5.5** *The sequences $(\tilde{u}_1^{(n)})_{n \in \mathbb{N}}$ and $(\tilde{u}_2^{(n)})_{n \in \mathbb{N}}$ produced by the algorithm* (23) *are bounded, i.e., there exists a constant $\tilde{C} > 0$ such that $\|\tilde{u}_i^{(n)}\|_2 \leq \tilde{C}$, for all $n \in \mathbb{N}$, and for $i = 1, 2$.*

*Proof* From the boundedness of $(u^{(n)})_{n \in \mathbb{N}}$ we have

$$\left\|\tilde{u}_i^{(n)}\right\|_2 = \left\|\chi_i u^{(n)}\right\|_2 \leq \kappa \left\|u^{(n)}\right\|_2 \leq \tilde{C} \quad \text{for } i = 1, 2.$$

From Remark 5.3 (iii) we can also show the following auxiliary lemma.

**Lemma 5.6** *The sequences $(\eta_1^{(n,L)})_{n\in\mathbb{N}}$ and $(\eta_2^{(n,M)})_{n\in\mathbb{N}}$ are bounded.*

*Proof* From previous considerations we know that

$$u_1^{(n,L)} = S_\alpha \, z_1^{(n,L-1)} + \tilde{u}_2^{(n-1)} - \eta_1^{(n,L)} \quad - \tilde{u}_2^{(n-1)},$$

$$u_2^{(n,M)} = S_\alpha \, z_2^{(n,M-1)} + u_1^{(n,L)} - \eta_2^{(n,M)} \quad - u_1^{(n,L)}.$$

Assume $(\eta_1^{(n,L)})_{n\in\mathbb{N}}$ were unbounded, then by Remark 5.3 (iii), also $S_\alpha(z_1^{(n,L-1)} + \tilde{u}_2^{(n-1)} - \eta_1^{(n,L)})$ would be unbounded. Since $(\tilde{u}_2^{(n)})_{n\in\mathbb{N}}$ and $(u_1^{(n,L)})_{n\in\mathbb{N}}$ are bounded by Lemma 5.5 and formula (29), we have a contradiction. Thus $(\eta_1^{(n,L)})_{n\in\mathbb{N}}$ has to be bounded. With the same argument we can show that $(\eta_2^{(n,M)})_{n\in\mathbb{N}}$ is bounded.

We can eventually show the convergence of the algorithm to minimizers of $\mathcal{J}$.

**Theorem 5.7** (Convergence to minimizers) *Assume $T < 1$. Then accumulation points of the sequence $(u^{(n)})_{n\in\mathbb{N}}$ produced by algorithm (23) are minimizers of $\mathcal{J}$. If $\mathcal{J}$ has a unique minimizer then the sequence $(u^{(n)})_{n\in\mathbb{N}}$ converges to it.*

*Proof* Let us denote $u^{(\infty)}$ the limit of a subsequence. For simplicity, we rename such a subsequence by $(u^{(n)})_{n\in\mathbb{N}}$. From Lemma 5.5 we know that $(\tilde{u}_1^{(n)})_{n\in\mathbb{N}}$, $(\tilde{u}_2^{(n)})_{n\in\mathbb{N}}$ and consequently $(u_1^{(n,L)})_{n\in\mathbb{N}}$, $(u_2^{(n,M)})_{n\in\mathbb{N}}$ are bounded. So the limit $u^{(\infty)}$ can be written as

$$u^{(\infty)} = u_1^{(\infty)} + u_2^{(\infty)} = \tilde{u}_1^{(\infty)} + \tilde{u}_2^{(\infty)} \tag{31}$$

where $u_1^{(\infty)}$ is the limit of $(u_1^{(n,L)})_{n\in\mathbb{N}}$, $u_2^{(\infty)}$ is the limit of $(u_2^{(n,M)})_{n\in\mathbb{N}}$, and $\tilde{u}_i^{(\infty)}$ is the limit of $(\tilde{u}_i^{(n)})_{n\in\mathbb{N}}$ for $i = 1, 2$. Now we show that $\tilde{u}_2^{(\infty)} = u_2^{(\infty)}$. By using the triangle inequality, from (29) it directly follows that

$$\left\| u_2^{(n+1,M)} - \tilde{u}_2^{(n)} \right\|_2 \to 0, \quad n \to \infty. \tag{32}$$

Moreover, since $\chi_2 \in V_2$ is a fixed vector which is independent of $n$, we obtain from Proposition 5.4 (ii) that

$$\left\| \chi_2(u^{(n)} - u^{(n+1)}) \right\|_2 \to 0, \quad n \to \infty,$$

and hence

$$\left\| \tilde{u}_2^{(n)} - \tilde{u}_2^{(n+1)} \right\|_2 \to 0, \quad n \to \infty. \tag{33}$$

Putting (32) and (33) together and noting that

$$\left\| u_2^{(n+1,M)} - \tilde{u}_2^{(n)} \right\|_2 + \left\| \tilde{u}_2^{(n)} - \tilde{u}_2^{(n+1)} \right\|_2 \geq \left\| u_2^{(n+1,M)} - \tilde{u}_2^{(n+1)} \right\|_2,$$

we have

$$\left\| u_2^{(n+1,M)} - \tilde{u}_2^{(n+1)} \right\|_2 \to 0, \quad n \to \infty, \tag{34}$$

which means that the sequences $(u_2^{(n,M)})_{n\in\mathbb{N}}$ and $(\tilde{u}_2^{(n)})_{n\in\mathbb{N}}$ have the same limit, i.e., $\tilde{u}_2^{(\infty)} = u_2^{(\infty)}$, which we denote by $u_2^{(\infty)}$. Then from (34) and (31) it directly follows that $\tilde{u}_1^{(\infty)} = u_1^{(\infty)}$.

As in the proof of the oblique thresholding theorem we set

$$F_1\left( u_1^{(n+1,L)} \right) := \left\| u_1^{(n+1,L)} - z_1^{(n+1,L)} \right\|_2^2 + 2\alpha \left\| \nabla u_1^{(n+1,L)} + \tilde{u}_2^{(n)} \right\|_1 (\Omega_1),$$

where

$$z_1^{(n+1,L)} := \left. \left( u_1^{(n+1,L-1)} + T^*\left( g - T\tilde{u}_2^{(n)} - Tu_1^{(n+1,L-1)} \right) \right) \right|_1.$$

The optimality condition for $u_1^{(n+1,L)}$ is

$$0 \in \partial_{V_1} F_1\left( u_1^{(n+1,L)} \right) + 2\eta_1^{(n+1,L)},$$

where

$$\eta_1^{(n+1,L)} = (\mathrm{Tr}|_1)^* \mathrm{Tr}|_1\left( z_1^{(n+1,L)} + P_{\alpha K}\left( \eta_1^{(n+1,L)} - z_1^{(n+1,L)} - \tilde{u}_2^{(n)} \right) \right).$$

In order to use the characterization of elements in the subdifferential of $|\nabla u|(\Omega)$, i.e., Proposition 5.2, we have to rewrite the minimization problem for $F_1$. More precisely, we define

$$\hat{F}_1\left( \xi_1^{(n+1,L)} \right) := \left\| \left( \xi_1^{(n+1,L)} - \tilde{u}_2^{(n)} \right)\Big|_1 - z_1^{(n+1,L)} \right\|_2^2 + 2\alpha \left\| \nabla \xi_1^{(n+1,L)} \right\| (\Omega_1),$$

for $\xi_1^{(n+1,L)} \in V_1$ with $\mathrm{Tr}|_1 \xi_1^{(n+1,L)} = \tilde{u}_2^{(n)}$. Then the optimality condition for $\xi_1^{(n+1,L)}$ is

$$0 \in \partial \hat{F}_1\left( \xi_1^{(n+1,L)} \right) + 2\eta_1^{(n+1,L)}. \tag{35}$$

Note that indeed $\xi_1^{(n+1,L)}$ is optimal if and only if $u_1^{(n+1,L)} = \left( \xi_1^{(n+1,L)} - \tilde{u}_2^{(n)} \right)\Big|_1$ is optimal.

Analogously we define

$$\hat{F}_2\left(\xi_2^{(n+1,M)}\right) := \left\|\xi_2^{(n+1,M)} - u_1^{(n+1,L)} - z_2^{(n+1,M)}\right\|_2^2 + 2\alpha \left\|\nabla \xi_2^{(n+1,M)}\right\|_2 \quad (\Omega_2),$$

for $\xi_2^{(n+1,M)} \in V_2$ with $\mathrm{Tr}\,|_{\Gamma_2}\,\xi_2^{(n+1,M)} = u_1^{(n+1,L)}$, and the optimality condition for $\xi_2^{(n+1,M)}$ is

$$0 \in \partial \hat{F}_2\left(\xi_2^{(n+1,M)}\right) + 2\eta_2^{(n+1,M)}, \tag{36}$$

where

$$\eta_2^{(n+1,M)} = (\mathrm{Tr}\,|_{\Gamma_2})^* \mathrm{Tr}\,|_{\Gamma_2}\, z_2^{(n+1,M)} + P_{\alpha K}\left(\eta_2^{(n+1,M)} - z_2^{(n+1,M)} - u_1^{(n+1,L)}\right).$$

Let us recall that now we are considering functionals as in Proposition 5.2 with $\varphi(s) = s$, $T = I$, and $\Gamma = \Gamma_i$, $i = 1, 2$. From Proposition 5.2 and Remark 5.3 we get that $\xi_1^{(n+1,L)}$, and consequently $u_1^{(n+1,L)}$ is optimal, i.e., $-2\eta_1^{(n+1,L)} \in \partial \hat{F}_1(\xi_1^{(n+1,L)})$, if and only if there exists an $M_1^{(n+1)} = (M_{0,1}^{(n+1)}, \bar{M}_1^{(n+1)}) \in V_1 \times V_1^d$ with $|\bar{M}_1^{(n+1)}(x)| \le 2\alpha$ for all $x \in \Omega_1$ such that

$$\left\langle \bar{M}_1^{(n+1)}(x), \nabla\left(u_1^{(n+1,L)} + \tilde{u}_2^{(n)}\right)(x) \right\rangle_{\mathbb{R}^d}$$
$$+ 2\alpha \left|\nabla\left(u_1^{(n+1,L)} + \tilde{u}_2^{(n)}\right)(x)\right| = 0, \tag{37}$$

$$-2\left(u_1^{(n+1,L)}(x) - z_1^{(n+1,L)}(x)\right) - \mathrm{div}\,\bar{M}_1^{(n+1)}(x) - 2\eta_1^{(n+1,L)}(x) = 0, \tag{38}$$

for all $x \in \Omega_1$. Analogously we get that $\xi_2^{(n+1,M)}$, and consequently $u_2^{(n+1,M)}$ is optimal, i.e., $-2\eta_2^{(n+1,M)} \in \partial \hat{F}_2(\xi_2^{(n+1,M)})$, if and only if there exists an $M_2^{(n+1)} = (M_{0,2}^{(n+1)}, \bar{M}_2^{(n+1)}) \in V_2 \times V_2^d$ with $|\bar{M}_2^{(n+1)}(x)| \le 2\alpha$ for all $x \in \Omega_2$ such that

$$\left\langle \bar{M}_2^{(n+1)}(x), \nabla\left(u_1^{(n+1,L)} + u_2^{(n+1,M)}\right)(x) \right\rangle_{\mathbb{R}^d}$$
$$+ 2\alpha \left|\nabla\left(u_1^{(n+1,L)} + \tilde{u}_2^{(n+1,M)}\right)(x)\right| = 0, \tag{39}$$

$$-2\left(u_2^{(n+1,M)}(x) - z_2^{(n+1,M)}(x)\right) - \mathrm{div}\,\bar{M}_2^{(n+1)}(x) - 2\eta_2^{(n+1,M)}(x) = 0, \tag{40}$$

for all $x \in \Omega_2$. Since $(\bar{M}_1^{(n)}(x))_{n \in \mathbb{N}}$ is bounded for all $x \in \Omega_1$ and $(\bar{M}_2^{(n)}(x))_{n \in \mathbb{N}}$ is bounded for all $x \in \Omega_2$, there exist convergent subsequences $(\bar{M}_1^{(n_k)}(x))_{k \in \mathbb{N}}$ and $(\bar{M}_2^{(n_k)}(x))_{k \in \mathbb{N}}$. Let us denote $\bar{M}_1^{(\infty)}(x)$ and $\bar{M}_2^{(\infty)}(x)$ the respective limits of the sequences. For simplicity we rename such sequences by $(\bar{M}_1^{(n)}(x))_{n \in \mathbb{N}}$ and $(\bar{M}_2^{(n)}(x))_{n \in \mathbb{N}}$.

Note that, by Lemma 5.6 (or simply from (38) and (40)) the sequences $(\eta_1^{(n,L)})_{n\in\mathbb{N}}$ and $(\eta_2^{(n,M)})_{n\in\mathbb{N}}$ are also bounded. Hence there exist convergent subsequences which we denote, for simplicity, again by $(\eta_1^{(n,L)})_{n\in\mathbb{N}}$ and $(\eta_2^{(n,M)})_{n\in\mathbb{N}}$ with limits $\eta_i^{(\infty)}$, $i = 1, 2$. By taking the limits for $n \to \infty$ in (37)–(40) we obtain

$$\left\langle \bar{M}_1^{(\infty)}(x), \; \nabla\left(u_1^{(\infty)} + u_2^{(\infty)}\right)(x)\right\rangle_{\mathbb{R}^d} + 2\alpha\left|\nabla\left(u_1^{(\infty)} + u_2^{(\infty)}\right)(x)\right| = 0,$$

for all $x \in \Omega_1$,

$$-2\left(u_1^{(\infty)}(x) - z_1^{(\infty)}(x)\right) - \operatorname{div} \bar{M}_1^{(\infty)}(x) - 2\eta_1^{(\infty)}(x) = 0, \quad \text{for all } x \in \Omega_1,$$

$$\left\langle \bar{M}_2^{(\infty)}(x), \; \nabla\left(u_1^{(\infty)} + u_2^{(\infty)}\right)(x)\right\rangle_{\mathbb{R}^d} + 2\alpha\left|\nabla\left(u_1^{(\infty)} + u_2^{(\infty)}\right)(x)\right| = 0,$$

for all $x \in \Omega_2$,

$$-2\left(u_2^{(\infty)}(x) - z_2^{(\infty)}(x)\right) - \operatorname{div} \bar{M}_2^{(\infty)}(x) - 2\eta_2^{(\infty)}(x) = 0, \quad \text{for all } x \in \Omega_2,$$

Since $\operatorname{supp} \eta_1^{(\infty)} = \Omega_1$ and $\operatorname{supp} \eta_2^{(\infty)} = \Omega_2$ we have

$$\left\langle \bar{M}_1^{(\infty)}(x), \nabla(u^{(\infty)})(x)\right\rangle_{\mathbb{R}^d} + 2\alpha|\nabla(u^{(\infty)})(x)| = 0, \quad \text{for all } x \in \Omega_1,$$

$$-2T^*((Tu^{(\infty)})(x) - g(x)) - \operatorname{div} \bar{M}_1^{(\infty)}(x) = 0, \quad \text{for all } x \in \Omega_1 \setminus \Gamma_1, \tag{41}$$

$$\left\langle \bar{M}_2^{(\infty)}(x), \nabla(u^{(\infty)})(x)\right\rangle_{\mathbb{R}^d} + 2\alpha|\nabla(u^{(\infty)})(x)| = 0, \quad \text{for all } x \in \Omega_2,$$

$$-2T^*((Tu^{(\infty)})(x) - g(x)) - \operatorname{div} \bar{M}_2^{(\infty)}(x) = 0, \quad \text{for all } x \in \Omega_2 \setminus \Gamma_2. \tag{42}$$

Observe now that from Proposition 5.2 we also have that $0 \in \mathcal{J}(u^{(\infty)})$ if and only if there exists $M^{(\infty)} = (M_0^{(\infty)}, \bar{M}^{(\infty)})$ with $|\bar{M}^{(\infty)}(x)| \leq 2\alpha$ for all $x \in \Omega$ such that

$$\left\langle \bar{M}^{(\infty)}(x), \nabla(u^{(\infty)})(x)\right\rangle_{\mathbb{R}^d} + 2\alpha|\nabla(u^{(\infty)})(x)| = 0, \quad \text{for all } x \in \Omega,$$

$$-2T^*((Tu^{(\infty)})(x) - g(x)) - \operatorname{div} \bar{M}^{(\infty)}(x) = 0, \quad \text{for all } x \in \Omega. \tag{43}$$

Note that $\bar{M}_j^{(\infty)}(x)$, $j = 1, 2$, for $x \in \Omega_1 \cap \Omega_2$ satisfies both (41) and (42). Hence let us choose

$$M^{(\infty)}(x) = \begin{cases} M_1^{(\infty)}(x) & \text{if } x \in \Omega_1 \setminus \Gamma_1, \\ M_2^{(\infty)}(x) & \text{if } x \in (\Omega_2 \setminus \Gamma_1) \cup \Gamma_1. \end{cases}$$

With this choice of $M^{(\infty)}$ equations (41)–(43) are valid and hence $u^{(\infty)}$ is optimal in $\Omega$.

*Remark 5.8* (i) If $\nabla u^{(\infty)}(x) = 0$ for $x \in \Omega_j$, $j = 1, 2$, then $\bar{M}_j^{(\infty)}$ is given as in equation (22) by

$$\bar{M}_j^{(\infty)}(x) = -2\alpha \frac{\nabla u^{(\infty)}|_{\Omega_j}(x)}{\left|\nabla u^{(\infty)}|_{\Omega_j}(x)\right|}.$$

(ii) The boundedness of the sequences $(\tilde{u}_1^{(n)})_{n\in\mathbb{N}}$ and $(\tilde{u}_2^{(n)})_{n\in\mathbb{N}}$ has been technically used for showing the existence of an optimal decomposition $u^{(\infty)} = u_1^{(\infty)} + u_2^{(\infty)}$ in the proof of Theorem 5.7. Their boundedness is guaranteed as in Lemma 5.5 by the use of the partition of the unity $\{\chi_1, \chi_2\}$. Let us emphasize that there is no way of obtaining the boundedness of the local sequences $(u_1^{(n,L)})_{n\in\mathbb{N}}$ and $(u_2^{(n,M)})_{n\in\mathbb{N}}$ otherwise. In Fig. 6 we show that the local sequences can become unbounded in case we do not modify them by means of the partition of the unity.

(iii) Note that for deriving the optimality condition (43) for $u^{(\infty)}$ we combined the respective conditions (41) and (42) for $u_1^{(\infty)}$ and $u_2^{(\infty)}$. In doing that, we strongly took advantage of the overlapping property of the subdomains, hence avoiding a fine analysis of $\eta_1^{(\infty)}$ and $\eta_2^{(\infty)}$ on the interfaces $\Gamma_1$ and $\Gamma_2$. This is the major advantage of this analysis with respect to the one provided in [26] for nonoverlapping domain decompositions.

## 6 A parallel algorithm and its convergence

The parallel version of the previous algorithm (23) reads as follows: Pick an initial $V_1 + V_2 \ni \tilde{u}_1^{(0)} + \tilde{u}_2^{(0)} := u^{(0)} \in \mathcal{H}$, for example $\tilde{u}_i^{(0)} = 0, i = 1, 2$, and iterate

$$
\begin{cases}
\begin{cases}
u_1^{(n+1,0)} = \tilde{u}_1^{(n)} \\
u_1^{(n+1,\ell+1)} = \arg\min_{\substack{u_1 \in V_1 \\ \mathrm{Tr}|_{\Gamma_1} u_1 = 0}} \mathcal{J}_1^s\left(u_1 + \tilde{u}_2^{(n)}, u_1^{(n+1,\ell)}\right), \ell = 0, \ldots, L-1,
\end{cases} \\
\begin{cases}
u_2^{(n+1,0)} = \tilde{u}_2^{(n)} \\
u_2^{(n+1,m+1)} = \arg\min_{\substack{u_2 \in V_2 \\ \mathrm{Tr}|_{\Gamma_2} u_2 = 0}} \mathcal{J}_2^s\left(\tilde{u}_1^{(n)} + u_2, u_2^{(n+1,m)}\right), m = 0, \ldots, M-1,
\end{cases} \\
u^{(n+1)} := \frac{u_1^{(n+1,L)} + u_2^{(n+1,M)} + u^{(n)}}{2} \\
\tilde{u}_1^{(n+1)} := \chi_1 \cdot u^{(n+1)} \\
\tilde{u}_2^{(n+1)} := \chi_2 \cdot u^{(n+1)}
\end{cases}
\tag{44}
$$

We are going to propose similar convergence results as for the sequential algorithm.

**Proposition 6.1** (Convergence properties) *Let us assume that $\gamma_T < 1$. The parallel algorithm (44) produces a sequence $(u^{(n)})_{n\in\mathbb{N}}$ in $\mathcal{H}$ with the following properties:*

(i) *$\mathcal{J}(u^{(n)}) > \mathcal{J}(u^{(n+1)})$ for all $n \in \mathbb{N}$ (unless $u^{(n)} = u^{(n+1)}$);*

(ii) *$\lim_{n\to\infty} \|u^{(n+1)} - u^{(n)}\|_2 = 0$;*

(iii) *the sequence $(u^{(n)})_{n\in\mathbb{N}}$ has subsequences which converge in $\mathcal{H}$.*

*Proof* With the same argument as in the proof of Theorem 5.4, we obtain

$$
\mathcal{J}(u^{(n)}) - \mathcal{J}\left(u_1^{(n+1,L)} + \tilde{u}_2^{(n)}\right) \geq C \sum_{\ell=0}^{L-1} \left\| u_1^{(n+1,\ell+1)} - u_1^{(n+1,\ell)} \right\|_2^2
$$

and

$$\mathcal{J}(u^{(n)}) - \mathcal{J}\left(\tilde{u}_1^{(n)} + u_2^{(n+1,M)}\right) \geq C \sum_{m=0}^{M-1} \left\| u_2^{(n+1,m+1)} - u_2^{(n+1,m)} \right\|_2^2.$$

Hence, by summing and halving

$$\mathcal{J}(u^{(n)}) - \frac{1}{2} \left[ \mathcal{J}\left(u_1^{(n+1,L)} + \tilde{u}_2^{(n)}\right) + \mathcal{J}\left(\tilde{u}_1^{(n)} + u_2^{(n+1,M)}\right) \right]$$

$$\geq \frac{C}{2} \left[ \sum_{\ell=0}^{L-1} \left\| u_1^{(n+1,\ell+1)} - u_1^{(n+1,\ell)} \right\|_2^2 + \sum_{m=0}^{M-1} \left\| u_2^{(n+1,m+1)} - u_2^{(n+1,m)} \right\|_2^2 \right].$$

We recall that $\mathcal{J}(u^{(n)}) = \| Tu^{(n)} - g \|_2^2 + 2\alpha|\nabla u^{(n)}|(\Omega)$ (and $T$ is linear). Then, by the standard inequality $(a^2 + b^2) \geq \frac{1}{2}(a + b)^2$ for $a, b > 0$, we have

$$\left\| Tu^{(n+1)} - g \right\|_2^2 = \left\| T\left( \frac{u_1^{(n+1,L)} + u_2^{(n+1,M)} + u^{(n)}}{2} \right) - g \right\|_2^2$$

$$\leq \frac{1}{2} \left\| T\left(u_1^{(n+1,L)} + \tilde{u}_2^{(n)}\right) - g \right\|_2^2 + \frac{1}{2} \left\| T\left(\tilde{u}_1^{(n)} + u_2^{(n+1,M)}\right) - g \right\|_2^2.$$

Moreover we have

$$|\nabla(u^{(n+1)})|(\Omega) \leq \frac{1}{2} \left[ \left| \nabla\left(u_1^{(n+1,L)} + \tilde{u}_2^{(n)}\right) \right|(\Omega) + \left| \nabla\left(\tilde{u}_1^{(n)} + u_2^{(n+1,M)}\right) \right|(\Omega) \right].$$

By the last two inequalities we immediately show that

$$\mathcal{J}(u^{(n+1)}) \leq \frac{1}{2} \left[ \mathcal{J}\left(u_1^{(n+1,L)} + \tilde{u}_2^{(n)}\right) + \mathcal{J}\left(\tilde{u}_1^{(n)} + u_2^{(n+1,M)}\right) \right],$$

hence

$$\mathcal{J}(u^{(n)}) - \mathcal{J}(u^{(n+1)})$$

$$\geq \frac{C}{2} \left[ \sum_{\ell=0}^{L-1} \left\| u_1^{(n+1,\ell+1)} - u_1^{(n+1,\ell)} \right\|_2^2 + \sum_{m=0}^{M-1} \left\| u_2^{(n+1,m+1)} - u_2^{(n+1,m)} \right\|_2^2 \right]$$

$$\geq 0. \tag{45}$$

Since the sequence $(\mathcal{J}(u^{(n)}))_{n\in\mathbb{N}}$ is monotonically decreasing and bounded from below by 0, it is also convergent. From (45) and the latter convergence we deduce

$$\sum_{\ell=0}^{L-1} \left\| u_1^{(n+1,\ell+1)} - u_1^{(n+1,\ell)} \right\|_2^2 + \sum_{m=0}^{M-1} \left\| u_2^{(n+1,m+1)} - u_2^{(n+1,m)} \right\|_2^2 \to 0,$$

$$n \to \infty. \tag{46}$$

In particular, by again using $(a^2 + b^2) \geq \frac{1}{2}(a + b)^2$ for $a, b > 0$ and the triangle inequality, we also have

$$u^{(n)} - u^{(n+1)} \ _2 \to 0, \quad n \to \infty. \tag{47}$$

The rest of the proof follows analogous arguments as in that of Proposition 5.4.

Analogous results as the one stated in Lemma 5.5 and Lemma 5.6 also hold in the parallel case. With these preliminary results the following theorem follows:

**Theorem 6.2** (Convergence to minimizers) *Assume* $T < 1$. *Then accumulation points of the sequence* $(u^{(n)})_{n \in \mathbb{N}}$ *produced by algorithm* (44) *are minimizers of* $\mathcal{J}$. *If* $\mathcal{J}$ *has a unique minimizer then the sequence* $(u^{(n)})_{n \in \mathbb{N}}$ *converges to it.*

*Proof* Note that $u^{(n+1)}$ is the average of the current iteration and the previous, i.e.,

$$u^{(n+1)} = \frac{u_1^{(n+1,L)} + u_2^{(n+1,M)} + u^{(n)}}{2}.$$

Observe that the sequences $(u_1^{(n+1,L)})_{n \in \mathbb{N}}$, $(u_2^{(n+1,M)})_{n \in \mathbb{N}}$ and $(u^{(n)})_{n \in \mathbb{N}}$ are bounded. Hence there exist convergent subsequences. By taking the limit for $n \to \infty$ we obtain

$$u^{(\infty)} = \frac{u_1^{(\infty)} + u_2^{(\infty)} + u^{(\infty)}}{2}$$

which is equivalent to

$$u^{(\infty)} = u_1^{(\infty)} + u_2^{(\infty)}.$$

With this observation the rest of the proof follows analogous arguments as in that of Theorem 5.7.

## 7 Applications and numerics for the sequential implementation

In this section we shall present the application of the sequential and parallel algorithms (23) and (44) for the minimization of $\mathcal{J}$ in one and two dimensions. In particular, we show how to implement the dual method of Chambolle [10] in order to compute the orthogonal projection $P_{\alpha K}(g)$ in the oblique thresholding, and we give a detailed explanation of the domain decompositions used in the numerics. Furthermore, we present numerical examples for image inpainting, i.e., the recovery of missing parts of images by minimal total variation interpolation, and compressed sensing [6–8,21], the nonadaptive compressed acquisition of images for a classical toy problem inspired by magnetic resonance imaging (MRI) [7,29]. The numerical examples of this section and respective Matlab codes can be found at [41].

### 7.1 Computation of $P_{\alpha K}(g)$

To solve the subiterations in (5) we compute the minimizer by means of oblique thresholding. More precisely, let us denote $u_2 = \tilde{u}_2^{(n)}$, $u_1 = u_1^{(n+1,\ell+1)}$, and $z_1 = u_1^{(n+1,\ell)} + T^*(g - Tu_2 - Tu_1^{(n+1,\ell)})$. We shall compute the minimizer $u_1$ of the first subminimization problem by

$$u_1 = (I - P_{\alpha K})(z_1 + u_2 - \eta) - u_2 \in V_1,$$

for some $\eta \in V_1$ with supp $\eta = \Omega_1$ which fulfills

$$\mathrm{Tr}\,|_{\Omega_1}(\eta) = \mathrm{Tr}\,|_{\Omega_1}(z_1 + P_{\alpha K}(\eta - z_1 - u_2)).$$

Hence the element $\eta \in V_1$ is a limit of the corresponding fixed point iteration

$$\eta^{(0)} \in V_1, \quad \mathrm{supp}\,\eta^{(0)} = \Omega_1,$$
$$\eta^{(m+1)} = (\mathrm{Tr}\,|_{\Omega_1})^*\,\mathrm{Tr}\,|_{\Omega_1}\left(z_1 + P_{\alpha K}\left(\eta^{(m)} - z_1 - u_2\right)\right), \quad m \geq 0. \quad (48)$$

Here $K$ is defined as in Section 2, i.e.,

$$K = \left\{ \mathrm{div}\, p : p \in \mathcal{H}^d, |p(x)|_\infty \leq 1, \quad \forall x \in \Omega_1 \right\}.$$

To compute the projection onto $\alpha K$ in the oblique thresholding we use an algorithm proposed by Chambolle in [10]. His algorithm is based on considerations of the convex conjugate of the total variation and on exploiting the corresponding optimality condition. It amounts to compute $P_{\alpha K}(g)$ approximately by $\alpha \,\mathrm{div}\, p^{(n)}$, where $p^{(n)}$ is the $n$th iterate of the following semi-implicit gradient descent algorithm:

Choose $\tau > 0$, let $p^{(0)} = 0$ and, for any $n \geq 0$, iterate

$$p^{(n+1)}(x) = \frac{p^{(n)}(x) + \tau(\nabla(\mathrm{div}\, p^{(n)} - g/\alpha))(x)}{1 + \tau\,|(\nabla(\mathrm{div}\, p^{(n)} - g/\alpha))(x)|}.$$

For $\tau > 0$ sufficiently small, i.e., $\tau < 1/8$, the iteration $\alpha \,\mathrm{div}\, p^{(n)}$ was shown to converge to $P_{\alpha K}(g)$ as $n \to \infty$ (compare [10, Theorem 3.1]). Let us stress that we propose this algorithm here, just for the ease of its presentation and its implementation; its choice for the approximation of projections is of course by no means a restriction and one may want to implement other recent, and perhaps faster strategies, e.g., [11,16,27,32,39].

### 7.2 Domain decompositions

In one dimension the domain $\Omega$ is a set of $N$ equidistant points on an interval $[a, b]$, i.e., $\Omega = \{a = x_1, \ldots, x_N = b\}$ and is split into two overlapping intervals $\Omega_1$ and $\Omega_2$. Let $|\Omega_1 \cap \Omega_2| =: G$ be the size of the overlap of $\Omega_1$ and $\Omega_2$. Then we set

**Fig. 1** Overlapping domain decomposition in 1D
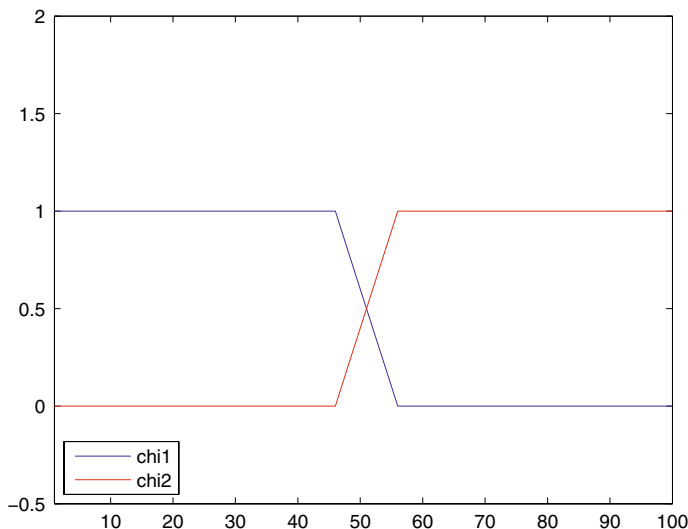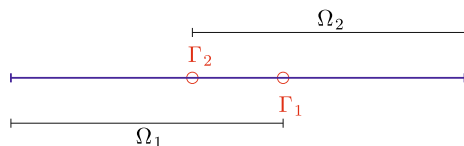




**Fig. 2** Auxiliary functions $\chi_1$ and $\chi_2$ for an overlapping domain decomposition with two subdomains

$\Omega_1 = \{a = x_1, \ldots, x_{n_1}\}$ and $\Omega_2 = \{x_{n_1-G+1}, \ldots, x_N = b\}$ with $|\Omega_1| := n_1 = \frac{N+G}{2}$. The interfaces $\Gamma_1$ and $\Gamma_2$ are located at the indexes $i = n_1$ and $n_1 - G + 1$, respectively (cf. Fig. 1). The auxiliary functions $\chi_1$ and $\chi_2$ can be chosen in the following way (cf. Fig. 2):

$$\chi_1(x_i) = \begin{cases} 1 & x_i \in \Omega_1 \setminus \Omega_2, \\ 1 - \frac{1}{G}(i - (n_1 - G)) & x_i \in \Omega_1 \cap \Omega_2, \end{cases}$$

$$\chi_2(x_i) = \begin{cases} 1 & x_i \in \Omega_2 \setminus \Omega_1, \\ \frac{1}{G}(i - (n_1 - G + 1)) & x_i \in \Omega_1 \cap \Omega_2. \end{cases}$$

Note that $\chi_1(x_i) + \chi_2(x_i) = 1$ for all $x_i \in \Omega$ (i.e for all $i = 1, \ldots, N$).

In two dimensions the domain $\Omega$, i.e., the set of $N_1 \times N_2$ equidistant points on the 2-dimensional rectangle $[a, b] \times [c, d]$, is split in an analogous way with respect to its rows. In particular we have that $\Omega_1$ and $\Omega_2$ consist of equidistant points on $[a, x_{n_1}] \times [c, d]$ and $[x_{n_1-G+1}, b] \times [c, d]$, respectively, compare Fig. 3. The splitting in more than two domains is done similarly:

Set $\Omega = \Omega_1 \cup \ldots \cup \Omega_{\mathcal{N}}$, the domain $\Omega$ decomposed into $\mathcal{N}$ domains $\Omega_i$, $i = 1, \ldots, \mathcal{N}$, where $\Omega_i$ and $\Omega_{i+1}$ are overlapping for $i = 1, \ldots, \mathcal{N} - 1$. Let

**Fig. 3** Decomposition of the image into two domains $\Omega_1$ and $\Omega_2$



$$a = x_1$$

$$\Omega_1 \cap \Omega_2$$

$$x_{n_1-G} \qquad \text{— — } \Gamma_2 \text{ — —}$$

$$\Omega_1 \cap \Omega_2$$

$$x_{n_1+1} \qquad \text{— — } \Gamma_1 \text{ — —}$$

$$\Omega_2 \cap \Omega_1$$

$$b = x_N$$
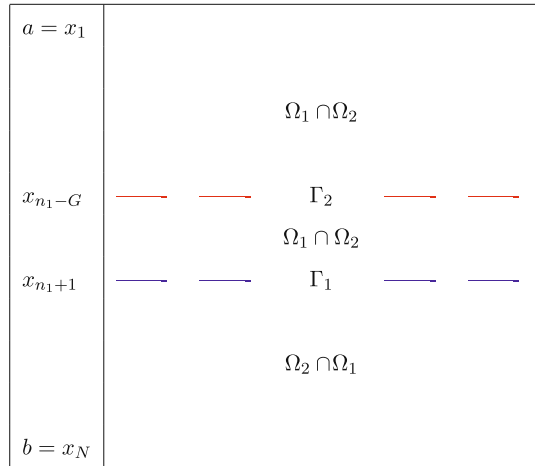
$|\Omega_i \cap \Omega_{i+1}| =: G$ for every $i = 1, \ldots, \mathcal{N} - 1$. Set $s = \lceil N_1/\mathcal{N} \rceil$. Then

$$\Omega_1 = [a = x_1, \ldots, x_{s+\frac{G}{2}}] \times \{c = y_1, \ldots, y_{N_2} = d\}$$

for $i = 2 : \mathcal{N} - 1$

$$\Omega_i = [x_{(i-1)s-\frac{G}{2}+1}, \ldots, x_{is+\frac{G}{2}}] \times \{c = y_1, \ldots, y_{N_2} = d\}$$

end

$$\Omega_{\mathcal{N}} = [x_{(\mathcal{N}-1)s-\frac{G}{2}+1}, x_{N_1}] \times \{c = y_1, \ldots, y_{N_2} = d\}.$$

The auxiliary functions $\chi_i$ can be chosen in an analogous way as in the one dimensional case:

$$\chi_i(x_{i_1}, y_{i_2}) = \begin{cases} \frac{1}{G}(i_1 - ((i-1)s - G/2 + 1)) & (x_{i_1}, y_{i_2}) \in \Omega_{i-1} \cap \Omega_i \\ 1 & (x_{i_1}, y_{i_2}) \in \Omega_i \setminus (\Omega_{i-1} \cup \Omega_{i+1}) \\ 1 - \frac{1}{G}(i_1 - (is - G/2)) & (x_{i_1}, y_{i_2}) \in \Omega_i \cap \Omega_{i+1} \end{cases}$$

for $i = 1, \ldots, \mathcal{N}$ with $\Omega_0 = \Omega_{\mathcal{N}+1} = \emptyset$.

To compute the fixed point $\eta$ of (15) in an efficient way, we make the following considerations, which allow to restrict the computation from $\Omega_1$ to a relatively small stripe around the interface. The fixed point $\eta$ is actually supported on $\Omega_1$ only, i.e., $\eta(x) = 0$ in $\Omega_1 \setminus \hat{\Omega}_1$. Hence, we restrict the fixed point iteration for $\eta$ to a relatively small stripe $\hat{\Omega}_1 \subset \Omega_1$. Analogously, one implements the minimization of $\eta_2$ on $\hat{\Omega}_2$. A similar trick was also used in [26] to compute suitable Lagrange multipliers at the interfaces of the nonoverlapping domains. However, there we needed to consider larger "bilateral stripes" around the support of the multiplier, making the numerical computation slightly more demanding for that algorithm.
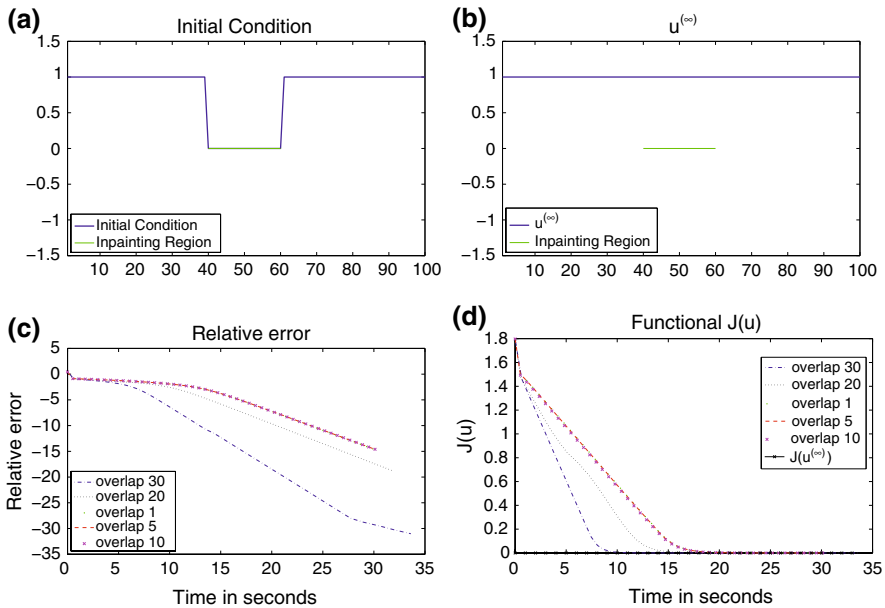
**Fig. 4** We present a numerical experiment related to the interpolation of a 1D signal by total variation minimization. The original signal is only provided outside of the green subinterval. The initial datum $g$ is shown in **a**. As expected, the minimizer $u^{(\infty)}$ is the constant vector 1, as shown in **b**. In **c** and **d** we display the rates of decay of the relative error and of the value of $\mathcal{J}$, respectively, for applications of the algorithm (23) with different sizes $G = 1, 5, 10, 20, 30$ of the overlapping region of two subintervals

## 7.3 Numerical experiments

### 7.3.1 Sequential implementation

In the following we present numerical examples for the sequential algorithm (23) in two particular applications: signal interpolation/image inpainting, and compressed sensing. The scope of the section is to illustrate by simple examples the main properties of the algorithms, as proven in our theoretical analysis. In particular, we emphasize the monotonicity properties of the algorithms with respect to the energy $\mathcal{J}$, the boundedness of the iterations due to the implementation of BUPUs, and the robustness in correctly computing minimizers independently of the size of overlapping regions.

In the numerical experiments the value for the parameter $\alpha$ has been chosen experimentally, i.e., we chose the value which gave the best compromise between visual quality of the minimizer and computational time of the algorithm. Note however, that there exist more systematic ways to choose an optimal value for $\alpha$, where the choice depends both on the data noise level and the exact solution of the problem, cf., e.g., [23] for a general approach in regularized inverse problems, or [12] for a discussion of the correspondence between the noise level and $\alpha$ in the case of total variation minimization.

In Figs. 4 and 5 we show a partially corrupted 1D signal on an interval     of 100 sampling points, with a loss of information on an interval $D \subset$     . The domain $D$ of

**(a)**

Initial Condition



**(b)**

$u^{(\infty)}$



**(c)**

Relative error



**(d)**
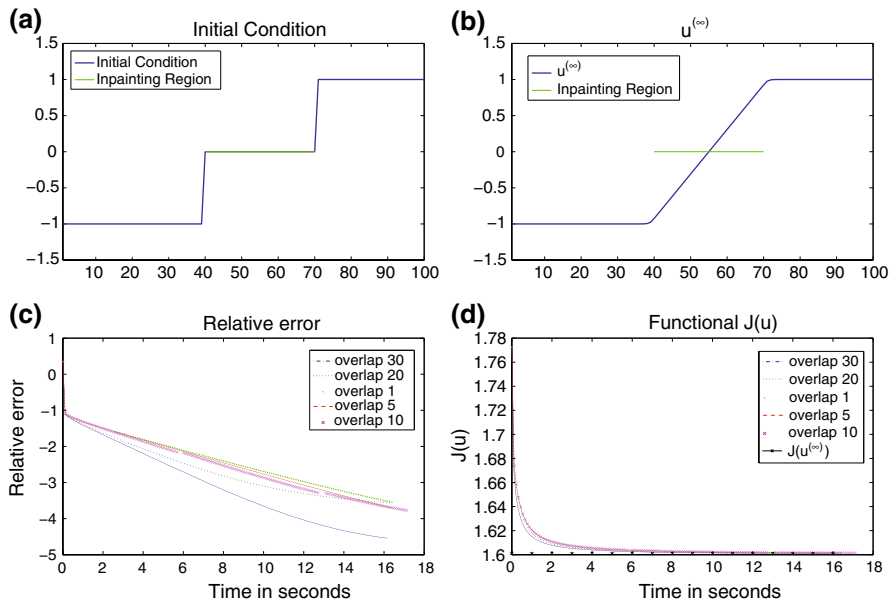
Functional J(u)



**Fig. 5** We show a second example of total variation interpolation in 1D. The initial datum $g$ is shown in **a**. As expected, a minimizer $u^{(\infty)}$ is (nearly) a piecewise linear function, as shown in **b**. In **c** and **d** we display the rates of decay of the relative error and of the value of $\mathcal{J}$ respectively, for applications of the algorithm (23) with different sizes $G = 1, 5, 10, 20, 30$ of the overlapping region of two subintervals

the missing signal points is marked in green. These signal points are reconstructed by total variation interpolation, i.e., minimizing the functional $\mathcal{J}$ in (3) with $\alpha = 0.4$ and $Tu = 1_{\backslash D} \cdot u$, where $1_{\backslash D}$ is the indicator function of $\Omega \setminus D$. A minimizer $u^{(\infty)}$ of $\mathcal{J}$ is precomputed with an algorithm working on the whole interval $\Omega$ without any decomposition. We show also the decay of relative error and of the value of the energy $\mathcal{J}$ for applications of algorithm (23) on two subdomains and with different overlap sizes $G = 1, 5, 10, 20, 30$. The fixed points $\eta$ are computed on a small interval $\hat{\Omega}_i, i = 1, 2$, of size 2. These results confirm the behavior of the algorithm (23) as predicted by the theory; the algorithm monotonically decreases $\mathcal{J}$ and computes a minimizer, independently of the size of the overlapping region. A larger overlapping region does not necessarily imply a slower convergence. In these figures we compare the speed in terms of CPU time. In Fig. 6 we also illustrate the effect of implementing the BUPU within the domain decomposition algorithm. In this case, with datum $g$ as in Fig. 5, we chose $\alpha = 1$ and an overlap of size $G = 10$. The fixed points $\eta$ are computed on a small interval $\hat{\Omega}_i, i = 1, 2$, respectively, of size 6.

Figure 7 shows an example of the domain decomposition algorithm (23) for total variation inpainting. As for the 1D example in Figs. 4, 5, and 6, the operator $T$ is a multiplier, i.e., $Tu = 1_{\backslash D} \cdot u$, where $\Omega$ denotes the rectangular image domain and $D \subset \Omega$ the missing domain in which the original image content got lost. The regularization parameter $\alpha$ is fixed at the value $10^{-2}$. In Fig. 7 the missing domain $D$ is the black writing which covers parts of the image. Here, the image domain of size $449 \times 570$ pixels
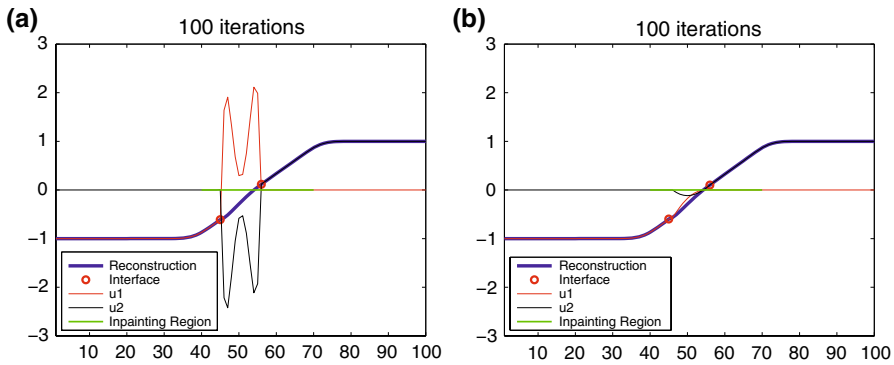
**Fig. 6** Here we present two numerical experiments related to the interpolation of a 1D signal by total variation minimization. The original signal is only provided outside of the green subinterval. On the *left* we show an application of algorithm (23) when no correction with the partition of unity is provided. In this case, the sequence of the local iterations $u_1^{(n)}, u_2^{(n)}$ is unbounded. On the *right* we show an application of algorithm (23) with the use of the partition of unity which enforces the uniform boundedness of the local iterations $u_1^{(n)}, u_2^{(n)}$
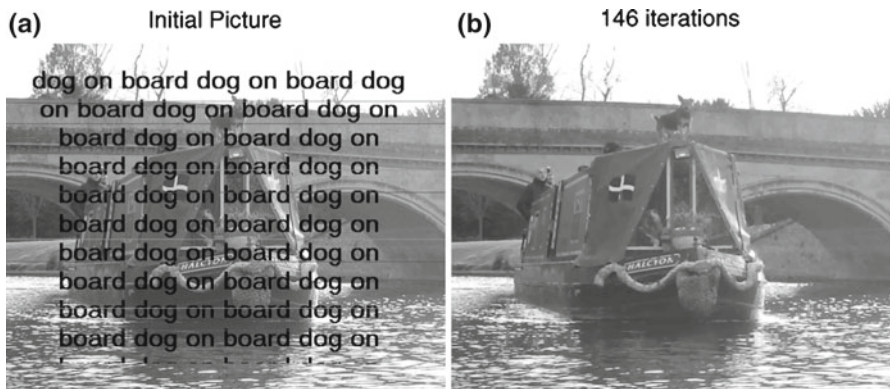


**Fig. 7** This figure shows an application of algorithm (23) for image inpainting. In this simulation the problem was split into five subproblems on overlapping subdomains

is split into five overlapping subdomains with an overlap size $G = 28 \times 570$. Further, the fixed points $\eta$ are computed on a small stripe $\hat{}_i, i = 1, \ldots, 5$, respectively, of size $6 \times 570$ pixels.

Finally, in Fig. 8 we illustrate the successful application of our domain decomposition algorithm (23) for a compressed sensing problem. Here, we consider a medical-type image (the so-called *Logan–Shepp phantom*) and its reconstruction from only partial Fourier data. In this case the linear operator $T = S \circ \mathcal{F}$, where $\mathcal{F}$ denotes the $2D$ Fourier matrix and $S$ is a *downsampling operator* which selects only a few frequencies as output. We minimize $\mathcal{J}$ with $\alpha$ set at $0.4 \times 10^{-2}$. In the application of algorithm (23), the image domain of size $256 \times 256$ pixels is split into four overlapping subdomains with an overlap size $G = 20 \times 256$. The fixed points $\eta$'s are computed in a small stripe $\hat{}_i, i = 1, \ldots, 4$, respectively, of size $6 \times 256$ pixels.

Sampling domain in the frequency plane
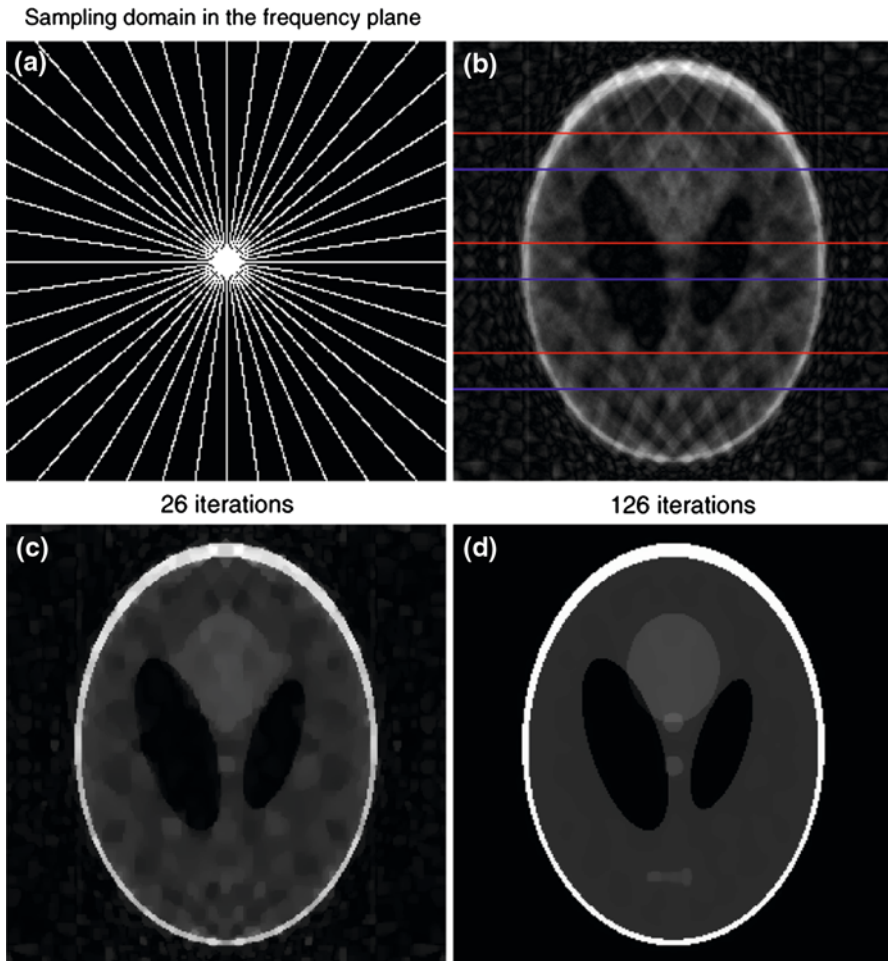


**Fig. 8** We show an application of algorithm (23) in a classical compressed sensing problem for recovering piecewise constant medical-type images from given partial Fourier data. In this simulation the problem was split via decomposition into four overlapping subdomains. On the *top-left* figure, we show the sampling data of the image in the Fourier domain. On the *top-right* the back-projection provided by the sampled frequency data together with the highlighted partition of the physical domain into four subdomains is shown. The *bottom* figures present intermediate iterations of the algorithm, i.e., $u^{(26)}$ and $u^{(125)}$

*Remark 7.1* The optimization of the different parameters of the algorithm, namely, the number of subdomains, the extent of the overlapping regions, the number of internal iterations, and the relationship with the rate of convergence are a very challenging problem, and a matter of current investigation.

### 7.3.2 Parallel implementation

In this section we show that already with relatively large images, classical methods do not scale well and may converge slowly, whereas the parallel implementation of

**Fig. 9** Image size $5616 \times 3744$
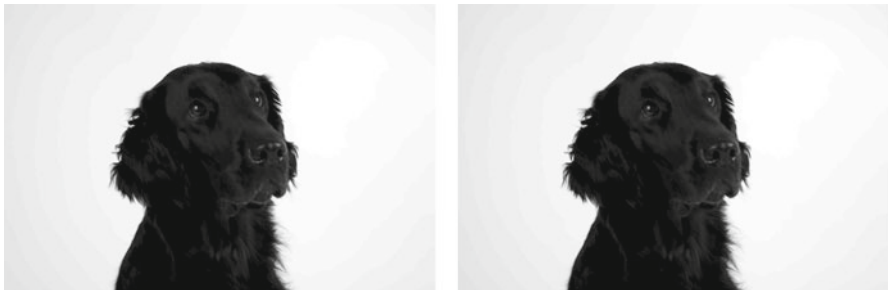


**Fig. 10** The computed approximate minimizers for the regularization parameter $\alpha = 0.1$, 100 outer iterations and 3 inner iterations, for 4 domains (*left picture*) and 16 domains (*right picture*)

the algorithm on a multiple processor computer allows for significant reductions of the CPU time, which improves with the number of subdomains. For a fair comparison we utilize Chambolle's algorithm both for the solution on the whole domain and on subdomains. We expect that the use of other algorithms [11,16,27,32,39] may change quantitatively the results but not qualitatively. Of course, considering problems of even larger size, for example in higher dimension, can only further promote and favour the use of parallel strategies.

In Fig. 9 we depict an image of size $5616 \times 3744$, which has been vandalized by superimposing a text. In Fig. 10 we show the results due to the application of the parallel algorithm (44) for performing an inpainting procedure, acting on 4 and 16 subdomains.

We do not dispose of a minimizer in this case (except for considering as an approximate minimizer one of the iterations $u^{(n)}$ for $n$ very large) and the value of $\mathcal{J}$ is not a good indicator of the quality of image restoration. Being the scope of the minimization the approximate recovery of the original image, which we may denote Org, we use as a stopping criterion $\|u^{(n)} - \text{Org}\|_2^2 < \epsilon$, for a prescribed tolerance . While the algorithm applied on the whole domain does not reach the prescribed accuracy after more than 6 h of running time, the computation with multiple subdomains can reach

**Table 1** Regularization parameter $\alpha = 0.1$, 3 inner iterations on the subdomains

|                     | 1 domain     | 4 domains   | 16 domains  |
| ------------------- | ------------ | ----------- | ----------- |
| CPU time            | 23086.68 s   | 6531.94 s   | 1583.52 s   |
| No. outer iterations | 1,000       | 10          | 10          |

The stopping criterion for all three algorithms is when the squared $L^2$-norm of the difference between the current minimizer and the original image $\|u^{(n)} - \text{Org}\|_2^2$ gets below $= 0.0048$

the result in less than half an hour. We also emphasize that in these experiments the computational time decreases linearly with the number of subdomains, showing that the computation of the Lagrange multipliers, used in our algorithm in order to correctly interface the patches, has a nearly negligible cost with respect to the minimizations on the subdomains, see Table 1.

### Appendix A: Proof of Proposition 5.2

It is clear that $\zeta \in \partial \mathcal{J}_\varphi(u)$ if and only if $u = \arg\min_{v \in \mathcal{H}}\{\mathcal{J}_\varphi(v) - \langle \zeta, v \rangle_\mathcal{H}\}$. Let us consider the following variational problem:

$$\inf_{v \in \mathcal{H}}\{\mathcal{J}_\varphi(v) - \langle \zeta, v \rangle_\mathcal{H}\} = \inf_{v \in \mathcal{H}}\|Tv - g\|_2^2 + 2\alpha\varphi(|\nabla v|)(\Omega) - \langle \zeta, v \rangle_\mathcal{H} \quad (\mathcal{P})$$

We denote such an infimum by $\inf(\mathcal{P})$. Now we compute $(\mathcal{P}^*)$ the dual of $(\mathcal{P})$. Let $\mathcal{F} : \mathcal{H} \to \mathbb{R}, \mathcal{G} : \mathcal{H} \times \mathcal{H}^d \to \mathbb{R}, \mathcal{G}_1 : \mathcal{H} \to \mathbb{R}, \mathcal{G}_2 : \mathcal{H}^d \to \mathbb{R}$, such that

$$\mathcal{F}(v) = -\langle \zeta, v \rangle_\mathcal{H},$$
$$\mathcal{G}_1(w_0) = \|w_0 - g\|_2^2,$$
$$\mathcal{G}_2(\bar{w}) = 2\alpha\varphi(|\bar{w}|)(\Omega),$$
$$\mathcal{G}(w) = \mathcal{G}_1(w_0) + \mathcal{G}_2(\bar{w}),$$

with $w = (w_0, \bar{w}) \in \mathcal{H} \times \mathcal{H}^d$. Then the dual problem of $(\mathcal{P})$ is given by (cf. [22, p 60])

$$\sup_{p^* \in \mathcal{H} \times \mathcal{H}^d}\{-\mathcal{F}^*(\Lambda^* p^*) - \mathcal{G}^*(-p^*)\}, \quad (\mathcal{P}^*)$$

where $\Lambda : \mathcal{H} \to \mathcal{H} \times \mathcal{H}^d$ is defined by

$$\Lambda v = (Tv, (\nabla v)^1, \ldots, (\nabla v)^d),$$

and $\Lambda^*$ is its adjoint. We denote the supremum in $(\mathcal{P}^*)$ by $\sup(\mathcal{P}^*)$. Using the definition of the conjugate function we compute $\mathcal{F}^*$ and $\mathcal{G}^*$. In particular

$$
\begin{aligned}
\mathcal{F}^*(\Lambda^* p^*) = \sup_{v \in \mathcal{H}} \ \langle \Lambda^* p^*, v \rangle_{\mathcal{H}} - \mathcal{F}(v) &= \sup_{v \in \mathcal{H}} \langle \Lambda^* p^* + \zeta, v \rangle_{\mathcal{H}} \\
&= \begin{cases} 0 & \Lambda^* p^* + \zeta = 0 \\ \infty & \text{otherwise} \end{cases},
\end{aligned}
$$

where $p^* = (p_0^*, \bar{p}^*)$ and

$$
\begin{aligned}
\mathcal{G}^*(p^*) &= \sup_{w \in \mathcal{H} \times \mathcal{H}^d} \ \langle p^*, w \rangle_{\mathcal{H} \times \mathcal{H}^d} - \mathcal{G}(w) \\
&= \sup_{w = (w_0, \bar{w}) \in \mathcal{H} \times \mathcal{H}^d} \langle p_0^*, w_0 \rangle_{\mathcal{H}} + \langle \bar{p}^*, \bar{w} \rangle_{\mathcal{H}^d} - \mathcal{G}_1(w_0) - \mathcal{G}_2(\bar{w}) \\
&= \sup_{w_0 \in \mathcal{H}} \langle p_0^*, w_0 \rangle_{\mathcal{H}} - \mathcal{G}_1(w_0) + \sup_{\bar{w} \in \mathcal{H}^d} \langle \bar{p}^*, \bar{w} \rangle_{\mathcal{H}^d} - \mathcal{G}_2(\bar{w}) \\
&= \mathcal{G}_1^* \left( p_0^* \right) + \mathcal{G}_2^*(\bar{p}^*).
\end{aligned}
$$

We have that

$$\mathcal{G}_1^*(p_0^*) = \left\langle \frac{p_0^*}{4} + g, p_0^* \right\rangle_{\mathcal{H}},$$

and (see [22])

$$\mathcal{G}_2^*(\bar{p}^*) = 2\alpha \varphi_1^* \left( \frac{|\bar{p}^*|}{2\alpha} \right) (\Omega),$$

if $\frac{|\bar{p}^*(x)|}{2\alpha} \in \text{Dom } \varphi_1^*$, where $\varphi_1^*$ is the conjugate function of $\varphi_1$ defined by

$$\varphi_1(s) := \varphi(|s|) \ s \in \mathbb{R}.$$

For simplicity we include in Appendix B the explicit computation of these conjugate functions. So we can write $(\mathcal{P}^*)$ in the following way

$$\sup_{p^* \in \mathcal{K}} \ - \left\langle \frac{-p_0^*}{4} + g, -p_0^* \right\rangle_{\mathcal{H}} - 2\alpha \varphi_1^* \left( \frac{|\bar{p}^*|}{2\alpha} \right) (\Omega), \tag{49}$$

where

$$\mathcal{K} = \left\{ p^* \in \mathcal{H} \times \mathcal{H}^d : \frac{|\bar{p}^*(x)|}{2\alpha} \in \text{Dom } \varphi_1^* \text{ for all } x \in \Omega, \Lambda^* p^* + \zeta = 0 \right\}.$$

The function $\varphi_1$ also fulfills assumption $(A_\varphi)$(ii) (i.e., there exists $c_1 > 0$, $b \geq 0$ such that $c_1 z - b \leq \varphi_1(z) \leq c_1 z + b$, for all $z \in \mathbb{R}^+$). The conjugate function of $\varphi_1$ is given by $\varphi_1^*(s) = \sup_{z \in \mathbb{R}} \{\langle s, z \rangle - \varphi_1(z)\}$. Using the previous inequalities and that $\varphi_1$ is even (i.e., $\varphi_1(z) = \varphi_1(-z)$ for all $z \in \mathbb{R}$) we have

$$
(\sup_{z \in \mathbb{R}}\{\langle s, z \rangle - c_1|z| + b\} \geq) \sup_{z \in \mathbb{R}}\{\langle s, z \rangle - \varphi_1(z)\} \geq \sup_{z \in \mathbb{R}}\{\langle s, z \rangle - c_1|z| - b\}
$$

$$
= \begin{array}{ll} -b & \text{if } |s| \leq c_1 \\ \infty & \text{else} \end{array}. \tag{50}
$$

In particular, one can see that $s \in \mathrm{Dom}\,\varphi_1^*$ if and only if $|s| \leq c_1$.

From $\Lambda^* p^* + \zeta = 0$ we obtain

$$
\langle \Lambda^* p^*, \omega \rangle_{\mathcal{H}} + \langle \zeta, \omega \rangle_{\mathcal{H}} = \langle p^*, \Lambda \omega \rangle_{\mathcal{H}^{d+1}} + \langle \zeta, \omega \rangle_{\mathcal{H}}
$$
$$
= \langle p_0^*, T\omega \rangle_{\mathcal{H}} + \langle \bar{p}^*, \nabla \omega \rangle_{\mathcal{H}^d} + \langle \zeta, \omega \rangle_{\mathcal{H}} = 0 \quad \text{for all } \omega \in \mathcal{H}.
$$

Then, since $\langle \bar{p}^*, \nabla \omega \rangle_{\mathcal{H}^d} = \langle -\mathrm{div}\,\bar{p}^*, \omega \rangle_{\mathcal{H}}$ (see Sect. 2), we have

$$
T^* p_0^* - \mathrm{div}\,\bar{p}^* + \zeta = 0.
$$

Hence we can write $\mathcal{K}$ in the following way

$$
\mathcal{K} = \left\{ p^* = (p_0^*, \bar{p}^*) \in \mathcal{H} \times \mathcal{H}^d : \frac{|\bar{p}^*(x)|}{2\alpha} \leq c_1 \text{ for all } x \in \Omega, \ T^* p_0^* - \mathrm{div}\,\bar{p}^* + \zeta = 0 \right\}.
$$

We now apply the duality results from [22, Theorem III.4.1], since the functional in $(\mathcal{P})$ is convex, continuous with respect to $\Lambda v$ in $\mathcal{H} \times \mathcal{H}^d$, and $\inf(\mathcal{P})$ is finite. Then $\inf(\mathcal{P}) = \sup(\mathcal{P}^*) \in \mathbb{R}$ and $(\mathcal{P}^*)$ has a solution $M = (M_0, \bar{M}) \in \mathcal{K}$.

Let us assume that $u$ is a solution of $(\mathcal{P})$ and $M$ is a solution of $(\mathcal{P}^*)$. From $\inf(\mathcal{P}) = \sup(\mathcal{P}^*)$ we get

$$
\|Tu - g\|_2^2 + 2\alpha\varphi(|\nabla u|)(\Omega) - \langle \zeta, u \rangle_{\mathcal{H}} = -\left\langle \frac{-M_0}{4} + g, -M_0 \right\rangle_{\mathcal{H}}
$$
$$
- 2\alpha\varphi_1^*\left(\frac{|\bar{M}|}{2\alpha}\right)(\Omega), \tag{51}
$$

where $M = (M_0, \bar{M}) \in \mathcal{H} \times \mathcal{H}^d$, $\frac{|\bar{M}(x)|}{2\alpha} \leq c_1$ and $T^* M_0 - \mathrm{div}\,\bar{M} + \zeta = 0$, which verifies the direct implication of (20). In particular

$$
-\langle \zeta, u \rangle_{\mathcal{H}} = \langle T^* M_0, u \rangle_{\mathcal{H}} - \langle \mathrm{div}\,\bar{M}, u \rangle_{\mathcal{H}} = \langle M_0, Tu \rangle_{\mathcal{H}} + \langle \bar{M}, \nabla u \rangle_{\mathcal{H}^d},
$$

and

$$\|Tu - g\|_2^2 + \langle M_0, Tu \rangle_{\mathcal{H}} + \langle \bar{M}, \nabla u \rangle_{\mathcal{H}^d} + 2\alpha\varphi(|\nabla u|)(\Omega) + \langle \frac{-M_0}{4} + g, -M_0 \rangle_{\mathcal{H}}$$

$$+ 2\alpha\varphi_1^* \left( \frac{|\bar{M}|}{2\alpha} \right)(\Omega) = 0. \tag{52}$$

Let us write (52) again in the following form

$$\sum_{x \in} |(Tu - g)(x)|^2 + \sum_{x \in} M_0(x)(Tu)(x) + \sum_{x \in} \sum_{j=1}^{d} \bar{M}^j(x)(\nabla u)^j(x)$$

$$+ \sum_{x \in} 2\alpha\varphi(|(\nabla u)(x)|)$$

$$+ \sum_{x \in} \left( \frac{-M_0(x)}{4} + g(x) \right)(-M_0(x)) + \sum_{x \in} 2\alpha\varphi_1^* \left( \frac{|\bar{M}(x)|}{2\alpha} \right) = 0. \tag{53}$$

Now we have

1. $2\alpha\varphi(|(\nabla u)(x)|) + \sum_{j=1}^{d} \bar{M}^j(x)(\nabla u)^j(x) + 2\alpha\varphi_1^*(\frac{|\bar{M}(x)|}{2\alpha}) \geq 2\alpha\varphi(|(\nabla u)(x)|) - \sum_{j=1}^{d} |\bar{M}^j(x)||(\nabla u)^j(x)| + 2\alpha\varphi_1^*(\frac{|\bar{M}(x)|}{2\alpha}) \geq 0$ by the definition of $\varphi_1^*$, since $2\alpha\varphi_1^*(\frac{|\bar{M}(x)|}{2\alpha}) = \sup_{S=(S^1,...,S^d)\in\mathbb{R}^d}\{\sum_{j=1}^{d} |\bar{M}^j(x)||S^j| - 2\alpha\varphi(|S|)\}$.

2. $|(Tu - g)(x)|^2 + M_0(x)(Tu)(x) + ((\frac{-M_0(x)}{4} + g(x))(-M_0(x))) = ((Tu)(x) - g(x))^2 + M_0(x)((Tu)(x) - g(x)) + (\frac{M_0(x)}{2})^2 = ((Tu)(x) - g(x) + \frac{M_0(x)}{2})^2 \geq 0$.

Hence condition (52) reduces to

$$2\alpha\varphi(|(\nabla u)(x)|) + \sum_{j=1}^{d} \bar{M}^j(x)(\nabla u)^j(x) + 2\alpha\varphi_1^* \left( \frac{|\bar{M}(x)|}{2\alpha} \right) = 0 \quad \text{for all } x \in \Omega, \tag{54}$$

$$-M_0(x) = 2((Tu)(x) - g(x)) \quad \text{for all } x \in \Omega. \tag{55}$$

Conversely, if such an $M = (M_0, \bar{M}) \in \mathcal{H} \times \mathcal{H}^d$ with $\frac{|\bar{M}(x)|}{2\alpha} \leq c_1$ exists, which fulfills conditions (19)–(21), it is clear from previous considerations that Eq. (51) holds. Let us denote the functional on the left side of (51) by

$$P(u) := \|Tu - g\|_2^2 + 2\alpha\varphi(|\nabla u|)(\Omega) - \langle \zeta, u \rangle_{\mathcal{H}},$$

and the functional on the right side of (51) by

$$P^*(M) := - \langle \frac{-M_0}{4} + g, -M_0 \rangle_{\mathcal{H}} - 2\alpha\varphi_1^* \left( \frac{|\bar{M}|}{2\alpha} \right)(\Omega).$$

We know that the functional $P$ is the functional of ($\mathcal{P}$) and $P^*$ is the functional of ($\mathcal{P}^*$). Hence inf $P = \inf(\mathcal{P})$ and sup $P^* = \sup(\mathcal{P}^*)$. Since $P$ is convex, continuous with respect to $\Lambda u$ in $\mathcal{H} \times \mathcal{H}^d$, and $\inf(\mathcal{P})$ is finite we know from duality results [22, Theorem III.4.1] that $\inf(\mathcal{P}) = \sup(\mathcal{P}^*) \in \mathbb{R}$. We assume that $M$ is no solution of ($\mathcal{P}^*$), i.e., $P^*(M) < \sup(\mathcal{P}^*)$, and $u$ is no solution of ($\mathcal{P}$), i.e, $P(u) > \inf(\mathcal{P})$. Then we have that

$$P(u) > \inf(\mathcal{P}) = \sup(\mathcal{P}^*) > P^*(M).$$

Thus (51) is valid if and only if $M$ is a solution of ($\mathcal{P}^*$) and $u$ is a solution of ($\mathcal{P}$) which amounts to saying that $\zeta \in \partial \mathcal{J}_\varphi(u)$.

If additionally $\varphi$ is differentiable and $|(\nabla u)(x)| = 0$ for $x \in$ , we show that we can compute $\bar{M}(x)$ explicitly. From Eq. (19) (respectively, (54)) we have

$$2\alpha \varphi_1^* \left( \frac{|-\bar{M}(x)|}{2\alpha} \right) = -\langle \bar{M}(x), (\nabla u)(x) \rangle_{\mathbb{R}^d} - 2\alpha \varphi(|(\nabla u)(x)|). \qquad (56)$$

From the definition of conjugate function we have

$$
\begin{aligned}
2\alpha \varphi_1^* \left( \frac{|-\bar{M}(x)|}{2\alpha} \right) &= 2\alpha \sup_{t \in \mathbb{R}} \left( \left\langle \frac{|-\bar{M}(x)|}{2\alpha}, t \right\rangle - \varphi_1(t) \right) \\
&= 2\alpha \sup_{t \geq 0} \left( \left\langle \frac{|-\bar{M}(x)|}{2\alpha}, t \right\rangle - \varphi_1(t) \right) \\
&= 2\alpha \sup_{t \geq 0} \sup_{\substack{S \in \mathbb{R}^d \\ |S| = t}} \left( \left\langle \frac{-\bar{M}(x)}{2\alpha}, S \right\rangle_{\mathbb{R}^d} - \varphi_1(|S|) \right) \\
&= \sup_{S \in \mathbb{R}^d} \left( \langle -\bar{M}(x), S \rangle_{\mathbb{R}^d} - 2\alpha \varphi(|S|) \right). \qquad (57)
\end{aligned}
$$

Now, if $|(\nabla u)(x)| = 0$ for $x \in$ , then it follows from (56) that the supremum is taken on in $S = |(\nabla u)(x)|$ and we have

$$\nabla_S (-\langle \bar{M}(x), S \rangle_{\mathbb{R}^d} - 2\alpha \varphi(|S|)) = 0,$$

which implies

$$\bar{M}^j(x) = -2\alpha \frac{\varphi'(|(\nabla u)(x)|)}{|(\nabla u)(x)|} (\nabla u)^j(x) \quad j = 1, \dots, d,$$

and verifies (22). This finishes the proof.

## Appendix B: Computation of conjugate functions

Let us compute the conjugate function of the convex function $\mathcal{G}_1(w_0) = \| w_0 - g \|_2^2$. From Definition 5.1 we have

$$\mathcal{G}_1^* \left( p_0^* \right) = \sup_{w_0 \in \mathcal{H}} \langle w_0, p_0^* \rangle_{\mathcal{H}} - \mathcal{G}_1(w_0) = \sup_{w_0 \in \mathcal{H}} \{ \langle w_0, p_0^* \rangle_{\mathcal{H}} - \langle w_0 - g, w_0 - g \rangle_{\mathcal{H}} \}.$$

We set $H(w_0) := \langle w_0, p_0^* \rangle_{\mathcal{H}} - \langle w_0 - g, w_0 - g \rangle_{\mathcal{H}}$. To get the maximum of $H$ we compute the Gâteaux-differential at $w_0$ of $H$,

$$H'(w_0) = p_0^* - 2(w_0 - g),$$

and we solve $H'(w_0) = 0$, since $H''(w_0) < 0$, and get $w_0 = \frac{p_0}{2} + g$, being the maximizer of $H$. Thus we have that

$$\sup_{w_0 \in \mathcal{H}} H(w_0) = \langle \frac{p_0^*}{4} + g, p_0^* \rangle_{\mathcal{H}} = \mathcal{G}_1^*(p_0^*).$$

Now we are going to compute the conjugate function of $\mathcal{G}_2(\bar{w}) = 2\alpha\varphi(|\bar{w}|)(\Omega)$. Associated to our notations we define the space $\mathcal{H}_0^+ = \mathbb{R}_0^{+N_1 \times \cdots \times N_d}$. From Definition 5.1 we have

$$\mathcal{G}_2^*(\bar{p}^*) = \sup_{\bar{w} \in \mathcal{H}^d} \{ \langle \bar{w}, \bar{p}^* \rangle_{\mathcal{H}^d} - 2\alpha\varphi(|\bar{w}|)(\Omega) \}$$

$$= \sup_{t \in \mathcal{H}_0^+} \sup_{\substack{\bar{w} \in \mathcal{H}^d \\ |\bar{w}(x)| = t(x)}} \{ \langle \bar{w}, \bar{p}^* \rangle_{\mathcal{H}^d} - 2\alpha\varphi(|\bar{w}|)(\Omega) \}$$

$$= \sup_{t \in \mathcal{H}_0^+} \{ \langle t, |\bar{p}^*| \rangle_{\mathcal{H}} - 2\alpha\varphi(t)(\Omega) \}.$$

If $\varphi$ were an even function then

$$\sup_{t \in \mathcal{H}_0^+} \{ \langle t, |\bar{p}^*| \rangle_{\mathcal{H}} - 2\alpha\varphi(t)(\Omega) \} = \sup_{t \in \mathcal{H}} \{ \langle t, |\bar{p}^*| \rangle_{\mathcal{H}} - 2\alpha\varphi(t)(\Omega) \}$$

$$= 2\alpha \sup_{t \in \mathcal{H}} \langle t, \frac{|\bar{p}^*|}{2\alpha} \rangle_{\mathcal{H}} - \varphi(t)(\Omega)$$

$$= 2\alpha\varphi^* \left( \frac{|\bar{p}^*|}{2\alpha} \right)(\Omega),$$

where $\varphi^*$ is the conjugate function of $\varphi$.

Unfortunately $\varphi$ is not even in general. To overcome this difficulty we have to choose a function which is equal to $\varphi(s)$ for $s \geq 0$ and does not change the supremum

for $s < 0$. For instance, one can choose $\varphi_1(s) = \varphi(|s|)$ for $s \in \mathbb{R}$. Then we have

$$\sup_{t \in \mathcal{H}_0^+} \{\langle t, |\bar{p}^*|\rangle_{\mathcal{H}} - 2\alpha\varphi(t)(\Omega)\} = \sup_{t \in \mathcal{H}}\{\langle t, |\bar{p}^*|\rangle_{\mathcal{H}} - 2\alpha\varphi_1(t)(\Omega)\}$$

$$= 2\alpha \sup_{t \in \mathcal{H}} \left\langle t, \frac{|\bar{p}^*|}{2\alpha} \right\rangle_{\mathcal{H}} - \varphi_1(t)(\Omega)$$

$$= 2\alpha\varphi_1^* \left(\frac{|\bar{p}^*|}{2\alpha}\right)(\Omega),$$

where $\varphi_1^*$ is the conjugate function of $\varphi_1$. Note that one can also choose $\varphi_1(s) = \varphi(s)$ for $s \geq 0$ and $\varphi_1(s) = \infty$ for $s < 0$.

## References

1. Ambrosio, L., Fusco, N., Pallara, D.: Functions of Bounded Variation and Free Discontinuity Problems. Oxford Mathematical Monographs, xviii. Clarendon Press, Oxford (2000)
2. Aubert, G., Kornprobst, P.: Mathematical Problems in Image Processing. Partial Differential Equations and the Calculus of Variation. Springer, Berlin (2002)
3. Bauschke, H.H., Borwein, J.M., Lewis, A.S.: The method of cyclic projections for closed convex sets in Hilbert space. Recent developments in optimization theory and nonlinear analysis (Jerusalem, 1995), pp 1–38, Contemp. Math., 204, Amer. Math. Soc., Providence (1997)
4. Braides, A.: -Convergence for Beginners, No.22 in Oxford Lecture Series in Mathematics and Its Applications. Oxford University Press, Oxford (2002)
5. Bruck, R.E., Reich, S.: Nonexpansive projections and resolvents of accretive operators in Banach spaces. Houston J. Math. **3**(4), 459–470 (1977)
6. Candès, E.J.: Compressive sampling, Int. Congress of Mathematics, **3**, pp. 1433–1452 (2006)
7. Candès, E.J., Romberg, J., Tao, T.: Exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theory **52**(2), 489–509 (2006)
8. Candès, E.J., Tao, T.: Near optimal signal recovery from random projections: universal encoding strategies? IEEE Trans. Inf. Theory **52**(12), 5406–5425 (2006)
9. Carstensen, C.: Domain decomposition for a non-smooth convex minimalization problems and its application to plasticity. Numer. Linear Algebra Appl. **4**(3), 177–190 (1998)
10. Chambolle, A.: An algorithm for total variation minimization and applications. J. Math. Imaging Vis. **20**(1–2), 89–97 (2004)
11. Chambolle, A., Darbon, J.: On total variation minimization and surface evolution using parametric maximum flows. Int. J. Comput. Vis. **84**(3), 288–307 (2009)
12. Chambolle, A., Lions, P.-L.: Image recovery via total variation minimization and related problems. Numer. Math. **76**(2), 167–188 (1997)
13. Chan, T.F., Golub, G.H., Mulet, P.: A nonlinear primal-dual method for total variation-based image restoration. SIAM J. Sci. Comput. **20**(6), 1964–1977 (1999)
14. Chan, T.F., Mathew, T.P.: Domain decomposition algorithms. Acta Numer. **3**, 61–143 (1994)
15. Combettes, P.L., Wajs, V.R.: Signal recovery by proximal forward-backward splitting. Multiscale Model. Simul. **4**(4), 1168–1200 (2005)
16. Darbon, J., Sigelle, M.: A fast and exact algorithm for total variation minimization. IbPRIA 2005 **3522**(1), 351–359 (2005)
17. Daubechies, I., Defrise, M., De Mol, C.: An iterative thresholding algorithm for linear inverse problems. Comm. Pure Appl. Math. **57**(11), 1413–1457 (2004)
18. Daubechies, I., DeVore, R.A., Fornasier, M., Güntürk, C.S.: Iteratively re-weighted least squares minimization for sparse recovery. Comm. Pure Appl. Math. **63**(1), 1–38 (2010)
19. Daubechies, I., Teschke, G., Vese, L.: Iteratively solving linear inverse problems under general convex constraints. Inverse Probl. Imaging **1**(1), 29–46 (2007)
20. Dobson, D., Vogel, C.R.: Convergence of an iterative method for total variation denoising. SIAM J. Numer. Anal. **34**(5), 1779–1791 (1997)

21. Donoho, D.L.: Compressed sensing. IEEE Trans. Inf. Theory **52**(4), 1289–1306 (2006)
22. Ekeland, I., Temam, R.: Convex analysis and variational problems. Translated by Minerva Translations, Ltd., London., Studies in Mathematics and its Applications, vol. 1. North-Holland Publishing Company/American Elsevier Publishing Company, Inc., Oxford/New York (1976)
23. Engl, H.W., Hanke, M., Neubauer, A.: Regularization of inverse problems. Mathematics and its Applications (Dordrecht), vol. 375. Kluwer , Dordrecht (1996)
24. Evans, L.C., Gariepy, R.F.: Measure Theory and Fine Properties of Functions. CRC Press,  (1992)
25. Fornasier, M.: Domain decomposition methods for linear inverse problems with sparsity constraints. Inverse Problems **23**(6), 2505–2526 (2007)
26. Fornasier, M., Schönlieb, C.-B.: Subspace correction methods for total variation and $\ell_1-$minimization. SIAM J. Numer. Anal. **47**(5), 3397–3428 (2009)
27. Goldstein, T., Osher, S.: The split Bregman method for $L^1$ regularized problems. SIAM J. Imaging Sci. **2**(2), 323–343 (2009)
28. Hiriart-Urruty, J.-B.H., Lemaréchal, C.: Convex Analysis and Minimization Algorithms I, vol. 305 of Grundlehren der mathematischen Wissenschaften. Springer, Berlin (1996)
29. Lustig, M., Donoho, D., Pauly, J.M.: Sparse MRI: The application of compressed sensing for rapid MR imaging. Magn. Reson. Med. **58**(6), 1182–1195 (2007)
30. Müller, J.: Parallel Methods for Nonlinear Imaging Techniques, Master thesis, University of Münster, (2008)
31. Nesterov, Y.: Smooth minimization of non-smooth functions. Math. Program. Ser. A **103**, 127–152 (2005)
32. Osher, S., Burger, M., Goldfarb, D., Xu, J., Yin, W.: An iterative regularization method for total variation-based image restoration. Multiscale Model. Simul. **4**(2), 460–489 (2005)
33. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Phys. D **60**(1–4), 259–268 (1992)
34. Schönlieb, C.-B.: Total variation minimization with an $H^{-1}$ constraint, CRM Series 9, Singularities in Nonlinear Evolution Phenomena and Applications proceedings, Scuola Normale Superiore Pisa 2009, 201–232
35. Tai, X.-C., Tseng, P.: Convergence rate analysis of an asynchronous space decomposition method for convex minimization. Math. Comp. **71**(239), 1105–1135 (2004)
36. Tai, X.-C., Xu, J.: Global convergence of subspace correction methods for convex optimization problems. Math. Comp. **71**(237), 105–124 (2002)
37. Vese, L.: A study in the BV space of a denoising-deblurring variational problem. Appl. Math. Optim. **44**, 131–161 (2001)
38. Warga, J.: Minimizing certain convex functions. J. Soc. Indust. Appl. Math. **11**, 588–593 (1963)
39. Weiss, P., Blanc-Féraud, L., Aubert, G.: Efficient schemes for total variation minimization under constraints in image processing. SIAM J. Sci. Comput. **31**(3), 2047–2080 (2009)
40. Xu, J., Tai, X.-C., Wang, L.-L.: A Two-level domain decomposition method for image restoration, UCLA Computational and Applied Mathematics Report 09-92, November 2009
41. Matlab code and numerical experiments of the methods provided in this paper can be downloaded at the web-page: http://homepage.univie.ac.at/carola.schoenlieb/webpage_tvdode/tv_dode_numerics.htm