WAVELET DECOMPOSITION METHOD FOR L_2/TV-IMAGE DEBLURRING

M. FORNASIER*, Y. KIM†, A. LANGER‡, AND C.-B. SCHÖNLIEB§

Abstract. In this paper we show additional properties of the limit of the sequence produced by the subspace correction algorithm proposed by Fornasier and Schönlieb [24] for L_2/TV-minimization problems. Inspired by the work of Vonesch and Unser [34], we adapt and specify this algorithm to the case of an orthogonal wavelet space decomposition and for deblurring problems.

Key words. Image deblurring, wavelet decomposition method, convex optimization, oblique-thresholding, total variation minimization, alternating minimization

AMS subject classifications. 65K10, 65M32, 49M27, 68U10, 90C25, 49J40, 42C40

1. Introduction. In image processing, one is interested in the restoration of an observed image, which is corrupted by a measurement device. Let $\Omega = [0,1]^2$ and $T : L_2(\Omega) \to L_2(\Omega)$ be a blur operator modelled as a convolution $Tu = u * \kappa$, with kernel $\kappa \in L_1(\Omega)$. Then the ideal observed noiseless image $\tilde{g}$ can be described as

$$\tilde{g} = Tu,$$

where $u \in L_2(\Omega)$ is the unknown image, which we would like to reconstruct. In general, the observed data is additionally corrupted by noise e, i.e.,

$$g = Tu + e. \tag{1.1}$$

We are in particular interested in the recovery of u from the given noisy observed image g when the operator T is not invertible or ill-conditioned, and regularization techniques are required [18].

Images can be well approximated using the superposition of few wavelets [14, 29]. Hence we make the realistic assumption that u can be represented by a sparse wavelet expansion, i.e., for a given wavelet basis $\{\psi_\lambda : \lambda \in \Lambda\}$ indexed by a countable set Λ the image u can be well approximated by a series expansion with few nonvanishing coefficients of the form

$$u \approx Su_\Lambda = \sum_{\lambda \in \Lambda} u_\lambda \psi_\lambda,$$

where $u_\Lambda = (u_\lambda)_{\lambda \in \Lambda} \in \ell_2(\Lambda)$ and $S : \ell_2(\Lambda) \to L_2(\Omega)$ is a bounded linear operator, called the synthesis operator. It is acknowledged that the simultaneous minimization of the least-squares discrepancy to data and of the ℓ_1-norm of coefficients promotes sparsity [15]. Hence we consider the minimization of the functional

$$J(u_\Lambda) = \|Au_\Lambda - g\|^2_{L_2(\Omega)} + 2\alpha \|u_\Lambda\|_{\ell_1(\Lambda)} = \|Au_\Lambda - g\|^2_{L_2(\Omega)} + 2\alpha \sum_{\lambda \in \Lambda} |u_\lambda| \tag{1.2}$$

with respect to the vector of wavelet coefficients $u_\Lambda = (u_\lambda)_{\lambda \in \Lambda}$, where $\alpha > 0$ is a fixed regularization parameter, and $A = T \circ S : \ell_2(\Lambda) \to L_2(\Omega)$ is the composition of the synthesis map S and the operator T. In order to address this minimization with respect to u_Λ, one can use, for instance, the so-called *iterative soft-thresholding algorithm* [15]: pick an initial $u_\Lambda^{(0)} \in \ell_2(\Lambda)$ and iterate

$$u_\Lambda^{(n+1)} = \mathbb{S}_\alpha(u_\Lambda^{(n)} + A^*(g - Au_\Lambda^{(n)})), \quad n \geq 0, \tag{1.3}$$

*Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstrasse 69, A-4040, Linz, Austria Email: massimo.fornasier@oeaw.ac.at

†Department of Mathematics, University of California, Irvine, 410X, Rowland Hall, Irvine, CA 92697-3875, USA Email: yuno1123@math.uci.edu

‡Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstrasse 69, A-4040, Linz, Austria Email: andreas.langer@oeaw.ac.at

§Department of Applied Mathematics and Theoretical Physics (DAMTP), University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom Email: C.B.Schoenlieb@damtp.cam.ac.uk

where $\mathbb{S}_\alpha : \ell_2(\Lambda) \to \ell_2(\Lambda)$ is defined componentwise by $\mathbb{S}_\alpha(v) = (S_\alpha v_\lambda)_{\lambda\in\Lambda}$, and

$$S_\alpha(v) = \begin{cases} v - \operatorname{sign}(v)\alpha & |v| > \alpha \\ 0 & \text{otherwise} \end{cases}$$

is the so-called *soft-thresholding operator*. The strong convergence of the algorithm in 1.3 to minimizers of J is proved in [15]. In [5] it was shown that under additional conditions on the operator A or on minimizers of 1.2 the algorithm in 1.3 converges linearly, although with a rather poor rate in general, see [22] for a broader discussion. There exist several alternative approaches, that promise to solve ℓ_1-minimization with fast convergence [16, 20, 27, 3]. One way to accelerate the speed of convergence of minimizing iterative soft-thresholding algorithms for large-scale problems was proposed in [21]. There a domain decomposition method for ℓ_1-norm minimization was introduced and analyzed. The main idea of this algorithm is to decompose the index set Λ into two (or more) disjoint sets Λ_i, $i = 1, 2, \ldots$, such that $\Lambda = \Lambda_1 \cup \Lambda_2$. Associated with this decomposition we define $\mathcal{V}_i = \{u_\Lambda \in \ell_2(\Lambda) : \operatorname{supp}(u_\Lambda) \subset \Lambda_i\}$ for $i = 1, 2$. Then we minimize J by using the following alternating algorithm: pick an initial $\mathcal{V}_1 \oplus \mathcal{V}_2 \ni u_{\Lambda_1}^{(0)} + u_{\Lambda_2}^{(0)} := u_\Lambda^{(0)}$, for example $u^{(0)} = 0$, and iterate

$$\left\{ \begin{array}{l} u_{\Lambda_1}^{(n+1)} \approx \arg\min_{u_{\Lambda_1}\in\mathcal{V}_1} J(u_{\Lambda_1} + u_{\Lambda_2}^{(n)}) \\ u_{\Lambda_2}^{(n+1)} \approx \arg\min_{u_{\Lambda_2}\in\mathcal{V}_2} J(u_{\Lambda_1}^{(n+1)}, u_{\Lambda_2}) \\ u_\Lambda^{(n+1)} := u_{\Lambda_1}^{(n+1)} + u_{\Lambda_2}^{(n+1)}, \end{array} \right. \tag{1.4}$$

where u_{Λ_i} is supported on Λ_i only, $i = 1, 2$. We observe that the ℓ_1-norm splits additively

$$\|u_{\Lambda_1} + u_{\Lambda_2}\|_{\ell_1(\Lambda)} = \|u_{\Lambda_1}\|_{\ell_1(\Lambda_2)} + \|u_{\Lambda_2}\|_{\ell_1(\Lambda_2)},$$

and hence the subproblems in (1.4) are of the same kind as the original problem (1.2), i.e., for example for the problem on Λ_1 we have

$$\arg\min_{u_{\Lambda_1}\in\mathcal{V}_1} J(u_{\Lambda_1} + u_{\Lambda_2}^{(n)}) = \arg\min_{u_{\Lambda_1}\in\mathcal{V}_1} \|A_{\Lambda_1} u_{\Lambda_1} - (g - A_{\Lambda_2} u_{\Lambda_2}^{(n)})\|_{L_2(\Omega)}^2 + 2\alpha\|u_{\Lambda_1}\|_{\ell_1(\Lambda_1)},$$

where A_{Λ_i} is the restriction of the matrix A to the columns indexed by Λ_i. Therefore, for solving the subminimization problems of (1.4) we can use one of the before mentioned methods, for example again the iterative thresholding algorithm:

$$u_{\Lambda_i}^{(\ell+1,n+1)} = \mathbb{S}_\alpha(u_{\Lambda_i}^{(\ell,n+1)} + A_{\Lambda_i}^*((g - A_{\Lambda_{\hat{i}}} u_{\Lambda_{\hat{i}}}^{(n)}) - A_{\Lambda_i} u_{\Lambda_i}^{(\ell,n+1)})), \quad \hat{i} \in \{1, 2\} \setminus \{i\}. \tag{1.5}$$

Great advantages of this domain decomposition algorithm are that we can solve instead of one large problem several smaller problems, which might lead to an acceleration of convergence with a reduction of overall computational cost, and that it can be easily parallelized. Convergence of both the sequential and the parallel versions of this algorithm is proven in [21]. The same method was used in [34] by Vonesch and Unser with minor modifications, specifically by using Haar wavelets for deblurring (or deconvolution) problems, where cyclic updates of the different resolution levels were combined with the preconditioning effect of subband-specific parameters. The effectiveness of this method was shown by solving multidimensional image deconvolution problems, as 3D fluorescence microscopy. We give a brief and intuitive explanation of the reason why this multilevel method works so well for deblurring problems: wavelet space decompositions split the function space into orthogonal subspaces $\mathcal{V}_i$. Note that T is just a convolution operator with kernel κ or a multiplier $\hat{\kappa}$ in the Fourier domain, where the $\mathcal{V}_i$'s represent nearly disjoint dyadic subbands, and we have that all A_{Λ_i} are also nearly orthogonal, i.e., $A_{\Lambda_i}^* A_{\Lambda_{\hat{i}}} \approx 0$ for $i \neq \hat{i}$. Hence each subiteration (1.5) of the algorithm in (1.4) is (nearly) restricted to one of the $\mathcal{V}_i$, independent of other subiterations, and converges fast as $A_{\Lambda_i}^* A_{\Lambda_i}$ is a well-conditioned operator. This is the case whenever the Fourier transform $\hat{\kappa}$ is, for example, a slowly decaying function on the subband associated with $\mathcal{V}_i$, see Figure 1.1.

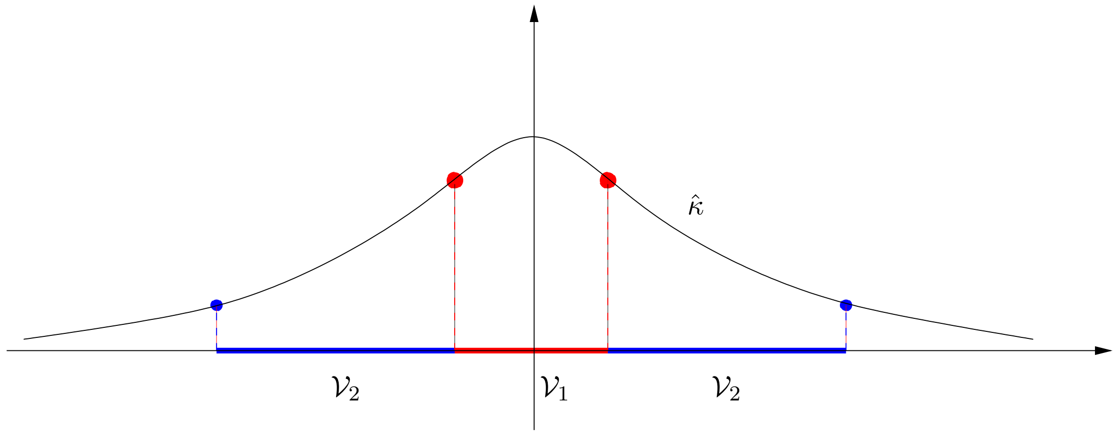

Fig. 1.1. *We depict a slowly decaying envelope of the Fourier transform $\hat{\kappa}$ of a kernel κ. The spaces $\mathcal{V}_1$ and $\mathcal{V}_2$ are two orthogonal spaces, obtained by a wavelet decomposition and associated to nearly disjoint subbands. Restricted on the subband associated to $\mathcal{V}_i$, the function $\hat{\kappa}$, essentially representing the spectrum of the matrix A_{Λ_i}, can be intuitively understood as bounded from above and below, providing the well-conditioning of the operator $A^*_{\Lambda_i}A_{\Lambda_i}$.*

To gain maximal performance of the algorithm in (1.4) we need to introduce preconditioner constants for each subiteration respectively, i.e., instead of considering $I - A^*_{\Lambda_i}A_{\Lambda_i}$ we take iteration operators

$$I - \frac{1}{\alpha_i} A^*_{\Lambda_i}A_{\Lambda_i},$$

for $\alpha_i \geq \|A_{\Lambda_i}\|^2$.

The main goal of this paper is to transpose these observations on preconditioning effects of alternating algorithms based on wavelet decompositions to the deblurring model where the term $\|u_\Lambda\|_{\ell_1(\Lambda)}$ in (1.2) is substituted by the *total variation* of the function u. We recall that for $u \in L_1(\Omega)$

$$V(u,\Omega) := \sup\left\{ \int_\Omega u \operatorname{div} \varphi \, dx : \varphi \in \left[C^1_c(\Omega)\right]^2, \|\varphi\|_\infty \leq 1 \right\}$$

is the variation of u. Moreover, $u \in BV(\Omega)$, the space of bounded variation functions [1, 19] if and only if $V(u,\Omega) < \infty$. In this case, we denote $|Du|(\Omega) = V(u,\Omega)$ the total variation of the finite Radon measure Du, the derivative of u in the distributional sense. The space $BV(\Omega)$ endowed with the norm $\|u\|_{BV(\Omega)} := \|u\|_{L_1(\Omega)} + |Du|(\Omega)$ is a Banach space. The minimization of the total variation is a well-understood regularization for preserving edges of images. Rudin, Osher, and Fatemi [32] proposed the minimization of functionals with total variation constraints as a regularization technique for image denoising. From this pioneering work, total variation minimization became a standard tool in image processing, also for more sophisticated problems, such as deblurring, superresolution, inpainting etc. [2, 8, 9, 17, 33]. We also refer to [10] for an extensive introduction to the use of total variation in imaging.

Our reason for expecting that the preconditioning effects observed by Vonesch and Unser [34] for Haar wavelet-based regularization will take place also in total variation regularization of deblurring problems stems from the well-known near characterization of BV in terms of wavelets [12, 13]: the BV-norm of a bivariate function u is in fact nearly equivalent to the ℓ_1-norm of its bivariate Haar wavelet coefficients u_Λ. More precisely, there exist constants $c_1, c_2 \in \mathbb{R}^+$ such that

$$c_1\|u_\Lambda\|_{\ell_{1+\delta}} \leq \|u\|_{L_1(\Omega)} + |Du|(\Omega) \leq c_2\|u_\Lambda\|_{\ell_1}, \text{ for all } u \in BV(\Omega), \tag{1.6}$$

and for all $\delta > 0$. Actually these inequalities result in embeddings of BV with respect to suitable Besov spaces:

$$B^1_{1,1} \subset BV \subset B^{1,w}_{1,1}.$$

We refer the interested reader to [13] for more details.

Because of this observation and the above mentioned preconditioning mechanism for a deblurring operator in connection with a wavelet space decomposition, we are interested in the minimization of the functional

$$\mathcal{J}(u) = \|Tu - g\|^2_{L_2(\Omega)} + 2\alpha |Du|(\Omega), \tag{1.7}$$

by using a suitably adapted wavelet-based multilevel algorithm.

1.1. Our approach. Domain decomposition and subspace correction methods for functionals of the form (1.7) were already proposed in [23, 24]. There some of the authors of this paper mainly focused on the splitting of the physical domain Ω into smaller subdomains $\Omega = \bigcup_i \Omega_i$ and studied an alternating minimization algorithm on each subspace. Nevertheless, the validity of the algorithm proposed in [24] is not restricted in principle to orthogonal decompositions of the space resulting from splittings of the physical domain Ω, but can also be applied to more abstract orthogonal decompositions of the function space, e.g., a wavelet space decomposition as we have it in mind here.

Let φ be a scaling function generating a multiresolution analysis $(V_i)_{i\in\mathbb{Z}}$ and ψ a corresponding wavelet function. Then we obtain

$$L_2(\Omega) = \overline{\bigcup_{i\in\mathbb{Z}} V_i} = \overline{V_i \oplus \bigoplus_{j=i}^{\infty} W_j} = \overline{\bigoplus_{j\in\mathbb{Z}} W_j},$$

where W_j is the wavelet space corresponding to the j-th level generated by the basis

$$\{\psi_\lambda : \lambda \in \Lambda_j\},$$

and Λ_j denotes the set of indices for the j-th level, see [11, 14] for more details. Moreover, W_j is the orthogonal complement of V_j in V_{j+1}, i.e., we have

$$V_{j+1} = V_j \oplus W_j. \tag{1.8}$$

In particular we may decompose $L_2(\Omega)$ in the following way

$$L_2(\Omega) = V_0 \oplus V_0^{\perp} = V_0 \oplus \overline{\left(\bigoplus_{j=0}^{\infty} W_j\right)}$$

and denote $\mathcal{V}_1 := V_0$ and $\mathcal{V}_2 := V_0^{\perp} = \overline{\bigoplus_{j=0}^{\infty} W_j}$. Associated with this wavelet decomposition into two subspaces the minimization of (1.7) can be carried out by the alternating subspace correction method proposed in [24], which reads as follows: pick an initial $\mathcal{V}_1 \oplus \mathcal{V}_2 \ni u_1^{(0)} + u_2^{(0)} := u^{(0)}$, for example $u^{(0)} = 0$, and iterate

$$\begin{cases} u_1^{(n+1)} \approx \arg\min_{u_1 \in \mathcal{V}_1} \mathcal{J}(u_1 + u_2^{(n)}) \\ u_2^{(n+1)} \approx \arg\min_{u_2 \in \mathcal{V}_2} \mathcal{J}(u_1^{(n+1)} + u_2) \\ u^{(n+1)} := u_1^{(n+1)} + u_2^{(n+1)}. \end{cases} \tag{1.9}$$

In [24] an implementation of this algorithm was suggested, which guaranteed to decrease the objective energy $\mathcal{J}$ monotonically. Convergence to minimizers of $\mathcal{J}$ could be proven only under technical conditions, which are in general not fulfilled, as also illustrated by numerical examples in [24].

In this paper we show additional properties of the limit of the sequence produced by the algorithm in (1.9) and obtain an additional condition under which the obtained limit is indeed the expected minimizer. Nevertheless, this condition cannot be ensured to hold always for any operator T. In particular we are able to construct a counterexample, which shows that in general we cannot expect convergence of the algorithm in (1.9) to a minimizer of $\mathcal{J}$, even for the simplest case of the identity operator $T = I$. Despite this quite special negative result, we show in this paper that an orthogonal wavelet space decomposition for deblurring problems works in practice very efficiently, as already observed by Vonesch and Unser in their study related to ℓ_1-regularization [34]. In particular, with the help of the newly obtained condition of convergence, we are able to show in our numerical examples that the sequence produced by this algorithm in fact numerically converge to a minimizer of $\mathcal{J}$.

Throughout the paper we eventually work on a finite dimensional space by considering a finite regular mesh as a discretization of Ω. Hence we consider instead of the continuous functional (1.7) its discrete approximation, for ease again denoted by $\mathcal{J}$ in (3.1). Note that the discrete approximation (3.1) Γ-converges to the continuous functional (1.7) (see [4]), and has the same singular nature as the continuous problem. For simplicity we will limit ourselves to decompose our problem only into two orthogonal subspaces $\mathcal{V}_1$ and $\mathcal{V}_2$, which is by no means a restriction, as a generalization to a multiple decomposition is straightforward, see [24, Remark 5.3]. However, we stress also that in our numerical experiments the beneficial effect of preconditioning seems not to improve significantly by considering multiple decompositions, see Section 6.

The paper is organized as follows. The main notations used throughout the paper are given in Section 2. In Section 3 we describe the algorithm in (1.9), specified to a wavelet space decomposition. The convergence of the algorithm is investigated in Section 4, where we show properties of the limit of the sequence produced by the algorithm. Additionally we construct a counterexample to show that convergence cannot be obtained in general. Section 5 contains the proof of the main results. In Section 6, we show numerical examples for total variation deblurring which illustrate our findings.

2. Notations. Since we are mainly interested in image deblurring problems, it is sufficient to us to introduce our main notations for a discretization in $[0,1]^2$ only. We assume now that Ω is a 2-dimensional mesh in $[0,1]^2$ of size $N_1 \times N_2$, where $N_1, N_2 \in \mathbb{N}$. The considered function space is $\mathcal{H} = \mathbb{R}^{N_1 \times N_2}$, with corresponding norm

$$\|u\|_{\mathcal{H}} = \|u\|_2 = \left(\sum_{x \in \Omega} |u(x)|^2 \right)^{1/2}.$$

Then the discrete gradient ∇u is the vector of the finite differences on the mesh, given by

$$(\nabla u)(x) = ((\nabla u)^1(x), (\nabla u)^2(x))$$

where

$$(\nabla u)^1(x_{i,j}) = \begin{cases} u(x_{i+1,j}) - u(x_{i,j}) & \text{if } i < N_1 \\ 0 & \text{if } i = N_1, \end{cases}$$

and

$$(\nabla u)^2(x_{i,j}) = \begin{cases} u(x_{i,j+1}) - u(x_{i,j}) & \text{if } j < N_2 \\ 0 & \text{if } j = N_2, \end{cases}$$

for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$. Then the discrete *total variation* of u is defined by

$$|\nabla u|(\Omega) := \sum_{x \in \Omega} |\nabla u(x)|.$$

where $|y| = \sqrt{y_1^2 + y_2^2}$ for every $y = (y_1, y_2) \in \mathbb{R}^2$.

For an operator Q we denote by Q^* its adjoint. Further we introduce the *discrete divergence* $\operatorname{div} : \mathcal{H}^2 \to \mathcal{H}$ defined, in analogy with the continuous setting, by $\operatorname{div} = -\nabla^*$ (∇^* is the adjoint of the discrete gradient ∇). The discrete divergence operator is explicitly given by

$$(\operatorname{div} p)(x_{i,j}) = \begin{cases} p^1(x_{i,j}) - p^1(x^1_{i-1,j}) & \text{if } 1 < i < N_1 \\ p^1(x_{i,j}) & \text{if } i_1 = 1 \\ -p^1(x_{i-1,j}) & \text{if } i_1 = N_1 \end{cases} + \begin{cases} p^2(x_{i,j}) - p^2(x_{i,j-1}) & \text{if } 1 < j < N_2 \\ p^2(x_{i,j}) & \text{if } j = 1 \\ -p^2(x_{i,j-1}) & \text{if } j = N_2, \end{cases}$$

for every $p = (p^1, p^2) \in \mathcal{H}^2$. Further we define the closed convex set

$$K := \left\{ \operatorname{div} p : p \in \mathcal{H}^2, |p(x)| \leq 1 \text{ for all } x \in \Omega \right\},$$

where $|p(x)| = \sqrt{(p^1(x))^2 + (p^2(x))^2}$, and denote $P_K(u) = \arg\min_{v \in K} \|u - v\|_2$ the *orthogonal projection onto* K. We will also denote by $\langle \cdot, \cdot \rangle_{\mathbb{R}^2}$ the scalar product in $\mathbb{R}^2$.

3. Description of the Algorithm.

3.1. Preconditioning. We are interested in solving by the multilevel algorithm in (1.9) the minimization of the discrete functional $\mathcal{J} : \mathcal{H} \to \mathbb{R}$ defined by

$$\mathcal{J}(u) = \|Tu - g\|_2^2 + 2\alpha |\nabla u|(\Omega), \tag{3.1}$$

where $T : \mathcal{H} \to \mathcal{H}$ is a blur operator with kernel κ, $g \in \mathcal{H}$ is a given datum, and $\alpha > 0$ is a fixed regularization parameter. Furthermore, it is convenient that we assume $\|T\| < 1$, which is not a restriction, as a proper rescaling of the problem yields the desired setting, and does not change the minimization problem. In order to guarantee the existence of minimizers for (3.1) we assume that $\mathcal{J}$ is coercive in $\mathcal{H}$, i.e., there exists a constant $C > 0$ such that $\{u \in \mathcal{H} : \mathcal{J}(u) \leq C\}$ is nonempty and bounded in $\mathcal{H}$. It is well known that if $1 \notin \ker(T)$ then this coercivity condition is satisfied, see [33, Proposition 3.1]. In addition, if T is injective, for instance, if κ is a Gaussian or an *averaging* convolution kernel (see Section 6), then (3.1) has unique minimizer.

We can identify $\mathcal{H}$ with the sequences of samples $(u(x))_{x\in\Omega}$ of a function u on $[0,1]^2$, and with V_1, the first scaling space of a multiresolution analysis, by means of the map $(u(x))_{x\in\Omega} \to \sum_{\lambda\in\Lambda_1} u(x_\lambda)\varphi_{1,\lambda}$, where $\varphi_{1,\lambda}$ is a properly dilated scaling function, and $(x_\lambda)_{\lambda\in\Lambda}$ is a suitable rearrangement of the nodes of the mesh Ω. Moreover, by property (1.8), we have the orthogonal splitting $\mathcal{H} = V_1 = V_0 \oplus W_0$. Of course, we may obtain further levels of decomposition

$$\mathcal{H} = V_j \oplus \left(\bigoplus_{i=j}^{0} W_i \right) \quad j \in \mathbb{Z}^-.$$

For simplicity we restrict ourselves to a decomposition into two subspaces $\mathcal{V}_1 := V_0$ and $\mathcal{V}_2 := W_0$ only. We define

$$\pi_{\mathcal{V}_i} : \mathcal{H} \to \mathcal{V}_i,$$

the orthogonal projection onto $\mathcal{V}_i$, for $i = 1, 2$. Then every $u \in \mathcal{H}$ has a unique representation $u = \pi_{\mathcal{V}_1}(u) + \pi_{\mathcal{V}_2}(u)$. In the sequel we denote $u_i = \pi_{\mathcal{V}_i}(u)$, for $i = 1, 2$. Moreover we introduce *surrogate functionals* on $\mathcal{V}_1 \oplus \mathcal{V}_2$ for $a \in \mathcal{V}_i$ and for $i = 1, 2$ by

$$\mathcal{J}_i(u_1, u_2; a) = \mathcal{J}(u_1 + u_2) + \alpha_i \|u_i - a\|_2^2 - \|T(u_i - a)\|_2^2, \tag{3.2}$$

where α_1, α_2 are positive constants chosen as specified below in order to ensure convergence of the subminimization iteration

$$u_i^{(n+1,\ell+1)} = \arg\min_{u_i \in \mathcal{V}_i} \mathcal{J}_i(u_1, u_2; u_i^{(n+1,\ell)}), \quad \ell > 0, \tag{3.3}$$

to a minimizer of the corresponding subproblem of (1.9), i.e.,

$$\arg\min_{u_i\in\mathcal{V}_i} \mathcal{J}(u_1+u_2),$$

for $i=1,2$. Let us further define the synthesis operators $S_1:\ell_2\to\mathcal{V}_1$ via the orthonormal basis for $\mathcal{V}_1$ and $S_2:\ell_2\to\mathcal{V}_2$ via the orthonormal basis for $\mathcal{V}_2$. That is $u_1=S_1(u_{\Lambda_1})$ and $u_2=S_2(u_{\Lambda_2})$ for $u_{\Lambda_1}=(u_\lambda)_{\lambda\in\Lambda_1}$ the scaling function coefficients and $u_{\Lambda_2}=(u_\lambda)_{\lambda\in\Lambda_2}$ the wavelet coefficients. Since S_1,S_2 are isometries, we know that

$$\|T_{\mathcal{V}_i}(u_i-a)\|_2^2=\|T_{\mathcal{V}_i}S_i(u_{\Lambda_i}-a_{\Lambda_i})\|_2^2 \quad\text{and}\quad \|u_i-a\|_2^2=\|u_{\Lambda_i}-a_{\Lambda_i}\|_{\ell_2}^2,$$

where $a=S_1(a_{\Lambda_1})$ or $a=S_2(a_{\Lambda_2})$ and $T_{\mathcal{V}_i}$ denotes the operator T restricted to the subspace $\mathcal{V}_i$, for $i=1,2$. Because of these observations it makes sense to choose

$$1\ge\alpha_i>\|T_{\mathcal{V}_i}S_i\|^2 \tag{3.4}$$

for $i=1,2$. Then we obtain

$$\|T_{\mathcal{V}_i}(u^i-a)\|_2^2=\|T_{\mathcal{V}_i}S_i(u_{\Lambda_i}-a_{\Lambda_i})\|^2\le\|T_{\mathcal{V}_i}S_i\|^2\|u_{\Lambda_i}-a_{\Lambda_i}\|_{\ell_2}^2<\alpha_i\|u_i-a\|_2^2.$$

Notice that with constants α_i as in (3.4), we have for $n=1,2,\dots,$

$$\begin{aligned}\mathcal{J}(u^{(n)})&\le\mathcal{J}_2(u_1^{(n,L)},u_2^{(n,M)};u_2^{(n-1,M)})\le\mathcal{J}(u_1^{(n,L)}+u_2^{(n-1,M)})\\&\le\mathcal{J}_1(u_1^{(n,L)},u_2^{(n-1,M)};u_1^{(n-1,L)})\le\mathcal{J}(u^{(n-1)}).\end{aligned}\tag{3.5}$$

3.2. An alternating minimization. A simple calculation shows that $\mathcal{J}_i$ can be written in the following form:

$$\begin{aligned}\mathcal{J}_i(u_i,u_{\hat{i}};a)&=\|T(u_i+u_{\hat{i}})-g\|_2^2+2\alpha|\nabla(u_i+u_{\hat{i}})|(\Omega)+\alpha_i\|u_i-a\|_2^2-\|T(u_i-a)\|_2^2\\&=\alpha_i\|u_i-z\|_2^2+2\alpha|\nabla(u_i+u_{\hat{i}})|(\Omega)+\phi(a,g,u_{\hat{i}}),\end{aligned}$$

where

$$z_i=\pi_{\mathcal{V}_i}a+\frac{1}{\alpha_i}\pi_{\mathcal{V}_i}(T^*(g-T(u_{\hat{i}}+a)))$$

and ϕ is a function depending only on $a,g,u_{\hat{i}}$, and $\hat{i}\in\{1,2\}\setminus\{i\}$. Hence,

$$\arg\min_{u_1\in\mathcal{V}_1}\mathcal{J}_1(u_1,u_2;a)=\arg\min_{u_1\in\mathcal{V}_1}\|u_1-z_1\|_2^2+2\beta_1|\nabla(u_1+u_2)|(\Omega) \tag{3.6}$$

$$\arg\min_{u_2\in\mathcal{V}_2}\mathcal{J}_2(u_1,u_2;a)=\arg\min_{u_2\in\mathcal{V}_2}\|u_2-z_2\|_2^2+2\beta_2|\nabla(u_1+u_2)|(\Omega) \tag{3.7}$$

where $\beta_i=\alpha/\alpha_i$, for $i=1,2$.

In order to address the subminimization problems (3.6) and (3.7) we have to solve a constrained optimization problem of the type

$$\arg\min_{\Pi u=0}\mathcal{J}(u),$$

where Π is a linear bounded operator, specifically an orthogonal projection. More precisely, we have to solve, respectively,

$$\arg\min_{\pi_{\mathcal{V}_2}u_1=0}\mathcal{J}_1(u_1,u_2;a)\quad\text{and}\quad\arg\min_{\pi_{\mathcal{V}_1}u_2=0}\mathcal{J}_1(u_1,u_2;a).$$

There exist a variety of methods that solve this type of constrained minimization problems, as the Augmented Lagrangian Method [26], and its adaptations known under the name of the Bregman iterations [6, 7, 25, 28, 30, 31, 35, 36, 37]. Here, for simplicity, we use the *Iterative Oblique*

Thresholding algorithm as proposed in the work [24]. Before stating the theorem which recalls the main idea of this algorithm, we introduce the notion of a subdifferential.

Definition 3.1. *For a convex function $F : \mathcal{H} \to \mathbb{R}$, we define the* subdifferential *of F at $u \in \mathcal{H}$, as the set valued function*

$$\partial F(u) := \{u^* \in \mathcal{H} : \langle u^*, v - u\rangle + F(u) \leq F(v) \quad \forall v \in \mathcal{H}\}.$$

It is obvious from this definition that $0 \in \partial F(u)$ if and only if u is a minimizer of F. We focus, for instance, on the minimization on $\mathcal{V}_1$, and similar statements hold symmetrically for the minimization on $\mathcal{V}_2$.

Theorem 3.2. *(Oblique Thresholding, [24]) For $u_2 \in \mathcal{V}_2$ and for $z_1 \in \mathcal{V}_1$ the following statements are equivalent:*

(i) $u_1^* = \arg\min_{u_1 \in \mathcal{V}_1} \|u_1 - z_1\|_2^2 + 2\beta_1 |\nabla(u_1 + u_2)|(\Omega)$,

(ii) there exists $\eta_1 \in Range(\pi_{\mathcal{V}_2})^* \simeq \mathcal{V}_2$ such that $0 \in u_1^* - (z - \eta_1) + \beta_1 \partial|\nabla(u_1^* + u_2)|(\Omega)$.

(iii) there exists $\eta_1 \in \mathcal{V}_2$ such that $u_1^* = (I - P_{\beta_1 K})(z + u_2 - \eta_1) - u_2 \in \mathcal{V}_1$,

(iv) there exists $\eta_1 \in \mathcal{V}_2$ such that $\eta_1 = \pi_{\mathcal{V}_2} P_{\beta_1 K}(\eta_1 - (z + u_2))$.

The existence of $\eta_1 \in \mathcal{V}_2$ as in the previous theorem is shown in [24, Proposition 4.6]. Moreover, the iteration (3.3) for $i = 1$ can be explicitely rewritten as

$$u_1^{(\ell+1)} = (I - P_{\beta_1 K})\left(u_1^{(\ell)} + \frac{1}{\alpha_1}\pi_{\mathcal{V}_1} T^*(g - Tu_2 - Tu_1^{(\ell)}) + u_2 - \eta_1^{(\ell)}\right) - u_2,$$

where $\eta_1^{(\ell)} \in V_2$ is any solution of the fixed point iteration

$$\eta_1 = \pi_{\mathcal{V}_2} P_{\beta_1 K}\left(\eta_1 - (u_1^{(\ell)} + \frac{1}{\alpha_1}\pi_{\mathcal{V}_1} T^*(g - Tu_2 - Tu_1^{(\ell)}) + u_2)\right).$$

The computation of $\eta_1^{(\ell)}$ can be in fact implemented as the limit of the following fixed point algorithm

$$\eta_1^{(0,\ell)} \in V_2, \quad \eta_1^{(m+1,\ell)} = \pi_{\mathcal{V}_2} P_{\beta_1 K}\left(\eta_1^{(m,\ell)} - (u_1^{(\ell)} + \frac{1}{\alpha_1}\pi_{\mathcal{V}_1} T^*(g - Tu_2 - Tu_1^{(\ell)}) + u_2)\right), \quad m \geq 0.$$

For the subspace $\mathcal{V}_2$ one can formulate analogous statements just by adjusting the notations accordingly.

Let us return to our sequential algorithm in (1.9) and express it explicitly as follows: pick an initial $\mathcal{V}_1 \oplus \mathcal{V}_2 \ni u_1^{(0,L)} + u_2^{(0,M)} := u^{(0)}$ and iterate for $n = 0, 1, 2, \ldots$,

$$\begin{cases} \begin{cases} u_1^{(n+1,0)} = u_1^{(n,L)} \\ u_1^{(n+1,\ell+1)} = \arg\min_{u_1 \in \mathcal{V}_1} \mathcal{J}_1(u_1, u_2^{(n,M)}; u_1^{(n+1,\ell)}) \quad \ell = 0, \ldots, L-1 \end{cases} \\ \begin{cases} u_2^{(n+1,0)} = u_2^{(n,M)} \\ u_2^{(n+1,m+1)} = \arg\min_{u_2 \in \mathcal{V}_2} \mathcal{J}_2(u_1^{(n+1,L)}, u_2; u_2^{(n+1,m)}) \quad m = 0, \ldots, M-1 \end{cases} \\ u^{(n+1)} := u_1^{(n+1,L)} + u_2^{(n+1,M)}. \end{cases} \tag{3.8}$$

Note that we prescribe a finite number $L, M \in \mathbb{N}$ of inner iterations for each subspace respectively. Then from (3.8) we obtain sequences $(u_1^{(n,L)})_n$, $(u_2^{(n,M)})_n$ and $(z_1^{(n,L)})_n$, $(z_2^{(n,M)})_n$ such that

$$z_1^{(n+1,L)} = u_1^{(n,L)} + \frac{1}{\alpha_1}\pi_{\mathcal{V}_1}(T^*(g - T(u_1^{(n,L)} + u_2^{(n,M)}))) \tag{3.9}$$

$$z_2^{(n+1,M)} = u_2^{(n,M)} + \frac{1}{\alpha_2}\pi_{\mathcal{V}_2}(T^*(g - T(u_1^{(n+1,L)} + u_2^{(n,M)}))). \tag{3.10}$$

Note that

$$u_1^{(n+1,L)} = \arg\min_{u \in \mathcal{V}_1} \|u - z_1^{(n+1,L)}\|_2^2 + 2\beta_1|\nabla(u + u_2^n)|(\Omega)$$

and

$$u_2^{(n+1,M)} = \arg\min_{u \in \mathcal{V}_2} \|u - z_2^{(n+1,M)}\|_2^2 + 2\beta_2|\nabla(u_1^{(n+1,L)} + u)|(\Omega).$$

4. Main Result. We do not pursue the analysis of the convergence of the algorithm in (3.8), as its proof is exactly the same as in [24, Theorem 5.1]. We would like to investigate instead further equivalent conditions for the limits of the sequences produced by this algorithm to be minimizers of $\mathcal{J}$.

THEOREM 4.1. *We collect properties of minimizers of $\mathcal{J}$ and limits of the algorithm in* (3.8) *in the following statements.*

a) Let $\zeta, u \in \mathcal{H}$. Then $\zeta \in \partial\mathcal{J}(u)$ if and only if there exists $(\xi_0, \xi) \in \mathcal{H} \times \mathcal{H}^2$ such that
1. $\|\xi\|_\infty \leq \alpha$,
2. $\langle \xi(x), \nabla u(x)\rangle_{\mathbb{R}^2} + \alpha|\nabla u(x)| = 0$ *for all* $x \in \Omega$,
3. $T^*\xi_0 - \operatorname{div}(2\xi) + \zeta = 0$,
4. $-\xi_0 = 2(Tu - g)$.

In particular u is a minimizers if and only if the conditions 1.-4. hold for $\zeta = 0$.

b) Let $(u^{(n)})_n$ be a sequence produced by (3.8)*. Then for a strongly convergent subsequence of $(u^{(n)} = u_1^{(n,L)} + u_2^{(n,M)})_n$ with limit $u^{(\infty)} = u_1^{(\infty)} + u_2^{(\infty)}$, we have*

$$u_1^{(\infty)} = \arg\min_{u \in \mathcal{V}_1} \|u - z_1^{(\infty)}\|_2^2 + 2\beta_1|\nabla(u + u_2^{(\infty)})|(\Omega), \tag{4.1}$$

$$u_2^{(\infty)} = \arg\min_{u \in \mathcal{V}_2} \|u - z_2^{(\infty)}\|_2^2 + 2\beta_2|\nabla(u_1^{(\infty)} + u)|(\Omega), \tag{4.2}$$

$$z_1^{(\infty)} = u_1^{(\infty)} + \frac{1}{\alpha_1}\pi_{\mathcal{V}_1}(T^*(g - Tu^{(\infty)})), \tag{4.3}$$

$$z_2^{(\infty)} = u_2^{(\infty)} + \frac{1}{\alpha_2}\pi_{\mathcal{V}_2}(T^*(g - Tu^{(\infty)})), \tag{4.4}$$

where $\beta_i = \alpha/\alpha_i$, for $i = 1, 2$. Moreover let us denote $z^{(\infty)} = u^{(\infty)} + T^*(g - Tu^{(\infty)})$. Then, $u^{(\infty)}$ is a minimizer of (3.1) *if and only if*

$$u^{(\infty)} = \arg\min_{u \in \mathcal{H}}\{\mathcal{F}(u) := \|u - z^{(\infty)}\|_2^2 + 2\alpha|\nabla u|(\Omega)\}. \tag{4.5}$$

Before proving the previous statements we add some comments on the possibility of verification of the minimality condition (4.5). Let $F(u_1, u_2) = \mathcal{F}(u_1 + u_2)$ for $u_1 \in \mathcal{V}_1$ and $u_2 \in \mathcal{V}_2$. Then (4.1) and (4.2) imply

$$F(u_1^{(\infty)}, u_2^{(\infty)}) \leq \arg\min_{\substack{v_1 \in \mathcal{V}_1 \\ v_2 \in \mathcal{V}_2}} \Big\{F(v_1, u_2^{(\infty)}), F(u_1^{(\infty)}, v_2)\Big\}. \tag{4.6}$$

Unfortunately, (4.6) may not imply that $u^{(\infty)} = u_1^{(\infty)} + u_2^{(\infty)}$ is a minimizer of (4.5) and eventually of (3.1). We propose the following univariate counterexample, which also shows that the algorithm in (3.8) may fail to converge to a minimizing solution. For simplicity, we return to the continuous setting and we assume that Ω is the interval $[-1, 2]$, and $g = \chi_{[0,1/2)}$. We consider univariate Haar wavelets, i.e., let $\varphi_0 = \chi_{[0,1)}$ and $\psi_0 = \chi_{[0,1/2)} - \chi_{[1/2,1)}$. Then we have

$$g = \frac{1}{2}\varphi_0 + \frac{1}{2}\psi_0.$$

We can prove the following proposition.

PROPOSITION 4.2. *Let $0 < \alpha < 1/8$ and $\mathcal{V}_1$ be the subspace of $L_2([-1, 2])$ generated by $\{\varphi_0(x - k) : k \in \{-1, 0, 1\}\}$ and $\mathcal{V}_2$ be the subspace of $L_2([-1, 2])$ generated by $\{\psi_{j,k}(x) = 2^{j/2}\psi_0(2^j x - k) : j \in \mathbb{Z}_+ \cup \{0\}, k \in \{-2^j, \ldots, 2^j\}\}$, then*

$$u_1^{(\infty)} = \frac{1 - 4\alpha}{2}\varphi_0, \quad u_2^{(\infty)} = \frac{1 - 4\alpha}{2}\psi_0,$$

which satisfy

$$\arg\min_{\substack{u_1\in\mathcal{V}_1\\ u_2\in\mathcal{V}_2}} F(u_1,u_2) < F(u_1^{(\infty)},u_2^{(\infty)}) \le \arg\min_{\substack{v_1\in\mathcal{V}_1\\ v_2\in\mathcal{V}_2}} \left\{F(v_1,u_2^{(\infty)}), F(u_1^{(\infty)},v_2)\right\} \tag{4.7}$$

where

$$\begin{aligned} F(u_1,u_2) &= \mathcal{F}(u_1+u_2) = \|u_1+u_2-g\|_2^2 + 2\alpha|\nabla(u_1+u_2)|([-1,2]) \\ &= \left\|u_1 - \frac{1}{2}\varphi_0\right\|_2^2 + \left\|u_2 - \frac{1}{2}\psi_0\right\|_2^2 + 2\alpha|\nabla(u_1+u_2)|([-1,2]). \end{aligned}$$

Proof. We prove the result by showing that the algorithm in (3.8), starting with $u^{(0)}=0$, stops by converging to $u^{(\infty)} = u_1^{(\infty)} + u_2^{(\infty)}$ in finite iterations, and that (4.7) holds. Let $u_1^{(0)} = u_2^{(0)} = 0$. Then

$$u_1^{(1)} = \arg\min_{u\in\mathcal{V}_1} \left\|u - \frac{1}{2}\varphi_0\right\|_2^2 + 2\alpha|\nabla u|([-1,2]). \tag{4.8}$$

Then $u_1^{(1)} = a\varphi_0$ for some $a>0$ and

$$\begin{aligned} \left\|u - \frac{1}{2}\varphi_0\right\|_2^2 + 2\alpha|\nabla u|([-1,2]) &= \left\|a\varphi_0 - \frac{1}{2}\varphi_0\right\|_2^2 + 2\alpha a|\nabla\varphi_0|([-1,2]) \\ &= \left(a-\frac{1}{2}\right)^2 + 4\alpha a = \left(a + \frac{4\alpha-1}{2}\right)^2 + 2\alpha - 4\alpha^2. \end{aligned}$$

Since $\alpha < 1/8$, (4.8) attains its minimum when

$$a = \frac{1-4\alpha}{2}, \text{ i.e., } u_1^{(1)} = \frac{1-4\alpha}{2}\varphi_0.$$

Now, we solve

$$u_2^{(1)} = \arg\min_{u\in\mathcal{V}_2} \left\|u - \frac{1}{2}\psi_0\right\|_2^2 + 2\alpha|\nabla(u_1^{(1)}+u)|([-1,2]). \tag{4.9}$$

It is not hard to see that $u_2^{(1)} = b\psi_0$ for some $b>0$. If we assume $b \le \frac{1-4\alpha}{2}$, then

$$\begin{aligned} \left\|u - \frac{1}{2}\psi_0\right\|_2^2 + 2\alpha|\nabla(u_1^{(1)}+u)|([-1,2]) &= \left(b-\frac{1}{2}\right)^2 + 2\alpha\Big(\frac{1-4\alpha}{2} + b + 2b + \frac{1-4\alpha}{2} - b\Big) \\ &= \left(b + \frac{4\alpha-1}{2}\right)^2 + 4\alpha - 12\alpha^2 \ge 4\alpha - 12\alpha^2, \end{aligned}$$

which is minimized when $b = \frac{1-4\alpha}{2}$. On the other hand, if we assume $b \ge \frac{1-4\alpha}{2}$, then since $0 < \frac{1-8\alpha}{2} < \frac{1-4\alpha}{2} \le b$,

$$\begin{aligned} \left\|u - \frac{1}{2}\psi_0\right\|_2^2 + 2\alpha|\nabla(u_1^{(1)}+u)|([-1,2]) &= \left(b-\frac{1}{2}\right)^2 + 2\alpha\Big(\frac{1-4\alpha}{2} + b + 2b - \frac{1-4\alpha}{2} + b\Big) \\ &= \left(b + \frac{8\alpha-1}{2}\right)^2 + 4\alpha - 16\alpha^2 \ge 4\alpha - 12\alpha^2, \end{aligned}$$

which is also minimized when $b = \frac{1-4\alpha}{2}$. Hence

$$u_2^{(1)} = \frac{1-4\alpha}{2}\psi_0.$$

Now, we solve

$$u_1^{(2)} = \arg\min_{u\in\mathcal{V}_1} \left\|u - \frac{1}{2}\varphi_0\right\|_2^2 + 2\alpha|\nabla(u+u_2^{(1)})|([-1,2]).$$

It is easy to see that $u_1^{(2)} = a\varphi_0$ for some $a > 0$. If we assume $a \le \frac{1-4\alpha}{2}$, then since $\frac{1-4\alpha}{2} \le \frac{1}{2}$,

$$\begin{aligned}\Big\|u - \frac{1}{2}\varphi_0\Big\|_2^2 + 2\alpha|\nabla(u + u_2^{(1)})|([-1,2]) &= \Big(a - \frac{1}{2}\Big)^2 + 2\alpha\Big(a + \frac{1-4\alpha}{2} + (1-4\alpha) + \frac{1-4\alpha}{2} - a\Big)\\ &= \Big(a - \frac{1}{2}\Big)^2 + 4\alpha(1-4\alpha) \ge 4\alpha - 12\alpha^2,\end{aligned}$$

which is minimized when $a = \frac{1-4\alpha}{2}$. On the other hand, if we assume $a \ge \frac{1-4\alpha}{2}$, then

$$\begin{aligned}\Big\|u - \frac{1}{2}\varphi_0\Big\|_2^2 + 2\alpha|\nabla(u + u_2^{(1)})|([-1,2]) &= \Big(a - \frac{1}{2}\Big)^2 + 2\alpha\Big(a + \frac{1-4\alpha}{2} + (1-4\alpha) + a - \frac{1-4\alpha}{2}\Big)\\ &= \Big(a + \frac{4\alpha-1}{2}\Big)^2 + 4\alpha - 12\alpha^2 \ge 4\alpha - 12\alpha^2,\end{aligned}$$

which is also minimized when $a = \frac{1-4\alpha}{2}$. We finally obtain

$$u_1^{(2)} = \frac{1-4\alpha}{2}\varphi_0 = u_1^{(1)}.$$

Therefore, after only one step of the algorithm in (3.8), we have

$$u_1^{(\infty)} = \frac{1-4\alpha}{2}\varphi_0, \quad u_2^{(\infty)} = \frac{1-4\alpha}{2}\psi_0.$$

It is now easy to see that $u_1^{(\infty)}, u_2^{(\infty)}$ satisfy (4.6) and

$$F(u_1^{(\infty)}, u_2^{(\infty)}) = 4\alpha - 8\alpha^2.$$

However, if $u = a\chi_{[0,1/2)} = \frac{a}{2}\varphi_0 + \frac{a}{2}\psi_0$, then

$$\begin{aligned}\mathcal{F}(u) &= \|u-g\|_2^2 + 2\alpha|\nabla u|([-1,2]) = (a-1)^2\|\chi_{[0,1/2)}\|_2^2 + 2\alpha\cdot 2a\\ &= \frac{1}{4}(a-1)^2 + 4\alpha a = \frac{1}{4}(a + (8\alpha-1))^2 + 4\alpha - 16\alpha^2.\end{aligned}$$

Since $0 < \alpha < 1/8$, if we set $u_0 = (1-8\alpha)\chi_{[0,1/2)} = \frac{1-8\alpha}{2}\varphi_0 + \frac{1-8\alpha}{2}\psi_0$, then

$$\min_{u_1 \in \mathcal{V}_1, u_2 \in \mathcal{V}_2} F(u_1, u_2) \le \mathcal{F}(u_0) = 4\alpha - 16\alpha^2 < 4\alpha - 8\alpha^2 = \mathcal{F}(u_1^{(\infty)} + u_2^{(\infty)}) = F(u_1^{(\infty)}, u_2^{(\infty)}).$$

□

Theorem 4.1-a) also provides us with the following useful characterization.

Corollary 4.3. *The subdifferential of* $\alpha\partial|\nabla u|(\Omega)$ *is fully characterized by*

$$\begin{aligned}\alpha\partial|\nabla u|(\Omega) &= \{\mathrm{div}(\xi) \in \mathcal{H} : \|\xi\|_\infty \le \alpha,\ \langle \xi(x), \nabla u(x)\rangle_{\mathbb{R}^2} + \alpha|\nabla u|(x) = 0 \textit{ for all } x \in \Omega\}\\ &= \{\mathrm{div}(\xi) \in \mathcal{H} : -\mathrm{div}(\xi) = P_{\alpha K}(-u - \mathrm{div}(\xi))\}.\end{aligned}$$

Proof. If we consider $T = I$ in Theorem 4.1-a), then $\tilde\zeta \in \alpha\partial|\nabla u|(\Omega)$ if and only if $\zeta = 2(\tilde\zeta + u - g) \in \partial\mathcal{J}(u)$ if and only if there exists $(\xi_0, \xi) \in \mathcal{H} \times \mathcal{H}^2$ such that

1. $\|\xi\|_\infty \le \alpha$,
2. $\langle \xi(x), \nabla u(x)\rangle_{\mathbb{R}^2} + \alpha|\nabla u(x)| = 0$ for all $x \in \Omega$,
3. $\tilde\zeta = \mathrm{div}(\xi)$.

Hence,

$$\alpha\partial|\nabla u|(\Omega) = \{\mathrm{div}(\xi) \in \mathcal{H} : \|\xi\|_\infty \le \alpha,\ \langle \xi(x), \nabla u(x)\rangle_{\mathbb{R}^2} + \alpha|\nabla u|(x) = 0 \text{ for all } x \in \Omega\}.$$

We also notice that

$$\mathrm{div}(\xi) \in \alpha|\nabla u|(\Omega) \text{ if and only if } 0 \in u - (u + \mathrm{div}(\xi)) + \alpha\partial|\nabla u|(\Omega),$$

which is equivalent to

$$u = \arg\min_v \|v - (u + \operatorname{div}(\xi))\|_2^2 + 2\alpha|\nabla v|(\Omega),$$

that is,

$$-u = \arg\min_v \|v + (u + \operatorname{div}(\xi))\|_2^2 + 2\alpha|\nabla v|(\Omega).$$

By [24, Examples 4.2.2], the latter optimality problem is equivalent to

$$-u = (I - P_{\alpha K})(-u - \operatorname{div}(\xi)),$$

that is,

$$-\operatorname{div}(\xi) = P_{\alpha K}(-u - \operatorname{div}(\xi)).$$

Therefore, we also have

$$\alpha\partial|\nabla u|(\Omega) = \{\operatorname{div}(\xi) \in \mathcal{H} : -\operatorname{div}(\xi) = P_{\alpha K}(-u - \operatorname{div}(\xi))\}.$$

□

5. Proof of Theorem 4.1.

a) The proof of this statement, which characterizes the minimizers of $\mathcal{J}$, can be found in [23, Appendix A].

b) For simplicity, we rename a convergent subsequence again by $(u^{(n)} = u_1^{(n,L)} + u_2^{(n,M)})_n$. Equations (4.3) and (4.4) follow directly from (3.9) for $n \to \infty$. Furthermore, it is also easy to see that for any $u_1 \in \mathcal{V}_1$,

$$\begin{aligned}\|u_1^{(\infty)} - z_1^{(\infty)}\|_2^2 + 2\beta_1|\nabla u^{(\infty)}|(\Omega) &= \lim_{n\to\infty} \|u_1^{(n+1,L)} - z_1^{(n+1,L)}\|_2^2 + 2\beta_1|\nabla(u_1^{(n+1,L)} + u_2^{(n,M)})|(\Omega)\\ &\le \lim_{n\to\infty} \|u_1 - z_1^{(n+1,L)}\|_2^2 + 2\beta_1|\nabla(u_1 + u_2^{(n,M)})|(\Omega)\\ &= \|u_1 - z_1^{(\infty)}\|_2^2 + 2\beta_1|\nabla(u_1 + u_2^{(\infty)})|(\Omega).\end{aligned}$$

The second limit is a consequence of [24, formula (5.7)], which states the asymptotic regularity of the sequence, i.e.,

$$\left(\sum_{\ell=0}^{L-1} \|u_1^{(n+1,\ell+1)} - u_1^{(n+1,\ell)}\|_2^2 + \sum_{m=0}^{M-1} \|u_2^{(n+1,m+1)} - u_2^{(n+1,m)}\|_2^2\right) \to 0, \quad n \to \infty. \tag{5.1}$$

Hence, we have

$$u_1^{(\infty)} = \arg\min_{u\in\mathcal{V}_1} \|u - z_1^{(\infty)}\|_2^2 + 2\beta_1|\nabla(u + u_2^{(\infty)})|(\Omega).$$

With the same argument one obtains (4.2). By Theorem 3.2 the optimality conditions (4.1) and (4.2) are equivalent to

$$\begin{aligned}0 &\in u_1^{(\infty)} - (z_1^{(\infty)} - \eta_1^{(\infty)}) + \beta_1\partial|\nabla u^{(\infty)}|(\Omega),\\ 0 &\in u_2^{(\infty)} - (z_2^{(\infty)} - \eta_2^{(\infty)}) + \beta_2\partial|\nabla u^{(\infty)}|(\Omega).\end{aligned}$$

Then by Corollary 4.3 there exist ξ_1, ξ_2 such that

$$\operatorname{div}(\xi_1) = -u_1^{(\infty)} + (z_1^{(\infty)} - \eta_1^{(\infty)}), \tag{5.2}$$

$$\operatorname{div}(\xi_2) = -u_2^{(\infty)} + (z_2^{(\infty)} - \eta_2^{(\infty)}), \tag{5.3}$$

and with the following additional properties

1. $\|\xi_1\|_\infty \le \beta_1$, $\|\xi_2\|_\infty \le \beta_2$ and
2. $\langle \xi_i(x), \nabla u^{(\infty)}(x)\rangle_{\mathbb{R}^2} + \beta_i|\nabla u^{(\infty)}(x)| = 0$ for all $x \in \Omega$ and $i = 1, 2$.

Multiplying (5.2) by α_1 and (5.3) by α_2 yields

$$-\alpha_1 u_1^{(\infty)} + \alpha_1 z_1^{(\infty)} - \alpha_1 \eta_1^{(\infty)} - \alpha_1 \operatorname{div}(\xi_1) = 0,$$
$$-\alpha_2 u_2^{(\infty)} + \alpha_2 z_2^{(\infty)} - \alpha_2 \eta_2^{(\infty)} - \alpha_2 \operatorname{div}(\xi_2) = 0.$$

If we sum up the last two equations we obtain

$$-\alpha_1 u_1^{(\infty)} + \alpha_1 z_1^{(\infty)} - \alpha_2 u_2^{(\infty)} + \alpha_2 z_2^{(\infty)} - \operatorname{div}(\alpha_1\xi_1) - \operatorname{div}(\alpha_2\xi_2) - (\alpha_1\eta_1^{(\infty)} + \alpha_2\eta_2^{(\infty)}) = 0 \quad (5.4)$$

From Theorem 3.2 we have that

$$\eta_1^{(\infty)} = \pi_{\mathcal{V}_2} P_{\beta_1 K}(\eta_1^{(\infty)} - (z_1^{(\infty)} + u_2^{(\infty)})) \text{ and } \eta_2^{(\infty)} = \pi_{\mathcal{V}_1} P_{\beta_2 K}(\eta_2^{(\infty)} - (z_2^{(\infty)} + u_1^{(\infty)}))$$

and it follows then from (5.2), (5.3), and Corollary 4.3 that

$$\alpha_1\eta_1^{(\infty)} = \pi_{\mathcal{V}_2}(-\operatorname{div}(\alpha_1\xi_1)) = \pi_{\mathcal{V}_2} P_{\alpha K}(-u^{(\infty)} - \operatorname{div}(\alpha_1\xi_1)), \quad (5.5)$$
$$\alpha_2\eta_2^{(\infty)} = \pi_{\mathcal{V}_1}(-\operatorname{div}(\alpha_2\xi_2)) = \pi_{\mathcal{V}_1} P_{\alpha K}(-u^{(\infty)} - \operatorname{div}(\alpha_2\xi_2)). \quad (5.6)$$

Plugging (5.5) and (5.6) in (5.4) and using the definition of $z_1^{(\infty)}$ and $z_2^{(\infty)}$ yield

$$\begin{aligned} 0 &= -T^*(Tu^{(\infty)} - g) - \operatorname{div}(\alpha_1\xi_1) - \operatorname{div}(\alpha_2\xi_2) + (\pi_{\mathcal{V}_2}\operatorname{div}(\alpha_1\xi_1) + \pi_{\mathcal{V}_1}\operatorname{div}(\alpha_2\xi_2)) \\ &= -T^*(Tu^{(\infty)} - g) - (\pi_{\mathcal{V}_1}\operatorname{div}(\alpha_1\xi_1) + \pi_{\mathcal{V}_2}\operatorname{div}(\alpha_2\xi_2)) \end{aligned}$$

Therefore, if there exists ξ such that $\operatorname{div}(\xi) \in \alpha\partial|\nabla u^{(\infty)}|(\Omega)$ and

$$\operatorname{div}(\xi) = \pi_{\mathcal{V}_1}\operatorname{div}(\alpha_1\xi_1) + \pi_{\mathcal{V}_2}\operatorname{div}(\alpha_2\xi_2), \quad (5.7)$$

then ξ also satisfies

1. $\|\xi\|_\infty \le \alpha$,
2. $\langle \xi(x), \nabla u^{(\infty)}(x)\rangle_{\mathbb{R}^2} + \alpha|\nabla u^{(\infty)}|(x) = 0$ for all $x \in \Omega$,
3. $T^*\xi_0 - \operatorname{div}(2\xi) = 0$,
4. $-\xi_0 = 2(Tu^{(\infty)} - g)$.

The existence of such ξ is a necessary and sufficient condition for $u^{(\infty)}$ to be a minimizer by a). Then ξ satisfies

$$-u^{(\infty)} + z^{(\infty)} = T^*(g - Tu^{(\infty)}) = -\alpha_1 u_1^{(\infty)} + \alpha_1 z_1^{(\infty)} - \alpha_2 u_2^{(\infty)} + \alpha_2 z_2^{(\infty)} = \operatorname{div}(\xi),$$

that is,

$$-z^{(\infty)} = -u^{(\infty)} - \operatorname{div}(\xi) \quad \text{and} \quad \operatorname{div}(\xi) \in \alpha|\nabla u^{(\infty)}|(\Omega),$$

where $z^{(\infty)} := u^{(\infty)} + T^*(g - Tu^{(\infty)})$. Note that for $i = 1, 2$,

$$\pi_{\mathcal{V}_i} z^{(\infty)} = (1 - \alpha_i)u_i^{(\infty)} + \alpha_i z_i^{(\infty)}.$$

By $\operatorname{div}(\xi) \in \alpha\partial|\nabla u^{(\infty)}|(\Omega)$ and Corollary 4.3, this is equivalent to

$$u^{(\infty)} - z^{(\infty)} = -\operatorname{div}(\xi) = P_{\alpha K}(-u^{(\infty)} - \operatorname{div}(\xi)) = P_{\alpha K}(-z^{(\infty)})$$

Hence,

$$-u^{(\infty)} = (I - P_{\alpha K})(-z^{(\infty)}) = \arg\min_u \|u + z^{(\infty)}\|_2^2 + 2\alpha|\nabla u|(\Omega).$$

which proves the theorem.

The proof of Theorem 4.1 provides us with another characterization of $u^{(\infty)}$ being a minimizer of (3.1) by ξ_1, ξ_2 in (5.2), (5.3).

COROLLARY 5.1. *Let $\alpha_1 \leq 1$, $\alpha_2 \leq 1$. The limit $u^{(\infty)}$, obtained in Theorem 4.1 b), is a minimizer of* (3.1) *if and only if there exist ξ_1, ξ_2 in* (5.2), (5.3) *with* $\mathrm{div}(\alpha_1\xi_1) = \mathrm{div}(\alpha_2\xi_2)$.

Proof. First let us prove the statement for $\alpha_1 = \alpha_2 = 1$: If $u^{(\infty)}$ is a minimizer of (3.1), then Theorem 4.1 and [24, Examples 4.2.2] say that

$$u^{(\infty)} = (I - P_{\alpha K})(z^{(\infty)}).$$

Since $\alpha_1 = \alpha_2 = 1$, we obtain

$$z_1^{(\infty)} = \pi_{\mathcal{V}_1} z^{(\infty)}, \quad z_2^{(\infty)} = \pi_{\mathcal{V}_2} z^{(\infty)}.$$

We then can rephrase this in two different ways as follows.

$$\begin{aligned} u_1^{(\infty)} &= (I - P_{\alpha K})(z_1^{(\infty)} + u_2^{(\infty)} - (u_2^{(\infty)} - z_2^{(\infty)})) - u_2^{(\infty)}, \\ \text{or} \quad u_2^{(\infty)} &= (I - P_{\alpha K})(z_2^{(\infty)} + u_1^{(\infty)} - (u_1^{(\infty)} - z_1^{(\infty)})) - u_1^{(\infty)}. \end{aligned}$$

By Theorem 3.2, we can take

$$\eta_1^{(\infty)} = u_2^{(\infty)} - z_2^{(\infty)}, \quad \eta_2^{(\infty)} = u_1^{(\infty)} - z_1^{(\infty)}.$$

This implies $\mathrm{div}(\xi_1) = \mathrm{div}(\xi_2)$ from (5.2) and (5.3). On the other hand, if $\mathrm{div}(\xi_1) = \mathrm{div}(\xi_2)$, then (5.7) implies that $u^{(\infty)}$ is a minimizer of (3.1).

Now let us prove the statement for $\alpha_1, \alpha_2 \leq 1$: Suppose that $u^{(\infty)}$ is a minimizer of (3.1). Then Theorem 4.1 b) says that

$$u^{(\infty)} = (I - P_{\alpha K})(z^{(\infty)}) \text{ if and only if } \mathrm{div}(\xi) = -u^{(\infty)} + z^{(\infty)} \in \alpha\partial|\nabla u^{(\infty)}|(\Omega) \text{ for some } \xi.$$

By the above considerations, we know that there exist $\eta_1^{\infty,1}, \eta_2^{\infty,1}$ such that

$$\eta_1^{\infty,1} = u_2^{(\infty)} - \pi_{\mathcal{V}_2} z^{(\infty)} = \alpha_2 u_2^{(\infty)} - \alpha_2 z_2^{(\infty)}, \quad \eta_2^{\infty,1} = u_1^{(\infty)} - \pi_{\mathcal{V}_1} z^{(\infty)} = \alpha_1 u_1^{(\infty)} - \alpha_1 z_1^{(\infty)}.$$

and

$$-u_1^{(\infty)} + (\pi_{\mathcal{V}_1} z^{(\infty)} - \eta_1^{\infty,1}) = \mathrm{div}(\xi) = -u_2^{(\infty)} + (\pi_{\mathcal{V}_2} z^{(\infty)} - \eta_2^{\infty,1}).$$

Let $\eta_i^{\infty,\alpha_i} = \frac{\eta_i^{\infty,1}}{\alpha_i}$, $\xi_i^{\alpha_i} = \frac{\xi}{\alpha_i}$ for $i = 1, 2$. Then

$$\begin{aligned} \mathrm{div}(\xi_1^{\alpha_1}) &= -u_1^{(\infty)} + (z_1^{(\infty)} - \eta_1^{\infty,\alpha_1}), \\ \mathrm{div}(\xi_2^{\alpha_2}) &= -u_2^{(\infty)} + (z_2^{(\infty)} - \eta_2^{\infty,\alpha_2}). \end{aligned}$$

Moreover one can see that $\mathrm{div}(\xi_i^{\alpha_i}) \in \beta_i\partial|\nabla u^{(\infty)}|(\Omega)$ for $i = 1, 2$. Hence if we let $\xi_1 = \xi_1^{\alpha_1}$ and $\xi_2 = \xi_2^{\alpha_2}$, then $\mathrm{div}(\alpha_1\xi_1) = \mathrm{div}(\xi) = \mathrm{div}(\alpha_2\xi_2)$.

On the other hand, if there exist ξ_1, ξ_2 satisfying $\mathrm{div}(\alpha_1\xi_1) = \mathrm{div}(\alpha_2\xi_2)$ in (5.2), (5.3), then by (5.7), we know that the limit $u^{(\infty)}$ is a minimizer of (3.1). ∎

6. Numerical Validation. In this section we illustrate the performance of the algorithm in (3.8) for the minimization of (3.1) when T is a blur operator with *averaging* kernel κ supported on 3×3 pixels and uniform values $1/9$. The function space is split into $N \in \mathbb{N}$ orthogonal spaces by a wavelet space decomposition such that

$$\mathcal{H} = V_{2-N} \oplus \left(\bigoplus_{j=2-N}^{0} W_j \right)$$

and we set $\mathcal{V}_1 := V_{2-N}$ and $\mathcal{V}_i := W_{2-i}$ for $i = 2, 3, \ldots, N$. Note that for $N = 1$ we have that $\mathcal{V}_1 = \mathcal{H}$ and thus we have no splitting. In order to gain maximal performance, the preconditioner constants are always chosen as

$$\alpha_i = \|T_{\mathcal{V}_i} S_i\|^2, \tag{6.1}$$

for $i = 1, \ldots, N$, as already discussed in detail for $N = 2$ in Section 3.1.

In our numerical examples we only consider decompositions by using Haar wavelets. In this case it is easy to see that the preconditioner constant for the scale space V_{2-N} is simply $\alpha_1 = \|T\|$ and the preconditioner constants for the wavelet spaces W_j, $j = 0, \ldots, 2-N$, are strictly smaller than the norm of T.

The implementation of the algorithm is done as suggested and discussed in [24]. That is the subiterations in (3.8) are solved by computing the minimizers by means of oblique thresholding, cf. Theorem 3.2. For the computation of the orthogonal projection onto $\beta_i K$, $i = 1, \ldots, N$, in the oblique thresholding we use an algorithm proposed by Chambolle in [8].

6.1. Experiments. In our examples we stop the algorithm in (3.8) as soon as the energy $\mathcal{J}$ reaches a significant level, i.e.,

$$\mathcal{J}(u^*) \leq \epsilon \tag{6.2}$$

where u^* denotes the first iterate for which (6.2) is fulfilled and ϵ is an estimate of the minimal energy.

In Figure 6.1 we show an image of size 156×156 pixels, which was corrupted by the blur operator T as above. In order to deblur this image we split the function space of the image into orthogonal subspaces via a wavelet space decomposition and compute its solution by the algorithm with $\alpha = 10^{-5}$ and stopping criterion (6.2) with $\epsilon = 0.04$. The computed result for 4 subspaces is shown in Figure 6.1 on the right hand side.

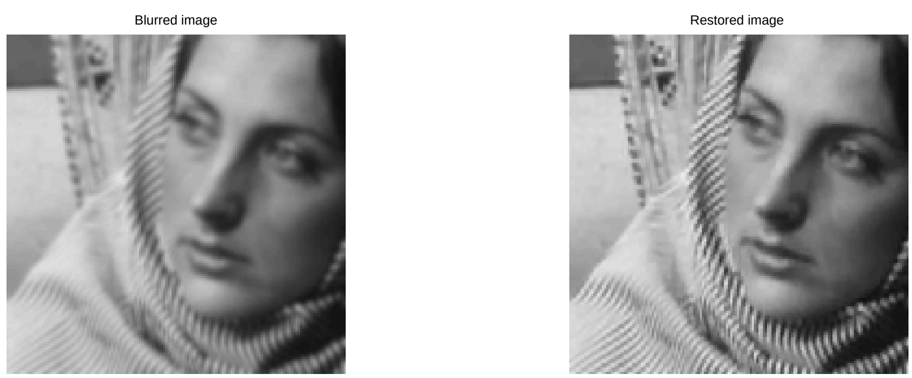

Fig. 6.1. *On the left we depict an image, blurred with an averaging kernel. On the right we show the corresponding solution computed on 4 orthogonal subspaces by the algorithm in* (3.8) *with* $\alpha = 10^{-5}$ *and stopping criterion* (6.2) *with* $\epsilon = 0.04$.

N	1	2	3	4	5	6
Iterations	525	12	8	6	6	6
CPU (s)	40.80	2.59	2.37	2.34	3.29	4.05

Table 6.1

Performance of the wavelet decomposition algorithm in (3.8) *for image deblurring (uniform kernel) with energy-stopping criterion* (6.2) *with* $\epsilon = 0.04$*: the number of iterations and CPU time in seconds are shown with respect to the number* N *of subspace decompositions.*

With the same setting as above we solve this specific deblurring problem with the algorithm in (3.8) for different numbers of subspaces and compare its performance with respect to the needed iterations and computational time in Table 6.1. Note that for $N = 1$ we solve this problem without any decomposition on the whole space $\mathcal{H}$. We see in Table 6.1 that the performance in this case is clearly the worst. When we solve the same problem with a decomposition into two or more wavelet spaces only a very few iterations are needed to reach the stopping criterion. Additionally a decomposition into only 2 subspaces leads to a significant speed-up in computational time, cf. also Figure 6.4.

By using Lemma 5.1 we check for a splitting into 2 orthogonal subspaces whether the sequential algorithm numerically converges to a minimizer by looking at

$$\| \operatorname{div}(\alpha_1 \xi_1^{(n)}) - \operatorname{div}(\alpha_2 \xi_2^{(n)}) \|, \tag{6.3}$$

where

$$\begin{aligned} \operatorname{div}(\xi_1^{(n)}) &= -u_1^{(n)} + (z_1^{(n)} - \eta_1^{(n)}) \\ \operatorname{div}(\xi_2^{(n)}) &= -u_2^{(n)} + (z_2^{(n)} - \eta_2^{(n)}). \end{aligned}$$

In Figure 6.2 we plot the decay of this norm discrepancy, indicator of the distance from convergence to a minimizer, with respect to the iterations n. The indicator numerically converges to zero for n increasing and the algorithm numerically converges to a minimizer of the original problem.

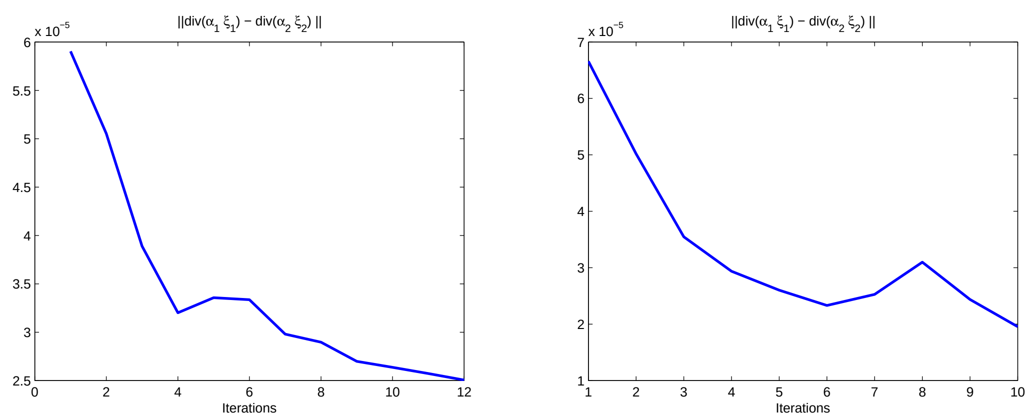

FIG. 6.2. *We plot* $\| \operatorname{div}(\alpha_1 \xi_1^{(n)}) - \operatorname{div}(\alpha_2 \xi_2^{(n)}) \|$ *for the problem of Figure 6.1 (left) and Figure 6.3 (right) in view of Lemma 5.1 in order to check whether the algorithm is indeed converging.*

In Figure 6.3 we depict another example of an image deblurring problem, where the image of size 279×285 pixels was blurred by the averaging kernel from above. The image is again recovered via the algorithm in (3.8) by splitting the function space $\mathcal{H}$ into orthogonal wavelet spaces. We take as the stopping criterion (6.2) with $\epsilon = 0.058$ and as a regularization parameter $\alpha = 10^{-5}$. In Table 6.2 we show the behaviour of the algorithm for different numbers of subspaces.

N	1	2	3	4	5	6
Iterations	405	10	8	7	7	7
CPU (s)	86.94	8.37	11.37	13.75	17.59	22.24

TABLE 6.2

Performance of the wavelet decomposition algorithm in (3.8) *for image deblurring (uniform kernel) with energy-stopping criterion* (6.2) *with* $\epsilon = 0.058$*: the number of iterations and CPU time in seconds are shown with respect to the number* N *of subspace decompositions.*

Again we see from the numerical results that with a decomposition into 2 subspaces the speed of convergence increases dramatically as depicted in Figure 6.4.

Fig. 6.3. *On the left we show an image, blurred by an averaging kernel. On the right we show the corresponding solution computed alternating on 3 orthogonal subspaces by the algorithm in* (3.8) *with* $\alpha = 10^{-5}$ *and stopping criterion* (6.2) *with* $\epsilon = 0.058$.

Let us display in Figure 6.5 also the "distance" between the obtained estimate and the original image. Therefore we recall the definition of *Signal-to-Error-Ratio Gain* [34] given by

$$\mathrm{SERG} = 20 \log_{10} \frac{\|g - org\|}{\|u^* - org\|},$$

where *org* denotes the original image before blurring. In Figure 6.5 we show the evolution of this measure with respect to the time for both mentioned deblurring problems for $N = 1$ (no splitting) and for $N = 2$ (splitting into 2 subspaces).

Acknowledgments. Massimo Fornasier, Yunho Kim, and Andreas Langer acknowledge the financial support provided by the FWF project Y 432-N15 START-Preis *Sparse Approximation and Optimization in High Dimensions*. Carola-B. Schönlieb acknowledges the financial support provided by the DFG Graduiertenkolleg 1023 *Identification in Mathematical Models: Synergy of Stochastic and Numerical Methods*, the Wissenschaftskolleg (Graduiertenkolleg, Ph.D. program) of the Faculty for Mathematics at the University of Vienna (funded by the Austrian Science Fund FWF) and the FFG project no. 813610 *Erarbeitung neuer Algorithmen zum Image Inpainting*. Further, this publication is based on work supported by Award No. KUK-I1-007-43 , made by King Abdullah University of Science and Technology (KAUST). The results of the paper also contribute to the project WWTF Five senses-Call 2006, *Mathematical Methods for Image Analysis and Processing in the Visual Arts*.

REFERENCES

[1] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs. Oxford: Clarendon Press. xviii, 2000.

[2] G. Aubert and L. Vese, *A variational method in image recovery*, SIAM J. Numer. Anal. 34 (1997) pp. 1948–1979.

[3] A. Beck and M. Teboulle, *Fast iterative shrinkage-thresholding algorithm for linear inverse problems*, SIAM J. Imaging Sciences, 2 (2009), pp. 183–202.

[4] A. Braides, Γ-*Convergence for Beginners*, No.22 in Oxford Lecture Series in Mathematics and Its Applications. Oxford University Press, 2002.

[5] K. Bredies and D. Lorenz, *Linear convergence of iterative soft-thresholding*, J. Fourier Anal. Appl. 14 (2008), pp. 813–837.

[6] J.-F. Cai, S. Osher, and Z. Shen, *Linearized Bregman Iterations for compressed sensing*, Math. Comp., 78 (2009), pp. 1515–1536.

[7] J.-F. Cai, S. Osher, and Z. Shen, *Convergence of the Linearized Bregman Iteration for* ℓ_1*-norm minimization*, Math. Comp., 78 (2009), pp. 2127–2136.

[8] A. Chambolle, *An algorithm for total variation minimization and applications*. J. Math. Imaging Vision, 20 (2004), pp. 89–97.

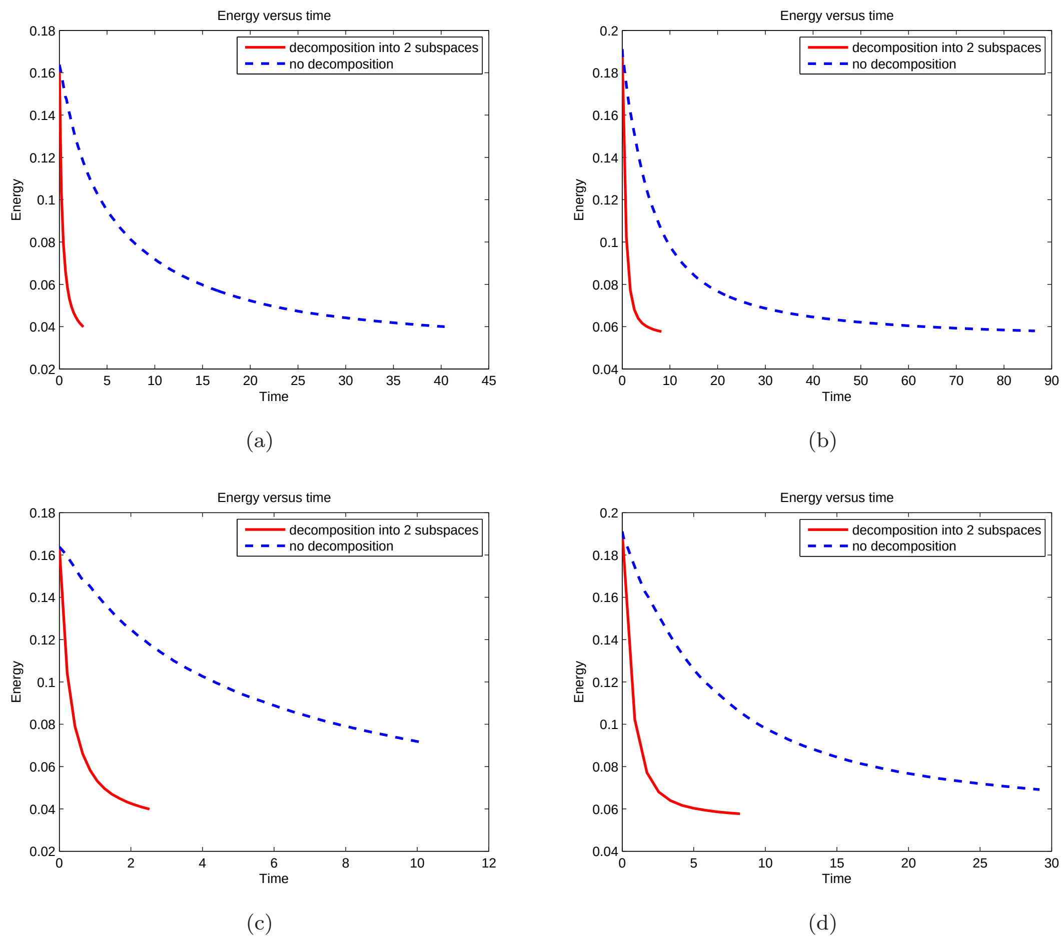

FIG. 6.4. *We show for N=1 and N=2 the decay of the energy. For the deblurring problem of Figure 6.1 the decays are plotted in (a) and zoomed in to the first 10 seconds in (c). In (b) and zoomed for the first 25 seconds in (d) the energies for the deblurring problem of Figure 6.3 are depicted.*

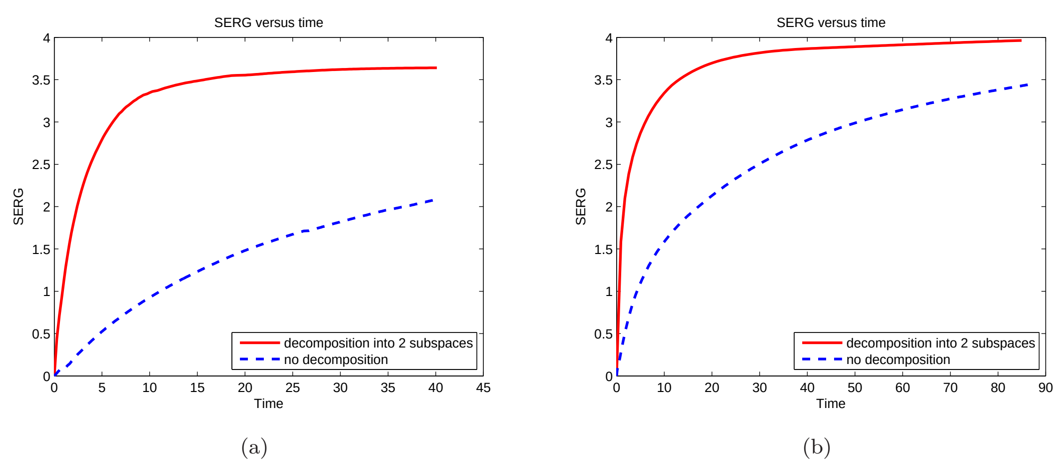

FIG. 6.5. *We show for N=1 and N=2 the evolution of the quality measure SERG: in (a) for the deblurring problem of Figure 6.1 and in (b) for the deblurring problem of Figure 6.3.*

[9] A. Chambolle and P.-L. Lions, *Image recovery via total variation minimization and related problems.*, Numer. Math., 76 (1997), pp. 167–188.
[10] A. Chambolle, V. Caselles, D. Cremers, M. Novaga, and T. Pock, *An Introduction to Total Variation for Image Analysis*, Theoretical Foundations and Numerical Methods for Sparse Recovery (M. Fornasier, ed.), Radon Series on Computational and Applied Mathematics, De Gruyter Verlag, 2010, pp. 263–340.
[11] A. Cohen, *Numerical Analysis of Wavelet Methods.*, Studies in Mathematics and its Applications 32. Amsterdam: North-Holland., 2003.
[12] A. Cohen, W. Dahmen, I. Daubechies, and R. DeVore, *Harmonic Analysis of the space BV*, Rev. Mat. Iberoamericana, 19 (2003), pp. 235–263.
[13] A. Cohen, R. DeVore, P. Petrushev, and H. Xu, *Nonlinear Approximation and the space $BV(\mathbb{R}^2)$*, Amer. J. Math. 121 (1999), pp. 587–628.
[14] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, 1992.
[15] I. Daubechies, M. Defrise, and C. De Mol, *An iterative thresholding algorithm for linear inverse problems*, Comm. Pure Appl. Math. 57 (2004), pp. 1413–1457.
[16] I. Daubechies, M. Fornasier, and I. Loris, *Acceleration of the projected gradient method for linear inverse problems with sparsity constraints*, J. Fourier Anal. Appl., 14 (2008), pp. 764–792.
[17] D. C. Dobson and C. R. Vogel, *Convergence of an iterative method for total variation denoising*, SIAM J. Numer. Anal. 34 (1997), pp. 1779–1791.
[18] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Mathematics and its Applications (Dordrecht). 375. Dordrecht: Kluwer Academic Publishers., 1996.
[19] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, CRC Press, 1992.
[20] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, *Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems*, IEEE Journal of Selected Topics in Signal Processing, Special Issue on Convex Optimization Methods for Signal Processing, 1 (2007), pp. 586–598.
[21] M. Fornasier, *Domain decomposition methods for linear inverse problems with sparsity constraints*, Inverse Problems, 23 (2007), pp. 2505–2526.
[22] M. Fornasier, *Numerical Methods for Sparse Recovery*, Theoretical Foundations and Numerical Methods for Sparse Recovery (M. Fornasier, ed.), Radon Series on Computational and Applied Mathematics, De Gruyter Verlag, 2010.
[23] M. Fornasier, A. Langer and C.-B. Schönlieb, *A convergent overlapping domain decomposition method for total variation minimization*, Numer. Math. 116 (2010), pp. 645–685.
[24] M. Fornasier and C.-B. Schönlieb, *Subspace correction methods for total variation and ℓ_1-minimization*, SIAM J. Numer. Anal., 47 (2009), pp. 3397–3428.
[25] T. Goldstein, S. Osher, *The split Bregman method for L_1 regularized problems*, UCLA CAM Report 08-29, 2008.
[26] K. Ito and K. Kunisch, *Lagrange Multiplier Approach to Variational Problems and Applications*, Series: Advances in Design and Control (No. 15) SIAM, 2008.
[27] S. J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, *A Method for ℓ_1-regularized least squares*, IEEE Journal of Selected Topics in Signal Processing, 1 (2007), pp. 606–617.
[28] A. Langer and M. Fornasier, *Analysis of the adaptive iterative Bregman algorithm*, preprint, 2010.
[29] S. Mallat, *A Wavelet Tour of Signal Processing. 2nd Ed.*, San Diego, CA: Academic Press., 1999.
[30] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, *An iterative regularization method for total variation-based image restoration*, Multiscale Model. Simul., 4 (2005), pp. 460–489.
[31] S. Osher, Y. Mao, B. Dong, W. Yin, *Fast Linearized Bregman Iteration for compressed sensing and sparse denoising*, Comm. Math. Sci., 8 (2010), pp. 93–111.
[32] L. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D, 60 (1992), pp. 259–268.
[33] L. Vese, *A study in the BV space of a denoising-deblurring variational problem*, Appl Math Optim 44 (2001), pp. 131–161.
[34] C. Vonesch and M. Unser, *A Fast multilevel algorithm for wavelet-regularized image restoration*, IEEE Transactions on Image Processing, 18 (2009), pp. 509523.
[35] W. Yin, *Analysis and generalization of the linearized Bregman method*, Rice University CAAM Technical Report TR09-02, 2009.
[36] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, *Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing*, SIAM J. Imag. Sci., 1 (2008), pp. 143–168.
[37] X. Zhang, M. Burger, X. Bresson, and S. Osher, *Bregmanized nonlocal regularization for deconvolution and sparse reconstruction*, SIAM J. Imag. Sci., to appear.