TOTAL VARIATION MINIMIZATION WITH AN $H^{-1}$ CONSTRAINT

CAROLA-BIBIANE SCHÖNLEB

Abstract. This paper is concerned with the numerical minimization of total variation functionals with an $H^{-1}$ constraint. We present an algorithm for its solution which is based on the dual formulation of the total variation and show its application in several areas of image processing.

1. Introduction and Motivations

Let $\Omega \subset \mathbb{R}^2$ be a bounded and open domain with Lipschitz boundary. For a given function $g \in L^2(\Omega)$ we are interested in the numerical realization of the following minimization problem

$$
\min_{u \in BV}(\Omega) \ J(u) = |Du|_\Omega + \frac{1}{2\lambda} \|Tu - g\|_{H^{-1}},
$$

where $T \in \mathcal{L}(L^2(\Omega))$ is a bounded linear operator and $\lambda > 0$ is a tuning parameter. The function $|Du|_\Omega$ is the total variation of $u$ and $\|\cdot\|_{-1}$ is the norm in $H^{-1}(\Omega)$, the dual of $H^1_0(\Omega)$.

Problem (1.1) is a model for minimizing the total variation of a function $u$ which obeys an $H^{-1}$ constraint, i.e., $\|Tu - g\|_{-1}$ is small, for a given function $g \in L^2(\Omega)$. In the terminology of inverse problems this means that from an observed datum $g$ one wants to determine the original function $u$, from which a priori we know that $Tu = g$ and $u$ has some regularity properties modeled by the total variation and the $H^{-1}$ norm. Minimization problems like this have important applications in a wide range of image processing tasks. We give an overview of such in the following subsections.

The main interest of this paper is the numerical solution of (1.1). In [14] Chambolle proposes an algorithm to solve total variation minimization with an $L^2$ constraint and $T = Id$. This approach was extended to more general operators $T$ in a subsequent work [7]. A generalization to other norms than $L^2$, including the case of the $H^{-1}$ norm for the case $T = Id$, was proposed in [5, 6]. In the following we shall report the generalization of Chambolle’s algorithm for the case of an $H^{-1}$ norm in the problem. Moreover, we present strategies to extend the use of this algorithm from problems with $T = Id$ to problems (1.1) with an arbitrary linear operator $T$. Additional to the theory of this algorithm we present applications of it in image processing, in particular for image denoising, image decomposition and inpainting. Finally we show how to speed up the numerical computations by considering a domain decomposition approach for our problem.

For now, let us start our discussion with some motivations for considering $TV - H^{-1}$ minimization.

1.1. Total Variation Minimization in Image Processing. In a wide range of image processing tasks one encounters the situation that the observed image $g$ is corrupted, e.g., by noise or blur, or that the features of interest in the image are hidden. Now the challenge is to recover the original image $u$, i.e., the hidden image features, from the observed datum $g$. In mathematical terms this means that one has to solve an inverse problem $Tu = g$, where $T$ models the process through which the image $u$ went before observation. In the case of an operator $T$ with unbounded inverse, this problem is ill-posed. In such cases one modifies the problem by introducing some additional a-priori information on $u$, usually in terms of a regularizing term, e.g., the total variation of $u$.
Many such problems in image processing are formulated as minimization problems, cf. [3] for a general overview on this topic. More precisely, let $\Omega \subset \mathbb{R}^2$ be an open and bounded domain with Lipschitz boundary, $B_1, B_2$ two Banach spaces over $\Omega$ and $g \in B_1$ be the given image. A general variational approach in image processing can be then written as

$$
(1.2) \quad J(u) = R(u) + \frac{1}{2\lambda} \| Tu - g \|_{B_1}^2 \to \min_{u \in B_2},
$$

where $\lambda > 0$ is the tuning parameter of the problem and $T \in \mathcal{L}(B_1)$ is a bounded linear operator. The term $R : B_2 \to \mathbb{R}$ denotes the regularizing term which smooths the image $u$ and represents some kind of a priori information about the minimizer $u$. The term $\| Tu - g \|_{B_1}$ is the so called fidelity term of the approach which forces the minimizer $u$ to stay close to the given image $g$ (how close is dependent on the size of $\lambda$). In the case of image denoising and decomposition the operator $T = Id$, the identity in $B_1$. For deblurring problems $T$ will be a symmetric kernel, e.g., a Gaussian, and for image inpainting $T$ denotes the characteristic function of a subdomain of $\Omega$, cf. Section 1.3 for details.

In general $B_2 \subset B_1$ signifying the smoothing effect of the regularizing term on the minimizer $u \in B_2$.

Under certain regularity assumptions on a minimizer $u$ of the functional $J$, the minimizer fulfills a so called optimality condition of (1.2), i.e., the corresponding Euler-Lagrange equation. This is, that $u$ is the variation of $u$ partial differential equation with certain boundary conditions on the total variation of the finite Radon measure $Du$ process. We denote it by $u(., t)$. At time $t = 0$, $u(., t = 0) = g \in B_1$ is the given image. It is then transformed through a process that can be written

$$
(1.3) \quad -\lambda \nabla R(u) + T^*(g - Tu) = 0, \quad \text{in} \quad \Omega,
$$

which is a partial differential equation with certain boundary conditions on $\partial \Omega$. Thereby $\nabla R$ denotes the Fréchet derivative of $R$ over $B_1 = B^2(\Omega)$. The evolutionary version of (1.3) is the so called steepest-descent or gradient flow approach. More precisely, a minimizer $u$ of (1.2) is embedded in an evolution process. We denote it by $u(., t)$. At time $t = 0$, $u(., t = 0) = g \in B_1$ is the given image. It is then transformed through a process that can be written

$$
(1.4) \quad u_t = -\lambda \nabla R(u) + T^*(g - Tu), \quad \text{in} \quad \Omega.
$$

Given a variational formulation (1.2), the steepest-descent approach is used to numerically compute a minimizer of $J$. Thereby (1.4) is iteratively solved until one is close enough to a minimizer of $J$, i.e., $u^{k+1} - u^k$ is sufficiently small for two subsequent iterates $u^k$ and $u^{k+1}$.

In other situations we encounter equations that do not come from variational principles, such as Cahn-Hilliard- and TV-$H^{-1}$ inpainting, cf. [10], [9], [13]. Then the image processing approach is directly given in the form of (1.4).

In what follows we shall concentrate on imaging approaches (1.2) which use the total variation as a regularizing term $R$, i.e., evolutionary approaches (1.4) where $\nabla R(u)$ is replaced by elements of the subdifferential of the total variation of $u$. We recall that for $u \in L^1_{\text{loc}}(\Omega)$

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

is the variation of $u$ and that $u \in BV(\Omega)$ (the space of bounded variation functions, [2, 25]) if and only if $V(u, \Omega) < \infty$, see [2, Proposition 3.6]. In such a case, $|Du|(\Omega) = V(u, \Omega)$, where $|Du|(\Omega)$ is the total variation of the finite Radon measure $Du$, the derivative of $u$ in the sense of distributions. The subdifferential $\partial |Du|(\Omega)$ is defined as

$$
\partial |Du|(\Omega) := \partial_{BV(\Omega)}(x) |Du|(\Omega) := \left\{ u^* \in BV(\Omega)' : \langle u^*, v - u \rangle + |Du|(\Omega) \leq |Dv|(\Omega) \quad \forall v \in BV(\Omega) \right\},
$$

where $BV(\Omega)'$ denotes the dual space of $BV(\Omega)$. It is obvious from this definition that $0 \in \partial |Du|(\Omega)$ if and only if $u$ is a minimizer of $|D\cdot|(\Omega)$.

The minimization of energies with total variation constraints, i.e.,

$$
(1.5) \quad \min_{u \in BV(\Omega)} \left\{ |Du|(\Omega) + \frac{1}{2\lambda} \| Tu - g \|_H^2 \right\},
$$

is the tuning parameter of the problem and $T \in \mathcal{L}(B_1)$ is a bounded linear operator. The term $R : B_2 \to \mathbb{R}$ denotes the regularizing term which smooths the image $u$ and represents some kind of a priori information about the minimizer $u$. The term $\| Tu - g \|_{B_1}$ is the so called fidelity term of the approach which forces the minimizer $u$ to stay close to the given image $g$ (how close is dependent on the size of $\lambda$). In the case of image denoising and decomposition the operator $T = Id$, the identity in $B_1$. For deblurring problems $T$ will be a symmetric kernel, e.g., a Gaussian, and for image inpainting $T$ denotes the characteristic function of a subdomain of $\Omega$, cf. Section 1.3 for details.
for a given \( g \in H \), where \( H \) is a suitable Hilbert space, e.g., \( H = L^2(\Omega) \), traces back to the first uses of such a functional model in noise removal in digital images as proposed by Rudin, Osher, and Fatemi [36]. There the operator \( T \) is just the identity. Extensions to more general operators \( T \) and numerical methods for the minimization of the functional appeared later in several important contributions [15, 23, 4, 39, 14]. From these pioneering and very successful results, the scientific output related to total variation minimization and its applications in signal and image processing increased dramatically in the last decade.

One of those outputs is TV-\( H^{-1} \) minimization (1.1), i.e., (1.5) with \( H = H^{-1}(\Omega) \). This approach found growing interest in recent years due to several advantages compared to the \( L^2 \) constrained problem, cf. the following subsections and, for instance, [35], [29], [13]. Thereby the norm in \( H^{-1}(\Omega) := (H^1_0(\Omega))^* \), i.e., the dual of \( H^1_0(\Omega) \) is defined by

\[
\|f\|_{-1}^2 = \|\nabla \Delta^{-1} f\|_{L^2(\Omega)}^2 = \int_\Omega (\nabla \Delta^{-1} f)^2 \, dx.
\]

The operator \( \Delta^{-1} \) denotes the inverse to the negative Dirichlet Laplacian, i.e., \( u = \Delta^{-1} f \) is the unique solution to

\[
\begin{align*}
-\Delta u &= f & \text{in } \Omega \\
u &= 0 & \text{on } \partial \Omega.
\end{align*}
\]

Note that the existence and uniqueness of minimizers for (1.1) is guaranteed. In fact the existence of a unique solution of (1.1) follows as a special case from the following theorem.

**Theorem 1.1.** (Theorem 3.1 in [29]) Given \( \Omega \subset \mathbb{R}^2 \), open, bounded and connected, with Lipschitz boundary, \( g \in H^{-s}(\mathbb{R}^2) \) (\( s > 0 \)), \( g = 0 \) outside of \( \Omega \), \( \lambda > 0 \), and \( T \in L^2(\Omega) \) an injective continuous linear operator such that \( T1_\Omega \neq 0 \), then the minimization problem

\[
\min_{u \in BV(\Omega)} |Du| (\Omega) + \frac{1}{2\lambda} \|T u - g\|_{L^2(\Omega)}^2, \quad s > 0,
\]

where \( \|\cdot\|_{L^2} \) denotes the norm in \( H^{-s}(\mathbb{R}^2) \), has a unique solution in \( BV(\Omega) \).

**Proof.** The proof is a standard application of methods from variational calculus and can be found, e.g., in [29]. The main ingredients in the proof, in order to guarantee compactness, are the Poincaré-Wirtinger inequality (cf. [1] for instance) which bounds the \( L^2 \) - and the \( L^1 \) - norm by the total variation, and the fact that \( L^2(\Omega) \) can be embedded in \( H^{-1}(\Omega) \).

In the subsequent two subsections we shall present two main applications of TV-\( H^{-1} \) minimization (1.1) for image denoising/decomposition, and image inpainting.

### 1.2. TV - \( H^{-1} \) Minimization for Image Denoising and Image Decomposition.

In taking \( T = Id \) in (1.2), we encounter two interesting areas in image processing, namely image denoising and image decomposition. Thereby, in many image denoising models, a given noisy image \( g \) is decomposed into its piecewise smooth part \( u \) and its oscillatory, noisy part \( v \), i.e., \( g = u + v \). Similar, in image decomposition, the piecewise smooth part \( u \) represents the structure/cartoon part of the image, and the oscillatory part \( v \) the texture part of the image. We call the latter task also cartoon-texture decomposition. The most famous model within this range is the TV - \( L^2 \) denoising model proposed by Rudin, Osher and Fatemi [36]

\[
\mathcal{J}(u) = |Du| (\Omega) + \frac{1}{2\lambda} \|u - g\|_{L^2(\Omega)}^2 \rightarrow \min_{u \in BV(\Omega)}.
\]

This model produces very good results for removing noise and preserving edges in structure images, meaning images without texture-like components, i.e., high oscillatory edges. Unfortunately it fails in the presence of the latter. Namely it cannot separate pure noise, i.e., well oscillatory components, from high oscillatory edges but removes both equally.
To overcome this situation, Y. Meyer [32] suggested to replace the $L^2$–fidelity term by a weaker norm. Namely he proposes the following model:

$$\mathcal{J}(u) = |Du|_1(\Omega) + \frac{1}{2\lambda} \| u - g \|_* \to \min_{u \in BV(\Omega)} ,$$

where the $\| \cdot \|_*$ is defined as follows.

**Definition 1.1.** Let $G$ denote the Banach space consisting of all generalized functions $g(x,y)$ which can be written as

$$g(x,y) = \nabla \cdot (\vec{f}(x,y)), \quad \vec{f} = (f_1,f_2), \quad f_1, f_2 \in L^\infty(\Omega), \quad \vec{f} \cdot \vec{n} = 0 \text{ on } \partial \Omega,$

where $\vec{n}$ is the unit normal on $\partial \Omega$. Then $\| f \|_*$ is the induced norm of $G$ defined as

$$\| f \|_* = \inf_{g = \nabla \cdot \vec{f}} \left\| \sqrt{f_1^2 + f_2^2} \right\|_{L^\infty(\Omega)} .$$

In fact, the space $G$ is the dual space of $W_0^{1,1}(\Omega)$. In [32] Meyer further introduces two other spaces with similar properties than $G$ but we are not going into detail here. We only mention that these spaces are intrinsically appropriate for modeling textured or oscillatory patterns and in fact they provide norms which are smaller for such than the $L^2$ norm. The drawback of Y. Meyers model is that it cannot be solved directly with respect to the minimizer $u$ and therefore has to be approximated, cf. [40]. Thereby the $*$–norm is replaced by

$$\frac{1}{2\mu} \left\| (u + \nabla \cdot \vec{f}) - g \right\|_{L^2(\Omega)}^2 + \frac{1}{\lambda} \left\| \sqrt{f_1^2 + f_2^2} \right\|_{L^p(\Omega)} ,$$

with $\lambda, \mu > 0$ and $p \geq 1$. In the case $p = 2$ the second term in the above expression is equivalent to the $H^{-1}$ norm. In particular $v = \nabla \cdot \vec{f}$ corresponds to $v \in H^{-1}(\Omega)$. Indeed, for $v \in H^{-1}(\Omega)$, there is a unique $v^* \in H_0^1(\Omega)$ such that

$$\| v \|_{L_2(\Omega)}^2 = \| \nabla v^* \|_{L_2(\Omega)}^2 = \| \nabla \Delta^{-1} v \|_{L_2(\Omega)}^2 = \left\| \sqrt{f_1^2 + f_2^2} \right\|_{L^2(\Omega)}^2 .$$

Limiting to the case $p = 2$ and the limit $\mu \to 0$ the TV-$H^{-1}$ denoising model was created, cf. [35], [33], [29], i.e.,

$$|Du|_1(\Omega) + \frac{1}{2\lambda} \| u - g \|_{L_2(\Omega)} \to \min_{u \in BV(\Omega)} .$$

Numerical experiments showed that (1.8) gives much better results than (1.7) under the presence of oscillatory data, cf. [35] and [29]. In Section 3 we will present some numerical results that support this claim.

1.3. TV-$H^{-1}$ Inpainting. Another important task in image processing is the process of filling in missing parts of damaged images based on the information gleaned from the surrounding areas. It is essentially a type of interpolation and is called inpainting. In this case the operator $T$ in (1.2) is the characteristic function of a subdomain of $\Omega$.

Now, let $g$ be the given image defined on the image domain $\Omega$. The problem is to reconstruct the original image $u$ in the damaged domain $D \subset \Omega$, called inpainting domain. The general variational approach (1.2) for this task then reads

$$\mathcal{J}(u) = R(u) + \frac{1}{2\lambda} \| \chi_{\Omega \setminus D}(u - g) \|^2_{B_1} \to \min_{u \in B_2},$$

where

$$\chi_{\Omega \setminus D}(x) = \begin{cases} 1 & \Omega \setminus D \\ 0 & D, \end{cases}$$

and

$$\chi_{\Omega \setminus D}(x) = \begin{cases} 1 & \Omega \setminus D \\ 0 & D, \end{cases}$$
is the characteristic function of $\Omega \setminus D$. Here the role of the regularizing term $R(u)$ is to fill in the image content into the missing domain $D$, e.g., by diffusion and/or transport. Due to the characteristic function $\chi_{\Omega \setminus D}$, the fidelity term of the inpainting approach has impact on the minimizer $u$ only outside of the inpainting domain only.

After the pioneering works of Masnou and Morel [31], and Bertalmio et al [8], the basic variational inpainting model is the $TV-L^2$ model, where as before $R(u) = |Du|_{\Omega} \approx \int_{\Omega} |\nabla u| \, dx$, $B_1 = L^2(\Omega)$ and $B_2 = BV(\Omega)$, cf. [18, 16, 37, 36]. A variational model with a regularizing term containing higher order derivatives is the Euler elastica model [20, 19, 31] where $R(u) = \int_{\Omega} \left( a + b \left( \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right) \right)^2 \right) |\nabla u| \, dx$ with positive weights $a$ and $b$. Other examples to be mentioned for (1.9) are the active contour model based on Mumford and Shah’s segmentation [38], and the inpainting scheme based on the Mumford-Shah-Euler image model [24], only to give a rough overview. For a more complete introduction to image inpainting we refer to [20].

Now second order variational inpainting methods (where the order of the method is determined by the derivatives of highest order in the corresponding Euler-Lagrange equation), like TV inpainting, have drawbacks as in the connection of edges over large distances (Connectivity Principle, cf. Figure 1) and the smooth propagation of level lines (sets of image points with constant grayvalue) into the damaged domain (Staircasing Effect, cf. Figure 2).

Figure 1. Two examples of Euler elastica inpainting compared with TV inpainting. In the case of large aspect ratios the TV inpainting fails to comply to the Connectivity Principle. Figure from [20].

This is due to the penalization of the length of the level lines within the minimizing process with a second order regularizing term, connecting level lines from the boundary of the inpainting domain via the shortest distance (linear interpolation). The regularizing term $R(u) \approx \int_{\Omega} |\nabla u| \, dx$ in the $TV-L^2$ inpainting approach, for instance, can be interpreted via the coarea formula which gives

$$\min_{u} \int_{\Omega} |\nabla u| \, dx \iff \min_{\Gamma_{\lambda}} \int_{-\infty}^{\infty} \text{length}(\Gamma_{\lambda}) \, d\lambda,$$

where $\Gamma_{\lambda} = \{ x \in \Omega : u(x) = \lambda \}$ is the level line for the grayvalue $\lambda$. If we consider on the other hand the regularizing term in the Euler elastica inpainting approach, the coarea formula reads

$$\min_{u} \int_{\Omega} \left( a + b \left( \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right) \right)^2 \right) |\nabla u| \, dx \iff \min_{\Gamma_{\lambda}} \int_{-\infty}^{\infty} a \text{length}(\Gamma_{\lambda}) + b \text{curvature}^2(\Gamma_{\lambda}) \, d\lambda.$$
Despite the presence of high curvature TV inpainting truncates the circle inside the inpainting domain (linear interpolation of level lines, i.e., Staircasing Effect). Depending on the weights $a$ and $b$ Euler’s elastica inpainting returns a smoothly restored object, taking the curvature of the circle into account. Figure from [19].

Thereby not only the length of the level lines but also their curvature is penalized (where the penalization of each depends on the ratio $b/a$). This results in a smooth continuation of level lines over the inpainting domain also over large distances, compare Figure 1 and 2. The performance of higher order inpainting methods, such as Euler elastica inpainting, can also be interpreted via the second boundary condition, necessary for the well-posedness of the corresponding Euler-Lagrange equation of fourth order. Not only the grayvalues of the image are specified on the boundary of the inpainting domain but also the gradient of the image function, namely the direction of the level lines are given.

In an attempt to solve both the connectivity principle and the staircasing effect resulting from second order image diffusions, a number of third and fourth order diffusions has been suggested for image inpainting. One of them is TV-$H^{-1}$ inpainting. Thereby the inpainted image $u$ of $g \in L^2(\Omega)$, shall evolve via

$$u_t = \lambda \Delta p + \chi_{\Omega \setminus D}(g - u), \quad p \in \partial |Du| (\Omega),$$

where $\partial |Du| (\Omega)$ denotes the subdifferential of the total variation. A similar form of this inpainting approach appeared the first time in [13]. In Section 3 we will present some numerical results for this approach.

Note that (1.12) does not follow a variational principle. Let $p \in \partial |Du| (\Omega)$. Then, in fact the functional

$$|Du| (\Omega) + \frac{1}{2\lambda} \left\| \chi_{\Omega \setminus D}(u - g) \right\|_{-1}^2$$

exhibits the optimality condition

$$0 = \lambda p + \chi_{\Omega \setminus D} \Delta^{-1}(\chi_{\Omega \setminus D}(u - g)),$$

which splits into

$$0 = \lambda p \quad \text{in } D$$

$$0 = \lambda p + \Delta^{-1}(u - g) \quad \text{in } \Omega \setminus D.$$  

Hence the minimization of (1.13) translates into a second order diffusion inside the inpainting domain $D$, whereas a stationary solution of (1.12) fulfills

$$0 = \lambda \Delta p \quad \text{in } D$$

$$0 = \lambda \Delta p + (g - u) \quad \text{in } \Omega \setminus D.$$
1.4. Numerical Solution for TV-$H^{-1}$ Minimization. The numerical solution of TV-$H^{-1}$ approaches is a challenging task on its own. Generally speaking, the computational costs are very high because of the high (fourth!) differential order of the corresponding optimality condition, induced by the $H^{-1}$ norm in the problem. Solving (1.1) by an explicit steepest descent iteration for instance, results in a stepsize of order $O(\Delta x^4)$, where $\Delta x$ is the spatial stepsize in $\Omega$. In particular in the case when $T \neq Id$ even existing semi-implicit schemes are on the one hand unconditionally stable, i.e., the stepsize for the iterations can be chosen arbitrarily large, but their convergence to a minimizer is still slow, depending on the size of the tuning parameter $\lambda$. This will be discussed in more detail at the end of this section. However, it is clear, that the numerical solution of TV-$H^{-1}$ approaches in general poses the challenge of limiting the computational costs.

Now, the numerical solution of TV-$H^{-1}$ minimization depends on the specific problem at hand. In [29] Lieu and Vese proposed a numerical method to solve TV-$H^{-1}$ denoising/decomposition (1.8) by using the Fourier representation of the $H^{-1}$ norm on the whole $\mathbb{R}^d$, $d \geq 1$. Thereby the space $H^{-1} (\mathbb{R}^d)$ is defined as a Hilbert space equipped with the inner product

$$\langle f, g \rangle_{-1} = \int \left(1 + |\xi|^2\right)^{-1} \hat{f} \cdot \hat{g} \, d\xi$$

and associated norm $\|f\|_{-1} = \sqrt{\langle f, f \rangle_{-1}}$, cf. also [22]. Here $\hat{g}$ denotes the Fourier transform of $g$ in $L^2(\mathbb{R}^d)$, i.e.,

$$(1.14) \quad \hat{g}(y) := \frac{1}{(2\pi)^{2d}} \int_{\mathbb{R}^d} e^{-ixy} g(x) \, dx, \quad y \in \mathbb{R}^d,$$

and $\bar{g}$ the complex conjugate of $g$. $H^{-1}(\mathbb{R}^d)$ is the dual space of $H^1(\mathbb{R}^d)$. Note that we consider functions $g$ defined in $\mathbb{R}^d$ rather than on a bounded domain which can be done by considering zero extensions of the image function. With this definition of the $H^{-1}$ norm the corresponding optimality condition for (1.8) reads

$$\lambda p + \left[2 \text{Re}\left\{\frac{\bar{\hat{g}} - \bar{\hat{u}}}{(1 + |\xi|^2)^{-1}}\right\}\right] = 0 \quad \text{in } \Omega$$

$$\sum_{\nabla u} \cdot \vec{n} = 0 \quad \text{on } \partial \Omega$$

$$u = 0 \quad \text{outside } \Omega,$$

where $\hat{(g)}$ denotes the inverse Fourier transform of $g$, defined in analogy to (1.14), $\text{Re}$ denotes the real part of a complex number, and $\vec{n}$ is the outward pointing unit normal vector on $\partial \Omega$. For the numerical computation of a solution of (1.15), we approximate an element $p$ of the subdifferential of $|\nabla u|(\Omega)$, by its relaxed version

$$(1.16) \quad p \approx \nabla \cdot (\nabla u / |\nabla u|_c),$$

where $|\nabla u|_c = \sqrt{|\nabla u|^2 + \epsilon}$.

Equation (1.15) leads to solve a second order PDE rather than a fourth order PDE, resulting in a better CFL condition for the numerical scheme, cf. [29]. The dual approach for TV-$H^{-1}$ denoising and decomposition [5, 6] shall be presented in section 2.

In the case of TV-$H^{-1}$ inpainting the situation is completely different since (1.12) does not fulfill a variational principle, cf. Section 1.3. In [11] and [13] the authors used a convexity splitting scheme to solve (1.12). Convexity splitting algorithms, proposed by Eyre in [26], are usually applied to gradient flows and provide a semi-implicit scheme for the discretization in time. Roughly said convexity splitting means to solve a gradient system

$$\begin{cases}
  u_t = -\nabla J(u) & \text{in } \Omega, \\
  u(., t = 0) = u_0 & \text{in } \Omega,
\end{cases}$$
with initial condition \( u_0 \in \mathbb{R}^N \) and a functional \( J \) from \( \mathbb{R}^N \) into \( \mathbb{R} \), where \( N \) is the dimension of the data space, by a semi-implicit algorithm of the form: Pick an initial \( u^0 = u_0 \) and iterate for \( k \geq 0 \)

\[
u^{k+1} - u^k = \tau \left( \nabla J_c(u^k) - \nabla J_e(u^{k+1}) \right),
\]

where \( u^k \) approximates the exact solution \( u(k\tau) \) (where \( \tau \) denotes the timestep) and \( J_c, J_e \) are strictly convex and chosen such that

\[
J(u) = J_c(u) - J_e(u).
\]

Under certain assumptions this discretization approach is unconditionally stable and relatively easy to apply to a large range of variational problems. Although most higher-order inpainting schemes arent gradient flows, among them (1.12), this method can still be applied in a modified form. For more details to the application of convexity splitting algorithms in higher order inpainting compare [11]. For the application of convexity splitting to (1.12) an element \( p \in \partial TV(u) \) is replaced by its relaxed version (1.16), namely we want to solve

\[
(1.17) \quad u_t = -\lambda \Delta (\nabla \cdot (\frac{\nabla u}{|\nabla u|_e})) + \chi_{\Omega \setminus D}(g-u).
\]

In [11] the authors propose the following splitting for the TV-\( H^{-1} \) inpainting equation. The regularizing term in (1.17) can be modeled by a gradient flow in \( H^{-1} \) of the energy

\[
J^1(u) = \int_{\Omega} |\nabla u|_e \, dx.
\]

We split \( J^1 \) in \( J^1_c - J^1_e \) with

\[
J^1_c(u) = \int_{\Omega} \frac{C_1}{2} |\nabla u|^2 \, dx, \quad J^1_e(u) = \int_{\Omega} -|\nabla u|_e + \frac{C_1}{2} |\nabla u|^2 \, dx.
\]

The fitting term is a gradient flow in \( L^2 \) of the energy

\[
J^2(u) = \frac{1}{2\lambda} \int_{\Omega} \chi_{\Omega \setminus D}(u-g)^2 \, dx
\]

and is splitted into \( J^2 = J^2_c - J^2_e \) with

\[
J^2_c(u) = \int_{\Omega} \frac{C_2}{2} |u|^2 \, dx, \quad J^2_e(u) = \int_{\Omega} -\frac{1}{2\lambda} \chi_{\Omega \setminus D}(u-g)^2 + \frac{C_2}{2} |u|^2 \, dx.
\]

For the splittings discussed above the resulting time-stepping scheme is

\[
\frac{u^{k+1} - u^k}{\tau} = -\nabla H^{-1}(J^1_c(u^{k+1}) - J^1_c(u^k)) - \nabla L^2(J^2_c(u^{k+1}) - J^2_c(u^k)),
\]

where \( \nabla_{H^{-1}} \) and \( \nabla_{L^2} \) represent the Fréchet derivatives with respect to the \( H^{-1} \) inner product and the \( L^2 \) inner product respectively. This translates to a numerical scheme of the form

\[
(1.18) \quad \frac{u^{k+1} - u^k}{\tau} + C_1 \Delta^2 u^{k+1} + C_2 u^k = C_1 \Delta^2 u^k - \Delta (\nabla \cdot (\frac{\nabla u^k}{|\nabla u^k|_e})) + C_2 u^k + \frac{1}{\lambda} \chi_{\Omega \setminus D}(g-u^k),
\]

where \( \Delta^2 = \Delta \Delta \). In order to make the scheme unconditionally stable, the constants \( C_1 \) and \( C_2 \) have to be chosen so that \( J^i_c, J^i_e, i = 1, 2 \), are all convex. The condition turns out to be \( C_1 > \frac{1}{\tau} \) and \( C_2 > 1/\lambda \). Since usually in inpainting tasks \( \lambda \) is chosen comparatively small, e.g., \( \lambda = 10^{-3} \), the condition on \( C_2 \) makes the numerical scheme (1.18), although unconditionally stable, quite slow.

In the following we are going to present a method introduced by Chambolle [14] for TV-\( L^2 \) minimization (1.7) and its generalization for the TV-\( H^{-1} \) case (1.1). This algorithm will give us the opportunity to address TV-\( H^{-1} \) minimization in a general way.
2. An Algorithm for TV-$H^{-1}$ Minimization

2.1. Preliminaries. Throughout this section $\|\cdot\|$ denotes the norm in $X = L^2(\Omega)$ in the continuous setting, i.e., the Euclidean norm in $X = \mathbb{R}^{N \times M}$ in the discrete setting. In the discrete setting the continuous image domain $\Omega = [a, b] \times [c, d] \subset \mathbb{R}^2$ is approximated by a finite grid $\{a = x_1 < \ldots < x_N = b\} \times \{c = y_1 < \ldots < y_M = d\}$ with equidistant step-size $h = x_{i+1} - x_i = \frac{b-a}{N} = \frac{d-c}{M} = y_{j+1} - y_j$ equal to 1 (one pixel). The digital image $u$ is an element in $X$. We denote $u(x_i, y_j) = u_{i,j}$ for $i = 1, \ldots, N$ and $j = 1, \ldots, M$.

Further we define $Y = X \times X$ with Euclidean norm $\|\cdot\|_Y$ and inner product $\langle \cdot, \cdot \rangle_Y$. Moreover the operators gradient $\nabla$, divergence $\nabla \cdot$ and Laplacian $\Delta$ in the discrete setting are defined as follows:

The gradient $\nabla u$ is a vector in $Y$ given by forward differences

$$(\nabla u)_{i,j} = ((\nabla_x u)_{i,j}, (\nabla_y u)_{i,j}),$$

with

$$(\nabla_x u)_{i,j} = \begin{cases} u_{i+1,j} - u_{i,j} & \text{if } i < N \\ 0 & \text{if } i = N \end{cases},$$

$$(\nabla_y u)_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j} & \text{if } j < M \\ 0 & \text{if } j = M \end{cases},$$

for $i = 1, \ldots, N$, $j = 1, \ldots, M$.

We further introduce a discrete divergence $\nabla \cdot : Y \to X$ defined, by analogy with the continuous setting, by $\nabla \cdot = -\nabla^* (\nabla^* \text{ is the adjoint of the gradient } \nabla)$. That is, the discrete divergence operator is given by backward differences like

$$(\nabla \cdot p)_{i,j} = \begin{cases} p^x_{i,j} - p^x_{i-1,j} & \text{if } 1 < i < N \\ p^x_{i,j} & \text{if } i = 1 \\ -p^x_{i-1,j} & \text{if } i = N \end{cases} + \begin{cases} p^y_{i,j} - p^y_{i,j-1} & \text{if } 1 < j < M \\ p^y_{i,j} & \text{if } j = 1 \\ -p^y_{i,j-1} & \text{if } j = M \end{cases},$$

for every $p = (p^x, p^y) \in Y$.

Finally we define the discrete Laplacian as $\Delta = \nabla \cdot \nabla$, i.e.,

$$(\Delta u)_{i,j} = \begin{cases} u_{i+1,j} - 2u_{i,j} + u_{i-1,j} & \text{if } 1 < i < N \\ u_{i+1,j} - u_{i,j} & \text{if } i = 1 \\ u_{i-1,j} - u_{i,j} & \text{if } i = N \\ u_{i,j+1} - 2u_{i,j} + u_{i,j-1} & \text{if } 1 < j < M \\ u_{i,j+1} - u_{i,j} & \text{if } j = 1 \\ u_{i,j-1} - u_{i,j} & \text{if } j = M, \end{cases}$$

and its inverse operator $\Delta^{-1}$, as in the continuous setting (1.6), i.e., $u = \Delta^{-1} f$ is the unique solution of

$$\begin{cases} -(\Delta u)_{i,j} = f_{i,j} & 1 < i < N, 1 < j < M \\ u_{i,j} = 0 & i = 1, N; j = 1, M. \end{cases}$$

Moreover, without always indicating it, when in the discrete setting, instead of minimizing

$$J(u) = \frac{1}{2\lambda} \|u - g\|_{-1}^2 + |Du| (\Omega),$$

we consider the discretized functional

$$J^h(u) := \sum_{1 \leq i \leq N, 1 \leq j \leq M} \frac{1}{2\lambda} (\nabla \Delta^{-1} (u - g))_{i,j}^2 + \left\| (\nabla u)_{i,j} \right\|,$$

with $|y| = \sqrt{y_1^2 + y_2^2}$ for every $y = (y_1, y_2) \in \mathbb{R}^2$ and some step-size $h$. As already pointed out in [14], the functional $J^h$ multiplied by $h$ converges as $h \to 0$ in the $\Gamma$ sense to $J$, cf. [12] for details.
2.2. Chambolle’s Algorithm for Total Variation Minimization. In [14] Chambolle proposes an algorithm to numerically compute a minimizer of

$$J(u) = \frac{1}{2\lambda} \|u - g\|_{L^2(\Omega)}^2 + |Du|_{L^2(\Omega)} + |Du|_{\Omega}.$$ 

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 the minimizer $u$ of $J$ as

$$u = g - P_{\lambda K}(g),$$

where $P_{\lambda K}$ denotes the orthogonal projection over $L^2(\Omega)$ on the convex set $K$ which is the closure of the set

$$\{ \nabla \cdot \xi : \xi \in C^1_c(\Omega; \mathbb{R}^2), |\xi(x)| \leq 1 \forall x \in \mathbb{R}^2 \}.$$ 

To numerically compute the projection $P_{\lambda K}(g)$ he uses a fixed point algorithm. All this will be explained in more detail in the context of TV-$H^{-1}$ minimization in the following subsection.

2.3. A Generalization of Chambolle’s Algorithm for TV-$H^{-1}$ Minimization. The main contribution of this paper is to generalize Chambolle’s algorithm to the case of an $H^{-1}$ constrained minimization of the total variation where $T$ is an arbitrary linear and bounded operator. In short we shall see how to solve (1.1) using a similar strategy as in [14]. We start with solving the simplified problem when $T = Id$, as also proposed in [5, 6], and as a second step present a method how to use this solution in order to solve the general case (1.1). Hence for the time being we consider the minimization problem

$$\min_u \{ J(u) = |Du|_{\Omega} + \frac{1}{2\lambda} \|u - g\|_{\Omega}^2 \}.$$ 

We proceed by exploiting the optimality condition of (2.19), i.e.,

$$0 \in \partial |Du|_{\Omega} + \Delta^{-1}(u - g) \frac{1}{\lambda}.$$ 

This can be rewritten as

$$\frac{\Delta^{-1}(g - u)}{\lambda} \in \partial |Du|_{\Omega}.$$ 

Since

$$s \in \partial f(x) \iff x \in \partial f^*(s),$$ 

where $f^*$ is the conjugate (or Fenchel transform) of $f$, it follows

$$u \in \partial |D|_{(\Omega)^*} \left( \frac{\Delta^{-1}(g - u)}{\lambda} \right).$$ 

Here

$$|D|_{(\Omega)^*}(v) = \chi_K(v) = \begin{cases} 0 & \text{if } v \in K \\ +\infty & \text{otherwise}, \end{cases}$$

where $K$ is the closure of the set

$$\{ \nabla \cdot \xi : \xi \in C^1_c(\Omega; \mathbb{R}^2), |\xi(x)| \leq 1 \forall x \in \mathbb{R}^2 \},$$ 

as before. Rewriting the above inclusion again we have

$$\frac{g}{\lambda} \in \frac{u - g}{\lambda} + \frac{1}{\lambda} |D|_{(\Omega)^*} \left( \frac{\Delta^{-1}(g - u)}{\lambda} \right),$$

i.e., with $w = \Delta^{-1}(g - u)/\lambda$ it reads

$$0 \in (-\Delta w - g/\lambda) + \frac{1}{\lambda} |D|_{(\Omega)^*}(w) \quad \text{in } \Omega,$$

$$w = 0 \quad \text{on } \partial \Omega.$$
In other words \( w \) is a minimizer of
\[
\frac{\|w - \Delta^{-1}g/\lambda\|_{H^1_0(\Omega)}^2}{2} + \frac{1}{\lambda} |D| (\Omega)^*(w),
\]
where \( H^1_0(\Omega) = \{ v \in H^1(\Omega) : v = 0 \text{ on } \partial \Omega \} \) and \( \|v\|_{H^1_0(\Omega)} = \|\nabla v\| \). Because of (2.21), for \( w \) being a minimizer of the above functional it is necessary that \( |D| (\Omega)^*(w) = 0 \), i.e., \( w \in K \). Hence a minimizer \( w \) fulfills
\[
w = p_K^1(\Delta^{-1}g/\lambda),
\]
where \( p_K^1 \) is the orthogonal projection on \( K \) over \( H^1_0(\Omega) \), i.e.,
\[
p_K^1(u) = \operatorname{argmin}_{v \in K} \|u - v\|_{H^1_0(\Omega)}.
\]
Hence the solution \( u \) of problem (2.19) is given by
\[
u = g + \Delta (p_K^1(\Delta^{-1}g)) \cdot \]
where \( -\Delta \) denotes the zero Dirichlet Laplacian as before.

Computing the nonlinear projection \( p_K^1(\Delta^{-1}g) \) amounts to solve the following problem:
\[
(2.22) \quad \min \left\{ \left\| (\nabla (\lambda \nabla \cdot p - \Delta^{-1}g))_{ij} \right\| : p \in Y, |p_{i,j}| \leq 1 \forall i = 1, \ldots, N; j = 1, \ldots, M \right\}.
\]
Analogous to [14] we use the Karush-Kuhn-Tucker conditions for the above constrained minimization. Then there exist \( \alpha_{i,j} \geq 0 \) such that the corresponding Euler-Lagrange equation reads
\[
\left[ \nabla (\Delta (\lambda \nabla \cdot p - \Delta^{-1}g))_{ij} \right] + \alpha_{i,j} p_{i,j} = 0, \quad \forall i = 1, \ldots, N; j = 1, \ldots, M,
\]
where either \( \alpha_{i,j} > 0 \) and \( |p_{i,j}| = 1 \) or \( |p_{i,j}| < 1 \) and \( \alpha_{i,j} = 0 \). Now, following the arguments in [14], in both cases this yields
\[
\alpha_{i,j} = \left| \nabla \Delta (\nabla \cdot p^n - \Delta^{-1}g/\lambda)_{ij} \right|, \quad \forall i = 1, \ldots, N; j = 1, \ldots, M.
\]
Then the gradient descent algorithm for solving (2.22) reads: for an initial \( p^0 = 0 \), iterate for \( n \geq 0 \)
\[
(2.23) \quad p_{i,j}^{n+1} = \frac{p_{i,j}^n - \tau (\nabla \Delta (\nabla \cdot p^n - \Delta^{-1}g/\lambda))_{ij}}{1 + \tau |(\nabla \Delta (\nabla \cdot p^n - \Delta^{-1}g/\lambda))_{ij}|}.
\]
Redoing the convergence proof in [14] we end up with a similar result. Essentially the same proof can be found in [5, 6].

**Theorem 2.1.** Let \( \tau \leq 1/64 \). Then, \( \lambda \nabla \cdot p^n \) converges to \( p_K^1(\Delta^{-1}g) \) as \( n \to \infty \).

**Proof.** The proof works similar to the proof in [14]. For the sake of completeness and clarity we will present the detailed proof here, keeping close to the notation in [14]. By induction we easily see that for every \( n \geq 0 \), \( |p_{i,j}^n| \leq 1 \) for all \( i, j \). Indeed, starting with \( p^0 \), with \( |p_{i,j}^0| \leq 1 \) for all \( i = 1, \ldots, N; j = 1, \ldots, M \), we have
\[
|p_{i,j}^{n+1}| \leq \frac{|p_{i,j}^n| + \tau \left| (\nabla \Delta (\nabla \cdot p^n - \Delta^{-1}g/\lambda))_{ij} \right|}{1 + \tau |(\nabla \Delta (\nabla \cdot p^n - \Delta^{-1}g/\lambda))_{ij}|} \leq 1.
\]
Now, let us fix an \( n \geq 0 \) and consider \( \left\| \nabla (\nabla \cdot p^{n+1} - \Delta^{-1}(g/\lambda)) \right\|_{Y} \). We want to show that this norm is decreasing with \( n \). In what follows we will abbreviate \( \|\| \) by \( \| \| \). We have
\[
\left\| \nabla (\nabla \cdot p^{n+1} - \Delta^{-1}(g/\lambda)) \right\|^2 = \|\nabla (\nabla \cdot p^{n+1} - p^n) + \nabla (\nabla \cdot p^n - \Delta^{-1}(g/\lambda))\|^2
\]
\[
= \|\nabla (\nabla \cdot p^n - \Delta^{-1}(g/\lambda))\|^2 + 2 \langle \nabla \cdot (p^{n+1} - p^n), \nabla (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)) \rangle
\]
\[
+ \|\nabla (\nabla \cdot (p^{n+1} - p^n))\|^2.
\]
Inserting $\eta = (p^{n+1} - p^n)/\tau$ in the above equation and integrating by parts in the second term we get

$$
\|\nabla (\nabla \cdot p^{n+1} - \Delta^{-1}(g/\lambda))\|^2 = \|\nabla (\nabla \cdot p^n - \Delta^{-1}(g/\lambda))\|^2 + 2\tau \langle \eta, \nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)) \rangle + \tau^2 \|\nabla \cdot \eta\|^2.
$$

By further estimating $\|\nabla \nabla \cdot \eta\| \leq \kappa \|\eta\|$, where $\kappa = \|\|\nabla \nabla \cdot \|\| = \sup_{\|p\| \leq 1} \|\nabla \nabla \cdot p\|$ the norm of the operator $\nabla \nabla \cdot : Y \to Y$, we deduce

$$
\|\nabla (\nabla \cdot p^{n+1} - \Delta^{-1}(g/\lambda))\|^2 \leq \|\nabla (\nabla \cdot p^n - \Delta^{-1}(g/\lambda))\|^2 + 2\tau \langle \eta, \nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)) \rangle + \kappa^2 \tau \|\eta\|^2.
$$

The operator norm $\kappa$ will be bounded at the end of the proof. For now we are going to show that the term multiplied by $\tau$ is always negative as long as $p^{n+1} \neq p^n$ and $\tau \leq 1/\kappa^2$, and hence that $\|\nabla (\nabla \cdot p^n - \Delta^{-1}(g/\lambda))\|^2$ is decreasing. To do so we consider

$$
(2.24) \quad 2 \langle \eta, \nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)) \rangle + \kappa^2 \|\eta\|^2 = \sum_{i,j \leq N} 2\eta_{i,j} \left(\nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda))\right)_{i,j} + \kappa^2 \|\eta\|^2.
$$

Now, from the fixed point equation we have

$$
\eta_{i,j} = -\left(\langle (\nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)))_{i,j} + (\nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)))_{i,j} \cdot p^{n+1}_{i,j}\right).
$$

Setting $\rho_{i,j} = \left| (\nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)))_{i,j} \cdot p^{n+1}_{i,j}\right|$ and inserting the above expression for $\eta_{i,j}$ into (2.24) we have for every $i, j$

$$
2\rho_{i,j} (\nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)))_{i,j} + \kappa^2 \|\rho_{i,j}\|^2 = \left(\kappa^2 \tau - 1\right) \|\rho_{i,j}\|^2 + \|\rho_{i,j}\|^2.
$$

Since $|p^{n+1}_{i,j}| \leq 1$ it follows that $|\rho_{i,j}| \leq \left| (\nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)))_{i,j}\right|$, and hence

$$
2\rho_{i,j} (\nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)))_{i,j} + \kappa^2 \|\rho_{i,j}\|^2 \leq \left(\kappa^2 \tau - 1\right) \|\rho_{i,j}\|^2.
$$

The last term is negative or zero if and only if $\kappa^2 \tau - 1 \leq 0$. Hence, if

$$
\tau \leq 1/\kappa^2,
$$

we see that $\|\nabla (\nabla \cdot p^n - \Delta^{-1}(g/\lambda))\|$ is nonincreasing with $n$. For $\tau < 1/\kappa^2$ it is immediately clear that the norm is even decreasing, unless $\eta = 0$, that is, $p^{n+1} = p^n$. The same holds for $\kappa^2 \tau = 1$.

Indeed, in this case, if $\|\nabla (\nabla \cdot p^{n+1} - \Delta^{-1}(g/\lambda))\|^2 = \|\nabla (\nabla \cdot p^n - \Delta^{-1}(g/\lambda))\|^2$ it follows that

$$
0 = 2 \langle \eta, \nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)) \rangle + \tau \|\nabla \nabla \cdot \eta\|^2
\leq 2 \langle \eta, \nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)) \rangle + \kappa^2 \|\eta\|^2
\leq \sum_{i,j \leq N} -\left| (\nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)))_{i,j}\right|^2 + |\rho_{i,j}|^2.
$$

and therefore $\left| (\nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)))_{i,j}\right| \leq |\rho_{i,j}|$. Since in turn

$$
|\rho_{i,j}| \leq \left| (\nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)))_{i,j}\right| \quad \text{we deduce} \quad |\rho_{i,j}| = \left| (\nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)))_{i,j}\right| \quad \text{for each} \quad i, j.
$$

But this can only be if either $\left| (\nabla \Delta (\nabla \cdot p^n - \Delta^{-1}(g/\lambda)))_{i,j}\right| = 0$ or $|p^{n+1}_{i,j}| = 1$. In both cases, the fixed point iteration (2.23) yields $p^{n+1}_{i,j} = p^n_{i,j}$ for all $i, j$.

Now, since $\|\nabla (\nabla \cdot p^n - \Delta^{-1}(g/\lambda))\|$ is decreasing with $n$, the norm is uniformly bounded and hence there exists an $m \geq 0$ such that

$$
m = \lim_{n \to \infty} \|\nabla (\nabla \cdot p^n - \Delta^{-1}(g/\lambda))\|.
$$
Moreover the sequence \( p^n \) has converging subsequences. Let \( \bar{p} \) be the limit of a subsequence \((p^{n_k})\) and \( \bar{p}' \) be the limit of \((p^{n_{k+1}})\). Inserting \( p^{n+1} \) and \( p^n \) into the fixed point equation (2.23) and passing to the limit we have

\[
\bar{p}_{i,j}' = \frac{\bar{p}_{i,j} - \tau (\nabla \Delta (\nabla \cdot \bar{p} - \Delta^{-1}g/\lambda))_{i,j}}{1 + \tau |(\nabla \Delta (\nabla \cdot \bar{p} - \Delta^{-1}g/\lambda))_{i,j}|}.
\]

Repeating the previous calculations we see that since \( m = \| \nabla (\nabla \cdot \bar{p} - \Delta^{-1}g/\lambda) \| \leq \| \nabla (\nabla \cdot \bar{p} - \Delta^{-1}g/\lambda) \| \), it must be that \( \bar{p}_{i,j} = (\bar{p}_{i,j}' - \bar{p}_{i,j}) / \tau = 0 \) for every \( i, j \), that is, \( \bar{p} = \bar{p}' \).

Hence \( \bar{p} \) is a fixed point of (2.23), i.e.,

\[
(\nabla \Delta (\nabla \cdot \bar{p} - \Delta^{-1}g/\lambda))_{i,j} + |(\nabla \Delta (\nabla \cdot \bar{p} - \Delta^{-1}g/\lambda))_{i,j}| \bar{p}_{i,j} = 0, \quad \forall i = 1, \ldots, N; \quad j = 1, \ldots, M,
\]

which is the Euler equation for a solution of (2.22). One can deduce that \( \bar{p} \) solves (2.22) and that \( \lambda \nabla \cdot \bar{p} \) is the projection \( P^K_1(\Delta^{-1}g) \). Since this projection is unique, we deduce that all the sequence \( \lambda \nabla \cdot p^n \) converges to \( P^K_1(\Delta^{-1}g) \). The theorem is proved if we can show that \( \kappa^2 \leq 64 \). By definition

\[
\kappa = \| \nabla \nabla \cdot \| = \sup_{\|p\|_1 \leq 1} \| \nabla \nabla \cdot p \|.
\]

Then for every \( i, j \), we have

\[
\| \nabla \nabla \cdot p \|^2 = \sum_{1 \leq i, j \leq N} |(\nabla \nabla \cdot p)_{i,j}|^2.
\]

For more clarity let us set \( u := \nabla \cdot p \in X \) for now. With the convention that \( p_{0,j} = p_{N,j} = p_{i,0} = p_{i,M} = 0 \) we get

\[
\| \nabla \nabla \cdot p \|^2 = \sum_{1 \leq i, j \leq N} \left[ |(\nabla u)_{i,j}|^2 + |(\nabla u)_{i,j}|^2 \right] = \sum_{1 \leq i, j \leq N} \left[ |(\nabla \cdot p)_{i+1,j} - (\nabla \cdot p)_{i,j}|^2 + \sum_{1 \leq i, j \leq N} (|(\nabla \cdot p)_{i,j+1} - (\nabla \cdot p)_{i,j}|)^2 \right] = \sum_{1 \leq i, j \leq N} \left[ |(\nabla \cdot p)_{i,j}|^2 + |(\nabla \cdot p)_{i+1,j}|^2 + |(\nabla \cdot p)_{i,j+1}|^2 \right] \leq 8 \sum_{1 \leq i, j \leq N} \left[ |p^0_{i,j}|^2 + |p^0_{i+1,j}|^2 + |p^0_{i,j+1}|^2 \right] \leq 64 \cdot \| u \|^2 \leq 64.
\]

Remark 2.1. In our numerical computations we stop the fixed point iteration (2.23) as soon as the distance between the iterates is small enough, i.e.,

\[
\| p^{n+1} - p^n \| \leq \varepsilon \cdot \| p^{n+1} \|,
\]

where \( \varepsilon \) is a chosen error bound.

Then, in summary, to minimize (2.19) we apply the following algorithm

**Algorithm (P)**

- For an initial \( p^0 = 0 \), iterate (2.23) until (2.25);
- Set \( P^1_K(\Delta^{-1}g) = \lambda \nabla \cdot p^0 \), where \( \hat{n} \) is the first iterate of (2.23) which fulfills (2.25);
- Compute a minimizer \( u \) of (2.19) by
  \[
  u = g + \Delta \left( P^1_K(\Delta^{-1}g) \right) = g + \Delta \left( \lambda \nabla \cdot p^0 \right).
  \]
The second step is to use the presented algorithm for (2.19) in order to solve (1.1), i.e.,
\[ \min_u \{ J(u) = |Du| (\Omega) + \frac{1}{2\lambda} \| Tu - g \|_2^2 \}. \]

To do so we first approximate a minimizer of (1.1) iteratively by a sequence of minimizers of, what we call, surrogate functionals \( J_s \). This approach is inspired by similar methods used, e.g., in [7], [21].

Let \( \tau > 0 \) be a fixed stepsize. Starting with an initial condition \( u^0 = g \), we solve for \( k \geq 0 \)
\begin{equation}
(2.26) \quad u^{k+1} = \arg\min_u J_s(u, u^k) = |Du| (\Omega) + \frac{1}{2\tau} \| u - u^k \|_2^2 + \frac{1}{2\lambda} \| u - (g + (Id - T)u^k) \|_2^2.
\end{equation}

Note that a function \( u \) for which \( J_s(u, u^k) = J(u) \), i.e., a fixed point of \( J_s \), is a potential minimizer for \( J \). A rigorous derivation of convergence properties is still missing and is a matter of future investigation. Note however that in the case of image inpainting, i.e., \( T = \chi_{\Omega \setminus D} \) and \( g \) is replaced by \( \chi_{\Omega \setminus D} g \), the optimality condition of (2.26) indeed describes a fourth order diffusion inside of the inpainting domain \( D \). Hence, in this case, minimizing (2.26) rather describes the behaviour of solutions of the inpainting approach (1.12) than directly minimizing (1.1), cf. also Subsection 1.3 and especially (1.13). Despite the missing theory, the numerical results obtained by using this scheme for inpainting issues suggest its correct asymptotic behaviour, see Section 3.2.

Now, the corresponding optimality condition to (2.26) reads
\[ 0 \in \partial |Du| (\Omega) + \frac{1}{\tau} \Delta^{-1} (u - u^k) + \frac{1}{\lambda} \Delta^{-1} (u - (g + (Id - T)u^k)), \]
which can be rewritten as
\[ \Delta^{-1} \left( \frac{g_1 - u}{\tau} + \frac{g_2 - u}{\lambda} \right) \in \partial |Du| (\Omega), \]
where \( g_1 = u^k, g_2 = g + (Id - T)u^k \). Setting
\[ g = \frac{g_1 + g_2}{\lambda + \tau}, \quad \mu = \frac{\lambda + \tau}{\lambda \tau}, \]
we end up with the same inclusion as (2.20), i.e.,
\[ \frac{\Delta^{-1} (g - u)}{\mu} \in \partial |Du| (\Omega), \]
and Algorithm (P) for solving (2.19) can be directly applied.

3. Applications

In this section we present applications of our new algorithm for solving (1.1) for image denoising, decomposition and inpainting, and present numerical results obtained. For comparison, we also present results for the \( TV - L^2 \) model in [36] on the same images.

Now, in order to compute the minimizer \( u \) of (1.1), we have the following algorithm.

Algorithm \( TV - H^{-1} \):
- In the case \( T = Id \) directly apply Algorithm (P) to compute a minimizer of (1.1).
- In the case \( T \neq Id \) iterate (2.26) by solving Algorithm (P) in every iteration step until the two subsequent iterates \( u^k \) and \( u^{k+1} \) are sufficiently close.

Note that in our numerical examples \( e \) in (2.25) is chosen to be \( 10^{-4} \).
3.1. Image Denoising and Decomposition. In the case of image denoising and image decomposition the operator $T = I_d$ and thus Algorithm (P) can be directly applied. For image denoising the signal to noise ratio (SNR) is computed as

$$SNR = 20 \log \left( \frac{\langle g \rangle}{\sigma} \right),$$

with $\langle g \rangle$ the average value of the pixels $g_{i,j}$ and $\sigma$ the standard deviation of the noise. For our numerical results the parameter $\lambda$ in (1.1) was chosen so that the best residual-mean-squared-error (RMSE) is obtained. We define the RMSE as

$$RMSE = \frac{1}{NM} \sqrt{\sum_{1 \leq i \leq N, 1 \leq j \leq M} (u_{i,j} - \hat{u}_{i,j})^2},$$

where $\hat{u}$ is the original image without noise, cf. [29]. Numerical examples for image denoising with $TV-H^{-1}$ minimization and their comparison with the results obtained by the $TV-L^2$ approach are presented in Figures 3-6. In both examples the superiority of the $TV-H^{-1}$ minimization approach with respect to the separation of noise and edges is clearly visible.

We also apply (1.1) for texture removal in images, i.e., image decomposition, and compare the numerical results with those of the $TV-L^2$ approach, cf. Figure 7. The cartoon-texture decomposition in this example works better in the case of $TV-H^{-1}$ minimization, since this approach differentiates between small oscillations and strong edges, better than the $TV-L^2$ approach.

3.2. Image Inpainting. In order to apply our algorithm to $TV-H^{-1}$ inpainting we follow the method of surrogate functionals from Section 2. In fact it turns out that a fixed point of the corresponding optimality condition of (2.26) with $T = \chi_{\Omega\setminus D}$ and $g$ is replaced by $\chi_{\Omega\setminus D}g$ is indeed a stationary solution of (1.12). This approach is also motivated by the fixed point approach used in [13] in order to prove existence of a stationary solution of (1.12). Hence a stationary solution to (1.12) can be computed iteratively by the following algorithm: Take $u^0 = g$, with any trivial (zero) expansion to the inpainting domain, and solve for $k \geq 0$

$$\min_{u \in BV (\Omega)} \left\{ |Du| (\Omega) + \frac{1}{2\tau} \|u - u^k\|_{-1}^2 + \frac{1}{2\lambda} \|u - \chi_{\Omega\setminus D}g - (1 - \chi_{\Omega\setminus D})u^k\|_{-1}^2 \right\} \to u^{k+1},$$

$$
Figure 4. Denoising results for the image of a horse in Figure 3. Results from the $TV - L^2$ denoising model compared with $TV - H^{-1}$ denoising with $\lambda = 0.05$ for both.
TOTAL VARIATION MINIMIZATION WITH AN $H^{-1}$ CONSTRAINT

Figure 5. Image of the roof of a house in Scotland and its noisy version with additive white noise

for positive iteration steps $\tau > 0$. Now, as before, let $g_1 = u^k$, $g_2 = \chi_{\Omega_1} \Omega g + (1 - \chi_{\Omega_1} \Omega) u^k$ and

$$g = \frac{g_1 + g_2 \tau}{\lambda + \tau},$$

$$\mu = \frac{\lambda + \tau}{\lambda + \tau},$$

then we end up with the same inclusion as (2.20), i.e.,

$$\frac{\Delta^{-1}(g - u)}{\mu} \in \partial |Du|(\Omega),$$

and Algorithm (P) can be directly applied. Compare Figure 8 for a numerical example.

3.3. Domain Decomposition for TV-$H^{-1}$ Minimization. As already discussed in Section 1.4 one of the drawbacks of using TV-$H^{-1}$ minimization in applications is its slow numerical performance. Also with our new algorithm, i.e., Algorithm (P), we are conditioned to timesteps $\tau \leq 1/64$. If now additionally the data dimension is large, e.g., when we have to process 2D images of high resolution, of sizes $3000 \times 3000$ pixels for instance, or even 3D image data, each iteration step itself is computationally expensive and we are far away from real-time computations. Total variation approaches for example already turned out to be an effective tool for the reconstruction of medical images as the ones gained from PET (Positron Emission Tomography) measurements (cf. [28], for instance). These imaging approaches need to deal with 3D or even 4D image data (including time dependence) in a fast and robust way. Motivated by this we were thinking about alternative ways to reduce the dimensionality of the data and hence speed up the reconstruction process. This is the interest of this section where we present a domain decomposition approach to be applied to TV-$H^{-1}$ minimization models.

Domain decomposition methods were introduced as techniques for solving partial differential equations based on a decomposition of the spatial domain of the problem into several subdomains. This means that instead of solving one big problem, i.e., the equation on the whole domain $\Omega$, a number of local subproblems is solved, i.e., the equation on $\Omega_1, \ldots, \Omega_N$ where $\Omega = \bigcup_{i=1}^N \Omega_i$. Thereby the main outcome of this method is a dimension reduction of the problem whose impact is shown when computed on parallel processors. In [27] the authors proposed a domain decomposition algorithm which is applicable to TV-$L^2$ minimization. The difficulty of doing this for total variation approaches is that solutions may be discontinuous, and hence their correct treatment on the interfaces of the domain decomposition patches isn’t straightforward. In the following we will roughly present a method how to modify the algorithm in [27] for TV-$H^{-1}$ minimization. In particular, we want to investigate splittings of the domain $\Omega$ into arbitrary nonoverlapping domains $\Omega = \Omega_1 \cup \Omega_2$ with $\Omega_1 \cap \Omega_2 = \emptyset$. For simplicity,
Figure 6. Denoising results for the image of the roof in Figure 5. Results from the $TV - L^2$ denoising model with $\lambda = 0.05$ compared with $TV - H^{-1}$ denoising with $\lambda = 0.01$.

we shall only consider a splitting in two domains. However note that, as in [27], the algorithm can be easily extended to more than two subdomains.

Following the notation in [27], let $\mathcal{H} = L^2(\Omega)$ and $V_i = L^2(\Omega_i)$, where $\mathcal{H} = V_1 \oplus V_2$. Let further be $\mathcal{H}^\psi = BV(\Omega)$ and $V_i^\psi = BV(\Omega_i)$, $i = 1, 2$. We take the norm in $BV(\Omega)$ to be $\|\cdot\|_{BV(\Omega)} = |D\cdot|(\Omega) + \|\cdot\|_{L^2(\Omega)}$. Let us further denote $\Pi_{V_i}$ the orthogonal projection onto $V_i$. This setting is the
Figure 7. Decomposition into cartoon and texture of an synthetic image. Results from the $TV-L^2$ model with $\lambda = 1$ and $TV-H^{-1}$ minimization with $\lambda = 0.1$. 
same as the one for $TV - L^2$ minimization in Example 2.1 in [27], with the only difference that the fidelity term in (1.1) is minimized in the weaker $H^{-1}$ norm instead of the norm in $L^2(\Omega)$. However, considering (1.1) within this setting, we fulfill all the necessary properties, assumed in [27], in order to apply their theory to our case. In what follows we will sketch the main arguments for this.

Now, we want to minimize $J$ in (1.1) by the following alternating algorithm: Pick an initial $V_1 \oplus V_2 \ni u_0 = u_0^0 + u_0^2 := u_0^0 \in BV(\Omega)$, for example $u_0^0 = 0$, and iterate

$$
\begin{align*}
\frac{u_1^{n+1}}{\approx} &= \arg\min_{u_1 \in V_1} J(u_1 + u_2) \\
\frac{u_2^{n+1}}{\approx} &= \arg\min_{u_2 \in V_2} J(u_1^{n+1} + u_2) \\
\frac{u_1^{n+1}}{:=} &= u_1^{n+1} + u_2^{n+1}.
\end{align*}
$$

In [27] this algorithm is implemented by solving the subspace minimization problems via an oblique thresholding iteration. Lets present the strategy in short for the $TV - H^{-1}$ case. The subproblem on $\Omega_1$ reads

$$
(3.28) \quad \arg\min_{u_1 \in V_1} \left\{ J(u_1 + u_2) = |D(u_1 + u_2)|(\Omega) + \frac{1}{2}\lambda \| Tu_1 - (g - Tu_2) \|^2_{-1} \right\}.
$$

As in Section 2 in (2.26), we introduce a sequence of surrogate functionals for this subminimization problem, i.e., let $u_0^0 = 0$, for $k \geq 0$ let

$$
J_k^*(u_1 + u_2, u_k^1) = |D(u_1 + u_2)|(\Omega) + \frac{1}{2\tau} \| u_1 - u_k^1 \|^2_{-1} + \frac{1}{2\lambda} \| u_1 - (g - Tu_2 + (Id - T)u_k^1) \|^2_{-1}.
$$

We want to realize an approximate solution to (3.28) by using the following algorithm: For $u_0^1 \in BV(\Omega_1)$,

$$
(3.29) \quad \frac{u_{1}^{k+1}}{=} = \arg\min_{u_1 \in V_1} J_k^*(u_1 + u_2, u_k^1), \quad k \geq 0.
$$

Problem (3.29) can be reformulated as

$$
\frac{u_{1}^{k+1}}{\approx} = \arg\min_{u \in BV(\Omega)} \{ F(u), \Pi_{V_2}(u) = 0 \},
$$

with $F(u) = J_k^*(u + u_2, u_k^1)$. Like in [27] we are going to use the following theorem:

**Theorem 3.1.** (Theorem 4.3 in [27]) We consider the following problem

$$
(3.30) \quad \arg\min_{x \in V} \{ F(x) : G(x) = 0 \},
$$

where $G : V \to \mathbb{R}$ is a bounded linear operator on $V$. If $F$ is continuous in a point of $\ker G$ and $G^*$ has closed range in $V$, then a point $x_0 \in \ker G$ is an optimal solution of (3.30) if and only if

$$
\partial F(x_0) \cap \text{Range} \ G^* \neq \emptyset.
$$
Now, since \( L^2(\Omega) \subset L^2(\mathbb{R}^2) \subset H^{-1}(\Omega) \) (by zero extensions of functions on \( \Omega \) to \( \mathbb{R}^2 \)), our functional \( F \) is continuous on \( V_1^p \subset V_1 = \ker \Pi_{V_1} \) in the norm topology of \( BV(\Omega) \). Further \( \Pi_{V_2}[BV(\Omega)] \) is a bounded and surjective map with closed range in the norm topology of \( BV(\Omega) \), i.e., \( (\Pi_{V_2}[BV(\Omega)])^\ast \) is injective and the Range \( (\Pi_{V_2}[BV(\Omega)])^\ast \cong (BV(\Omega))' \) is closed. By applying Theorem 3.1, we know that the optimality of \( u_1^{k+1} \) is equivalent to the existence of an \( \eta \in \text{Range}(\Pi_{V_2}[BV(\Omega)])^\ast \cong (BV(\Omega))' \) such that

\[-\eta \in \partial_{BV(\Omega)} F(u_1^{k+1}),\]

where \( \partial_{BV(\Omega)} \) denotes the subdifferential of \( F \) on \( BV(\Omega) \). Now

\[\partial_{BV(\Omega)} F(u_1^{k+1}) = \frac{1}{\mu} \Delta^{-1}(u_1^{k+1} - z) + \partial_{BV(\Omega)} \left| D(u_1^{k+1} + u_2) \right|(\Omega)\]

where

\[z = \frac{z_1 \lambda + z_2 \tau}{\lambda + \tau}, \quad \mu = \frac{\lambda \tau}{\lambda + \tau}, \]

with \( z_1 = u_1^k, \ z_2 = g - Tw_2 + (Id - T)u_1^k \). Then the optimality of \( u_1^{k+1} \) is equivalent to

\[0 \in \frac{1}{\mu} \Delta^{-1}(u_1^{k+1} - z) + \eta + \partial_{BV(\Omega)} \left| D(u_1^{k+1} + u_2) \right|(\Omega)\]

The latter is equivalent to

\[u_1^{k+1} + u_2 \in \partial_{BV(\Omega)} \left| D_\lambda \right|(\Omega)^\ast \left( \frac{1}{\mu} \Delta^{-1}(z - u_1^{k+1}) - \eta \right),\]

i.e.

\[\frac{u_2 + z}{\mu} \in \frac{z - u_1^{k+1}}{\mu} + \frac{1}{\mu} \partial_{BV(\Omega)} \left| D_\lambda \right|(\Omega)^\ast \left( \frac{1}{\mu} \Delta^{-1}(z - u_1^{k+1}) - \eta \right).\]

By letting \( w = \Delta^{-1}(z - u_1^{k+1})/\mu - \eta \) we have

\[0 \in (\Delta(w + \eta) - (u_2 + z)/\mu) + \frac{1}{\mu} \partial_{BV(\Omega)} \left| D_\lambda \right|(\Omega)^\ast(w),\]

or, in other words, \( w \) is a minimizer of

\[\frac{\|w - (\Delta^{-1}(u_2 + z)/\mu - \eta)\|_{H_0^1(\Omega)}^2}{2} + \frac{1}{\mu} |D_\lambda| \left(\Omega\right)^\ast(w).\]

Following the same procedure as in Section 2 we get that

\[w = \mathbb{P}_K(\Delta^{-1}(u_2 + z)/\mu - \eta),\]

where \( \mathbb{P}_K \) denotes the orthogonal projection on \( K \) over \( H_0^1(\Omega) \) like in Section 2. Then a minimizer \( u_1^{k+1} \) of (3.29) can be computed as

\[u_1^{k+1} = -\Delta \left( (Id - \mathbb{P}_K^\ast) (\Delta^{-1}(z + u_2) - \mu \eta) - u_2 \right).\]

By applying \( \Pi_{V_2} \) to both sides of the latter equality we get

\[0 = \mu \Delta \eta + \Pi_{V_2} \left[ \Delta \mathbb{P}_K^\ast (\Delta^{-1}(u_2 + z) - \mu \eta) \right].\]

Assuming necessary zero boundary conditions on \( \partial \Omega \) the resulting fixed point equation for \( \eta \) reads

\[\eta = \frac{1}{\mu} \Pi_{V_2} \left[ \mathbb{P}_K^\ast (\mu \eta - \Delta^{-1}(u_2 + z)) \right].\]

Like in [27] this fixed point can be computed via the iteration

\[\eta^0 \in V_2, \quad \eta^{m+1} = \frac{1}{\mu} \Pi_{V_2} \left[ \mathbb{P}_K^\ast (\mu \eta^m - \Delta^{-1}(u_2 + z)) \right], \quad m \geq 0.\]
In sum we solve (1.1) by the alternating subspace minimizations: Pick an initial \( V_1 \oplus V_2 \ni u^0 \), \( L^1 + u^0, M^2 := u^0 \in B_{V_1}(\Omega) \), for example \( u^0 = 0 \), and iterate

\[
\begin{aligned}
&\left\{ \begin{array}{l}
  u_1^{n+1,0} = u_1^n, \\
  u_1^{n+1,\ell+1} = \arg\min_{u_1 \in V_1} \mathcal{J}_1^s(u_1 + u_2^{n,M}, u_1^{n+1,\ell}) \\
  u_2^{n+1,0} = u_2^n, \\
  u_2^{n+1,m+1} = \arg\min_{u_2 \in V_2} \mathcal{J}_2^s(u_1^{n+1,L} + u_2^{n+1,M}, u_2^{n+1,m})
\end{array} \right. \quad \ell = 0, \ldots, L - 1 \\
u^{n+1} := u_1^{n+1,L} + u_2^{n+1,M},
\end{aligned}
\]

where each subminimization problem is computed by the oblique thresholding algorithm

**Oblique Thresholding for TV \(- H^{-1}\) Minimization**

Start with an initial condition \( u^0 = 0 \) and iterate

\[
u^{k+1} = -\Delta \left( I - \mathbb{P}_{1,\mu K} \right) \left( \Delta^{-1}(z + u_j) - \mu \eta - u_j \right), \quad i = 1, 2, \ i \neq j,
\]

where \( \eta \) in (3.32) is computed via the fixed point iteration

\[
\eta^{m+1} = \frac{1}{\mu} \Pi_{V_2} \left[ \mathbb{P}_{1,\mu K} \left( \mu \eta^m - \Delta^{-1}(u_j + z) \right) \right], \quad m \geq 0.
\]

As before the projection \( \mathbb{P}_{1,\mu K} \) is computed by **Algorithm (P)**. Note that, as in [27], a parallel version of (3.31) can be obtained by a slight modification of the update, i.e., \( u^{n+1} := (u^n + u_1^{n+1,L} + u_2^{n+1,M})/2 \).

In Figures 9 and 10 the given image was divided in four subdomains, marked by the red lines, and the image is inpainted via (3.31).

**Acknowledgments**

This paper was initiated during my participation at the workshop *Singularities in nonlinear evolution phenomena and applications*, held at the Centro di Ricerca Matematica Ennio De Giorgi on May 26-29, 2008, organized by Sisto Baldo, Matteo Novaga and Giandomenico Orlandi. The author thanks the research center and the workshop organizers for their hospitality and a very interesting workshop. Further, CBS would like to thank Massimo Fornasier for a sequence of fruitful discussions and Antonin Chambolle for his very helpful suggestions concerning the generalization of his algorithm to the \( H^{-1} \) constrained case.
The author also acknowledges the following funds for their financial support. CBS is partially supported by the project WWTF Five senses-Call 2006, *Mathematical Methods for Image Analysis and Processing in the Visual Arts* project nr. CI06 003, by the FFG project *Erarbeitung neuer Algorithmen zum Image Inpainting* project nr. 813610 and by KAUST (King Abdullah University of Science and Technology).

**References**


Carola-Bibiane Schönlieb

Department of Applied Mathematics and Theoretical Physics (DAMTP),
Centre for Mathematical Sciences,
Wilberforce Road,
Cambridge CB3 0WA,
United Kingdom.

E-mail address: c.b.s.schonlieb@damtp.cam.ac.uk