INVERSE TOTAL VARIATION FLOW

M. BURGER[1], K. FRICK[2,†], S. OSHER[3], AND O. SCHERZER[2]

[1]Institut für Numerische und Angewandte Mathematik.
Fachbereich Mathematik und Informatik, Westfälische Wilhelms Universität Münster.
Einsteinstrasse 62, D 48149 Münster, Germany.

[2]Institut für Informatik, Universität Innsbruck.
Technikerstrasse 21a, A 6020, Innsbruck, Austria.

[3]Department of Mathematics, UCLA.
520 Portola Plaza, Los Angeles, CA 90095-1555.

martin.burger@uni-muenster.de, {klaus.frick, otmar.scherzer}@uibk.ac.at, sjo@math.ucla.edu

† Corresponding author.

Abstract. In this paper we analyze iterative regularization with the Bregman distance of the total variation semi norm. Moreover, we prove existence of a solution of the corresponding flow equation as introduced in [8] in a functional analytical setting using methods from convex analysis. The results are generalized to variational denoising methods with L^p-norm fit-to-data terms and Bregman distance regularization term. For the associated flow equations well–posedness is derived using recent results on metric gradient flows from [2]. In contrast to previous work the results of this paper apply for the analysis of variational denoising methods with the Bregman distance under adequate noise assumptions. Besides from the theoretical results we introduce a level set technique based on Bregman distance regularization for denoising of surfaces and demonstrate the efficiency of this method.

1. Introduction

There are at least two evolutionary concepts based on partial differential equations for data filtering: *Scale space methods* with parabolic partial differential equations approximate data u^δ (for instance images), defined on a domain Ω, by the solution of

$$\frac{\partial u}{\partial t} = -A(u) \quad \text{in } \Omega \tag{1a}$$

$$u(0) = u^\delta \tag{1b}$$

$$\frac{\partial u}{\partial \nu} = 0 \quad \text{on } \partial\Omega \tag{1c}$$

with $A(u) := -\mathrm{div}\left(g(|\nabla u|)\frac{\nabla u}{|\nabla u|}\right)$ and ν the normal vector to the boundary of $\partial\Omega$. For given $t_0 > 0$, $u(t_0)$ is considered an approximation and filtered version of u^δ. The value of t_0 controls the amount of filtering.

In *semi group* theory the solution of (1) is defined iteratively. The initialization consists in setting $u_0 = u^\delta$. Then, with $\hat{g}$ satisfying $g(x) = \hat{g}'(x)$ the functionals

$$F_k(u) = \frac{1}{2}\|u - u_k\|^2_{\mathrm{L}^2} + \alpha \int_\Omega \hat{g}(|\nabla u|)\,\mathrm{d}x, \qquad k = 0, 1, \dots, \tag{2}$$

Date: November 13, 2006.

are minimized iteratively and the minimizer is denoted by u_{k+1}. Classical semi group theory assumes that $\hat{g} : [0, \infty) \to [0, \infty)$ is convex, satisfies a growth and a coercivity condition, in which case the minimizer of the functional F_k is unique, belongs to a Sobolev space and satisfies $u_{k+1} = (\mathrm{Id} + \alpha A)^{-1}(u_k)$. From the semi group generation theorem (cf. [12]) it follows that the limit

$$u(t) = \lim_{N \to \infty} \left(\mathrm{Id} + \frac{t}{N} A \right)^{-N} (u^\delta) \tag{3}$$

exists for all $t \in [0, \infty)$ and that it solves (1). The operator $\left(\mathrm{Id} + \frac{t}{N} A\right)^{-N}$ corresponds to making N iterative regularization steps with a regularization parameter $\alpha = t/N$, that is, $u^N = \left(\mathrm{Id} + \frac{t}{N} A\right)^{-N} (u^\delta)$. In other words iterative regularization and the implicit Euler method correspond if the regularization parameter α and the time discretization Δt are identified. The correspondence between diffusion filtering and iterated regularization has been analyzed numerically and analytically in [25].

Inverse scale space methods as introduced in [18] are defined as the semigroups corresponding to iterative regularization

$$u_{k+1} = \mathrm{argmin} \left\{ \frac{1}{2} \left\| u - u^\delta \right\|_{\mathrm{L}^2}^2 + \alpha \int_\Omega \hat{g}(|\nabla u - \nabla u_k|) \, \mathrm{d}x \right\} . \tag{4}$$

Here one typically initializes $u_0 = 0$ or $u_0 = \int_\Omega u^\delta \, \mathrm{d}x$ and u_{k+1} satisfies the Euler-Lagrange equation

$$u_{k+1} - u^\delta = \alpha \mathrm{div} \left(g(|\nabla u - \nabla u_k|) \frac{\nabla u - \nabla u_k}{|\nabla u - \nabla u_k|} \right) . \tag{5}$$

In particular for $\hat{g}(x) = \frac{1}{p} |x|^p$ it follows that

$$u_{k+1} - u^\delta = \alpha \mathrm{div} \left(|\nabla (u_{k+1} - u_k)|^{p-2} \nabla (u_{k+1} - u_k) \right) , \tag{6a}$$

$$u(0) = u_0 . \tag{6b}$$

Identifying the regularization parameter α and a time discretization Δt via

$$\alpha = \frac{1}{(\Delta t)^{p-1}} \tag{7}$$

equation (6) can be considered as an implicit time step of the following flow equation

$$\begin{aligned} u - u^\delta &= \Delta^p \left(\frac{\partial u}{\partial t} \right) , \\ u(0) &= u^\delta . \end{aligned} \tag{8}$$

Here $\Delta^p(u) = \mathrm{div} \left(|\nabla u|^{p-2} \nabla u \right)$ is the p-Laplacian. It is obvious that (7) degenerates for $p = 1$ and thus for $\alpha \to \infty$ it cannot be claimed that $\Delta t \to 0$. Consequently (8) is not properly defined for $p = 1$.

For $\hat{g}(x) = x$, and $\hat{u} = u - u^\delta$, (6) the inverse scale space method becomes *Showalter's method* (see e.g. [16]) which for denoising and gradient evaluation applications reads as follows

$$\triangle^{-1} \hat{u} = \frac{\partial \hat{u}}{\partial t} \tag{9a}$$

$$\hat{u}(0) = -u^\delta \; (+u_0) \tag{9b}$$

where Δ^{-1} is the solution operator corresponding to Laplace's equation with homogeneous Neumann boundary data.

The name inverse scale space is motivated from the fact that

$$\lim_{t\to\infty} u(t) = u^\delta, \quad \lim_{t\to 0^+} u(t) = u_0 \ . \tag{10}$$

That is the method "inverts" the axiom of fidelity in scale space theory, which asserts that $u(t) \to u^\delta$ for $t \to 0$.

Scale space methods are very well investigated and analyzed. In particular for image processing *total variation regularization* (cf. [23]), which consists in minimization of the functional

$$F(u) = \frac{1}{2} \left\| u - u^\delta \right\|_{\mathrm{L}^2}^2 + \alpha |\mathrm{D}u|(\Omega) \,, \tag{11}$$

and the associated *total variation flow equation*

$$\frac{\partial u(t)}{\partial t} = \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) \quad \text{in } \Omega \tag{12a}$$

$$u(0) = u^\delta \tag{12b}$$

$$\frac{\partial u}{\partial \nu} = 0 \quad \text{on } \partial\Omega \tag{12c}$$

have attracted much attention, since they allow for discontinuous solutions, which is considered an inherent property of image intensities (due to the appearance of edges). For some analytical work related to total variation regularization we refer for instance to [1] and [10]). Existence and uniqueness of a *weak* solution of (12) is shown in [3].

Inverse scale space methods have not attracted as much attention. The reason, as can be seen from (7), is that the original concept does not apply for $p = 1$ and hence, in particular, does not allow for discontinuous minimizers. On the other hand, they are attractive alternatives due to the possible derivation of stopping criteria dependent on the noise and possible generalizations for rather general tasks in imaging and inverse problems. Total variation inverse scale space methods have been derived in [22] employing the concept of *Bregman distance regularization.* This method consists in computing first a minimizer u_1 of (11). The updates are determined successively by calculating

$$u_{k+1} = \mathop{\mathrm{argmin}}_{u\in\mathrm{L}^2} \left\{ \frac{1}{2} \left\| u - u^\delta \right\|_{\mathrm{L}^2}^2 + \alpha \left(|\mathrm{D}u|(\Omega) - \langle s, u \rangle \right) \right\} \ , \tag{13}$$

where s is an element of the subgradient of the total variation semi norm in u_1. Introducing the Bregman distance with respect to $|\mathrm{D}u|(\Omega)$ defined by

$$D(u, \tilde{u}) := |\mathrm{D}u|(\Omega) - |\mathrm{D}\tilde{u}|(\Omega) - \int_\Omega s(u - \tilde{u})$$

(cf Definition 1) allows to characterize u_{k+1} in (13) as

$$u_{k+1} = \mathop{\mathrm{argmin}}_{u\in L^2(\Omega)} \left\{ \frac{1}{2} \left\| u - u^\delta \right\|_{\mathrm{L}^2}^2 + \alpha D(u, u_k) \right\} \ . \tag{14}$$

As we show in this paper, this iterative process satisfies a *discrete inverse scale space property*, that is $u_k \to u^\delta$ for $k \to \infty$. Moreover, we derive an according flow equation, which satisfies the inverse scale space property.

This paper is organized as follows. In Section 2 we analyze (13), showing well-posedness (that is existence of minimizers) and study the asymptotic behavior of the sequence of minimizers $\{u_k\}_{k\in\mathbb{N}}$. The analysis allows to generalize results from [22] to be applicable for data perturbed by (for instance white) noise. In Section 3 we analyze the flow equation corresponding to (14) (that is when $\alpha \to \infty$) which has first been introduced in [8]. We show existence of a solution in a functional analytical setting using methods from convex analysis and results from [2]. The flow equation turns out to satisfy (10). In Section 5 we motivate iterative total variation Bregman distance regularization as a variational method for penalizing a distance between level sets of u_{k+1} and u_k, motivating the use of this method for image and surface denoising.

Notation. In this paper we use the following notation and make the assumptions, which are not stated explicitly any further afterwards:

(1) $\Omega \subseteq \mathbb{R}^n$ is open, bounded with Lipschitz boundary $\partial\Omega$.
(2) We denote by $\mathcal{D}(\Omega) = \mathcal{C}_c^\infty(\Omega)$ the space of continuously differentiable functions with compact support in Ω.
(3) For $p \in (1,\infty)$ p_* denotes its conjugate, which satisfies $\frac{1}{p} + \frac{1}{p_*} = 1$. Let μ be a positive measure on Ω and $m \in \mathbb{N}$. $\mathrm{L}^p(\Omega, \mathbb{R}^m; \mu)$ denotes the space of p times μ-integrable functions $f : \Omega \to \mathbb{R}^m$ with the norm

$$\|f\|_{\mathrm{L}^p(\Omega,\mathbb{R}^m;\mu)} = \left(\int_\Omega |f|^p \, \mathrm{d}\mu \right)^{\frac{1}{p}} .$$

If μ is the n dimensional Lebesgue measure we simply write $\mathrm{L}^p(\Omega)^m$ and if $m = 1$ we write $\mathrm{L}^p(\Omega)$. In this case we abbreviate $\|f\|_{\mathrm{L}^p(\Omega)} = \|f\|_{\mathrm{L}^p}$. The dual of $\mathrm{L}^p(\Omega, \mathbb{R}^m; \mu)$ is $\mathrm{L}^{p_*}(\Omega, \mathbb{R}^m; \mu)$. The dual pairing with respect to $\mathrm{L}^{p_*}(\Omega)$ and $\mathrm{L}^p(\Omega)$ is denoted by $\langle \cdot, \cdot \rangle$.
(4) $\mathrm{L}^\infty(\Omega, \mathbb{R}^n; \mu)$ denotes the space of essentially bounded functions.
(5) For $p \in (1,\infty)$ we define

$$K^p(\Omega) = \overline{\left\{ \mathrm{div}\,(z) : z \in \mathcal{D}(\Omega)^n, \|z\|_{\mathrm{L}^\infty(\Omega)^n} \leq 1 \right\}}, \tag{15}$$

where the closure is taken with respect to the strong topology on $\mathrm{L}^p(\Omega)$. Note that $K^p(\Omega)$ is convex and therefore also closed w.r.t. the weak topology in $\mathrm{L}^p(\Omega)$.

2. Iterative regularization with Bregman Distances

For the analysis of iterative Bregman distance regularization, motivated in (13), we require several results from convex analysis and functional analysis, which are reviewed in the beginning of this section.

2.1. Review on Results from Convex Analysis.

Definition 1. Let $\mathcal{J} : \mathrm{L}^p(\Omega) \to \mathbb{R} \cup \{+\infty\}$ be a convex and proper functional.

(1) An element $s \in \mathrm{L}^{p_*}(\Omega)$ lies in the *subgradient* $\partial\mathcal{J}(u_0)$ of $\mathcal{J}$ at $u_0 \in \mathrm{L}^p(\Omega)$ if

$$\mathcal{J}(u) - \mathcal{J}(u_0) - \langle s, u - u_0 \rangle \geq 0 \text{ for all } u \in \mathrm{L}^p(\Omega) .$$

(2) The *Bregman* distance of $\mathcal{J}$ at $u, \tilde{u} \in \mathrm{L}^p(\Omega)$ with respect to $s \in \partial\mathcal{J}(\tilde{u}) \subseteq \mathrm{L}^{p_*}(\Omega)$ is defined by

$$D^s_{\mathcal{J}}(u, \tilde{u}) := \mathcal{J}(u) - \mathcal{J}(\tilde{u}) - \langle s, u - \tilde{u} \rangle . \tag{16}$$

(3) The *Legendre-Fenchel* conjugate of a convex functional $\mathcal{J} : \mathrm{L}^p(\Omega) \to \mathbb{R} \cup \{+\infty\}$ is the functional $\mathcal{J}^* : \mathrm{L}^{p_*}(\Omega) \to \mathbb{R} \cup \{+\infty\}$

$$u^* \mapsto \mathcal{J}^*(u^*) := \sup_{u \in \mathrm{L}^p(\Omega)} \{\langle u^*, u \rangle - \mathcal{J}(u)\} .$$

From the definition of the subdifferential it follows that $u \mapsto D^s_J(u, \tilde{u})$ is non negative.

Example 1. The duality mapping defined by

$$\mathfrak{J}_p : \mathrm{L}^p(\Omega) \to \mathrm{L}^{p_*}(\Omega) , \quad u \to |u|^{p-2} u$$

satisfies:

(1)
$$\int_\Omega \mathfrak{J}_p(u)u \, \mathrm{d}x = \|u\|_{\mathrm{L}^p}^p = \|\mathfrak{J}_p(u)\|_{\mathrm{L}^{p_*}}^{p_*} .$$

(2) $\mathfrak{J}_p^{-1} = \mathfrak{J}_{p_*}$. In particular $\mathfrak{J}_2 \equiv \mathrm{Id}$.

(3) $\mathfrak{J}_p(u)$ is the subdifferential of $\|\cdot\|_{L^p}^p$ at $u \in L^p(\Omega)$, that is

$$\mathfrak{J}_p(u) = \partial \|\cdot\|_{\mathrm{L}^p}^p (u) . \tag{17}$$

Lemma 1. *Assume that* $p > 1$. *For* $u \in L^p(\Omega)$ *let*

$$J(u) := \sup_{\substack{z \in \mathcal{D}(\Omega)^n \\ \|z\|_{L^\infty(\Omega)^n} \leq 1}} \int_\Omega u \ div(z) \ dx = \begin{cases} |Du|(\Omega) & \text{if } u \in BV(\Omega) \\ +\infty & \text{else} \end{cases}$$

the total variation semi norm *of* u. *Then*

$$J^*(v) = \chi_{K^{p_*}(\Omega)}(v) = \begin{cases} 0 & \text{if } v \in K^{p_*}(\Omega) \\ +\infty & \text{else} . \end{cases} \tag{18}$$

Proof. Let $u \in \mathrm{L}^p(\Omega)$, then

$$\begin{aligned}\left(\chi_{K^{p_*}(\Omega)}\right)^* (u) &= \sup_{v \in K^{p_*}(\Omega)} \int_\Omega uv \ \mathrm{d}x \\ &= \sup_{\substack{z \in \mathcal{D}(\Omega)^n \\ \|z\|_{\mathrm{L}^\infty(\Omega)^n} \leq 1}} \int_\Omega u \ \mathrm{div}\,(z) \ \mathrm{d}x = J(u).\end{aligned}$$

Since the set $K^{p_*}(\Omega)$ is convex, the function $\chi_{K^{p_*}(\Omega)}$ is convex. Moreover, since $\chi_{K^{p_*}(\Omega)}$ is lower semicontinuous ([7, Ex. 2.8.2]) it follows that $\chi_{K^{p_*}(\Omega)}(v) = \left(\chi_{K^{p_*}(\Omega)}\right)^{**}(v) = J^*(v)$ (see e.g. [14, Chap. 1, Prop. 4.1]). ☐

2.2. **Properties of $K^p(\Omega)$.** As it turned out in the previous section, the Legendre Fenchel conjugate J^* of the total variation seminorm defined on $\mathrm{L}^p(\Omega)$ is the characteristic function of the set $K^{p_*}(\Omega)$. Y. Meyer [21] introduced the *G-norm* which has proven to be a feasable instrument to discribe this duality. We summarize the basic facts of this theory and provide some generalizations (see also [4]).

For $1 < p < \infty$ we introduce the subspaces

$$X_p(\Omega) = \{z \in \mathrm{L}^\infty(\Omega)^n \ : \ \mathrm{div}\,(z) \in \mathrm{L}^p(\Omega)\} . \tag{19}$$

From [3, Thm. C.3.] we know that there exists a linear trace operator $T_p : X_p(\Omega) \to \mathrm{L}^\infty(\partial\Omega)^n$ such that

$$\|T_p(z)\|_{\mathrm{L}^\infty(\partial\Omega)^n} \leq \|z\|_{\mathrm{L}^\infty(\Omega)^n}$$

and

$$\int_\Omega u \, \mathrm{div}\,(z) \ \mathrm{d}x + \int_\Omega \nabla uz \ \mathrm{d}x = \int_{\partial\Omega} T_p(z) u \ \mathrm{d}\mathcal{H}^{n-1} \tag{20}$$

for all $z \in X_p(\Omega)$ and $u \in \mathrm{W}^{1,1}(\Omega) \cap \mathrm{L}^{p_*}(\Omega)$. Moreover we define the space

$$\mathrm{L}^p_\diamondsuit(\Omega) = \left\{ v \in \mathrm{L}^p(\Omega) \ : \ \int_\Omega v \ \mathrm{d}x = 0 \right\} . \tag{21}$$

Proposition 1. *Let* $1 < p < \infty$, *then*

$$L^p_\diamondsuit(\Omega) = \overline{div\,(\mathcal{D}(\Omega)^n)} \tag{22}$$

where the closure is taken w.r.t. the strong topology on $L^p(\Omega)$.

Proof. Let $\{z_k\} \subseteq \mathcal{D}(\Omega)^n$ such that $\mathrm{div}\,(z_k) \to v$ w.r.t. the $\mathrm{L}^p(\Omega)$-norm. Then $\int_\Omega \mathrm{div}\,(z_k)\ \mathrm{d}x = 0$ for all k and thus

$$\int_\Omega v\ \mathrm{d}x = \int_\Omega \lim_{k\to\infty} \mathrm{div}\,(z_k)\ \mathrm{d}x = \lim_{k\to\infty} \int_\Omega \mathrm{div}\,(z_k)\ \mathrm{d}x = 0.$$

If conversely $v \in \mathrm{L}^p_\diamondsuit(\Omega)$ there exist a sequence $v_k \in \mathcal{C}^\infty(\overline{\Omega}) \cap \mathrm{L}^p_\diamondsuit(\Omega)$ such that $v_k \to v$ in $\mathrm{L}^p(\Omega)$ (From the density of $\mathcal{D}(\Omega)$ in $\mathrm{L}^p(\Omega)$ the existence of an approximating sequence $\{v_k\} \subseteq \mathcal{D}(\Omega)$ follows; Then $v_k - \frac{1}{|\Omega|}\int_\Omega v_k\ \mathrm{d}x \to v$). We abbreviate $\bar{p} = \max(p,n)$ and observe that $v_k \in \mathrm{L}^{\bar{p}}_\diamondsuit(\Omega)$ for all $k \in \mathbb{N}$. From [6, Thm.3] it follows that there exists $z_k \in \mathcal{C}(\overline{\Omega})^n \cap \mathrm{W}^{1,\bar{p}}_0(\Omega)^n$ such that

$$\mathrm{div}\,(z_k) = v_k.$$

Therefore we can approximate each z_k by an element $\tilde{z}_k \in \mathcal{D}(\Omega)^n$ such that

$$\|\mathrm{div}\,(\tilde{z}_k - z_k)\|_{\mathrm{L}^p} \le C\,\|\mathrm{div}\,(\tilde{z}_k - z_k)\|_{\mathrm{L}^{\bar{p}}} \le \frac{1}{k}$$

for a constant $C > 0$. This implies that $\mathrm{div}\,(\tilde{z}_k) \to v$. □

The *G-norm* of an element v as introduced by Meyer [21] and Aubert and Aujol [4] is defined as the infimum of $\|z\|_{\mathrm{L}^\infty(\Omega)^n}$ taken over all z such that $\mathrm{div}\,(z) = v$. In order to provied a generalization to our case it is important to note that the above proof does *not* imply existence of an element $z \in \mathrm{L}^\infty(\Omega)^n$ such that $\mathrm{div}\,(z) = v$. However from (20) it follows that at least

$$\ker(T_p) \subseteq \mathrm{L}^p_\diamondsuit(\Omega).$$

and [6, Thm. 3] proves

Lemma 2. *If $p \ge n$ we have that*

$$\mathit{ker}\,(T_p) = L^p_\diamondsuit(\Omega).$$

Definition 2. Let $1 < p < \infty$ and $v \in \mathrm{L}^p_\diamondsuit(\Omega)$. The G-norm of v is defined as

$$\|v\|_* = \inf\left\{\liminf_{k\to\infty} \|z_k\|_{\mathrm{L}^\infty(\Omega)^n} \ :\ \{z_k\} \subseteq \mathcal{D}(\Omega)^n,\ \lim_{k\to\infty} \|\mathrm{div}\,(z_k) - v\|_{\mathrm{L}^p} = 0\right\} \tag{23}$$

Moreover we introduce the space

$$\mathcal{L}^p_\diamondsuit(\Omega) = \left\{v \in \mathrm{L}^p_\diamondsuit(\Omega)\ :\ \|v\|_* < \infty\right\}. \tag{24}$$

The linearity of $\mathrm{div}\,(\cdot)$ and the norm properties of $\|\cdot\|_{\mathrm{L}^\infty(\Omega)^n}$ impliy that $\left(\mathcal{L}^p_\diamondsuit(\Omega), \|\cdot\|_*\right)$ is a normed space. As we will show in Proposition 2 below $\mathcal{L}^p_\diamondsuit(\Omega)$ contains those elements of $\mathrm{L}^p_\diamondsuit(\Omega)$ which can be represented as $\mathrm{div}\,(z)$ with $z \in \mathrm{L}^\infty(\Omega)^n$.

Proposition 2. *For every $v \in \mathcal{L}^p_\diamondsuit(\Omega)$ there exists $z \in \mathit{ker}\,(T_p)$ such that*

$$\mathit{div}\,(z) = v,\ \|z\|_{L^\infty(\Omega)^n} \le \|v\|_*.$$

Proof. We follow the proof of [4, Lemma 2.3]. Let $v \in \mathcal{L}^p_\diamondsuit(\Omega)$, i.e. $\|v\|_* < \infty$. From the definition of the G-norm (23) it follows that there exists a sequence $\{z_k\} \subseteq \mathcal{D}(\Omega)^n$ such that

$$\lim_{k\to\infty} \|\mathrm{div}\,(z_k) - v\|_{\mathrm{L}^p} = 0, \quad \lim_{k\to\infty} \|z_k\|_{\mathrm{L}^\infty(\Omega)^n} = \|v\|_*.$$

Then the boundedness of $\{z_k\}$ implies the existence of an element $z \in \mathrm{L}^\infty(\Omega)^n$ such that – up to an extraction of a subsequence – z_k converges weakly* to z in $\mathrm{L}^\infty(\Omega)^n$.

From (20) it follows that for every $\phi \in \mathcal{C}^\infty(\overline{\Omega})$

$$\int_\Omega \phi v \, \mathrm{d}x = \lim_{k\to\infty} \int_\Omega \phi \operatorname{div}(z_k) \, \mathrm{d}x = \lim_{k\to\infty} - \int_\Omega \nabla\phi \, z_k \, \mathrm{d}x$$
$$= - \int_\Omega \nabla\phi \, z \, \mathrm{d}x = \int_\Omega \phi \operatorname{div}(z) \, \mathrm{d}x - \int_{\partial\Omega} \phi T_p(z) \, \mathrm{d}\mathcal{H}^{n-1}.$$

Choosing $\phi \in \mathcal{D}(\Omega)$ yields $\operatorname{div}(z) = v$ in the sense of distributions and since $v \in \mathrm{L}^p(\Omega)$ it follows that $\operatorname{div}(z) \in \mathrm{L}^p(\Omega)$. For an arbitrary $\phi \in \mathcal{C}^\infty(\overline{\Omega})$ it then follows that

$$\int_{\partial\Omega} \phi \, T_p(z) \, \mathrm{d}\mathcal{H}^{n-1} = 0$$

which shows that $T_p(z) = 0$ $\mathcal{H}^{n-1}$- a.e. on $\partial\Omega$ (application of Whitney's Extension Theorem [17, Chap. 6.5 Thm. 1]). Finally, weak* lower semicontinuity of $\|\cdot\|_{\mathrm{L}^\infty(\Omega)^n}$ implies

$$\|z\|_{\mathrm{L}^\infty(\Omega)^n} \leq \lim_{k\to\infty} \|z_k\|_{\mathrm{L}^\infty(\Omega)^n} = \|v\|_* \,.$$

□

A straightforward consequence of (23) is that

$$K^p(\Omega) \subseteq \left\{ v \in \mathcal{L}^p_\Diamond(\Omega) \ : \ \|v\|_* \leq 1 \right\}.$$

Indeed, if $\{z_k\} \subseteq \mathcal{D}(\Omega)^n$ such that $\|z_k\|_{\mathrm{L}^\infty(\Omega)^n} \leq 1$ and $\operatorname{div}(z_k) \to v \in \mathcal{L}^p_\Diamond(\Omega)$ then

$$\|v\|_* \leq \liminf_{k\to\infty} \|z_k\|_{\mathrm{L}^\infty(\Omega)^n} \leq 1.$$

We further investigate the relation between the topology induced by $\|\cdot\|_*$ and the weak topology on $\mathrm{L}^p_\Diamond(\Omega)$.

Example 2. Let $\Omega = [0, \pi]$ and $p = p_* = 2$. We define

$$z_k(x) = \frac{1}{2k} \sin\left[(2k)^2 x\right].$$

Since $z_k(0) = z_k(\pi) = 0$, $v_k = z_k' \in \mathcal{L}^2_\Diamond(\Omega)$ and we have that

$$\|v_k\|_* = \|z_k\|_{\mathrm{L}^\infty} \leq \frac{1}{k}$$

i.e. $v_k \to 0$ w.r.t. $\|\cdot\|_*$. However $v_k = 2k \cos\left[(2k)^2 x\right]$ is unbounded in $\mathrm{L}^2(\Omega)$.

This example shows that convergence w.r.t. $\|\cdot\|_*$ does not imply weak convergence w.r.t. $\mathrm{L}^p(\Omega)$. However under additional assumtions the result is true:

Proposition 3. *Let $\{v_k\}_{k\in\mathbb{N}}$ be a sequence in $\mathcal{L}^p_\Diamond(\Omega)$ and $v \in \mathcal{L}^p_\Diamond(\Omega)$ such that*

$$\sup_{k\in\mathbb{N}} \|v_k\|_{L^p} < \infty, \ \lim_{k\to\infty} \|v_k - v\|_* = 0$$

Then $v_k \rightharpoonup v$ in $L^p(\Omega)$.

Proof. Since $\{v_k\}$ is bounded in $\mathrm{L}^p(\Omega)$ we can extract a subequence $\{v_{k(l)}\}$ such that

$$v_{k(l)} \rightharpoonup \hat{v}, \quad \text{in } \mathrm{L}^p(\Omega) \tag{25}$$

for some $\hat{v} \in \mathrm{L}^p_\Diamond(\Omega)$. Moreover let $v \in \mathcal{L}^p_\Diamond(\Omega)$ such that $v_k \to v$ w.r.t. $\|\cdot\|_*$. Then – using Proposition 2 – we can choose $\{z_{k(l)}\} \subseteq \mathrm{L}^\infty(\Omega)^n$ such that $T_p(z_{k(l)}) = 0$, $\operatorname{div}\left(z_{k(l)}\right) = v_{k(l)} - v$ and $\left\|z_{k(l)}\right\|_{\mathrm{L}^\infty(\Omega)^n} \leq \left\|v_{k(l)} - v\right\|_*$ for all $l \in \mathbb{N}$. This together with (20) shows that $\int_\Omega u(v_{k(l)}) - v) \, \mathrm{d}x = - \int_\Omega \nabla u z_{k(l)} \, \mathrm{d}x$ for all $u \in \mathrm{W}^{1,1}(\Omega) \cap \mathrm{L}^{p_*}(\Omega)$ and therefore

$$\lim_{k\to\infty} \int_\Omega u v_{k(l)} \, \mathrm{d}x = \int_\Omega u v \, \mathrm{d}x.$$

This together with (25) shows $\hat{v} = v$. Consequentely every subsequence of $\{v_k\}$ has in turn a weakly convergent subsequence with limit v. This implies that $v_k \rightharpoonup v$. □

We conclude this section with two results that are based on the analysis in [3, Appendix C].

Lemma 3. *For all $u \in BV(\Omega) \cap L^{p_*}(\Omega)$ and $v \in \mathcal{L}^p_\diamond(\Omega)$ there exists a $|Du|$ - measurable function $\theta(v, Du, \cdot) : \Omega \to \mathbb{R}$ that satisfies*

$$\int_\Omega uv \ dx = \int_\Omega \theta(v, Du, x) \ d|Du|$$

and

$$\|\theta(v, Du, \cdot)\|_{L^\infty(\Omega, \mathbb{R}, |Du|)} \leq \|v\|_* .$$

Moreover let $f : \mathbb{R} \to \mathbb{R}$ be Lipschitz continuous and strictly increasing, then

$$\theta(v, D(f \circ u), x) = \theta(v, Du, x),$$

for almost every $x \in \Omega$ w.r.t. to both the measure $|Du|$ and the measure $|D(f \circ u)|$.

Proof. According to Proposition 2 there exists a $z \in \ker(T_p)$ such that $\mathrm{div}\,(z) = v$ and $\|z\|_{\mathrm{L}^\infty(\Omega)^n} \leq \|v\|_*$. From Theorems C.7 and C.9 in [3] it follows that there exists a function $\theta(v, \mathrm{D}u, \cdot) \in \mathrm{L}^\infty(\Omega, \mathbb{R}, |\mathrm{D}u|)$ such that

$$\int_\Omega uv \ \mathrm{d}x = \int_\Omega u \mathrm{div}\,(z) \ \mathrm{d}x = \int_\Omega \theta(v, \mathrm{D}u, x) \ \mathrm{d}|\mathrm{D}u|$$

and

$$\|\theta(v, \mathrm{D}u, \cdot)\|_{\mathrm{L}^\infty(\Omega, \mathbb{R}, |\mathrm{D}u|)} \leq \|z\|_{\mathrm{L}^\infty(\Omega)^n} \leq \|v\|_*$$

The last assertion follows directly from [3, Cor. 16]. □

Corollary 1. *For all $u \in BV(\Omega) \cap L^{p_*}(\Omega)$ and $v \in \mathcal{L}^p_\diamond(\Omega)$ we have that*

$$\left| \int_\Omega uv \ dx \right| \leq |Du|(\Omega) \|v\|_* .$$

Note that for those $u \in \mathrm{BV}(\Omega) \cap \mathrm{L}^{p_*}(\Omega)$ satisfying $\mathfrak{J}_{p_*}(u) \in \mathcal{L}^p_\diamond(\Omega)$ Corollary 1 implies that

$$\|u\|^{p_*}_{\mathrm{L}^{p_*}} = \int_\Omega u \mathfrak{J}_{p_*}(u) \ \mathrm{d}x \leq |\mathrm{D}u|(\Omega) \|\mathfrak{J}_{p_*}(u)\|_* . \tag{26}$$

According to [21], *simple functions* are defined as those satisfying (26) with equality.

2.3. **Analysis of Iterative Bregman Distance Regularization.** Iterative Bregman distance regularization, as motivated in (13), is defined as follows

$$u_{k+1} := \operatorname*{argmin}_{u \in \mathrm{L}^p(\Omega)} I(\alpha; u, u_k) , \text{ where } I(\alpha; u, \tilde{u}) := \left\{ \frac{1}{p} \left\| u - u^\delta \right\|^p_{\mathrm{L}^p} + \alpha D^{\tilde{v}}_J(u, \tilde{u}) \right\} ,$$

where $\tilde{v}$ is an element of the subgradient of J at $\tilde{u}$. The algorithm is uniquely defined up to the choice of the element $\tilde{v}$ in the subgradient of J at $\tilde{u}$. Following and extending the work in [22], which considered $p = 2$, we make the effective choice of the subgradient element as follows:

Algorithm 1.
- Let $u_0 \in \mathrm{BV}(\Omega)$ and $v_0 \in \partial J(u_0)$.
- For $k = 0, 1, \ldots$

$$\begin{aligned} u_{k+1} &:= \operatorname*{argmin}_{u \in \mathrm{L}^p(\Omega)} I(\alpha; u, u_k). \\ v_{k+1} &:= v_k + \frac{1}{\alpha} \mathfrak{J}_p(u^\delta - u_{k+1}) . \end{aligned}$$

The goals of this subsection are to prove well–posedness of Algorithm 1 and to investigate the asymptotic behavior of the iterated minimizers u_k. By showing that $u_k \to u^\delta$ we justify the terminology *inverse scale space method*.

Theorem 1. *Assume that $u^\delta \in L^p(\Omega)$. Then for each $k \in \mathbb{N}$ there exists a unique minimizer $u_k \in L^p(\Omega) \cap BV(\Omega)$ of I_k and a subgradient $v_k \in J(u_k) \subseteq L^{p_*}(\Omega)$ such that*

$$\alpha v_k + \mathfrak{J}_p(u_k - u^\delta) = \alpha v_{k-1}. \tag{27}$$

Proof. Let $\tilde{u} \in \mathrm{L}^p(\Omega)$ and $s \in \partial J(\tilde{u})$. We show weak lower semicontinuity and coercivity of I_k. Then, existence of a minimizer follows from [13, Chap. 3, Thm. 1.1].

Since both $u \to J(u)$ (see for instance [1, Thm. 2.3]) and $u \to \langle s, u \rangle$ are weakly lower semicontinuous on $\mathrm{L}^p(\Omega)$, the Bregman distance

$$u \mapsto D_J^s(u, \tilde{u})$$

is weakly lower semicontinuous. Therefore $I(\alpha; \cdot, u_k)$ is weakly lower semicontinuous on $\mathrm{L}^p(\Omega)$. It remains to show that $I(\alpha; \cdot, u_k)$ is coercive on $\mathrm{L}^p(\Omega)$, that is

$$I(\alpha; u, u_k) \geq \beta \|u\|_{\mathrm{L}^p}^p + \gamma$$

for some $\beta > 0$ and $\gamma \in \mathbb{R}$. We verify the assertion for the functional $I(\alpha; \cdot, u_0)$. For $k \geq 1$ the assertion can be proven analogously taking into account that $u_k \in \mathrm{L}^p(\Omega) \cap \mathrm{BV}(\Omega)$ and $v_k \in \partial J(u_k)$. Since J is convex, $v_0 \in \partial J(u_0)$ is equivalent to

$$J(u) \geq J(u_0) + \langle v_0, u - u_0 \rangle \text{ for all } u \in \mathrm{L}^p(\Omega).$$

Therefore,

$$\begin{aligned} \frac{1}{p} \left\| u - u^\delta \right\|_{\mathrm{L}^p}^p + \alpha(J(u) - \langle v_0, u \rangle) &\geq \frac{1}{p} \left\| u - u^\delta \right\|_{\mathrm{L}^p}^p + \alpha(J(u_0) - \langle v_0, u_0 \rangle) \\ &\geq \frac{1}{p} \left| \|u\|_{\mathrm{L}^p} - \left\| u^\delta \right\|_{\mathrm{L}^p} \right|^p + \alpha(J(u_0) - \langle v_0, u_0 \rangle). \end{aligned}$$

Thus $I(\alpha; \cdot, u_0)$ is L^p-coercive and we can apply [13, Chap. 3, Thm. 1.1] and conclude that there exists a minimizer $u_1 \in \mathrm{BV}(\Omega) \cap \mathrm{L}^p(\Omega)$ which satisfies the Euler-Lagrange equation

$$v_1 := \frac{\mathfrak{J}_p(u^\delta - u_1)}{\alpha} \in \partial J(u_1).$$

□

Lemma 4. *For every $k \in \mathbb{N}$ we have*

$$\left\| u_{k+1} - u^\delta \right\|_{L^p} \leq \left\| u_k - u^\delta \right\|_{L^p}. \tag{28}$$

Proof. (See [22, Prop. 3.2]). Since the Bregman distance is nonnegative we have

$$\begin{aligned} \frac{1}{p} \left\| u_{k+1} - u^\delta \right\|_{\mathrm{L}^p}^p &\leq \frac{1}{p} \left\| u_{k+1} - u^\delta \right\|_{\mathrm{L}^p}^p + \alpha D_J^{v_{k+1}}(u_{k+1}, u_k) \\ &= I(\alpha; u_{k+1}, u_k) \leq I(\alpha; u_k, u_k) = \frac{1}{p} \left\| u_k - u^\delta \right\|_{\mathrm{L}^p}^p. \end{aligned}$$

□

Using Lemma 1 the dual formulation of Algorithm 1 can be derived. We consider the dual functional of I with respect to $v \in \mathrm{L}^{p_*}(\Omega)$, defined as follows:

$$I^*(\alpha; v, \tilde{v}) := \frac{1}{p_*} \|v - \tilde{v}\|_{\mathrm{L}^{p_*}}^{p_*} + \frac{1}{\alpha^{p_*-1}} \left(J^*(v) - \left\langle u^\delta, v \right\rangle \right), \quad \tilde{v} \in \mathrm{L}^{p_*}(\Omega). \tag{29}$$

Theorem 2. *Assume that $p > 1$ and $u^\delta \in L^p(\Omega)$. Then u_k, v_k as defined in Algorithm 1 satisfy*

$$v_k = \underset{v \in K^{p_*}(\Omega)}{\mathit{argmin}}\ I^*(\alpha; v, v_{k-1}) \in K^{p_*}(\Omega)\ . \tag{30}$$

Moreover, in particular, for $p = 2$

$$u_k = (u^\delta + \alpha v_{k-1}) - \alpha \pi_{K^2(\Omega)} \left(\frac{u^\delta}{\alpha} + v_{k-1} \right), \tag{31}$$

where $\pi_{K^2(\Omega)}(u)$ denotes the orthogonal projection of $u \in L^2(\Omega)$ onto $K^2(\Omega)$.

Proof. The proof is along the lines of [9] where the dual formulation of total variation regularization has been derived.

From Lemma 1 it follows that

$$I^*(\alpha; v, \tilde{v}) = \begin{cases} \frac{1}{p_*} \|v - \tilde{v}\|_{\mathrm{L}^{p_*}}^{p_*} - \frac{1}{\alpha^{p_*-1}} \left\langle u^\delta, v \right\rangle & \text{if } v \in K^{p_*}(\Omega) \\ +\infty & \text{else} \end{cases} \tag{32}$$

The functional I^* is strictly convex and weakly lower semicontinuous with respect to v and thus $I^*(\alpha; \cdot, v_{k-1})$ attains a unique minimizer $\tilde{v}_k$. From Lemma 1 it follows that $\tilde{v}_k \in K^{p_*}(\Omega)$. It remains to show that $v_k = \tilde{v}_k$. From the definition of v_k in Algorithm 1 and Theorem 1 it follows that

$$v_k = v_{k-1} - \frac{1}{\alpha} \mathfrak{J}_p(u_k - u^\delta) \in \partial J(u_k)\ . \tag{33}$$

Then, from the duality relation (see for instance [14]) it follows that

$$u_k \in \partial J^* \left(v_{k-1} - \frac{1}{\alpha} \mathfrak{J}_p(u_k - u^\delta) \right). \tag{34}$$

Moreover, since (33) is equivalent to

$$-\alpha(v_k - v_{k-1}) = \mathfrak{J}_p(u_k - u^\delta)$$

it follows by applying Example 1 (3) that

$$\mathfrak{J}_{p_*}(v_k - v_{k-1}) \alpha^{p_*-1} - u^\delta = -u_k. \tag{35}$$

Combination of (33), (34) and (35) shows that

$$0 \in \mathfrak{J}_{p_*}(v_k - v_{k-1}) \alpha^{p_*-1} - u^\delta + \partial J^*(v_k)\ . \tag{36}$$

Therefore, v_k minimizes the functional $I^*(\alpha; \cdot, v_{k-1})$, which together with the fact that the minimizer is unique implies that $v_k = \tilde{v}_k$.

For $p = 2$ minimization of $I^*(\alpha; \cdot, v_{k-1})$ is equivalent to minimizing the functional

$$v \to \frac{1}{2} \left\| v - \left(v_{k-1} + \frac{u^\delta}{\alpha} \right) \right\|_{\mathrm{L}^2}^2 + \frac{1}{\alpha} J^*(v)\ .$$

Therefore, from Lemma 1 it follows that $v_k = \pi_{K^2(\Omega)}(\frac{u^\delta}{\alpha} + v_{k-1})$ and together with (33) we see that

$$u_k = (u^\delta + \alpha v_{k-1}) - \alpha \pi_{K^2(\Omega)} \left(\frac{u^\delta}{\alpha} + v_{k-1} \right).$$

□

In the dual formulation (30) there exists an equation for v_k which is independent of u_k. In contrast to Algorithm (1), where the variables u_k and v_k are coupled.

We now show the *discrete inverse fidelity property* of Algorithm 1 which means that the sequence u_k approach the original (noisy) data as $k \to \infty$.

Theorem 3. *Let $u^\delta \in L^p(\Omega)$ and $\{u_k\}$ as defined in Algorithm 1. Then we have that*

$$\lim_{k\to\infty} \left\| u_k - u^\delta \right\|_{L^p} = 0.$$

Proof. Let $k > 1$. Since $u_k \in \partial J^*(v_k)$ we have that $\langle v_{k-1} - v_k, u_k \rangle \leq 0$ and consequentely that

$$\begin{aligned} \left\langle v_{k-1} - v_k, u^\delta \right\rangle &\leq \left\langle v_{k-1} - v_k, u^\delta - u_k \right\rangle \\ &= - \left\langle v_k - v_{k-1}, u^\delta - u_k \right\rangle \\ &= -\frac{1}{\alpha} \left\| u_k - u^\delta \right\|_{\mathrm{L}^p}^p \end{aligned}$$

where the last equality follows from (33). This estimate combined with Lemma 4 gives

$$\begin{aligned} \left\| u_k - u^\delta \right\|_{\mathrm{L}^p}^p &\leq \frac{1}{k} \sum_{j=1}^k \left\| u_j - u^\delta \right\|_{\mathrm{L}^p}^p \\ &\leq \frac{\alpha}{k} \sum_{j=1}^k \left\langle v_j - v_{j-1}, u^\delta \right\rangle \\ &= \alpha \left(\left\langle \frac{v_k}{k}, u^\delta \right\rangle - \left\langle \frac{v_0}{k}, u^\delta \right\rangle \right). \end{aligned}$$

Hence it suffices to show that $\frac{v_k}{k}$ converges weakly to 0 as $k \to \infty$. Since $v_k \in K^{p_*}(\Omega)$ (that is $\|v_k\|_* \leq 1$) for all $k > 0$ we immediately see that

$$\lim_{k\to\infty} \left\| \frac{v_k}{k} \right\|_* = 0.$$

Moreover, Lemma 4 and (33) applied iteratively imply that

$$\begin{aligned} \left\| \frac{v_k}{k} \right\|_{\mathrm{L}^{p_*}}^{p_*} &= \frac{1}{k} \left\| v_{k-1} - \frac{1}{\alpha} \mathfrak{J}_p (u_k - u^\delta) \right\|_{\mathrm{L}^{p_*}}^{p_*} \\ &= \frac{1}{k} \left\| v_0 - \frac{1}{\alpha} \sum_{j=1}^k \mathfrak{J}_p (u_j - u^\delta) \right\|_{\mathrm{L}^{p_*}}^{p_*} \\ &\leq \frac{\|v_0\|_{\mathrm{L}^{p_*}}^{p_*}}{k} + \frac{1}{\alpha k} \sum_{j=1}^k \left\| u_j - u^\delta \right\|_{\mathrm{L}^p}^p \\ &\leq \frac{\|v_0\|_{\mathrm{L}^{p_*}}^{p_*}}{k} + \frac{1}{\alpha} \left\| u_0 - u^\delta \right\|_{\mathrm{L}^p}^p . \end{aligned} \tag{37}$$

We finally apply Proposition 3 and get

$$\frac{v_k}{k} \rightharpoonup 0, \qquad \text{in } \mathrm{L}^{p_*}(\Omega).$$

□

3. Continuous Inverse Scale Space Flow

In this section we derive the gradient flow equation associated with (30) which is the dual formulation of the iterative regularization method introduced in Algorithm 1. The analysis is based on results from [2]. There the authors describe the explicit construction of solutions of gradient flow equations in metric spaces $(\mathcal{S}, d)$ w.r.t. functionals $\phi : \mathcal{S} \to \overline{\mathbb{R}}$. The analysis is very general and allows for weaker topologies than the metric d in order to show convergence of discrete solutions.

A *partition* of the interval $[0,\infty)$ is a sequence of positive numbers $\vec{\tau} = \{\tau^k\}_{k\in\mathbb{N}}$ such that

$$|\vec{\tau}| = \sup_{k\in\mathbb{N}} \tau^k < \infty$$

and the associated sequence $\{t^k_{\vec{\tau}}\}_{k\in\mathbb{N}_0}$ defined by $\tau^k = t^k_{\vec{\tau}} - t^{k-1}_{\vec{\tau}}$ satisfies

$$0 = t^0_{\vec{\tau}} < t^1_{\vec{\tau}} < \cdots < t^k_{\vec{\tau}} < \dots,$$
$$\sum_{j=1}^{\infty} \tau^j = +\infty.$$

For a given partition $\vec{\tau}$ and $u_0 \in \mathrm{BV}(\Omega)\cap \mathrm{L}^p(\Omega)$ resp. $v_0 \in \partial J(u_0)$ we set for $k = 1,2,\dots$

$$\begin{aligned} &U^0_{\vec{\tau}} = u_0, && V^0_{\vec{\tau}} = v_0 \\ &U^k_{\vec{\tau}} = \operatorname*{argmin}_{u\in \mathrm{L}^p(\Omega)} I\left(\frac{1}{\tau^k}; u, U^{k-1}_{\vec{\tau}}\right), && V^k_{\vec{\tau}} = \operatorname*{argmin}_{v\in \mathrm{L}^{p_*}(\Omega)} I^*\left(\frac{1}{\tau^k}; v, V^{k-1}_{\vec{\tau}}\right) \end{aligned} \tag{38}$$

The discrete values of $U^k_{\vec{\tau}}$ resp. $V^k_{\vec{\tau}}$ are extended to a piecewise constant function $\overline{U}_{\vec{\tau}}(t)$ resp. $\overline{V}_{\vec{\tau}}(t)$ for $t\in[0,\infty)$ defined by

$$\overline{U}_{\vec{\tau}}(t) = U^k_{\vec{\tau}}, \quad \overline{V}_{\vec{\tau}}(t) = V^k_{\vec{\tau}}, \tag{39}$$

whenever $t \in \left(t^{k-1}_{\vec{\tau}}, t^k_{\vec{\tau}}\right]$. In order to show existence of a limiting function we apply recent results in [2]. We adopt the notation therein and rewrite the dual functional

$$I^*(\alpha; v, \tilde{v}) = \frac{1}{p_*} d(v,\tilde{v})^{p_*} + \frac{1}{\alpha^{p_*-1}}\phi(v)$$

where $d(v,\tilde{v}) = \|v-\tilde{v}\|_{\mathrm{L}^{p_*}}$ and $\phi(v) = \chi_{K^{p_*}(\Omega)}(v) - \langle u^\delta, v\rangle$.

Theorem 4. *Assume that $\{\vec{\tau}_l\}_{l\in\mathbb{N}}$ is a sequence of partitions of $[0,\infty)$ such that $\lim_{l\to\infty}|\vec{\tau}_l| = 0$. Then there exists a function $v \in \mathcal{C}(0,\infty; K^{p_*}(\Omega))$, which is uniformly continuous, differentiable almost everywhere in $[0,\infty)$ and satisfies*

$$\|v'\|_{L^{p_*}} \in L^{p_*}_{loc}(0,\infty) \tag{40}$$

such that

$$v(t) = \lim_{l\to\infty} \overline{V}_{\vec{\tau}_l}(t) \text{ for all } t\in[0,\infty) \tag{41}$$

in the weak topology of $L^{p_*}(\Omega)$.

Proof. The weak $\mathrm{L}^{p_*}(\Omega)$ topology satisfies the topological assumptions in Section 2.1 in [2]. That is ϕ is weakly lower semicontinuous on $\mathrm{L}^{p_*}(\Omega)$ (condition 2.1a) and from [14, Chap. 3 Prop. 1.1] we see that

$$\inf_{v\in\mathrm{L}^{p_*}(\Omega)} I^*(\alpha; v, 0) \geq -\inf_{u\in\mathrm{L}^p(\Omega)} I(\alpha; u, 0) > -\infty,\ \forall\alpha>0.$$

which shows the coercivity assumption 2.1b. Weak compactness of $\mathrm{L}^{p_*}(\Omega)$ bounded subsets of sublevels of ϕ (assumption 2.1c) follows by the Banach - Alaoglu Theorem [20, Thm. 2.6.18]. Hence we can apply [2, Prop. 2.2.3] which directly proves the assertion. The fact that $v(t)\in K^{p_*}(\Omega)$ then follows from weak closedness of $K^{p_*}(\Omega)$ and (41). □

Proposition 4 (Growth property)**.** *Let v be as in Theorem 4. Then v satisfies*

$$\|v(t)\|^{p_*}_{L^{p_*}} \leq \|v_0\|^{p_*}_{L^{p_*}} + \left\|u_0 - u^\delta\right\|^p_{L^p} t, \quad \textit{for all } t\geq 0.$$

Proof. Let $k, l \in \mathbb{N}$ and $t \in \left(t_{\vec{\tau}_l}^{k-1}, t_{\vec{\tau}_l}^{k}\right]$. From (38) and (33) we conclude that

$$\begin{aligned}\overline{V}_{\vec{\tau}_l}(t) &= V_{\vec{\tau}_l}^{k}\\ &= V_{\vec{\tau}_l}^{k-1} - \tau_l^k \mathfrak{J}_p\left(U_{\vec{\tau}_l}^{k} - u^\delta\right)\\ &= v_0 - \sum_{j=1}^{k} \tau_l^j \mathfrak{J}_p\left(U_{\vec{\tau}_l}^{j} - u^\delta\right).\end{aligned}$$

From Lemma 4 it then follows that

$$\begin{aligned}\left\|\overline{V}_{\vec{\tau}_l}(t)\right\|_{\mathrm{L}^{p_*}}^{p_*} &\le \|v_0\|_{\mathrm{L}^{p_*}}^{p_*} + \sum_{j=1}^{k} \tau_l^k \left\|U_{\vec{\tau}_l}^{j} - u^\delta\right\|_{\mathrm{L}^p}^{p}\\ &\le \|v_0\|_{\mathrm{L}^{p_*}}^{p_*} + \left\|u_0 - u^\delta\right\|_{\mathrm{L}^p}^{p} t_{\vec{\tau}_l}^{k-1} + \left\|u_0 - u^\delta\right\|_{\mathrm{L}^p}^{p} \tau_l^k\\ &\le \|v_0\|_{\mathrm{L}^{p_*}}^{p_*} + \left\|u_0 - u^\delta\right\|_{\mathrm{L}^p}^{p} t + \mathcal{O}(|\vec{\tau}_l|).\end{aligned}$$

If $\vec{\tau}_l$ is a sequence such that $\overline{V}_{\vec{\tau}_l}(t) \rightharpoonup v(t)$ as in Theorem 4, weak lower semicontinuity of $\|\cdot\|_{\mathrm{L}^{p_*}}$ finally shows that

$$\|v(t)\|_{\mathrm{L}^{p_*}}^{p_*} \le \liminf_{t\to\infty} \left\|\overline{V}_{\vec{\tau}_l}(t)\right\|_{\mathrm{L}^{p_*}}^{p_*} \le \|v_0\|_{\mathrm{L}^{p_*}}^{p_*} + \left\|u_0 - u^\delta\right\|_{\mathrm{L}^p}^{p} t.$$

□

Remark 1 (The case $p = 2$)**.** In the Hilbert space setting we obtain an even sharper result as in Theorem 4. Note that the (unique) minimizer $\tilde{v}$ of $I^*(\alpha; v, w)$ also minimizes

$$\bar{I}^*(\alpha; v, w) = \frac{\alpha}{2}\|v - w\|_{\mathrm{L}^2}^2 + J^*(v) - \left\langle v, u^\delta\right\rangle.$$

Since for $v_0, v_1, w \in \mathrm{L}^2(\Omega)$ and $t \in [0,1]$ we have that

$$\|(1-t)v_0 + tv_1 - w\|_{\mathrm{L}^2}^2 = (1-t)\|v_0 - w\|_{\mathrm{L}^2}^2 + t\|v_1 - w\|_{\mathrm{L}^2}^2 - t(1-t)\|v_0 - v_1\|_{\mathrm{L}^2}^2$$

we obtain for $v_0, v_1, w \in K^2(\Omega)$ that

$$\begin{aligned}\bar{I}^*(\alpha; (1-t)v_0 + tv_1, w) &= \frac{\alpha}{2}\|(1-t)v_0 + tv_1 - w\|_{\mathrm{L}^2}^2 - \langle (1-t)v_0 + tv_1, w\rangle\\ &= \frac{\alpha}{2}\left((1-t)\|v_0 - w\|_{\mathrm{L}^2}^2 + t\|v_1 - w\|_{\mathrm{L}^2}^2\right)\\ &\quad - \frac{\alpha}{2}t(1-t)\|v_0 - v_1\|_{\mathrm{L}^2}^2 - (1-t)\left\langle v_0, u^\delta\right\rangle - t\left\langle v_1, u^\delta\right\rangle\\ &= (1-t)\bar{I}^*(\alpha; v_0, w) + t\bar{I}^*(\alpha; v1, w) - \frac{\alpha}{2}t(1-t)\|v_0 - v_1\|_{\mathrm{L}^2}^2.\end{aligned}$$

Hence $\bar{I}^*$ satisfies the convexity assumption [2, 4.0.1] and we can apply [2, Thm. 4.2.2.] to obtain that $\overline{V}_{\vec{\tau}(l)} \to v$ uniformly on each bounded interval $[0, T]$ as $l \to \infty$. Particularly it follows that

$$\lim_{l\to\infty} \overline{V}_{\vec{\tau}(l)} = v(t), \qquad \text{for all } t \in [0, \infty)$$

in the strong topology on $\mathrm{L}^2(\Omega)$.

So far we proved that the iterated minimizers of the dual problem of Algorithm 1 converge to a continuous function of time as the time discretization goes to zero (that is the regularization parameter α converges to $+\infty$). It remains to investigate under which conditions these functions satisfy gradient flow equations w.r.t. to ϕ. To this end we introduce the operator $A : K^{p_*}(\Omega) \to \mathrm{L}^p(\Omega)$ which is defined by

$$A(v) = \operatorname*{argmin}_{u\in\partial\phi(v)} \|u\|_{\mathrm{L}^p}. \tag{42}$$

and the *slope*

$$|A|(v) := \begin{cases} \|A(v)\|_{\mathrm{L}^p} & \text{if } v \in K^{p_*}(\Omega) \\ +\infty & \text{else} \end{cases} .$$

In order to show existence of a gradient flow for the dual problem (38), $|A|$ needs to satisfy a lower semicontinuity condition.

Proposition 5. *Let $v \in L^{p_*}(\Omega)$. Then*

$$|A|(v) = \inf \left\{ \liminf_{k\to\infty} |A|(v_k) \;:\; v_k \rightharpoonup v,\, \{v_k\} \subseteq K^{p_*}(\Omega) \right\}.$$

The right hand side of the above equation is referred to as the *relaxed slope* and generally does not coincide with the slope $|A|$. Proposition 5 follows directly from the closedness of the subgradient $\partial\phi$ (or equivalently of ∂J^*) according to [2, Lmm. 2.3.6.].

Lemma 5. *The subgradient $\partial J^* \subseteq L^{p_*}_\diamondsuit(\Omega) \times L^p(\Omega)$ is weakly - weakly closed, that is for all sequences $\{v_k\} \subseteq L^{p_*}_\diamondsuit(\Omega)$ and $\{u_k\} \subseteq L^p(\Omega)$ the following implication holds*

$$\left.\begin{aligned} \sup_{k\in\mathbb{N}} J^*(v_k) < \infty,\, v_k \rightharpoonup v \text{ in } L^{p_*}(\Omega) \\ u_k \in \partial J^*(v_k),\, u_k \rightharpoonup u \text{ in } L^p(\Omega) \end{aligned}\right\} \Rightarrow u \in \partial J^*(v).$$

Proof. Since $J^*(v_k) < \infty$ it follows that $v_k \in K^{p_*}(\Omega)$. Therefore, since $u_k \in \partial J^*(v_k)$ we have that $\langle u_k, w - v_k\rangle \le 0$ for all $w \in K^{p_*}(\Omega)$ that is

$$J(u_k) = \langle u_k, v_k\rangle < \infty.$$

From weak convergence of $\{v_k\}$ and $\{u_k\}$ it follows that there exists a constant $C > 0$ such that $\|u_k\|_{\mathrm{L}^p} \|v_k\|_{\mathrm{L}^{p_*}} \le C$ for all $k > 0$ and consequently

$$\sup_{k\in\mathbb{N}} J(u_k) = \sup_{k\in\mathbb{N}} \int_\Omega u_k v_k \,\mathrm{d}x \le C. \tag{43}$$

In other words the sequence $\{u_k\}$ is bounded in $\mathrm{BV}(\Omega)$. From compact imbedding $\mathrm{BV}(\Omega) \hookrightarrow \mathrm{L}^1(\Omega)$ we conclude that every subsequence of $\{u_k\}$ has a strongly L^1 convergent subsequence with limit u. Thus $u_k \to u$ strongly in $\mathrm{L}^1(\Omega)$ and $u \in \mathrm{BV}(\Omega)$.

Let $\lambda \ge 0$. We introduce the truncation operator $S_\lambda(r) := (\lambda - (\lambda - |r|)^+)\,\mathrm{sign}(r)$ and set $S_\lambda^\varepsilon = (S_\lambda * G_\varepsilon)$ where G_ε denotes the Gaussian kernel with standard deviation $\varepsilon > 0$. Note that $S_\lambda^\varepsilon : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous (with Lipschitz constant ≤ 1) and strictly increasing. Lemma 3 implies that for every $k > 0$ there exists a $|\mathrm{D}u_k|$ - measurable function $\theta_k(v_k, \mathrm{D}u_k, \cdot) : \Omega \to \mathbb{R}$ such that $\theta_k(v_k, \mathrm{D}u_k, \cdot) \le 1$ $|\mathrm{D}u_k|$ - a.e. $x \in \Omega$,

$$\int_\Omega v_k u_k \,\mathrm{d}x = \int_\Omega \theta_k(v_k, \mathrm{D}u_k, x) \,\mathrm{d}|\mathrm{D}u_k| \tag{44}$$

and

$$\theta_k(v_k, \mathrm{D}(S_\lambda^\varepsilon(u_k)), x) = \theta_k(v_k, \mathrm{D}u_k, x), \quad |\mathrm{D}u_k| - \text{ a.e.} \tag{45}$$

Note that (45) also holds for $|\mathrm{D}S_\lambda^\varepsilon(u_k)|$ - almost every $x \in \Omega$. Since $\theta_k(v_k, \mathrm{D}u_k, \cdot) \le 1$ $|\mathrm{D}u_k|$ - a.e. and $\langle v_k, u_k\rangle = J(u_k)$ we conclude from (44) that $\theta_k(v_k, \mathrm{D}u_k, x) = 1$ for $|\mathrm{D}u_k|$ - almost every $x \in \Omega$ and consequently from (45)

$$\int_\Omega v_k S_\lambda^\varepsilon(u_k) \,\mathrm{d}x = \int_\Omega \theta_k(v_k, \mathrm{D}(S_\lambda^\varepsilon(u_k)), x) \,\mathrm{d}|\mathrm{D}S_\lambda^\varepsilon(u_k)| = J\left(S_\lambda^\varepsilon(u_k)\right).$$

Since $\|u_k - u\|_{\mathrm{L}^1} \to 0$ it follows that $\|S_\lambda^\varepsilon(u_k) - S_\lambda^\varepsilon(u)\|_{\mathrm{L}^1} \to 0$ as $k \to \infty$ from the Lipschitz continuity of S_λ^ε and $S_\lambda^\varepsilon(u_k)$ is uniformly bounded in $\mathrm{L}^\infty(\Omega)$ (by λ).

The Vitali Convergence Theorem (see e.g. [15, Thm. VI.5.6]) hence yields $\lim_{k\to\infty}\|S_\lambda^\varepsilon(u_k)-S_\lambda^\varepsilon(u)\|_{\mathrm{L}^{p'}}=0$ for all $p'<\infty$ and since $\langle\cdot,\cdot\rangle$ is strongly - weakly continuous on $\mathrm{L}^p(\Omega)\times\mathrm{L}^{p_*}(\Omega)$ we get that

$$\lim_{k\to\infty} J\left(S_\lambda^\varepsilon(u_k)\right)=\lim_{k\to\infty}\int_\Omega S_\lambda^\varepsilon(u_k)v_k\ \mathrm{d}x=\int_\Omega S_\lambda^\varepsilon(u)v\ \mathrm{d}x\le J\left(S_\lambda^\varepsilon(u)\right).$$

Together with the weak lower semicontinuity of J this implies that $J\left(S_\lambda^\varepsilon(u)\right)=\langle S_\lambda^\varepsilon(u),v\rangle$ and thus $S_\lambda^\varepsilon(u)\in\partial J^*(v)$. Since $S_\lambda^\varepsilon\to S_\lambda$ uniformly on $\mathbb{R}$ as $\varepsilon\to 0$ we can choose $\varepsilon(\lambda)>0$ such that

$$\sup_{x\in\mathbb{R}}\left|S_\lambda^{\varepsilon(\lambda)}-S_\lambda(x)\right|<\lambda^{-1}.$$

and therefore

$$\begin{aligned}2^{-p}\left\|S_\lambda^{\varepsilon(\lambda)}(u)-u\right\|_{\mathrm{L}^p}^p&\le\left\|S_\lambda^{\varepsilon(\lambda)}(u)-S_\lambda(u)\right\|_{\mathrm{L}^p}^p+\|S_\lambda(u)-u\|_{\mathrm{L}^p}^p\\&\le|\Omega|\,\lambda^{-p}+\int_{\{|u|\ge\lambda\}}|u|^p\ \mathrm{d}x.\end{aligned}$$

Therefore $S_\lambda^{\varepsilon(\lambda)}(u)\to u$ in $\mathrm{L}^p(\Omega)$. Then, from strong closedness of $\partial J^*(v)$ (see e.g. [14, Thm. I.5.1]) it follows that $u\in\partial J^*(v)$. □

Theorem 5. *Let v be as in Theorem 4. Then*

$$\mathfrak{J}_{p_*}(v') = -A(v(t)) \,. \tag{46a}$$

$$v(0) = v_0. \tag{46b}$$

for a.e. $t\in[0,\infty)$.

Proof. From Proposition 5 we know, that the slope $|A|$ coincides with its relaxed slope, that is the weakly-lower semicontinuous envelope of $|A|$ (adopting the terminology of [2] this means that the relaxed slope is a *strong upper gradient* for ϕ). Moreover $u_0\in\mathrm{BV}(\Omega)\cap\mathrm{L}^p(\Omega)$ and $v_0\in\partial J(u_0)$ implies $v_0\in D(\phi)$. Hence we can directly apply [2, Thm 2.3.3.] which proves the assertion.

□

It is rather straight-forward to see that $A(v)$ provides a subgradient of the functional $F:\mathrm{L}^{p_*}(\Omega)\to\mathbb{R}$,

$$F(v):=-\int_\Omega u^\delta v\ \mathrm{d}x+J^*(v), \tag{47}$$

and hence, (46) can be interpreted as a gradient flow for F in $\mathrm{L}^{p_*}(\Omega)$. We shall return to the use of F as a Lyapunov functional for the flow in the discussion of its multiscale properties.

From the definition of the operator A it follows that for a solution v of (46) there exists a unique function u satisfying

$$A(v(t))=u(t)-u^\delta \text{ and } u(t)\in\partial J^*(v(t)) \tag{48}$$

for all $t\in[0,\infty)$ and from duality we conclude that $v(t)\in\partial J(u(t))$. From these results it is evident that (46) is equivalent to

$$\mathfrak{J}_{p_*}\left(v'\right)(t)=u^\delta-u(t), \qquad v(t)\in\partial J(u(t)), \tag{49a}$$

$$v(0)=v_0 \qquad u(0)=u_0. \tag{49b}$$

We now focus on basic properties of the solutions of (49).

Proposition 6 (Regularity)**.** *Assume that* $u^\delta \in L^p(\Omega)$*. Then, the solutions* (u,v) *of (49) satisfy*

$$v \in \mathcal{C}([0,\infty), L^{p_*}(\Omega))\,, \quad u \in L^\infty([0,\infty), L^p(\Omega))\,.$$

Moreover, $u(t) \in BV(\Omega)$ *for all* $t \in [0,\infty)$*.*

Proof. The property of v follows from Theorem 4. Moreover, from the definition of $K^{p_*}(\Omega)$ and the fact that $v(t) \in K^{p_*}(\Omega)$ it follows that $\chi_{K^{p_*}(\Omega)}(v(t)) = 0$ for all $t \in [0,\infty)$. Application of [2, Rem. 1.3.3.] shows that

$$\|v'(t)\|_{\mathrm{L}^{p_*}}^{p_*} = \langle u^\delta, v'(t)\rangle \le \left\|u^\delta\right\|_{\mathrm{L}^p} \cdot \|v'(t)\|_{\mathrm{L}^{p_*}}\,, \qquad \text{a.e. in } [0,\infty) \tag{50}$$

and consequently

$$\|v'(t)\|_{\mathrm{L}^{p_*}} \le \left\|u^\delta\right\|_{\mathrm{L}^p}^{\frac{1}{p_*-1}}\,. \tag{51}$$

Thus the boundedness of $u(t)$ follows from (49). It remains to show that $u(t)$ has finite total variation. To this end we choose $0 \le t < \infty$ and note that

$$\begin{aligned}\partial J^*(v(t)) &= \{u \in \mathrm{L}^p(\Omega) : \langle u, w\rangle \le \langle u, v(t)\rangle \text{ for all } w \in K^{p_*}(\Omega)\}\\ &= \{u \in \mathrm{L}^p(\Omega) : \sup_{w\in K^{p_*}(\Omega)} \langle u, w\rangle \le \langle u, v(t)\rangle\}\\ &= \{u \in \mathrm{L}^p(\Omega) : J(u) = \langle u, v(t)\rangle\}.\end{aligned}$$

This shows that

$$J(u(t)) \le \|u(t)\|_{\mathrm{L}^p} \cdot \|v(t)\|_{\mathrm{L}^{p_*}} < \infty.$$

□

Theorem 6 (Uniqueness)**.** *If* $p = 2$*, then (49) has a unique solution* (v,u)*.*

Proof. The mapping $v \to \chi_{K^2(\Omega)}(v) - \langle u^\delta, v\rangle$ is convex and therefore

$$\partial\left(\chi_{K^2(\Omega)}(v) - \langle u^\delta, v\rangle\right) = \partial\chi_{K^2(\Omega)}(v) - u^\delta$$

is monotone (see e.g. [7]). That is

$$\langle u_1 - u_2, v_1 - v_2\rangle \ge 0 \text{ for all } v_i \in K^2(\Omega)\,, u_i \in \partial\chi_{K^2(\Omega)}(v_i) - u^\delta, i = 1,2.$$

Assume that there exist two solutions $v, \hat{v}$ of (46), that is,

$$\begin{aligned}-v'(t) &\in \partial\chi_{K^2(\Omega)}(v(t)) - u^\delta\,,\\ -\hat{v}'(t) &\in \partial\chi_{K^2(\Omega)}(\hat{v}(t)) - u^\delta\,.\end{aligned}$$

From the monotonicity of the right hand sides it follows that

$$\langle -v'(t) + \hat{v}'(t), v(t) - \hat{v}(t)\rangle \ge 0$$

and therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}\|v(t) - \hat{v}(t)\|_{\mathrm{L}^2}^2 = \langle v'(t) - \hat{v}'(t), v(t) - \hat{v}(t)\rangle \le 0.$$

This shows that $v = \hat{v}$. □

So far the results have shown that iterating the dual of iterative Bregman distance regularization (30) gives a implicit Euler scheme for the flow equation (46). It remains to investigate the relation between the piecewise constant functions $\overline{U}_{\vec{\tau}_l}$ and the function u introduced in (48).

Theorem 7. *Let* $\{\vec{\tau}_l\}_{l\in\mathbb{N}}$ *be a sequence of partitions of* $[0,\infty)$ *such that* $\lim_{l\to\infty}|\vec{\tau}_l| = 0$*. Then*

$$\lim_{l\to\infty}\left\|\overline{U}_{\vec{\tau}_l}(t) - u(t)\right\|_{L^p} = 0, \qquad a.e. \ \text{in } [0,\infty)\,. \tag{52}$$

Proof. For a given partition $\vec{\tau}$ of $[0,\infty)$ let $V_{\vec{\tau}}^k$ be as in (38), the iterative minimizers of the dual problem. We define a piecewise linear function $\tilde{V}_{\vec{\tau}} : [0,\infty) \to \mathrm{L}^{p_*}(\Omega)$ that linearly interpolates the discrete values $V_{\vec{\tau}}^k$. That is

$$\tilde{V}_{\vec{\tau}}(t) = \frac{1}{\tau_k}\left(\left(t_{\vec{\tau}}^k - t\right) V_{\vec{\tau}}^{k-1} + \left(t - t_{\vec{\tau}}^{k-1}\right) V_{\vec{\tau}}^k\right) \text{ for } t \in \left(t_{\vec{\tau}}^{k-1}, t_{\vec{\tau}}^k\right].$$

We proceed by estimating the difference between the linearly interpolated function $\tilde{V}_{\vec{\tau}}$ and the piecewise constant function $\overline{V}_{\vec{\tau}}$. For each $t \in \left(t_{\vec{\tau}}^{k-1}, t_{\vec{\tau}}^k\right]$ we have

$$\begin{aligned} \left\|\tilde{V}_{\vec{\tau}}(t) - \overline{V}_{\vec{\tau}}(t)\right\|_{\mathrm{L}^{p_*}} &= \left\|\frac{1}{\tau_k}\left(\left(t_{\vec{\tau}}^k - t\right) V_{\vec{\tau}}^{k-1} + \left(t - t_{\vec{\tau}}^{k-1}\right) V_{\vec{\tau}}^k\right) - V_{\vec{\tau}}^k\right\|_{\mathrm{L}^{p_*}} \\ &= \frac{t_{\vec{\tau}}^k - t}{\tau_k}\left\|V_{\vec{\tau}}^{k-1} - V_{\vec{\tau}}^k\right\|_{\mathrm{L}^{p_*}} \leq \left\|V_{\vec{\tau}}^{k-1} - V_{\vec{\tau}}^k\right\|_{\mathrm{L}^{p_*}}. \end{aligned} \tag{53}$$

From the dual formulation (29) of the Bregman distance regularization it follows that for all $v \in K^{p_*}(\Omega)$

$$\frac{1}{p^*}\left\|V_{\vec{\tau}}^{k-1} - V_{\vec{\tau}}^k\right\|_{\mathrm{L}^{p_*}}^{p_*} - \tau_k^{p_*-1}\left\langle u^\delta, V_{\vec{\tau}}^k\right\rangle \leq \frac{1}{p^*}\left\|V_{\vec{\tau}}^{k-1} - v\right\|_{\mathrm{L}^{p_*}}^{p_*} - \tau_k^{p_*-1}\left\langle u^\delta, v\right\rangle.$$

With the choice $v = V_{\vec{\tau}}^{k-1}$ it follows that

$$\begin{aligned} \frac{1}{p^*}\left\|V_{\vec{\tau}}^k - V_{\vec{\tau}}^{k-1}\right\|_{\mathrm{L}^{p_*}}^{p_*} &\leq \tau_k^{p_*-1}\left\langle u^\delta, V_{\vec{\tau}}^{k-1} - V_{\vec{\tau}}^k\right\rangle \\ &\leq \tau_k^{p_*-1}\left\|u^\delta\right\|_{\mathrm{L}^p} \cdot \left\|V_{\vec{\tau}}^{k-1} - V_{\vec{\tau}}^k\right\|_{\mathrm{L}^{p_*}} \end{aligned}$$

which implies that $\left\|V_{\vec{\tau}}^{k-1} - V_{\vec{\tau}}^k\right\|_{\mathrm{L}^{p_*}} = \mathcal{O}(|\vec{\tau}|)$.

Let $\{\vec{\tau}_l\}_{l\in\mathbb{N}}$ be a sequence of partitions of $[0,\infty)$ such that $\lim_{l\to\infty}|\vec{\tau}_l| = 0$ and $\overline{V}_{\vec{\tau}_l}(t) \rightharpoonup v(t)$ for all $t \in [0,\infty)$. Together with (53) it follows that for all $t \in [0,\infty)$ and all $u \in \mathrm{L}^p(\Omega)$

$$\begin{aligned} \lim_{l\to\infty}\int_\Omega \tilde{V}_{\vec{\tau}_l}(t) u \,\mathrm{d}x &= \lim_{l\to\infty}\int_\Omega \left(\tilde{V}_{\vec{\tau}_l}(t) - \overline{V}_{\vec{\tau}_l}(t)\right) u \,\mathrm{d}x \\ &\quad + \lim_{l\to\infty}\int_\Omega \overline{V}_{\vec{\tau}_l}(t) u \,\mathrm{d}x = \int_\Omega v(t) u \,\mathrm{d}x. \end{aligned}$$

Let $I \subseteq [0,\infty)$ be a compact intervall and set $\Omega_I = I \times \Omega$. This shows that for every $t \in [0,\infty)$ there exists a positive number $M_t > 0$ such that $\left\|\tilde{V}_{\vec{\tau}_l}(t)\right\|_{\mathrm{L}^{p_*}} \leq M_t$ for all $l \in \mathbb{N}$. Since $t \mapsto \left\|\tilde{V}_{\vec{\tau}_l}(t)\right\|_{\mathrm{L}^{p_*}}$ is continuous we have that $M = \sup_{t\in I} M_t < \infty$ and it follows that

$$\left\|\tilde{V}_{\vec{\tau}_l}(t)\right\|_{\mathrm{L}^{p_*}} \leq M \text{ for all } t \in I \text{ and } l \in \mathbb{N}.$$

Then, by using the Lebesgue Dominated Convergence Theorem (see for instance [24, Thm. 1.34]) it follows that for all $u \in \mathrm{L}^p(\Omega_I)$

$$\begin{aligned} \lim_{l\to\infty}\int_{\Omega_T} \tilde{V}_{\vec{\tau}_l} u \,\mathrm{d}t \otimes \mathrm{d}x &= \int_0^T \lim_{l\to\infty}\int_\Omega \tilde{V}_{\vec{\tau}_l}(x,t) u(x,t) \,\mathrm{d}x \,\mathrm{d}t \\ &= \int_0^T \int_\Omega v(x,t) u(x,t) \,\mathrm{d}x \,\mathrm{d}t \end{aligned} \tag{54}$$

or in other words that $\tilde{V}_{\vec{\tau}_l} \rightharpoonup v$ in $\mathrm{L}^{p_*}(\Omega_I)$. Moreover, since

$$\frac{\partial \tilde{V}_{\vec{\tau}_l}(t)}{\partial t} = \frac{V_{\vec{\tau}_l}^k - V_{\vec{\tau}_l}^{k-1}}{t_{\vec{\tau}_l}^k - t_{\vec{\tau}_l}^{k-1}}, \quad t \in \left(t_{\vec{\tau}_l}^{k-1}, t_{\vec{\tau}_l}^k\right) \tag{55}$$

we can apply [2, Thm. 2.3.3.], from which it follows that

$$\lim_{l\to\infty} \left\| \frac{\partial \tilde{V}_{\vec{\tau}_l}}{\partial t} \right\|_{\mathrm{L}^{p_*}(\Omega_I)} = \|v'\|_{\mathrm{L}^{p_*}(\Omega_I)} . \tag{56}$$

Therefore the sequence $\left\{ \frac{\partial \tilde{V}_{\vec{\tau}_l}}{\partial t} \right\}_{l\in\mathbb{N}}$ is bounded in $\mathrm{L}^{p_*}(\Omega_I)$ and consequently $\{\tilde{V}_{\vec{\tau}_l}\}_{l\in\mathbb{N}}$ is bounded in the space

$$\mathcal{Y} := \left\{ y \in \mathrm{L}^{p_*}(\Omega_I) : y' = \frac{\partial y}{\partial t} \in \mathrm{L}^{p_*}(\Omega_I) \right\}.$$

Since the space $\mathcal{Y}$ with norm

$$\|y\|_{\mathcal{Y}} = \|y\|_{\mathrm{L}^{p_*}(\Omega_I)} + \|y'\|_{\mathrm{L}^{p_*}(\Omega_I)}$$

is a reflexive Banach space (see for instance [26, Chap. III]), there exists a weakly convergent subsequence $\{\tilde{V}_{\vec{\tau}_{k(l)}}\}_{l\in\mathbb{N}}$ and a limit $\hat{v} \in \mathcal{Y}$ satisfying

$$\left. \begin{aligned} \tilde{V}_{\vec{\tau}_{k(l)}} &\rightharpoonup \hat{v}, \\ \frac{\partial \tilde{V}_{\vec{\tau}_{k(l)}}}{\partial t} &\rightharpoonup \hat{v}' \end{aligned} \right\} \text{ as } l \to \infty \quad \text{in } \mathrm{L}^{p_*}(\Omega_I). \tag{57}$$

Together with (54) this shows $\hat{v} = v$. From (56) and (57) and the Radon-Riesz property of $\mathrm{L}^{p_*}(\Omega_I)$ (see [20, Chap. 2.5]) strong convergence of $\left\{ \frac{\partial \tilde{V}_{\vec{\tau}_{k(l)}}}{\partial t} \right\}_{l\in\mathbb{N}}$ follows. We further conclude (see for instance [17, Chap. 1.3 Thm. 5]) that there exists another subsequence, which for the simplicity of notation we again denote by $k(l)$, such that

$$\left\| \frac{\partial \tilde{V}_{\vec{\tau}_{k(l)}}}{\partial t}(t) - v'(t) \right\|_{\mathrm{L}^{p_*}} \to 0 \text{ for almost all } t \in I . \tag{58}$$

Moreover, from (35) and (55) we obtain for $t \in \left(t^{k-1}_{\vec{\tau}_{k(l)}}, t^k_{\vec{\tau}_{k(l)}}\right]$

$$\begin{aligned} \overline{U}_{\vec{\tau}_{k(l)}}(t) = U^k_{\vec{\tau}_{k(l)}} &= u^\delta - \left[\frac{1}{\tau^k_{k(l)}} \right]^{p_*-1} \mathfrak{J}_{p_*}\left(V^k_{\vec{\tau}_{k(l)}} - V^{k-1}_{\vec{\tau}_{k(l)}} \right) \\ &= u^\delta - \mathfrak{J}_{p_*}\left(\frac{V^k_{\vec{\tau}_{k(l)}} - V^{k-1}_{\vec{\tau}_{k(l)}}}{\tau^k_{\vec{\tau}_{k(l)}}} \right) \\ &= u^\delta - \mathfrak{J}_{p_*}\left(\frac{\partial \tilde{V}_{\vec{\tau}_{k(l)}}(t)}{\partial t} \right) \end{aligned} \tag{59}$$

and from (33) that

$$\overline{V}_{\vec{\tau}_{k(l)}}(t) \in \partial J(\overline{U}_{\vec{\tau}_{k(l)}}(t)). \tag{60}$$

From (58) and the norm-norm continuity of $\mathfrak{J}_{p_*}$ ([11, Thm 2.16]) it follows that the right hand side of (59) converges to a function $\tilde{u}(t) \in \mathrm{L}^p(\Omega)$ for almost all $t \in I$ that satisfies

$$\mathfrak{J}_{p_*}(v') = u^\delta - \tilde{u}(t), \text{ almost everywhere in } I.$$

The graph of ∂J is closed in $\mathrm{L}^p(\Omega) \times \mathrm{L}^{p_*}(\Omega)$ with respect to the strong topology on $\mathrm{L}^p(\Omega)$ and weak topology on $\mathrm{L}^{p_*}(\Omega)$ (cf. [14, Chap. I Cor. 5.1]). That is, the set $\{(u,v) : v \in \partial J(u), u \in \mathrm{L}^p(\Omega)\}$ satisfies: Let $u_k \in \mathrm{L}^p(\Omega)$ and $v_k \in \partial J(u_k) \subseteq \mathrm{L}^{p_*}(\Omega)$ satisfying $v_k \rightharpoonup v$ in $\mathrm{L}^{p_*}(\Omega)$ and $u_k \to u$ in $\mathrm{L}^p(\Omega)$, then $v \in \partial J(u)$. Therefore $(v, \tilde{u})$ is a solution of (49) and since for every solution v of (46) the function u as in (48) is unique it follows that $\tilde{u} = u$ on I.

In other words this shows that every subsequence of $\overline{U}_{\vec{\tau}_l}$ has an almost everywhere convergent subsequence with limit u (restricted to I) which means that

$$\lim_{l\to\infty} \left\|\overline{U}_{\vec{\tau}_l}(t) - u(t)\right\|_{\mathrm{L}^p} = 0, \qquad \text{a.e. in } I.$$

Choosing $I = [k, k+1]$ for $k \in \mathbb{Z}$ finally shows (52). □

Corollary 2 (Monotonicity)**.** *Let $u^\delta \in L^p(\Omega)$. If u is a solution of (49) we have*

$$\left\|u(s) - u^\delta\right\|_{L^p} \le \left\|u(t) - u^\delta\right\|_{L^p} \tag{61}$$

for almost all s, t in $[0,\infty)$ satisfying $s > t$.

Proof. Let $T > t$ and $\{\vec{\tau}_l\}_{l\in\mathbb{N}}$ be a partition of $[0,\infty)$ such that $\lim_{l\to\infty} |\vec{\tau}_l| = 0$ and that (52) holds a.e. in $[0,T]$ (particularly for s and t). Then there exists an index l_0 such that for all $l > l_0$ there exist $k(l), \tilde{k}(l) \in \mathbb{N}$ satisfying $t \in (t_{\vec{\tau}_l}^{k(l)-1}, t_{\vec{\tau}_l}^{k(l)}]$, $s \in (t_{\vec{\tau}_l}^{\tilde{k}(l)-1}, t_{\vec{\tau}_l}^{\tilde{k}(l)}]$ and

$$t_{\vec{\tau}_l}^{k(l)} < t_{\vec{\tau}_l}^{\tilde{k}(l)-1}.$$

Then it follows from Lemma 4 that

$$\left\|\overline{U}_{\vec{\tau}_l}(s) - u^\delta\right\|_{\mathrm{L}^p} \le \left\|\overline{U}_{\vec{\tau}_l}(t) - u^\delta\right\|_{\mathrm{L}^p}, \qquad \text{for all } l > l_0.$$

With this we obtain the estimate

$$\begin{aligned}
\left\|u(s) - u^\delta\right\|_{\mathrm{L}^p} &\le \left\|\overline{U}_{\vec{\tau}_l}(s) - u^\delta\right\|_{\mathrm{L}^p} + \left\|\overline{U}_{\vec{\tau}_l}(s) - u(s)\right\|_{\mathrm{L}^p} \\
&\le \left\|\overline{U}_{\vec{\tau}_l}(t) - u^\delta\right\|_{\mathrm{L}^p} + \left\|\overline{U}_{\vec{\tau}_l}(s) - u(s)\right\|_{\mathrm{L}^p} \\
&\le \left\|\overline{U}_{\vec{\tau}_l}(t) - u(t)\right\|_{\mathrm{L}^p} + \left\|\overline{U}_{\vec{\tau}_l}(s) - u(s)\right\|_{\mathrm{L}^p} + \left\|u(t) - u^\delta\right\|_{\mathrm{L}^p}.
\end{aligned}$$

Taking the limit $l \to \infty$ shows (61). □

In this section we showed that there exists a solution (v, u) of the *inverse total variation flow equation* (49). The sequences $\{u_k\}$ and $\{v_k\}$ in Algorithm 1 can be considered as numerical approximations of u and v respectively, corresponding to the time step size

$$\Delta t = \frac{1}{\alpha^{p_*-1}}.$$

It is important to note that the generation of the (dual) flow equation (46) is independent of the minimizers of the primal variational problem in Algorithm 1. In other words, the function u in (48) is established *artificially* and a connection to Algorithm 1 is a priori *not* obvious. Theorem 7 finally provides this relation.

4. Multiscale Properties

In the following we discuss the multiscale properties of the inverse total variation flow. From the interpretation of (49) as an inverse scale space method, we expect that large scales are reconstructed for small times, while finer scales take a longer time to be included in the reconstruction. Our numerical results below will confirm this behavior. Moreover we provide some examples giving a theoretical justification. Throughout the whole section we make the natural assumption that $u(0) = 0$ and $v(0) = 0$.

In a linear inverse scale space method we would expect that the reconstruction approaches the image continuously in time. This is not true for the inverse total variation flow as the next result shows, the continuous evolution rather appears for the dual variable v:

Theorem 8. *Let u^δ be normalized such that $\mathfrak{J}_p(u^\delta) \in \mathcal{L}_\Diamond^{p_*}(\Omega)$. For $t\left\|\mathfrak{J}_p(u^\delta)\right\|_* \le 1$, a solution of* (49) *is given by*

$$u(t) = 0, \qquad v(t) = t\ \mathfrak{J}_p(u^\delta). \tag{62}$$

Proof. In the following let $t \leq \frac{1}{\|\mathfrak{J}_p(u^\delta)\|_*}$. We verify that (v,u) defined by (62) is indeed a solution of (49). Since $v' = \mathfrak{J}_p(u^\delta) = \mathfrak{J}_p(u^\delta - u)$ it follows that

$$\mathfrak{J}_{p_*}(v') = u^\delta - u.$$

After Proposition 2 there exists $z \in \ker(T_{p_*})$ such that $\operatorname{div}(z) = \mathfrak{J}_p(u^\delta)$ and $\|z\|_{L^\infty(\Omega)^n} \leq \left\|\mathfrak{J}_p(u^\delta)\right\|_*$. Then, for all $\varphi \in C_c^1(\Omega)$ and all $t \in (0,T)$ we have

$$\begin{aligned} J(\varphi) - J(u(t)) - \int_\Omega v(t)\ (\varphi - u(t))\ \mathrm{d}x &= J(\varphi) - \int_\Omega t\ \mathfrak{J}_p(u^\delta)\ \varphi\ \mathrm{d}x \\ &= \int_\Omega (|\nabla\varphi| + t\ z\ \nabla\varphi)\ \mathrm{d}x \\ &\geq (1 - t\left\|\mathfrak{J}_p(u^\delta)\right\|_*)\int_\Omega |\nabla\varphi|\ \mathrm{d}x \geq 0. \end{aligned}$$

By standard continuity and density arguments we can now extend the inequality

$$J(\varphi) - J(u(t)) - \int_\Omega v(t)\ (\varphi - u(t))\ \ \mathrm{d}x \geq 0$$

to all $\varphi \in L^p(\Omega)$ and hence, $v(t) \in \partial J(0) = \partial J(u(t))$. Thus, (v,u) is a solution of (49). □

Theorem 8 shows that the reconstruction does not change in an initial stage of the evolution up to time $t_* = \frac{1}{\|\mathfrak{J}_p(u^\delta)\|_*}$, while the dual variable changes at linear rate in time.

In order to illustrate the behavior for $t > t_*$, we consider simple functions. We start with an auxiliary result (see also [5],[3]).

Lemma 6. *Let $u \in L^p(\Omega)$ be a simple function. Then*

$$\frac{\mathfrak{J}_p(u)}{\|\mathfrak{J}_p(u)\|_*} \in \partial J(u).$$

Proof. For an arbitrary $w \in L^p(\Omega)$ we have

$$\begin{aligned} J(u) + \left\langle \frac{\mathfrak{J}_p(u)}{\|\mathfrak{J}_p(u)\|_*}, w - u \right\rangle &= J(u) + \frac{1}{\|\mathfrak{J}_p(u)\|_*}\left(\langle \mathfrak{J}_p(u), w\rangle - \langle \mathfrak{J}_p(u), u\rangle\right) \\ &= J(u) + \frac{1}{\|\mathfrak{J}_p(u)\|_*}\left(\langle \mathfrak{J}_p(u), w\rangle - \|u\|_{L^p}^p\right) \\ &= J(u) + \frac{1}{\|\mathfrak{J}_p(u)\|_*}\langle \mathfrak{J}_p(u), w\rangle - J(u) \\ &\leq \frac{1}{\|\mathfrak{J}_p(u)\|_*}\|\mathfrak{J}_p(u)\|_*\, J(w) = J(w). \end{aligned}$$

□

According to [19], such simple functions are the building blocks of cartoon images. For the inverse total variation flow they are also fundamental, since they can be recovered in finite time:

Theorem 9. *Let u^δ be a simple function. Then, for $t\left\|\mathfrak{J}_p(u^\delta)\right\|_* \geq 1$, a solution of* (49) *is given by*

$$u(t) = u^\delta, \qquad v(t) = t_*\ \mathfrak{J}_p(u^\delta). \tag{63}$$

Proof. We have noticed in Lemma 6 that for simple functions the inclusion

$$v(t) = t_*\ \mathfrak{J}_p(u^\delta) = \frac{\mathfrak{J}_p(u^\delta)}{\|\mathfrak{J}_p(u^\delta)\|_*} \in \partial J(u^\delta) = \partial J(u(t))$$

holds. Moreover, $\mathfrak{J}_{p_*}(v')(t) = 0 = u^\delta - u(t)$ is obviously satisfied. Hence (v, u) is a solution of (49). □

It is also easy to see that the converse of Theorem 9 is true, i.e., if $\mathfrak{J}_p(u^\delta) \in \mathcal{L}^{p_*}_\diamond(\Omega)$ and (63) holds for $t > t_*$, then u^δ is simple. Hence, the simple images are the ones that can be reconstructed at time t_*.

Remark 2. Combining Theorems 8 and 9 shows that for simple initial data u^δ a solution (v, u) of (49) is given by

$$v : t \to \begin{cases} t\ \mathfrak{J}_p(u^\delta) & \text{if } 0 \le t \le t_*, \\ t_*\ \mathfrak{J}_p(u^\delta) & \text{else.} \end{cases} \qquad u : t \to \begin{cases} 0 & \text{if } 0 \le t \le t_*, \\ u^\delta & \text{else.} \end{cases}$$

This shows that the regularity result in Proposition 6 and the growth property in Proposition 4 are sharp.

Typical simple images in spatial dimension one are piecewise constant functions, that is, for $\Omega = (-1, 1)$ the function

$$f(x) = \begin{cases} -1 & \text{if } x < 0, \\ +1 & \text{if } x > 0. \end{cases}$$

is simple (for any $p \in (1, \infty)$), and satisfies $\|\mathfrak{J}_p(f)\|_* = 1$. Hence, this simple step function will be reconstructed by the inverse total variation flow at time $t_* = 1$.

A bit more instructive is the analysis for piecewise constant signal with a positive part in the middle, which has two scale parameters, namely its height and width:

Example 3. Let $\Omega = (-1, 1)$ and let, for $H > 0$, $R \in (0, 2)$,

$$f(x) = \frac{H}{C} \begin{cases} -1 & \text{if } |x| \ge \frac{R}{2}, \\ \left(\frac{2}{R} - 1\right)^{p_* - 1} & \text{if } |x| < \frac{R}{2} \end{cases}$$

with $C = 1 + \left(\frac{2}{R} - 1\right)^{p_*-1}$. In this case we have $\int_{-1}^{1} \mathfrak{J}_p(f)\ \mathrm{d}x = 0$ and $\mathfrak{J}_p(f) = \frac{dz}{dx}$ for

$$z(x) = \left(\frac{H}{C}\right)^{p-1} \begin{cases} -(x+1) & \text{if } x \le -\frac{R}{2}, \\ \left(\frac{2}{R} - 1\right) x & \text{if } |x| < \frac{R}{2} \\ (1 - x) & \text{if } x \ge \frac{R}{2}. \end{cases}$$

The function z satisfies $z(-1) = z(1) = 0$ and $\|\mathfrak{J}_p(f)\|_* = \|z\|_\infty = \left(\frac{H}{C}\right)^{p-1} \left(1 - \frac{R}{2}\right)$ and consequently the reconstruction time is given by $t_* = \left(\frac{C}{H}\right)^{p-1} \left(\frac{2}{2-R}\right)$. In Figure 1 the G-norm of $\mathfrak{J}_p(f)$ and the reconstruction time for f are plotted against the variable R for fixed $H = 1$ and $p = 2$.

We note that the smaller the spatial features are (that is the smaller $\|\mathfrak{J}_p(f)\|_*$ is) the longer it takes to recover the signal. By increasing the width of the peak the reconstruction time decreases until the positive part of the signal ($|x| < R/2$) equals half the interval length. Beyond this point the negative parts of the signal ($|x| \ge R/2$) behave like peaks and therefore require a larger reconstruction time.

We finally turn to the large-time properties of the flow (46) in the case of arbitrary images. In the previous section, we have interpreted the inverse total variation flow as a gradient flow according to $F : \mathrm{L}^{p_*}(\Omega) \to \overline{\mathbb{R}}$ (cf. (47))

$$F(v) := -\int_\Omega u^\delta v\ \mathrm{d}x + J^*(v).$$

Below we prove monotone decrease of the dual functional F, which also yields a convergence of solutions u of (49) to u^δ. Theorem 10 can be considered as the continuous version of Theorem 3.

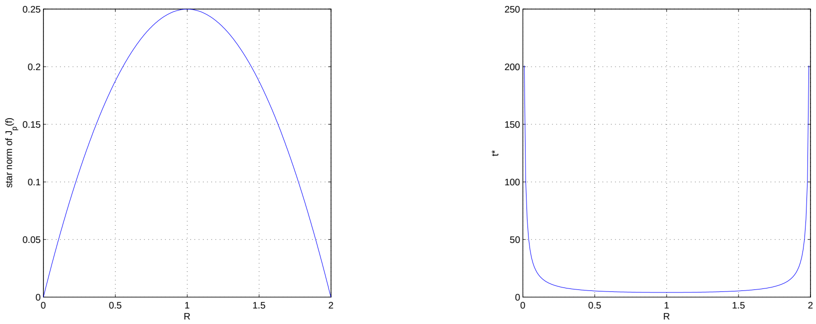

FIGURE 1. Left: the star norm of $\mathfrak{J}_p(f)$ represents the size of spatial features. Right: reconstruction time t_* for f.

Theorem 10 (Inverse Fidelity). *Let $u^\delta \in L^p(\Omega)$, let (v,u) be a solution of (49) and let F be defined by (47). Then $t \mapsto F(v(t))$ is non-increasing (decreasing if $u(t) \neq u^\delta$) and we have*

$$\lim_{t\to\infty} \left\|u(t) - u^\delta\right\|_{L^p} = 0. \tag{64}$$

If, in addition $u^\delta \in BV(\Omega)$, then we get the convergence rate

$$\left\|u(t) - u^\delta\right\|_{L^p} \leq \left(\frac{J(u^\delta)}{t}\right)^{\frac{1}{p}}, \qquad \textit{for a.e. } t \in [0,\infty).$$

Proof. Let $s,t > 0$ and $u(t) \in \partial J^*(v(t))$. Then it follows that $\langle v(s) - v(t), u(t)\rangle \leq 0$ and together with the definition of F and the fact that $J^*(v(t)) = J^*(v(s)) = 0$ we have that

$$\begin{aligned} F(v(t)) - F(v(s)) &= \langle v(s) - v(t), u^\delta\rangle \\ &\leq \langle v(s) - v(t), u^\delta\rangle - \langle v(s) - v(t), u(t)\rangle \\ &= -\langle v(t) - v(s), u^\delta - u(t)\rangle, \end{aligned} \tag{65}$$

for almost every s,t. From Theorem 4 it follows that v is differentiable at almost all $t \in [0,\infty)$ and therefore $t \to F(v(t))$ is differentiable a.e. in $[0,\infty)$. Together with (65) it becomes clear that

$$\lim_{s\uparrow t} \frac{F(v(t)) - F(v(s))}{t-s} = \frac{\partial}{\partial t} F(v(t)) \leq -\left\langle \frac{\partial v}{\partial t}(t), u(t) - u^\delta \right\rangle = -\left\|u(t) - u^\delta\right\|^p_{\mathrm{L}^p},$$

for a.e. $t \in [0,\infty)$. Hence, we may conclude that

$$F(v(t)) - F(v(\tau)) \leq -\int_\tau^t \|u(s) - u^\delta\|^p_{\mathrm{L}^p}\, ds \leq 0.$$

With $\tau = 0$ and $v(0) = 0$ this gives

$$\int_0^t \|u(s) - u^\delta\|^p_{\mathrm{L}^p}\, ds \leq -F(v(t)).$$

Combining this inequality with Corollary 2 yields

$$\left\|u(t) - u^\delta\right\|^p_{\mathrm{L}^p} \leq \frac{1}{t}\int_0^t \left\|u(s) - u^\delta\right\|^p_{\mathrm{L}^p}\, ds \leq \int_\Omega \frac{v(t)}{t} u^\delta \,\mathrm{d}x.$$

In order to show (64) we prove that $\frac{v(t)}{t}$ weakly converges to 0 as $t \to \infty$. Since $v(t) \in K^{p_*}(\Omega)$ we have that

$$\lim_{t\to\infty} \left\| \frac{v(t)}{t} \right\|_* = \lim_{t\to\infty} \frac{\|v(t)\|_*}{t} \leq \lim_{t\to\infty} \frac{1}{t} = 0.$$

Therefore it suffices to show that $\left\{ \frac{v(t)}{t} \right\}_{t\geq 0}$ is bounded in $\mathrm{L}^{p_*}(\Omega)$ after Lemma 3 which directly follows from Proposition 4. Finally, since $v(t) \in K^{p_*}(\Omega)$ it follows for $u^\delta \in \mathrm{BV}(\Omega)$ that (cf. Lemma 1)

$$-F(v(t)) = \langle v(t), u^\delta \rangle \leq \sup_{v \in K^{p_*}(\Omega)} \langle v, u^\delta \rangle = J(u^\delta).$$

□

Theorem 10 provides some global information about the speed of reconstruction in relation to the smoothness of the image. The difference between the reconstruction and the image decays like $\sqrt[p]{J(u^\delta)/t}$, and which is decreasing with the total variation of u^δ. Thus, also in this global sense smooth images are reconstructed faster than nonsmooth ones.

5. Numerical Simulations

In this section we give some argumentation for applying Algorithm 1 for image and surface denoising. The efficiency of this method for image restoration has already been investigated in [22] and [8].

In the upcoming section we show that iterative Bregman regularization is a feasible technique for denoising of surfaces that are represented as level sets of suitable level set functions. Since an image can be considered as the union of its level lines the following argumentation is also applicable on image restoration.

5.1. **Surface Denoising.**

Let $u \in \mathrm{BV}(\Omega)$ and the zero super level set

$$C = \{x \in \Omega : u(x) > 0\}.$$

Assume that

$$|\partial C|(\Omega) := |\mathrm{D}\chi_C|(\Omega) < \infty .$$

For $x \in \Omega$, we consider $\delta(x)$ a realization of a random variable $\Delta(x)$. With the function

$$u^\delta(x) = u(x) + \delta(x) \text{ for all } x \in \Omega . \tag{66}$$

we associate a noisy surface as the boundary of the super level set of u^δ

$$\partial C^\delta \text{ with } C^\delta = \{x \in \Omega : u^\delta(x) > 0\}. \tag{67}$$

Let $\{u_k\}_{k\in\mathbb{N}}$ be defined by Algorithm 1. We see from the co-area formula in $\mathrm{BV}(\Omega)$ (see e.g. [17, Chap. 5.5 Thm. 1]) that for all $u \in \mathrm{BV}(\Omega) \cap \mathrm{L}^p(\Omega)$

$$\begin{aligned} I(\alpha; u_1, 0) =& \frac{1}{p} \int_\Omega \left| u_1 - u^\delta \right|^p + \alpha |\mathrm{D}u_1|(\Omega) \\ =& \frac{1}{p} \left\| u_1 - u^\delta \right\|_{\mathrm{L}^p}^p + \alpha \int_{-\infty}^{\infty} |\partial\{x \in \Omega : u_1(x) > c\}|(\Omega) \,\mathrm{d}c \\ \leq& \frac{1}{p} \left\| u - u^\delta \right\|_{\mathrm{L}^p}^p + \alpha \int_{-\infty}^{\infty} |\partial\{x \in \Omega : u(x) > c\}|(\Omega) \,\mathrm{d}c . \end{aligned} \tag{68}$$

Equation (68) shows that α enforces a trade off between the amount of displacements of u^δ and u_1 and smoothness of the level sets of u_1 in average. We explain the second Bregman iteration step:

$$
\begin{aligned}
u_2 &= \operatorname*{argmin}_{u\in \mathrm{L}^p(\Omega)} I(\alpha; u, u_1) \\
&= \operatorname*{argmin}_{u\in \mathrm{L}^p(\Omega)} \left\{ \frac{1}{p} \left\| u - u^\delta \right\|_{\mathrm{L}^p}^p + \alpha(J(u) - \langle v_1, u\rangle) \right\}.
\end{aligned} \tag{69}
$$

Let $u \in \mathrm{L}^p(\Omega) \cap \mathrm{BV}(\Omega)$ and $v \in K^{p_*}(\Omega)$. Then there exists $z \in \ker(T_{p_*})$ such that $\mathrm{div}\,(z) = v$ and $\|z\|_{\mathrm{L}^\infty(\Omega)^n} \le 1$. After Lemma 3 we can find a function $\theta(v,u,\cdot) \in \mathrm{L}^\infty(\Omega, \mathbb{R}, |\mathrm{D}u|)$ such that

$$\int_\Omega u \mathrm{div}\,(z)\ \mathrm{d}x = \int_\Omega uv\ \mathrm{d}x = \int_\Omega \theta(v,u,x) \mathrm{d}|\mathrm{D}u|$$

and $\|\theta(v,u,\cdot)\|_{\mathrm{L}^\infty(\Omega,\mathbb{R},|\mathrm{D}u|)} \le \|v\|_* \le 1$. Thus there exists an angle $\gamma(v,u,x) \in [0,\pi)$ for all $x \in \Omega$ such that $\cos(\gamma(v,u,x)(x)) = \theta(v,u,x)$ $|\mathrm{D}u|$ – a.e. in Ω and that

$$J(u) - \langle v, u\rangle = \int_\Omega (1 - \cos(\gamma(v,u,x)))\mathrm{d}|\mathrm{D}u|.$$

It remains to investigate the geometrical meaning of $\gamma(v,u,\cdot)$. To this end we apply [3, Thm c.14] and obtain that

$$\theta(v,u,\cdot) = z \cdot \frac{\mathrm{D}u}{|\mathrm{D}u|}, \quad |\mathrm{D}^a u| - \text{ a.e. in } \Omega$$

where $\frac{\mathrm{D}u}{|\mathrm{D}u|}$ denotes the density of $\mathrm{D}u$ w.r.t. $|\mathrm{D}u|$ and $|\mathrm{D}^a u|$ the absolut continuous part of the measure $|\mathrm{D}u|$. In other words $\gamma(v,u,x)$ can be considered as the (generalized) angle between $z(x)$ and the normalized gradient of u at $x \in \Omega$.

With this we rewrite (69) and obtain

$$
\begin{aligned}
u_2 &= \operatorname*{argmin}_{u\in \mathrm{L}^p(\Omega)} \left\{ \frac{1}{p} \left\| u - u^\delta \right\|_{\mathrm{L}^p}^p + \alpha(J(u) - \langle v_1, u\rangle) \right\} \\
&= \operatorname*{argmin}_{u\in \mathrm{L}^p(\Omega)} \left\{ \left\| u - u^\delta \right\|_{\mathrm{L}^p}^p + \alpha \int_\Omega (1 - \cos\gamma(z_1, u, x))\ \mathrm{d}|\mathrm{D}u| \right\}.
\end{aligned}
$$

where $\mathrm{div}\,(z_1) = v_1$. It is important to note that $\langle v_1, u_1\rangle = |\mathrm{D}u_1|(\Omega)$ and therefore $\gamma(v_1, u_1, \cdot) = 0$ $|\mathrm{D}^a u_1|$ a.e. in Ω, i.e. z_1 is parallel to the normalized gradient of u_1. Therefore, the second step of iterative Bregman regularization consists in simultaneously minimizing the distance between u_2 of u^δ and the angle between unit normal vectors of the level sets of u_1 and u_2.

5.2. **Computational Realization.** We describe the computational realization used for iterative total variation flow regularization (with $p = 2$).

We assume that

$$\|\delta\|_{\mathrm{L}^2} \le \delta'. \tag{70}$$

The pseudo code for iterative Bregman distance regularization, Algorithm 1, in the case $p = 2$ reads as follows (see for instance [8]):

Algorithm 2. (1) $w_0 = 0$, $u_0 = 0$, $k = 0$

(2) `while` $\left\|u_{k+1} - u^\delta\right\|_{\mathrm{L}^2} \ge \delta'$

$$
\begin{aligned}
u_{k+1} &= \operatorname*{argmin}_{u\in \mathrm{L}^2(\Omega)} \left\{ \frac{1}{2} \left\| u - (u^\delta + w_k) \right\|_{\mathrm{L}^2}^2 + \alpha|\mathrm{D}u|(\Omega) \right\}, \\
w_{k+1} &= w_k + (u^\delta - u_{k+1}), \\
k &\leftarrow k + 1.
\end{aligned}
$$

In the case $p = 2$ iterative Bregman distance regularization simplifies to subsequent minimization of the ROF functional with data $u^\delta + w_k$.

Let

$$\Omega = \prod_{1\le i\le n} [a_i, b_i], \quad a_i < b_i \in \mathbb{R}$$

be an n–dimensional cuboid on which we consider a regular grid $\{x_i\}_{1\le i\le N}$ and n-linear ansatzfunctions $\{\psi_i\}_{1\le i\le N}$.

For a function $u \in \mathrm{L}^2(\Omega)$ we denote by

$$\hat{u} = \sum_{i=1}^{N} u_i \psi_i$$

the best approximation on $\mathcal{L} = \mathrm{span}\{\psi_i : i = 1, \dots, N\}$ of u in in the L^2 sense.

For minimization of the ROF-functional with data $u^\delta + w_k$ (that is the k-th iteration step of Bregman distance regularization) we solve the weak form of the Euler-Lagrange equation for u_i using a linear finite element method:

$$\sum_{i=1}^{N} \int_\Omega u_i\psi_i\psi_j + \alpha \frac{u_i}{\sqrt{|\nabla \hat{u}|^2 + \varepsilon^2}} \nabla\psi_i \nabla\psi_j \, \mathrm{d}x = \sum_{i=1}^{N} \int_\Omega (u_i^\delta + (w_k)_i)\psi_i\psi_j \, \mathrm{d}x, \quad 1 \le j \le N. \tag{71}$$

Here $\varepsilon > 0$ is a small positive parameter adapted to the grid size. The nonlinear equation (71) is solved by the following fixed point procedure

(1) Set $\hat{u}_0 = \hat{u}^\delta$.
(2) `for` $k = 1, 2, \dots$ solve

$$A_k u^{k+1} = b_k,$$

where

$$(A_k)_{ij} = \int_\Omega \psi_i\psi_j + \frac{\alpha}{\sqrt{|\nabla u^k|^2 + \varepsilon^2}} \nabla\psi_i\nabla\psi_j \, \mathrm{d}x$$

$$(b_k)_j = \sum_{i=1}^{N} \int_\Omega (u_i^\delta + (w_k)_i)\psi_i\psi_j \, \mathrm{d}x.$$

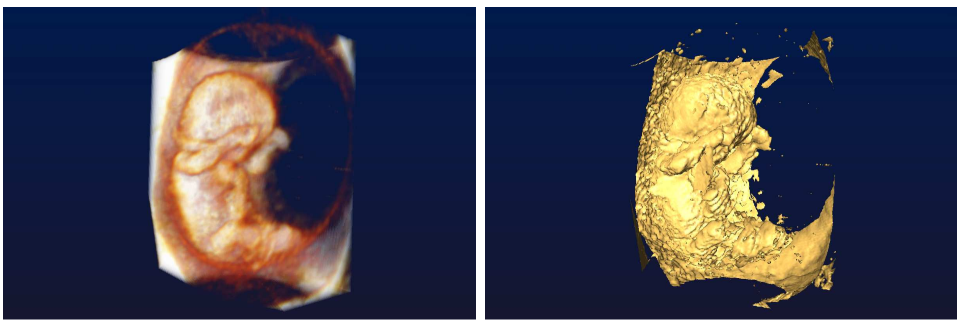

FIGURE 2. Volumetic view and level set of original data.

The following example shows 3D ultrasound data of a fetus with a resolution of $93 \times 186 \times 158$. A volumetric view as well as one of the (noisy) level set of the fetus data are displayed in Figure 2. The data is scaled between 0 and 1 and we performed Algorithm 2 with $\alpha = 0.5$ and $\varepsilon = 10^{-4}$. That is the minimizers

u_k correspond to solutions of (49) at times $t = 2k$. Figure 3 shows the denoised isosurfaces at steps $1, \ldots, 6$.

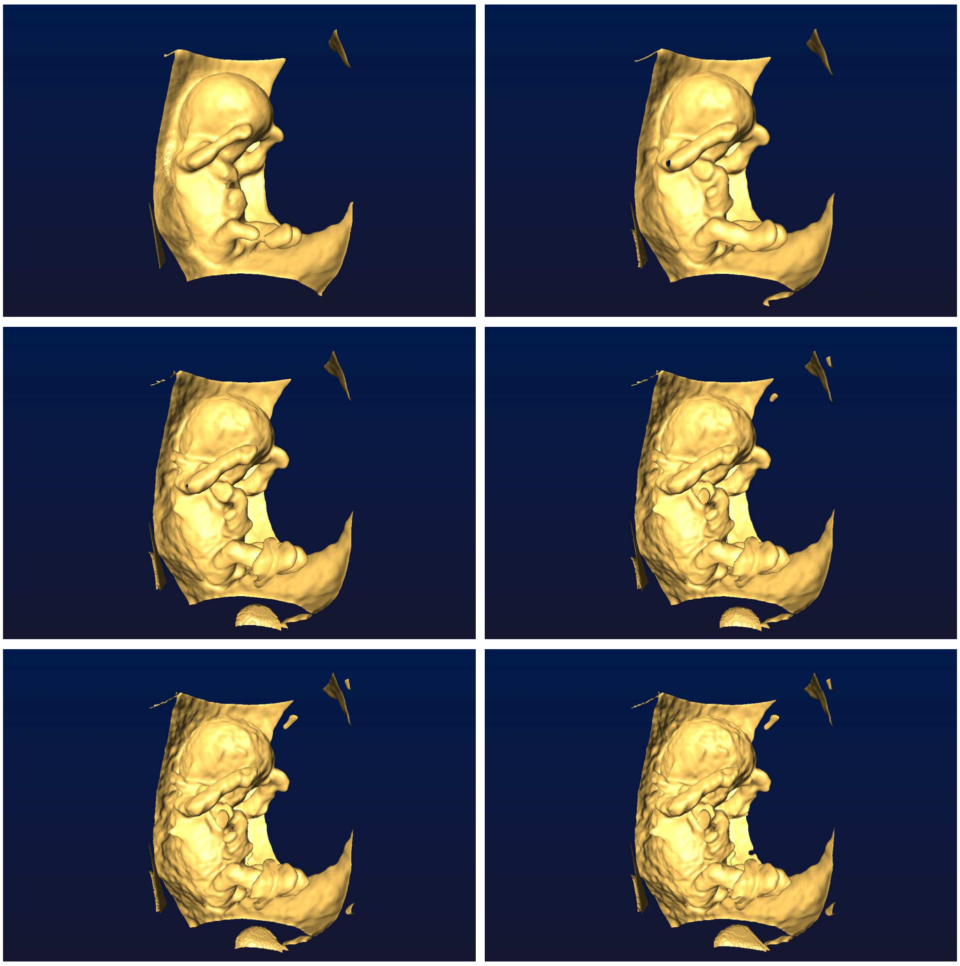

FIGURE 3. Denoised surfaces: Steps $1, \ldots, 6$ with $\alpha = 0.5$.

Acknowledgments

The work of M.B. has been supported by the Austrian Science Foundation FWF through project SFB F 013 / 08 and the Johann Radon Institute for Computational and Applied Mathematics (Austrian Academy of Sciences). The work of O.S. and K.F. has been supported by the Austrian Science Foundation (FWF) Projects Y-123INF, FSP 9203-N12 and FSP 9207-N12. The work of S.O. has been supported by NSF grants DMS-0312222, ACI-0321917 and DMI-0327077, and NIH Grant U54RR021813.

The authors also want to thank GE Medical Systems Kretztechnik AG for providing the ultrasound data as well as the referees for helpful comments.

References

[1] Robert Acar and Curtis Vogel. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Problems*, 10:1217–1229, 1994.

[2] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows*. Lectures in mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.

[3] Fuensanta Andreu-Vaillo, Vicent Caselles, and José M. Mazón. *Parabolic quasilinear equations minimizing linear growth functionals*, volume 223 of *Progress in Mathematics*. Birkhäuser, Basel, 2004.

[4] Gilles Aubert and Jean-François Aujol. Modeling very oscillationg signals. application to image processing. *Appl. Math. Optim.*, 51(2):163–182, 2005.

[5] Giovanni Bellettini, Vicent Caselles, and Matteo Novaga. The total variation flow in $\mathbb{R}^n$. *J. Differential Equations*, 184(2):475–525, 2002.

[6] Jean Bourgain and Haïm Brézis. On the equation $\operatorname{div} y = f$ and application to control of phases. *J. Amer. Math. Soc.*, 16(2):393–426, 2003.

[7] Haïm Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. Number 5 in North-Holland Mathematics Studies. North-Holland Publishing Co, Amsterdam-London, 1973.

[8] Martin Burger, Guy Gilboa, Stanley Osher, and Jinjun Xu. Nonlinear inverse scale space methods for image restoration. *Comm. Math. Sciences*, 4:179–212, 2006.

[9] Antonin Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vision*, 20(1–2):89–97, 2004.

[10] Antonin Chambolle and Pierre-Louis Lions. Image recovery via total variation minimization and related problmes. *Numer. Math.*, 76(2):167–188, 1997.

[11] Ioana Cioranescu. *Geometry of Banach spaces, duality mappings and nonlinear problems*. Kluwer Academic Publishers Group, Dordrecht, 1990.

[12] Michael G. Crandall and Thomas M. Liggett. Generation of semi-groups of nonlinear transformations on general banach spaces. *Amer. J. Math.*, 93:265–298, 1971.

[13] Bernard Dacorogna. *Direct methods in the calculus of variations*. Springer-Verlag, Berlin, 1989.

[14] Ivar Ekeland and Roger Temam. *Convex analysis and variational problems*, volume 1 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam-Oxford, 1976.

[15] Jürgen Elstrodt. *Maß- und Integrationstheorie*. Springer, Berlin - Heidelberg - New York, 1996.

[16] Heinz Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*. Mathematics and its Applications. Kluwer Academic Publishers Group, Dordrecht, 1996.

[17] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[18] Chuck W. Groetsch and Otmar Scherzer. Non-stationary iterated tikhonov - morozov method and third-order differential equations for the evaluation of unbounded operators. *Math. Methods Appl. Sci.*, 23(15):1287–1300, 2000.

[19] Ali Haddad and Yves Meyer. Variational methods in image processing. Technical Report 04-52, UCLA, 2004. CAM Reports.

[20] Robert E. Megginson. *An introduction to Banach space theory*, volume 183 of *Graduate texts in mathematics*. Springer, New York, 1989.

[21] Yves Meyer. Oscillating patterns in image processing and nonlinear evolution equations. *University Lecture Series*, 22, 2001.

[22] Stanley Osher, Martin Burger, Donald Goldfarb, Jinjun Xu, and Wotao Yin. An iterative regularization method for total variation based image restoration. *Multiscale Model Simul.*, 4(2):460–489, 2005.

[23] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60:259–268, 1992.

[24] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[25] Otmar Scherzer and Joachim Weickert. Relations between regularization and diffusion filtering. *J. Math. Imaging Vision*, 12(1):43–63, 2000.

[26] Roger Temam. *Navier Stokes equations*, volume 2 of *Studies in Mathematics and Applications*. North Holland Publisher Co., Amsterdam, New York, Oxford, 1977.