INVERSE TOTAL VARIATION FLOW

M. BURGER, K. FRICK, S. OSHER, AND O. SCHERZER

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

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

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 $u^\delta$ (for instance images), defined on a domain $\Omega$, by the solution of

\begin{align}
\frac{\partial u}{\partial t} &= -A(u) \quad \text{in } \Omega \\
u(0) &= u^\delta \\
\frac{\partial u}{\partial \nu} &= 0 \quad \text{on } \partial \Omega
\end{align}

with $A(u) := -\text{div} \left( g(|\nabla u|) \frac{\nabla u}{|\nabla u|} \right)$ and $\nu$ 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 $u^\delta$. 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 $g(x) = \hat{g}'(x)$ the functionals

\begin{equation}
F_k(u) = \frac{1}{2} \|u - u_k\|_{L^2}^2 + \alpha \int_{\Omega} \hat{g}(|\nabla u|) \, dx, \quad k = 0, 1, \ldots,
\end{equation}

Date: November 13, 2006.
are minimized iteratively and the minimizer is denoted by $u_{k+1}$. Classical semi

group theory assumes that $\hat{g} : [0, \infty) \rightarrow [0, \infty)$ 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 $u_{k+1} = (\text{Id} + \alpha A)^{-1}(u_k)$. From the semi

group generation theorem (cf. [12]) it follows that the limit

$$u(t) = \lim_{N \rightarrow \infty} \left( \text{Id} + \frac{t}{N} A \right)^{-N}(u^\delta)$$

exists for all $t \in [0, \infty)$ and that it solves (1). The operator $(\text{Id} + \frac{t}{N} A)^{-N}$ corre-

sponds to making $N$ iterative regularization steps with a regularization para-

meter $\alpha = t/N$, that is, $u^N = (\text{Id} + \frac{t}{N} A)^{-N}(u^\delta)$. In other words iterative regularization

and the implicit Euler method correspond if the regularization parameter $\alpha$ and

the time discretization $\Delta t$ 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

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

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

Lagrange equation

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

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

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

Identifying the regularization parameter $\alpha$ and a time discretization $\Delta t$ via

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

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

$$u - u^\delta = \Delta^p \left( \frac{\partial u}{\partial t} \right),$$

$$u(0) = u^\delta .$$

Here $\Delta^p(u) = \text{div} \left( \| \nabla u \|^{p-2} \nabla u \right)$ is the $p$-Laplacian. It is obvious that (7) degener-

ates for $p = 1$ and thus for $\alpha \rightarrow \infty$ it cannot be claimed that $\Delta t \rightarrow 0$. 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

$$\Delta^{-1} \hat{u} = \frac{\partial \hat{u}}{\partial t},$$

$$\hat{u}(0) = -u^\delta (+u_0)$$

where $\Delta^{-1}$ is the solution operator corresponding to Laplace’s equation with ho-
mogeneous Neumann boundary data.
The name inverse scale space is motivated from the fact that
\[ \lim_{t \to \infty} u(t) = u^\delta, \quad \lim_{t \to 0^+} u(t) = u_0. \]
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 mini-
mization of the functional
\[ F(u) = \frac{1}{2} \|u - u^\delta\|_{L^2}^2 + \alpha |Du|(\Omega), \]
and the associated total variation flow equation
\begin{align}
\frac{\partial u(t)}{\partial t} &= \text{div} \left( \frac{\nabla u}{|\nabla u|} \right) \quad \text{in } \Omega \\
u(0) &= u^\delta \\
\frac{\partial u}{\partial \nu} &= 0 \quad \text{on } \partial \Omega
\end{align}
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 \textit{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 \textit{Bregman distance regularization}.
This method consists in computing first a minimizer \( u_1 \) of (11). The updates are
determined successively by calculating
\[ u_{k+1} = \arg\min_{u \in L^2(\Omega)} \left\{ \frac{1}{2} \|u - u^\delta\|_{L^2}^2 + \alpha \left( |Du|(\Omega) - \langle s, u \rangle \right) \right\}, \]
where \( s \) is an element of the subgradient of the total variation semi norm in \( u_1 \).
Introducing the Bregman distance with respect to \( |Du|(\Omega) \) defined by
\[ D(u, \bar{u}) := |Du|(\Omega) - |D\bar{u}|(\Omega) - \int_{\Omega} s(u - \bar{u}) \]
(cf Definition 1) allows to characterize \( u_{k+1} \) in (13) as
\[ u_{k+1} = \arg\min_{u \in L^2(\Omega)} \left\{ \frac{1}{2} \|u - u^\delta\|_{L^2}^2 + \alpha D(u, u_k) \right\}. \]
As we show in this paper, this iterative process satisfies a \textit{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 $\mathcal{D}(\Omega) = C_c^\infty(\Omega)$ the space of continuously differentiable functions with compact support in $\Omega$.
3. For $p \in (1, \infty)$ $p^*$ denotes its conjugate, which satisfies $\frac{1}{p} + \frac{1}{p^*} = 1$. Let $\mu$ be a positive measure on $\Omega$ and $m \in \mathbb{N}$. $L^p(\Omega, \mathbb{R}^m; \mu)$ denotes the space of $p$ times $\mu$-integrable functions $f : \Omega \to \mathbb{R}^m$ with the norm
\[
\|f\|_{L^p(\Omega, \mathbb{R}^m; \mu)} = \left( \int_\Omega |f|^p \, d\mu \right)^{\frac{1}{p}}.
\]
If $\mu$ is the $n$ dimensional Lebesgue measure we simply write $L^p(\Omega)^m$ and if $m = 1$ we write $L^p(\Omega)$.

4. The dual of $L^p(\Omega, \mathbb{R}^m; \mu)$ is $L^{p^*}(\Omega, \mathbb{R}^m; \mu)$. The dual pairing with respect to $L^p(\Omega)$ and $L^{p^*}(\Omega)$ is denoted by $\langle \cdot, \cdot \rangle$.
5. $L^\infty(\Omega, \mathbb{R}^n; \mu)$ denotes the space of essentially bounded functions.

6. For $p \in (1, \infty)$ we define
\[
J_p^\ast(u, \tilde{u}) := J(u) - J(\tilde{u}) - \langle s, u - \tilde{u} \rangle.
\]

7. The Legendre-Fenchel conjugate of a convex functional $J : L^p(\Omega) \to \mathbb{R} \cup \{+\infty\}$ is the functional $J^* : L^{p^*}(\Omega) \to \mathbb{R} \cup \{+\infty\}$
\[
u^* \mapsto J^*(\nu^*) := \sup_{u \in L^p(\Omega)} \{ \langle \nu^*, u \rangle - J(u) \}.
\]
From the definition of the subdifferential it follows that $u \mapsto D_J(u, \tilde{u})$ is non negative.

**Example 1.** The duality mapping defined by
\[
\mathcal{J}_p : L^p(\Omega) \to L^{p^*}(\Omega), \quad u \mapsto |u|^{p-2} u
\]
satisfies:
(1) \[ \int_{\Omega} J_p(u) \, dx = \|u\|_{L^p}^p = \|J_p(u)\|_{L^p}^p. \]

(2) \( J_p^{-1} = J_p. \) In particular \( J_2 \equiv \text{Id}. \)

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

\[ J_p(u) = \partial \|u\|_{L^p}^p. \]

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

\[ J(u) := \sup_{z \in \mathcal{D}(\Omega)^n} \int_{\Omega} u \, \text{div}(z) \, dx \]

the total variation semi norm of \( u \). Then

\[ J^*(v) = \chi_{K_p^*(\Omega)}(v) = \begin{cases} 0 & \text{if } v \in K_p^*(\Omega) \\ +\infty & \text{else} \end{cases}. \]

Proof. Let \( u \in L^p(\Omega) \), then

\[ (\chi_{K_p^*(\Omega)})^*(u) = \sup_{v \in K_p^*(\Omega)} \int_{\Omega} uv \, dx \]

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 \( \chi_{K_p^*(\Omega)}(v) = (\chi_{K_p^*(\Omega)})^{**}(v) = J^*(v) \) (see e.g. [14, Chap. 1, Prop. 4.1]). \( \square \)

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 \( L^p(\Omega) \) is the characteristic function of the set \( K_p^*(\Omega) \). Y. Meyer [21] introduced the \( G \)-norm which has proven to be a feasible instrument to describe 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

\[ X_p(\Omega) = \{ z \in L^\infty(\Omega)^n : \text{div}(z) \in L^p(\Omega) \} \]

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

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

and

\[ \int_{\Omega} u \, \text{div}(z) \, dx + \int_{\Omega} \nabla u z \, dx = \int_{\partial \Omega} T_p(z) u \, d\mathcal{H}^{n-1} \]

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

\[ L^p_\mathcal{D}(\Omega) = \left\{ v \in L^p(\Omega) : \int_{\Omega} v \, dx = 0 \right\}. \]

Proposition 1. Let \( 1 < p < \infty \), then

\[ L^p_\mathcal{D}(\Omega) = \overline{\text{div}(\mathcal{D}(\Omega)^n)} \]

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 \( \text{div} (z_k) \to v \) w.r.t. the \( L^p(\Omega) \)-norm. Then
\[
\int_{\Omega} \text{div} (z_k) \, dx = \lim_{k \to \infty} \int_{\Omega} \text{div} (z_k) \, dx = 0.
\]
If conversely \( v \in L_p^e(\Omega) \) there exist a sequence \( v_k \in \mathcal{C}^\infty(\overline{\Omega}) \cap L^p(\Omega) \) such that \( v_k \to v \) in \( L^p(\Omega) \) (From the density of \( \mathcal{D}(\Omega) \) in \( L^p(\Omega) \) the existence of an approximating sequence \( \{v_k\} \subseteq \mathcal{D}(\Omega) \) follows; Then \( v_k - \frac{1}{k^n} \int_{\Omega} v_k \, dx \to v \)). We abbreviate \( \bar{p} = \max(p, n) \) and observe that \( v_k \in L^p_\bar{p}(\Omega) \) for all \( k \in \mathbb{N} \). From [6, Thm.3] it follows that there exists \( z_k \in \mathcal{C}(\overline{\Omega})^n \cap W^{1, \bar{p}}_0(\Omega)^n \) such that
\[
\text{div} (z_k) = v_k.
\]
Therefore we can approximate each \( z_k \) by an element \( \tilde{z}_k \in \mathcal{D}(\Omega)^n \) such that
\[
\|\text{div} (\tilde{z}_k - z_k)\|_{L^p} \leq C \|\text{div} (\tilde{z}_k - z_k)\|_{L^p} \leq \frac{1}{k}
\]
for a constant \( C > 0 \). This implies that \( \text{div} (\tilde{z}_k) \to v \). \( \square \)

The \textit{G-norm} of an element \( v \) as introduced by Meyer [21] and Aubert and Aujol [4] is defined as the infimum of \( \|z\|_{L^\infty(\Omega)^n} \) taken over all \( z \) such that \( \text{div} (z) = v \). In order to provide a generalization to our case it is important to note that the above proof does not imply existence of an element \( z \in L^\infty(\Omega)^n \) such that \( \text{div} (z) = v \). However from (20) it follows that at least
\[
\ker (T_p) \subseteq L_p^e(\Omega).
\]
and [6, Thm. 3] proves

**Lemma 2.** If \( p \geq n \) we have that
\[
\ker (T_p) = L_p^e(\Omega).
\]

**Definition 2.** Let \( 1 < p < \infty \) and \( v \in L_p^e(\Omega) \). The G-norm of \( v \) is defined as
\[
\|v\|_* = \inf \left\{ \liminf_{k \to \infty} \|z_k\|_{L^\infty(\Omega)^n} : \{z_k\} \subseteq \mathcal{D}(\Omega)^n, \lim_{k \to \infty} \|\text{div} (z_k) - v\|_{L^p} = 0 \right\}
\]
Moreover we introduce the space
\[
\mathcal{L}_p^e(\Omega) = \left\{ v \in L_p^e(\Omega) : \|v\|_* < \infty \right\}.
\]

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

**Proposition 2.** For every \( v \in \mathcal{L}_p^e(\Omega) \) there exists \( z \in \ker (T_p) \) such that
\[
\text{div} (z) = v, \quad \|z\|_{L^\infty(\Omega)^n} \leq \|v\|_*.
\]

**Proof.** We follow the proof of [4, Lemma 2.3]. Let \( v \in \mathcal{L}_p^e(\Omega) \), 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
\[
\lim_{k \to \infty} \|\text{div} (z_k) - v\|_{L^p} = 0, \quad \lim_{k \to \infty} \|z_k\|_{L^\infty(\Omega)^n} = \|v\|_*.
\]
Then the boundedness of \( \{z_k\} \) implies the existence of an element \( z \in L^\infty(\Omega)^n \) such that – up to an extraction of a subsequence – \( z_k \) converges weakly* to \( z \) in \( L^\infty(\Omega)^n \).
From (20) it follows that for every $\phi \in C^\infty(\overline{\Omega})$

$$\int_{\Omega} \phi v \, dx = \lim_{k \to \infty} \int_{\Omega} \phi \, \text{div} (z_k) \, dx = \lim_{k \to \infty} -\int_{\Omega} \nabla \phi \, z_k \, dx$$

$$= -\int_{\Omega} \nabla \phi \, z \, dx = \int_{\Omega} \phi \, \text{div} (z) \, dx - \int_{\partial \Omega} \phi T_p(z) \, d\mathcal{H}^{n-1}.$$  

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

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

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 $\|\cdot\|_{L^\infty(\Omega)}$ implies

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

\[\square\]

A straightforward consequence of (23) is that

$$K^n(\Omega) \subseteq \{ v \in L^p_0(\Omega) : \|v\|_* \leq 1 \}.$$  

Indeed, if $\{z_k\} \subseteq D(\Omega)^n$ such that $\|z_k\|_{L^\infty(\Omega)} \leq 1$ and $\text{div} (z_k) \to v \in L^p_0(\Omega)$ then

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

We further investigate the relation between the topology induced by $\|\cdot\|_*$ and the weak topology on $L^p_0(\Omega)$.

**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 $z_k(0) = z_k(\pi) = 0$, $v_k = z_k' \in L^2_0(\Omega)$ and we have that

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

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

This example shows that convergence w.r.t. $\|\cdot\|_*$ does not imply weak convergence w.r.t. $L^2(\Omega)$. However under additional assumptions the result is true:

**Proposition 3.** Let $\{v_k\}_{k \in \mathbb{N}}$ be a sequence in $L^p_0(\Omega)$ and $v \in L^p_0(\Omega)$ 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 $L^p(\Omega)$ we can extract a subsequence $\{v_{k(l)}\}$ such that

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

for some $\hat{v} \in L^p_0(\Omega)$. Moreover let $v \in L^p_0(\Omega)$ such that $v_k \rightharpoonup v$ w.r.t. $\|\cdot\|_*$. Then $\hat{v} \in L^p(\Omega)$ and $\|z_{k(l)}\|_{L^\infty(\Omega)} \leq \|v_{k(l)} - v\|_*$ for all $l \in \mathbb{N}$. This together with (20) shows that $\int_{\Omega} u (v_{k(l)}) - v \, dx = -\int_{\Omega} \nabla u z_{k(l)} \, dx$ for all $u \in W^{1,1}(\Omega) \cap L^p(\Omega)$ and therefore

$$\lim_{k \to \infty} \int_{\Omega} u v_{k(l)} \, dx = \int_{\Omega} uv \, dx.$$
This together with (25) shows $\hat{v} = v$. Consequently every subsequence of $\{ v_k \}$ has in turn a weakly convergent subsequence with limit $v$. This implies that $v_k \rightarrow v$. □

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 $v \in L^p_0(\Omega)$ there exists a $|Du|$-measurable function $\theta(v, Du, \cdot) : \Omega \rightarrow \mathbb{R}$ that satisfies

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

and

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

Moreover let $f : \mathbb{R} \rightarrow \mathbb{R}$ 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 $\text{div}(z) = v$ and $\|z\|_{L^\infty(\Omega) \ast} \leq \|v\|_\ast$. From Theorems C.7 and C.9 in [3] it follows that there exists a function $\theta(v, Du, \cdot) \in L^\infty(\Omega, \mathbb{R}, |Du|)$ such that

$$
\int_{\Omega} uv \, dx = \int_{\Omega} \text{div}(z) \, dx = \int_{\Omega} \theta(v, Du, x) \, d|Du|
$$

and

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

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

**Corollary 1.** For all $u \in BV(\Omega) \cap L^p(\Omega)$ and $v \in L^p_0(\Omega)$ we have that

$$
\left| \int_{\Omega} uv \, dx \right| \leq |Du|(\Omega) \|v\|_\ast.
$$

Note that for those $u \in BV(\Omega) \cap L^p(\Omega)$ satisfying $J_{p, \lambda}(u) \in L^p_0(\Omega)$ Corollary 1 implies that

$$
\|u\|_{L^p_\ast} = \int_{\Omega} u J_{p, \lambda}(u) \, dx \leq |Du|(\Omega) \|J_{p, \lambda}(u)\|_\ast.
$$

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

$$
\begin{align*}
 u_{k+1} &:= \arg\min_{u \in L^p(\Omega)} I(\alpha; u, u_k), \quad \text{where } I(\alpha; u, \tilde{u}) := \left\{ \frac{1}{p} \|u - u^\delta\|_{L^p}^p + \alpha D^\delta_J(u, \tilde{u}) \right\}, \\
 v_{k+1} &:= v_k + \frac{1}{\alpha} \nabla J(u^\delta - u_{k+1}).
\end{align*}
$$

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 $u_0 \in BV(\Omega)$ and $v_0 \in \partial J(u_0)$.
- For $k = 0, 1, \ldots$
  
  $u_{k+1} := \arg\min_{u \in L^p(\Omega)} I(\alpha; u, u_k)$.
  
  $v_{k+1} := v_k + \frac{1}{\alpha} \nabla J(u^\delta - u_{k+1})$. 

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 \textit{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
\[
\alpha v_k + J'_p(u_k - u^\delta) = \alpha v_{k-1}.
\]

**Proof.** Let \( \tilde{u} \in L^p(\Omega) \) 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 \( L^p(\Omega) \), the Bregman distance
\[
u \mapsto D^*_p(u, \tilde{u})
\]
is weakly lower semicontinuous. Therefore \( I(\alpha; \cdot, u_k) \) is weakly lower semicontinuous on \( L^p(\Omega) \). It remains to show that \( I(\alpha; \cdot, u_k) \) is coercive on \( L^p(\Omega) \), that is
\[I(\alpha; u, u_k) \geq \beta \|u\|_{L^p}^p + \gamma\]
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 \( u_k \in L^p(\Omega) \cap BV(\Omega) \) and \( v_k \in \partial J(u_k) \). Since \( J \) is convex, \( v_0 \in \partial J(u_0) \) is equivalent to
\[J(u) \geq J(u_0) + \langle v_0, u - u_0 \rangle \text{ for all } u \in L^p(\Omega).
\]

Therefore,
\[
\frac{1}{p} \left\| u - u^\delta \right\|_{L^p}^p + \alpha (J(u) - \langle v_0, u \rangle) \geq \frac{1}{p} \left\| u - u^\delta \right\|_{L^p}^p + \alpha (J(u_0) - \langle v_0, u_0 \rangle)
\]
\[
\geq \frac{1}{p} \left\| u \right\|_{L^p}^p - \left\| u^\delta \right\|_{L^p}^p + \alpha (J(u_0) - \langle v_0, u_0 \rangle).
\]

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 \( u_1 \in BV(\Omega) \cap L^p(\Omega) \) which satisfies the Euler-Lagrange equation
\[
v_1 := \frac{3p(u^\delta - u_1)}{\alpha} \in \partial J(u_1).
\]

**Lemma 4.** For every \( k \in \mathbb{N} \) we have
\[
\|u_{k+1} - u^\delta\|_{L^p} \leq \|u_k - u^\delta\|_{L^p}.
\]

**Proof.** (See [22, Prop. 3.2]). Since the Bregman distance is nonnegative we have
\[
\frac{1}{p} \left\| u_{k+1} - u^\delta \right\|_{L^p}^p \leq \frac{1}{p} \left\| u_{k+1} - u^\delta \right\|^p_{L^p} + \alpha D^{*k+1}_J(u_{k+1}, u_k)
\]
\[
= I(\alpha; u_{k+1}, u_k) \leq I(\alpha; u_k, u_k) = \frac{1}{p} \left\| u_k - u^\delta \right\|^p_{L^p}.
\]

Using Lemma 1 the dual formulation of Algorithm 1 can be derived. We consider the dual functional of \( J \) with respect to \( v \in L^{p^*}(\Omega) \), defined as follows:
\[
I^*(\alpha; v, \tilde{v}) := \frac{1}{p^*} \|v - \tilde{v}\|_{L^{p^*}}^{p^*} + \frac{1}{\alpha p^* - 1} \left( J^*(v) - \langle u^\delta, v \rangle \right), \quad \tilde{v} \in L^{p^*}(\Omega).
\]
Theorem 2. Assume that $p > 1$ and $u^\delta \in L^p(\Omega)$. Then $u_k, v_k$ as defined in Algorithm 1 satisfy
\begin{equation}
\tag{30}
v_k = \arg \min_{v \in K^{p^*}(\Omega)} I^*(\alpha; v, v_{k-1}) \in K^{p^*}(\Omega).
\end{equation}
Moreover, in particular, for $p = 2$
\begin{equation}
\tag{31}
u_k = (u^\delta + \alpha v_{k-1}) - \alpha \pi_{K^2(\Omega)} \left( \frac{u^\delta}{\alpha} + v_{k-1} \right),
\end{equation}
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
\begin{equation}
\tag{32}
I^*(\alpha; v, \bar{v}) = \begin{cases} \frac{1}{p^*} \|v - \bar{v}\|_{L^{p^*}}^2 - \frac{1}{\alpha p^*} \langle u^\delta, v \rangle & \text{if } v \in K^{p^*}(\Omega) \\ +\infty & \text{else} \end{cases}
\end{equation}
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 $\hat{v}_k$. From Lemma 1 it follows that $\hat{v}_k \in K^{p^*}(\Omega)$. It remains to show that $v_k = \hat{v}_k$. From the definition of $v_k$ in Algorithm 1 and Theorem 1 it follows that
\begin{equation}
\tag{33}
v_k = v_{k-1} - \frac{1}{\alpha} \partial p(u_k - u^\delta) \in \partial J(u_k).
\end{equation}

Then, from the duality relation (see for instance [14]) it follows that
\begin{equation}
\tag{34}
u_k \in \partial J^* \left( v_{k-1} - \frac{1}{\alpha} \partial p(u_k - u^\delta) \right).
\end{equation}
Moreover, since (33) is equivalent to
\[-\alpha (v_k - v_{k-1}) = \partial p(u_k - u^\delta)\]
it follows by applying Example 1 (3) that
\begin{equation}
\tag{35}
\partial p(v_k - v_{k-1}) \alpha p^*-1 - u^\delta = -u_k.
\end{equation}
Combination of (33), (34) and (35) shows that
\begin{equation}
\tag{36}
0 \in \partial p(v_k - v_{k-1}) \alpha p^*-1 - u^\delta + \partial J^*(v_k).
\end{equation}
Therefore, $v_k$ minimizes the functional $I^*(\alpha; \cdot, v_{k-1})$, which together with the fact that the minimizer is unique implies that $v_k = \hat{v}_k$.

For $p = 2$ minimization of $I^*(\alpha; \cdot, v_{k-1})$ is equivalent to minimizing the functional
\[v \to \frac{1}{2} \left\| v - \left( v_{k-1} + \frac{u^\delta}{\alpha} \right) \right\|_{L^2}^2 + \frac{1}{\alpha} J^*(v).\]
Therefore, from Lemma 1 it follows that $v_k = \pi_{K^2(\Omega)}(u^\delta + v_{k-1})$ and together with (33) we see that
\[u_k = (u^\delta + \alpha v_{k-1}) - \alpha \pi_{K^2(\Omega)} \left( \frac{u^\delta}{\alpha} + v_{k-1} \right).\]

\hfill \Box

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 $u^δ \in L^p(\Omega)$ and $\{u_k\}$ as defined in Algorithm 1. Then we have that
\[
\lim_{k \to \infty} \|u_k - u^δ\|_{L^p} = 0.
\]

Proof. Let $k > 1$. Since $u_k \in \partial J^*(v_k)$ we have that
\[
\langle v_{k-1} - v_k, u_k \rangle = -\langle v_k - v_{k-1}, u^δ - u_k \rangle = -\frac{1}{\alpha} \|u_k - u^δ\|_{L^p}^2,
\]
where the last equality follows from (33). This estimate combined with Lemma 4 gives
\[
\|u_k - u^δ\|_{L^p}^p \leq \frac{1}{k} \sum_{j=1}^{k} \|u_j - u^δ\|_{L^p}^p \leq \frac{\alpha}{k} \sum_{j=1}^{k} \langle v_j - v_{j-1}, u^δ \rangle = \frac{\alpha}{k} \left( \langle v_k, u^δ \rangle - \langle v_0, u^δ \rangle \right).
\]
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 $\|v_k\|_s \leq 1$) for all $k > 0$ we immediately see that
\[
\lim_{k \to \infty} \left\| \frac{v_k}{k} \right\|_s = 0.
\]
Moreover, Lemma 4 and (33) applied iteratively imply that
\[
\left\| \frac{v_k}{k} \right\|_{L^p}^p \leq \frac{1}{k} \left\| v_{k-1} - \frac{1}{\alpha} 3_p(u_k - u^δ) \right\|_{L^p}^p \leq \frac{1}{k} \left\| v_0 - \frac{1}{\alpha} \sum_{j=1}^{k} 3_p(u_j - u^δ) \right\|_{L^p}^p \leq \frac{\|v_0\|_{L^p}^p}{k} + \frac{1}{\alpha k} \sum_{j=1}^{k} \|u_j - u^δ\|_{L^p}^p \leq \frac{\|v_0\|_{L^p}^p}{k} + \frac{1}{\alpha} \left\| u_0 - u^δ \right\|_{L^p}^p.
\]
We finally apply Proposition 3 and get
\[
\frac{v_k}{k} \to 0, \quad \text{in } L^p(\Omega).
\]

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 $\phi : \mathcal{S} \to \mathbb{R}$. 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 \(\bar{\tau} = \{\tau^k\}_{k \in \mathbb{N}}\) such that
\[
|\bar{\tau}| = \sup_{k \in \mathbb{N}} \tau^k < \infty
\]
and the associated sequence \(\{t^k_{\bar{\tau}}\}_{k \in \mathbb{N}_0}\) defined by \(\tau^k = t^k_{\bar{\tau}} - t^{k-1}_{\bar{\tau}}\) satisfies
\[
0 = t^0_{\bar{\tau}} < t^1_{\bar{\tau}} < \cdots < t^k_{\bar{\tau}} < \cdots,
\]
\[
\sum_{j=1}^{\infty} \tau^j = +\infty.
\]
For a given partition \(\bar{\tau}\) and \(u_0 \in \text{BV}(\Omega) \cap L^p(\Omega)\) resp. \(v_0 \in \partial J(u_0)\) we set for \(k = 1, 2, \ldots\)
\[
U^0_{\bar{\tau}} = u_0, \quad V^0_{\bar{\tau}} = v_0
\]
\[
U^k_{\bar{\tau}} = \arg\min_{u \in L^p(\Omega)} I\left(\frac{1}{\tau^k}; u, U^{k-1}_{\bar{\tau}}\right), \quad V^k_{\bar{\tau}} = \arg\min_{v \in L^p(\Omega)} I^*\left(\frac{1}{\tau^k}; v, V^{k-1}_{\bar{\tau}}\right)
\]
The discrete values of \(U^k_{\bar{\tau}}\) resp. \(V^k_{\bar{\tau}}\) are extended to a piecewise constant function \(U_{\bar{\tau}}(t)\) resp. \(V_{\bar{\tau}}(t)\) for \(t \in [0, \infty)\) defined by
\[
\bar{u}_{\bar{\tau}}(t) = U^k_{\bar{\tau}}, \quad \bar{v}_{\bar{\tau}}(t) = V^k_{\bar{\tau}},
\]
whenever \(t \in (t^k_{\bar{\tau} - 1}, t^k_{\bar{\tau}}]\). 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
\[
I^*(\alpha; v, \bar{v}) = \frac{1}{p_*} d(v, \bar{v})^{p_*} + \frac{1}{\alpha^{p_*-1}} \phi(v)
\]
where \(d(v, \bar{v}) = \|v - \bar{v}\|_{L^p}\) and \(\phi(v) = \chi_{K^{p_*}(\Omega)}(v) - \langle u^d, v \rangle\).

**Theorem 4.** Assume that \(\{\bar{\tau}\}_{\mathbb{N}}\) is a sequence of partitions of \([0, \infty)\) such that \(\lim_{t \to \infty} |\bar{\tau}| = 0\). Then there exists a function \(v \in C(0, \infty; K^{p_*}(\Omega))\), which is uniformly continuous, differentiable almost everywhere in \([0, \infty)\) and satisfies
\[
\|v'\|_{L^p} \in L^p_{\text{loc}}(0, \infty)
\]
such that
\[
v(t) = \lim_{t \to \infty} \bar{v}_{\bar{\tau}}(t) \text{ for all } t \in [0, \infty)
\]
in the weak topology of \(L^p(\Omega)\).

**Proof.** The weak \(L^p(\Omega)\) topology satisfies the topological assumptions in Section 2.1 in [2]. That is \(\phi\) is weakly lower semicontinuous on \(L^p(\Omega)\) (condition 2.1a) and from [14, Chap. 3 Prop. 1.1] we see that
\[
\inf_{v \in L^p(\Omega)} I^*(\alpha; v, 0) \geq - \inf_{u \in L^p(\Omega)} I(\alpha; u, 0) > -\infty, \forall \alpha > 0.
\]
which shows the coercivity assumption 2.1b. Weak compactness of \(L^{p_*}(\Omega)\) bounded subsets of sublevels of \(\phi\) (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
\[
\|v(t)\|_{L^p} \leq \|v_0\|_{L^p} + \|u_0 - u^d\|_{L^p} t, \quad \text{for all } t \geq 0.
\]
Proof. Let \( k, l \in \mathbb{N} \) and \( t \in \left[ t_l \frac{k}{n}, t_{l+1} \right] \). From (38) and (33) we conclude that

\[
\nabla \bar{\varphi}_l(t) = V_l^k = V_l^{k-1} - \tau_l \mathcal{J}_p \left( U_l^k - u^\delta \right) = v_0 - \sum_{j=1}^k \tau_l \mathcal{J}_p \left( U_l^j - u^\delta \right).
\]

From Lemma 4 it then follows that

\[
\begin{align*}
\left\| \nabla \bar{\varphi}_l(t) \right\|_{L^p}^p & \leq \left\| v_0 \right\|_{L^p}^p + \sum_{j=1}^k \left\| U_l^j - u^\delta \right\|_{L^p}^p \\
& \leq \left\| v_0 \right\|_{L^p}^p + \left\| v_0 - u^\delta \right\|_{L^p}^p + \left\| u_0 - u^\delta \right\|_{L^p}^p \tau_l^k \\
& \leq \left\| v_0 \right\|_{L^p}^p + \left\| u_0 - u^\delta \right\|_{L^p}^p t + O(\bar{\varphi}_l). \end{align*}
\]

If \( \bar{\varphi}_l \) is a sequence such that \( \nabla \bar{\varphi}_l(t) \to v(t) \) as in Theorem 4, weak lower semicontinuity of \( \left\| . \right\|_{L^p} \), finally shows that

\[
\left\| v(t) \right\|_{L^p}^p \leq \liminf_{t \to \infty} \left\| \nabla \bar{\varphi}_l(t) \right\|_{L^p}^p \leq \left\| v_0 \right\|_{L^p}^p + \left\| u_0 - u^\delta \right\|_{L^p}^p t.
\]

\( \square \)

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

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

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

\[
\left\| (1 - t)v_0 + tv_1 - w \right\|_{L^2}^2 = (1 - t) \left\| v_0 - w \right\|_{L^2}^2 + t \left\| v_1 - w \right\|_{L^2}^2 - t(1 - t) \left\| v_0 - v_1 \right\|_{L^2}^2,
\]

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

\[
\begin{align*}
\tilde{I}^*(\alpha; (1 - t)v_0 + tv_1, w) &= \frac{\alpha}{2} \left\| (1 - t)v_0 + tv_1 - w \right\|_{L^2}^2 - \langle (1 - t)v_0 + tv_1, w \rangle \\
&= \frac{\alpha}{2} \left( (1 - t) \left\| v_0 - w \right\|_{L^2}^2 + t \left\| v_1 - w \right\|_{L^2}^2 \right) \\
&= \frac{\alpha}{2} \left( (1 - t) \left\| v_0 - v_1 \right\|_{L^2}^2 - (1 - t) \langle v_0, u^\delta \rangle - t \langle v_1, u^\delta \rangle \right) \\
&= (1 - t)I^*(\alpha; v_0, w) + tI^*(\alpha; v_1, w) - \frac{\alpha}{2} t(1 - t) \left\| v_0 - v_1 \right\|_{L^2}^2.
\end{align*}
\]

Hence \( \tilde{I}^* \) satisfies the convexity assumption \([2, 4.0.1]\) and we can apply \([2, \text{ Thm. 4.2.2.}]\) to obtain that \( \nabla \tilde{\varphi}_l(t) \to v \) uniformly on each bounded interval \([0, T] \) as \( l \to \infty \). Particularly it follows that

\[
\lim_{l \to \infty} \nabla \tilde{\varphi}_l(t) = v(t), \quad \text{for all} \ t \in [0, \infty)
\]

in the strong topology on \( L^2(\Omega) \).

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 \( \alpha \) converges to \( +\infty \)). It remains to investigate under which conditions these functions satisfy gradient flow equations w.r.t. \( \phi \).

To this end we introduce the operator \( A : K^p(\Omega) \to L^p(\Omega) \) which is defined by

\[
(42) \quad A(v) = \argmin_{w \in \partial \phi(v)} \left\| u \right\|_{L^p}.
\]
and the slope
\[ |A|(v) := \begin{cases} \|A(v)\|_{L^p} & \text{if } v \in K^{p^*}(\Omega) \\ +\infty & \text{else} \end{cases}. \]

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
\[ |A|(v) = \operatorname*{lim inf}_{k \to \infty} |A|(v_k) : v_k \to v, \{ v_k \} \subseteq K^{p^*}(\Omega). \]

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 \( \partial J^* \)) according to [2, Lmm. 2.3.6.].

**Lemma 5.** The subgradient \( \partial J^* \subseteq L^{p^*_0}(\Omega) \times L^p(\Omega) \) is weakly - weakly closed, that is for all sequences \( \{ v_k \} \subseteq L^{p^*_0}(\Omega) \) and \( \{ u_k \} \subseteq L^p(\Omega) \) the following implication holds
\[ \sup_{k \in \mathbb{N}} J^*(v_k) < \infty, v_k \to v \text{ in } L^{p^*}(\Omega), \quad u_k \in \partial J^*(v_k), u_k \to u \text{ in } L^p(\Omega) \implies u \in \partial J^*(v). \]

**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 \( \langle u_k, w - v_k \rangle \leq 0 \) 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 \( \| u_k \|_{L^p} \| v_k \|_{L^{p^*}} \leq C \) for all \( k > 0 \) and consequently
\[ \sup_{k \in \mathbb{N}} J(u_k) = \sup_{k \in \mathbb{N}} \int_{\Omega} u_k v_k \, dx \leq C. \]

In other words the sequence \( \{ u_k \} \) is bounded in \( \text{BV}(\Omega) \). From compact imbedding \( \text{BV}(\Omega) \hookrightarrow L^1(\Omega) \) 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 \( L^1(\Omega) \) and \( u \in \text{BV}(\Omega) \).

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

\[ \int_{\Omega} v_k u_k \, dx = \int_{\Omega} \theta_k(v_k, D u_k, x) \, d|D u_k| \]

and
\[ \theta_k(v_k, D(S_{\lambda_k}^r(u_k)), x) = \theta_k(v_k, D u_k, x), \quad |D u_k| - \text{ a.e. } x \in \Omega, \]

(45)
\[ \theta_k(v_k, D(S_{\lambda_k}^r(u_k)), x) = \theta_k(v_k, D u_k, x), \quad |D u_k| - \text{ a.e.} \]

Note that (45) also holds for \( |D S_{\lambda_k}^r(u_k)| \) - almost every \( x \in \Omega \). Since \( \theta_k(v_k, D u_k, \cdot) \leq 1 \)

\[ \int_{\Omega} v_k S_{\lambda_k}^r(u_k) \, dx = \int_{\Omega} \theta_k(v_k, D(S_{\lambda_k}^r(u_k)), x) \, d|D S_{\lambda_k}^r(u_k)| = J(S_{\lambda_k}^r(u_k)). \]

Since \( \| u_k - u \|_{L^1} \to 0 \) it follows that \( \| S_{\lambda_k}^r(u_k) - S_{\lambda_k}^r(u) \|_{L^1} \to 0 \) as \( k \to \infty \) from the Lipschitz continuity of \( S_{\lambda}^r \) and \( S_{\lambda_k}^r(u_k) \) is uniformly bounded in \( L^\infty(\Omega) \) (by \( \lambda \)).
Together with the weak lower semicontinuity of $J$ Theorem 5. Therefore $S$ for a.e. $(46b)$ and therefore $p$

From Proposition 5 we know, that the slope Proof. $\lim_{k \to \infty} J(S^\varepsilon(u_k)) = \lim_{k \to \infty} \int_\Omega S^\varepsilon(u_k)v_k \ dx = \int_\Omega S^\varepsilon(u)v \ dx \leq J(S^\varepsilon(u)).$

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

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

$$2^{-p}\|S^\varepsilon(\lambda)(u) - u\|_L^p \leq \|S^\varepsilon(\lambda)(u) - S_\lambda(u)\|_L^p + \|S_\lambda(u) - u\|_L^p \leq |\lambda|^{-p} + \int_{\{|u| \geq \lambda\}} |u|^p \ dx.$$ Therefore $S^\varepsilon(\lambda)(u) \to u$ in $L^p(\Omega)$. Then, from strong closedness of $\partial J^*(v)$ (see e.g. [14, Thm. 1.5.1]) it follows that $u \in \partial J^*(v)$. □

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

$$\begin{align*}
\text{3}_{p^*}(v') & = -A(v(t)) \quad \text{(46a)} \\
v(0) & = v_0. \quad \text{(46b)}
\end{align*}$$

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 $\phi$). Moreover $u_0 \in BV(\Omega) \cap L^p(\Omega)$ 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 $F : L^{p^*}(\Omega) \to \mathbb{R},$

$$F(v) := -\int_\Omega u^\delta v \ dx + J^*(v), \quad \text{(47)}$$

and hence, (46) can be interpreted as a gradient flow for $F$ in $L^{p^*}(\Omega)$. 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

$$A(v(t)) = u(t) - u^\delta \text{ and } u(t) \in \partial J^*(v(t)) \quad \text{for all } t \in [0, \infty) \text{ and from duality we conclude that } v(t) \in \partial J(u(t)).$$

From these results it is evident that (46) is equivalent to

$$\begin{align*}
\text{3}_{p^*}(v') & = u^\delta - u(t), \quad \text{for all } t \in [0, \infty) \text{ and from duality we conclude that } v(t) \in \partial J(u(t)) \quad \text{(49a)} \\
v(0) & = v_0, \quad u(0) = u_0. \quad \text{(49b)}
\end{align*}$$

We now focus on basic properties of the solutions of (49).
Theorem 6 (Regularity). Assume that $u^\delta \in L^p(\Omega)$. Then, the solutions $(u, v)$ of (49) satisfy
\[ v \in C([0, \infty), L^{p^*}(\Omega)), \quad u \in L^\infty([0, \infty), L^p(\Omega)). \]
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
\[ \|v'(t)\|_{L^{p^*}} = \langle u^\delta, v'(t) \rangle \leq \|u^\delta\|_{L^p}, \quad \text{a.e. in } [0, \infty) \]
and consequently
\[ \|v'(t)\|_{L^{p^*}} \leq \|u^\delta\|_{L^p}^{\frac{1}{p^*}}. \]
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 $0 \leq t < \infty$ and note that
\[
\partial J^*(v(t)) = \{ u \in L^p(\Omega) : \langle u, w \rangle \leq \langle u, v(t) \rangle \text{ for all } w \in K^{p^*}(\Omega) \} \\
= \{ u \in L^p(\Omega) : \sup_{w \in K^{p^*}(\Omega)} \langle u, w \rangle \leq \langle u, v(t) \rangle \} \\
= \{ u \in L^p(\Omega) : J(u) = \langle u, v(t) \rangle \}.
\]
This shows that
\[ J(u(t)) \leq \|u(t)\|_{L^p} \cdot \|v(t)\|_{L^{p^*}} < \infty. \]

\[ \square \]

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
\[
\langle u_i - u_2, v_1 - v_2 \rangle \geq 0 \quad \text{for all } v_i \in K^2(\Omega), u_i \in \partial \chi_{K^2(\Omega)}(v_i) - u^\delta, i = 1, 2.
\]
Assume that there exist two solutions $v, \hat{v}$ of (46), that is,
\[
-v'(t) \in \partial \chi_{K^2(\Omega)}(v(t)) - u^\delta, \\
-\hat{v}'(t) \in \partial \chi_{K^2(\Omega)}(\hat{v}(t)) - u^\delta.
\]
From the monotonicity of the right hand sides it follows that
\[
\langle -v'(t) + \hat{v}'(t), v(t) - \hat{v}(t) \rangle \geq 0
\]
and therefore
\[
\frac{d}{dt} \|v(t) - \hat{v}(t)\|_{L^2}^2 = \langle v'(t) - \hat{v}'(t), v(t) - \hat{v}(t) \rangle \leq 0.
\]
This shows that $v = \hat{v}$. \[ \square \]

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 $\bar{U}_{\hat{\tau}_i}$ and the function $u$ introduced in (48).

Theorem 7. Let $\{\bar{\tau}_i\}_{i \in \mathbb{N}}$ be a sequence of partitions of $[0, \infty)$ such that $\lim_{i \to \infty} \|\bar{\tau}_i\| = 0$. Then
\[ \lim_{i \to \infty} \|\bar{U}_{\bar{\tau}_i}(t) - u(t)\|_{L^p} = 0, \quad \text{a.e. in } [0, \infty). \]
Proof. For a given partition $\bar{\tau}$ of $[0, \infty)$ let $V^{k}_{\bar{\tau}}$ be as in (38), the iterative minimizers of the dual problem. We define a piecewise linear function $\bar{V}_{\bar{\tau}} : [0, \infty) \rightarrow L^{p} (\Omega)$ that linearly interpolates the discrete values $V^{k}_{\bar{\tau}}$. That is

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

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

$$\left\| \bar{V}_{\bar{\tau}}(t) - V_{\bar{\tau}}(t) \right\|_{L^{p}_{\Omega}} = \left\| \frac{1}{\tau_k} \left( (t^k - t) V^{k-1}_{\bar{\tau}} + (t - t^{k-1}) V^k_{\bar{\tau}} \right) - V^k_{\bar{\tau}} \right\|_{L^{p}_{\Omega}}$$

$$= \frac{t^k - t}{\tau_k} \left\| V^{k-1}_{\bar{\tau}} - V^k_{\bar{\tau}} \right\|_{L^{p}_{\Omega}} \leq \left\| V^{k-1}_{\bar{\tau}} - V^k_{\bar{\tau}} \right\|_{L^{p}_{\Omega}}.$$

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

$$\frac{1}{p^*} \left\| V^{k-1}_{\bar{\tau}} - V^k_{\bar{\tau}} \right\|_{L^{p*}_{\Omega}} - \tau_k^{p^*-1} \left\langle u^\delta, V^k_{\bar{\tau}} \right\rangle \leq \frac{1}{p^*} \left\| V^{k-1}_{\bar{\tau}} - v \right\|_{L^{p*}_{\Omega}} - \tau_k^{p^*-1} \left\langle u^\delta, v \right\rangle.$$

With the choice $v = V^{k-1}_{\bar{\tau}}$ it follows that

$$\frac{1}{p^*} \left\| V^{k-1}_{\bar{\tau}} - V^k_{\bar{\tau}} \right\|_{L^{p*}_{\Omega}} \leq \tau_k^{p^*-1} \left\langle u^\delta, V^{k-1}_{\bar{\tau}} - V^k_{\bar{\tau}} \right\rangle$$

$$\leq \tau_k^{p^*-1} \left\| u^\delta \right\|_{L^{p*}_{\Omega}} \left\| V^{k-1}_{\bar{\tau}} - V^k_{\bar{\tau}} \right\|_{L^{p*}_{\Omega}}.$$

which implies that $\left\| V^{k-1}_{\bar{\tau}} - V^k_{\bar{\tau}} \right\|_{L^{p*}_{\Omega}} = O(\lvert \bar{\tau} \rvert)$.

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

$$\lim_{\bar{\tau} \to \infty} \int_{\Omega} \bar{V}_{\bar{\tau}}(t) u \, dx = \lim_{\bar{\tau} \to \infty} \int_{\Omega} \left( \bar{V}_{\bar{\tau}}(t) - V_{\bar{\tau}}(t) \right) u \, dx$$

$$+ \lim_{\bar{\tau} \to \infty} \int_{\Omega} V_{\bar{\tau}}(t) u \, dx = \int_{\Omega} v(t) u \, dx.$$

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 $\left\| \bar{V}_{\bar{\tau}}(t) \right\|_{L^{p*}_{\Omega}} \leq M_t$ for all $\bar{\tau} \in \mathbb{N}$. Since $t \mapsto \left\| \bar{V}_{\bar{\tau}}(t) \right\|_{L^{p*}_{\Omega}}$ is continuous we have that $M = \sup_{t \in I} M_t < \infty$ and it follows that

$$\left\| \bar{V}_{\bar{\tau}}(t) \right\|_{L^{p*}_{\Omega}} \leq M$$

for all $t \in I$ and $\bar{\tau} \in \mathbb{N}$.

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

$$\lim_{\bar{\tau} \to \infty} \int_{\Omega_I} \bar{V}_{\bar{\tau}}u \, dt \, dx = \int_{0}^{T} \lim_{\bar{\tau} \to \infty} \int_{\Omega} \bar{V}_{\bar{\tau}}(x,t)u(x,t) \, dx \, dt$$

$$= \int_{0}^{T} \int_{\Omega} v(x,t)u(x,t) \, dx \, dt$$

or in other words that $\bar{V}_{\bar{\tau}} \to v$ in $L^{p}(\Omega_I)$. Moreover, since

$$\frac{\partial \bar{V}_{\bar{\tau}}(t)}{\partial t} = \frac{V^{k}_{\bar{\tau}} - V^{k-1}_{\bar{\tau}}}{\tau_{\bar{\tau}}^{k-1} - \tau_{\bar{\tau}}^{k}} \quad \text{for } t \in (\tau_{\bar{\tau}}^{k-1}, \tau_{\bar{\tau}}^{k})$$
we can apply [2, Thm. 2.3.3.], from which it follows that
\begin{equation}
\lim_{l \to \infty} \left\| \frac{\partial V_{\tilde{k}}}{\partial t} \right\|_{L^p(\Omega)} = \left\| v' \right\|_{L^p(\Omega)}.
\end{equation}
Therefore the sequence \( \left\{ \frac{\partial V_{\tilde{k}}}{\partial t} \right\}_{l \in \mathbb{N}} \) is bounded in \( L^p(\Omega) \) and consequently \( \{ V_{\tilde{k}} \}_{l \in \mathbb{N}} \) is bounded in the space
\[ Y := \left\{ y \in L^p(\Omega) : y' = \frac{\partial y}{\partial t} \in L^p(\Omega) \right\}. \]
Since the space \( Y \) with norm
\[ \| y \|_Y = \| y \|_{L^p(\Omega)} + \| y' \|_{L^p(\Omega)} \]
is a reflexive Banach space (see for instance [26, Chap. III]), there exists a weakly convergent subsequence \( \{ V_{\tilde{k}(l)} \}_{l \in \mathbb{N}} \) and a limit \( \tilde{v} \in Y \) satisfying
\begin{equation}
V_{\tilde{k}(l)} \to \tilde{v}, \quad \frac{\partial V_{\tilde{k}(l)}}{\partial t} \to \tilde{v}' \quad \text{as} \quad l \to \infty \quad \text{in} \quad L^p(\Omega).
\end{equation}
Together with (54) this shows \( \tilde{v} = v \). From (56) and (57) and the Radon-Riesz property of \( L^p(\Omega) \) (see [20, Chap. 2.5]) strong convergence of \( \left\{ \frac{\partial V_{\tilde{k}(l)}}{\partial t} \right\}_{l \in \mathbb{N}} \) 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
\begin{equation}
\left\| \frac{\partial V_{\tilde{k}(l)}}{\partial t}(t) - v'(t) \right\|_{L^p} \to 0 \quad \text{for almost all} \quad t \in I.
\end{equation}
Moreover, from (35) and (55) we obtain for \( t \in \left( t_{\tilde{k}(l)}^{k-1}, t_{\tilde{k}(l)}^k \right] \)
\[ \overline{V}_{\tilde{k}(l)}(t) = U_{\tilde{k}(l)}^k = u^\delta - \left[ \frac{1}{t_{\tilde{k}(l)}^k} - \frac{1}{t_{\tilde{k}(l)}^{k-1}} \right] 3_{p_*} \left( V_{\tilde{k}(l)}^k - V_{\tilde{k}(l)}^{k-1} \right) = u^\delta - 3_{p_*} \left( \frac{V_{\tilde{k}(l)}^k - V_{\tilde{k}(l)}^{k-1}}{t_{\tilde{k}(l)}^k} \right) = u^\delta - 3_{p_*} \left( \frac{\partial V_{\tilde{k}(l)}(t)}{\partial t} \right) \]
and from (33) that
\begin{equation}
\overline{V}_{\tilde{k}(l)}(t) \in \partial J(\overline{V}_{\tilde{k}(l)}(t)).
\end{equation}
From (58) and the norm-norm continuity of \( 3_{p_*} \) ([11, Thm 2.16]) it follows that the right hand side of (59) converges to a function \( \tilde{u}(t) \in L^p(\Omega) \) for almost all \( t \in I \) that satisfies
\[ 3_{p_*}(v') = u^\delta - \tilde{u}(t), \quad \text{almost everywhere in} \quad I. \]
The graph of \( \partial J \) is closed in \( L^p(\Omega) \times L^p(\Omega) \) with respect to the strong topology on \( L^p(\Omega) \) and weak topology on \( L^p(\Omega) \) (cf.[14, Chap. I Cor. 5.1]). That is, the set \( \{ (u, v) : v \in \partial J(u), u \in L^p(\Omega) \} \) satisfies: Let \( u_k \in L^p(\Omega) \) and \( v_k \in \partial J(u_k) \subseteq L^p(\Omega) \) satisfying \( v_k \to v \) in \( L^p(\Omega) \) and \( u_k \to u \) in \( L^p(\Omega) \), 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 $U_{\xi}$ has an almost everywhere convergent subsequence with limit $u$ (restricted to $I$) which means that

$$
\lim_{l \to \infty} \|U_{\xi}(t) - u(t)\|_{L^p} = 0, \quad \text{a.e. in } I.
$$

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

\begin{proof}

Let $T > t$ and $\{\xi_l\}_{l \in \mathbb{N}}$ be a partition of $[0, \infty)$ such that $\lim_{l \to \infty} |\xi_l| = 0$ 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 $k(l), k(l) \in \mathbb{N}$ satisfying $t \in (t_{\xi_{k(l)}}^{k(l)-1}, t_{\xi_{k(l)}}^{k(l)})$, $s \in (t_{\xi_{k(l)}}^{k(l)-1}, t_{\xi_{k(l)}}^{k(l)})$ and

$$
t_{\xi_{k(l)}}^{k(l)} < t_{\xi_{k(l)}}^{k(l)-1}.
$$

Then it follows from Lemma 4 that

$$
\|U_{\xi}(s) - u^\delta\|_{L^p} \leq \|U_{\xi}(t) - u^\delta\|_{L^p}, \quad \text{for all } l > l_0.
$$

With this we obtain the estimate

$$
\|u(s) - u^\delta\|_{L^p} \leq \|U_{\xi}(s) - u^\delta\|_{L^p} + \|U_{\xi}(s) - u(s)\|_{L^p}
$$

$$
\leq \|U_{\xi}(t) - u^\delta\|_{L^p} + \|U_{\xi}(s) - u(s)\|_{L^p}
$$

$$
\leq \|U_{\xi}(t) - u(t)\|_{L^p} + \|U_{\xi}(s) - u(s)\|_{L^p} + \|u(t) - u^\delta\|_{L^p}.
$$

Taking the limit $l \to \infty$ shows (61).
\end{proof}

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

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

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 \textit{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 $v$.

\begin{thm}

Let $u^\delta$ be normalized such that $J_p(u^\delta) \in L^p_0(\Omega)$. For $t \|J_p(u^\delta)\|_\ast \leq 1$, a solution of (49) is given by

$$
u(t) = 0, \quad v(t) = t J_p(u^\delta).
$$
\end{thm}
Proof. In the following let $t \leq \frac{1}{\|\mathcal{J}_p(u^\delta)\|_*}$. We verify that $(v, u)$ defined by (62) is indeed a solution of (49). Since $v' = \mathcal{J}_p(u^\delta) = \mathcal{J}_p(u^\delta - u)$ it follows that
\[ \mathcal{J}_p(u') = u^\delta - u. \]

After Proposition 2 there exists $z \in \ker(T_{p_*})$ such that $\text{div}(z) = \mathcal{J}_p(u^\delta)$ and $\|z\|_{L^\infty(\Omega)^n} \leq \|\mathcal{J}_p(u^\delta)\|_*$. Then, for all $\varphi \in C^1_c(\Omega)$ and all $t \in (0, T)$ we have
\[ J(\varphi) - J(u(t)) - \int_{\Omega} v(t) (\varphi - u(t)) \, dx = J(\varphi) - \int_{\Omega} t \mathcal{J}_p(u^\delta) \, \varphi \, dx \]
\[ = \int_{\Omega} (|\nabla \varphi| + t \, z \, \nabla \varphi) \, dx \]
\[ \geq (1 - t \|\mathcal{J}_p(u^\delta)\|_*) \int_{\Omega} |\nabla \varphi| \, dx \geq 0. \]

By standard continuity and density arguments we can now extend the inequality
\[ J(\varphi) - J(u(t)) - \int_{\Omega} v(t) (\varphi - u(t)) \, dx \geq 0 \]
to all $\varphi \in L^p(\Omega)$ and hence, $v(t) \in \partial J(u(t))$. Thus, $(v, u)$ is a solution of (49). \qed

Theorem 8 shows that the reconstruction does not change in an initial stage of the evolution up to time $t_* = \frac{1}{\|\mathcal{J}_p(u^\delta)\|_*}$, 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
\[ \frac{\mathcal{J}_p(u)}{\|\mathcal{J}_p(u)\|_*} \in \partial J(u). \]

**Proof.** For an arbitrary $w \in L^p(\Omega)$ we have
\[ J(u) + \left( \frac{\mathcal{J}_p(u)}{\|\mathcal{J}_p(u)\|_*}, w - u \right) = J(u) + \frac{1}{\|\mathcal{J}_p(u)\|_*} \left( \langle \mathcal{J}_p(u), w \rangle - \langle \mathcal{J}_p(u), u \rangle \right) \]
\[ = J(u) + \frac{1}{\|\mathcal{J}_p(u)\|_*} \left( \langle \mathcal{J}_p(u), w \rangle - \|u\|_{L^p}^p \right) \]
\[ = J(u) + \frac{1}{\|\mathcal{J}_p(u)\|_*} \langle \mathcal{J}_p(u), w \rangle - J(u) \]
\[ \leq \frac{1}{\|\mathcal{J}_p(u)\|_*} \|\mathcal{J}_p(u)\|_* J(w) = J(w). \]
\qed

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 $u^\delta$ be a simple function. Then, for $t \|\mathcal{J}_p(u^\delta)\|_* \geq 1$, a solution of (49) is given by
\[ u(t) = u^\delta, \quad v(t) = t_* \mathcal{J}_p(u^\delta). \]

**Proof.** We have noticed in Lemma 6 that for simple functions the inclusion
\[ v(t) = t_* \mathcal{J}_p(u^\delta) = \frac{\mathcal{J}_p(u^\delta)}{\|\mathcal{J}_p(u^\delta)\|_*} \in \partial J(u^\delta) = \partial J(u(t)) \]
holds. Moreover, \( \mathcal{J}_{p_\ast}(v')(t) = 0 = u^\delta - u(t) \) 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 \( \mathcal{J}_{p_\ast}(u^\delta) \in L^p_0(\Omega) \) and (63) holds for \( t > t_\ast \), then \( u^\delta \) is simple. Hence, the simple images are the ones that can be reconstructed at time \( t_\ast \).

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

\[
v : t \rightarrow \begin{cases} t \mathcal{J}_p(u^\delta) & \text{if } 0 \leq t \leq t_\ast, \\ t_\ast \mathcal{J}_p(u^\delta) & \text{else.} \end{cases} \\
u : t \rightarrow \begin{cases} 0 & \text{if } 0 \leq t \leq t_\ast, \\ u^\delta & \text{else.} \end{cases}
\]

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 \( \|\mathcal{J}_p(f)\|_* = 1 \). Hence, this simple step function will be reconstructed by the inverse total variation flow at time \( t_\ast = 1 \).

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 \( H > 0, R \in (0, 2) \),

\[
f(x) = \frac{H}{C} \begin{cases} -1 & \text{if } |x| \geq \frac{R}{2}, \\ (\frac{H}{R} - 1)^{p_\ast - 1} & \text{if } |x| < \frac{R}{2}, \end{cases}
\]

with \( C = 1 + (\frac{H}{R} - 1)^{p_\ast - 1} \). In this case we have \( \int_{-1}^{1} \mathcal{J}_p(f) \, dx = 0 \) and \( \mathcal{J}_p(f) = \frac{H}{R} \) for

\[
z(x) = \left( \frac{H}{C} \right)^{p_\ast - 1} \begin{cases} -(x + 1) & \text{if } x \leq -\frac{R}{2}, \\ (\frac{H}{R} - 1) - x & \text{if } |x| < \frac{R}{2}, \\ (1 - x) & \text{if } x \geq \frac{R}{2}. \end{cases}
\]

The function \( z \) satisfies \( z(-1) = z(1) = 0 \) and \( \|\mathcal{J}_p(f)\|_* = \|z\|_\infty = (\frac{H}{C})^{p_\ast - 1} \left( 1 - \frac{H}{R} \right) \) and consequently the reconstruction time is given by \( t_\ast = (\frac{H}{C})^{p_\ast - 1} \left( \frac{2}{R - R/2} \right) \). In Figure 1 the G-norm of \( \mathcal{J}_p(f) \) 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 \( \|\mathcal{J}_p(f)\|_* \) 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 \( |x| \geq R/2 \) 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 \( F : L^p(\Omega) \to R \) (cf. (47))

\[
F(v) := -\int_\Omega u^\delta v \, dx + J^\ast(v).
\]

Below we prove monotone decrease of the dual functional \( F \), which also yields a convergence of solutions \( u \) of (49) to \( u^\delta \). Theorem 10 can be considered as the continuous version of Theorem 3.
Theorem 10 (Inverse Fidelity). Let $u^δ ∈ L^p(Ω)$, let $(v, u)$ be a solution of (49) and let $F$ be defined by (47). Then $t ↦ F(v(t))$ is non-increasing (decreasing if $u(t) ≠ u^δ$) and we have

$$\lim_{t→∞} \|u(t) - u^δ\|_{L^p} = 0.$$  

If, in addition $u^δ ∈ BV(Ω)$, then we get the convergence rate

$$\|u(t) - u^δ\|_{L^p} ≤ \left( \frac{J(u^δ)}{t} \right)^\frac{p}{p-1}, \quad \text{for a.e. } t ∈ [0, ∞).$$

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

$$F(v(t)) - F(v(s)) = \langle v(s) - v(t), u^δ \rangle$$

$$≤ \langle v(s) - v(t), u^δ \rangle - \langle v(s) - v(t), u(t) \rangle$$

$$= -\langle v(t) - v(s), u^δ - u(t) \rangle,$$

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

$$\lim_{s→t} \frac{F(v(t)) - F(v(s))}{t-s} = \frac{∂}{∂t} F(v(t)) ≤ -\left( \frac{∂v}{∂t}(t), u(t) - u^δ \right) = -\|u(t) - u^δ\|_{L^p}^{p},$$

for a.e. $t ∈ [0, ∞)$. Hence, we may conclude that

$$F(v(t)) - F(v(τ)) ≤ -\int_τ^t \|u(s) - u^δ\|_{L^p}^p \, ds ≤ 0.$$

With $τ = 0$ and $v(0) = 0$ this gives

$$\int_0^t \|u(s) - u^δ\|_{L^p}^p \, ds ≤ -F(v(t)).$$

Combining this inequality with Corollary 2 yields

$$\|u(t) - u^δ\|_{L^p}^p ≤ \frac{1}{t} \int_0^t \|u(s) - u^δ\|_{L^p}^p \, ds ≤ \int_Ω \frac{v(t)}{t} u^δ \, dx.$$
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

\[
\lim_{t \to \infty} \frac{\| v(t) \|}{t} = \lim_{t \to \infty} \frac{\| v(t) \|}{t} \leq \lim_{t \to \infty} \frac{1}{t} = 0.
\]

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

\[
-F(v(t)) = \langle v(t), u^\delta \rangle \leq \sup_{v \in K^p(\Omega)} \langle v, u^\delta \rangle = J(u^\delta).
\]

\[\square\]

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 \( \sqrt{J(u^\delta)/t} \), and which is decreasing with the total variation of \( u^\delta \). 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 \( u \in BV(\Omega) \) and the zero super level set

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

Assume that

\[
|\partial C|(\Omega) := |D\chi_C|(\Omega) < \infty .
\]

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

\[
u^\delta(x) = u(x) + \delta(x) \quad \text{for all } x \in \Omega,
\]

we associate a noisy surface as the boundary of the super level set of \( u^\delta \)

\[
\partial C^\delta \quad \text{with} \quad C^\delta = \{ x \in \Omega : u^\delta(x) > 0 \}.
\]

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

\[
I(\alpha; u_1, 0) = \frac{1}{p} \int_{\Omega} |u_1 - u^\delta|^p + \alpha |D u_1|(\Omega)
\]

(66)

\[
= \frac{1}{p} \int_{\Omega} |u_1 - u^\delta|^p + \alpha \int_{-\infty}^{\infty} |\partial \{ x \in \Omega : u_1(x) > c \}|(\Omega) \, dc
\]

\[
\leq \frac{1}{p} \int_{-\infty}^{\infty} |u - u^\delta|^p + \alpha \int_{-\infty}^{\infty} |\partial \{ x \in \Omega : u(x) > c \}|(\Omega) \, dc .
\]

(67)
Equation (68) shows that $\alpha$ enforces a trade off between the amount of displacements of $u^\delta$ and $u_1$ and smoothness of the level sets of $u_1$ in average. We explain the second Bregman iteration step:

$$u_2 = \operatorname{arg min}_{u \in L^p(\Omega)} I(\alpha; u, u_1)$$

(69)

$$= \operatorname{arg min}_{u \in L^p(\Omega)} \left\{ \frac{1}{p} \| u - u^\delta \|_{L^p}^p + \alpha (J(u) - \langle v_1, u \rangle) \right\}.$$  

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

$$\int_\Omega u \text{div} \left( z \right) \, dx = \int_\Omega uv \, dx = \int_\Omega \theta(v, u, x) \, d|Du|$$

and $\| \theta(v, u, \cdot) \|_{L^\infty(\Omega, \mathbb{R}, |Du|)} \leq \| v \|_\ast \leq 1$. Thus there exists an angle $\gamma(v, u, x) \in [0, \pi)$ for all $x \in \Omega$ such that $\cos(\gamma(v, u, x)) = \theta(v, u, x) \ |Du| - \text{a.e. in } \Omega$ and that

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

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

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

where $\frac{Du}{|Du|}$ denotes the density of $Du$ w.r.t. $|Du|$ and $|D^\alpha u|$ the absolute continuous part of the measure $|Du|$. 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

$$u_2 = \operatorname{arg min}_{u \in L^p(\Omega)} \left\{ \frac{1}{p} \| u - u^\delta \|_{L^p}^p + \alpha (J(u) - \langle v_1, u \rangle) \right\}$$

$$= \operatorname{arg min}_{u \in L^p(\Omega)} \left\{ \| u - u^\delta \|_{L^p}^p + \alpha \int_\Omega (1 - \cos(\gamma(z_1, u, x))) \, d|Du| \right\}.$$  

where $\text{div} \left( z_1 \right) = v_1$. It is important to note that $\langle v_1, u_1 \rangle = |Du_1|(\Omega)$ and therefore $\gamma(v_1, u_1, \cdot) = 0$ $|D^\alpha u_1|$ a.e. in $\Omega$, 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 $u^\delta$ 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

$$\| \delta \|_{L^2} \leq \delta'.$$

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. $u_0 = 0$, $u_0 = 0$, $k = 0$

2. while $\| u_{k+1} - u^\delta \|_{L^2} \geq \delta'$

$$u_{k+1} = \operatorname{arg min}_{u \in L^2(\Omega)} \left\{ \frac{1}{2} \| u - (u^\delta + w_k) \|_{L^2}^2 + \alpha |Du|(\Omega) \right\},$$

$$w_{k+1} = w_k + (u^\delta - u_{k+1}),$$

$k \leftarrow k + 1.$
In the case $p = 2$ iterative Bregman distance regularization simplifies to subsequent minimization of the ROF functional with data $u^\delta + w_k$.

Let

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

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

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

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

the best approximation on $L^2 = \text{span}\{\psi_i : i = 1, \ldots, N\}$ of $u$ 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:

\begin{equation}
\sum_{i=1}^N \int_{\Omega} u_i \psi_i \psi_j + \alpha \frac{u_i}{\sqrt{|\nabla u_i|^2 + \varepsilon^2}} \nabla \psi_i \nabla \psi_j \, dx = \sum_{i=1}^N \int_{\Omega} (u_i^\delta + (w_k)_i) \psi_i \psi_j \, dx, \quad 1 \leq j \leq N.
\end{equation}

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 $k = 1, 2, \ldots$ solve $A_k u^{k+1} = b_k$, where

$$A_k = \int_{\Omega} \psi_i \psi_j + \frac{\alpha}{\sqrt{|\nabla u_i|^2 + \varepsilon^2}} \nabla \psi_i \nabla \psi_j \, dx$$

$$(b_k)_j = \sum_{i=1}^N \int_{\Omega} (u_i^\delta + (w_k)_i) \psi_i \psi_j \, dx.$$
$u_k$ correspond to solutions of (49) at times $t = 2k$. Figure 3 shows the denoised isosurfaces at steps 1,...,6.

![Figure 3. Denoised surfaces: Steps 1,..., 6 with $\alpha = 0.5$.](image)

**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**


