

Inverse Scale Spaces for Nonlinear Regularization

Johan Lie, Jan M. Nordbotten

April 19, 2006

Abstract

Error minimization of global functionals provides a natural setting for analyzing image processing and regularization. This approach leads to scale spaces, which in the continuous formulation are the solution of nonlinear partial differential equations. In this work we derive properties for a class of inverse scale-space methods. The main contribution of this paper is the development of a proof that the methods considered are convergent for convex regularization operators. The proof is based on energy methods and the Bregman distance. Further, estimates for convergence toward a clean image with noisy forcing data is provided in terms of both the L_2 norm and Bregman distances. This leads to natural estimates of optimal stopping scale for the inverse scale space method. These analytical results are discussed in the context of a numerical example.

1 Introduction

During the last two decades, variational image processing using partial differential equations has been intensively studied (see e.g. [2, 7]). One common task is noise removal or regularization, i.e from noisy data $f = g + \eta$, with η representing noise, one would like to reconstruct the underlying data g . This task can be handled by solving the minimization problem

$$\min_u \{J(u) + \lambda H(f, u)\}, \quad (1)$$

where $J(u)$ is a convex regularization functional, and $H(f, u)$ is a convex fidelity functional [6, 15, 20]. The variable u is defined as a function of the image coordinates and time throughout. In Section 3 we will show J and H commonly encountered in applications.

A minimum of (1) is the steady-state of the partial differential equation

$$\partial_t u = -\partial_u J(u) + \lambda \partial_u H(f, u), \quad (2)$$

with $u(x, t) = f$ for $t = 0$. [6, 15, 20]. Throughout this paper ∂_x will refer to the derivative (for functions) or subgradients (for functionals) with respect to the variable x . Thus $\partial_u J$ and $\partial_u H$ denote elements of the subgradients (or generalized gradients) of J and H respectively [1, 17].

For convex functionals, this system is known to approach a steady state as $t \rightarrow \infty$, which leads to a sequence $\{u_k\}_{k=1}^N$, where u_k is gradually more regular for increasing k . The notion of regularity however depends on both J and H [6]. Note that if $\lambda = 0$ in Equation (2), the set of images $\{u_k\}$ is said to form a scale space, with t as the scale [2, 21].

Recently, methodology from the inverse problem community has been used to construct inverse scale spaces [4, 5, 14, 22]. These scale spaces are highly related to Tikhonov type of regularization of ill-posed inverse problems [12, 13]. The idea as devised in [4, 5] is to construct a sequence of images u_k by solving the constrained partial differential equation

$$\partial_t p = -\partial_u H(f, u), \quad p \in \partial_u J(u), \quad (3)$$

with a constant initial condition u_0 . For this system it can be proved that $\lim_{k \rightarrow \infty} u_k = f$, and one may choose K in such a way that u_K is a regularized version of f [4]. We will later refer to this method as the direct inverse scale space method.

Inverse scale spaces for image denoising were motivated by Lysaker, Osher and Tai in [16]. Later, iterative refinement techniques have been studied in a large number of works including [18, 19]. While the direct inverse scale space

has many desirable theoretical properties, it is difficult to compute numerically, particularly in higher dimensions [4].

An interesting numerical method for inverse scale space modeling called relaxed inverse flow was devised in [5]. This method approximates the solution to Equation (1) by solving the coupled system of constrained partial differential equations

$$\begin{aligned}\partial_t u &= -p(u) + \lambda q(f + v, u) \\ \partial_t v &= \alpha q(f, u)\end{aligned}\quad (4)$$

with $p \in \partial_u J(u)$ and $q \in \partial_u H(f, u)$. The scalar function v is dependent on the image coordinates and time.

The system of equations (4) has the obvious advantage over Equation (3) that the time derivative appears on the primary variables u and v of the image, which are more regular than the subgradient p . The model numerically handles the challenging case of Total Variation flow, but convergence of the method was proved only when p is linear [4]. In this paper we prove that the relaxed inverse flow in fact is convergent for a class of linear and nonlinear operators, using energy-methods and Bregman distances [3, 8, 10]. In the context of image regularization, the Bregman distance has previously been used in numerical computations [19]. We continue by providing estimates of the behavior of the flow toward the unknown clean image, which leads to stopping criteria for the method. These stopping criteria are compared to the Discrepancy Principle commonly used in applications. Further, in Section 3 we extend the numerical investigation of the relaxed inverse flow before we conclude the paper in Section 4.

2 Convergence Estimates of the relaxed inverse flow

In this section we first prove convergence of the flow using energy methods and Bregman distances [3, 8, 10]. Subsequently, we give estimates for the solution of equations (4) to the clean image.

2.1 Convergence to the forcing data

The relaxed inverse flow corresponding to the minimization problem (1) is stated in equations

(4). If we fix the functional $H = \langle f, u \rangle$ as the L_2 inner product of f and u then we have the subgradient $q(f, u) = \partial_u H(f, u) = f - u$, and Equations (4) have the stationary solution

$$u = f, \quad (5)$$

$$v = \frac{p(f)}{\lambda}. \quad (6)$$

Previously it has been shown that the flow in Eq. (4) converge to this steadystate solution for linear p [4, 5]. To show that u and v converge to the stationary solution for a more general p we will first derive an energy for the system (4). Multiply equations (4) with test functions ω_u and ω_v , ($\omega_u, \omega_v \in C_0^\infty$) and integrate over the domain:

$$\begin{aligned}\int_{\Omega} \omega_u u_t dx &= \\ \int_{\Omega} [-p(u) + \lambda(f + v - u)] \omega_u dx, &\quad (7)\end{aligned}$$

$$\int_{\Omega} \omega_v v_t dx = \alpha \int_{\Omega} (f - u) \omega_v dx. \quad (8)$$

Where throughout this section $p(u) = \partial_u J(u)$. Define the deviations from the stationary solutions as:

$$\tilde{u} = u - f, \quad (9)$$

$$\tilde{v} = v - \frac{p(f)}{\lambda}. \quad (10)$$

Inserting equations (9) and (10) into equations (7) and (8) while choosing $\omega_u = \tilde{u}/\lambda$ and $\omega_v = \tilde{v}/\alpha$ we obtain:

$$\begin{aligned}\frac{1}{\lambda} \int_{\Omega} \tilde{u} \tilde{u}_t dx &= \\ \int_{\Omega} \left[\frac{p(f) - p(u)}{\lambda} - \tilde{u} + \tilde{v} \right] \tilde{u} dx, &\quad (11)\end{aligned}$$

$$\frac{1}{\alpha} \int_{\Omega} \tilde{v} \tilde{v}_t dx = - \int_{\Omega} \tilde{u} \tilde{v} dx. \quad (12)$$

We sum these equations to eliminate the mixed term, and get

$$\partial_t e = \frac{1}{\lambda} \int_{\Omega} [p(f) - p(u)] \tilde{u} dx - \int_{\Omega} \tilde{u}^2 dx, \quad (13)$$

where the energy of the solution is defined as

$$e = \frac{1}{2} \left[\frac{1}{\lambda} \|\tilde{u}\|_2^2 + \frac{1}{\alpha} \|\tilde{v}\|_2^2 \right]. \quad (14)$$

From the definition, e can never be negative, and $e = 0$ if and only if u and v are equal to the stationary solution. We are therefore interested in the sign of the right hand side of equation (13). We introduce the Bregman distance D in terms of a functional J [8, 19]

$$D(u, v) = J(u) - J(v) - \langle \partial J(v), u - v \rangle, \quad (15)$$

The Bregman distance can be seen as the difference between $J(u)$ and the first order approximation to $J(u)$ from $J(v)$, and is thus non-negative for continuously differentiable convex J . It is however not a distance in the sense of a metric. The triangle inequality does not hold, and D is not symmetric. For a strictly convex J , $D(u, v) = 0$ implies $u = v$, but not for a general J . The introduction of the Bregman distance allows us to rewrite Equation (13),

$$\partial_t e = - \left(\frac{1}{\lambda} D_s(f, u) + \|\tilde{u}\|_2^2 \right). \quad (16)$$

Here we have introduced the symmetric distance

$$D_s(u, v) = D(u, v) + D(v, u). \quad (17)$$

The sum of the elements in the parenthesis on the left hand side of Equation (16) are strictly positive for $u \neq f$. Further, since e is bounded below by zero, we have that as $t \rightarrow \infty$, the derivative of the energy must satisfy $\lim_{t \rightarrow \infty} \partial_t e = 0$. Since both the elements in the parenthesis are non-negative, they must tend to zero individually, and we have the following convergence result of the relaxed inverse scale space flow defined by Equations (4) subject to the convexity of the functional J .

Theorem 2.1. If $p(u)$ is the continuous sub-gradient of a convex functional $J(u)$, then the energy e as defined in Equation (14) will be strictly decreasing. For initial conditions satisfying $e(0) < \infty$, the energy will dissipate $\lim_{t \rightarrow \infty} \partial_t e = 0$, and in the limit both $\|f - u\|_2 \rightarrow 0$ and $D(f, u) \rightarrow 0$.

The operators p considered herein are sub-gradients of convex functionals on the domain of continuous functions. However, the requirement that e must be bounded places restrictions on the admissible forcing data f .

2.2 Convergence to the clean image in L_2

The forcing data f consists of the clean image g and noise η . Theorem 2.1 gives us proof of con-

vergence to the forcing data, however what we are really interested in is the behavior in terms of the clean data. In particular, if we have a good estimate of the error $\|u - g\|_2$, this will provide us with a criteria for an optimal stopping time t^* . We will now derive inequalities providing estimates for distances $\mathcal{D}(u, g)$, where the distance \mathcal{D} will in this section be the L_2 norm, and in the next section the symmetric sum of Bregman distances defined above.

Define the deviation of u with respect to the clean image g as

$$\hat{u} = u - g, \quad (18)$$

and the associated energy

$$E = \frac{1}{2} \left[\frac{1}{\lambda} \|\hat{u}\|_2^2 + \frac{1}{\alpha} \|\tilde{v}\|_2^2 \right]. \quad (19)$$

The time derivative of this energy is obtained by choosing $\omega_u = \hat{u}/\lambda$ and $\omega_v = \tilde{v}/\alpha$ in equations (7) and (8):

$$\begin{aligned} \partial_t E &= \partial_t e + \int_{\Omega} \left(v - \frac{p(u)}{\lambda} - \tilde{u} \right) \eta dx \quad (20) \\ &\leq \partial_t e + \left\| v - \frac{p(u)}{\lambda} - \tilde{u} \right\|_{\frac{\beta}{\beta-1}} \|\eta\|_{\beta}, \quad (21) \end{aligned}$$

where we have used Hölder's Inequality. We see that in the last line the L_{β} norm of the noise appears. We will treat this as a known quantity for some value of β . This is more formally stated as

Corollary 2.2. If $p(u)$ is the continuous sub-gradient of a convex functional $J(u)$, then the energy E as defined in Equation 19 will be decreasing at least as long as

$$\frac{D_s(f, u) + \lambda \|\hat{u}\|_2^2}{\|p(u) + \lambda(\tilde{u} - v)\|_{\frac{\beta}{\beta-1}}} \geq \|\eta\|_{\beta},$$

for all $\beta \geq 0$.

For $\eta = 0$, Corollary 2.2 trivially reduces to Theorem 2.1. Corollary 2.2 comes closer to giving information about the convergence to the true image, however it does not provide us with direct information about $\|\hat{u}\|_2$. We can obtain this information from Inequality (21) by inserting the definitions of the energies e and E :

Corollary 2.3. If $p(u)$ is the continuous sub-gradient of a convex functional $J(u)$, then the L_2 norm of the deviation from the true image $\|\hat{u}\|_2$ satisfies

$$\partial_t \|\hat{u}\|_2^2 \leq \partial_t \|\tilde{u}\|_2^2 + 2 \|p(u) + \lambda(\tilde{u} - v)\|_{\frac{\beta}{\beta-1}} \|\eta\|_{\beta}.$$

2.3 Convergence to the clean image in the Bregman distance

For applications, it may be desirable to have an estimate of the deviation $u - g$ in the Bregman distance D_s . This is also an interesting metric, since it is known that $D(u, g)$ is decreasing when $\|\eta\|_2 \leq \|u - g\|_2$ for the direct model stated in Equation (3) as shown in [4]. This property is known as the Discrepancy Principle.

By definition, the time derivative of the Bregman distance $D(g, u)$ can be written as

$$\begin{aligned} \partial_t D(g, u) &= -\partial_t J(u) - \partial_t \langle p(u), u - g \rangle \quad (22) \\ &= \partial_t D(f, u) + \langle \partial_t p(u), g - f \rangle \quad (23) \end{aligned}$$

By applying Hölder's Inequality we can bound the last term on the right hand side of the equation to obtain the following lemma

Lemma 2.4. If $p(u)$ is the continuous subgradient of a convex functional $J(u)$, the Bregman distance $D(g, u)$ satisfies

$$\partial_t D(g, u) \leq \partial_t D(f, u) + \|\eta\|_\beta \|\partial_t p(u)\|_{\frac{\beta}{\beta-1}}.$$

We have the equivalent derivation for the symmetric sum of Bregman distances

$$\partial_t D_s(u, g) = \partial_t [\langle p(u) - p(g), u - g \rangle] = \quad (24)$$

$$\partial_t [D_s(u, f) + \langle p(f) - p(g), u \rangle + \langle p(u), f - g \rangle]. \quad (25)$$

By applying Hölder's Inequality we can bound the last term on the right hand side of the equation to obtain

Lemma 2.5. If $p(u)$ is the continuous subgradient of a convex functional $J(u)$, the sum of Bregman distances $D_s(u, g)$ satisfies

$$\begin{aligned} \partial_t D_s(u, g) &\leq \partial_t D_s(u, f) + \\ &\|\eta\|_\beta \|\partial_t p(u)\|_{\frac{\beta}{\beta-1}} + \langle p(f) - p(g), \partial_t u \rangle. \end{aligned}$$

The applicability of this lemma is dependent on the availability of an approximation of the last term. If we consider the linear case, where $p(u) = -\nabla^2 u = \partial_{\frac{1}{2}} \|\nabla u\|_2^2$, we can for $\nabla(f - g) = \partial_t \nabla u = 0$ on $\partial\Omega$ simplify Equation (25) such that we obtain

$$\partial_t D_s(u, g) = \partial_t [D_s(u, f) + 2\langle p(u), \eta \rangle] \quad (26)$$

$$\leq \partial_t D_s(u, f) + 2\|\eta\|_\beta \|\partial_t p(u)\|_{\frac{\beta}{\beta-1}}. \quad (27)$$

For this special case we see that we obtain a bound on the Bregman distance which only depend on the L_β norm of the noise. We thus have

Corollary 2.6. If $p(u)$ is the continuous subgradient of the convex functional $J(u) = \frac{1}{2} \|\nabla u\|_2^2$, the sum of Bregman distances $D_s(u, g) = D(u, g) + D(g, u)$ satisfies

$$\partial_t D_s(u, g) \leq \partial_t D_s(u, f) + 2\|\eta\|_\beta \|\partial_t p(u)\|_{\frac{\beta}{\beta-1}}.$$

2.4 Stopping criteria

It is of interest to obtain an estimate for t^* defined such that $\|\hat{u}(t^*)\|_2 = \min_{t \in \mathbb{R}^+} \mathcal{D}(u, g)$. Here the distance \mathcal{D} refers to a distance measure of interest. Define t' as the time when

$$\partial_t \mathcal{D}(u, g) = 0. \quad (28)$$

We may expect that t' should be a good estimate of t^* , where $\partial_t \mathcal{D}$ can be estimated from the preceding sections. However, as observed by Burger et al [4], the both distances $\mathcal{D} = \|\hat{u}\|_2$ and $\mathcal{D} = D_s(u, g)$ may be oscillating such that t' is multivalued. They go on to show that for idealized cases oscillations can be avoided for parameters α and λ satisfying $\lambda \geq 4\alpha$. We here point out that this can be seen directly from equations (4) by considering them in zero spatial dimensions. For this special case, the spatial derivative vanishes, and the equations reduce to a single second order equation in time

$$0 = \partial_{tt} u + \lambda \partial_t u + \lambda \alpha u - \lambda \alpha f. \quad (29)$$

We see from the characteristics of the equation that oscillatory solutions appear for $\lambda < 4\alpha$.

The above discussion together with extensive experiments, as well as the results reported in [4] indicate that we can make the hypothesis

Hypothesis 2.7. There exists a critical value of λ dependent on α and the choice of distance \mathcal{D} , denoted λ_c such that for all for all λ satisfying

$$0 < \lambda \leq \lambda_c \leq 4\alpha \quad (30)$$

the distance \mathcal{D} will have at most one local minimum.

Under Hypothesis 2.7 and a choosing \mathcal{D} such that we can apply either Corollary 2.3, Lemma 2.5, or Lemma 2.4 we choose to define t^* as the minimum value of the (still possibly multivalued) variable t' .

In Section 3, we refer to the stopping criteria derived from Corollary 2.3 and Lemma 2.4 as the L_2 and Bregman criteria respectively. The corresponding stopping times are denoted by t'_{L_2} and t'_D . We further compare with the stopping time t'_η of the discrepancy principle of the corresponding direct inverse scale space flow.

3 Numerical Implementation and Experiments

In this section we investigate the behavior and show details of numerical implementation of a specific realization of the relaxed inverse flow, namely the relaxed inverse Total Variation (TV) flow. After a short introduction to the TV denoising model, we show some essential properties of the relaxed inverse TV scale space. Then we proceed to show numerical realizations of the various stopping criteria introduced in section 2.

The TV (TVL_2) model introduced by Rudin, Osher and Fatemi is a model of great importance in PDE image processing [20]. The model has the property that it preserves edges, but removes random oscillations and thus is suitable for image regularization purposes, see e.g [23]. Using this model, the regularization functional from (1) reads

$$J(u) = \int_{\Omega} |\nabla u| dx. \quad (31)$$

The subgradient of J is given by

$$p(u) = -\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right). \quad (32)$$

We restrict ourselves to show details of the numerical implementation of this flow in two spatial dimensions. The generalization to higher dimensions is however straight forward.

All derivatives are approximated using finite difference schemes. Superscripts denotes iteration index, while subscripts denotes spatial derivatives. P and Q denotes numerical approximations to p and q respectively. Time derivatives are approximated by forward Euler schemes, while spatial derivatives are approximated using forward and backward schemes which jointly form central difference schemes in

a standard way.

$$P(u^n) = -D_x^- \left(\frac{D_x^+ u^n}{\sqrt{(D_x^+ u^n)^2 + (D_y^+ u^n)^2 + \epsilon}} \right) - D_y^- \left(\frac{D_y^+ u^n}{\sqrt{(D_x^+ u^n)^2 + (D_y^+ u^n)^2 + \epsilon}} \right), \quad (33)$$

where ϵ is a regularization parameter. D^+ and D^- denotes forward and backward differences in the direction of the subscript. The finite difference schemes used here are consistent with the schemes from [9]. For the fidelity-term we use

$$Q(f, u^n) = f - u^n. \quad (34)$$

Thus the numerical scheme for the forward flow can be written analogously as the equations (4)

$$\begin{aligned} u^{n+1} &= u^n + \Delta t (-P(u^n) + \lambda Q(f + v^n, u^n)) \\ v^{n+1} &= v^n + \Delta t \alpha Q(f, u^n) \end{aligned} \quad (35)$$

By appropriately changing the terms P and Q , the scheme (35) can be used for other regularization flows. In order to be consistent with the theory introduced in section 2, we need to pick P as a numerical approximation to the subgradient of a convex functional $J(u)$ and Q as the L_2 fidelity measure.

3.1 Basic properties of relaxed inverse total variation flow

We will first illustrate the basic properties of the inverse scale space model. A more detailed treatment over a wide range of examples can be found in the papers by Burger et.al [4, 5]. In all simulations shown herein, we initialize u as the mean of f , and $v = 0$. This is a reasonable initial condition on u , reflecting that we are considering an inverse scale space method. The initial condition is consistent with the presentation in [4]. We point out that this initial condition (on u) corresponds to the steady state solution ($t = \infty$) of the forward scale space flow with $\lambda = 0$. In all simulations, we use the regularization parameter $\lambda = \frac{1}{10}$, and motivated by Hypothesis 2.7 we use $\alpha = \frac{\lambda}{3}$. Choosing an optimal regularization parameter is not in the scope of this paper, however we indicate that visually better looking results can be achieved by tuning the regularization parameter according to the noise level in f , and a priori knowledge of the

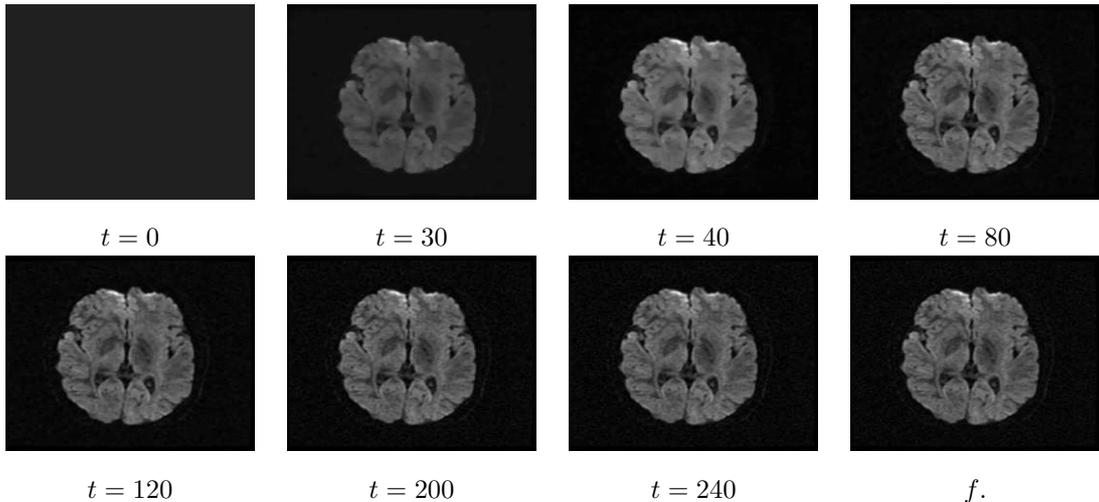


Figure 1: An illustration of the inverse scale space, with u evolving from the mean value of f at $t = 0$ towards f at $t = \infty$.

actual problem that is solved. The size of the regularization parameter should be somehow inversely proportional to the noise level. We do not study the exact relationship between λ_{opt} and $\|\eta\|_2$ in this paper, although optimal parameter selection is an important issue in regularization methods. For a treatment on optimal regularization parameter selection (although in a forward setting), see [11].

In Figure 1, the relaxed inverse TVL_2 model is used for regularization of a noisy Magnetic Resonance (MR) image. By studying Figure 1, we see that the relaxed inverse scale space method performs as intuitively expected, i.e large scales are evolving into u earlier than small scales. Note that u eventually converge to f if no stopping criteria is enforced, as proved in Section 2. By inspection of the time evolution, as shown in Figure 1, we see that after some time noise is beginning to evolve into the regularized image u . Therefore a stopping criteria must be chosen in such a way that the flow will stop at a time where u is more regular than f and still close to g . Such stopping criteria were discussed such stopping criteria theoretically in Section 2.4, and we look at numerical realizations in the next subsection.

3.2 Optimal stopping criteria.

The relaxed inverse flow must be stopped at an appropriate time in order to perform as a practical regularization method. In Section 2.4 a set of stopping criteria are discussed. In the

current section we investigate how the theoretically derived stopping criteria perform, both in terms of stopping the flow at a time which gives a visually appealing result, and in terms of predicting the actual minimum of the corresponding distance functions. In order to quantitatively compare the theoretical predictions t^* to the numerical stopping times t' we introduce the notation $[t] = t'/t^*$. If $[t] \approx 1$, the numerical stopping time approximates the theoretical stopping time well. In this section we discuss the stopping criteria corresponding to approximate minima of the distances $D(g, u)$ and $\|g - u\|_2$ from Section 2. We also compare the proposed stopping criteria with the commonly used Discrepancy Principle stopping criterium, see e.g [4].

In Figure 2 the MR image from Figure 1 is processed with the relaxed total variation inverse flow. The norms $\|u - f\|_2$ and $\|u - g\|_2$ are not monotone, as expected from the discussions in Section 2.4 and in the paper [4], since $\lambda < 4\alpha$. We show the result from stopping times t'_D, t'_{L_2} and t'_η corresponding to the stopping criteria from Section 2.4. In order to differentiate between the local and global minima of $\|u - g\|_2$ we denote the local minimum by $t'_{L_2^{loc}}$. For this specific application, we observe that the stopping times satisfy $t'_{L_2^{loc}} < t'_D < t'_{L_2} < t'_\eta$, which is investigated in more detail in Figure 4. We immediately see that criteria t'_D, t'_{L_2} and t'_η as shown in Figure 2(b)-(d) respectively give similar results, with only minute differences. Fur-

ther, we find all the regularized images $u(t'_D)$, $u(t'_{L_2})$ and $u(t'_\eta)$ visually appealing.

The above discussion can be seen clearer from Figure 3, where we have shown a profile of the true image, together with the corresponding profiles from Figure 2. In Figure 2(a) the transect shown in Figure 3 is indicated with a bright vertical line. The difference between $u(t'_D)$ and $u(t'_\eta)$ is also clearer in this image. In this specific application the stopping time t'_η seems better than t'_D .

To see the relationship between the different stopping times depicted in Figure 4 we show the energies of the system, together with various stopping times. As expected, the energy e of the system is always decreasing. Further, we see that both $\|u - g\|_2$ and $\|u - f\|_2$ have local minima, which is to be expected from the discussion surrounding Hypothesis 2.7. This local minimum is the reason that the first L_2 stopping time is significantly earlier than those obtained by considering the Bregman distance, i.e the image is over-regularized.

The relationship between the optimal stopping times and the approximate stopping times is more clearly understood from Table 1. Here we have given the ratio of approximate to optimal stopping times $[t]$. As expected from Corollary 2.3 and Lemma 2.4, we see that $[t]$ is strictly less than 1. From the exponential nature of the solution revealed by Equation (29), we expect that capturing the correct magnitude of t^* , will produce good results. Therefore, it is encouraging that for the example considered herein, t^* is never more than two times t' , and indeed, as discussed above, we see from figures 1 and 3 that the solution changes little between the different stopping times $u(t'_D)$ and $u(t'_\eta)$. This leads us to conclude that the stopping criteria derived herein have merit in practical applications.

4 Conclusion

In this work we have proved that the relaxed inverse flow given in Eq. (4) from [4, 5] is convergent for convex regularization functionals, using Bregman distances and the energy of the system. Thus the flow comprises an inverse scale space flow. The proposed energy e of Equation (4) is decreasing for all $u \neq f$. Further, we derive stopping criteria based on estimates of the time derivatives $\partial_t \|u - g\|_2$ and $\partial_t D(g, u)$. This makes it possible to stop the flow when u is

close to the true image, even if only an estimate of the noise level is provided. Our numerical experiments indicate that relaxed inverse scale space methods are well suited for image regularization purposes, and that the stopping criteria derived herein, are applicable to real images. For the images we have considered, the proposed stopping criteria based on estimates of the minimum value of $D(g, u)$ together with the standard Discrepancy Principle stopping criterium have given the best results. However, a more detailed analysis of optimal stopping times is a topic for future research.

Acknowledgements

This work was completed in part while visiting the Computational and Applied Mathematics group, Department of Mathematics at University of California, Los Angeles. We especially thank Stanley Osher for interesting discussions and for introducing the authors to Bregman distances. This research was funded in part through grant 159973 (BeMatA) from the Norwegian Research Council.

References

- [1] H. ATTOUCH, G. BUTTAZZO, AND G. MICHAILLE, *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*, vol. 6 of MPS-SIAM Series on Optimization, SIAM, 2006.
- [2] G. AUBERT AND P. KORNPORST, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, vol. 147 of Applied Mathematical Sciences, Springer-Verlag, 2002.
- [3] L. M. BREGMAN, *Relaxation method for finding common points of convex sets and its application in optimization*, Reports of the USSR Academy of Sciences, 171 (1966), pp. 1019–1022.
- [4] M. BURGER, G. GILBOA, S. OSHER, AND J. XU, *Nonlinear inverse scale space methods*, Communications in Mathematical Sciences, 4 (2006), pp. 179–212.
- [5] M. BURGER, S. OSHER, J. XU, AND G. GILBOA, *Nonlinear inverse scale space*

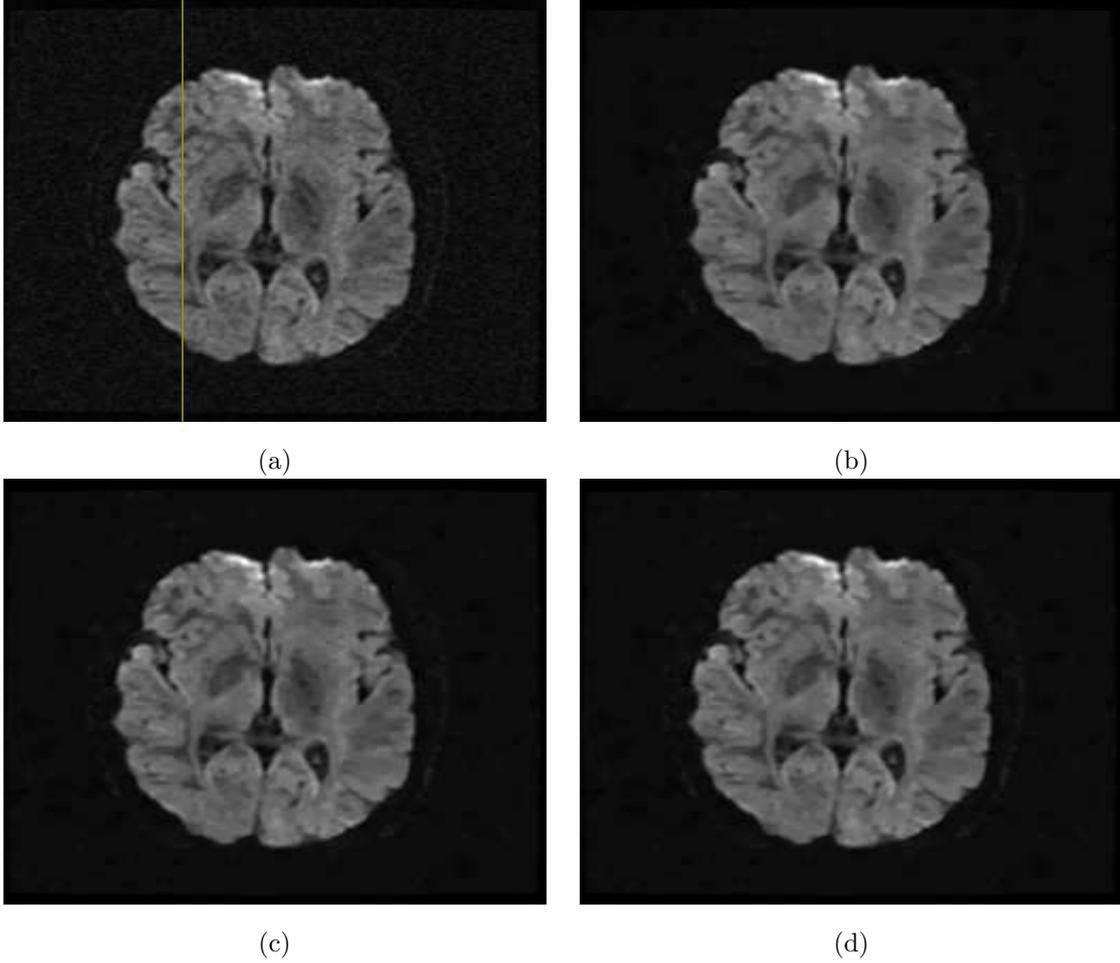


Figure 2: A 2D slice of a noisy MR image (a) is denoised using the relaxed inverse flow. (b)-(d) depicts u at stopping times t'_D , t'_{L_2} and t'_η respectively. The bright vertical line in (a) indicates the profile used in Figure 3.

Criterion	t^*	t'	$[t] = t'/t^*$
$D(g, u)$	85	63.2	0.74
$\ u - g\ _2^{loc}$	19,1	15,50	0.81
$\ u - g\ _2$	110	61.2	0.56
$\ \eta\ _2$		83.5	

Table 1: Stopping times t' estimated from the stopping criteria of Section 2.4 are compared with the times t^* corresponding to the true minima of the functionals involved.

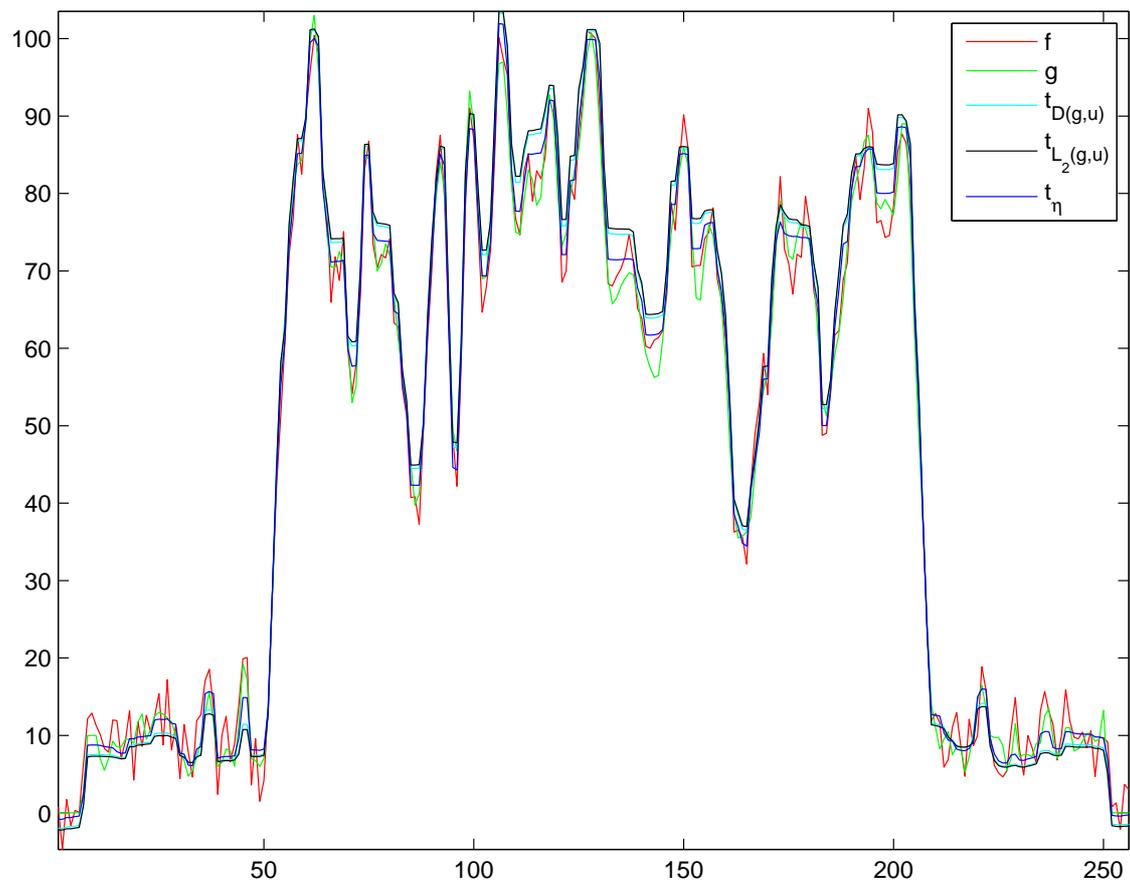


Figure 3: In this figure we show 1D profiles of u at various stopping times corresponding to the stopping criteria of Section 2.4.

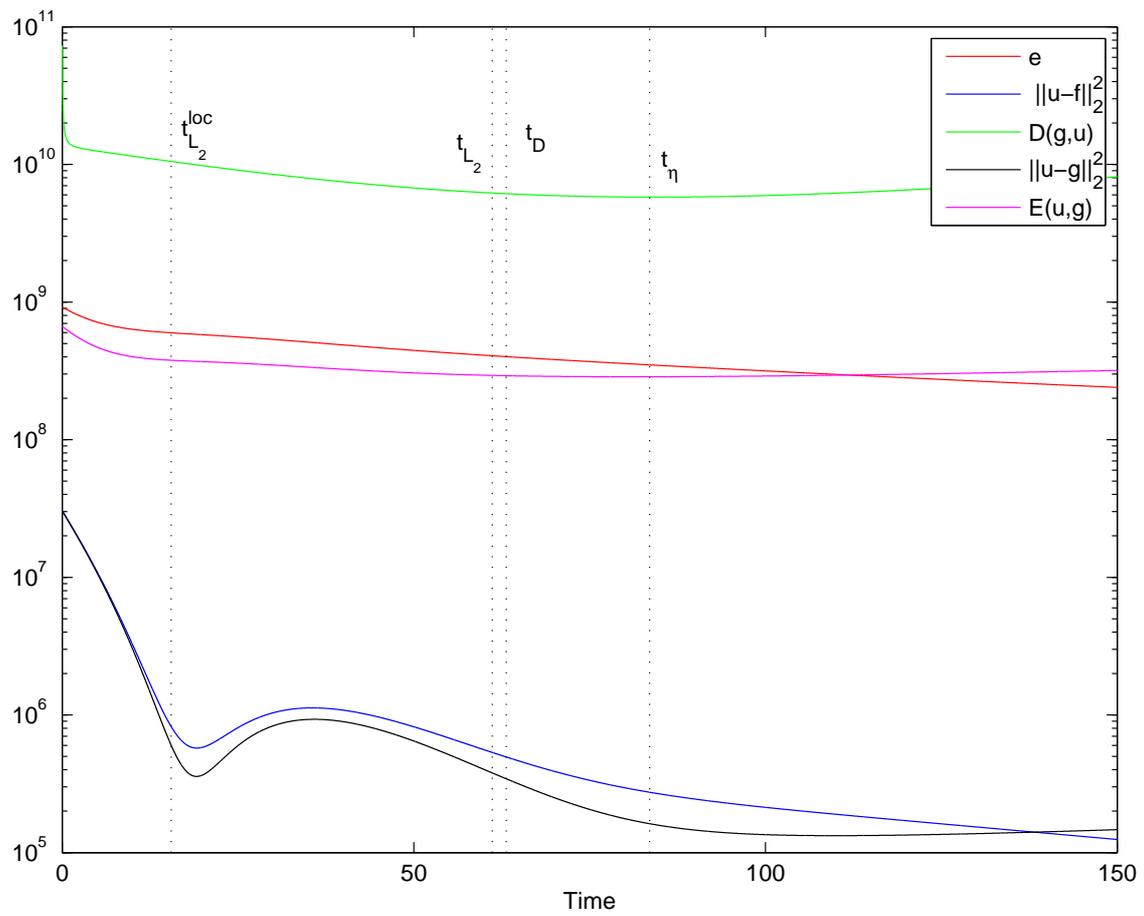


Figure 4: Energies and stopping times corresponding to the stopping criteria from Section 2.4.

- methods for image restoration*, Lecture Notes in Computer Science, 3752 (2005), pp. 25–36.
- [6] T. F. CHAN AND S. ESEDOGLU, *Aspects of total variation regularized L_1 function approximation*, SIAM Journal of Applied Mathematics, 65 (2005), pp. 1817–1837.
- [7] T. F. CHAN AND J. SHEN, *Image Processing And Analysis: Variational, PDE, Wavelet, And Stochastic Methods*, Society for Industrial and Applied Mathematics, 2005.
- [8] G. CHEN AND M. TEBoulLE, *Convergence analysis of a proximal-like minimization algorithm using Bregman functions*, SIAM J. Optimization, 3 (1993), pp. 538–543.
- [9] S. ESEDOGLU AND R. TSAI, *Threshold dynamics for the piecewise constant Mumford-Shah functional*, UCLA Computational and Applied Mathematics Report (04-63), (2004).
- [10] L. C. EVANS, *Partial Differential Equations*, AMS, 1998.
- [11] G. GILBOA, N. SOCHEN, AND Y. Y. ZEEVI, *Estimation of optimal PDE-based denoising in the SNR sense*, UCLA Computational and Applied Mathematics Report, 05-48 (2005).
- [12] M. HANKE AND P. C. HANSEN, *Regularization methods for large-scale problems*, Surveys Math. Indust., 3 (1993), pp. 253–315.
- [13] P. C. HANSEN, *Deconvolution and regularization with toeplitz matrices*, Numerical Algorithms, 29 (2002), pp. 323–378.
- [14] R. KIMMEL, R. MALLADI, AND N. SOCHEN, *Image processing via the Beltrami operator*, in Proc. of third Asian Conf. on Computer Vision, Hong Kong, January 1998.
- [15] M. LYSAKER, *Variational PDE techniques for image processing problems*, PhD thesis, 2004.
- [16] M. LYSAKER, S. OSHER, AND X. C. TAI, *Noise removal using smoothed normals and surface fitting*, IEEE Trans. Image Processing, 13 (2004), pp. 1345–1357.
- [17] J. NOCEDAL AND S. J. WRIGHT, *Numerical Optimization*, Springer Series in Operations Research, Springer-Verlag, New York, 1999.
- [18] S. OSHER, M. BURGER, D. GOLDFARB, J. XU, AND W. YIN, *Using geometry and iterated refinement for inverse problems (1): Total variation based image restoration*, UCLA Computational and Applied Mathematics Report, 04-13 (2004).
- [19] ———, *An iterative regularization method for total variation-based image restoration*, Multiscale Model. Simul., 4 (2005), pp. 460–489.
- [20] L. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithm*, Physica D., 60 (1992), pp. 259–268.
- [21] G. SAPIRO, *Geometric Partial Differential Equations and Image Analysis*, Cambridge University Press, January 2001.
- [22] O. SCHERZER AND C. GROETSCH, *Inverse scale space theory for inverse problems*, in Scale-Space and Morphology in Computer Vision : Third International Conference, Scale-Space 2001, M. Kerckhove, ed., vol. 2106 of Lecture Notes in Computer Science, Springer-Verlag GmbH, 2001.
- [23] D. STRONG AND T. F. CHAN, *Edge-preserving and scale-dependent properties of total variation regularization*, Inverse Problems, 19 (2003), pp. 165–187.