A NONLINEAR INVERSE SCALE SPACE METHOD FOR A CONVEX MULTIPLICATIVE NOISE MODEL

JIANING SHI[†] AND STANLEY OSHER[‡]

Abstract. We are motivated by a recently developed nonlinear inverse scale space method for image denoising [5, 6], whereby noise can be removed with minimal degradation. The additive noise model has been studied extensively, using the ROF model [23], an iterative regularization method [21], and the inverse scale space flow [5, 6]. However, the multiplicative noise model has not been studied thoroughly yet. Earlier total variation models for the multiplicative noise cannot easily be extended to the inverse scale space, due to the lack of global convexity.

In this paper, we review existing multiplicative models and present a new total variation framework for the multiplicative noise model, which is globally strictly convex. We extend this convex model to the nonlinear inverse scale space flow, and its corresponding relaxed inverse scale space flow. We demonstrate the convergence of the flow for the multiplicative noise model, as well as its regularization effect and its relation to the Bregman distance. We investigate the properties of the flow, and study the dependence on flow parameters. The numerical results show an excellent denoising effect and significant improvement over earlier multiplicative models.

Key words. inverse scale space, total variation, multiplicative noise, denoising, Bregman distance

1. Introduction. Image denoising is an important problem of interest to the mathematical community and with wide applications in fields ranging from computer vision to medical imaging. A variety of methods have been proposed over the last decades, including traditional filtering, wavelets, stochastic approaches, and partial differential equations (PDE) based variational methods. We refer the readers to [9] for a review of various methods. The additive noise model has been extensively studied, using the Rudin-Osher-Fatemi (ROF) model [23], an iterative regularization method [21], and the inverse scale space (ISS) flow [5, 6]. However, the multiplicative noise has not been studied thoroughly yet. In this paper, we obtain a new convex multiplicative noise model and extend it to the nonlinear inverse scale space.

In the additive noise model, one sought to recover the signal u which was corrupted by additive noise v, i.e., $f=u+v$. The ROF model [23] introduced total variational minimization to image processing, and defined the solution as follows:

$$u=\underset{u\in BV(\Omega)}{\arg\min}\left\{|u|_{BV}+\frac{\lambda}{2}\|f-u\|_{L^2}^2\right\} \tag{1.1}$$

for some scale parameter $\lambda>0$, where $BV(\Omega)$ denotes the space of functions with bounded variation on Ω, equipped with the BV seminorm which is formally given by $|u|_{BV}=\int_\Omega|\nabla u|$, also referred to as the total variation (TV) of u. Successful implementations of this minimization problem include the Euler-Lagrange equation by gradient descent on the original minimization problem [23], second-order cone programming [15], duality [8], and an extremely fast method based on graph cuts [11]. The choice of the positive scale constant λ in the ROF model is important, in that large λ corresponds to a small amount of noise removal while small λ can result in oversmoothing the image. The ROF model was extended to an iterative regularization method based on the Bregman distance in [21], motivated by Meyer's analysis in [19], where he defined texture, "highly oscillatory patterns in an image", as elements of the dual space of $BV(\Omega)$. In order to preserve the texture information Meyer suggested a modified variational problem using the space $(G,\|\cdot\|_*)$, the dual space for the closure of the Schwartz space in $BV(\Omega)$, equipped with the $\|\|_*$, see [19], as:

$$u=\underset{u\in BV(\Omega)}{\arg\min}\left\{|u|_{BV}+\lambda\|f-u\|_*\right\}. \tag{1.2}$$

*Research supported by NSF grants DMS-0312222, and ACI-0321917 and NIH G54 RR021813.

[†]Department of Biomedical Engineering, Columbia University, 351 Engineering Terrace, MC 8904, 530 W 120th St, New York, NY 10027, USA. Email:js2615@columbia.edu

[‡]Department of Mathematics, UCLA, 520 Portola Plaza, Los Angeles, CA 90095, USA. Email: sjo@math.ucla.edu

Meyer also used his analysis to quantify and explain the observed loss of contrast in the ROF model.

The iterative regularization model, introduced in [21], improved the quality of regularized solutions in terms of texture preservation as well as the signal-to-noise ratio (SNR). The nonlinear inverse scale space (ISS) method devised in [5, 6] formulates the iterative regularization method as a continuous time flow and provides a better temporal resolution, a better stopping criterion and is usually much faster to implement than the iterative regularization. Both methods have solutions which start from an initial condition $u_0(x)=0$, (assuming $\int f=0$), come close to the true u, then approach the noisy image f. The idea behind these two methods is that larger scales converge faster than smaller ones, where scale can be precisely defined using Meyer's $\|\|_*$. The inverse scale space method has proven to yield better denoising results than standard ROF. In fact it yields improved denoising results among PDE based methods [4, 5, 6, 9, 21, 23].

Multiplicative noise has not been as well studied. We consider the problem of seeking the true signal u from a noisy signal f, corrupted by multiplicative noise η. This arises in medical imaging, e.g., magnetic field inhomogeneity in MRI [14, 20], speckle noise in ultrasound [28], and speckle noise in synthetic aperture radar (SAR) images [2, 27]. Denoising of speckled ultrasound images has been tackled e.g. in [1] and more general noise models, which are correlated with the signal amplitude (such as multiplicative noise), have been studied in [24]. The first total variation approach to solving the multiplicative model was presented in [22], which used a constrained optimization approach with two Lagrange multipliers. The authors also considered blurry and noisy images. However, their fitting term is nonconvex, in general, which leads to difficulties in using the iterative regularization or the inverse scale space method. In [2] a multiplicative model was introduced involving BV, with a fitting term derived from a maximum a posteriori (MAP) estimation. However, the fitting team was only convex for $u\in(0,2f)$. Some interesting analysis for the convergence, existence, and uniqueness of the solution was done, though the numerical results exhibited some loss of contrast.

Motivated by the effectiveness of the inverse scale space, we derive a total variation framework that is globally convex. We further construct a new inverse scale space method for the multiplicative noise based on this convex functional.

2. Review of multiplicative models. We are motivated by the following problem in image restoration. In what follows we will always require $f>0,\ u>0$. Given a noisy image $f:\Omega\to\mathbb{R}$, where Ω is a bounded open subset of $\mathbb{R}^2$, we want to obtain a decomposition

$$f=u\eta, \tag{2.1}$$

where u is the true signal and η the noise. We would like to denoise the signal while preserving the maximum information about u. In image processing, such information is manifested in a large part by keeping the sharpness of edges, since conventional denoising methods tend to blur the image by smearing the edge information or by creating ringing artifacts at edges.

We assume that we have some prior information about the mean and variance of the multiplicative noise,

$$\begin{aligned}\frac{1}{N}\int\eta&=1,\\ \frac{1}{N}\int(\eta-1)^2&=\sigma^2,\end{aligned} \tag{2.2}$$

where $N=\int 1$. This states that the mean of the noise is equal to 1, and the variance is equal to σ^2.

The procedure introduced in [22] sought the solution of a constrained optimization problem in the space of bounded variation, implemented by the gradient projection method involving Euler-Lagrange equations and artificial time evolution as in [23]. The following optimization problem was studied,

$$u=\underset{u\in BV(\Omega)}{\arg\min}\Big\{J(u)+\lambda H(u,f)\Big\}, \tag{2.3}$$

where $J(u)=\int_\Omega |\nabla u|$, and $H(u,f)$ is a fidelity term for a known signal or image f. In the case of additive noise, $H(u,f)=\frac{1}{2}\|f-u\|^2_{L^2}$, whereas $H(u,f)$ takes a more sophisticated form for multiplicative noise. A fidelity term $H(u,f)$ was used consisting of two integrals with two Lagrange multipliers $H(u,f)=\lambda_1\int\frac{f}{u}+\lambda_2\int(\frac{f}{u}-1)^2$. The initial data was chosen to satisfy the constraints

$$\begin{aligned}\frac{1}{N}\int\frac{f}{u(0)}&=1,\\ \frac{1}{N}\int\left(\frac{f}{u(0)}-1\right)^2&=\sigma^2,\end{aligned}\tag{2.4}$$

where $N=\int_\Omega 1$.

A gradient projection method was introduced to make sure the two constraints are always satisfied during the evolution, which reduces to requiring

$$\frac{\partial}{\partial t}\int\frac{f}{u}=-\int\frac{f}{u^2}u_t=0,\tag{2.5}$$

and

$$\frac{\partial}{\partial t}\int\left(\frac{f}{u}-1\right)^2=0=\frac{\partial}{\partial t}\int\left(\frac{f}{u}\right)^2=-2\int\frac{f^2}{u^3}u_t.\tag{2.6}$$

Thus we evolve the following Euler-Lagrange equation:

$$u_t=\nabla\cdot\frac{\nabla u}{|\nabla u|}+\lambda_1\frac{f}{u^2}+\lambda_2\frac{f^2}{u^3},\tag{2.7}$$

where the values of λ_1 and λ_2 can be found dynamically using the gradient projection method. However, for this to be a convex model we need $\lambda_1,\lambda_2\geq 0$, which is not necessarily the case as t evolves. Moreover, if we fix $\lambda_1,\lambda_2>0$ then the corresponding minimization problem will lead us to a sequence of constant functions u approaching $+\infty$.

Multiplicative noise, in various imaging modalities, is often not necessarily Gaussian noise. In the speckle noise model, such as synthetic aperture radar (SAR) imagery, the noise is treated as gamma noise with mean equal to one. The distribution of the noise η takes the form of $g(\eta)$,

$$g(\eta)=\frac{L^L}{\Gamma(L)}\eta^{L-1}exp(-L\eta)\mathbf{1}_{\{\eta>=0\}}.\tag{2.8}$$

Based on this, Aubert and Aujol [2] formulated the following minimization problem:

$$u=\arg\min_{u\in S(\Omega)}\left\{J(u)+\lambda\int\left(\log u+\frac{f}{u}\right)\right\},\tag{2.9}$$

where $J(u)=\int_\Omega|\nabla u|$, and new fitting function $H(u,f)=\int(\log u+\frac{f}{u})$ is strictly convex for $u\in(0,2f)$. The derivation of this functional is based on maximum a posteriori (MAP) on $p(u|f)$, assuming the noise η follows a gamma law with mean one, and $p(u)$ follows a Gibbs prior. Some interesting analysis of this model was done in [2].

For denoising, the inverse scale space method has so far been only applied to the additive noise model. In this paper we will present a new total variation functional for the multiplicative noise model, which is globally convex. We will also extend the total variation minimization approach to inverse scale space for our new model.

3. Inverse scale space. In this section, we review some key results for inverse scale space. An iterative algorithm based on total variation was introduced in [21], which was preceeded by several interesting and related pieces of work [16, 25, 26]. The inverse scale space is based on the iterative

algorithm, by taking the time step to the limit of 0, formally introduced in [5, 6] on dealing with additive noise and deconvolution. We review some key results for inverse scale space in this section.

Following the treatment of Burger et al [5, 6], the total variation based minimization Eqn. (2.3) can be extended into an iterative algorithm. Introduced in [21], the authors propose to solve a sequence of u_k,

$$u_k = \arg\min_{u \in BV(\Omega)} \left\{ D(u, u_{k-1}) + \lambda H(u, f) \right\}, \tag{3.1}$$

where u_k is the primal variable, and $D(\cdot,\cdot)$ is the Bregman distance related to a differentiable functional $J : \mathbb{R}^n \to \mathbb{R}$, formally defined by

$$D_J^p(u, v) = J(u) - J(v) - \langle u - v, p \rangle, \quad p \in \partial J(v). \tag{3.2}$$

The initial condition u_0 is a constant satisfying $\int \partial_u H(u_0, f) = 0$ and $p_0 = 0$. Here $\partial_u H(u, f)$ is the subgradient of $H(u, f)$ with respect to u, and $p \in \partial J(u)$ is an element in the subgradient of J at u. We will also use the notation p_k, which is an element of the subgradient of J at u_k.

For a continuously differentiable functional there exists a unique element p in the subdifferential and consequently a unique Bregman distance. For a nonsmooth and not strictly convex functional such as total variation, one can still obtain convergence of the reconstruction, as long as the functional has suitable lower-semicontinuity and compactness properties in the weak-* topology of $BV(\Omega)$ (and also in $L^2(\Omega)$ by compact embedding).

One can rewrite the above equation and arrive at

$$u_k = \arg\min_{u \in BV(\Omega)} \left\{ J(u) - \langle u, p_{k-1} \rangle + \lambda H(u, f) \right\}. \tag{3.3}$$

The Euler-Lagrange equation for Eqn. (3.3) is

$$p_k - p_{k-1} + \lambda \partial_u H(u_k, f) = 0.$$

Such an iterative regularization is taken to the limit by letting $\lambda = \Delta t$, $u_k = u(k\Delta t)$, then dividing by Δt and letting $\Delta t \to 0$. This leads to a continuous flow in the inverse scale space,

$$\partial_t p = -\partial_u H(u, f), \qquad p \in \partial J(u) \tag{3.4}$$

with the initial conditions $p_0 = 0$, $u_0 = c_0$, where c_0 is a constant satisfying $\int \partial_u H(c_0, f) = 0$. This was introduced in [21] and analyzed in [4].

This elegant formulation of the inverse scale space flow is not straightforward to compute in two or more dimensions, since the relation between p and u is complicated for TV and other nonlinear cases. However, in 1D one can solve this directly by regularizing the total variation term $J(u)$ by the following

$$J_\epsilon(u) = \int_\Omega \sqrt{|\nabla u|^2 + \epsilon}, \tag{3.5}$$

where $\epsilon > 0$ is a small constant.

A relaxed form of the inverse scale space flow was introduced, which involves the following coupled equations:

$$\begin{aligned} \partial_t u &= -p + \lambda(-\partial_u H(u, f) + v), \\ \partial_t v &= -\alpha \partial_u H(u, f), \end{aligned} \tag{3.6}$$

with $p \in \partial J(u)$ and initial conditions $u_0 = c_0$, $v_0 = 0$.

Some examples of the inverse scale space flows using different choices for $J(u)$ and $H(u, f)$ are as follows:

- *Linear model:* $J(u)=\frac{1}{2}\|\nabla u\|_2^2$, $H(u,f)=\frac{1}{2}\|f-u\|_{L^2}^2$.

$$\begin{aligned} u_t &= \Delta u+\lambda(f-u+v),\\ v_t &= \alpha(f-u). \end{aligned} \tag{3.7}$$

- *ROF model:* $J(u)=\int_\Omega|\nabla u|$, $H(u,f)=\frac{1}{2}\|f-u\|_{L^2}^2$.

$$\begin{aligned} u_t &= \nabla\cdot\tfrac{\nabla u}{|\nabla u|}+\lambda(f-u+v),\\ v_t &= \alpha(f-u). \end{aligned} \tag{3.8}$$

- *$TV-L^1$ model:* $J(u)=\int_\Omega|\nabla u|$, $H(u,f)=\|f-u\|_{L^1}$.

$$\begin{aligned} u_t &= \nabla\cdot\tfrac{\nabla u}{|\nabla u|}+\lambda(\mathrm{sign}(f-u)+v),\\ v_t &= \alpha\,\mathrm{sign}(f-u). \end{aligned} \tag{3.9}$$

  Note that $H(u,f)$ is not strictly convex or smooth here and sign is the notation for an element in the subgradient of $H(u,f)$.
- *Deconvolution by ROF:* $J(u)=\int_\Omega|\nabla u|$, $H(u,f)=\frac{1}{2}\|f-K*u\|_{L^2}^2$.

$$\begin{aligned} u_t &= \nabla\cdot\tfrac{\nabla u}{|\nabla u|}+\lambda\left(\hat{K}*(f-K*u)+v\right),\\ v_t &= \alpha\hat{K}*(f-K*u). \end{aligned} \tag{3.10}$$

Here K is a real blurring kernel, $\hat{K}(x,y)=K(-x,-y)$ and $*$ denotes convolution [10].

So far the Eqn. (3.8) - (3.10) have not been studied analytically. However a nice *a priori* inequality was obtained in [18].

4. New methodology. The difficulty of extending previous total variation models for multiplicative noise to the inverse scale space method is due to the lack of global convexity. We will overcome this difficulty via the following approaches.

4.1. Logarithm transform. The simplest idea is to take the log of both sides of Eqn. (2.1),

$$\log f=\log u+\log\eta,$$

which essentially converts the multiplicative problem into an additive one. The additive noise problem has already been successfully treated using the ROF method [23], the Bregman iterative method based on the ROF model [21], as well as with inverse scale space and its relaxed version [5, 6].

Now we consider seeking $w=\log u$ based on the noisy observation $\log f$. We obtain the following total variation minimization, using the BV norm of w. We no longer use the BV norm of u, and this gives us a convex optimization problem:

$$w=\arg\min_{w\in BV(\Omega)}\left\{J(w)+\frac{\lambda}{2}\|w-\log f\|_{L^2}^2\right\}, \tag{4.1}$$

where $J(w)=|w|_{BV}$, and the fidelity term $H(w,f)=\frac{1}{2}\|w-\log f\|_{L^2}^2$. Note that a nice property about $H(w,f)$ is its strict convexity. It follows that the corresponding relaxed inverse scale space flow can be expressed as,

$$\begin{aligned} w_t &= \nabla\cdot\left(\tfrac{\nabla w}{|\nabla w|}\right)+\lambda(\log f-w+v),\\ v_t &= \alpha(\log f-w), \end{aligned} \tag{4.2}$$

with $v(0)=0$, $w(0)=c_0$, for $c_0=\int\log f/\int 1$.

This method gives excellent results, as we shall see in Section 6. However we will also use a more general model, described as follows, which gives equally excellent results.

4.2. Convex total variation minimization for a general multiplicative noise model. We consider a rather general total variation formulation for the multiplicative noise:

$$u = \arg\min_{u \in BV(\Omega)} \left\{ J(u) + \lambda \int \left(a\frac{f}{u} + \frac{b}{2}(\frac{f}{u})^2 + c\log u \right) \right\}, \tag{4.3}$$

where $J(u)$ is the total variation of u and a, b, c are nonnegative constants. This formulation seems to include all previous models [2, 12, 22].

- When $c = 0$, this resembles the (constrained) model of [22]. It is generally different because a and b might be negative in [22].
- When $b = 0$ and $a = c$, this reduces to the model in [2].
- When $a = 0$ and $b = c$, this reduces to the model in [12].

We can solve this total variation problem by gradient descent, resulting in the following unsteady Euler-Lagrange equation,

$$u_t = \nabla \cdot \frac{\nabla u}{|\nabla u|} + \lambda\big(a\frac{f}{u^2} + b\frac{f^2}{u^3} - c\frac{1}{u}\big). \tag{4.4}$$

We require the following conditions:

- The fidelity term $H(u, f)$ has a minimum at $u = f$, which implies that $c = a + b$.
- The fidelity term $H(u, f)$ is convex for u near f, which means $a + 2b > 0$. However, this term will not be convex for $u >> f$.

Due to the lack of global convexity, we abandon this approach and proceed by letting $w = \log(u)$. We also replace $J(u) = J(\exp(w))$ by $J(w)$. By using such a partial transformation of variables, the fidelity term $H(w, f)$ is rendered convex, while the total variation term is left unaltered. Such replacement is reasonable since the mapping $w \to e^w$ is monotically increasing. Convexity of the objective function allows us to take advantage of the excellent regularization ability of the inverse scale space. It is noteworthy that we use the BV norm of w, instead of u. The regularization of the BV norm therefore takes place in the logarithmic sense for the image, penalizing on $\log(u)$ instead of u. In this logarithmic case, the amount of regularization depends on the image intensity, resulting in strong smoothing for image values near 0, and less smoothing for large intensities. This is somewhat visible in Figures 6.1 and 6.2.

One can obtain the following total variation based minimization,

$$w = \arg\min_{w \in BV(\Omega)} \left\{ J(w) + \lambda \int \left(af\exp(-w) + \frac{b}{2}f^2\exp(-2w) + (a+b)w \right) \right\}. \tag{4.5}$$

We aim to solve this minimization problem for w and recover the signal by $u = \exp(w)$.

The fidelity term $H(w, f) = \int (af\exp(-w) + \frac{b}{2}f^2\exp(-2w) + (a+b)w)$ is globally strictly convex. Using gradient descent and the Euler-Lagrange equation for this total variation based problem we obtain:

$$w_t = \nabla \cdot \frac{\nabla w}{|\nabla w|} + \lambda\big(af\exp(-w) + bf^2\exp(-2w) - (a+b)\big). \tag{4.6}$$

We choose the initial condition $w(0)$ defined below so that $\int \partial_w H(c_0, f) = 0$.

In order to perform iterative regularization on this model, we define the following sequence $\{w_k\}$ together with $\{p_k\}$.

- Set $w_0 = c_0$, $p_0 = 0$,
- Given w_{k-1} and $p_{k-1} \in \partial J(w_{k-1})$, $k \geq 1$, perform the following two steps:
  (i) compute $w_k = \arg\min_w Q_k(w)$ with

$$Q_k : w \longmapsto J(w) - J(w_{k-1}) - \langle p_{k-1}, w - w_{k-1} \rangle + \lambda H(w, f),$$

  where $H(w, f) = \int \left(af\exp(-w) + \frac{b}{2}f^2\exp(-2w) + (a+b)w \right)$ and $\langle \cdot, \cdot \rangle$ denotes the usual duality product.

(ii) update the dual variable

$$p_k = p_{k-1} + \lambda \partial_w H(w,f) \in \partial J(w_k),$$

where $\partial_w H(w,f) = -af\exp(-w) - bf^2\exp(-2w) + (a+b)$.

- Increase k by 1 and continue.

Such an iterative procedure can be further extended to the nonlinear inverse scale space (ISS) flow, see Eqn. (3.4). We obtain the following ISS model:

$$\begin{aligned} \frac{\partial p}{\partial t} &= af\exp(-w) + bf^2\exp(-2w) - (a+b), \\ p &\in \partial J(w), \end{aligned} \tag{4.7}$$

with $w(0) = c_0, p(0) = 0$.

For the initialization of the ISS, $w(0) = c_0$, we use the fact that

$$\int_\Omega \partial_t p = 0, \tag{4.8}$$

which means that for all time, $\int_\Omega \partial_w H(w(t),f)$ is time invariant. This expression vanishes as $t \to 0$, leading to

$$\int \big(-af\exp(-c_0) - bf^2\exp(-2c_0) + (a+b) \big) = 0.$$

This gives us the initial condition for the inverse scale space flow:

$$c_0 = \begin{cases} -\log\left(\dfrac{\int 1}{\int f}\right) & \text{if } b = 0, \\ -\log\left(\dfrac{-a\int f + \sqrt{(a\int f)^2 + 4b\int f^2 \int a+b}}{2b\int f^2}\right) & \text{if } b \neq 0. \end{cases} \tag{4.9}$$

We can further obtain the relaxed inverse scale space flow (RISS) for the multiplicative noise model:

$$\begin{aligned} \frac{\partial w}{\partial t} &= -p(w) + \lambda\big(af\exp(-w) + bf^2\exp(-2w) - (a+b) + v\big), \\ \frac{\partial v}{\partial t} &= \alpha\big(af\exp(-w) + bf^2\exp(-2w) - (a+b)\big), \\ p(w) &= -\nabla \cdot \frac{\nabla w}{|\nabla w|}, \end{aligned} \tag{4.10}$$

with initial conditions $w(0) = c_0$, and $v(0) = 0$.

Note that our model Eqn. (4.3) can easily be generalized to include any functional

$$H(u,f) = \int G\left(\frac{f}{u}\right) + c\log u \tag{4.11}$$

that satisfies the following conditions:

- $H(u,f)$ has a minimum and is convex at $u = f$, which means $c = G'(1)$ and $G''(1) + G'(1) > 0$.
- $H(e^w, f)$ is convex in w for all w.

Examples of such a functional include:

$$G(z) = \sum_{j=1}^{M} a_j \frac{z^j}{j}, \text{ with } \sum_{j=1}^{M} a_j = c$$

and for which $\sum_{j=1}^{M} a_j j x^j > 0$ for all $x > 0$.

5. Analysis of our convex model.

5.1. Review. In [21] the iterative regularization was analyzed in detail for the deblurring problem, wherein $J(u)=|u|_{BV}$ and $H(u,f)=\frac{1}{2}\|f-Au\|_{L^2}^2$, with A being the identity or a compact operator. The following results were obtained:
(a) well-definedness of the iterates: for each $k\in\mathbb{N}$, there exists a minimizer u_k of Q_k and there exists a subgradient $p_k\in\partial J(u_k)$ and $q_k=\partial_u H(u_k,f)=\lambda A^*(Au_k-f)$ such that $p_k+q_k=p_{k-1}$. If A has no nullspace, then the minimizer u_k is unique.
(b) the sequence obtained from the iterates $H(u_k,f)$ is non-increasing, and moreover

$$H(u_k,f)\leq H(u_k)+D^{p_{k-1}}(u_k,u_{k-1})\leq H(u_{k-1},f). \tag{5.1}$$

(c) if there exists a minimizer of $H(\cdot,f),\tilde{u}\in BV$, then

$$H(u_k,f)\leq H(\tilde{u},f)+\frac{J(\tilde{u})}{k} \tag{5.2}$$

and $u_k\rightarrow\tilde{u}$ in the weak-* topology in BV with

$$\|f-Au_k\|_{L^2}\leq\sqrt{\frac{J(\tilde{u})}{\lambda k}}, \tag{5.3}$$

(d) in the noisy case, suppose g is the true noise free image and $\tilde{u}$ is a minimizer of $H(\cdot,g)$ with $H(\tilde{u},g)=0$ and $H(\tilde{u},f)\leq\delta^2$ (δ is the noise level), then the key result is: as long as $H(u_k,f)\geq\delta^2$ (the residual lies above the noise level), the Bregman distance between u_k and $\tilde{u}$ is decreasing, more precisely,

$$D(\tilde{u},u_k)\leq D(\tilde{u},u_{k-1}). \tag{5.4}$$

Result (d) yields a natural stopping criterion, the so-called *generalized discrepancy principle*, which consisted in stopping the iteration at the index $k_*=k_*(\delta,f)$ given by

$$k_*=\max\{k\in\mathbb{N}:H(u_k,f)>\delta^2\}. \tag{5.5}$$

Note that the stopping index k_* is well-defined due to the monotone decrease of $H(u_k,f)$.

This gives insight into why Bregman iteration improves the contrast of recovered signal: the Bregman distance between $\tilde{u}$ and u_k gets smaller as the iterates approach the true noise free image, and then gets bigger as we get too close to the noisy image.

A variant of this procedure was proven to work in [17] with non-quadratic fidelity terms of the form $H(u,f)=h(u-f)$, where h is non-negative, convex and positively homogeneous, continuous with respect to weak convergence in BV, and $h(c)\neq 0$ for constant functions $c\neq 0$. It differs from the minimization in [5, 21], because in [17] the authors replaced λ by $\lambda 2^{k-1}$, resulting in the following total variation minimization

$$u_k=\underset{u\in BV(\Omega)}{\arg\min}\left\{D(u,u_{k-1})+\lambda 2^{k-1}H(u,f)\right\}. \tag{5.6}$$

Results analogous to those described above [21] were shown to be valid for the non-quadratic convex fidelity term.

This algorithm was inspired by Scherzer and Groetsch [16, 25], and Tadmor, Nezzar and Vese [26]. The authors proposed an interesting iterative algorithm,

$$u_k=\underset{u\in BV(\Omega)}{\arg\min}\left[\frac{\lambda}{2}2^k\int(f-u)^2+J(u-u_{k-1})\right]. \tag{5.7}$$

Note that Eqn. (5.7) is not the Bregman iteration defined in [3]. In the quadratic case, e.g.

$$J(u)=\frac{1}{2}\int_{\Omega}|\nabla u|^2, \tag{5.8}$$

the two methods are the same (if the 2^k are removed in Eqn. (5.7)); otherwise they are different.

Iterative regularization was also analyzed for wavelet based denoising in [29], where the iterative procedure was applied to the following total variation based minimization problem:

$$\tilde{u}=\arg\min_{\tilde{u}\in B_1^{1,1}(\Omega)}\left(\sum|\tilde{u}_i|+\frac{\lambda}{2}\sum(\tilde{f}_i-\tilde{u}_i)\right). \tag{5.9}$$

Note that $\tilde{f}_i$ represents the wavelet coefficients of the noisy image f, formally defined as $\tilde{f}=\{\tilde{f}_i\}=\{\langle f,\phi_i\rangle\}$, and one can reconstruct the image u from the wavelet coefficients $\tilde{u}_i$. $B_1^{1,1}(\Omega)$ stands for the Besov space, which is defined and discussed in e.g. [19].

Returning to the quadratic functionals mentioned above, Burger, Resmerita and He [7] obtained extremely interesting error estimates in the case where the source condition is satisfied, i.e., $J(\tilde{u})<\infty$.

For the exact data, let $\tilde{u}\in BV$ be a solution of $Au=f$ and assume there exists $\xi\in\partial J(\tilde{u})$ such that $\xi=A^*q$ for $q\in L^2(\Omega)$, which is a smoothness condition on the curvature of the level sets of $\tilde{u}$. Then one can obtain

$$D(\tilde{u},u_k)\le\frac{\|q\|^2}{2\lambda k}, \tag{5.10}$$

which implies that u_k approaches $\tilde{u}$ in the sense of Bregman distance.

Moreover, suppose the noisy data satisfy

$$\|f-f^{\delta}\|_{L^2}\le\delta, \tag{5.11}$$

then an *a priori* stopping rule $k_s(\delta)\sim\frac{1}{\delta}$ yields semi-convergence of the regularization method. Interested readers are referred to [7] for more details.

Analogous results were obtained for the inverse scale space flows in [5, 6, 29]. We will restrict ourselves to formal estimates of ISS for our model. For the relaxed inverse scale space, the only theoretical results so far were done (for A the identity operator) by Lie and Nordbotten [18]. They showed formally

$$\begin{aligned}&\frac{\partial}{\partial t}\frac{1}{2}\left[\frac{1}{\lambda}\|u-f\|_{L^2}^2+\frac{1}{\alpha}\|v-\frac{p(f)}{\lambda}\|_{L^2}^2\right]\\&\quad=-\frac{1}{\lambda}[D(f,u)+D(u,f)]-\|u-f\|_{L^2}^2,\end{aligned} \tag{5.12}$$

which yielded a nice decay of energy estimate, assuming $p(f)=\partial J(f)$ is in L^2.

5.2. Formal analysis of inverse scale space flow. In this section we assume existence of solution and are just concerned with *a priori* inequalities and qualitative behavior of our models.

Consider the inverse scale space flow

$$\frac{dp}{dt}=-\partial_w H(w,f), \tag{5.13}$$

where

$$\begin{aligned}H(w,f)=&\int\left(af\exp(-w)+\frac{b}{2}f^2\exp(-2w)+(a+b)w\right)\\&-\int(a+\frac{b}{2}+(a+b)\log f).\end{aligned} \tag{5.14}$$

We have subtracted a w independent term above so that the convex (in w) functional $H(w,f)$ takes on its minimum at $w=\log f$, with 0 as its minimum value.

We further assume that $J(w)$ is strictly convex, which is not true for BV. However in practice one approximates it using the functional $J_\epsilon(w)=\int_\Omega \sqrt{|\nabla w|^2+\epsilon}$, for $\epsilon>0$. This means that $p(w)$ has an inverse $w(p)$, and in fact $w(p)$ is the subgradient of the dual functional J^* defined as

$$J^*(p)=\sup_w\{\langle w,p\rangle - J(w)\}, \tag{5.15}$$

where J^* is strictly convex. Due to the lower semi-continuity and convexity (see Chap. I, Cor. 5.2 in [13]), we have

$$p\in\partial J(w)\Leftrightarrow w\in\partial J^*(p).$$

We assume $J^*\in C^2$, strictly convex and has an expression,

$$J^*(p)-J^*(q)=\langle p-q,\partial J^*(q)\rangle+\frac{1}{2}\langle p-q,\frac{\partial^2 J^*}{\partial p^2}(\tilde{q})(p-q)\rangle,$$

for $\tilde{q}$ which approaches q as p approaches q.

Hence we obtain the following

$$\begin{aligned}\frac{d}{dt}H(w,f)&=\langle\partial_w H(w,f),\frac{dw}{dt}\rangle\\ &=-\langle\partial_w H(w,f),\frac{\partial^2 J^*}{\partial p^2}(p)\partial_w H(w,f)\rangle\\ &\le -a\|\partial_w H(w,f)\|_{L^2}^2\end{aligned} \tag{5.16}$$

Here we assume $\partial J^*(p)$ is Fréchet differentiable with derivative $\frac{\partial^2 J^*}{\partial p^2}$, which is true for $J_\epsilon(w)$, see e.g. [5, 6]. This leads us to an estimate

$$\|w(t)-\log f\|_{L^2}\le c_1 e^{-c_2 t}\|w(0)-\log f\|_{L^2}, \tag{5.17}$$

where $c_1,c_2>0$ depend on a, as well as the upper and lower bounds of the function

$$B(w,f)=af\exp(-w)+2bf^2\exp(-2w)$$

for w near $\log f$. Such an estimate implies that we have L^2 convergence of $w(t)$ to $\log f$ as $t\to\infty$.

To obtain Eqn. (5.17) from Eqn. (5.16), we note that

$$H(w,f)=\int\frac{(w-\log f)^2}{2}B(\tilde{w},t),$$

and

$$\|\partial_w H(w,f)\|_{L^2}^2\le\int(w-\log f)^2B^2(\hat{w},f),$$

for $\tilde{w}$, $\hat{w}$ between w and $\log f$. Then (5.17) follows from Grönwall's inequality.

Furthermore, consider the exact data (noise free case), where $J(\log f)<\infty$, one can obtain

$$\begin{aligned}\frac{d}{dt}D(\log f,w)&=-\frac{d}{dt}J(w)-\langle\log f-w,\frac{dp}{dt}\rangle\\ &=\langle\log f-w,\partial_w H(w,f)\rangle\\ &\le -c_3\|\log f-w\|_{L^2}^2\end{aligned} \tag{5.18}$$

for $c_3>0$, c_3 is a lower bound on the function $B(w,f)$, defined above near $w=\log f$. These results concern convergence of $w(t)$ to the image $\log f$, when $J(\log f)<\infty$.

For the noisy data case, where $J(\log f)=\infty$, we proceed as follows. Define g as the true noise free image, and f the noisy image. Let $\log g$ be close to $\log f$ in the L^2 norm with $J(\log g)<\infty$, then one can obtain

$$\begin{aligned}\frac{d}{dt}D(\log g,w)&=\langle \log g-w,\partial_w H(w,f)\rangle\\&=\langle \log g-\log f,\partial_w H(w,f)\rangle+\langle \log f-w,\partial_w H(w,f)\rangle\\&\le -c_3\|\log f-w\|_{L^2}^2\\&+\frac{c_4}{2}\|\log g-\log f\|_{L^2}^2+\frac{1}{2c_4}c_5^2\|\log f-w\|_{L^2}^2.\end{aligned}\tag{5.19}$$

Here $c_4>0$ is arbitrary and c_5 is an upper bound on the function $B(w,f)$ near $w=\log f$. We take $c_4=c_5^2/c_3$ and arrive at:

$$\frac{d}{dt}D(\log g,w)=-\frac{c_3}{2}\|\log f-w\|_{L^2}^2+\frac{c_5^2}{2c_3}\|\log g-\log f\|_{L^2}^2<0\tag{5.20}$$

as long as

$$\|\log f-w(t)\|_{L^2}>\frac{c_5}{c_3}\|\log f-\log g\|_{L^2},\tag{5.21}$$

where c_3 is the lower bound of $B(w,f)$ near the neighborhood of $w=\log f$, and c_5 the upper bound.

Therefore the Bregman distance between $\log g$, where g is the "true noise free image", and $w(t)$ is decreasing until $w(t)$ gets closer in L^2 to $\log f$, to the point where the L_2 distance between $w(t)$ and $\log f$ is less than $\frac{\max B(w)}{\min B(w)}\|\log f-\log g\|_{L^2}$ (w near $\log f$).

In view of our estimates, "w near $\log f$" is the neighborhood

$$\|w-\log f\|_{L^2}\le c_1\|w(0)-\log f\|_{L^2}.\tag{5.22}$$

This means that we get monotonically closer, in the sense of Bregman distance, to the denoised solution until we get too close to the noisy image, in the L^2 norm.

5.3. Linear Analysis for relaxed inverse scale space. Our relaxed inverse scale space model [5, 6] can be written in a general form as:

$$\frac{\partial w}{\partial t}=-p(w)+\lambda\big(v-\partial_w H(w,f)\big),\tag{5.23}$$

$$\frac{\partial v}{\partial t}=-\alpha\partial_w H(w,f).\tag{5.24}$$

As in [5, 6] we approximate $p(w)$ by $p(w)=-\Delta w$, differentiate Eqn.(5.23) with respect to t, and substitute for $\partial_t v$ using Eqn.(5.24). This leads to

$$\partial_{tt}^2 w+(-\Delta+\lambda\partial_{ww}^2 H(w,f))\partial_t w+\lambda\alpha\partial_w H(w,f)=0.\tag{5.25}$$

For this analysis, we freeze w as a constant w_0, f as a constant f_0, take the Fourier transform, so $-\Delta$ is replaced by $|\xi|^2$. We therefore obtain the characteristic equation:

$$r^2+(\lambda\partial_w^2 H(w_0,f_0)+|\xi|^2)r+\lambda\alpha\partial_w H(w_0,f_0)=0,\tag{5.26}$$

with solutions

$$r_\pm=\frac{-(\lambda\alpha_w^2 H(w_0,f_0)+|\xi|^2)\pm\sqrt{(\lambda\partial_w^2 H(w_0 f_0)+|\xi|^2)^2-4\alpha\lambda\partial_w H(w_0,f_0)}}{2}.\tag{5.27}$$

Both roots will be real for all frequencies if

$$\alpha\partial_w H(w_0,f_0) < \frac{\lambda}{4}(\partial_w^2 H(w_0,f_0))^2. \tag{5.28}$$

Let $f_0\exp(-w_0)=C$. Thus, our restriction of the parameters to ensure monotone evolution of $w(t)$ towards $\log f$ is:

$$\alpha(C-1)(Cb+a+b) < \frac{\lambda}{4}(aC+2bC^2)^2. \tag{5.29}$$

In all of our calculations the convergence was monotone.

6. Numerical Results. We carry out numerical experiments for both RISS models: (a) the relaxed inverse scale space (RISS) for our global convex model, Eqn. (4.10), with the initial condition Eqn. (4.9), (b) taking the logarithm transform and use additive RISS, Eqn. (4.2). RISS is tested on both 1D and 2D cases, for different types of images. We demonstrate the convergence of the algorithm, as well as the regularization effect in terms of denoising. We study how the relation between λ and α affects the convergence rate of the flow. We also show the comparison of both RISS algorithms with the Rudin-Lions-Osher algorithm [23] numerically.

6.1. Relaxed Inverse Scale Space 1D. Figure 6.1 shows our test example in 1D, where the signal u is corrupted by multiplicative noise η. The mean of the noise is equal to 1, while the standard deviation is σ. In this example, we use $\sigma=0.1$. The goal is to recover the signal based on the corrupted signal f.

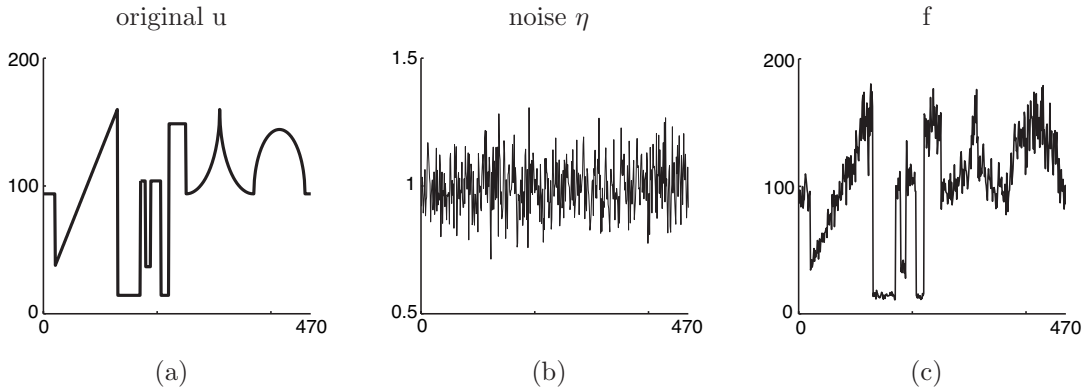


FIG. 6.1. *1D signal u is corrupted by multiplicative noise η, resulting in noisy signal f, (a) original signal u, (b) multiplicative noise η, with mean* 1*, and standard deviation* $\sigma=0.1$*. (c) noisy signal f.*

Figure 6.2 shows the numerical result of the relaxed inverse scale space (RISS), Eqn. (4.10) and Eqn. (4.9). These results show the restoration of the signal at different stopping indices: (a) the standard deviation of the restored noise component, defined as $\tilde{\sigma}=\frac{1}{\sqrt{N}}\|\frac{f}{\exp(w(t))}-1\|_{L^2}$, matches the prior information σ; (b) the L^2 norm between the restored signal and the true noise-free signal, $\|\exp(w(t))-g\|_{L^2}$, reaches its minimum; (c) the Bregman distance between the restored signal and the noise-free signal, defined as $D(\log(g),w(t))$, reaches its minimum. From now on we stop at the index where the variance of the recovered noise matches that of our prior knowledge.

The stopping criterion is related to the Bregman distance between the recovered signal and the ground truth. In the additive RISS model, when the variance of the recovered noise reaches the true value, this is also the point where the Bregman distance between the ground truth and the recovered signal reaches its minimum. In the multiplicative case, such relation is not as clean, as is predicted theoretically earlier in the paper, see Eqn. (5.21).

6.2. Parameter Study. We investigate how different pairs of values for a and b affect the algorithm. Figure 6.2 shows the result for $a=1$, $b=0$, Figure 6.3 for $a=0$, $b=1$, and Figure 6.4 for $a=1$, $b=1$. We demonstrate success of our algorithm for different pairs of a and b, providing equally good results in restoration and regularization of the signal, as well as contrast preservation.

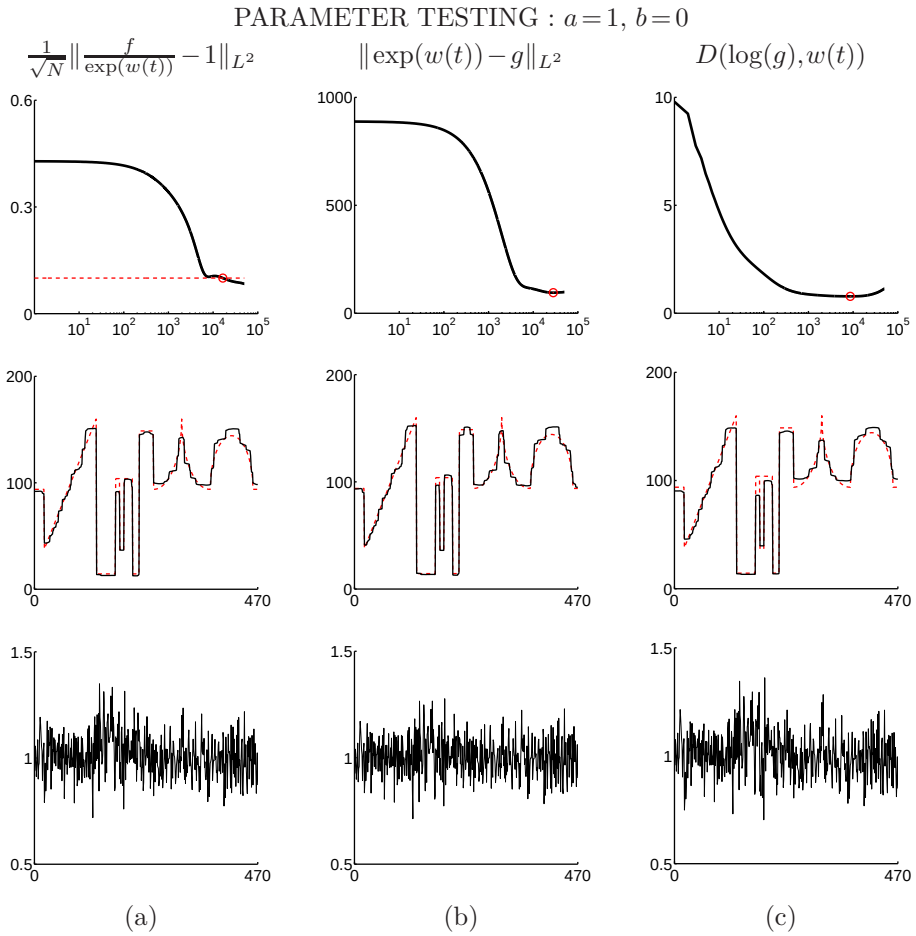


FIG. 6.2. *Numerical results for the 1D case using the relaxed inverse scale space flow. We observe the flow at different iterations: (a) recovered u using our stopping criterion based on our knowledge about the standard deviation for the multiplicative noise, (b) recovered u at the minimum of L^2 norm between u and ground truth g, (c) recovered u at the minimum of the Bregman distance $D(log(g),w(t))$. In all cases, $a=1$, $b=0$, $\lambda=1$, $\alpha=0.125$, $dt=0.0005$.*

Convergence of the algorithm is shown in Figure 6.5 for different values of a and b. We define the following energy function

$$e(t)=\frac{1}{2\lambda}H(w(t),f)+\frac{1}{2\alpha}\|v(t)-\frac{q}{\lambda}\|_{L^2}^2, \tag{6.1}$$

for the multiplicative relaxed inverse scale space flow, Eqn.(4.10). Our simulations show that convergence is satisfied in all three cases.

Convergence of the algorithm also depends on the ratio between λ and α, as predicted by our linear analysis of the relaxed inverse scale space flow, Eqns. (5.28)-(5.29). Numerically we show in Figure 6.6 the convergence results for different values of α. We show below that the convergence of the algorithm is achieved in all our simulations.

6.3. Comparison with other methods. We show numerical results for our multiplicative RISS algorithm, as well as the following algorithms: (a) Logarithm transform into the additive

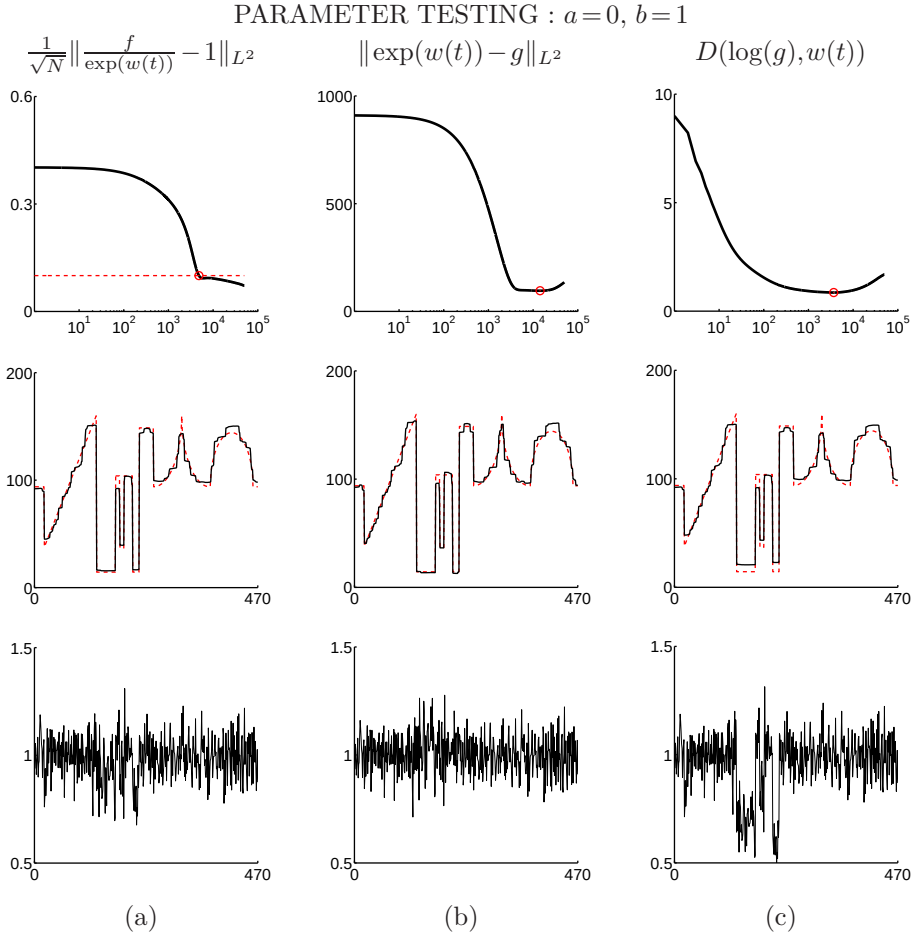


FIG. 6.3. *Numerical result for the 1D case using the relaxed inverse scale space flow. We observe the flow at different iterations: (a) recovered u using our stopping criterion based on our knowledge about the standard deviation for the multiplicative noise, (b) recovered u at the minimum of L^2 norm between u and ground truth g, (c) recovered u at the minimum of the Bregman distance $D(log(g),w(t))$. In all cases, $a=0$, $b=1$, $\lambda=1$, $\alpha=0.125$, $dt=0.0005$.*

problem, and then apply the additive RISS to solve the problem; (b) Rudin-Lions-Osher algorithm, which uses gradient projection to dynamically solve the total variation problem; (c) Gradient descent on our convex multiplicative model, see Eqn. (4.6); (d) Aubert-Aujol algorithm using a nonconvex fidelity term. Numerical experiments have been carried out to test two types of noise, Gaussian noise and speckle noise. We compares our multiplicative RISS algorithm with the three algorithms (a)-(c) mentioned above. The two best here and throughout our numerical experiments are the multiplicative RISS and the log transform RISS. In order to quantify the denoising effect, we define the Signal-to-noise Ratio SNR in dB as:

$$SNR=20\log_{10}\sqrt{\frac{\|g\|_{L^2}}{\|u-g\|_{L^2}}}. \tag{6.2}$$

Figure 6.7 shows a numerical comparison of our algorithm for the 1D case, where the noise takes the form of Gaussian noise, with mean of 1 and standard deviation of 0.1. Our mulplicative RISS (a) and logarithm transform additive RISS (b) show great results in signal restoration. The inverse scale space method essentially constructs a "path" from an initial condition to the noisy image, where we stop according to a fit-to-data prior information. This approach improves the solution

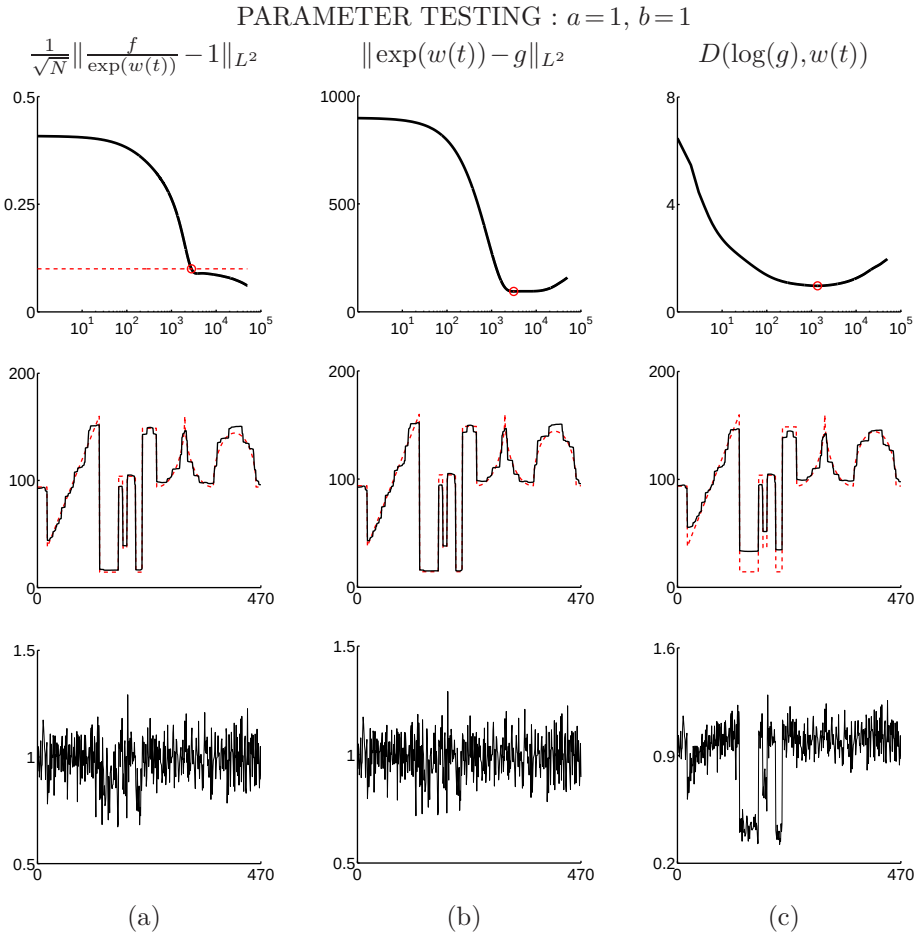


FIG. 6.4. *Numerical result for the 1D case using the relaxed inverse scale space flow. We observe the flow at different iterations: (a) recovered u using our stopping criterion based on our knowledge about the standard deviation for the multiplicative noise, (b) recovered u at the minimum of L^2 norm between u and ground truth g, (c) recovered u at the minimum of the Bregman distance $D(log(g),w(t))$. In all cases, $a=1$, $b=1$, $\lambda=1$, $\alpha=0.125$, $dt=0.0005$.*

quality compared to the traditional total variation based models.

We also study the RISS in 2D, and carry out numerical experiments on the cameraman and Barbara images. Figure 6.8 shows the numerical result of our multiplicative RISS algorithm, which shows excellent denoising effect as well as contrast restoration. Clearly our inverse scale space algorithm, either in the general multiplicative algorithm or directly taking advantage of the logarithm transform, outperform other methods in literature. Figure 6.9 shows the numerical results for Barbara, where the texture information is well preserved for both of our inverse scale space algorithms.

As an important numerical experiment, we study how our algorithm deals with speckle noise that takes bigger amplitude. The speckle noise used in our experiments follow a Gamma distribution, satisfying a mean of 1 and standard deviation of 0.2. Figure 6.10 shows a 1D example, and Figure 6.11 and Figure 6.12 show the numerical results for 2D images. In all cases, our inverse scale space method restores the best contrast as well as SNR. We can observe that the inverse scale space methods are very suitable for preserving textures, see the Barbara image example.

It is also interesting to note that our numerical algorithms are robust to both types of noise, since the physical model does not make explicit assumptions about the underlying noise distribution, as long as they satisfy the prior information Eqn. (2.2).

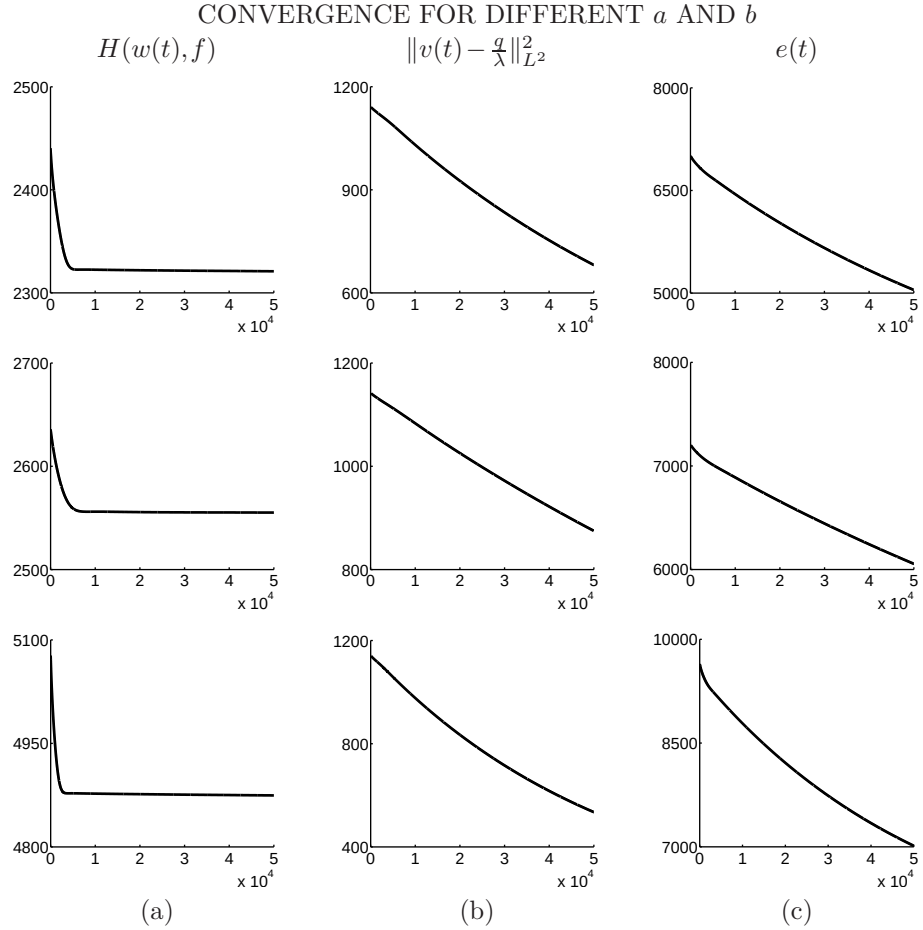


FIG. 6.5. *Convergence for the 1D relaxed inverse scale space flow. Three columns stand for the following energy functionals (a)* $e_1(t) = H(w(t), f)$ *for the multiplicative model, (b)* $e_2(t) = \|v(t) - \frac{q}{\lambda}\|^2_{L^2}$*, (c)* $e(t) = \frac{e_1(t)}{2\lambda} + \frac{e_2(t)}{2\alpha}$*. Three rows are results for flow with different values for a and b respectively, (top row)* $a = 0, b = 1$ *(middle row)* $a = 1, b = 0$*, (bottom row)* $a = 1$*,* $b = 1$*.*

7. Conclusions. In this paper, we study image restoration under multiplicative noise corruption. The multiplicative noise model suffers from non-convexity, which prevents from theoretical or numerical development into the inverse scale space methods. We propose in the paper two ways to overcome this difficulty. The first approach uses a Logarithm transform, which is pretty straightforward. We devote most of our paper to a second approach, where we formulate a general model for the multiplicative noise model. We make this model globally convex by change of variable. We further extend such a model to the nonlinear inverse scale space flow.

The excellent performance of the inverse scale space method, previously shown for the additive model, is extended to multiplicative noise models. Both multiplicative inverse scale space methods show excellent results in image denoising, in texture restoration, contrast preservation and SNR. We tested both of our inverse scale space models, which have performed equally well in terms of denoising. This is a strong indicator that improvement of denoising effect is mostly due to the Bregman boosting that inverse scale space offers, i.e., solution quality is improved by adding the noise back in the Bregman iteration [21]. The choice for the intrinsic models, i.e. the fidelity term $H(u, f)$, is less important, as long as it models the error norm between u and f reasonably well.

We don't make explicit assumptions about the underlying noise distribution in our models, which allows us to test this model on both the Gaussian and speckle noises. Not surprisingly, our

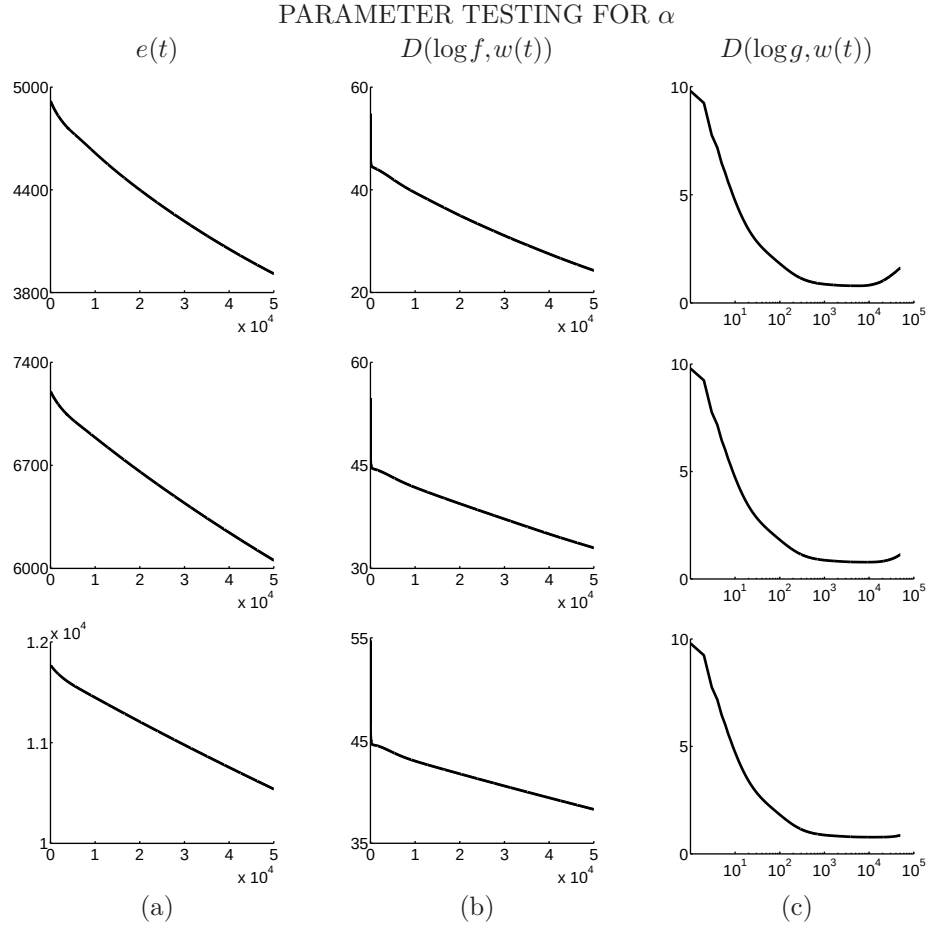


FIG. 6.6. *Parameter study for the relaxed inverse scale space flow in 1D. Three columns stand for (a) flow energy $e(t)$, (b) Bregman distance between noisy data f and the flow, $D(log(f),w(t)$, (c) Bregman distance between ground true g and the flow, $D(log(g),w(t))$. Three rows are numerical results for alpha of different values, from top to bottom, $\alpha = 0.25, 0.125, 0.0625$ respectively, while $\lambda = 1$, $a = 1$, $b = 0$ for all cases.*

model can handle both types of noise and has shown great numerical results. As part of the inverse problem, we assume some prior knowledge about the noise, i.e., the mean and variance of the noise. We take care of the mean of the noise by a proper choice for the initial condition for the inverse scale space flow, and the variance of the noise serves an *a priori* stopping criterion for the flow.

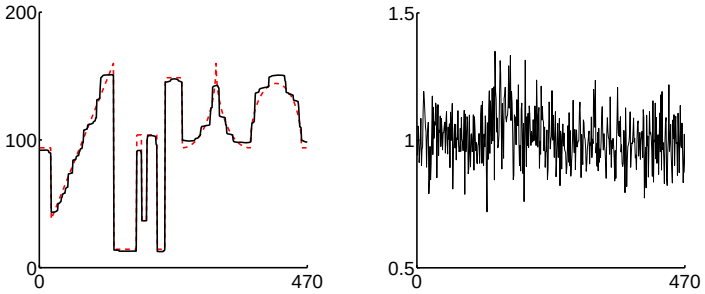
We have analyzed the multiplicative inverse scale space flow for its convergence, as well as regularization effect, which is closely related to the Bregman distance. Convergence of the relaxed inverse scale space flow is also analyzed by linearization. We show numerically such convergence is satisfied as long as the time step in the explicit scheme is small enough.

In our multiplicative inverse scale space flow, choice of the parameters is also investigated. We show that different pairs of a and b do not affect the algorithm behavior. However the choice of λ and α does affect the convergence rate of the algorithm, and the speed of the flow.

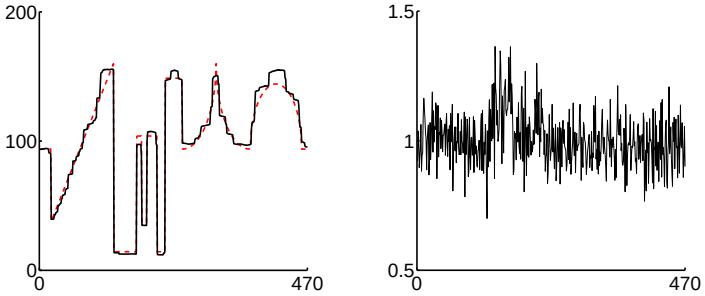
Acknowledgments. This work is supported by NSF grants DMS-0312222, and ACI-0321917 and NIH G54 RR021813. The authors thank Jinjun Xu (UCLA) for providing the test images used in the numerical experiments, Jerome Darbon (UCLA) for some useful discussions, and the unknown referees for constructive comments.

EXPERIMENTS USING GAUSSIAN NOISE

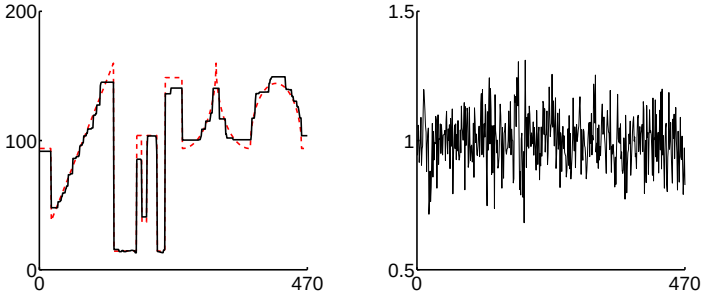
(a) Multiplicative RISS for our convex model



(b) Logarithm transform additive RISS



(c) Rudin-Lions-Osher algorithm



(d) Gradient descent for our convex model

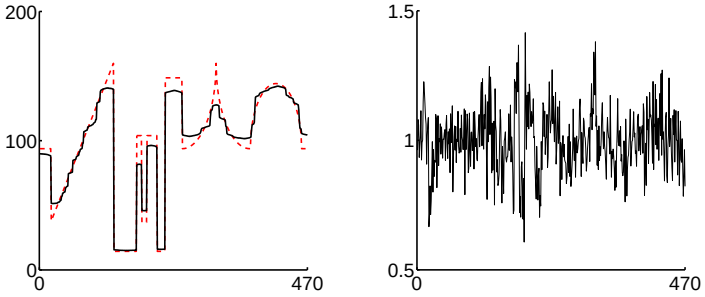


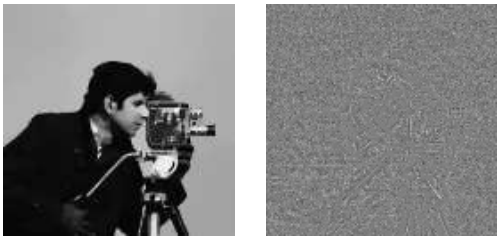
FIG. 6.7. *Comparison of our multiplicative RISS method with previous methods in literature for the multiplicative noise model. In all cases, (left) shows the recovered signal u, (right) shows the recovered noise η. (a) Multiplicative RISS method, where $a=1$, $b=0$, $\lambda=1$, $\alpha=0.125$, $dt=0.0005$. (b) Using the logarithm transform and then applying the additive RISS method, where $\lambda=1$, $dt=0.001$. (c) Rudin-Lions-Osher algorithm, where $dt=0.001$, λ_1 and λ_2 are calculated dynamically using the gradient projection method. (d) Gradient descent on our new globally convex multiplicative model, where $a=1$, $b=0$, $\lambda=1$, $dt=0.001$.*

EXPERIMENTS USING GAUSSIAN NOISE

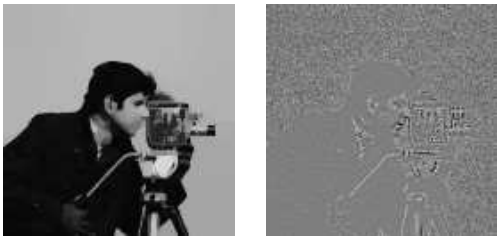
(a) original g (b) noisy f

(c) Multiplicative RISS, $SNR = 37\ dB$

(d) Logarithm transform additive RISS, $SNR = 37\ dB$



(e) Rudin-Lions-Osher algorithm, $SNR = 31\ dB$



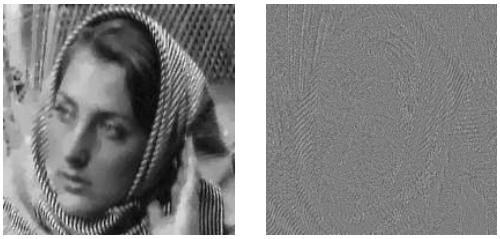
(f) Gradient descent for our convex model, $SNR = 35\ dB$

FIG. 6.8. *Numerical results for 2D image, case study using the "cameraman" image, size 130 × 130. (a) original image g, the ground truth, (b) noisy image f, (c) result using multiplicative RISS method, (left) recovered image u, (right) recovered noise η, multiplied by 128 for better visualization. In this simulation, the parameters are as follows:* $a = 1$, $b = 0$, $\lambda = 1$, $\alpha = 0.25$, $dt = 0.0002$. *Signal-to-noise ratio* $SNR = 37dB$. *(d) Using the logarithm transform and then applying the additive RISS method, where* $\lambda = 1$, $dt = 0.0002$. *(e) Rudin-Lions-Osher algorithm, where* $dt = 0.001$, λ_1 *and* λ_2 *are calculated dynamically using the gradient projection method. (f) Gradient descent on our new globally convex multiplicative model, where* $\lambda = 10$, $dt = 0.001$. *For (c)-(f), (left) shows the recovered image u, (right) shows the recovered noise multiplied by 128.*

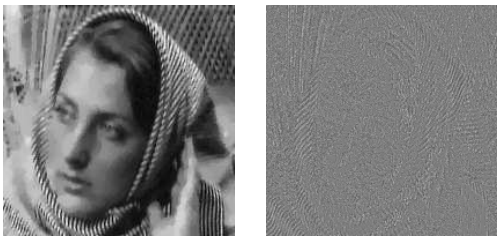
EXPERIMENTS USING GAUSSIAN NOISE

(a) original g (b) noisy f

(c) Multiplicative RISS, $SNR = 35\ dB$



(d) Logarithm transform additive RISS, $SNR = 35\ dB$



(e) Rudin-Lions-Osher algorithm, $SNR = 33\ dB$

(f) Gradient descent for our convex model, $SNR = 34\ dB$

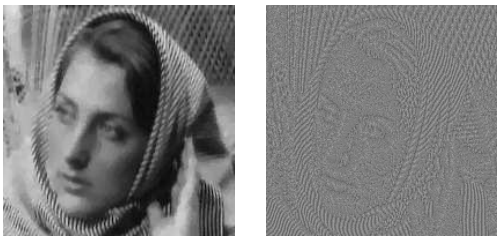
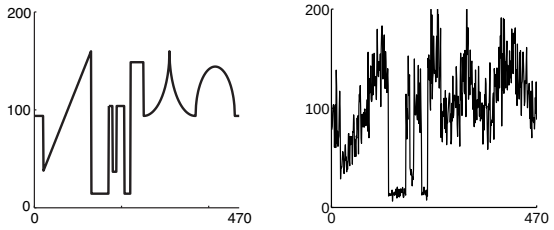


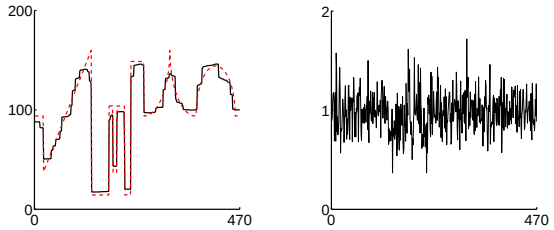
FIG. 6.9. *Numerical results for 2D image, case study using the "Barbara" image, size 200 × 200. (a) original image g, the ground truth, (b) noisy image f, (c) Our new multiplicative RISS algorithm, (left) shows the recovered image u, (right) shows the recovered noise η multiplied by 128. In this simulation, the parameters are as follows:* $a = 1$, $b = 0$, $\lambda = 1$, $\alpha = 0.25$, $dt = 0.001$. *Our simulation shows that the relaxed inverse scale space method preserves the texture information in high quality. (d) Using the logarithm transform and then applying the additive RISS method, where* $\lambda = 1$, $dt = 0.001$. *(e) Rudin-Lions-Osher algorithm, where* $dt = 0.001$, λ_1 *and* λ_2 *are calculated dynamically using the gradient projection method. (f) Gradient descent on our new globally convex multiplicative model, where* $\lambda = 10$, $dt = 0.001$.

EXPERIMENTS USING SPECKLE NOISE

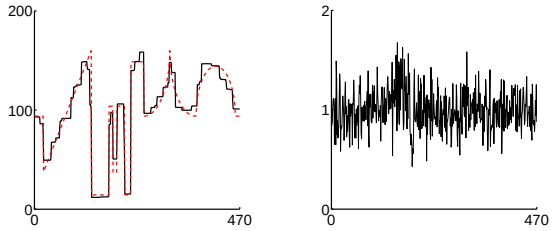
(a) original g (b) noisy f



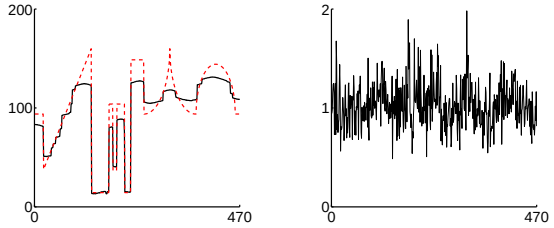
(c) Multiplicative RISS, $SNR = 11.73\ dB$



(d) Rudin-Lions-Osher algorithm, $\lambda = 10$, $SNR = 10.77\ dB$



(e) Aubert-Aujol algorithm, $\lambda = 50$, $SNR = 9.49\ dB$



(f) Aubert-Aujol algorithm, $\lambda = 80$, $SNR = 10.79\ dB$

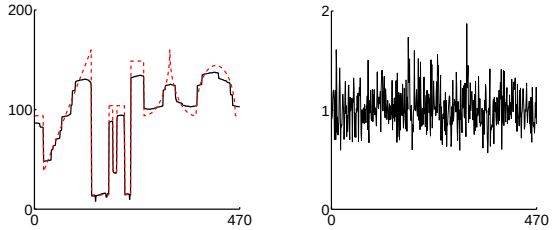
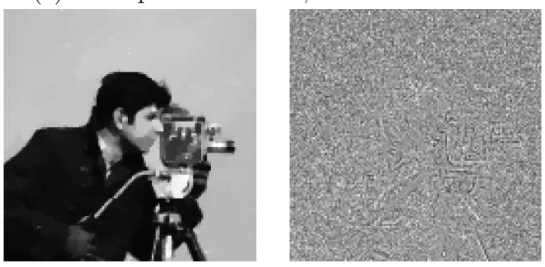


FIG. 6.10. *Numerical results for speckle noise, which follows a Gamma distribution, where the mean of the noise is 1, and standard deviation is 0.2. (a) Original signal. (b) Noisy signal corrupted by speckle noise. (c) Our multiplicative RISS algorithm,* $\lambda = 1$, $\alpha = 0.125$, $a = 1$, $b = 0$, $dt = 0.0005$. *(d) Rudin-Lions-Osher algorithm,* $\lambda = 10$, $dt = 0.01$. *(e) Aubert-Aujol algorithm,* $\lambda = 50$, $dt = 0.1$. *(f) Aubert-Aujol algorithm,* $\lambda = 80$, $dt = 0.1$.

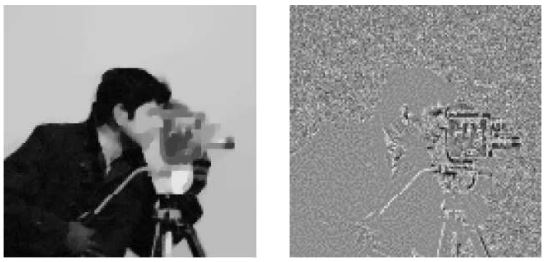
EXPERIMENTS USING SPECKLE NOISE

(a) original g (b) noisy f

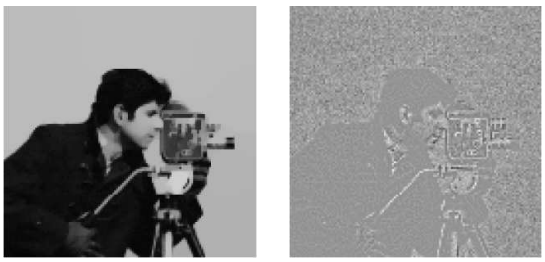
(c) Multiplicative RISS, $SNR = 16.73\ dB$



(d) Rudin-Lions-Osher algorithm, $\lambda = 500$, $SNR = 12.67\ dB$



(e) Aubert-Aujol algorithm, $\lambda = 500$, $SNR = 12.29\ dB$



(f) Aubert-Aujol algorithm, $\lambda = 1000$, $SNR = 15.13\ dB$

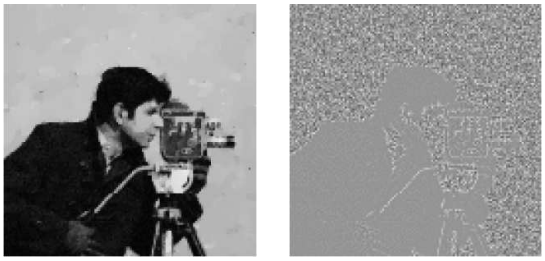
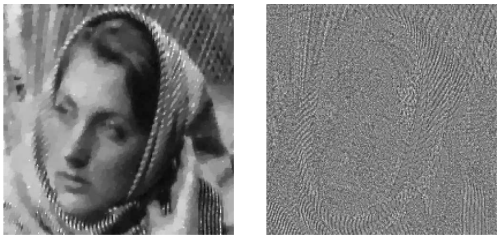


FIG. 6.11. *Numerical results for speckle noise, which follows a Gamma distribution, where the mean of the noise is 1, and standard deviation is 0.2. (a) Original image. (b) noisy image corrupted by speckle noise. (c) Our multiplicative RISS algorithm,* $\lambda = 1$, $\alpha = 0.125$, $a = 1$, $b = 0$, $dt = 0.001$. *(d) Rudin-Lions-Osher algorithm,* $\lambda = 500$, $dt = 0.01$. *(e) Aubert-Aujol algorithm,* $\lambda = 500$, $dt = 0.01$. *(f) Aubert-Aujol algorithm with different set of parameters,* $\lambda = 1000$, $dt = 0.01$.

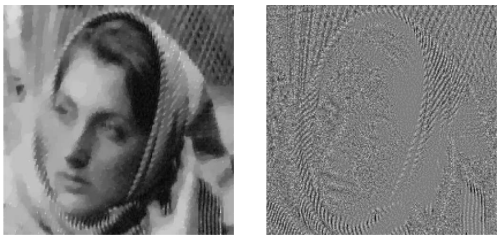
EXPERIMENTS USING SPECKLE NOISE

(a) original g (b) noisy f

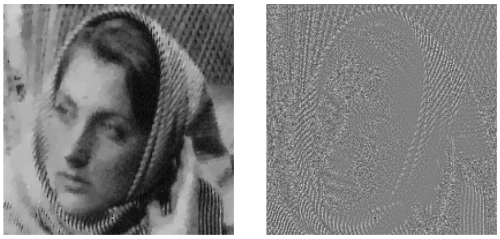
(c) Multiplicative RISS, $SNR = 15.32\ dB$



(d) Rudin-Lions-Osher algorithm, $\lambda = 1000$, $SNR = 14.43\ dB$



(e) Aubert-Aujol algorithm, $\lambda = 800$, $SNR = 13.84\ dB$



(f) Aubert-Aujol algorithm, $\lambda = 2000$, $SNR = 12.14\ dB$

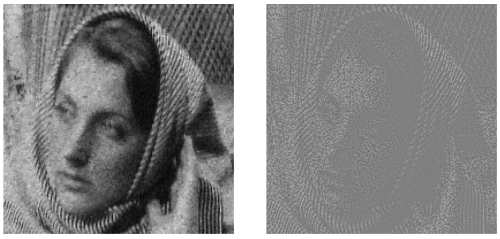


FIG. 6.12. *Numerical results for speckle noise, which follows a Gamma distribution, where the mean of the noise is 1, and standard deviation is 0.2. (a) Original image. (b) Noisy image corrupted by speckle noise. (c) Our multiplicative RISS algorithm,* $\lambda = 1$, $\alpha = 0.125$, $a = 1$, $b = 0$, $dt = 0.0005$. *(d) Rudin-Lions-Osher algorithm, (e) Aubert-Aujol algorithm,* $\lambda = 800$, $dt = 0.01$. *(f) Aubert-Aujol algorithm, with* $\lambda = 2000$, $dt = 0.001$.

REFERENCES

[1] F. Argenti and G. Torricelli, *Speckle suppression in ultrasonic images based on undecimated wavelets*, EURASIP Journal on Applied Signal Processing, 5 (2003), pp. 470-478.

[2] G. Aubert and J.-F. Aujol, *A variational approach to remove multiplicative noise*, SIAM J. Applied Mathematics, volume 68, number 4 (2008), pp. 925-946.

[3] L. M. Bregman, *The relaxation method for finding the common point of convex sets and its application to the solution of problems in convex programming*, USSR Comp. Math. and Math. Phys., 7 (1967), pp. 200-217.

[4] M. Burger, K. Frick, S. Osher, and O. Scherzer, *Inverse Total Variation Flow*, SIAM J. Multiscale Modeling Simulation., 6 (2007), pp. 366-395.

[5] M. Burger, G. Gilboa, S. Osher, and J. Xu, *Nonlinear inverse scale space methods*, Communications in Mathematical Sciences (CMS), 4 (2006), pp. 179-212.

[6] M. Burger, S. Osher, J. Xu, and G. Gilboa, *Nonlinear inverse scale space methods for image restoration*, VLSM, LNCS, 3752 (2005), pp. 85-96.

[7] M. Burger, E. Resmerita, and L. He, *Error estimation for Bregman iterations and inverse scale space methods in image restoration*, Computing, 81(2007), pp. 109-135.

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

[9] T. Chan and J. Shen, *Image Processing and Analysis: Variational, PDE, Wavelet and Stochastic Methods*, SIAM, 2005.

[10] T. Chan and C.K. Wong, *Total variation blind deconvolution*, IEEE Trans. Image Processing, 7 (1998), pp. 370-375.

[11] J. Darbon and M. Sigelle, *Image restoration with constrained total variation. Part 1, fast and exact minimization*, J. Math. Imaging and Vision, 26 (2006), pp. 261-276.

[12] J. Darbon, M. Sigelle and F. Tupin, *Contrast preservation filtering using nice-levelable functions and applications to SAR imaging*, Trans. on Image Proc., (2007) to appear.

[13] I. Ekeland and R. Temam, *Convex analysis and variational problems, volume 1 of Studies in Mathematics and its Applications*, North-Holland Publishing Co., Amsterdam-Oxford (1976).

[14] A. Fan, *Variational Approach to MR Bias Correction*, Master Thesis, MIT, 2003.

[15] D. Goldfarb and W. Yin, *Second order cone programming methods for total variation based image restoration*, SIAM J. Sci. Comput., 27 (2002), pp. 622-645.

[16] C. Groetsch and O. Scherzer, *Nonstationary iterated Tikhonov-Morozov method and third order differential equations for the evaluation of unbounded operators*, Math. Methods Appl. Sci., 23 (2000), pp. 1287-1300.

[17] L. He, M. Burger and S. Osher, *Iterative total variation regularization with non-quadratic fidelity*, J. Math. Imaging and Vision, 26 (2006), pp. 167-189.

[18] J. Lie and J-M. Nordbotten, *Inverse scale spaces for nonlinear regularization*, J. Math. Imaging and Vision, 27 (2007), pp. 41-50.

[19] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, AMS, Providence, RI, 2001.

[20] A. Montillo and J. Udupa and L. Axel and D. Metaxas, *Interaction between noise suppression and inhomogeneity correction in MRI*, Proceedings of SPIE, 5032 (2003), pp. 1025-1036.

[21] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, *An iterative regularization method for total variation-based image restoration*, SIAM J. Multiscale Modeling and Simulation, 4 (2005), pp. 460-489.

[22] L.I. Rudin, P.-L. Lions and S. Osher, *Multiplicative denoising and deblurring: theory and algorithms*, In: Geometric Level Set Methods in Imaging, Vision, and Graphics, chapter Multiplicative denoising and deblurring: theory and algorithms, eds. S. Osher, N. Paragios, 2003, pp. 103-120.

[23] L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259-268.

[24] O. Scherzer, *Explicit versus implicit relative error regularization on the space of functions of bounded variation*, Inverse problems, image analysis, and medical imaging, 313 (2002), pp. 171-198.

[25] O. Scherzer and C. Groetsch, *Inverse scale space theory for inverse problems*, in M.Kerckhove, editor, Scale-Space and Morphology in Computer Vision, Lecture Notes in Comput. Sci. 2106, Springer, New York, 2001, pp. 317-325.

[26] E. Tadmor, S. Nezzar, and L. Vese, *A multiscale image representation using hierarchical* (BV, L^2) *decompositions*, Multiscale Model. Simul., 2 (2004), pp. 554-579.

[27] M. Tur and C. Chin and J.W. Goodman, *When is speckle noise multiplicative?*, Applied Optics, 21 (1982), pp.1157-1159.

[28] R.F. Wagner and S.W. Smith and J.M. Sandrik and H. Lopez, *Statistics of speckle in ultrasound b-scans*, IEEE Trans. on Sonics and Ultrasonics, 30 (1983), pp. 156-163.

[29] J.J. Xu and S. Osher, *Regularization and nonlinear scale space applied to wavelet based denoising*, IEEE Trans on Image Proc., 16 (2007), pp. 534-544.