BLIND DECONVOLUTION USING TV REGULARIZATION AND BREGMAN ITERATION

LIN HE *, ANTONIO MARQUINA †, AND STANLEY J. OSHER ‡

Abstract. In this paper we formulate a new time dependent model for blind deconvolution based on a constrained variational model that uses the sum of the total variation norms of the signal and the kernel as a regularizing functional. We incorporate mass conservation and the nonnegativity of the kernel and the signal as additional constraints. We apply the recently discovered idea of iterative regularization to recover finer scales. We also present an analytical study of the model discussing uniqueness of the solution, convergence to steady state and a priori parameter estimation. We present a simple algorithmic implementation of the model and we perform a series of numerical experiments to show evidence of the good behavior of the numerical scheme and quality of the results.

Key words. Blind Deconvolution, Total Variation, Bregman Distance, Denoising, Deblurring.

1. Introduction. Given a blurry and noisy image $f : \Omega \to \mathbb{R}$,

$$f = k * u + n, \tag{1.1}$$

where Ω is a bounded open subset in $\mathbb{R}^2$, k is a convolution kernel with compact support (e.g., discrete Gaussian kernel), u is the original image and n is Gaussian white noise, our task is to recover the unknown u and/or k from f.

Given knowledge of the kernel k, one of the most successful approaches to solve the above problem is the TV regularization method proposed by L. Rudin, S. Osher and E. Fatemi ([2], [15], [16]). They solved the following constrained minimization problem:

$$u = \arg\min_u \left\{ F(u) := \frac{1}{2}||k * u - f||_2^2 + \alpha||u||_{BV} \right\}, \tag{1.2}$$

where $\alpha > 0$ is a scale parameter which balances the trade-off between removing noise or small details and keeping important features such as sharp edges.

If k is the delta function, $\delta * u = u$, (1.2) becomes a pure denoising algorithm

$$u = \arg\min_u \left\{ F(u) := \frac{1}{2}||u - f||_2^2 + \alpha||u||_{BV} \right\}, \tag{1.3}$$

which effectively recovers the image with edges from noisy data (cf. [2]).

For general convolution kernel k, the restoration of the image u is usually a numerically ill-posed problem. Vogel and Oman have demonstrated that the ill-conditioning in solving the Euler-Lagrange equation of (1.2) is due to the fact that the compact convolution operator, defined by k, has eigenvalues which cluster to zero. They proposed a preconditioned conjugate gradient method combined with the fixed point method. For details see [5]. There are some other related papers on this topic (cf. [17], [19] and [18]).

*UCLA Mathematics Department, Box 951555, Los Angeles, CA 90095-1555, USA. email: helin@math.ucla.edu

†Departamento de Matematica Aplicada, Universidad de Valencia, Spain, e-mail: marquina@uv.es. Research supported by DGICYT project BFM2001-2814 and a fellowship "Beca Salvador de Madariaga" from Ministerio de Educación y Ciencia, Spain.

‡UCLA Mathematics Department, Box 951555, Los Angeles, CA 90095-1555, USA. email: sjo@math.ucla.edu. Research supported by: NSF DMS 0312222, ACI-0321917 and NIH P20MH65166

Now we are interested in recovering the true image u and kernel k without knowing any *a priori* knowledge of k and u. This is what we call the blind deconvolution problem. Instead of minimizing only u in (1.2), we view it as a joint minimization problem with respect to both u and k. For example, You and Kaveh [11] proposed the following model

$$(u,k) = \arg\min_{u,k} \left\{ G(u,k) := \frac{1}{2}||k*u-f||_2^2 + \alpha_1||u||_{H^1}^2 + \alpha_2||k||_{H^1}^2 \right\}. \tag{1.4}$$

However, (1.4) cannot preserve sharp edges in the minimizer u. Moreover, as we will prove in the next section, (1.4) does not yield a local minimizer. In a later work [12] they, and independently Chan and Wong in [1], proposed a similar blind deconvolution method using TV regularization. The minimization problem was formulated as

$$(u,k) = \arg\min_{u,k} \left\{ F(u,k) := \frac{1}{2}||k*u-f||_2^2 + \alpha_1||u||_{BV} + \alpha_2||k||_{BV} \right\}. \tag{1.5}$$

The function $F(u,k)$ depends on both u and k. If we fix u or k and view F as a one-variable function depending only on k or u, respectively, then F is convex. But F is not jointly convex if it is viewed as a two-variable function. In [1], an alternating minimization (AM) algorithm was devised to solve the two Euler-Lagrange equations of u and k alternatively. The idea is to fix k first, solve u from $F_u(u,k)=0$, and then fix u, solve k from $F_k(u,k)=0$, repeatedly:

$$F_u(u,k) = \widehat{k} * (k*u-f) - \alpha_1 \nabla \cdot \frac{\nabla u}{|\nabla u|} = 0, \tag{1.6}$$

$$F_k(u,k) = \widehat{u} * (u*k-f) - \alpha_2 \nabla \cdot \frac{\nabla k}{|\nabla k|} = 0, \tag{1.7}$$

where $\hat{k}, \hat{u}$ are the L^2-conjugate of k and u respectively. The cosine preconditioned conjugate gradient method and the fixed-point method (cf. [19]) were applied to find the solution of each Euler-Lagrange equation. Although it has been numerically shown that the AM algorithm converges quickly and recovers good image features after only a few AM iterations, time delay effects are noticed in the convergent process since u and k are not updated at the same time. In our paper, we deal with (1.6) and (1.7) as a coupled system and evolve it by a time marching method, which is based on the gradient descent method.

Also, as noticed in [1], if (u,k) is a solution, so are $(-u,-k)$, $(\frac{\alpha_2}{\alpha_1}k, \frac{\alpha_1}{\alpha_2}u)$ and $(u(x \pm c), k(x \mp c))$ as well, for any constant vector $c \in \Omega$. We will modify the energy function (1.5) by adding additional constraints for the kernel and the image, removing this source of nonuniqueness. Another more elusive source of nonuniqueness may arise from the following example suggested by J.M. Morel: Let $k_1 = \frac{1}{2}\psi_{[-1,1]}$ where $\psi_{[-1,1]}$ is the characteristic function of the interval $[-1,1]$ and $k = k_1 * k_1$ (k_1 and k are convolution kernels). We choose $u = k_1$ and $f = u*k$. If $u_1 = u*k_1$ we have that $u_1 * k_1 = f$. Then, u, k and u_1, k_1 are different solutions of the blind deconvolution problem with the same energy, since $TV(u) = TV(u_1)$ and $TV(k) = TV(k_1)$.

Finally we will apply iterative regularization using the Bregman distance (cf. [14]) to our new blind deconvolution model. It was recently introduced to image processing by Osher et. al. in [3].

The ideal result of minimizing ROF model (1.3) would be to decompose f into a true signal u and the additive noise $v := f - u$. This is not usually obtained in

practice since it is expected to find some signal in v. In [7], Yves Meyer has analyzed this by introducing the dual of the BV space to extract both the u component in BV and the v component as an oscillating function (texture or noise). On the topic of extracting texture from images using variational methods, we also refer to ([9][10]).

In [7], Meyer introduced the norm of the dual of the BV space, which is called the * norm:

DEFINITION 1.1. *Let G denote the Banach space consisting of all generalized functions $v(x)$ which can be written as*

$$v(x) = \partial_1 g_1(x) + \partial_2 g_2(x) \tag{1.8}$$

where $g_1, g_2 \in L^\infty(R^2)$.

The norm $||v||_*$ of v in G is defined as the lower bound of all L^∞ norms of the functions $|g|$ where $g = (g_1, g_2)$, $|g| = \sqrt{(|g_1|^2 + |g_2|^2)(x)}$ and where the infimum is computed over all decompositions (1.8) of v.

Meyer showed that for the ROF minimization problem, if the norm of f in G does not exceed α, then the optimal ROF decomposition of f is given by

$$u = 0, v = f. \tag{1.9}$$

Otherwise,

$$||v||_* = \alpha \ , \int u(x)v(x)dx = \alpha||u||_{BV}. \tag{1.10}$$

The Euler-Lagrange equation from the ROF minimization is :

$$-\alpha\nabla \cdot \frac{\nabla u}{|\nabla u|} = f - u = v. \tag{1.11}$$

The term v was previously thrown away even though it is an element in G with its * norm $= \alpha$ when $||f||_* > \alpha$. In [3], a correction step was taken which added v back to f and then $f+v$ was processed as a new noisy data by the ROF minimization procedure again. Their iterative regularization procedure is: decompose $f + v_{k-1} = u_k + v_k$ by the ROF model, where $k \geq 1$ and $v_0 = 0$, until the first k for which $||f - u_k|| \leq \delta$ where δ is the L^2 norm of the noise. This iterative idea is equivalent to using the Bregman distance, see [3] for details.

2. Blind Deconvolution Models.

2.1. Existence of Minimizers. First, we will prove the existence of minimizers of $F(u,k)$ in (1.5). We will only present the sketch of the proof derived from the Kondrachov compactness theorem and lower semicontinuous property of the function $F(u,k)$.

Take a minimizing sequence function $\{u_j, k_j\}$ of (1.5), where $F(u_j, k_j)$ decreases to $\min_{u,k} F(u,k)$ which is denoted as M. Since $\{u_j\}$ has bounded BV norm and L^1 norm, by the Kondrachov compactness theorem, this sequence is precompact in $L^1(\Omega)$, which means that up to an extraction, there exists a function u that $u_j \to u$ in $L^1(\Omega)$. Then up to an extraction again, we can actually find $u_j \to u$ *a.e.*. Also from lower semicontinuity of BV norm, we have $\liminf_{j\to\infty} |u_j|_{BV} \geq |u|_{BV}$.

Now among this subsequence of $\{u_j\}$, correspondingly consider for $\{k_j\}$, the same conclusion holds that up to an extraction $k_j \to k$ *a.e.* and $\liminf_{j\to\infty} |k_j|_{BV} \geq |k|_{BV}$.

Since we have $u_j \to u$ *a.e.* and $k_j \to k$ *a.e.*, applying the Fatou Lemma, we obtain

$$\begin{aligned} \liminf_{j\to\infty} \int (k_j * u_j - f)^2 dx &\geq \int \liminf_{j\to\infty} (k_j * u_j - f)^2 dx \\ &= \int (k * u - f)^2 dx \end{aligned}$$

which means that (u, k) is one of the minimizers of (1.5).

Remark *If we replace the TV regularization term of u and k by $\int |\nabla u|^2 dx$ and $\int |\nabla k|^2 dx$ respectively, i.e., H^1 norm , then we still have the existence of minimizers from the Rellich compactness theorem and the fact that $L^2(\Omega) \subset L^1(\Omega)$ for bounded Ω.*

2.2. Analysis of the H^1 Regularization Blind Deconvolution Model. As we have mentioned, even though $F(u, f)$ is convex with respect to u and k respectively (if we fix one of them and consider F as a one-variable function), it is not generally convex for (u, k). To study the behavior of (1.4) we will use the Fourier transform to solve the Euler-Lagrange equations. We will conclude from simple analysis that there are no local minimizers for the kernel k and the image u, if $k(x) = k(-x)$.

If there is a local minimizer (u, k) from (1.4), then (u, k) should be the solution of the Euler-Lagrange equations:

$$\begin{aligned} G_u(u,k) &:= \widehat{k} * (k * u - f) - \alpha_1 \Delta u = 0 \\ G_k(u,k) &:= \widehat{u} * (u * k - f) - \alpha_2 \Delta k = 0 \end{aligned} \tag{2.1}$$

where $\widehat{k}$ is the adjoint of k. These two equations can be solved by using Fourier Transform. Solve for $G_u(u, k)$ at ξ point, we have

$$-\alpha_1 \xi^2 F(u) = |F(k)|^2 F(u) - F(\widehat{k})\, F(f),$$

i.e.

$$F(u) = \frac{F(\widehat{k})\, F(f)}{|F(k)|^2 + \alpha_1 \xi^2} \tag{2.2}$$

Similarly, we have for k

$$F(k) = \frac{F(\widehat{u})\, F(f)}{|F(u)|^2 + \alpha_2 \xi^2} \tag{2.3}$$

Plug $F(u)$ into $F(k)$, use

$$F(\widehat{u}) = \widehat{F(u)} = \frac{F(k)\, F(\widehat{f})}{|F(k)|^2 + \alpha_1 \xi^2},$$

and denote

$$X = |F(k)|^2 + \alpha_1 \xi^2,$$

then we obtain

$$F(k)\, \xi^2 (\alpha_2 X^2) = F(k)\, \xi^2 (\alpha_1 |F(f)|^2) \tag{2.4}$$

(2.4) means that, when $F(k)$ and ξ are not zero, $X^2 = \frac{\alpha_1}{\alpha_2} |F(f)|^2$
Since f is the observed image, α_1 and α_2 are constants, then

$$|F(k)|^2 = -\alpha_1 \xi^2 + \sqrt{\frac{\alpha_1}{\alpha_2}|F(f)|^2}$$

or

$$|F(k)|^2 = -\alpha_1 \xi^2 - \sqrt{\frac{\alpha_1}{\alpha_2}|F(f)|^2}$$

To guarantee $|F(k)|^2 \geq 0$, we can only choose the first expression when

$$\sqrt{\frac{\alpha_1}{\alpha_2}|F(f)|^2} \geq \alpha_1 \xi^2.$$

which means ξ is bounded by above. However, $|\xi|$ could go to ∞, which contradicts the assumption, since $f \in L^2$. Hence, we proved that if there are minimizers then they are not local minimizers.

Remark *The above conclusion does not mean there are no local minimizers for (1.5).*

2.3. Analysis of TV Regularization Blind Deconvolution Model. Following (1.9)(1.10) from Meyer's book (cf. [7]), we obtain a similar result for the deblurring ROF model:

$$u = \arg\min_u \left\{ F(u) := \frac{1}{2}||k * u - f||_2^2 + \alpha_1 ||u||_{BV} \right\} \tag{2.5}$$

LEMMA If the norm of $\widehat{k} * f$ in G does not exceed α_1, then the above energy model gives the minimizer u such that

$$u = 0, v = f. \tag{2.6}$$

Otherwise,

$$||\widehat{k} * v||_* = \alpha_1 \ , \int (k * u)v = \alpha_1 ||u||_{BV}. \tag{2.7}$$

Applying this lemma to our blind deconvolution TV regularization formula, we will also have the same conclusion for k.
Corollary If the norm of $\widehat{u} * f$ in G does not exceed α_2, then (1.5) gives the minimization k that

$$k = 0, v = f. \tag{2.8}$$

Otherwise,

$$||\widehat{u} * v||_* = \alpha_2 \ , \int (u * k)v = \alpha_2 ||k||_{BV}. \tag{2.9}$$

Combining these two interesting facts, if (1.5) has a minimizer u and k and for sufficiently small α_1 and α_2, we will get

$$\begin{aligned} \int (k*u)v &= \alpha_1||u||_{BV} \\ \int (u*k)v &= \alpha_2||k||_{BV} \end{aligned} \tag{2.10}$$

Using $\int (k*u)v = \int (u*k)v$, we have

$$\alpha_1||u||_{BV} = \alpha_2||k||_{BV}. \tag{2.11}$$

Remark *(2.11) will be a criteria which we use to choose α_1 and α_2.*

3. New Model and Numerical Algorithm .

3.1. New Model . As we have noted above, model (1.5) may not yield reasonable solutions due to the obvious fact that it does not have a unique solution. If (u,k) is a solution, so are $(-u,-k)$, $(\frac{\alpha_2}{\alpha_1}k, \frac{\alpha_1}{\alpha_2}u)$ and $(u(x\pm c), k(x\mp c))$, for any vector $c \in \Omega$. In order to obtain a unique reasonable solution, we include the following constraints on u and k:

$$\int k\ dx = 1, \quad \int u\ dx = w = \int f\ dx, \quad u(x), k(x) \geq 0 \tag{3.1}$$

Then our model is modified as the following:

$$(u,k) = \arg\min_{u,k} \left\{ \begin{array}{ll} F(u,k) := & \frac{1}{2}||k*u-f||_2^2 + \alpha_1||u||_{TV} + \alpha_2||k||_{TV} \\ & +\frac{\lambda_1}{2}(\int u\ dx - w)^2 + \frac{\lambda_2}{2}(\int |u| dx - w)^2 \\ & +\frac{\lambda_3}{2}(\int k\ dx - 1)^2 + \frac{\lambda_4}{2}(\int |k|\ dx - 1)^2 \end{array} \right\} \tag{3.2}$$

For λ_1, λ_2, λ_3 and λ_4 large, the constraints from (3.1) are approximately satisfied. We still need to prevent $(u(x\pm c), k(x\mp c))$ from being a solution, thus we continue to assume the symmetry of the kernel k. Since $\widehat{u}*(u*k)$ is symmetric if k is symmetric, we state that if we solve the Euler-Lagrange equation (1.6) and (1.7) with a symmetric kernel as the initial guess, the kernel won't deviate too much when the noise n is not too large relative to the noisy and blurry data f.

In practice (3.2) yields a great improvement in quality of image and stability and convergence of the algorithm. The idea is just to preserve the mass and the positivity of the kernel and the image. Also numerically this model is very easy to implement, the variational derivatives of these four additional terms are constants.

In general, (3.2) yields very accurate solutions. In [1], these constraints were imposed on (u,k) by modifying the results after every AM iteration by solving (1.6) and (1.7). Thus after every AM iteration the accuracy was degraded and some information was lost in the procedure.

3.2. Numerical Algorithm . The AM(Alternating Minimization) method was introduced in [1]. It was proved that for the H^1 blind deconvolution model (1.4) the alternating minimization procedure converges globally, but the solution depends on the initial guess in ([8]). The AM algorithm is stated as follows: Assume that we have u^m and k^m. Then solve for k^{m+1} and u^{m+1} completely by the cosine preconditioned conjugate gradient method and the fixed point method. See the following:

- Solve for k^{m+1}

$$\widehat{u}^m * (u^m * k^{m+1} - f) - \alpha_2 \nabla \cdot \frac{\nabla k^{m+1}}{|\nabla k^{m+1}|} = 0 \tag{3.3}$$

- Solve for u^{m+1}

$$\widehat{k}^{m+1} * (k^{m+1} * u^{m+1} - f) - \alpha_1 \nabla \cdot \frac{\nabla u^{m+1}}{|\nabla u^{m+1}|} = 0 \tag{3.4}$$

There is another variant of the AM algorithm which solves (3.4) first before (3.3). It has been shown in [8] that they have the same limits provided they have the same initial conditions. The initial guesses k_0 and u_0 are usually chosen to be the delta function and the observed image respectively.

It has been demonstrated numerically in [1] [11] that the AM algorithm converges quickly and produces a good restored image after a few AM iterations. However, we still observe time delay effects in the convergent process since we are not updating u and k simultaneously. In this paper we propose a time marching method to update u and k after every small time step iteration, based on a straightforward gradient descent idea. Suppose we have k^n and u^n, plugging them into (3.5), gives us k^{n+1}; plugging them into (3.6), gives us u^{n+1}. Thus we obtain u^{n+1} and k^{n+1}.

- Solve for k^{n+1}

$$\begin{aligned} \tfrac{k^{n+1}-k^n}{dt} = \quad & -\widehat{u}^n * (u^n * k^n - f) + \alpha_2 \nabla \cdot \frac{\nabla k^n}{|\nabla k^n|} \\ & -\lambda_3 \left(\textstyle\int |k^n| \, dx - 1\right) - \lambda_4 \left(\textstyle\int |k^n| \, dx - 1\right) \operatorname{sgn}(k^n) \end{aligned} \tag{3.5}$$

- Solve for u^{n+1}

$$\begin{aligned} \tfrac{u^{n+1}-u^n}{dt} = \quad & -\widehat{k}^n * (k^n * u^n - f) + \alpha_1 \nabla \cdot \frac{\nabla u^n}{|\nabla u^n|} \\ & -\lambda_1 \left(\textstyle\int |u^n| \, dx - w\right) - \lambda_2 \left(\textstyle\int |u^n| \, dx - w\right) \operatorname{sgn}(u^n) \end{aligned} \tag{3.6}$$

3.3. Application of Bregman Distance. Consider the generalized model:

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

where J and H are convex functionals of u, $J(u) \geq 0$.

DEFINITION 3.1. *Given a differentiable function φ, the Bregman distance (cf.[14]) is defined by*

$$D_\varphi(x, y) = \varphi(x) - \varphi(y) - < x - y, \partial\varphi(y) >, \tag{3.7}$$

where $< \cdot, \cdot >$ denotes the inner product in R^n and $\partial\varphi(y)$ is an element of the sub-gradient of φ at point y.

Since $J(u)$ is convex, $D_J(u, v)$ is also convex in u for each v. The quantity $D_J(u, v)$ is not a distance in the usual sense; e.g., in general, $D_J(u, v) \neq D_J(v, u)$ and also the triangle inequality does not hold. However, it is a measure of closeness in that $D_J(u, v) \geq 0$ and $D_J(u, v) = 0$ if $u = v$ (if and only if for strictly convex functionals).

Suppose u_0 is the minimizer of $min_u \{J(u) + H(u, f)\}$. An iterative procedure is given by the sequence of variational problems

$$u_m = \arg \min_{u \in BV(\Omega)} \{Q(u, u_{m-1}, f) =: H(u, f) + D_J(u, u_{m-1})\} \tag{3.8}$$

In this paper, we consider:

$$\begin{aligned} H(u, f) &= \tfrac{1}{2}||k * u - f||_2^2, \\ J(u) &= \alpha_1 ||u||_{BV} + \tfrac{\lambda_1}{2}(\int u \, dx - w)^2 + \tfrac{\lambda_2}{2}(\int |u| \, dx - w)^2. \end{aligned} \tag{3.9}$$

Remark *In $J(u)$, the terms $\alpha_2||k||_BV + \frac{\lambda_3}{2}(\int k \, dx - 1)^2 + \frac{\lambda_4}{2}(\int |k| \, dx - 1)^2$ are not included because when we minimize for fixed k, these terms are just constants.*

In the following we show that it is easy to program this algorithm.
Take the sub-gradient of (3.8) for our problem. For $m \geq 1$, we have :

$$H_u(u_m, f) + \partial J(u_m) - \partial J(u_{m-1}) = 0. \tag{3.10}$$

While for $m = 0$,

$$\widehat{k} * (k * u_0 - f) + \partial J(u_0) = 0;$$

Write $f - k * u_0 = v_0$, i.e.

$$\partial J(u_0) = \widehat{k} * v_0.$$

Plug the above equation into (3.10) for $m = 1$, it becomes

$$\widehat{k} * (k * u_1 - f) + \partial J(u_1) - \widehat{k} * v_0 = 0.$$

It can be simplified as

$$\widehat{k} * (k * u_1 - f - v_0) + \partial J(u_1) = 0.$$

Write $f + v_0 - k * u_1 = v_1$, we then have

$$\partial J(u_1) = \widehat{k} * v_1.$$

From above we can see, if we denote

$$v_m = f + v_{m-1} - k * u_m \ \forall m, \tag{3.11}$$

we will have

$$\partial J(u_m) = \widehat{k} * v_m. \tag{3.12}$$

In other words, to apply the Bregman iteration we only have to change the observed image from f to $f + v_m$ at the $(m + 1)$th iteration.

The general procedure of our numerical algorithm using the Bregman iteration is this: starting from a guess for the image u and the kernel k which are the observed image f and δ function correspondingly, we solve the Euler-Lagrange equation (3.5)and (3.6), and denote the solution as k_1 and u_{10}. Fixing the kernel k_1, we apply the Bregman iteration on u m times, and write the result as u_{1m}. This is called a round. We expect this to give a better image than u_{10} because $H(u_{1m}, f)$ will always decrease with m (shown in [3]). Furthermore, $H(u_{1m}, f)$ converges to zero for a pure deblurring problem, and for noisy data, the Bregman distance between u_{1m} and the true solution $\tilde{u}$ is decreasing as long as the residual $H(u_{1m}, f)$ lies above the noise level. These results were obtained in [3].

We next go back to the Euler-Lagrange equation (3.5) and (3.6) with what we expect to be a better guess k_1 and u_{1m}, and also repeat the iteration for another round. A better kernel than k_1 is expected. We repeat these rounds and stop as soon as the recovered image becomes noisy.

4. Explicit Numerical Scheme for the 2D Model. We can write our 2D model as:

$$\begin{aligned} u_t = \; & -\widehat{k} * (k * u - f) + \alpha_1 \frac{u_{xx}(u_y^2+\beta) - 2u_{xy}u_x u_y + u_{yy}(u_x^2+\beta)}{(u_x^2+u_y^2+\beta)^{\frac{3}{2}}} \\ & + \lambda_1 \left(\int |u| \, dx - w\right) + \lambda_2 \operatorname{sgn}(u)\left(\int |u| \, dx - w\right) \end{aligned} \tag{4.1}$$

$$\begin{aligned} k_t = \; & -\widehat{u} * (u * k - f) + \alpha_2 \frac{k_{xx}(k_y^2+\beta) - 2 * k_{xy}k_x k_y + k_{yy}(k_x^2+\beta)}{(k_x^2+k_y^2+\beta)^{\frac{3}{2}}} \\ & + \lambda_3 \left(\int |k| \, dx - 1\right) + \lambda_4 \operatorname{sgn}(k)\left(\int |k| \, dx - 1\right) \end{aligned} \tag{4.2}$$

using f and the delta function as initial guesses for u and k respectively, and homogeneous Neumann boundary conditions. The parameter $\beta > 0$ is the regularization parameter for the sgn function. It is chosen to be very small for a pure deblurring problem. We usually pick it as 0.01, and change it to $1.0e-6$ after the recovered image and kernel are improved. This usually happens after a couple rounds of the procedure.

We consider u_{ij}^n and k_{ij}^n the approximations to $u(x_i, y_i, t_n)$ and $k(x_i, y_i, t_n)$, respectively where $x_i = idx$, $y_i = jdy$, $t_n = ndt$.

We denote by

$$v_{ij}^n = [\widehat{k}^n * (k^n * u^n - f)]_{ij} + \lambda_1 \left[\int |u^n| \, dx - w\right] + \lambda_2 \operatorname{sgn}(u_{ij}^n)\left[\int |u^n| \, dx - w\right] \tag{4.3}$$

and

$$w_{ij}^n = [\widehat{u}^n * (u^n * k^n - f)]_{ij} + \lambda_3 \left[\int |k^n| \, dx - 1\right] + \lambda_4 \operatorname{sgn}(k_{ij}^n)\left[\int |k| \, dx - 1)\right] \tag{4.4}$$

where the convolutions are computed using the discrete cosine transform (DCT) to enforce the homogeneous Neumann boundary condition. The explicit scheme for the simultaneous evolution of the signal and the kernel reads as follows:

$$\frac{u_{ij}^{n+1} - u_{ij}^{n}}{dt} = -v_{ij}^{n} + \alpha_1 \, s_{ij}^{n}(u) \tag{4.5}$$

$$\frac{k_{ij}^{n+1} - k_{ij}^{n}}{dt} = -w_{ij}^{n} + \alpha_2 \, s_{ij}^{n}(k) \tag{4.6}$$

while the second order term $s_{ij}^{n}(u)$, (resp. $s_{ij}^{n}(k)$) is computed by means of

$$s_{ij}^{n}(u) := \frac{g_{ij}^{xx}(u)(g_{ij}^{y}(u)^2 + \beta) - 2g_{ij}^{xy}(u)g_{ij}^{x}(u)g_{ij}^{y}(u) + g_{ij}^{yy}(u)(g_{ij}^{x}(u)^2 + \beta)}{(g_{ij}^{x}(u)^2 + g_{ij}^{y}(u)^2 + \beta)^{\frac{3}{2}}}, \tag{4.7}$$

where

$$g_{ij}^{x}(u) = \frac{u_{i+1,j}^{n} - u_{i-1,j}^{n}}{2dx}, \tag{4.8}$$

$$g_{ij}^{y}(u) = \frac{u_{i,j+1}^{n} - u_{i,j-1}^{n}}{2dy}, \tag{4.9}$$

$$g_{ij}^{xx}(u) = \frac{u_{i+1,j}^{n} - 2u_{ij}^{n} + u_{i-1,j}^{n}}{dx^2}, \tag{4.10}$$

$$g_{ij}^{yy}(u) = \frac{u_{i,j+1}^{n} - 2u_{ij}^{n} + u_{i,j-1}^{n}}{dy^2}, \tag{4.11}$$

$$g_{ij}^{xy}(u) = \frac{u_{i+1,j+1}^{n} - u_{i-1,j+1}^{n} - u_{i+1,j-1}^{n} + u_{i-1,j-1}^{n}}{4dxdy}, \tag{4.12}$$

The expressions for $s_{ij}^{n}(k)$ are the same using k instead of u.

To speed up the convergence of k to the true kernel solution (which is a Gaussian kernel in our examples), we will compute the following equation in place of (4.5). This model is based on level set motion and was formulated by A. Marquina and S. Osher in [4]. It is very efficient and fast. Our model for k is expressed in terms of explicit partial derivatives as

$$\begin{aligned} k_t = \; & -\sqrt{k_x^2 + k_y^2}\,\{\widehat{u}^n * (u^n * k^n - f) + \alpha_2 \, \tfrac{k_{xx}(k_y^2+\beta) - 2 * k_{xy} k_x k_y + k_{yy}(k_x^2+\beta)}{(k_x^2+k_y^2+\beta)^{3/2}} \\ & + \lambda_3 \, (\textstyle\int |k| \; dx - 1) + \lambda_4 \, \mathrm{sgn}(k)(\int |k| \; dx - 1)\} \end{aligned} \tag{4.13}$$

The first order scheme reads as follows:

$$\frac{k_{ij}^{n+1} - k_{ij}^{n}}{dt} = -\sqrt{k g_{ij}^{x\,2} + k g_{ij}^{y\,2}}\,(w_{ij}^{n} + \alpha_2 \, s_{ij}^{n}(k)), \tag{4.14}$$

where kg_{ij}^x and kg_{ij}^y are defined as follows from Rouy and Tourin, [13],

$$kg_{ij}^{x\,2} = \max(\max(\frac{k_{ij}^n - k_{i-1,j}^n}{dx}, 0)^2, \min(\frac{k_{i+1,j}^n - k_{i,j}^n}{dx}, 0)^2) \tag{4.15}$$

$$kg_{ij}^{y\,2} = \max(\max(\frac{k_{ij}^n - k_{i,j-1}^n}{dy}, 0)^2, \min(\frac{k_{i,j+1}^n - k_{i,j}^n}{dy}, 0)^2) \tag{4.16}$$

if $w_{ij}^n > 0$, and

$$kg_{ij}^{x\,2} = \max(\min(\frac{k_{ij}^n - k_{i-1,j}^n}{dx}, 0)^2, \max(\frac{k_{i+1,j}^n - k_{i,j}^n}{dx}, 0)^2) \tag{4.17}$$

$$kg_{ij}^{y\,2} = \max(\min(\frac{k_{ij}^n - k_{i,j-1}^n}{dy}, 0)^2, \max(\frac{k_{i,j+1}^n - k_{i,j}^n}{dy}, 0)^2) \tag{4.18}$$

if $w_{ij}^n < 0$. For more details, please see ([4]).

The speed up procedure is only applied to the kernel evolution equation. The signal evolution equation (4.6) remains unchanged. Since we start with the blurred and noisy signal f, we need to increase total variation of the signal to restore u. If we use the speed up procedure for the signal it is impossible to recover local extrema because of the maximum principle.

5. Numerical Results. In this section we present some of the results we have obtained using our blind deconvolution TV regularization model with the iterated regularization procedure. We also compare this with the results obtained using the H^1 norm of k instead of the BV norm of k.

We are given a scaled noisy and blurred gray image f that takes value from [0,1]. We take the image f as the initial guess for u. We use δ function as the initial guess for the kernel k. Inside one round described above, we run 1000 iterations to solve the Euler-Lagrange equations for k and u, and then another 1000 iterations to update u by the Bregman iteration. For a 64 by 64 matrix image, this takes 90 seconds for a computer with a Mobile Pentium 4 1.8Ghz processor. Generally the full algorithm converges after 3 rounds, that means all the work takes less than five minutes.

We usually pick $\lambda_1 = \lambda_2 = 0.001$ and $\lambda_3 = \lambda_4 = 10$ to enforce the positivity of the kernel and the image and preserve their mass. Also, from (2.11), we know $\alpha_1||u||_{BV} = \alpha_2||k||_{BV}$, i.e., the choice of α_1 and α_2 are dependent on the proportion of the TV norm of the image and the kernel. Of course, we do not know the value of both norms until we try some preliminary experiments.

We use white Gaussian noise $\eta \sim \mathcal{N}(0, \sigma^2)$, $\|\eta\|_{L^2} \approx \sigma$.

$$SNR := 20 * \log_{10}\left(\frac{\|f - \bar{f}\|_{L^2}}{\|\eta - \bar{\eta}\|_{L^2}}\right)$$

is the signal-to-noise ratio, measured in decibels(dB). $\bar{f}$ and $\bar{\eta}$ are the means of f and η over Ω.

Figure 5.1 with $SNR = 1.47$ from (a) to (i) displays the results of our numerical algorithm including the Bregman iteration. We have run this experiment for three rounds. In every round we only apply the Bregman iteration to u once, thus, including the kernel, we obtain nine pictures. We observe that the iterated refinement procedure gives us more details of the image. It is sometimes the case that in the earlier rounds it also brings back some noise. However, as we do more rounds the recovered kernel gets better.

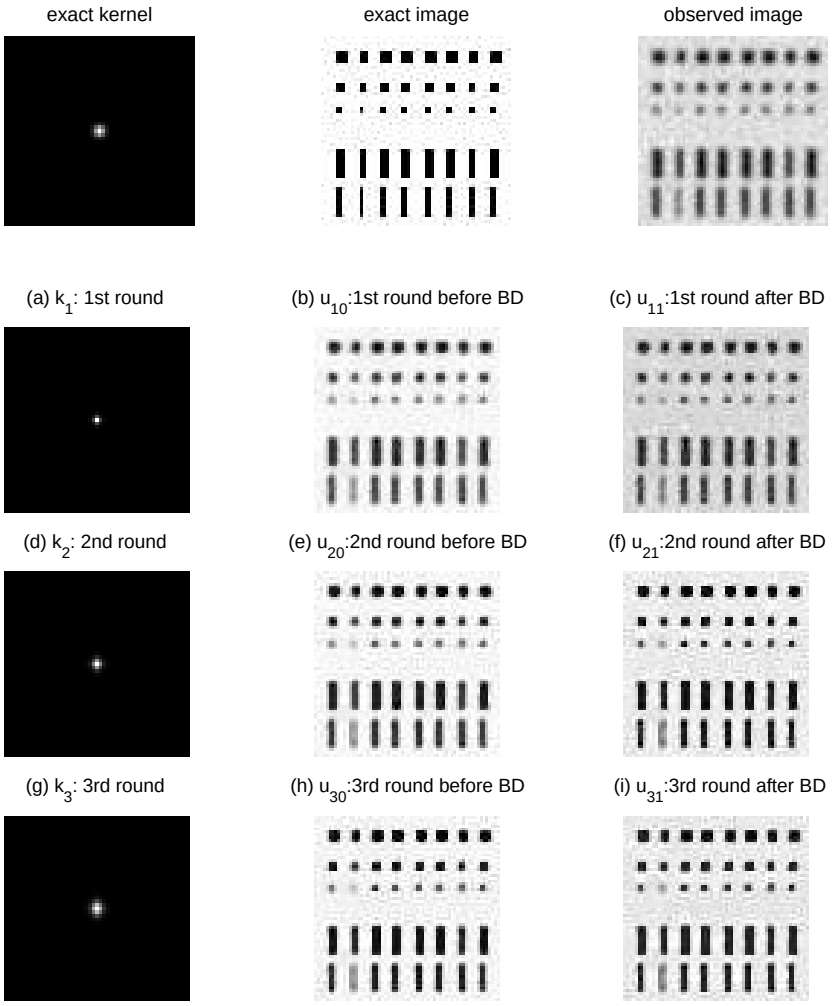


FIG. 5.1. *Example of numerical algorithm procedure, SNR = 1.47. We observe that the recovered kernel does improve after every round, and the Bregman iteration gives us more details of the image. However, the Bregman iteration brings back some noise in the first round.*

Figure 5.2 shows us how to choose good parameters α_1, α_2. The smaller the parameter, the more details we have of the picture in the very early rounds. This sometimes unfortunately includes noise, see (g)$\sim$(i). The bigger the parameter, the more we smooth the image u and obtain less detail, see (d) $\sim$ (f). Since the main purpose of applying the Bregman iteration is to bring back details without noise, the bigger parameter is preferred; however, one must remember that the noise will eventually come back in the later rounds, for proof and examples see [3].

Figure 5.3 is the noise free case. From [3] we know that the Bregman distance $D(u_m, \tilde{u})$ converges to zero, and we do see this. We have recovered more and more details as we repeatedly apply the iterated refinement regularization procedure. Comparing (e), (f) and (g), we see that (g) from deconvolution with the Bregman iteration is the best, but (e) from blind deconvolution with the Bregman iteration is as good

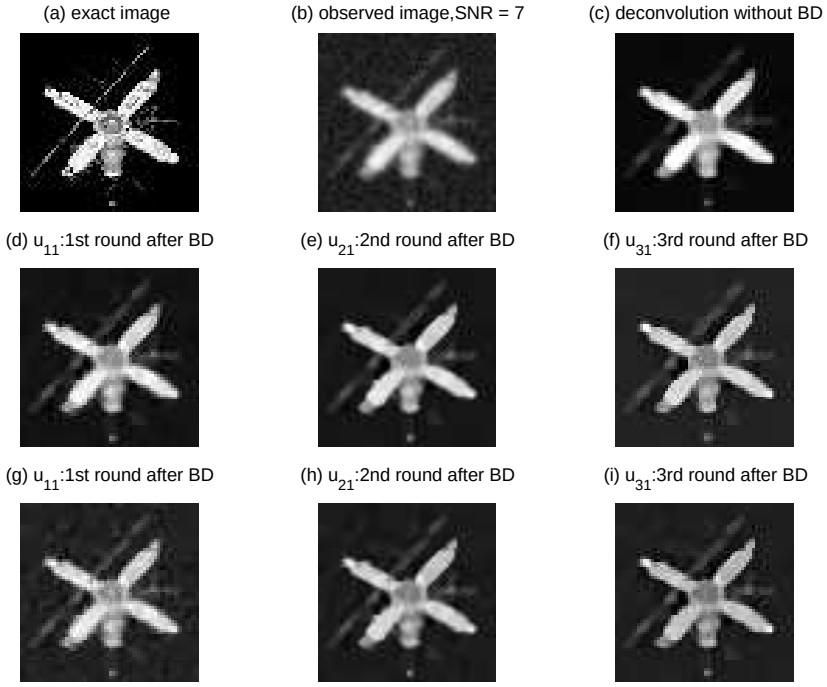


FIG. 5.2. *Compare different α_1, α_2. Notice that (d)$\sim$(f) with fewer details is smoother than (g)$\sim$(i).*

as (f) from deconvolution without the Bregman iteration.

In Figure 5.4 we apply our blind deconvolution model to a pure blurred image of the 256×256 resolution of the satellite image. The blurring kernel used was a 129×129 kernel, appearing Figure 5.7 (top right). We obtain a deconvolved image and a symmetric kernel shown on the bottom. Our resolution improves the results of previous tests done in [1]. The first and second Bregman iterates appear at the bottom of Figure 5.4, where we observe that details are no recovered in this case.

In Figure 5.5 we apply our blind deconvolution model to a blurred and noisy image of the 256×256 resolution of the satellite image. The size of the obtained blurring kernel is 129×129. This numerical test is a demanding example, since both blur and noise are strong. We obtain a deconvolved image and a symmetric kernel shown in Figure 5.5 (bottom). In Figure 5.6 we display two iterations of the iterative regularization procedure applied to the TV regularization of the signal. For the sake of comparison we display in Figure 5.7 the TV restoration of the original blurred and noisy signal, using the blurring kernel, and the first iterate of the iterative regularization applied to the TV equation. We observe in this last picture the presence of more details plus some noise.

In Figure 5.8 we perform blind deconvolution with the iterated refinement method on a blured and noisy 64×64 resollution satellite image using two diferent initial guesses for the kernel. (c) and (d) are obtained using the exact kernel as the initial kernel. Comparing (c) and (d) with (a) and (b) where the δ function is used as the initial kernel, we do not observe very remarkable differences. It is not surprising to

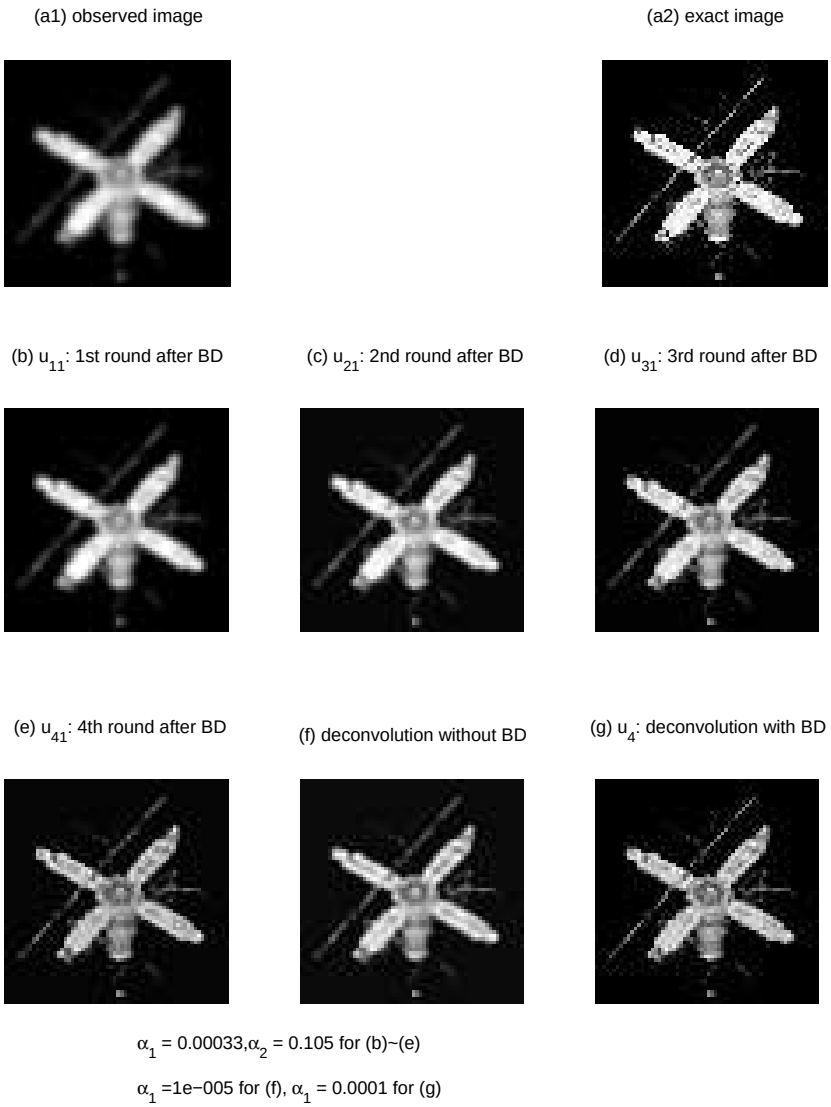


FIG. 5.3. *Noise free case: from (b) to (e), we see that more details are restored*

see that the algorithm converges faster for a better initial kernel k. It is possible that this model, with the symmetry assumption, converges globally, and the result does not depend on the initial guess.

In Figure 5.9, we show the results obtained by regularizing with the square of the H^1 norm of the kernel instead of using the TV norm of kernel. The observed image is taken from Figure 5.8 with $SNR = 2.6$. Everything else is kept the same, including the constraint terms and our algorithm. Figures (a) and (b) are from blind deconvolution without applying the Bregman iteration. Figure (c) and (d) are from blind deconvolution with the Bregman iteration. Even though we can not tell the differences between (b) and (d) by eye, the L^2 norm of the distance from the exact image *uexact* shows us the effectiveness of the Bregman iteration. Comparing the

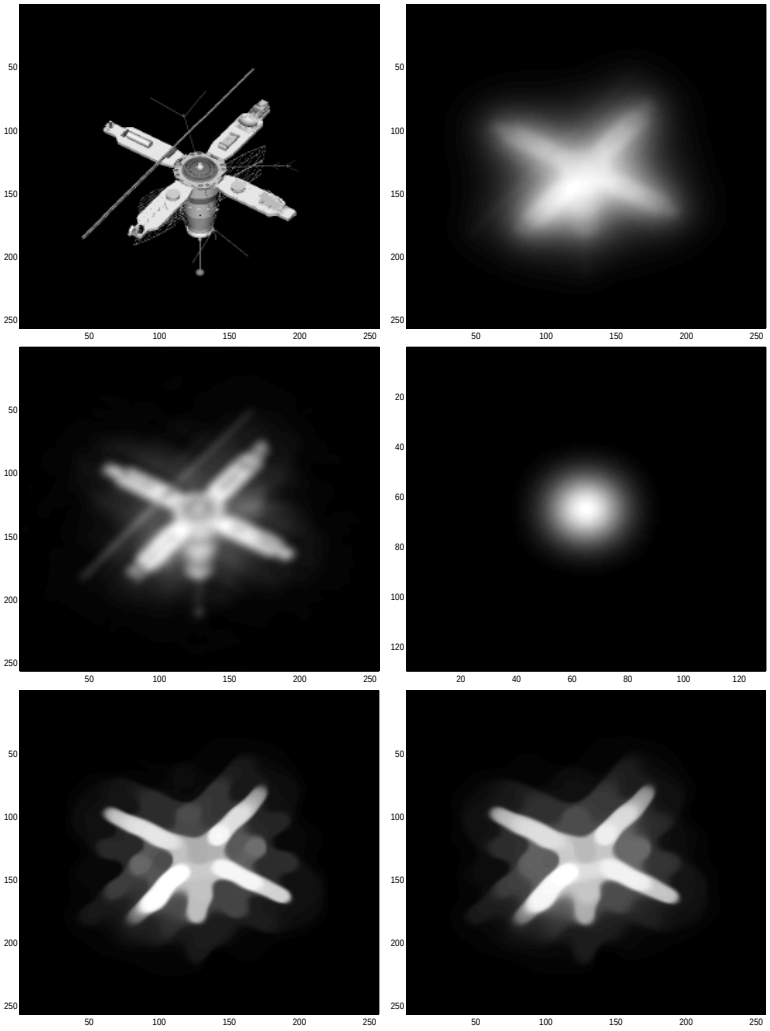


FIG. 5.4. *Top left, original image. Top right, pure blurred Image. Middle left, blind deconvolution: resulting signal after 3200 iterations of gradient descent for Euler-Lagrange Eqns. done simultaneously for the kernel and the signal. Middle right, resulting kernel. Bottom left, first Bregman iteration. Bottom right, second Bregman iteration*

results with Figure 5.8, we see that using the H^1 norm of the kernel is almost good as using the TV norm of kernel. However there is a small disadvantage to use the H^1 norm of the kernel, which is that we do not know how to choose good parameters for α_1 and α_2.

Figure 5.10 displays the results of applying our new model with the Bregman iteration to a blurred and noisy galaxy picture. The picture itself is complicated, because the bright stars could be treated as noise. The result from using blind deconvolution with the iterated regularization method (b), shows that we still recover a good image as in (a), where we know the exact kernel. (c) and (d) are presented to show again, that our model with the Bregman iteration does not depend on the initial guess, and using the H^1 norm of the kernel k can recover a good picture as well if we know how

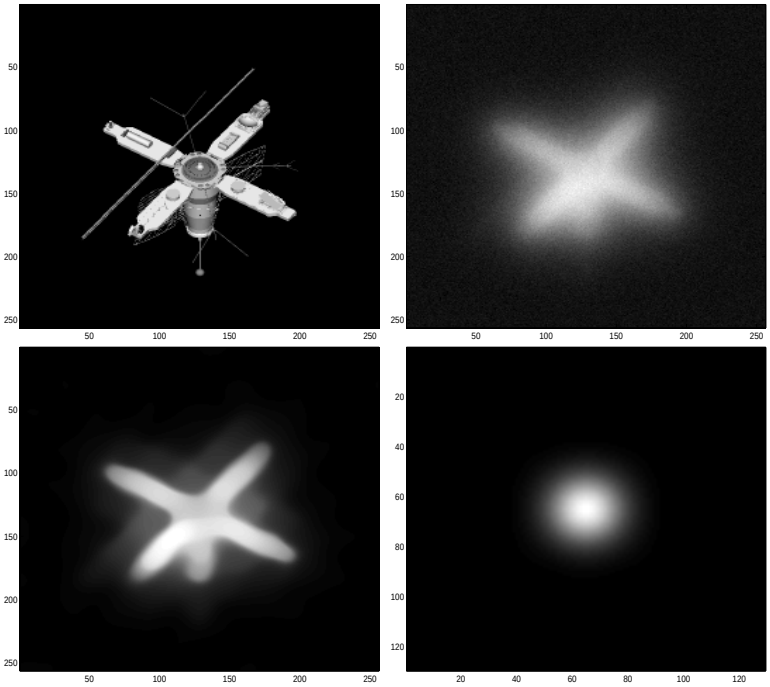


FIG. 5.5. *Top left, original image. Top right, original blurred and noisy signal. Bottom left, blind deconvolution: Resulting signal after 3300 iterations of gradient descent Euler-Lagrange eqns for the kernel and the signal starting with the delta function and the blurred and noisy signal. Bottom right, Resulting kernel*

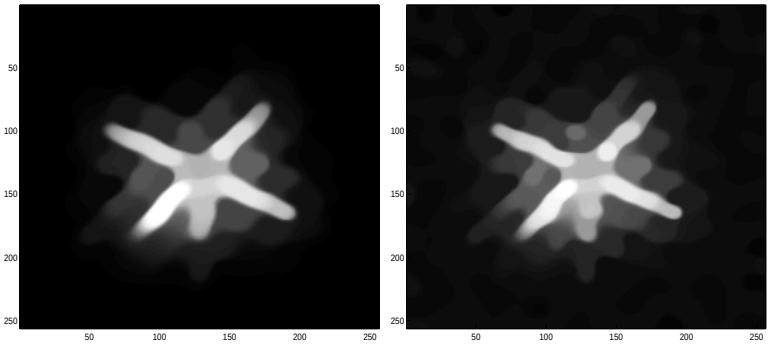


FIG. 5.6. *Left, first iteration of the Bregman iterative regularization for the TV model, starting with the resulting signal and kernel obtained from the blind deconvolution process. Right, second iteration of the Bregman iterative regularization for the TV model.*

to choose the value of the parameters α_1 and α_2.

REFERENCES

[1] Tony F. Chan and C.K. Wong, *Total Variation Blind Deconvolution*, IEEE Transactions on Image Processing, Vol 7, pp. 370-375, March 1998.

[2] Leonid I.Rudin, Stanley Osher and Emad Fatemi, *Nonlinear Total Variation Based Noise Removal Algorithms*, Physica D. 60(1992), pp. 259-268.

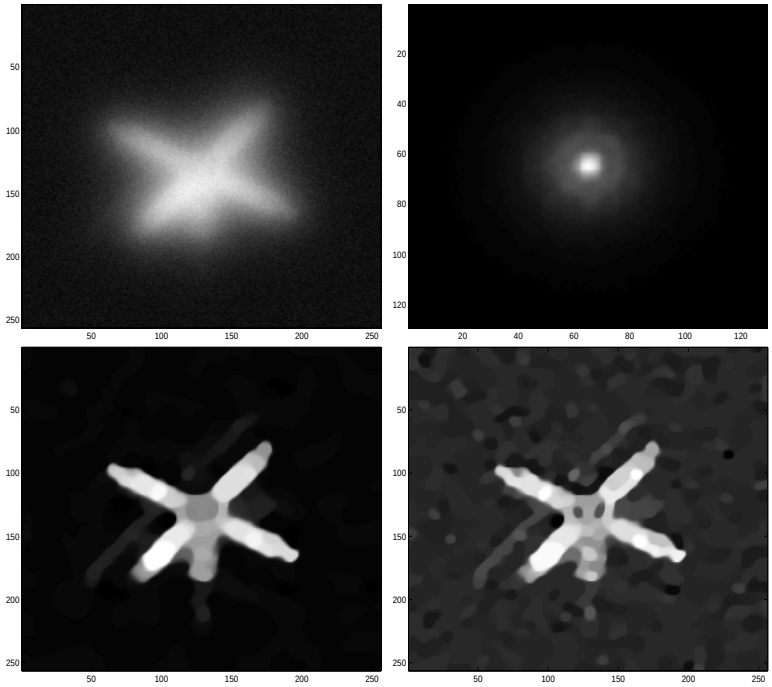


FIG. 5.7. *Left, TV deconvolution using the kernel and the blurred and noisy image: Top left: blurred and noisy image image. Top right: experimental blurring kernel. Bottom left: TV deconvolution with* $\alpha_1 = 0.00055$. *Bottom right, first iteration of the Bregman iterative regularization starting with the solution of the TV model*

[3] Stanley Osher, Martin Burger, Donald Goldfarb, Jinjun Xu and Wotao Yin, *An Iterated Regularizaiton Method for Total Variation Based on Image Restoration*, UCLA CAM Report, 04-13, (2004).

[4] Antonio Marquina and Stanley Osher, *Explicit Algorithms for a New Time Dependent Model Based on Level Set Motion for Nonlinear Deblurring and Noise Removal*, SIAM J. Sci. Comput. Vol 22, No2, pp.387-405, 2000.

[5] Curtis R. Vogel and Mary E. Oman, *Fast, Robust Total Variation-Based Reconstruction of Noisy, Blurred Images*, IEEE Transaction on Image Processing, Vol 7, No 6, June 1998.

[6] Gong Chen and Marc Teboulle, *Convergence Analysis of a Proximal-Like Minimization Algorithm using Bregman Functions* , SIAM J. Optimization, Vol 3, No 3, pp.538-543, August 1993.

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

[8] Tony F. Chan and C.K. Wong, *Convergence of the Alternating Minimization Algorithm for Blind Deconvolution*,

[9] Luminita A. Vese and Stanley J. Osher, *Modeling Textures with Total Variation Minimization and Oscillating Patterns in Image Processing*, J.Scientif. Comput. 19(2003), 553-572.

[10] J.-F. Aujol and A.Chambolle, *Dual Norms and Image Decomposition Models*, INRIA Rapport de Recherche #5130, (2004)

[11] Y.You and M.Kaveh, *A Regularization Approach to Joint Blur identification and Image Restoration*, IEEE Transaction on Image Processing, Vol 5 pp. 416-428, March 1996.

[12] Y.You and M.Kaveh, *Anisotropic Blind Image Restoration*, IEEE International Conference Image Proc. Lausanne '96.

[13] Rouy, E. and Tourin, A., *A Viscosity Solutions Approach to Shape-From-Shading*, SIAM J.Num. Anal 29, 867-884(1992).

[14] L.Bregman, *The Relaxation Method of Finding the Common Points of Convex Sets and its Application to the Solution of Problems in Convex Programming*, U.S.S.R. Comput. Math. and Math. Phys., 7(1967), pp. 200-217.

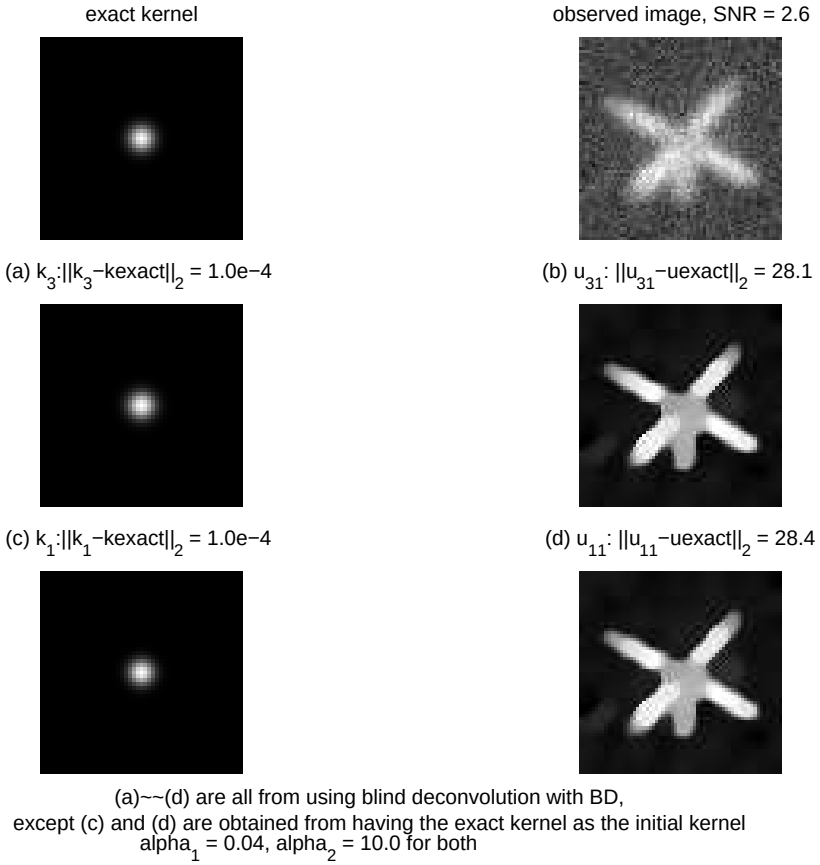


FIG. 5.8. *Comparison 1: We show that a better initial guess for the kernel does not improve the result.*

[15] Lions P.-L, Osher S.J., Rudin L. (1992): *Denoising and Deblurring Using Constrained Nonlinear Partial Differential Equation*. tech. reps., Cognitech Inc., Santa Monica, CA, submit. to SINUM.

[16] Rudin L., Osher S.J. (1994): *Total Variation based Image Restoration with Free Local Constraints*. in Proc. of the IEEE ICIP-94, vol. 1, Austin, TXm pp31-35.

[17] Vicente F.Candela, Antonio Marquina and Susana Serna (2003) *A Local Spectral Inversion of a Linearized TV Model for Denoising and Deblurring*, IEEE Transactions on Image Processing, Vol. 12, No 7, July 2003.

[18] Raymond H. Chan, Tony F. Chan, and Chiu-Kwong Wong, *Cosine Transform Based Preconditioners of Total Variation Deblurring*, UCLA CAM report, September 1997.

[19] R. Chan, T. Chan and C. Wong. *Cosine Transform Based Preconditioners for Total Variation Minimization Problems in Image Processing*. In S. Margenov and P. Vassilevski, editors. Iterative Methods in Linear Algebra, II, pp. 311-329, IMACS Series in Computational and Applied Math., V3, IMACS, NJ, 1996.

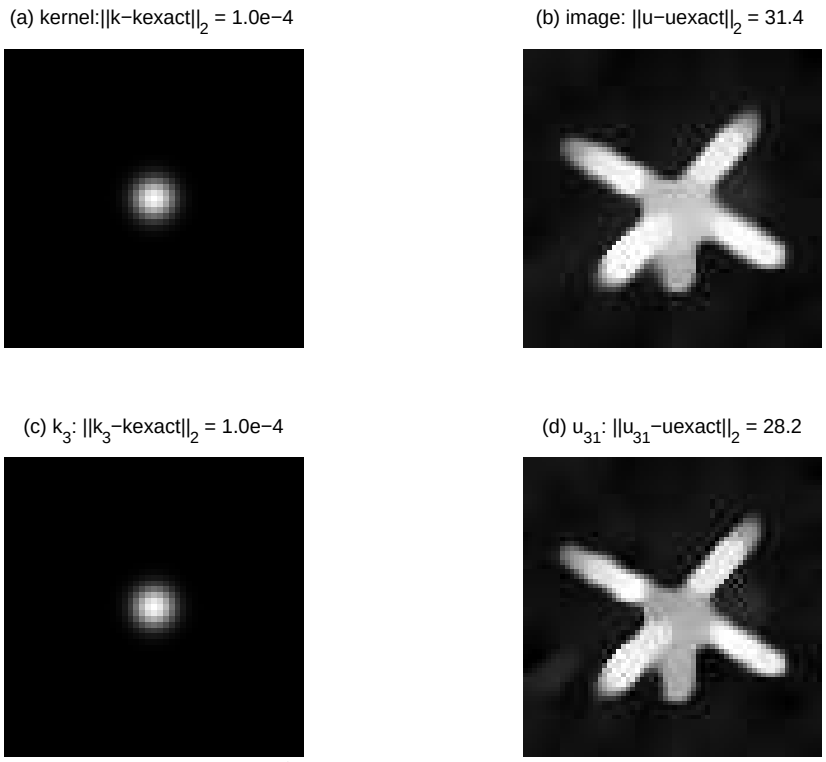
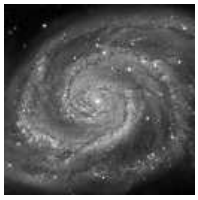
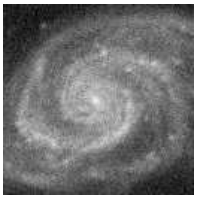


FIG. 5.9. *Comparison 2: We use the H^1 norm of the kernel. The result is comparable to that using the TV norm of kernel.*

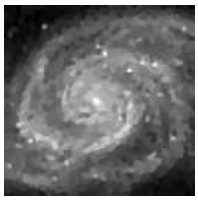
exact image



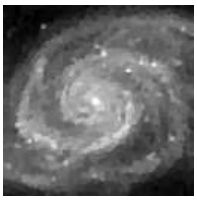
observed image, SNR = 9.0



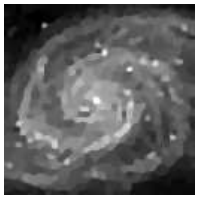
(a) u_4:deconvolution with BD



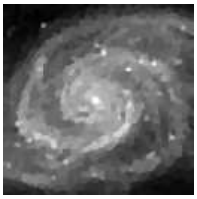
(b) u_{21}:blind deconvolution with BD



(c) u_{11}:using a better initial kernel



(d) u_{21}:using the H^1 norm of the kernel



α_1 = 0.05 for (a);α_1 = 0.01,α_2 = 3 for (b) and (d), α_1 =0.02 , α_2 = 6 for (c).
Both (c) and (d) are obtained from blind deconvolution with BD

FIG. 5.10. *Comparison 3*