Commun Nonlinear Sci Numer Simulat 20 (2015) 882–896

Contents lists available at ScienceDirect

Commun Nonlinear Sci Numer Simulat

journal homepage: www.elsevier.com/locate/cnsns

Inexact Bregman iteration for deconvolution of superimposed extended and point sources ☆

A. Benfenati, V. Ruggiero *

Dipartimento di Matematica e di Informatica, Università di Ferrara, Polo Scientifico Tecnologico, Blocco B, Via Saragat 1, I-44122 Ferrara, Italy

ARTICLE INFO

Article history:
Received 29 November 2013
Received in revised form 5 June 2014
Accepted 25 June 2014
Available online 12 July 2014

Keywords:
High contrast imaging
Poisson noise
Inexact Bregman iteration
Scaled Gradient Projection method

ABSTRACT

In this paper we consider the deconvolution of high contrast images consisting of very bright stars (point component) and smooth structures underlying the stars (diffuse component). A typical case is a weak diffuse jet line emission superimposed to a strong stellar continuum. In order to reconstruct the diffuse component, the original object can be regarded as the sum of these two components. When the position of the point sources is known, a regularization term can be introduced for the second component. An approximation of the original object can be obtained by solving a reduced variational problem whose unknowns are the intensities of the stars and the diffuse component. We analyze this problem when the detected image is corrupted by Poisson noise and Tikhonov-like regularization is used, giving conditions for the existence and the uniqueness of the solution. Furthermore, since only an overestimation of the regularization parameter is available, we propose to solve the variational problem by inexact Bregman iteration combined with a Scaled Gradient Projection method (SGP). Numerical simulations show that the images obtained with this approach enable us to reconstruct the original intensity distribution around the point source with satisfactory accuracy.

1. Introduction

Deblurring images corrupted by Poisson noise has recently been subject of an increasingly amount of works, due to the interest in biological and astronomical imaging (see, for example [1–7] and references therein).

In the framework of high resolution imaging from ground-based telescopes, the deconvolution and reconstruction process can be particularly challenging for high contrast images, i.e. when one is interested in detecting a weak signal close to a very bright star. This is the case, for instance, of the study of circumstellar structures around Young Stellar Objects (YSO). Previous studies on YSO simulations [8] shows that the reconstruction of the diffuse component by standard deconvolution methods is unsatisfactory in the region close to the emitting star. To overcome this difficulty, the basic idea is to consider the original object x as a superposition of two components, a point source x_P and an extended source x_E

$$x = x_P + x_E$$

☆ This research is supported by the project *Learning meets time: A new computational approach for learning in dynamic systems*, contract RBFR12M3AC, funded by the Italian Ministry of Education, University and Research *Futuro in Ricerca 2012* program and by the Italian Spinner2013 PhD project *High-complexity inverse problems in biomedical applications and social systems*.

* Corresponding author.
E-mail addresses: alessandro.benfenati@unife.it (A. Benfenati), valeria.ruggiero@unife.it (V. Ruggiero).

http://dx.doi.org/10.1016/j.cnsns.2014.06.045
1007-5704/

and to introduce different regularizations for the two components. A short review of the methods based on this approach for Gaussian data (least-squares formulation) is provided by [9]. In the framework of Poisson noise, the same idea was proposed by Lucy in [10] and, recently in [11,12]. In these last papers, the authors assume that the point source is a single star with known position and they propose to introduce a regularization term only for the extended source. Then, in the Bayesian framework, a maximum-likelihood solution can be obtained by minimizing with respect to x_P and x_E the following constrained nonlinear programming problem

$$\min_{x_P, x_E \geqslant 0} f_0(g; H(x_P + x_E) + b) + \beta f_1(x_E) \tag{1}$$

where $g \in \mathbb{R}^m$ is the detected image, $H \in \mathbb{R}^{m \times n}$ is the imaging matrix, modeling the distortion due to the acquisition system, and b is a given nonnegative constant background term. The objective function is the combination of two terms: the first one measures the discrepancy from the data g and, for Poisson noise, is expressed by the generalized Kullback–Leibler (KL) divergence (or Csiszár I-divergence) given by

$$f_0(g; Hx + b) = \sum_{i=1}^{m} g_i \log \frac{g_i}{(Hx + b)_i} + (Hx + b)_i - g_i \tag{2}$$

with the agreement $0 \log 0 = 0$. The second function f_1 is a regularization term, that enables to incorporate *a priori* information about the expected diffuse component of the solution into the reconstruction process. f_1 is a convex nonnegative function, weighted by the positive regularization parameter β, which has to be estimated. In [11], a Tikhonov-like regularization is used for the diffuse component and a scaled gradient method is proposed to obtain an approximate solution of (1). This approach depends on the choice of the regularization parameter β. In the case of regularization with Poisson data, this selection is basically an open problem. Recently, few discrepancy principles were proposed and discussed (see for example [7,2,13]); further, in [14,15], a constrained model for seminorm regularization is considered when some knowledge of level noise γ is available; this model allows to compute an estimate of β and a restored image by the minimization of a seminorm, under the constraint that the data fidelity is not greater than γ. But also for this formulation an accurate estimate of γ must be known.

A different approach is based on Bregman regularization. The Bregman approach allows to use an overestimated value of β and, in the framework of inverse problems, this procedure has the typical semi-convergence property of the iterative methods, described, for instance, in [16]. The procedure is proposed by Osher et al. in [17] to recover images corrupted by Gaussian noise and, later, by Brune et al. in [18] for dealing with Poisson data. Each iteration of this scheme requires to exactly compute the minimizer of a problem similar to (1), where f_1 is replaced by its Bregman distance at the current iterate. When the data fidelity function is the KL divergence, a closed formula for the minimizer may be unavailable and it could be extremely expensive to obtain exact solutions of these subproblems. On the other hand, in recent years the variational model (1) has been deeply investigated in order to design efficient iterative algorithms specifically tailored for different noise statistics and different regularization terms. Then in [19] an inexact version of the iterative procedure is proposed, where the *inexactness* in the inner subproblem solution of the Bregman iteration is controlled by a criterion that preserves its convergence and its features in image restoration problems.

In this work, we propose to apply the inexact Bregman iteration for the deconvolution of high contrast images with Poisson noise, where the regularization term depends only on the diffuse component and an overestimation of β is used; for the effectiveness of the procedure, the crucial issue becomes to devise an efficient algorithm for approximately solving the inner subproblems. For Tikhonov regularization, the iterates of the inexact algorithm can be efficiently obtained by the so-called Scaled Gradient Projection (SGP) method [20], but the approach can be extended to other regularization terms as, for example, the edge-preserving regularization [21]. For seminorm regularization, suitable inner solvers can be devised so that the convergence of the inexact Bregman approach is assured (for example [22,23]).

In order to show the effectiveness of the inexact Bregman iteration, we report the analysis of the numerical reconstructions obtained on a simulation of observations of a YSO jet [11]. We consider different star-jet intensity contrast cases by varying the magnitude of the star at each simulation. The obtained results appear very promising for the application to real cases.

The paper is organized as follows. In Section 2, we introduce the features of the problem (1), (2) and we derive a reduced formulation. We give the condition for the existence and the uniqueness of the solution for Tikhonov-like regularization term and we discuss the implementation of SGP method for its numerical solution. In Section 3, we recall the exact and inexact Bregman method ant its convergence properties and we discuss the application of the scheme for the problem (1) and (2). Finally in Section 4 we describe the results obtained by the inexact Bregman scheme on test-problems generated as described in [11] and kindly provided by the authors of this last paper.

2. Statement of the problem

A common problem in astrophysical imaging is to reconstruct two maps of a given region: one consisting of stars (point sources) and the other representing the smooth structures underlying the stars (extended sources). In this case, the object x to be reconstructed can be considered as the sum of two nonnegative components, corresponding to the point sources (x_P)

and to the extended sources (x_E). Following several recent papers [24,9,25], the two components can be reconstructed using different regularization terms. In the framework of Poisson noise, a similar approach is used in [10–12], with the additional assumption that the point component consists of some stars with known locations; since any unresolved star can be modeled as a delta function which is different from zero at a known position, the point component is zero everywhere except at the pixels corresponding to the stars. Then, for an object $x \in \mathbb{R}^n$, the point component can be expressed as

$$x_P = \sum_{j=1}^{q} \bar{x}_{P_j} \delta(v_j) \tag{3}$$

where $v_j, j = 1, 2, \ldots, q$ are the known positions of the q stars in the image array, while $\bar{x}_P = (\bar{x}_{P_1}, \ldots, \bar{x}_{P_q})^T \in \mathbb{R}^q$ are the positive intensities whose values have to be improved by the restoration method. The structure imposed on x_P implicitly works as a regularization term. On the other hand, in order to enforce the smoothness of x_E, it is convenient to introduce an explicit nonnegative and convex regularization term for the diffuse component, such as $f_1(x_E) = \frac{1}{2}\|Lx_E\|^2$, where $\|.\|$ denotes the l_2 norm and L is the identity or the first, second order finite difference operator.

As it is well-known [1], for images corrupted by Poisson noise, the discrepancy of the object x to be reconstructed from the detected data g is measured by the generalized Kullback–Leibler divergence (2), that in this case depends on $2n$ variables. When x_P is given by (3), we can rewrite the KL divergence so that it depends only on $q + n$ variables $\bar{x} \equiv \begin{pmatrix} \bar{x}_P \\ x_E \end{pmatrix}$:

$$f_0(g; \overline{H}\bar{x} + b) = \sum_{i=1}^{m} g_i \log \frac{g_i}{(\overline{H}\bar{x} + b)_i} + (\overline{H}\bar{x} + b)_i - g_i \tag{4}$$

where $\overline{H} = [\mathcal{H}\ H], \mathcal{H} = [h_{v_1}\ h_{v_2} \ldots, h_{v_q}]$ and h_j denotes the jth column of the matrix H. With respect to $\bar{x}$, the generalized KL is a nonnegative and convex function. Moreover, when the imaging matrix H is given in terms of a known space-invariant point spread function (PSF), normalized to unit volume, the features of H, such as

$$h_{ij} \geqslant 0 \quad H\mathbf{1} > 0 \quad H^T\mathbf{1} = \mathbf{1} \tag{5}$$

still apply for $\overline{H}$. Here, $\mathbf{1}$ denotes an array of suitable size with all entries equal to 1.

When $m = n$ and we assume periodic extension of the PSF, the matrix H satisfies the property

$$H\mathbf{1} = \mathbf{1} \tag{6}$$

As pointed out in [7], under the assumption (5) on $\overline{H}, f_0(g; \overline{H}\bar{x} + b)$ is a coercive function for $\bar{x} \geqslant 0$. For normalized PSF, the gradient and the Hessian of $f_0(g; \overline{H}\bar{x} + b)$ with respect to $\bar{x}$ assume the well known expressions

$$\begin{aligned} \nabla_{\bar{x}} f_0(g; \overline{H}\bar{x} + b) &= \mathbf{1} - \overline{H}^T \frac{g}{\overline{H}\bar{x} + b} \\ \nabla^2_{\bar{x}} f_0(g; \overline{H}\bar{x} + b) &= \overline{H}^T \operatorname{diag}\left(\frac{g}{(\overline{H}\bar{x} + b)^2}\right) \overline{H} \end{aligned} \tag{7}$$

where the operations between two vectors are intended componentwise.

Then, in the framework of Bayesian approach, a maximum a posteriori estimate of the object can be obtained by solving the following minimization problem

$$\min_{\bar{x} \geqslant 0} f_\beta(\bar{x}) \equiv f_0(g; \overline{H}\bar{x} + b) + \beta f_1(x_E) \tag{8}$$

with $f_1(x_E) = \frac{1}{2}\|Lx_E\|^2$. Since $f_\beta(\bar{x})$ is nonnegative, coercive and convex, there exists a solution for problem (8). Under suitable assumptions, this solution is unique. Indeed, we can prove that $f_\beta(\bar{x})$ is strictly convex, by showing that the intersection between the null spaces of the Hessian matrices of f_0 and f_1 is trivial.

Theorem 1. *Let $g \in \mathbb{R}^m$ be positive. Assume that the matrix H satisfies (5) and the column rank of $\mathcal{H} \in \mathbb{R}^{m \times q}$ is equal to q; then we have:*

a. *when the null space of L is trivial, the solution of the problem (8) is unique for any $\beta > 0$;*
b. *when the null space of L is $\{\alpha\mathbf{1}, \alpha \in \mathbb{R}\}$ and $H\mathbf{1} \notin \operatorname{span}(\mathcal{H})$, the solution of the problem (8) is unique for any $\beta > 0$.*

Proof. In both cases, we prove that the positive semidefinite matrix $\nabla^2_{\bar{x}} f_\beta(\bar{x})$, given by

$$\nabla^2_{\bar{x}} f_\beta(\bar{x}) = \overline{H}^T \operatorname{diag}\left(\frac{g}{(\overline{H}\bar{x} + b)^2}\right) \overline{H} + \beta \begin{pmatrix} 0 & 0 \\ 0 & L^T L \end{pmatrix} \tag{9}$$

is positive definite for $\beta > 0$ and consequently f_β is strictly convex for $\bar{x} \geqslant 0$. Indeed we can show that the intersection of the null spaces of $\nabla^2_{\bar{x}} f_0$ and $\beta \nabla^2_{\bar{x}} f_1$ is trivial.

a. Let $\mathcal{N}(L)=\{0\}$; since $\mathcal{N}(\nabla^2_{\bar{x}} f_0)=\mathcal{N}(\overline{H})$ and $\mathcal{N}(\nabla^2_{\bar{x}} f_1)=\left\{\begin{pmatrix}\bar{x}_P\\0\end{pmatrix}, 0\in\mathbb{R}^n\right\}$, we have to prove that for any $\bar{x}_p\neq 0$

$$\overline{H}\begin{pmatrix}\bar{x}_P\\0\end{pmatrix}\neq 0$$

Indeed, for the linear independence of the columns of $\mathcal{H}, \mathcal{H}\bar{x}_P\neq 0$. Then $\nabla^2_{\bar{x}} f_\beta(\bar{x})$ is positive definite for $\bar{x}\geqslant 0$.

b. When $\mathcal{N}(L)=\{\alpha\mathbf{1},\alpha\in\mathbb{R}\}, \mathcal{N}(\nabla^2_{\bar{x}} f_1)=\left\{\begin{pmatrix}\bar{x}_P\\\alpha\mathbf{1}\end{pmatrix}, \mathbf{1}\in\mathbb{R}^n\right\}$. Then, we have to prove that

$$\overline{H}\begin{pmatrix}\bar{x}_P\\\beta\alpha\mathbf{1}\end{pmatrix}\neq 0$$

Indeed, by contradiction, if we assume

$$\mathcal{H}\bar{x}_P+\beta\alpha H\mathbf{1}=0$$

there exists a vector $-\frac{1}{\beta\alpha}\bar{x}_P$ such that its components are the coordinates of $H\mathbf{1}$ with respect to the columns of $\mathcal{H}$; then $H\mathbf{1}\in\text{span}(\mathcal{H})$, in contradiction with the hypothesis. □

Remark. Part b. of the previous theorem holds even when the regularization term is an hypersurface (HS) potential [21] or a Markov random field regularization [26]. Indeed, also for these regularization terms, $\mathcal{N}(\nabla^2_{\bar{x}} f_1)=\left\{\begin{pmatrix}\bar{x}_P\\\alpha\mathbf{1}\end{pmatrix}\right\}$ [27].

The assumption $g>0$ is not very restrictive, since g includes also the background emission which, in the real applications, is different from zero. In any case, this restriction can be overcome by setting the zero values in the data to very small values (without side effects).

Moreover the assumption $\mathcal{N}(L)=\{\alpha\mathbf{1},\alpha\in\mathbb{R}\}$ holds for the first or second order finite difference operator L when standard boundary conditions are imposed on the image, as periodic or reflexive boundary conditions [27]; for zero boundary conditions or $L=I$, the null space of L is trivial.

When (6) holds, the assumption $H\mathbf{1}\notin\text{span}(\mathcal{H})$ is equivalent to require that $\{\alpha\mathbf{1},\alpha\in\mathbb{R}\}$ is not a subspace of $\text{span}(\mathcal{H})$; since q is small, it is easy to verify if there exists the unique solution of $\mathcal{H}x=\mathbf{1}$; for example, when $q=1$ and h_{v_1} is not a constant vector, $H\mathbf{1}\notin\text{span}(\mathcal{H})$. Moreover, when H is a nonsingular matrix, there does not exist x such that $\mathcal{H}x=\mathbf{1}$; indeed, in this case, by (6), there would be contradiction with the linear independence of the columns of H: $0=\mathbf{1}-\mathcal{H}x=H\mathbf{1}-\mathcal{H}x=\sum_{j\neq v_1,\dots v_q}h_j+\sum_{j=v_1,\dots v_q}h_j(1-x_j)$.

In order to obtain the numerical solution of the minimization problem (8), there exists a number of algorithms specifically tailored for the KL divergence combined with a differentiable regularization term (for a survey see [1]). In the class of first order methods requiring only function and gradient evaluations, the Scaled Gradient Projection (SGP) method is very efficient for image restoration from Poisson data [3,20]. This projected gradient-type method is based on a modification of the gradient direction by a suitable positive definite and diagonal scaling matrix D_k and a strategy of positive steplength selection $\alpha_k\in[\alpha_{min},\alpha_{max}]$ through an adaptive alternation of the Barzilai–Borwein rules [28,29]. In particular, at the kth iteration, the descent direction is computed as follows:

$$d_k=\max\{\bar{x}^{(k)}-\alpha_k D_k\nabla_{\bar{x}} f_\beta(\bar{x}^{(k)}),0\}-\bar{x}^{(k)} \tag{10}$$

where the componentwise operator *max* implements the projection on the nonnegative orthant. A monotone or nonmonotone line-search procedure assures the convergence to a stationary point. The main computational cost of each SGP iteration depends essentially on two matrix-vector products, required in the computation of the objective function and its gradient. The updating of the scaling matrix and the steplength parameter, as well as the projection onto the nonnegative orthant and the line-search procedure along the projected direction require only linear operations. In this case, the matrix-vector products $\overline{H}\bar{x}$ and $\overline{H}^T z$ require respectively:

- the matrix-vector product Hx_E and q sums of a scalar by a vector (SAXPY in BLAS1 [30,31]):

$$\overline{H}\bar{x}=\sum_{j=1}^{q}\bar{x}_{P_j}h_{v_j}+Hx_E \tag{11}$$

- the matrix-vector product $H^T z$ and q additional scalar-vector products (SDOT in BLAS1 [30,31]):

$$\overline{H}^T z=\begin{pmatrix}h_{v_1}^T z\\\vdots\\h_{v_q}^T z\\H^T z\end{pmatrix} \tag{12}$$

For space-invariant PSF, Hx_E and H^Tz can be obtained by standard FFT algorithms or other fast trigonometric transforms (in dependence on the boundary conditions imposed on the image) while we can determine $h_{v_j}\ j = 1,\ldots,q$ by an initial matrix-vector product He_j where e_j is the jth element of the canonical basis of $\mathbb{R}^n$ (by FFT or other fast transforms). Thus the computational complexity depends on the number of the matrix-vector products Hx_E, H^Tz. These features make SGP well-suited to be implemented also on graphic processing units (GPUs) or on multiprocessor systems [32].

A crucial point for the effectiveness of the method is the updating strategy of the scaling matrix D_k; we can decompose the gradient of f_β as the sum of a positive and a non positive part, i.e. $\nabla_x f_\beta(\bar{x}) = V(\bar{x}) - U(\bar{x})$, with $V(\bar{x}) > 0, U(\bar{x}) \geqslant 0$; here, $V(\bar{x}) = \mathbf{1} + \beta\begin{pmatrix} 0 \\ V_1(x_E) \end{pmatrix}$, with $V_1(x_E)$ equal to the positive part of the gradient of f_1 (for example $V_1(x_E) = x_E$ for $f_1(x_E) = \frac{1}{2}\|x_E\|^2$). Then D_k is chosen as a diagonal matrix with entries $d_i^{(k)}$ defined as follows:

$$d_i^{(k)} = \min\left\{\overline{L}, \max\left\{\underline{L}, \frac{\bar{x}_i^{(k)}}{V(\bar{x}^{(k)})_i}\right\}\right\} \tag{13}$$

where $\overline{L}$ and $\underline{L}$ are appropriate thresholds; in particular for the problem (8), these values can be chosen as follows:

$$\begin{aligned} \underline{L} &= \min\{z_i | z_i > 0\} \\ \overline{L} &= \max_{i=v_1,\ldots,v_q}\{g_i\} \end{aligned} \tag{14}$$

where $z = \tilde{x}H^T\mathrm{diag}\left(\frac{H^Tg}{H\tilde{x}+b}\right)$, i.e. z is obtained by an iteration of Lucy Richardson method, starting from $\tilde{x} = \left(\sum_{i=1}^m g_i - b\right)/m$.

We observe that the MC-RL method used in [11] for $f_1(x_E) = \frac{1}{2}\|x_E\|^2$ is equivalent to SGP method for $\alpha_k = 1$ and without line-search. The strategy for updating the steplength and the line-search procedure can be regarded as an acceleration technique for MC-RL.

3. Exact and inexact Bregman procedure for deblurring of Poisson data

In this section we resume the features of the exact and inexact Bregman procedure in the image restoration framework.

First of all, we recall that the Bregman distance of a proper convex function $F : \mathbb{R}^n \to \mathbb{R}$ between x and y is defined as follows

$$D^pF(x,y) = F(x) - F(y) - \langle p, x - y\rangle \tag{15}$$

where p is a subgradient of F at y and $\langle\cdot,\cdot\rangle$ denotes the canonical inner product of two vectors of $\mathbb{R}^n$.

3.1. The exact scheme

In a Bayesian framework, the problem of recovering an image from a vector g of detected noisy data is related to the solution of the following variational problem

$$\min f_0(g; Hx + b) + \beta f_1(x) \tag{16}$$

where f_0 is a measure of the discrepancy between $Hx + b$ and g and f_1 is a regularization term. We start the description of the Bregman iteration by considering the case where the given data are noise-free. If g^* denotes the noise-free data, we have to solve the following problem

$$\min f_0(g^*; Hx + b) + \beta f_1(x) \tag{17}$$

Here $f_0(g^*; Hx + b)$ can be regarded as a nonnegative convex penalty function for the set $\overline{S} = \{x \in \mathbb{R}^n; Hx + b = g^*, x \geqslant 0\}$, i.e. $f_0(g^*; Hx + b) \geqslant 0$ for any x and $f_0(g^*; Hx + b) = 0 \iff x \in \overline{S}$.

The aim of the Bregman method [33] is to determine a solution of the problem

$$\min_x f_1(x) \text{ subject to } f_0(g^*; Hx + b) = 0 \tag{18}$$

by solving a sequence of unconstrained subproblems where $\frac{1}{\beta}$ plays the role of a penalty parameter and $f_1(x)$ is replaced by its Bregman distance at the current iterate. More specifically, starting from a nonnegative vector $x^{(0)} \in \mathbb{R}^n$ such that $p^{(0)} \equiv 0 \in \partial f_1(x^{(0)})$, the Bregman procedure generates a sequence of iterates as follows

$$x^{(k+1)} = \mathrm{argmin}_x Q_k(x, p^{(k)}) \tag{19}$$

where

$$Q_k(x, p^{(k)}) = D^{p^{(k)}} f_1(x, x^{(k)}) + \frac{1}{\beta} f_0(g^*; Hx + b) \tag{20}$$

with $p^{(k)} \in \partial f_1(x^{(k)})$. Under suitable assumptions [17], it is proved that the sequence $\{x^{(k)}\}$ converges to a minimizer of $f_0(g^*; Hx+b)$, i.e. to a point x^* such that $f_0(g^*; Hx^*+b)=0$ or $Hx^*+b=g^*, x^* \geqslant 0$, while $\{D^{p^{(k)}} f_1(x^*, x^{(k)})\}$ is decreasing [17,34].

Furthermore, if some iterate $x^{(\bar{k})}$ of the Bregman procedure satisfies $f_0(g^*; Hx^{(\bar{k})}+b)=0$, then $x^{(\bar{k})}$ is a solution of (18) [34,19].

The decreasing behavior of the sequence $\{f_0(g^*; Hx^{(k)}+b)\}$ and the consequent convergence results are based on the following inequality that exploits the nonnegativity property of the Bregman distance of f_1:

$$\beta\left(D^{p^{(k)}} f_1(x, x^{(k)}) + D^{p^{(k-1)}} f_1(x^{(k)}, x^{(k-1)})\right) + f_0(g^*; Hx^{(k)}+b) \leqslant f_0(g^*; Hx+b) + \beta D^{p^{(k-1)}} f_1(x, x^{(k-1)}) \tag{21}$$

for any $x \geqslant 0$. For details about the assumptions and the proofs, see [17,34,19].

Now we observe that in real applications we have to solve problem (16) instead of (17), where g is a set of noisy data. Then, for the convergence properties of the Bregman method, a limit point of $\{x^{(k)}\}$ is a solution $\bar{x}$ of $f_0(g; Hx+b)=0$, while we are interested to a solution x^* of $f_0(g^*; Hx+b)=0$, where g^* is the noise-free data. Nevertheless we can observe that, in this case, the Bregman iteration has the typical *semi-convergence* behavior of the iterative methods for the solution of inverse problems, as described for example in [16]; the sequence $\{x^{(k)}\}$ first approaches the required solution x^* and then it goes away, converging toward $\bar{x}$. Indeed, following [17], if an estimate γ for the noise level is known, that is $f_0(g; Hx^*+b) \leqslant \gamma$, from (21) with $x=x^*$, it follows that, while $f_0(g; Hx^{(k)}+b) \geqslant \gamma$, the Bregman distance of the iterates from the object x^* decreases:

$$D^{p^{(k)}} f_1(x^*, x^{(k)}) \leqslant D^{p^{(k-1)}} f_1(x^*, x^{(k-1)}) \tag{22}$$

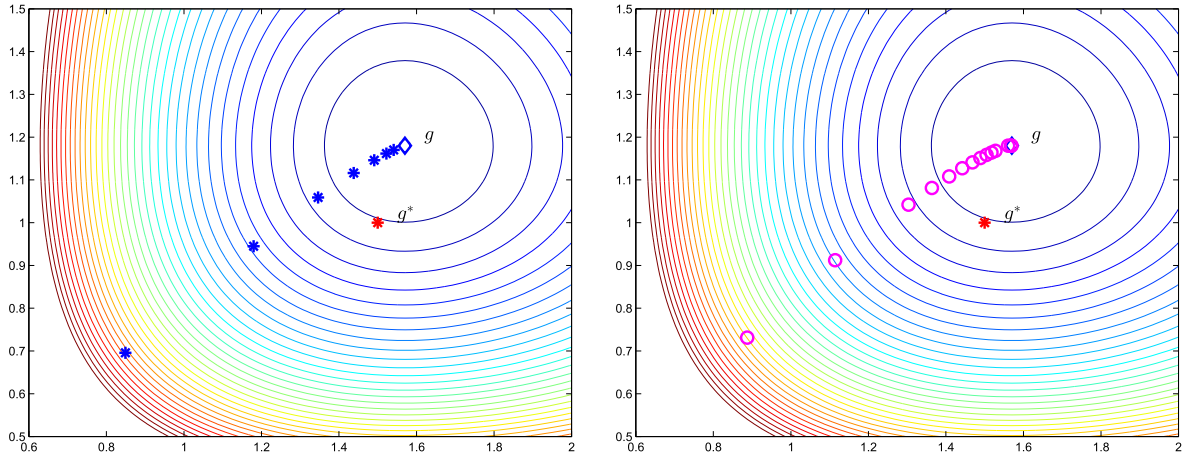
Consequently, a stopping criterion for the iterative procedure is to terminate at the iteration k^* such that

$$k^* = \max\left\{k | f_0(g; Hx^{(k)}+b) \geqslant \gamma\right\} \tag{23}$$

In the case of Gaussian noise, the Morozov discrepancy principle can be a reasonable stopping criterion. In the case of Poisson noise, it makes sense to stop the Bregman iteration if the generalized KL divergence of $Hx^{(k)}+b$ from the detected data g reaches the noise level. For an estimate of this noise level, see the discrepancy criterion in [7].

Further the Bregman approach, very similar to a penalization approach, enables us to use an overestimated value of β, avoiding the problem of an accurate selection of this parameter. Numerical results show that for a raw overestimation of β few iterations are required to obtain a satisfying solution and, consequently, few instances of subproblems (19), very similar to (16), have to be solved.

In order to clarify the regularization effect of the Bregman iteration, we introduce a simple example. Here H is the identity matrix of order 2, $b=0, f_0(g;x)$ is the KL divergence, $f_1 = \frac{1}{2}\|x\|^2$ and $g = (1.57,\ 1.18)^T$. We assume that g is a perturbation of $g^* = (1.5,\ 1)^T$ so that the Bregman iteration generates a sequence of iterates which converges to g tracing a trajectory that *goes* near g^*, as shown in Fig. 1(a). For this simple example, the inner subproblems can be exactly solved. In Table 1, we report the value of f_0 at the current iterate $x^{(k)}$, the Bregman distance $D^{p^{(k)}} f_1(g^*, x^{(k)}) = \frac{1}{2}\|g^* - x^{(k)}\|^2$ and the relative error $\rho^{(k)} = \frac{\|x^{(k)} - g^*\|}{\|g^*\|}$ with respect to g^*, that is the solution for data without perturbations. Following the above stopping criterion (23), $k=3$ should be the final iterate since $f_0(g; g^*) = 0.016916$. Nevertheless, we observe that the next iterate provides a vector closest to the solution of the unperturbed problem. Indeed the decreasing of the sequence $D^{p^{(k)}} f_1(g^*, x^{(k)})$ is assured



(a) exact scheme (b) inexact scheme

Fig. 1. Behavior of the exact and inexact Bregman scheme for a simple example.

Table 1
Exact and inexact Bregman scheme for a simple example, where $f_0(g;x)$ is the KL divergence, $f_1(x) = \frac{1}{2}\|x\|^2, g = (1.57,\ 1.18)^T, g^* = (1.5,\ 1)^T$.

k	$f_0(g;x^{(k)})$	$D^{p^{(k)}} f_1(g^*, x^{(k)})$	$\rho^{(k)}$
Exact scheme			
1	0.383195	0.258	0.399
2	0.085509	0.0528	0.180
3	0.024324	0.0136	0.0914
4	0.007676	0.0087	0.0731
5	0.002561	0.0107	0.0811
Inexact scheme			
1	3.029703	1.166	0.847
2	0.328989	0.224	0.371
3	0.118477	0.0782	0.219
4	0.034295	0.0202	0.111
5	0.019239	0.0126	0.0879
6	0.011214	0.0100	0.0786
7	0.006711	0.0098	0.0775
8	0.004093	0.0104	0.0802

until the data fidelity function is greater than the noise level, but it could happen that the sequence may continue to decrease for some iterations. Then, it is convenient to take into account also some iterates corresponding to values of f_0 slightly smaller than the noise level.

3.2. The inexact iterative procedure

When a closed formula for the solution of inner subproblem (19) is unavailable, we can obtain the new iterate of the Bregman procedure by means of an iterative solver. Some iterative solvers provide useful information about the gradient or an approximate subgradient of the function f_1 that can be exploited for the next iteration. Therefore, if we find a suitable stopping criterion for the inner solver that preserves the convergence property of the Bregman procedure, it is possible to reduce the total number of inner iterations, computing just an approximate minimizer of the inner subproblem.

In [19] an inexact Bregman iteration has been proposed in either cases of f_1 differentiable and non-differentiable. Since for the deconvolution of high contrast images we consider Tikhonov-like regularization terms, we restrict the discussion to the case of f_1 differentiable. Nevertheless, we remark that the inexact scheme enables to face also seminorm regularization terms through the introduction of the notions of ϵ-subgradient and *inexact* Bregman distance.

In order to resume the main features of the inexact Bregman procedure, we start considering problem (17), i.e. the problem with noise-free data g^*. We assume that the kth inner subproblem (19) is faced by an iterative algorithm. This inner solver generates sequences $\{x_l^{(k)}\}$ and $\{q_l^{(k)}\}, q_l^{(k)} \in \partial f_0(g^*, Hx_l^{(k)} + b)$, such that the vector $\eta_l^{(k)} = \frac{1}{\beta} q_l^{(k)} + \nabla f_1(x_l^{(k)}) - p^{(k)}$ is a subgradient of $Q_k(x, p^{(k)})$ at $x_l^{(k)}$ and we have

$$\lim_{l\to\infty} x_l^{(k)} = \bar{x}^{(k)} \quad \lim_{l\to\infty} q_l^{(k)} = \bar{q}^{(k)}$$

where $\bar{x}^{(k)}$ is a minimizer of the kth subproblem $Q_k(x, p^{(k)})$ and $\bar{q}^{(k)}$ is the related subgradient of f_0 at $\bar{x}^{(k)}$. As a consequence, given a tolerance $\mu_{k+1} > 0$, there exists an index $\bar{l}$ such that $\|\eta_l^{(k)}\| \leqslant \mu_{k+1}$; then, we set as approximate solution of the kth subproblem $x^{(k+1)} = x_l^{(k)}$, with $p^{(k+1)} = \nabla f_1(x^{(k+1)})$ and $q^{(k+1)} = q_l^{(k)}$.

When the above stopping criterion for the inner solver is coupled with the assumption $\sum_k \mu_k < \infty$, the convergence properties of the inexact scheme are preserved. For sake of completeness, we report Corollary 1 in [19] where one can find the details and the algorithm.

Proposition 1. *Let $f_0(g^*; Hx + b)$ and $f_1(x)$ be nonnegative, proper, closed and convex functions, with f_1 differentiable,* $\mathrm{dom}(f_0 \circ H) \subset \mathrm{dom}(f_1)$ *and the relative interiors of $f_0 \circ H$ and f_1 have at least a point in common. We assume that, for any k, there exists a minimizer of the subproblem (19) and that x^* is a minimizer of $f_0(g^*; Hx + b)$ such that $f_1(x^*) < \infty$. If for any $k \geqslant 0$ the inner solver determines $x^{(k+1)}, q^{(k+1)} \in \partial f_0(g^*; Hx^{(k+1)} + b)$ and $p^{(k+1)} = \nabla f_1(x^{(k+1)})$ so that the following condition on $\eta^{(k+1)} = \frac{1}{\beta} q^{(k+1)} + p^{(k+1)} - p^{(k)}$ holds*

$$\|\eta^{(k+1)}\| \leqslant \mu_{k+1} \tag{24}$$

with $\sum_{i=1}^{\infty} \mu_i < \infty$, then we have that

$$D^{p^{(k)}} f_1(x^*, x^{(k)}) \leqslant D^{p^{(k-1)}} f_1(x^*, x^{(k-1)}) + \langle \eta^{(k)}, x^{(k)} - x^* \rangle \tag{25}$$

and

$$f_0(g^*; Hx^{(k)} + b) \leqslant f_0(g^*; Hx^* + b) + \frac{\beta}{k}\left(f_1(x^*) - f_1(x^{(0)}) + \sum_{i=1}^{k} \langle \eta^{(i)}, x^{(i)} - x^* \rangle \right) \tag{26}$$

Moreover, if the level subsets of $f_0 \circ H$ are bounded, a limit point of the sequence $\{x^{(k)}\}$ is a minimizer of $f_0(g^*; Hx + b)$; if x^* is the unique minimizer of $f_0(g^*; Hx + b)$, then $x^{(k)} \to x^*$ as $k \to \infty$.

Remark. From the numerical point of view, an easily implementable choice for the tolerance in the criterion (24) is

$$\|\eta^{(k+1)}\| \leqslant \frac{c}{(k+1)^{\chi}} \quad k \geqslant 0 \tag{27}$$

where c and χ are positive constants, with $\chi > 1$. In this way, at the first iteration the tolerance is equal to the parameter c and, in the subsequent iterations, the stopping rule is gradually more severe; the parameter χ controls the increase of the inner accuracy. A practical rule to choose the value of c is to use a standard stopping criterion (related for example to the objective function) with a moderate tolerance in the inner solver at the first outer iteration; then set $c = \|\eta^{(1)}\|$; in this way, the only input parameter is χ. In the Section 4, we implement this rule.

When the noise-free data g^* is replaced by the detected noisy data g and we are interested to a solution x^* of $f_0(g^*; Hx + b) = 0$, rather than $f_0(g; Hx + b) = 0$, the previous remarks can be repeated again for the inexact scheme.

An inequality similar to (21) holds for any x:

$$-\langle \eta^{(k)}, x - x^{(k)} \rangle + D^{p^{(k)}} f_1(x, x^{(k)}) + D^{p^{(k-1)}} f_1(x^{(k)}, x^{(k-1)}) + \frac{1}{\beta} f_0(g; Hx^{(k)} + b) \leqslant \frac{1}{\beta} f_0(g; Hx + b) + D^{p^{(k-1)}} f_1(x, x^{(k-1)})$$

When an estimate of a noise level γ is known (that is $f_0(g; Hx^* + b) \leqslant \gamma$), from the previous inequality with $x = x^*$, we obtain

$$-\mu_k D + D^{p^{(k-1)}} f_1(x^{(k)}, x^{(k-1)}) + \frac{1}{\beta} f_0(g; Hx^{(k)} + b) \leqslant \frac{\gamma}{\beta} + D^{p^{(k-1)}} f_1(x^*, x^{(k-1)})$$

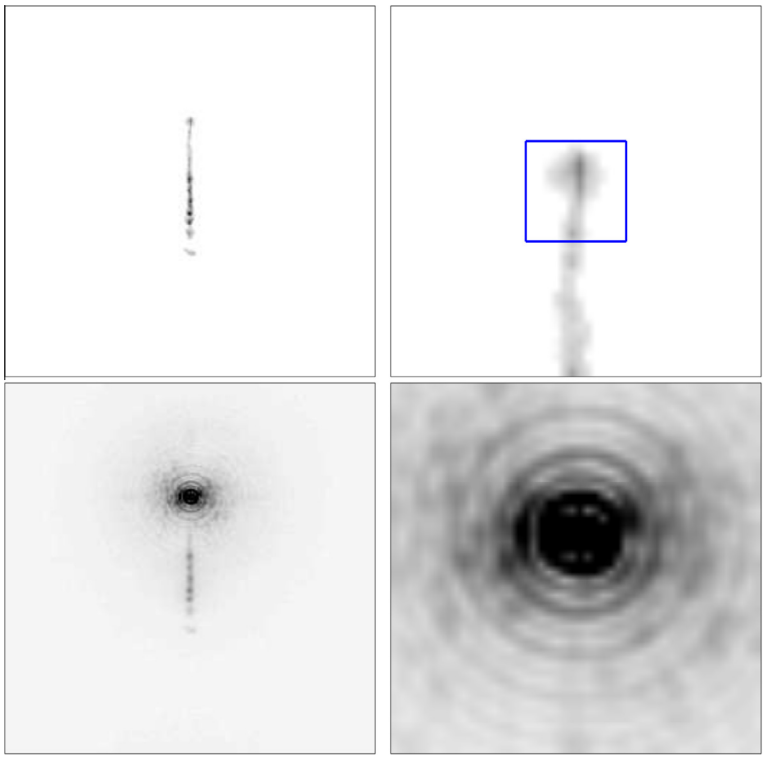


Fig. 2. *Upper panels:* original diffuse object x_E^* and a detailed view of the region around the star; a square box surrounding the star is emphasized. *Lower panels:* blurred image of test-problem *Star8* and a detailed view of the region around the star. All the images are in reverse gray sqrt-scale.

Table 2
Numerical results obtained by applying the discrepancy criterion; $\overline{\beta}$ is the solution of the discrepancy equation. The meaning of the parameters appearing in the table is defined within the text.

$f_1(x_E)=\frac{1}{2}\|Lx_E\|^2$	$\overline{\beta}$	*totit (T)*	ρ	ρ_w	ρ_P
Star8					
$L=I$	3.1×10^{-9}	5462(10)	0.24	0.72	5.1×10^{-4}
$L=L_1$	1.4×10^{-9}	5600(13)	0.23	0.58	2.4×10^{-4}
$L=L_2$	5.8×10^{-10}	4428(8)	0.23	0.52	1.5×10^{-4}
Star9					
$L=I$	3.0×10^{-9}	6286(11)	0.23	0.69	1.1×10^{-3}
$L=L_1$	1.4×10^{-9}	5828(10)	0.23	0.53	5.4×10^{-4}
$L=L_2$	5.3×10^{-10}	4064(9)	0.22	0.47	3.3×10^{-4}
Star10					
$L=I$	2.9×10^{-9}	6447(10)	0.23	0.66	3.1×10^{-3}
$L=L_1$	1.4×10^{-9}	5792(20)	0.22	0.50	1.7×10^{-3}
$L=L_2$	5.0×10^{-10}	3882(9)	0.22	0.45	1.2×10^{-3}
Star11					
$L=I$	2.7×10^{-9}	5458(9)	0.23	0.64	7.4×10^{-3}
$L=L_1$	1.3×10^{-9}	5664(8)	0.22	0.48	4.5×10^{-3}
$L=L_2$	4.5×10^{-10}	4228(20)	0.22	0.44	3.6×10^{-3}

where D is the diameter of the bounded level set $\{x|f_0(g;Hx+b)\leqslant f_0(g;Hx^{(0)}+b)\}$. Then a sufficient condition to assure a decreasing behavior for the Bregman distance $D^{p^{(k-1)}}f_1(x^*,x^{(k-1)})$ is that, while $f_0(g;Hx^{(k)}+b)\geqslant\gamma$, the following inequality holds:

$$\frac{1}{\beta}(f_0(g;Hx^{(k)}+b)-\gamma)+D^{p^{(k-1)}}f_1(x^{(k)},x^{(k-1)})=Q_{k-1}(x^{(k)},p^{(k-1)})-\frac{\gamma}{\beta}\geqslant\mu_k D$$

Then we observe a decreasing behavior of the sequence $\{D^{p^{(k)}}f_1(x^*,x^{(k)})\}$ until the objective function is greater than $\frac{\gamma}{\beta}$ within a tolerance dependent of the *inexactness* level required to the inner minimizer. As for the exact Bregman method, a discrepancy criterion can provide a reasonable stopping criterion. Nevertheless, the numerical experience shows that the term $-\mu_k D$ is a pessimistic lower bound of $\langle\eta^{(k)},x^*-x^{(k)}\rangle$; indeed, since in few iterations we observe a semi-convergence behavior, a practical criterion is to visually control the obtained approximation at any iteration.

For the simple example considered in the previous subsection, the results reported in Table 1 show that $D^{p^{(k)}}f_1(g^*,x^{(k)})$ is a decreasing sequence until $f_0(g;x^{(k)})>0.016916$ and even later. Fig. 1 shows the trajectory of the inexact scheme; since the number of outer iterations before reaching the noise level increases, the points on this trajectory are more numerous. The

Table 3
Numerical results obtained by applying SGP method to problem (8); $\tilde{\beta}$ corresponds to the best relative reconstruction error ρ for the diffuse component obtained by several trials with different values of β. The meaning of the parameters appearing in the table is defined within the text.

$f_1(x_E)=\frac{1}{2}\|Lx_E\|^2$	$\tilde{\beta}$	$\tilde{k}$	ρ	ρ_w	ρ_P	$f_0\left(g;\overline{H}\overline{x}^{(\tilde{k})}+b\right)$
Star8, $tol_{SGP}=10^{-8}, M=5$						
$L=I$	1.2×10^{-10}	2433	0.1583	0.4342	9.0×10^{-5}	30879.4
$L=L_1$	4.0×10^{-11}	2492	0.1585	0.3825	2.6×10^{-4}	30882.3
$L=L_2$	1.0×10^{-11}	3493	0.1578	0.3149	2.9×10^{-4}	30874.0
Star9, $tol_{SGP}=10^{-8}, M=5$						
$L=I$	1.6×10^{-10}	2942	0.1597	0.4632	5.6×10^{-5}	31024.6
$L=L_1$	5.0×10^{-11}	2594	0.1548	0.3072	4.4×10^{-4}	31032.4
$L=L_2$	1.0×10^{-11}	3172	0.1558	0.3399	7.9×10^{-4}	31020.4
Star10, $tol_{SGP}=0.5\times10^{-8}, M=5$						
$L=I$	9.0×10^{-11}	4423	0.1598	0.4264	3.9×10^{-4}	31124.4
$L=L_1$	5.0×10^{-11}	3253	0.1578	0.2760	3.9×10^{-4}	31143.2
$L=L_2$	8.0×10^{-12}	4107	0.1548	0.3318	1.4×10^{-3}	31124.9
Star11, $tol_{SGP}=0.5\times10^{-8}, M=5$						
$L=I$	8.0×10^{-11}	5096	0.1578	0.4179	1.5×10^{-3}	31306.7
$L=L_1$	5.0×10^{-11}	3020	0.1528	0.2603	7.2×10^{-4}	31331.1
$L=L_2$	7.0×10^{-12}	4019	0.1525	0.3279	3.2×10^{-3}	31312.9

Table 4
Numerical results obtained by the inexact Bregman iteration.

$f_1(x_E)=\frac{1}{2}\|Lx_E\|^2$	β	$\bar{k}$ (*totit*)	$\rho^{(\bar{k})}$	$\rho_w^{(\bar{k})}$	$\rho_P^{(\bar{k})}$	$f_0(g;\overline{H}\bar{x}^{(\bar{k})}+b)$
Star8						
$L=I$	3.1×10^{-9}	43 (3603)	0.1409	0.3926	3.469×10^{-4}	30856.4
$L=I$	1.1×10^{-9}	20 (4196)	0.1382	0.3318	3.510×10^{-4}	30854.0
$L=L_1$	1.4×10^{-9}	34 (5133)	0.1579	0.3678	1.223×10^{-4}	30854.2
$L=L_1$	0.5×10^{-9}	18 (4809)	0.1528	0.3495	3.212×10^{-4}	30852.0
$L=L_2$	5.8×10^{-10}	48 (6689)	0.1590	0.3344	1.224×10^{-4}	30854.9
$L=L_2$	3.8×10^{-10}	36 (5864)	0.1566	0.3072	1.665×10^{-4}	30853.8
Star9						
$L=I$	3×10^{-9}	26 (4036)	0.1473	0.3712	3.045×10^{-3}	30999.9
$L=I$	2×10^{-9}	20 (3590)	0.1410	0.3292	6.126×10^{-4}	30999.8
$L=L_1$	1.4×10^{-9}	41 (5755)	0.1472	0.2477	4.945×10^{-4}	30996.8
$L=L_1$	0.4×10^{-9}	16 (5287)	0.1535	0.3368	4.600×10^{-4}	30994.6
$L=L_2$	5.3×10^{-10}	94 (10113)	0.1478	0.2265	5.166×10^{-4}	30994.2
$L=L_2$	3.3×10^{-10}	63 (10071)	0.1489	0.2547	5.479×10^{-4}	30994.1
Star10						
$L=I$	2.9×10^{-9}	20 (4187)	0.1579	0.4139	3.140×10^{-4}	31108.6
$L=I$	1.9×10^{-9}	17 (3769)	0.1546	0.4289	5.737×10^{-4}	31108.1
$L=L_1$	1.4×10^{-9}	39 (6682)	0.1524	0.2791	5.297×10^{-4}	31107.3
$L=L_1$	0.4×10^{-9}	15 (4932)	0.1547	0.3306	7.536×10^{-4}	31105.4
$L=L_2$	5.0×10^{-10}	77 (6704)	0.1485	0.3056	1.448×10^{-3}	31105.8
$L=L_2$	4.0×10^{-10}	69 (9764)	0.1512	0.2684	9.411×10^{-4}	31105.1
Star11						
$L=I$	2.7×10^{-9}	26 (5673)	0.1512	0.3595	1.131×10^{-3}	31296.4
$L=I$	1.7×10^{-9}	20 (4621)	0.1442	0.3057	5.040×10^{-4}	31296.9
$L=L_1$	1.3×10^{-9}	54 (9057)	0.1481	0.2409	8.211×10^{-4}	31295.5
$L=L_1$	0.3×10^{-9}	13 (5553)	0.1469	0.2464	8.877×10^{-5}	31296.6
$L=L_2$	4.5×10^{-10}	87(10782)	0.1445	0.1845	1.251×10^{-4}	31295.8
$L=L_2$	2.5×10^{-10}	45 (10057)	0.1456	0.1856	6.999×10^{-5}	31296.4

function $f_0(g;x^{(k)})$ decreases less quickly and, consequently, near the critical level error 0.016916, the method provides more information that can lead to a better approximation of g^*.

The inexact Bregman iteration can be efficiently used for the reconstruction of high contrast range images, using SGP scheme as inner solver. Any inner subproblem is very similar to (8); indeed we have that

$$D^{p^{(k)}}f_1(\bar{x},\bar{x}^{(k)})=\frac{1}{2}\left\|L\left(x_E-x_E^{(k)}\right)\right\|^2 \tag{28}$$

where $\bar{x}^{(k)}=\begin{pmatrix}x_P^{(k)}\\ x_E^{(k)}\end{pmatrix}$ and L is the identity or the first, second order finite difference operator. Then, the solution of any inner subproblem exists and, under the assumption of Theorem 1, it is unique. As explained in Section 2, SGP method can be efficiently used as inner solver of the inexact scheme. For this application, since the Bregman iteration requires to follow the descent trajectory of f_0, to avoid a too fast decrease close to the levels of interest, it is crucial that the upper bound α_{max} of the steplength has a value lower than the one in the standard setting of SGP, i.e. 10 instead of 10^5.

The value of β can be set equal to an overestimation of the regularization parameter, obtained by some discrepancy criterion, such as that in [7]. The effectiveness of this approach is shown in the next section on a set of simulated test-problems.

4. Numerical experiments

This section is devoted to numerically evaluate the effectiveness of the inexact procedure based on Bregman iteration for recovering high contrast images corrupted by Poisson noise. The numerical experiments described in this section have been performed in MATLAB environment (R2012b), on a PC equipped with an Intel Core i7-3517U processor 1.9 GHz, 8 GB RAM.

We consider four test-problems, provided by the authors of [12], that are simulated Large Binocular Telescope (LBT) infrared narrow-band observations of a star-jet system. In all test-problems the point source is a single star ($q=1$). The images are obtained from an optical image taken with Hubble Space Telescope of the HH34 jet (Herbig–Haro objects) [35], downscaled by a factor 3 and immersed into a 256×256 array, surrounded by zeros. A pixel scale of 15 mas/pixel has been assumed as typical scale of a camera attached to a single LBT dish. In the pixel $(178,129)$ corresponding to the source, a point-like star with variable magnitude (from 8 to 11 mag) was added in order to simulate four different objects at different

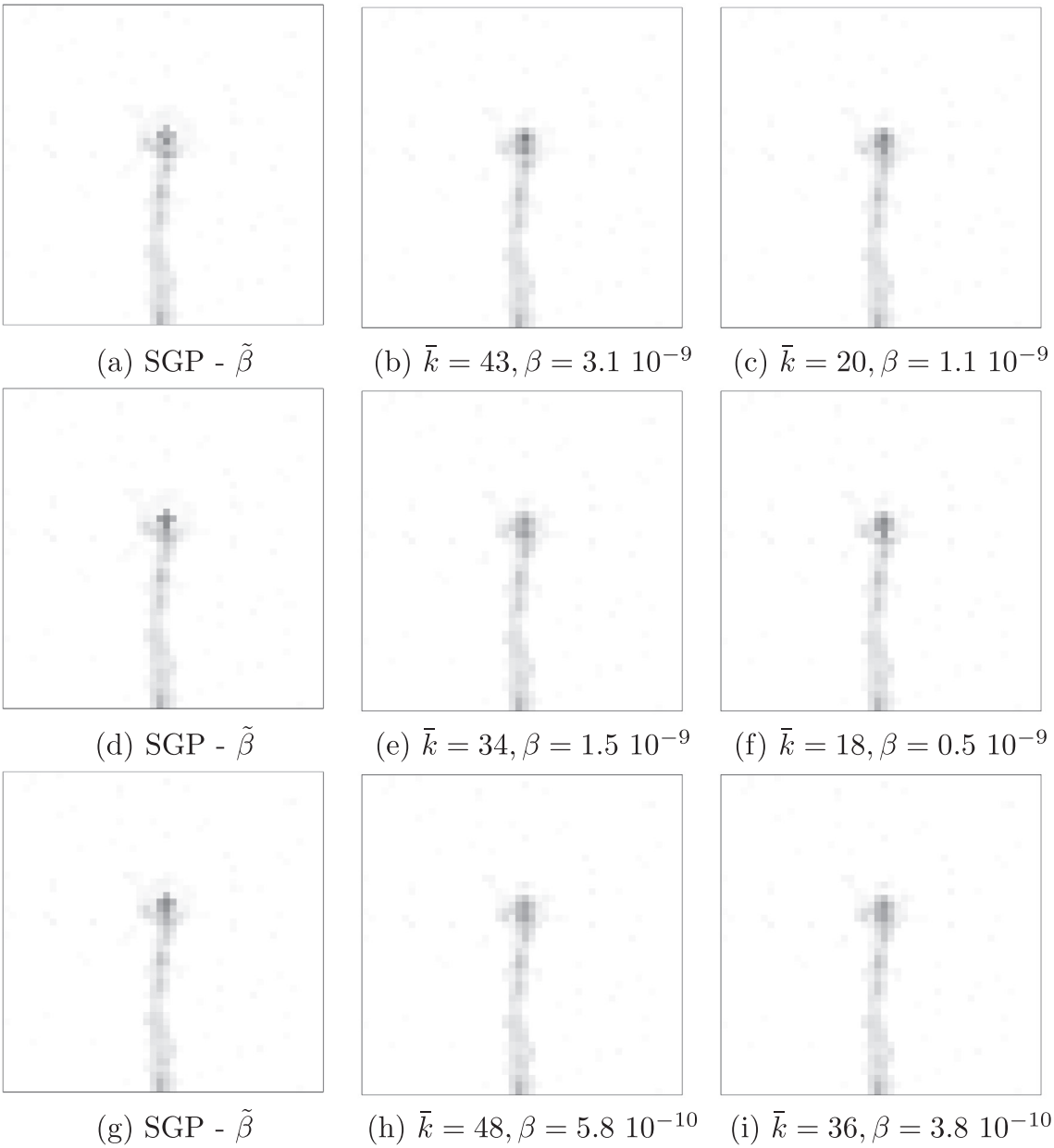


Fig. 3. Test-problem *Star8*: restored images obtained by solving (8) with $\tilde{\beta}$ (SGP method, see Table 3) and by inexact Bregman method; we show a region around the star. For the Bregman method, we show the reconstructions obtained at the iteration $\bar{k}$ (see Table 4), corresponding to the best value of the relative reconstruction error. The regularization term is given by $f_1(x_E) = \frac{1}{2}\|Lx_E\|^2$, with $L = I$ in (a)–(c) *panels*, $L = L_1$ in (d)–(f) *panels* and $L = L_2$ in (g)–(i) *panels*. All the images are in reverse gray sqrt-scale.

contrast. For each object, a 256×256 image has been obtained by convolving the object with an adaptive-optics corrected PSF with a Strehl ratio of 0.67. An average H-band sky brightness of 15.5 mag/arcsec^2 has been added as background emission and, finally, the results are corrupted by a mixture of Poisson noise and additive Gaussian noise with $\sigma = 10e^-$/ pixel. A more precise and detailed description on the image acquisition process is available in [12]. The four test-problems are denoted by *Star8, Star9, Star10, Star11*, in dependence of the magnitude of the point source. In Fig. 2 we show in reverse gray square root scale (sqrt-scale) the original diffuse object x_E^*, the deblurred image related to the test-problem *Star8* and a detailed view of the region around the star; a square 14×14 box around the point source is emphasized.

In the first numerical simulation, we compute a value for the regularization parameter β of the problem (8) through the discrepancy principle proposed in [7]. By a method combining bisection and secant iterations, we solve the discrepancy equation

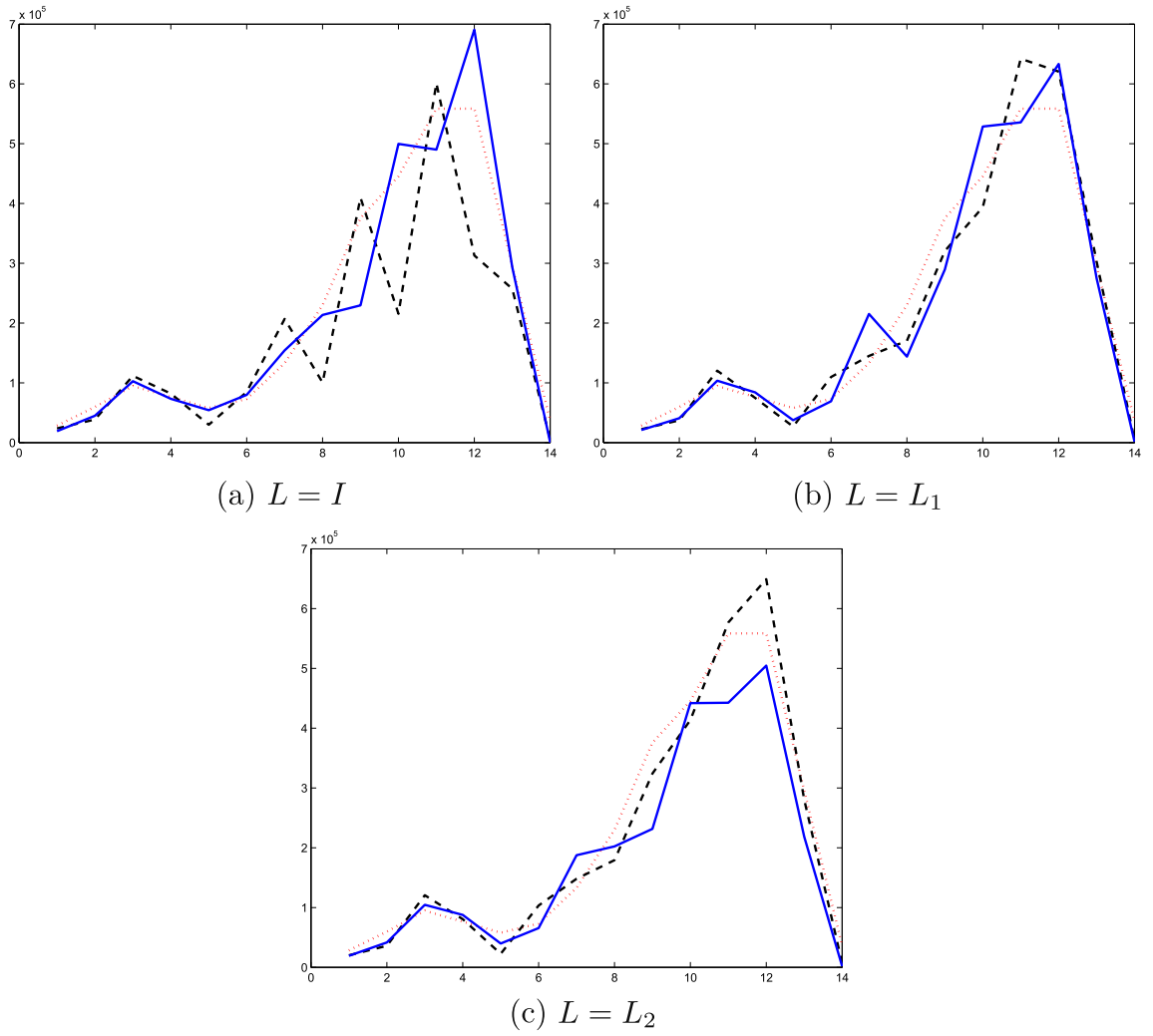


Fig. 4. Test-problem *Star8*: superposition of the line outs from column number 129 (from pixel 167 to pixel 180) for the restored images; the *dotted line* (red) is relative to the original diffuse object; the *dashed line* (black) is relative to the restored image obtained by solving (8) with $\bar{\beta}$ (SGP method, see Table 3); the blue lines are relative to the reconstruction obtained by inexact Bregman method (see Table 4) at the iteration k (*solid line*). The value of β are respectively 1.1×10^{-9} ($L = I$), 0.5×10^{-9} ($L = L_1$), 3.8×10^{-10} ($L = L_2$). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$$\Theta_g(\beta) = \frac{2}{m} f_0(g, \overline{H}\bar{x}_\beta + b) = 1 \tag{29}$$

where $\bar{x}_\beta$ is the solution of (8). For all test-problems, bisection iterations show that at the extremes of the interval $[10^{-10}, 10^{-8}]$, the discrepancy function $\Theta_g(\beta)$ changes the sign. In Table 2 we report the approximate solution $\overline{\beta}$ of (29) for the four test-problems in the case of different regularization terms. Here L_1 and L_2 denote the first and second order finite difference operators respectively. The minimization problems (8) involved in the evaluation of the discrepancy function $\Theta_g(\beta)$ are solved by SGP method; the Matlab code of SGP is an adapted version of that downloadable from *http://www.unife.it/prin/software*; the scheme is stopped when the relative difference between two consecutive values of the objective function is less than tol_{SGP}

$$|f_\beta(\bar{x}^{(k)}) - f_\beta(\bar{x}^{(k-1)})| \leqslant \text{tol}_{SGP} |f_\beta(\bar{x}^{(k)})| \tag{30}$$

and the mean of this relative difference over the last M iterations is less than $10 \cdot tol_{SGP}$. In these experiments, $\text{tol}_{SGP} = .5\ 10^{-9}$ and $M = 10$. In Table 2 we report also the total number *totit* of SGP iterations needed for the evaluations of $\Theta_g(\beta)$ and, in brackets, the number T of evaluations performed to obtain the numerical solution of (29) so that $|\Theta_g(\overline{\beta}) - 1| \leqslant 10^{-4}$ and the relative distance between the last two iterates is less than 10^{-3}. Moreover we report the relative reconstruction errors

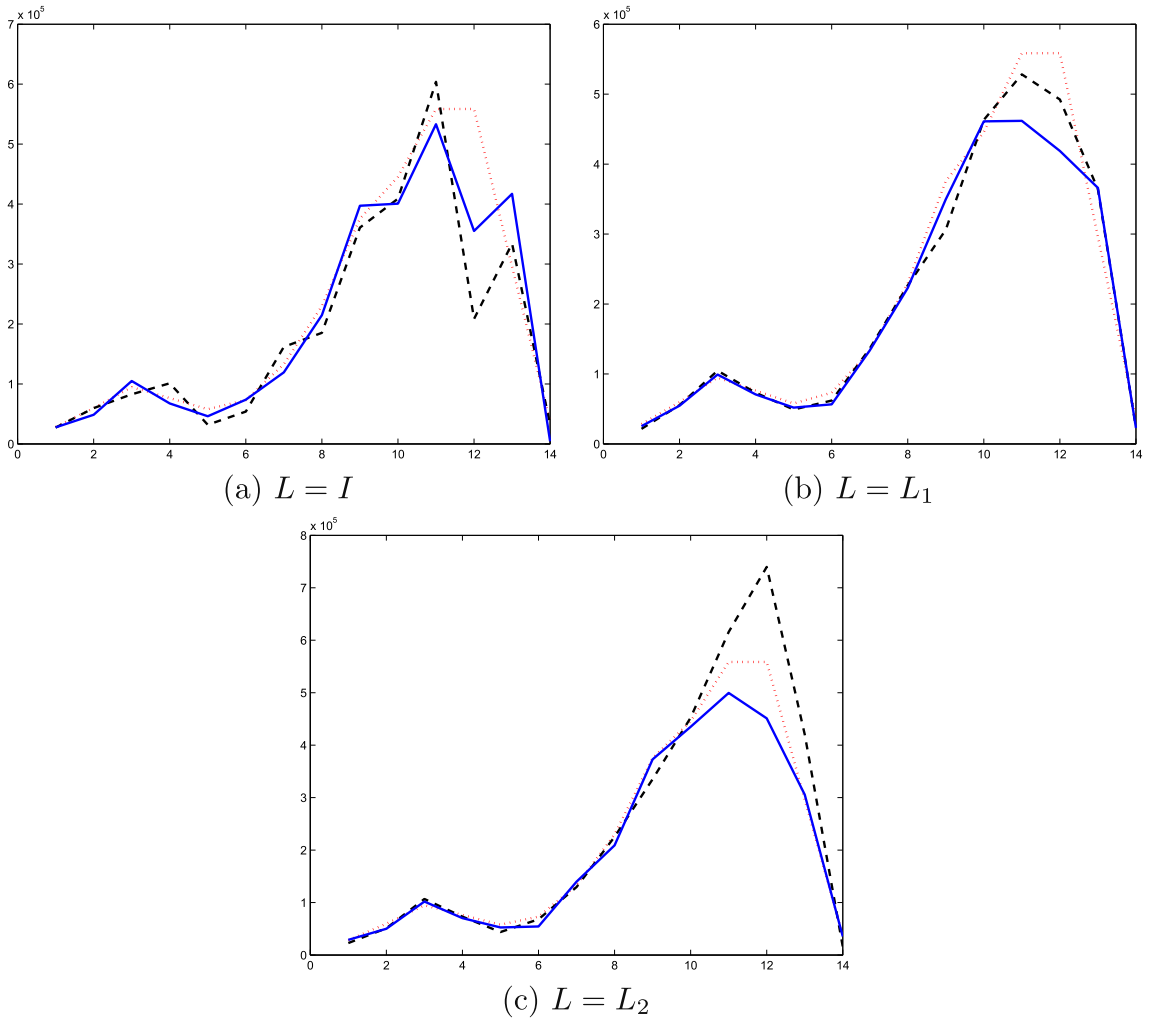


Fig. 5. Test-problem *Star11*: superposition of the line outs from column number 129 (from pixel 167 to pixel 180) for the restored images; the *dotted line* (red) is relative to the original diffuse object; the *dashed line* (black) is relative to the restored images obtained by solving (8) with $\bar{\beta}$ (SGP method, see Table 3); the blue lines are relative to the reconstruction obtained by inexact Bregman method (see Table 4) at the iteration $\bar{k}$ (*solid line*). The value of β are respectively 1.7×10^{-9} ($L = I$), 0.3×10^{-9} ($L = L_1$), 2.5×10^{-10} ($L = L_2$). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ρ, ρ_w, ρ_P for the image solution corresponding to $\overline{\beta}$. In particular, if $\bar{x}_\beta = \begin{pmatrix} \bar{x}_{P_\beta} \\ x_{E_\beta} \end{pmatrix}$ is the numerical solution of (8) for a given β, we denote by ρ the relative reconstruction error in l_2 norm of the diffuse component, given by

$$\rho = \frac{\|x_{E_\beta} - x_E^*\|}{\|x_E^*\|} \tag{31}$$

Since the astrophysical interest for high contrast images is focused on the region close to the star, we consider even a relative reconstruction error ρ_w, computed considering only the pixels of x_E^* and x_{E_β} within the 14×14 square window around the star (shown in Fig. 2). Furthermore, ρ_P is the relative error for the intensity of the point-like star. If we consider the reconstruction errors for the images obtained in the intermediate iterations of the scheme used to solve the nonlinear Eq. (29), we observe that for all test-problems the final value of $\overline{\beta}$ does not correspond to the solutions with minimum reconstruction error. In general, in the intermediate iterations, the reconstruction errors can be smaller than those in Table 2. A similar behavior can be observed for the other regularization terms.

Therefore, for any regularization term, we determine experimentally another value of the regularization parameter $\tilde{\beta}$; this value corresponds to the minimum relative reconstruction error ρ for the diffuse component obtained by several trials with different values of β. In Table 3 we report the results related to these values. Here $\tilde{k}$ denotes the number of SGP iterations needed to satisfy the stopping criterion.

In the second numerical simulation, we recover the high contrast images of the four test-problems by the inexact Bregman method combined as SGP as inner solver. The values $\overline{\beta}$ obtained from the discrepancy equation are used as an overestimation of the optimal value of the regularization parameter; the numerical experiments are carried out also for another value of β that is less than $\overline{\beta}$ and still greater than the optimal value of the regularization parameter. For the inner stopping rule, the sequence $\{\mu_k\}$ is chosen as in (27), setting the value of c as explained in Remark of Section 3: the first subproblem is solved by the standard stopping rule (30) of SGP (with moderate tolerance, i.e. $tol_{SGP} = 10^{-5}, M = 1$ or $M = 5$); then $c = \|\eta^{(1)}\|$ and $\chi = 1.5$. Furthermore, in order to avoid a too fast decrease of f_0 close to the levels of interest, the maximum steplength α_{max} is set equal to 10 until $\mu_k = c/k^{\chi} > 5\ 10^{-4}$; this happens within the first ten/twenty outer iterations; then $\alpha_{max} = 1000$. In Table 4 we report the results obtained with $f_1(x_E) = \frac{1}{2}\|Lx_E\|^2$; $\overline{k}$ denotes the outer iteration of the inexact Bregman scheme where we obtain the minimum relative reconstruction error $\rho^{(\overline{k})}$ of the diffuse component; we report in brackets also the total number of SGP iterations (*totit*) involved in $\overline{k}$ outer iterations, the relative error $\rho_w^{(\overline{k})}$ for the square window around the star, the relative error $\rho_P^{(\overline{k})}$ for the point component and the value of the KL function at the iterate $\bar{x}^{(\overline{k})}$. Table 4 shows that the inexact Bregman scheme enables us to obtain relative reconstruction errors at least comparable, but in the most cases better than those obtained by solving the variational problem with a suitable value of the regularization parameter. With regard to the box around the star, we observe that the values of $\rho_w^{(\overline{k})}$ are lower or approximately equal to those in Table 3. In Fig. 3 we show a magnification of the part of the jet close to the star in the restored images related to test-problem *Star8*; using the three different regularization terms, we show on the left panels the results obtained by solving with SGP the variational problem (8) with $\beta = \bar{\beta}$ (see Table 3). On the middle and right panels we show the same region obtained by the inexact Bregman method at the outer iteration $\bar{k}$ for two different overestimation of β (see Table 4). For all regularization terms, the box around the star appears less affected by artifacts when using the inexact Bregman scheme in confirmation of the results about $\rho_w^{(\overline{k})}$.

For the same test-problem, in Fig. 4 we show the superposition of the column number 129 (from pixel 167 to pixel 80) of the original object (dotted line) and of some restored images; in particular we compare the result (dashed line) obtained by solving (8) with $\beta = \bar{\beta}$ and the one by the inexact Bregman method. For this method, we report the result of the outer iteration with minimum error $\rho^{(\overline{k})}$ for the diffuse component (solid line). We show similar results in Fig. 5 for test-problem *Star11*. Table 4 and Figs. 3–5 highlight that a regularization term involving first order or second order finite differences provides in general a best quality in the reconstructions, also with high intensity of the star.

Furthermore we observe that the number of outer iterations needed to obtain a satisfactory restored image in general decreases with increasing the accuracy of the parameter estimate.

From the point of view of the efficiency, the inexact Bregman scheme is not too expensive; indeed, because of the moderate tolerance required in the determination of the outer iterates, the total number of SGP iterations is two or three times greater than the one required to solve with high accuracy the variational problem (8) with an optimal value of the regularization parameter.

5. Conclusions

In this paper we consider the deconvolution problem of high contrast images corrupted by Poisson noise. These images consist of very bright stars and other smooth structures underlying the stars. In order to investigate the regions around the stars, the original object is written as sum of two components, corresponding to the point sources x_P and to the extended sources x_E. When the point sources have known positions, the unknowns can be reduced to the intensities of the stars and the pixels of the diffuse component. In the framework of Bayesian approach, a maximum a posteriori estimate of the object can be obtained by solving a variational problem, where in presence of Poisson noise the data fidelity function is the generalized KL divergence and an explicit nonnegative and convex regularization term is introduced only for the diffuse component. Indeed the structure imposed on x_P implicitly works as a regularization. We derive the conditions that assure the existence and the uniqueness of the solution of the reduced optimization problem when a Tikhonov-like regularization is used. From the numerical point of view, the solution of the problem can be efficiently obtained by a Scaled Gradient Projection method, named SGP, specifically tailored for the KL divergence combined with a differentiable regularization term. Nevertheless, for Poisson noise, the selection of a suitable regularization parameter β is basically an open problem. Numerical investigations show that the discrepancy criterion introduced in [7] provides only an overestimation of the parameter. Then we propose to use this overestimation to restore the high contrast images by the inexact Bregman iteration combined with SGP method. The Bregman scheme requires the solution of a sequence of variational problems very similar to the original problem and it has the typical semiconvergence behavior of the iterative methods for the solution of inverse problems; the inexact scheme enables to solve with moderate accuracy the inner subproblems and simultaneously to monitor the decrease of data fidelity function near the noise level and the quality of the intermediate approximations of the restored image. The analysis of the numerical reconstructions obtained by the inexact Bregman scheme on a simulation of observations of a YSO jet appear very promising for the application to real cases. Indeed the relative errors in the restored images are at least comparable, but in the most cases better than those obtained by solving the variational problem with an accurate value of the regularization parameter. These results can be observed also in the regions around to the stars, that are very interesting for astrophysical investigations. Furthermore, a regularization term involving first order or second order finite differences provides a better quality in the reconstructions, also with high intensity of the star. The computational cost of

the inexact Bregman scheme is not too high, because of the moderate tolerance required in the determination of the outer iterates. Hence the reconstruction obtained at each iteration can be visually inspected with waiting times of a few seconds, above all on GPUs-equipped platforms when a GPU-based implementation of SGP is available. Future investigation could regard the introduction of other regularization functionals, such as Hyper Surface potential, and the implementation of the inexact Bregman procedure in a High Performance Computing framework, for example using a GPU-based architecture.

Acknowledgments

The authors thank Prof. M. Bertero for helpful suggestions and interesting discussions about the variational model and the numerical results; the authors are grateful to Dr. A. La Camera and Prof. P. Boccacci for providing the test-problems used in the simulations of the last section.

References

[1] Bertero M, Boccacci P, Desiderà G, Vicidomini G. Image deblurring with Poisson data: from cells to galaxies. Inverse Prob 2009;25:123006.
[2] Bardsley JM, Goldes J. Regularization parameter selection methods for ill-posed Poisson maximum likelihood estimation. Inverse Prob 2009;25:095005.
[3] Zanella R, Boccacci P, Zanni L, Bertero M. Efficient gradient projection methods for edge-preserving removal of Poisson noise. Inverse Prob 2009;25:045010.
[4] Brune C, Sawatzky A, Burger M. Scale space and variationale methods in computer vision. LNCS, vol. 5567. Springer; 2009. Ch. Bregman-TV-EM methods with application to optical nanoscopy, pp. 235–246.
[5] Figueiredo MAT, Bioucas-Dias JM. Restoration of Poissonian images using alternating direction optimization. IEEE Trans Image Process 2010;19(12):3133–45.
[6] Setzer S, Steidl G, Teuber T. Deblurring Poissonian images by split Bregman techniques. J Vis Commun Image R 2010;21:193–9.
[7] Bertero M, Boccacci P, Talenti G, Zanella R, Zanni L. A discrepancy principle for Poisson data. Inverse Prob 2010;26:10500.
[8] Ciliegi P, Camera AL, Desiderà G, Antoniucci S, Arcidiacono C, Lombini M, Diolaiti E, Bellocchi E, Mannucci F, Bertero M, Boccacci P, Lorenzetti D, Nisini B. Analysis of LBT LINC-NIRVANA simulated images of galaxies and young stellar objects. In: Schöller M, Danchi W, Delplancke F, editors. Optical and infrared interferometry. Proc SPIE, vol. 7013. p. 701335.
[9] Giovanelli JF, Coulais A. Positive deconvolution for superimposed extended source and point sources. Astron Astrophys 2005;439:402–12.
[10] Lucy LB. An iterative technique for the rectification of observed distributions. Astron J 1974;79:745–54.
[11] Camera AL, Antoniucci S, Bertero M, Boccacci P, Lorenzetti D, Nisini B. Image reconstruction for observations with an high dynamic range: LINC-NIRVANA simulations of a stellar jet. In: Delplancke F, Rajagopal FJK, Malbet F, editors. Optical and infrared interferometry iii. Proc SPIE, vol. 8455. p. 8455. 3D.
[12] Camera AL, Antoniucci S, Bertero M, Boccacci P, Lorenzetti D, Nisini B, Arcidiacono C. Reconstruction of high dynamic range images: Simulations of LBT observations of a stellar jet, a pathfinder study for future AO-assisted giant telescopes. Publ Astron Soc Pac 2014;126:180–93.
[13] Staglianò A, Boccacci P, Bertero M. Analysis of an approximate model for Poisson data reconstruction and a related discrepancy principle. Inverse Prob 2011;27:125003.
[14] Carlavan M, Blanc-Féraud L, Two constrained formulations for deblurring Poisson noisy images. In: Proc. IEEE international conference on image processing (ICIP), Brussels, Belgium.
[15] Teuber T, Steidl G, Chan RH. Minimization and parameter estimation for seminorm regularization models with I-divergence constraints. Inverse Prob 2013;29:035007.
[16] Engl HW, Hanke M, Neubauer A. Regularization of inverse problems. Dordrecht: Kluwer; 1996.
[17] Osher S, Burger M, Goldfarb D, Xu J, Yin W. An iterative regularization method for total variation-based image restoration. SIAM J Multiscale Model Simul 2005;4(2):460–89.
[18] Brune C, Sawatzky A, Burger M. Primal and dual Bregman methods with application to optical nanoscopy. Int J Comput Vis 2010;92(2):211–29.
[19] Benfenati A, Ruggiero V. Inexact Bregman iteration with an application to Poisson data reconstruction. Inverse Prob 2013;29:065016.
[20] Bonettini S, Zanella R, Zanni L. A scaled gradient projection method for constrained image deblurring. Inverse Prob 2009;25:015002.
[21] Charbonnier P, Blanc-Féraud L, Aubert G, Barlaud A. Deterministic edge-preserving regularization in computed imaging. IEEE Trans Image Process 1997;6:298–311.
[22] Chambolle A, Pock T. A first-order primal-dual algorithm for convex problems with applications to imaging. J Math Imaging Vision 2011;40:120–45.
[23] Bonettini S, Ruggiero V. An alternating extragradient method for total variation based image restoration from Poisson data. Inverse Prob 2011;27:095001.
[24] Mol CD, Defrise M, Inverse imaging with mixed penalties. In: Proceedings URSI EMTS 2004, Ed. PLUS Univ. Pisa, 2004, pp. 798–800.
[25] Ayasso H, Rodet T, Abergel A. A variational Bayesian approach for unsupervised super-resolution using mixture models of point and smooth sources applied to astrophysical map-making. Inverse Prob 2012;28(12):125005.
[26] Geman S, Geman D. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. IEEE Trans Pattern Anal Mach Intel 1984;6:721–41.
[27] S. Bonettini, V. Ruggiero, On the uniqueness of the solution of image reconstruction problems with Poisson data. In: T.S. et al., editor. Proceedings of ICNAAM 2010, AIP conference proceedings, vol. 1281. AIP; 2010. p. 1803–6.
[28] Zhan B, Gao L, Dai YH. Gradient methods with adaptive step-sizes. Comput Optim Appl 2006;35:69–86.
[29] Frassoldati G, Zanni L, Zanghirati G. New adaptive stepsize selections in gradient methods. J Ind Manag Optim 2008;4(2):299–312.
[30] Dongarra JJ, Croz JD, Hammarling S, Hanson RJ. An extended set of fortran basic linear algebra subprograms. ACM Trans Math Soft 1988;14:1–17.
[31] C. Moler, Matlab incorporates lapack, <http://www.mathworks.fr/company/newsletters/articles/matlab-incorporates-lapack.html>, [Retrieved 26 October 2013] (2000).
[32] Ruggiero V, Serafini T, Zanella R, Zanni L. Iterative regularization algorithms for constrained image deblurring on graphics processors. J Global Optim 2010;48:145–57.
[33] Bregman LM. The relaxation method of finding the common points of convex sets and its applications to the solution of problems in convex optimization. USSR Comput Math Math Phys 1967;7:200–17.
[34] Goldstein T, Osher S. The split Bregman algorithm for L1 regularized problems. SIAM J Imaging Sci 2009;2(2):323–43.
[35] Reipurth B, Heathcote S, Morse J, Hartigan P, Bally J. Hubble space telescope images of the HH34 jet and bow shock: structure and proper motions. Astron J 2002;123:362–81.