Inexact Bregman iteration with an application to Poisson data reconstruction

A. Benfenati and V. Ruggiero

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

E-mail: alessandro.benfenati@unife.it, valeria.ruggiero@unife.it

Abstract. This work deals with the solution of image restoration problems by an iterative regularization method based on the Bregman iteration. Any iteration of this scheme requires to exactly compute the minimizer of a function. However, in some image reconstruction applications, it is either impossible or extremely expensive to obtain exact solutions of these subproblems. In this paper, we propose an inexact version of the iterative procedure, where the *inexactness* in the inner subproblem solution is controlled by a criterion that preserves the convergence of the Bregman iteration and its features in image restoration problems. In particular, the method allows to obtain accurate reconstructions also when only an overestimation of the regularization parameter is known. The introduction of the *inexactness* in the iterative scheme allows to address image reconstruction problems from data corrupted by Poisson noise, exploiting the recent advances about specialized algorithms for the numerical minimization of the generalized Kullback–Leibler divergence combined with a regularization term. The results of several numerical experiments enable to evaluate the proposed scheme for image deblurring or denoising in presence of Poisson noise.

Keywords. Bregman iteration, Inexact Bregman iteration, Image restoration, Regularization parameter, Poisson noise.

1. Introduction

This paper concerns the iterative procedure based on the use of Bregman distance, that has gained interest as regularization method in image restoration problems. Mathematically, image restoration is an inverse and ill-posed problem that consists in finding an approximation of the original object $x^* \in R^n$ from a set $g \in R^m$ of detected data. We assume that the distortion due to the acquisition system is described by a given matrix $H \in \mathbb{R}^{m \times n}$. In the Bayesian framework [1, 2], an approximation of the original object x^* is obtained by solving the following constrained nonlinear programming problem

$$\min_{x \in C} f_0(g; x) + \beta f_1(x) \tag{1}$$

where C is the nonnegative orthant or a subset that describes some physical constraint on x (as, for example, the flux conservation) and the objective function is the combination of two terms: the first one is a convex nonnegative functional $f_0(g; x)$ which measures the discrepancy from the data g and has to be chosen according to the noise statistics, while the second one is a regularization term $f_1(x)$, weighted by a positive parameter β. $f_1(x)$ is a convex nonnegative function that enables to incorporate *a priori* information about the expected solution into the reconstruction process.
To recover images with sharp edges, Total Variation (TV) functional [3] is used. This regularization prefers piecewise constant functions and then it can produce a loss of contrast. In order to obtain a contrast enhancement, a procedure based on the Bregman iteration [4] is proposed by Osher et al. in [5] to recover images corrupted by Gaussian noise and, later, by Brune et al. in [6] for dealing with Poisson data.

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

$$D^p F(x, y) = F(x) - F(y) - < p, x - y > \tag{2}$$

where p is a subgradient of F at y and $< \cdot, \cdot >$ denotes the canonical inner product of two vectors of $\mathbb{R}^n$. The Bregman iteration consists in solving a sequence of subproblems, similar to (1) with $f_1(x)$ replaced by its Bregman distance at the current iterate, as follows:

$$x^{(k+1)} = \text{argmin}_x Q_k(x, p^{(k)}) \equiv f_0(g; x) + \beta D^{p^{(k)}} f_1(x, x^{(k)}) \quad k = 0, 1, ... \tag{3}$$

with $p^{(k)} \in \partial f_1(x^{(k)})$ and $p^{(0)} \equiv 0$; when $C \neq \mathbb{R}^n$ in (1), we can track back to an unconstrained problem, by redefining $f_0(g; x)$ as $f_0(g; x) + i_C(x)$ (or $f_1(x)$ as $f_1(x) + i_C(x)$), where $i_C(x)$ is the indicator function of the set C.

The Bregman iterative scheme enables to address another relevant issue of the model (1), that is the choice of the regularization parameter β. This selection is especially difficult in the case of Poisson noise (see for example [7, 8, 9, 10]). The Bregman approach allows to use an overestimated value of β. Under suitable assumptions, for noise–free data $g^* \in \mathbb{R}^m$, the sequence of the iterates obtained by (3) converges to a minimizer of $f_0(g^*; x)$, which in general coincides with the original object x^*; consequently, for noisy data, $\{x^{(k)}\}$ does not converge to x^*, but to a minimizer $\bar{x}$ of $f_0(g; x)$. If an estimate γ of the discrepancy between Hx^* and the noisy data g is known, for value of β moderately large, the Bregman distance of f_1 between x^* and the current iterate is decreasing as long as $f_0(g; x^{(k)}) \geq \gamma$ [5]. Then, for noisy data, the Bregman iteration has the typical semi-convergence property, described, for instance, in [11]. Numerical results show that, in general, for a raw overestimation of β, few iterations are required to obtain a satisfying solution and, consequently, few instances of the subproblem (3) have to be solved. Indeed, since the first step requires the solution of the original problem with a larger influence of the regularization term, the first iterate $x^{(1)}$ is an over regularized approximation of the solution. The additional information available in $x^{(1)}$ can be used at the second step, when we have to minimize the same data fidelity function combined with the Bregman distance of f_1 at $x^{(1)}$, that can be interpreted as the residual between the regularization term and its approximation around $x^{(1)}$. In this way, we obtain an enhanced approximation $x^{(2)}$ and so on [12, 13]. Looking back at the statistical model in [2], the Bregman distance of f_1 in the iterative procedure can be interpreted as a refinement of the a–priori probability, motivating the observed contrast enhancement.

The convergence and the properties of the Bregman iteration hold when these subproblems are exactly solved while in many applications a closed formula for the minimizer may be unavailable. 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.

The aim of this work is to propose a strategy to deal with the inexact solution of the inner subproblems, devising how we can preserve the convergence of the iterative procedure. The analysis is proposed in a discrete setting, since, in real applications, it

may be necessary to use iterative minimization algorithms for solving inner subproblems. As remarked in [14], the Bregman method can be viewed as a generalization of the well-known proximal point algorithm, where the Euclidean distance is replaced by the Bregman distance of a strictly convex and differentiable function (Bregman nonlinear proximal point algorithm). Indeed, when the regularization term is the quadratic penalization $\frac{1}{2}\|x\|_2^2$, the Bregman procedure is exactly the proximal point method with a constant parameter. In recent years, there has been a growing interest in inexact implementations of proximal point algorithms, starting from the seminal work of Rockafellar [15] (for a survey see [16] and reference therein). Inexact schemes are recently proposed also for the Alternating Direction Methods [17]. In this paper, the approach is similar to that in [18] where the notion of ϵ-subgradient is introduced to deal with inexact Bregman schemes; in our work, a criterion to monitoring the inexactness of the current iterate is devised; the updating rule for the subgradient of the function f_1 is eliminated and the subgradient or the ϵ-subgradient of f_1 (required in the subsequent subproblem) is obtained by the inner iterative solver used for the current subproblem. Furthermore, in this analysis, $f_1(x)$ is a proper lower semicontinuous convex function, not necessarily differentiable and, analogously, $f_0(g;x)$, so that the convergence of the method holds also for general image restoration problems (for example, for impulsive noise [19]). For the effectiveness of the scheme, the crucial issue becomes to devise an efficient algorithm for approximately solving the inner subproblems.

The introduction of the inexactness in the Bregman iteration allows to address image reconstruction problems with Poisson data (deblurring and denoising), exploiting the recent advances in devising efficient specialized algorithms for the Kullback Leibler divergence combined with a regularization term for the solution of the inner minimization subproblems (see for example [20, 21, 22, 23, 24, 25, 26, 27] and references therein). A further contribution of this paper is to carry out an experimental analysis, showing that the proposed scheme can be an effective tool for image restoration in presence of Poisson noise, when only an overestimation of the regularization parameter is known.

The paper is organized as follows. In section 2, in a discrete setting we recall the iterative method based on Bregman iteration and its convergence and we describe the proposed inexact version of the method, devising an inner stopping criterion that assures the convergence of the scheme. In section 3, we introduce the inexact Bregman method for image denoising and deblurring from data corrupted by Poisson noise and we discuss some methods for solving the inner subproblems. Finally in section 4 we describe the results of a set of numerical experiments concerning the proposed approach in the case of Poisson data.

2. The iterative procedure, based on the Bregman iteration

In this section we recall the iterative Bregman method, its convergence properties and the features of the procedure in the framework of image restoration problems. Then we introduce the inexact version and the assumptions for the convergence.

2.1. The problem and the iterative general scheme

The original problem considered by Bregman in [4] can be formulated as follows:

$$\min_x f_1(x) \text{ subject to } Hx = f \tag{4}$$

where f_1 is a proper, closed, convex and nonnegative function defined over $\mathbb{R}^n$, $H \in \mathbb{R}^{m\times n}$ is a given matrix and $f \in \mathbb{R}^m$. We assume that the active set $\{x \in \text{dom} f_1 | Hx = f\}$ is not empty. Using the classical penalty approach, we can replace the problem (4) with a sequence of unconstrained problems

$$\min_x q_k(x) \equiv f_1(x) + \frac{1}{\beta_k} f_0(x) \tag{5}$$

where β_k is a positive constant and $f_0(x)$ is a penalty function, that is $f_0(x) \geq 0$ for any x and $f_0(x) = 0$ if and only if $Hx = f$. It is well known that, when the subproblems $q_k(x)$ have a solution $x^{(k)}$ and $\{\beta_k\}$ is a decreasing sequence tending to 0 for $k \to \infty$, any limit point of $\{x^{(k)}\}$ is a solution of (4) (see [28, p.402]). Nevertheless a small value for β_k can make (5) extremely difficult to be solved numerically.
In the procedure based on the Bregman iteration [4], the value of β_k is kept constant, $\beta_k \equiv \beta$, and $f_1(x)$ is replaced by its Bregman distance at the current iterate. Furthermore we can choose $f_0(x)$ as a coercive and convex function, so that any subproblem admits a solution. For example, in the image restoration framework, $f_0(x)$ is a data fidelity function. Then the Bregman iteration modifies (5) by iteratively solving a sequence of subproblems according to the following scheme:

- given $x^{(0)}$ such that $p^{(0)} \equiv 0 \in \partial f_1(x^{(0)})$
- for $k = 0, 1, 2, ...$

$$\begin{aligned} x^{(k+1)} = \text{argmin}_x Q_k(x, p^{(k)}) \equiv D^{p^{(k)}} f_1(x, x^{(k)}) + \frac{1}{\beta} f_0(x) = \\ = f_1(x) - f_1(x^{(k)}) - < p^{(k)}, (x - x^{(k)}) > + \frac{1}{\beta} f_0(x) \end{aligned} \tag{6}$$

  with $p^{(k)} \in \partial f_1(x^{(k)})$

As observed in [14], the method can be described also as a generalization of the proximal point algorithm by the rule:

$$x^{(k+1)} = (\partial f_1 + \frac{1}{\beta} \partial f_0)^{-1} (\partial f_1(x^{(k)})$$

In the following proposition, we resume some features of the iterative method based on the Bregman iteration, already described in [5] for $f_0(x) = \frac{1}{2}\|Hx - f\|^2$. These results

hold for more general f_0 and f_1. For the sake of completeness, we report the proof in Appendix A.

Proposition 1 *Let $f_0(x)$ and $f_1(x)$ be nonnegative, proper, closed and convex functions, with $dom(f_0) \subseteq dom(f_1)$ and the relative interiors of f_0 and f_1 have at least a point in common. We assume that, for any k, there exists a minimizer $x^{(k+1)}$ of the subproblem (6); then, the following conditions hold:*

(a) there exist $p^{(k+1)} \in \partial f_1(x^{(k+1)})$ and $q^{(k+1)} \in \partial f_0(x^{(k+1)})$ such that

$$p^{(k+1)} = p^{(k)} - \frac{1}{\beta} q^{(k+1)} \tag{7}$$

(b) the sequence $f_0(x^{(k)})$ is monotonically non increasing and we have

$$f_0(x^{(k)}) \leq f_0(x^{(k)}) + \beta D^{p^{(k-1)}} f_1(x^{(k)}, x^{(k-1)}) \leq f_0(x^{(k-1)}) \tag{8}$$

(c) if there exists x such that $f_1(x) < \infty$, we have

$$\begin{aligned} &\beta \left(D^{p^{(k)}} f_1(x, x^{(k)}) + D^{p^{(k-1)}} f_1(x^{(k)}, x^{(k-1)}) \right) + f_0(x^{(k)}) \leq \\ &\leq f_0(x) + \beta D^{p^{(k-1)}} f_1(x, x^{(k-1)}) \end{aligned} \tag{9}$$

(d) if $\widehat{x}$ is a minimizer of $f_0(x)$ such that $f_1(\widehat{x}) < \infty$, we have that

$$D^{p^{(k)}} f_1(\widehat{x}, x^{(k)}) \leq D^{p^{(k-1)}} f_1(\widehat{x}, x^{(k-1)}) \tag{10}$$

and

$$f_0(x^{(k)}) \leq f_0(\widehat{x}) + \beta \frac{f_1(\widehat{x}) - f_1(x^{(0)})}{k} \tag{11}$$

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

Under suitable hypothesis, the above proposition guarantees the convergence of the minimizers of the subproblems (6) to a solution of $f_0(x) = 0$ (or $Hx = f$) while the sequence $\{D^{p^{(k)}} f_1(\widehat{x}, x^{(k)})\}$ is decreasing.
A crucial property to obtain these convergence results is the decreasing behavior of the sequence $\{f_0(x^{(k)})\}$, that follows from the nonnegativity of the Bregman distance of f_1 at the current iterate. From the numerical point of view, this requires that the updating rule (7) gives an exact subgradient at the current iterate. If this does not happen (for example, if $x^{(k)}$ is a rough approximation of the Q_{k-1}'s minimizer, $Q_k(x, p^{(k)})$ can assume negative values and $Q_k(x^{(k+1)}, p^{(k)})$ can be less than $f_0(x^{(k+1)})$ with the result that the sequence $\{f_0(x^{(k)})\}$ may have a non monotone behavior. Consequently, in the implementation of the method, the minimizer has to be obtained by a closed formula or by an iterative method with a very high accuracy.

As mentioned before, f_0 is a penalty function for the constraint $Hx = f$, that is $f_0(x) = 0$ if and only if $Hx = f$; several examples of functions satisfying this property

are the standard least squares $\frac{1}{2}\|Hx-f\|_2^2$, the l_1 norm $\|Hx-f\|_1$ or, under suitable assumptions, the generalized Kullback Leibler divergence $\sum_i f_i \log \frac{f_i}{(Hx)_i} + (Hx)_i - f_i$. We remark that, setting $y = Hx$, we can consider f_0 as a function of a variable y. Then $q^{(k+1)} \in H^* \partial_y f_0(Hx^{(k+1)})$ and we can write $q^{(k+1)} = H^* u^{(k+1)}$, with $u^{(k+1)} \in \partial_y f_0(Hx^{(k+1)})$, $k \geq 0$. Since $p^{(0)} = 0$, if we put $v^{(0)} = 0$, it is immediate to verify from (7) that, for $k \geq 0$, the following rule holds:

$$p^{(k+1)} = H^* v^{(k+1)} = H^*(v^{(k)} - \frac{1}{\beta} u^{(k+1)}) \tag{12}$$

Then the rule (7) can be substituted by

$$v^{(k+1)} = v^{(k)} - \frac{1}{\beta} u^{(k+1)} \tag{13}$$

and the definition of $x^{(k+1)}$ in (6) can be restated as

$$x^{(k+1)} = \text{argmin}_x \left(f_1(x) - < v^{(k)}, Hx > + \frac{1}{\beta} f_0(x) \right) \tag{14}$$

Using (14), in the following proposition we can prove that a solution $x^{\bar{k}}$ of $f_0(x) = 0$ obtained through (14)-(13) is indeed a solution of the original constrained problem (4) or its variant

$$\min_x f_1(x) \text{ subject to } f_0(x) = 0 \tag{15}$$

The proposition is a generalization of Theorem 2.2 in [29].

Proposition 2 *Let $f_0(x)$ be a convex function such that $f_0(x) = 0$ if and only if $Hx = f$. Suppose that some iterate $x^{\bar{k}}$ of the Bregman procedure satisfies $f_0(x^{\bar{k}}) = 0$. Then $x^{\bar{k}}$ is a solution of the constrained problem (4) (or (15)).*

Proof. Let $x^{\bar{k}}$ be such that $f_0(x^{\bar{k}}) = 0$ and

$$x^{\bar{k}} = \text{argmin}_x f_1(x) - < v^{\bar{k}}, Hx > + \frac{1}{\beta} f_0(x) \tag{16}$$

for a suitable $v^{\bar{k}}$. Let $\widehat{x}$ be a solution of the problem (4). Then $f_0(\widehat{x}) = 0$ and, for the hypothesis on f_0,

$$H\widehat{x} = f = Hx^{\bar{k}} \tag{17}$$

Since $x^{\bar{k}}$ satisfies (16), we have

$$f_1(x^{\bar{k}}) - < v^{\bar{k}}, Hx^{\bar{k}} > + \frac{1}{\beta} f_0(x^{\bar{k}}) \leq f_1(\widehat{x}) - < v^{\bar{k}}, H\widehat{x} > + \frac{1}{\beta} f_0(\widehat{x}) \tag{18}$$

Using (17) in (18) and taking into account that $f_0(x^{\bar{k}}) = f_0(\widehat{x}) = 0$, we have that

$$f_1(x^{\bar{k}}) \leq f_1(\widehat{x})$$

Because $\widehat{x}$ is a solution of the original optimization problem, this last inequality is an equality, showing that $x^{\bar{k}}$ solves (4). □

2.2. The iterative procedure for image restoration problems

When we solve an image restoration problem formulated as (1) by Bregman iteration, the data fidelity term $f_0(g;x)$ plays the role of penalty function for (4) or (15), with $f = g$; the array g is the detected noisy data and $H \in \mathbb{R}^{m\times n}$ is the imaging matrix. Then, thanks to Proposition 1, we can affirm that a limit point of $\{x^{(k)}\}$ is a solution $\bar{x}$ of $f_0(g;x) = 0$, while we are interested to a solution x^* of $f_0(g^*;x) = 0$ (here $f = g^*$), where g^* is the noise–free data. For this application, the Bregman iteration has the typical semi-convergence behavior of the iterative methods for the solution of inverse problems, as described for example in [11]; the sequence $\{x^{(k)}\}$ first approaches the required solution x^* and then it goes away, converging toward $\bar{x}$ [5]. Indeed, if an estimate γ for the noise level is known, that is $f_0(g;x^*) \le \gamma$, following the same argument used in [5], we can observe from (9) with $x = x^*$, that, while $f_0(g;x^{(k)}) \ge \gamma$, we have that the Bregman distance of the iterates from the object x^* decreases:

$$D^{p^{(k)}} f_1(x^*, x^{(k)}) \le D^{p^{(k-1)}} f_1(x^*, x^{(k-1)}) \tag{19}$$

Thank to (8), a stopping criterion for the iterative procedure is to terminate at the iteration k^* such that

$$k^* = \max\{k | f_0(g;x^{(k)}) \ge \gamma\} \tag{20}$$

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 Kullback Leibler divergence of $Hx^{(k)}$ and the detected data g reaches the noise level. For an estimate of this noise level, see the discrepancy criterion in [7].

2.3. The inexact iterative procedure

When a closed formula for the solution of inner minimization subproblem (6) is unavailable, at any step we can obtain an approximate solution by using an iterative solver with a severe stopping criterion. As a consequence, also for efficient methods, a huge number of iterations may be required. In this section we propose a strategy to deal with inexact solutions of the inner subproblems that preserve the convergence property of the iterative procedure. The crucial point of the proposed scheme is devising a suitable stopping criterion for the inner solver of subproblems (6).
To explain this novel scheme, we recall the basic definition and some results about the ϵ-subgradient of a proper convex function.

Definition 1 *[30, §23]. Let F be a proper convex function on $\mathbb{R}^n$. The ϵ-subdifferential of F at $x \in dom(F)$, defined for $\epsilon \in \mathbb{R}$, $\epsilon \ge 0$, is the set*

$$\partial_\epsilon F(x) = \{p \in \mathbb{R}^n : \ F(z) \ge F(x) + \langle p, z - x\rangle - \epsilon, \quad \forall z \in \mathbb{R}^n\}$$

A vector $p \in \partial_\epsilon F(x)$ is an ϵ-subgradient of F at x.

For $\epsilon = 0$ the definition of subdifferential is recovered while for $\epsilon > 0$ we have a larger set; furthermore, for $\epsilon_1 > \epsilon_2 > 0$, we have $\partial_{\epsilon_1} F(x) \supseteq \partial_{\epsilon_2} F(x) \supseteq \partial F(x)$.

Definition 2 *[30, §12]. The conjugate of a convex function F is the function F^* defined by*

$$F^*(y) = \sup_x (\langle x, y\rangle - F(x))$$

If F is lower semicontinuous and proper, then F^* is lower semicontinuous and $F^{**} = F$.

Proposition 3 *[27]. Let $F(x)$ be a proper lower semicontinuous convex function. Then, for every $x \in dom(F)$ and $p \in dom(F^*)$ we have $p \in \partial_\epsilon F(x)$, with $\epsilon = F(x) - (\langle p, x\rangle - F^*(p))$.*

We observe that $\epsilon = 0$ (that is $p \in \partial F(x)$) if and only if $F(x) = \langle p, x\rangle - F^*(p)$.
The computation of the Bregman distance of the function F between z and x requires a subgradient of F at x. When p is an ϵ-subgradient of F at x, we can introduce an *inexact* Bregman distance given by

$$\Delta_\epsilon^p F(z, x) = F(z) - F(x) - < p, z - x > + \epsilon \tag{21}$$

We have that, for any $z \in \mathbb{R}^n$, $\Delta_\epsilon^p F(z, x) \geq 0$. When $\epsilon = 0$ (that is $p \in \partial F(x)$), the definition of the Bregman distance is recovered and $\Delta_0^p F(z, x) = D^p F(z, x)$.
The inexact scheme can be stated as follows. Starting from $\tilde{p}^{(0)} = 0 \in \partial f_1(\tilde{x}^{(0)})$, at any step $k \geq 0$, we consider the subproblem

$$\begin{aligned} \min_x \tilde{Q}_k(x, \tilde{p}^{(k)}, \epsilon_k) &\equiv \frac{1}{\beta} f_0(x) + \Delta_{\epsilon_k}^{\tilde{p}^{(k)}} f_1(x, \tilde{x}^{(k)}) = \\ &= \frac{1}{\beta} f_0(x) + f_1(x) - f_1(\tilde{x}^{(k)}) - < \tilde{p}^{(k)}, x - x^{(k)} > + \epsilon_k \end{aligned} \tag{22}$$

where $\tilde{p}^{(k)} \in \partial_{\epsilon_k} f_1(\tilde{x}^{(k)})$ ($\epsilon_0 = 0$).

Algorithm - Inexact Bregman

Choose $\tilde{x}^{(0)}$ such that $\tilde{p}^{(0)} = 0 \in \partial f_1(\tilde{x}^{(0)})$, $\epsilon_0 = 0$, $\beta > 0$; choose sequences $\{\mu_k\}$ and $\{\nu_k\}$ such that $\sum_i \mu_i < \infty$ and $\sum_i i\nu_i < \infty$
For $k = 0, 1, 2, \ldots$ do the following steps:

Step 1. Determine by an iterative solver an approximate solution $\tilde{x}^{(k+1)}$ of the subproblem $\min_x \tilde{Q}_k(x, \tilde{p}^{(k)}, \epsilon_k)$ and the related $\tilde{q}^{(k+1)} \in \partial f_0(\tilde{x}^{(k+1)})$ and $\tilde{p}^{(k+1)} \in \partial_{\epsilon_{k+1}} f_1(\tilde{x}^{(k+1)})$ so that

$$\|\eta^{(k+1)}\| \leq \mu_{k+1} \qquad \text{and } \epsilon_{k+1} \leq \nu_{k+1}$$

with $\eta^{(k+1)} = \frac{1}{\beta}\tilde{q}^{(k+1)} + \tilde{p}^{(k+1)} - \tilde{p}^{(k)}$

Step 2. Terminate if a stopping criterion is satisfied

End

We suppose that the k–th inner subproblem is solved by an iterative algorithm, which enables us to compute an approximate solution $\tilde{x}^{(k+1)}$, a suitable subgradient $q^{(k+1)} \in \partial f_0(x^{(k+1)})$ and an ϵ_{k+1}-subgradient $\tilde{p}^{(k+1)}$ of f_1 at $\tilde{x}^{(k+1)}$, $k \geq 0$.
If f_1 is a differentiable function, we assume that the inner solver generates sequences $\{x_l^k\}$ and $\{q_l^k\}$ convergent to a minimizer $\bar{x}^k$ of $\tilde{Q}_k$ and a subgradient $\bar{q}^k \in \partial f_0(\bar{x}^k)$ respectively as $l \to \infty$. As a consequence, given $\mu_{k+1} > 0$, there exists an index $\bar{l}$ such that $\|\eta_l^k\| \leq \mu_{k+1}$, with $\eta_l^k = \frac{1}{\beta} q_l^k + \nabla f_1(x_l^k) - \tilde{p}^{(k)}$; then, we set as approximate solution of the $\tilde{Q}_k$ subproblem $\tilde{x}^{(k+1)} = x_{\bar{l}}^k$, with $\tilde{p}^{(k+1)} = \nabla f_1(\tilde{x}^{(k+1)})$, $\epsilon_{k+1} = 0$ and $q^{(k+1)} = q_{\bar{l}}^k$.
When f_1 is a non differentiable function, we can consider the primal–dual formulation of (22), given by

$$\min_x \max_y \Phi_k(x, y) \equiv \frac{1}{\beta} f_0(x) + < y, x > - f_1^*(y) - f_1(\tilde{x}^{(k)}) - < \tilde{p}^{(k)}, x - x^{(k)} > + \epsilon_k \quad (23)$$

and we can apply a suitable primal–dual method generating sequences $\{x_l^k\}$, $\{y_l^k\}$ convergent to a saddle point $(\bar{x}^k, \bar{y}^k)$ of the convex–concave proper function $\Phi_k(x, y)$. We recall that $(\bar{x}^k, \bar{y}^k)$ is a saddle point of $\Phi_k(x, y)$ if there exist $\bar{q}^k \in \partial f_0(\bar{x}^k)$ and $\bar{w}^k \in \partial f_1^*(\bar{y}^k)$ such that the following conditions hold:

$$\frac{1}{\beta} \bar{q}^k + \bar{y}^k - \tilde{p}^{(k)} = 0$$
$$\bar{x}^k = \bar{w}^k$$

Then, if we assume that the inner solver generates sequences $\{x_l^k\}$, $\{y_l^k\}$, $\{q_l^k\}$ which converge to $\bar{x}^k$, $\bar{y}^k$ and $\bar{q}^k$ respectively as $l \to \infty$, then the sequence of the dual iterates enables to compute an ϵ-subgradient of f_1 at the current x-iterate. Indeed, since $y_l^k \in \mathrm{dom} f_1^*$, by Proposition 3, we have that $y_l^k \in \partial_{\epsilon_l} f_1(x_l^k)$, with $\epsilon_l^k = f_1(x_l^k) - < y_l^k, x_l^k > + f_1^*(y_l^k)$ and the sequence $\{\epsilon_l^k\}$ converges to 0 as $l \to \infty$. Consequently, given $\mu_{k+1} > 0$ and $\nu_{k+1} > 0$, there exists an index $\bar{l}$ such that $\|\eta_{\bar{l}}^k\| \leq \mu_{k+1}$, with $\eta_l^k = \frac{1}{\beta} q_l^k + y_l^k - \tilde{p}^{(k)}$ and $\epsilon_{\bar{l}}^k \leq \nu_{k+1}$; then, we set as approximate solution of the $\tilde{Q}_k$ subproblem $\tilde{x}^{(k+1)} = x_{\bar{l}}^k$, with $\tilde{p}^{(k+1)} = y_{\bar{l}}^k$, $\epsilon_{k+1} = \epsilon_{\bar{l}}^k$ and $q^{(k+1)} = q_{\bar{l}}^k$.
In section 3 we show that, in image restoration problems, when f_1 is the discrete Total Variation function, the Alternating Extragradient Method (AEM) [26] enables to obtain an approximate solution $\tilde{x}^{(k+1)}$ of (22) and its ϵ_{k+1}-subgradient $\tilde{p}^{(k+1)}$; other schemes such as the primal–dual methods in [31] can be used as inner solvers.
From $q^{(k+1)} \in \partial f_0(\tilde{x}^{(k+1)})$ and $\tilde{p}^{(k+1)} \in \partial_{\epsilon_{k+1}} f_1(\tilde{x}^{(k+1)})$ we can compute an ϵ-subgradient of $\tilde{Q}_k(x, \tilde{p}^{(k)}, \epsilon_k)$ at $\tilde{x}^{(k+1)}$, as stated by the following Lemma.

Lemma 1 *Let $f_0(x)$ and $f_1(x)$ be nonnegative, proper, lower semicontinuous and convex functions, with $dom(f_0) \subset dom(f_1)$ and the relative interiors of f_0 and f_1 have at least a point in common. If $q^{(k+1)} \in \partial f_0(\tilde{x}^{(k+1)})$ and $\tilde{p}^{(k+1)} \in \partial_{\epsilon_{k+1}} f_1(\tilde{x}^{(k+1)})$, then the following vector*

$$\eta^{(k+1)} = \frac{1}{\beta} q^{(k+1)} + \tilde{p}^{(k+1)} - \tilde{p}^{(k)} \quad (24)$$

is an ϵ_{k+1}-subgradient of $\tilde{Q}_k$ at $\tilde{x}^{(k+1)}$, that is $\eta^{(k+1)} \in \partial_{\epsilon_{k+1}} \tilde{Q}_k(\tilde{x}^{(k+1)}, \tilde{p}^{(k)}, \epsilon_k)$.

Proof. For the convexity of $f_0(x)$ and the definition of $\tilde{p}^{(k+1)}$ as ϵ_{k+1}-subgradient of f_1 at $\tilde{x}^{(k+1)}$, for any $x \in \mathbb{R}^n$ we have

$$\begin{aligned}
&\tilde{Q}_k(\tilde{x}^{(k+1)}, \tilde{p}^{(k)}, \epsilon_k) + < \eta^{(k+1)}, x - \tilde{x}^{(k+1)} > = \\
&= \frac{1}{\beta} f_0(\tilde{x}^{(k+1)}) + f_1(\tilde{x}^{(k+1)}) - f_1(\tilde{x}^{(k)}) - < \tilde{p}^{(k)}, \tilde{x}^{(k+1)} - \tilde{x}^{(k)} > + \epsilon_k + \\
&+ \ < \frac{1}{\beta} q^{(k+1)} + \tilde{p}^{(k+1)} - \tilde{p}^{(k)}, x - \tilde{x}^{(k+1)} > \leq \\
&\leq \frac{1}{\beta} f_0(x) + f_1(x) + \epsilon_{k+1} - f_1(\tilde{x}^{(k)}) - < \tilde{p}^{(k)}, x - \tilde{x}^{(k)} > + \epsilon_k = \\
&= \tilde{Q}_k(x, \tilde{p}^{(k)}, \epsilon_k) + \epsilon_{k+1}
\end{aligned}$$

Then $\eta^{(k+1)} \in \partial_{\epsilon_{k+1}} \tilde{Q}_k(\tilde{x}^{(k+1)}, \tilde{p}^{(k)}, \epsilon_k)$. □

For the sequence $\{f_0(\tilde{x}^{(k+1)})\}$ generated by the inexact scheme, the monotonicity property does not hold, but it is replaced by

$$\begin{aligned}
\frac{1}{\beta} f_0(\tilde{x}^{(k+1)}) &\leq \tilde{Q}_k(\tilde{x}^{(k+1)}, \tilde{p}^{(k)}, \epsilon_k) = \\
&= \frac{1}{\beta} f_0(\tilde{x}^{(k+1)}) + \Delta_{\epsilon_k}^{\tilde{p}^{(k)}} f_1(\tilde{x}^{(k+1)}, \tilde{x}^{(k)}) \leq \frac{1}{\beta} f_0(\tilde{x}^{(k)}) + \epsilon_k \qquad (25)
\end{aligned}$$

where the first inequality follows from the nonnegativity of $\Delta_{\epsilon_k}^{\tilde{p}^{(k)}} f_1(\tilde{x}^{(k+1)}, \tilde{x}^{(k)})$. Obviously, when f_1 is differentiable, $\epsilon_k = 0$ and the monotonicity property is preserved. From the last inequality in (25), we have

$$-\epsilon_k \leq \Delta_{\epsilon_k}^{\tilde{p}^{(k)}} f_1(\tilde{x}^{(k+1)}, \tilde{x}^{(k)}) - \epsilon_k \leq \frac{1}{\beta}(f_0(\tilde{x}^{(k)}) - f_0(\tilde{x}^{(k+1)}))$$

for any $k \geq 0$. Multiplying this last inequality by k and summing for $i = 1, ..., k-1$, it follows:

$$-\sum_{i=1}^{k-1} i\epsilon_i \leq \frac{1}{\beta} \sum_{i=1}^{k-1} f_0(\tilde{x}^{(i)}) - \frac{k-1}{\beta} f_0(\tilde{x}^{(k)}) \qquad (26)$$

Furthermore, in view of (24), we obtain that (9) can be restated as follows:

$$\begin{aligned}
&\Delta_{\epsilon_k}^{\tilde{p}^{(k)}} f_1(x, \tilde{x}^{(k)}) + \Delta_{\epsilon_{k-1}}^{\tilde{p}^{(k-1)}} f_1(\tilde{x}^{(k)}, \tilde{x}^{(k-1)}) + \frac{1}{\beta} f_0(\tilde{x}^{(k)}) \leq \\
&\leq \frac{1}{\beta} f_0(x) + \Delta_{\epsilon_{k-1}}^{\tilde{p}^{(k-1)}} f_1(x, \tilde{x}^{(k-1)}) + \\
&+ < \eta^{(k)}, \tilde{x}^{(k)} - x > + \epsilon_k \qquad (27)
\end{aligned}$$

for any x such that $f_1(x) < \infty$. This inequality enables us to prove the convergence of the inexact iterative procedure, when a suitable stopping criterion is used to obtain approximate solutions of the inner subproblems.

Proposition 4 *Let $f_0(x)$ and $f_1(x)$ be nonnegative, proper, lower semicontinuous and convex functions, with $dom(f_0) \subset dom(f_1)$ and the relative interiors of f_0 and f_1 have*

at least a point in common. We assume that, for any k, there exists a minimizer of the subproblem (22) and that $\widehat{x}$ is a minimizer of $f_0(x)$ such that $f_1(\widehat{x}) < \infty$. If for any $k \geq 0$ the inner solver determines $\tilde{x}^{(k+1)}$, $q^{(k+1)} \in \partial f_0(\tilde{x}^{(k+1)})$ and $\tilde{p}^{(k+1)} \in \partial_{\epsilon_{k+1}} f_1(\tilde{x}^{(k+1)})$ so that the following condition on $\eta^{(k+1)} = \frac{1}{\beta} q^{(k+1)} + \tilde{p}^{(k+1)} - \tilde{p}^{(k)}$ and ϵ_{k+1} holds

$$\|\eta^{(k+1)}\| \leq \mu_{k+1} \quad \textit{and} \quad \epsilon_{k+1} \leq \nu_{k+1} \tag{28}$$

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

$$\Delta_{\epsilon_k}^{\tilde{p}^{(k)}} f_1(\widehat{x}, \tilde{x}^{(k)}) \leq \Delta_{\epsilon_{k-1}}^{\tilde{p}^{(k-1)}} f_1(\widehat{x}, \tilde{x}^{(k-1)}) + < \eta^{(k)}, \tilde{x}^{(k)} - \widehat{x} > + \epsilon_k \tag{29}$$

and

$$f_0(\tilde{x}^{(k)}) \leq f_0(\widehat{x}) + \frac{\beta}{k} \left(f_1(\widehat{x}) - f_1(\tilde{x}^{(0)}) + \sum_{i=1}^{k} < \eta^{(i)}, \tilde{x}^{(i)} - \widehat{x} > + \sum_{i=1}^{k} (i+1)\epsilon_i \right) \tag{30}$$

Moreover, if the level subsets of f_0 are bounded, a limit point of the sequence $\{\tilde{x}^{(k)}\}$ is a minimizer of $f_0(x)$; if $\widehat{x}$ is the unique minimizer of $f_0(x)$, then $\tilde{x}^{(k)} \to \widehat{x}$ as $k \to \infty$.

Proof. In view of (27) with $x = \widehat{x}$, since for any $k \geq 0$ we have that $\Delta_{\epsilon_k}^{\tilde{p}^{(k)}} f_1(\tilde{x}^{(k+1)}, \tilde{x}^{(k)}) \geq 0$ and $f_0(\tilde{x}^{(k)}) - f_0(\widehat{x}) \geq 0$, the inequality (29) holds.
Summing up the inequalities (27) computed at $\widehat{x}$ related to the first k steps, we have:

$$\begin{aligned} &\Delta_{\epsilon_k}^{\tilde{p}^{(k)}} f_1(\widehat{x}, \tilde{x}^{(k)}) + \sum_{i=1}^{k} \Delta_{\epsilon_{i-1}}^{\tilde{p}^{(i-1)}} f_1(\tilde{x}^{(i)}, \tilde{x}^{(i-1)}) + \frac{1}{\beta} \sum_{i=1}^{k} f_0(\tilde{x}^{(i)}) \leq \\ &\leq \frac{k}{\beta} f_0(\widehat{x}) + (f_1(\widehat{x}) - f_1(\tilde{x}^{(0)})) + \sum_{i=1}^{k} < \eta^{(i)}, \tilde{x}^{(i)} - \widehat{x} > + \sum_{i=1}^{k} \epsilon_i \end{aligned} \tag{31}$$

Combining (26) with (31), we obtain

$$\begin{aligned} &\frac{\beta}{k} \left(\Delta_{\epsilon_k}^{\tilde{p}^{(k)}} f_1(\widehat{x}, \tilde{x}^{(k)}) + \sum_{i=1}^{k} \Delta_{\epsilon_{i-1}}^{\tilde{p}^{(i-1)}} f_1(\tilde{x}^{(i)}, \tilde{x}^{(i-1)}) \right) + \\ &+ f_0(\tilde{x}^{(k)}) \leq f_0(\widehat{x}) + \frac{\beta}{k} \Big(f_1(\widehat{x}) - f_1(\tilde{x}^{(0)}) + \\ &\sum_{i=1}^{k} < \eta^{(i)}, \tilde{x}^{(i)} - \widehat{x} > + \sum_{i=1}^{k-1} (i+1)\epsilon_i + \epsilon_k \Big) \end{aligned} \tag{32}$$

Since $\Delta_{\epsilon_{i-1}}^{\tilde{p}^{(i-1)}} f_1(\tilde{x}^{(i)}, \tilde{x}^{(i-1)}) \geq 0$ for any i and $\Delta_{\epsilon_k}^{\tilde{p}^{(k)}} f_1(\widehat{x}, \tilde{x}^{(k)}) \geq 0$, we have that (30) follows. Furthermore, if we denote by D the diameter of the level set $\{x | f_0(x) \leq f_0(\tilde{x}^{(0)})\}$, by applying the Cauchy–Schwarz inequality and condition (28) to inequality (30), we obtain

$$f_0(\tilde{x}^{(k)}) \leq f_0(\widehat{x}) + \beta \frac{f_1(\widehat{x}) - f_1(\tilde{x}^{(0)})}{k} + \frac{\beta}{k} \left(D \sum_{i=1}^{k} \mu_i + \sum_{i=1}^{k} (i+1)\nu_i \right) \tag{33}$$

Since the sequence $\{\tilde{x}^{(k)}\}$ is bounded, there exists a subsequence of $\{\tilde{x}^{(k)}\}$ convergent to a limit point $\tilde{x}$. Since $D \sum_{i=1}^{\infty} \mu_i + \sum_{i=1}^{\infty} (i+1)\nu_i < \infty$, in view of (33), we have $f_0(\tilde{x}) \leq f_0(\widehat{x})$

for $k \to \infty$. Then, $\tilde{x}$ is a minimizer of $f_0(x)$. If $\widehat{x}$ is the unique minimizer of $f_0(x)$, then $\tilde{x}^{(k)} \to \widehat{x}$ as $k \to \infty$. □

Corollary 1 *Let f_1 be a differentiable function. Under the same hypotheses of the previous proposition, if for any $k \geq 0$ the inner solver determines $\tilde{x}^{(k+1)}$, $q^{(k+1)} \in \partial f_0(\tilde{x}^{(k+1)})$ and $\tilde{p}^{(k+1)} = \nabla f_1(\tilde{x}^{(k+1)})$ so that the following condition on $\eta^{(k+1)} = \frac{1}{\beta} q^{(k+1)} + \tilde{p}^{(k+1)} - \tilde{p}^{(k)}$ holds*

$$\|\eta^{(k+1)}\| \leq \mu_{k+1} \tag{34}$$

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

$$D^{\tilde{p}^{(k)}} f_1(\widehat{x}, \tilde{x}^{(k)}) \leq D^{\tilde{p}^{(k-1)}} f_1(\widehat{x}, \tilde{x}^{(k-1)}) + < \eta^{(k)}, \tilde{x}^{(k)} - \widehat{x} > \tag{35}$$

and

$$f_0(\tilde{x}^{(k)}) \leq f_0(\widehat{x}) + \frac{\beta}{k} \left(f_1(\widehat{x}) - f_1(\tilde{x}^{(0)}) + \sum_{i=1}^{k} < \eta^{(i)}, \tilde{x}^{(i)} - \widehat{x} > \right) \tag{36}$$

Moreover, if the level subsets of f_0 are bounded, a limit point of the sequence $\{\tilde{x}^{(k)}\}$ is a minimizer of $f_0(x)$; if $\widehat{x}$ is the unique minimizer of $f_0(x)$, then $\tilde{x}^{(k)} \to \widehat{x}$ as $k \to \infty$.

Remark. From the numerical point of view, an easily implementable choice for the tolerance in the criterion (28) (or (34)) is

$$\|\eta^{(k+1)}\| \leq \frac{c}{(k+1)^{\alpha}} \tag{37}$$

$$\epsilon_{k+1} \leq \frac{d}{(k+1)^{\vartheta}} \tag{38}$$

for $k \geq 0$, where c, d, α and ϑ are positive constants, with $\alpha > 1$ and $\vartheta > 2$. In this way, at the first iteration the tolerances are equal to the parameters c and d respectively and, in the subsequent iterations, the stopping rule is gradually more severe; the parameters α and ϑ control the increase of the inner accuracy. A practical rule to choose the values of c and d is to use a standard stopping criterion with a moderate tolerance in the inner solver at the first outer iteration; then set $c = \|\eta^{(1)}\|$, $d = \epsilon_1$, $\vartheta = 2.1$; in this way, the only input parameter is α. In the section 4, we implement this rule.

For this inexact scheme, we can repeat the considerations concerning the image restoration problems, where the discrepancy $f_0(g; x)$ from the detected data g plays the role of penalty function and $f_1(x)$ is the regularization term. We are interested to a solution x^* of $f_0(f; x) = 0$, where f is the noisy-free data, rather than $f_0(g; x) = 0$. We assume that an estimate of a noise level γ is known, that is $f_0(g; x^*) \leq \gamma$. Then, as long as $f_0(g; x^{(k)}) \geq \gamma$, from (27) with $x = x^*$, the Cauchy–Schwarz inequality and the

boundedness of the level set of $f_0(g;x)$, we have

$$-\nu_k - \mu_k D + \Delta_{\epsilon_k}^{\tilde{p}^{(k)}} f_1(x^*, \tilde{x}^{(k)}) + \Delta_{\epsilon_{k-1}}^{\tilde{p}^{(k-1)}} f_1(\tilde{x}^{(k)}, \tilde{x}^{(k-1)}) + \frac{1}{\beta} f_0(g;\tilde{x}^{(k)}) \leq \tag{39}$$

$$\leq \frac{\gamma}{\beta} + \Delta_{\epsilon_{k-1}}^{\tilde{p}^{(k-1)}} f_1(x^*, \tilde{x}^{(k-1)}) \tag{40}$$

Then a sufficient condition to assure a decreasing behavior for the *inexact* Bregman distance $\Delta_{\epsilon_{k-1}}^{p^{(k-1)}} f_1(x^*, \tilde{x}^{(k-1)})$ is that, while $f_0(g;x^{(k)}) \geq \gamma$ the following inequality holds:

$$\frac{1}{\beta}(f_0(g;\tilde{x}^{(k)}) - \gamma) + \Delta_{\epsilon_{k-1}}^{\tilde{p}^{(k-1)}} f_1(\tilde{x}^{(k)}, \tilde{x}^{(k-1)}) = \tilde{Q}_{k-1}(\tilde{x}^{(k)}, \tilde{p}^{(k-1)}, \epsilon_{k-1}) - \frac{\gamma}{\beta} \geq \mu_k D + \nu_k \tag{41}$$

This condition depends on the *exactness* level required to the inner minimizer: the numerical experience shows that the term $-\mu_k D - \nu_k$ is a pessimistic lower bound of $< \eta^{(k)}, x^* - \tilde{x}^{(k)} > -\epsilon_k$. Indeed in the numerical experiments of Section 4 we observe that, when f_1 is a differentiable function, $\Delta_{\epsilon_k}^{p^{(k)}} f_1(x^*, \tilde{x}^{(k)}) \equiv D^{p^{(k)}} f_1(x^*, \tilde{x}^{(k)})$ is a decreasing sequence until the relative reconstruction error decreases and sometimes ever later. When f_1 is a non differentiable function, $\Delta_{\epsilon_k}^{\tilde{p}^{(k)}} f_1(x^*, \tilde{x}^{(k)})$ is a decreasing sequence with a behavior very similar to a Bregman distance. In Appendix B we report the behavior of the exact and the inexact Bregman iteration on a simple example concerning a 1D test problem for which the true analytic solution is known.

As for the exact Bregman method, a discrepancy criterion can provide a reasonable stopping criterion. Nevertheless, the numerical experience shows that in general in few iterations we observe a semi-convergence behavior; then a practical criterion is to visually control the obtained approximation at any iteration.

3. Image restoration for Poisson data

In recent years, the image restoration from data corrupted by Poisson noise has received a considerable attention. Poisson noise occurs in all imaging processes where images are obtained by means of the count of particles, in general photons, arriving in the image domain (see [32] for a review). For data corrupted by Poisson noise, the problem (1) can be formulated as follows:

$$\min_{x \in \mathbb{R}^N} \sum_i \left(g_i \log \frac{g_i}{(Hx+b)_i} + (Hx+b)_i - g_i \right) + i_C(x) + \beta f_1(x) \tag{42}$$

where the data fidelity function f_0 is the sum of the generalized Kullback-Leibler (KL) divergence and the indicator function $i_C(x)$ of the nonnegative orthant C; the regularization term $f_1(x)$ can be selected according to the features of the application. Here b is a nonnegative constant background term. When $H = I$ we have a denoising problem, while, in the other cases, we deal with a deblurring problem. We assume that the imaging matrix H has nonnegative entries and it satisfies a normalization condition $H^T e = e$, where e is an array with all components equal to 1.

In the following we state the conditions for the convergence of the inexact iterative procedure for the KL function combined with different regularization terms.

First of all, we assume to consider proper, lower semicontinuous and convex functions as regularization term, as the quadratic penalization

$$f_1(x) = \|x\|^2 \tag{43}$$

where $\|\cdot\|$ is the usual Euclidean norm of a vector, the quadratic discrete gradient ∇

$$f_1(x) = \|\nabla x\|^2 \tag{44}$$

and, in the edge–preserving regularization framework, the discrete Total Variation (TV) functional [3]

$$f_1(x) = \sum_i \left\| (\nabla x)_i \right\| \tag{45}$$

or its smoothed version, known also as hypersurface (HS) potential [33]

$$f_1(x) = \sum_i \left\| \begin{pmatrix} (\nabla x)_i \\ \delta \end{pmatrix} \right\| \tag{46}$$

Since the KL function is proper, convex and lower semicontinuous on the nonnegative orthant, the subproblems (22) are proper, convex and lower semicontinuous. Furthermore, under the previous assumption on the imaging matrix H, $f_0(x)$ is a coercive function (see [25, 7]). As a consequence, the level sets of $f_0(x)$ are bounded.

Proposition 5

(a) For $f_1(x) = \|x\|^2$, the solution of any subproblem (22) exists and is unique.

(b) For $f_1(x)$ given by (44) or by HS potential (46), if H has nonnegative entries and $H^T e = e$, the solution of any subproblem (22) exists and is unique.

(c) For $f_1(x)$ given by the discrete TV (45), if H has nonnegative entries and $H^T e = e$, the solution of any subproblem (22) exists; the uniqueness of the solution is guaranteed if null$(H) = 0$ and the component of g are positive.

Proof.

(a) For any $\beta > 0$, the Hessian matrix of any subproblem (22) is positive definite on the domain. Then (22) is strictly convex and coercive.

(b) The proof is based on the observation that the intersection of the space of the constant images with the null space of H is just the zero vector. Furthermore the Hessian matrix of any subproblem (22) is positive definite on the domain (see [34]).

(c) See [25, Prop. 3].

□

When $f_1(x)$ is differentiable, the approximate solution $\tilde{x}^{(k+1)}$ of the subproblem (22) can be obtained by efficient differentiable optimization methods. 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 [20, 35]. This projected gradient-type method is based on a modification of the gradient direction

by a suitable positive definite and diagonal scaling matrix and an accurate strategy of steplength selection, based on an adaptive alternation of the Barzilai–Borwein rules [36, 37]. 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 monotone line-search along the projected direction require only linear operations. For details, see [20, 35].
For image restoration by TV regularization from Poisson data, several solvers have been recently proposed (see for example [25, 24, 23, 26, 27]). A suitable choice for the solver of subproblems (22) is the alternating extragradient method (AEM), since it does not involve the solution of an inner ROF model nor requires the solution of a system and it can be used both for denoising and deblurring problems. The global convergence of the scheme and an estimate of its convergence rate require only local Lipschitz continuity of the gradient of the primal-dual formulation of the objective function. Indeed, AEM is a first order scheme for the primal-dual formulation of the model KL-TV or its variant such as (23), that requires only matrix-vector products. In particular, AEM is an especially tailored extragradient–type method, based on three successive alternating (or Gauss-Seidel) projections, whose stepsize parameter α_l is adaptively computed, without requiring the knowledge of a Lipschitz constant. Then, given the primal-dual iterate (x_l, y_l), the $(l+1)$-th iteration of AEM is given by the following step:

$$\begin{aligned}
\tilde{y}_l &= \operatorname{argmin}_{y\in \operatorname{dom} f_1^*} \frac{1}{2}\|y - (y_l + \alpha_l \nabla x_l)\|^2 \\
x_{l+1} &= \operatorname{argmin}_{x\in \operatorname{dom} f_0 \cap \operatorname{dom} f_1} \frac{1}{2}\|x - (x_l + \alpha_l \nabla^T \tilde{y}_l)\|^2 \\
y_{l+1} &= \operatorname{argmin}_{y\in \operatorname{dom} f_1^*} \frac{1}{2}\|y - (y_l + \alpha_l \nabla x_{l+1})\|^2
\end{aligned} \tag{47}$$

From Proposition 3, since $y_{l+1} \in \operatorname{dom} f_1^*$, $\nabla^T y_{l+1}$ is an ϵ_{l+1}-subgradient of $f_1 \circ \nabla$ at x_{l+1}, with $\epsilon_{l+1} = (f_1 \circ \nabla)(x_{l+1}) - < y_{l+1}, \nabla x_{l+1} >$. In this case, f_1^* is the indicator function of the domain of the dual variable. Then, at the $(k+1)$-th outer iteration, if $\bar{l}$ is the iteration for which the conditions in (28) (or (37)-(38)) are satisfied, the outer iterate $\tilde{x}^{(k+1)}$ is set equal to $x_{\bar{l}}$, $\nabla^T y_{\bar{l}} = \tilde{p}^{(k+1)}$ is its ϵ_{k+1}-subgradient and $q^{(k+1)} = \nabla f_0(x^{(k+1)})$. For details on AEM, see [26].

4. Numerical experiments

This section is devoted to numerically evaluate the effectiveness of the inexact procedure based on Bregman iteration for recovering images corrupted by Poisson noise. The numerical experiments described in this section have been performed in MATLAB environment, on a server with a dual Intel Xeon QuadCore E5620 processor at 2,40 GHz, 12 Mb cache and 18 Gb of RAM. In the experiments we consider a set of test problems, where the simulated data are obtained by convolving the image with a Point Spread Function (PSF) and then by perturbing with Poisson noise; the Poisson noise

has been simulated by the `imnoise` function in the Matlab Image Processing Toolbox. The considered test problems are described in the following.

Deblurring test problems

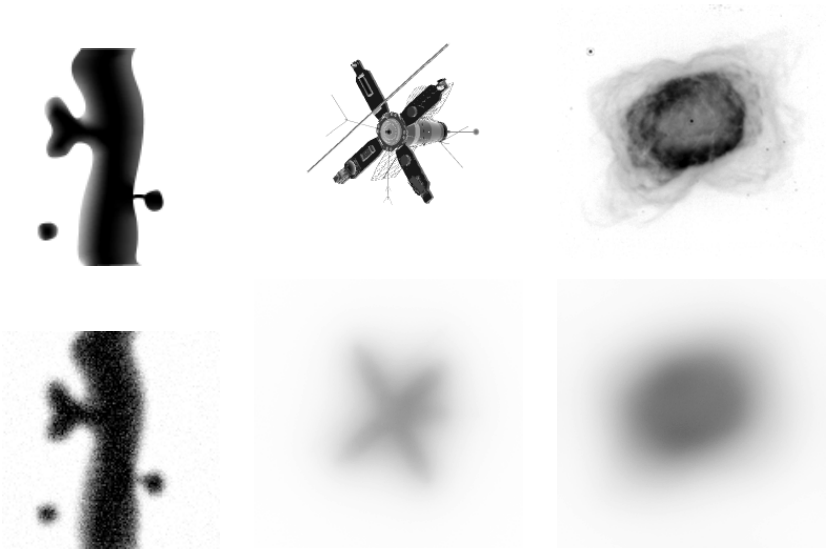
- *micro*: the original image is a phantom of size 128×128 described in [38]; its values are in the range $[1, 69]$ and the total flux is $2.9461\ 10^5$; the background term b is set to zero; to obtain the simulated data, a Gaussian PSF is used with standard deviation equal to $\sqrt{5}$ pixels in vertical and horizontal directions; the original and the simulated images are in figure 1 (a).
- *spacecraft*: the original image is a 256×256 image with sharp details, whose values are in the range $[0, 255]$ and the background term b is set to 1; following [9], the PSF used simulates that taken by a ground-based telescope and is downloaded from *http://www.mathcs.emory.edu/nagy/RestoreTools/index.html*; this test problem is denoted with *L-spacecraft*; a second test problem, named *H-spacecraft*, is generated by multiplying the object and the background by 100; then the image is convolved with the same PSF and perturbed with Poisson noise (see figure 1 (b));
- *NGB 7027*: the 256×256 original image is an example of a diffuse astronomical object; as for *H-spacecraft*, its values are in the range $[0, 25500]$ and the background term b is set to 100; the simulated data are obtained with the same PSF of *H-spacecraft* and then perturbed with Poisson noise (see figure 1 (c)).

Denoising test problem

- LCR phantom: the original image is obtained from the phantom described in [39]; it is an array 256×256, consisting in concentric circles of intensities 14, 27, 40, enclosed by a square frame of intensity 2, all on a background of intensity 1. The relative difference in Euclidean norm between the noisy image and the original one is 0.21273.

4.1. Efficiency of the inexact versus the exact procedure

In this section, we report some numerical tests, showing that the inexact version of the iterative procedure appears promising from the point of view of the efficiency. For sake of brevity, we report the results obtained in the case of *L-spacecraft*, but similar results have been obtained for the other deblurring test-problems.
In order to restore *L-spacecraft*, we consider the minimization of the KL function combined with the HS regularization (KL-HS model) and with the TV regularization (KL-TV model). The value of the regularization parameter related to the minimum relative reconstruction error with respect to the original image x^* is obtained experimentally by several run for both models and it is set as $\beta_{opt} = 1.63 \cdot 10^{-3}$. When the KL-HS model is solved with β_{opt} by SGP ($\delta = 0.1$), after 497 iterations we obtain a relative reconstruction error in Euclidean norm equal to 0.36572; in this case SGP is stopped when the relative difference between two consecutive values of the objective



(a) *micro* (b) *H-spacecraft* (c) *NGB 7027*

Figure 1. Deblurring problems: the original images are in the upper panels, while the blurred images are in the lower panels. All the images are shown in reverse gray scale.

function is less than $tol_{SGP} = 10^{-7}$ and the mean of this difference over the last 10 iterations is less than $10 \cdot tol_{SGP}$. For the KL-TV model with β_{opt}, AEM enables to obtain a relative reconstruction error in Euclidean norm equal to 0.36967 after 2042 iterations. AEM is stopped when the relative difference in Euclidean norm between two successive iterates is less than $tol_{AEM} = 4 \cdot 10^{-5}$.

In the first numerical experiment, the KL-HS and KL-TV models are solved by the exact and inexact iterative procedure, using $\beta = 10\beta_{opt}$ as regularization parameter and $x^{(0)} = \frac{\sum_i g_i}{m} - b$ as starting point; for the KL-HS model, the inner solver is SGP in both versions of the procedure, while for the model KL-TV, AEM is used. The Matlab codes of SGP and AEM are adapted versions of those downloadable from *http://www.unife.it/prin/software*. For the exact version, the inner subproblems Q_k in (6) have to be solved accurately enough to make sure that the update rule (7) gives $p^{(k+1)} \in \partial f_1(x^{(k+1)})$. If this condition is not satisfied, the nonnegativity of $D^{p^{(k+1)}} f_1(x, x^{(k+1)})$ is not assured and the inequality (8) may not apply. Then, for the exact iterative procedure, the previous described standard stopping criteria of SGP and AEM are used with a severe tolerance, given by $tol_{SGP} = 10^{-10}$ and $tol_{AEM} = 10^{-5}$ respectively.

For the inexact version, in the case of the KL-HS model the SGP inner solver uses (37) as stopping rule while for the KL-TV model AEM is stopped by (37)-(38). For the setting of the parameters c and d, following the Remark in section 2.3, in both cases, the first subproblem is solved by the standard stopping rules of SGP or AEM respectively (with moderate tolerance, i.e. $tol_{SGP} = 10^{-7}$ and $tol_{AEM} = 5 \cdot 10^{-4}$) and, then, we set

Table 1. *L-spacecraft* deblurring test problem: exact and inexact iterative methods for the KL-HS model ($\delta = 0.0134$), using SGP as inner solver. In the exact version $tol_{SGP} = 10^{-10}$; in the inexact version, in the first outer iteration $tol_{SGP} = 10^{-7}$ while in the subsequent iterations (37) is used as stopping rule with $c = 5.38$ and $\alpha = 1.5$. *time* denotes the execution time in seconds at the end of the current outer iteration.

	Exact iterative procedure-SGP			Inexact iterative procedure-SGP		
k	ρ_k	*it*	*time*	ρ_k	*it*	*time*
1	0.4880	4716	157.7	0.5001	1061	35.8
2	0.4012	7500	408.2	0.4265	4713	201.0
3	0.3801	2969	506.9	0.3936	3420	322.1
4	0.3647	7015	744.3	0.3779	3304	440.1
5	**0.3644**	2609	831.1	0.3697	3763	572.3
6	0.3655	3073	933.9	**0.3681**	2452	660.1
7	0.3696	3888	1064.4	0.3691	3153	771.9
8	0.3746	2907	1160.2	0.3735	2714	868.0

Table 2. *L-spacecraft* deblurring test problem: exact and inexact iterative methods for the KL-TV model, using AEM as inner solver. In the exact version $tol_{AEM} = 10^{-5}$; in the inexact version, in the first outer iteration $tol_{AEM} = 5 \cdot 10^{-4}$ while in the subsequent iterations (37)-(38) are used as stopping rules with $c = 19.1$, $d = 65.4$, $\vartheta = 2.1$ and $\alpha = 1.5$. *time* denotes the execution time in seconds at the end of the current outer iteration.

	Exact iterative procedure-AEM			Inexact iterative procedure-AEM		
k	ρ_k	*it*	*time*	ρ_k	*it*	*time*
1	0.47727	5373	206.4	0.54087	245	12.2
2	0.40557	5409	425.6	0.42557	1022	54.4
3	0.37766	4775	611.9	0.37946	2708	162.7
4	0.36610	4157	777.8	0.3682	2933	279.2
5	**0.36414**	3363	910.9	**0.36682**	2971	397.8
6	0.36518	2392	1003.3	0.36916	2828	511.9
7	0.36869	2625	1105.6	0.37127	2995	629.2
8	0.37258	1991	1184.6	0.37483	2631	736.0
9	0.37783	2118	1268.6	0.37997	3234	865.1

$c = \|\eta^{(1)}\|$ and $d = \epsilon_1$; furthermore we set $\vartheta = 2.1$.

In tables 1 and 2 we show the different behavior of the two versions of the iterative procedure. For any outer iteration k, we report the number of iterations *it* of the inner solver (SGP or AEM), the execution time *time* in seconds at the end of the current outer iteration and the relative reconstruction error $\rho_k = \frac{\|z^{(k)} - x^*\|}{\|x^*\|}$ with respect to the original image x^*, where $z^{(k)}$ is the outer k-th iterate $x^{(k)}$ or $\tilde{x}^{(k)}$ respectively.

With regard to the KL-HS model, table 1 shows that, for the exact iterative procedure, the minimum reconstruction error is obtained with 5 outer iterations while, for the

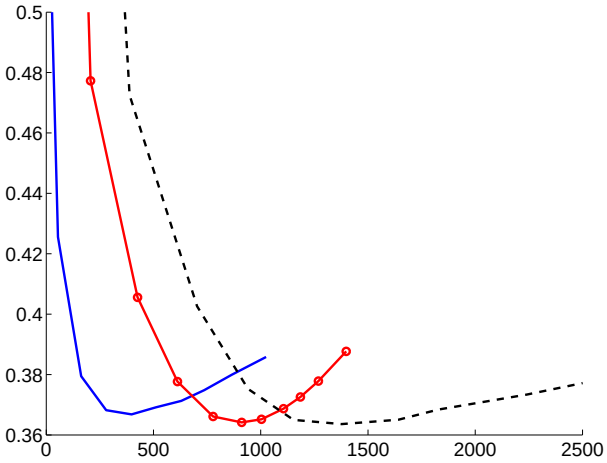


Figure 2. Test problem *L-spacecraft*: plot of the relative reconstruction error versus the execution time in seconds for the exact iterative procedure with EM-TV as inner solver (dashed line) or AEM (solid line with circle markers); the solid line is relative to the inexact version with AEM as inner solver.

inexact version, 6 iterations are necessary; nevertheless the inexact method allows to determine the restored image in a shorter time; indeed the total number of inner iterations for 5 outer iterations of the exact version is equal to 24809 against 18713 for 6 outer iterations of the inexact scheme.

In a similar way, as regards the KL-TV model, table 2 shows that, for the exact iterative procedure, the minimum reconstruction error is obtained with 5 outer iterations and 23077 total inner iterations; even for the inexact version, 5 outer iterations are necessary but only 9879 inner iterations are required.

In figure 2 we show the behavior of the reconstruction error ρ_k with respect to the execution time for the inexact and exact versions of the iterative procedure with AEM as inner solver and for the exact scheme with EM-TV as inner solver [6]. We remember that the Bregman iteration combined with EM-TV has a very complex structure since three iterative methods are nested one inside the other. For the implementation of EM-TV combined with Bregman iteration we refer to [6]; for any k-th outer iteration, we execute 1500 inner iterations of EM-TV. It should be noted that, since we are not able to find an inner stopping criterion for which the inequality (8) is verified, we have determined experimentally a minimum prefixed number of inner iterations assuring an approximately correct behavior for the KL and the objective function of the subproblems. Furthermore the inner step of EM-TV solver uses the Chambolle method [40], that is stopped when the maximum difference between two successive dual iterates is less than 10^{-2}.

Figure 2 shows that the inner stopping rules (37)-(38) of the inexact scheme increase the efficiency of Bregman iteration without affecting its features.

To deepen the meaning of the stopping rules in the inexact iterative procedure, we solve *L-spacecraft* with both models, using different value of α; the setting of the others

Table 3. *L-spacecraft* deblurring test problem: results obtained by the inexact iterative method with different values of α in the inner stopping rule. For the KL-HS model, the inner solver is SGP, while for KL-TV model, AEM is used.

α	k	ρ_k	*cum–it*
KL-HS model			
1.2	7	0.37039	20849
1.5	6	0.36809	18713
1.7	6	0.36736	22030
KL-TV model			
1.5	5	0.36682	9879
3	5	0.36721	12196
4	4	0.36665	16634

parameters is the same used for the experiments related to tables 1 and 2 with the inner tolerance for the first subproblem of the inexact procedure equal to $tol_{SGP} = 10^{-7}$ and $tol_{AEM} = 5 \cdot 10^{-4}$ respectively; in the subsequent outer iterations, the stopping rules (37)-(38) become gradually more severe. The parameter α allows to adjust how quickly must increase the accuracy required in a inner subproblem. The results of these numerical experiments are reported in table 3. For each value of α, we report the iteration k corresponding to minimum reconstruction error ρ_k, the value of ρ_k and the cumulative number of inner iterations *cum–it* performed in k outer iterations. As shown in table 3, a suitable choice of α corresponds to require a not too severe accuracy after the first 5-6 iterations and, in any case, the inexact version appears to be more efficient than the exact one in terms of inner iterations.

4.2. Behavior of the inexact procedure with overestimated regularization parameter: edge–preserving regularization

In this section, we describe some numerical tests for verifying the semi-convergence property of the inexact scheme. In particular, we consider the deblurring test problems *micro* and *H-spacecraft*; in both cases, we consider the models KL-HS and KL-TV. As in the previous section, the value of the regularization parameter related to the minimum relative reconstruction error with respect to the original image x^* is obtained experimentally, by using SGP method for the KL-HS model and AEM for the KL-TV model. As regards *micro*, when we solve the KL-HS model with $\delta = 10^{-2}$ by SGP, the minimum relative reconstruction error, equal to 0.0898, is obtained for $\beta_{opt} = 0.09$ in 1128 iterations. SGP method is stopped when the relative difference between two consecutive values of the objective function is less than $tol_{SGP} = 10^{-7}$ and the mean of this difference over the last 10 iterations is less than $10 \cdot tol_{SGP}$. When we solve the KL-TV model by AEM, the reconstruction error is equal to 0.0903 for $\beta_{opt} = 0.09$ and the number of iterations is 1728. The stopping criterion for AEM is that the relative difference in Euclidean norm between two successive iterates is less than $tol_{AEM} = 10^{-5}$.

Table 4. Inexact iterative method for deblurring problems: model KL-HS.

k	$D^{\tilde{p}^{(k)}} f_1(x^*, \tilde{x}^{(k)})$	ρ_k	it
micro; $tol_{SGP} = 10^{-5}$, $c = 14.1$, $\alpha = 2$			
1	2690.7	0.23227	474
2	2380.8	0.08939	1415
3	2320.5	0.08229	434
4	2292.3	**0.08046**	352
5	2286.0	0.08073	483
6	2283.9	0.08109	203
7	2292.2	0.08169	370
8	2312.2	0.08266	416
H-spacecraft; $tol_{SGP} = 10^{-7}$, $c = 39.6$, $\alpha = 1.5$			
1	2288 10^4	0.27868	1433
2	2199 10^4	0.26633	1397
3	2162 10^4	0.26077	1518
4	2140 10^4	0.25898	1086
5	2125 10^4	0.25803	1593
6	2115 10^4	**0.25791**	1411
7	2109 10^4	0.25819	1494
8	2107 10^4	0.25911	1640
9	2109 10^4	0.26030	1404

As regards *H-spacecraft*, the solution of the model KL-HS with $\delta = 1$ is obtained by SGP with a miminum relative reconstruction error equal to 0.2657, for $\beta_{opt} = 2.977\ 10^{-5}$ in 1196 iterations and the same stopping criterion of the previous test problem. For the model KL-TV, AEM allows to obtain a relative reconstruction error equal to 0.2663 with 6880 iterations (with β_{opt} and the same stopping criterion of the previous test problem).
In order to evaluate the semi-convergence behavior of the inexact scheme, we solve *micro* and *H-spacecraft* by the inexact procedure. The results obtained for the model KL-HS are shown in table 4, while table 5 shows the results for the KL-TV model. As described for the experiments in the previous section, at the first outer iteration the inner solvers (SGP and AEM) use standard stopping criteria with a weak tolerance; in particular, for the subsequent outer iterations, SGP uses the stopping criterium (37) with $c = \|\eta^{(1)}\|$ and, in the case of AEM, the stopping criteria are (37)-(38) with $c = \|\eta^{(1)}\|$, $d = \epsilon_1$ and $\vartheta = 2.1$.
The value of β is put in both cases equal to $10\beta_{opt}$ and the starting iterate is $\tilde{x}^{(0)} = \frac{\sum_i g_i}{m} - b$. In table 4 and 5 we report for each test problem the behavior of the relative reconstruction error ρ_k at any outer iteration of the inexact procedure and the value of Bregman distance $D^{\tilde{p}^{(k)}} f_1(x^*, \tilde{x}^{(k)})$ or inexact Bregman distance $\Delta^{\tilde{p}^{(k)}}_{\epsilon_k} f_1(x^*, \tilde{x}^{(k)})$. The symbol it denotes the number of iterations of the inner solver for each k–th outer iteration.
Tables 4-5 provide a numerical evidence of the semi-convergence of the inexact scheme

Table 5. Inexact iterative method for deblurring problems: model KL-TV.

k	$\Delta_{\epsilon_k}^{\tilde{p}^{(k)}} f_1(x^*, \tilde{x}^{(k)})$	ρ_k	it
micro; $tol_{AEM} = 10^{-4}$, $c = 0.76$, $d = 19.5$, $\alpha = 1.2$			
1	3467.5	0.16845	491
2	2952.1	0.09991	796
3	2782.2	0.09495	739
4	2711.5	**0.09319**	994
5	2676.9	0.09358	938
6	2658.8	0.09438	1013
7	2653.6	0.09512	1200
8	2655.3	0.09565	1209
H-spacecraft; $tol_{AEM} = 5\ 10^{-5}$, $c = 101.9$, $d = 775.0$, $\alpha = 1.5$			
1	2409 10^4	0.36301	2001
2	2293 10^4	0.32202	4394
3	2236 10^4	0.30739	4573
4	2192 10^4	0.29659	5813
5	2162 10^4	0.28828	6946
6	2140 10^4	0.28191	7555
7	2122 10^4	0.27719	7743
8	2107 10^4	0.27346	8193
9	2095 10^4	0.27041	8789
10	2086 10^4	0.26787	9545
11	2078 10^4	0.26571	10444
12	2072 10^4	0.26383	11983
13	2067 10^4	0.26220	15000
14	2064 10^4	0.26130	11300
15	2063 10^4	0.26057	12790
16	2062 10^4	0.26007	13617
17	2063 10^4	0.25980	15000
18	2065 10^4	**0.25974**	15000
19	2067 10^4	0.25990	15000
20	2070 10^4	0.26025	15000

when an overestimation of the regularization parameter is used. In general few iterations enable us to obtain satisfactory restored images.

For test problem *micro*, in figure 3(a), we compare the results obtained by solving the model KL-HS with the inexact iterative method and those obtained by SGP with β_{opt}; in particular we report an *log-error* image, computed pixel by pixel with the rule $\log |{x_{ij}}^* - \tilde{x}_{ij}^{(k)}|$, with $\log 0 = 0$. We introduce this graphical representation of the absolute error because the values of ρ_k are often similar and in the restored images the differences are not evident; the log operator can highlight the small absolute errors.

Figure 3(a) shows the *log-error* images obtained by the inexact method that corresponds to the minimum reconstruction error ($k = 4$, $\rho_4 = 0.08046$) and the one obtained by SGP (relative reconstruction error equal to 0.0898, with 1128 iterations, as reported

Table 6. Test problem *micro*: minimum errors versus iterations of the inexact iterative method with different values of β.

	KL-HS model				KL-TV model			
β	k	ρ_k	k	$D^{\tilde{p}^{(k)}} f_1(x^*, \tilde{x}^{(k)})$	k	ρ_k	k	$\Delta^{\tilde{p}^{(k)}}_{\epsilon_k} f_1(x^*, \tilde{x}^{(k)})$
$10\beta_{opt}$	4	0.08046	6	2283.9	4	0.09319	7	2653.6
$15\beta_{opt}$	7	0.08281	9	2446.8	6	0.09372	11	2653.4
$20\beta_{opt}$	11	0.08926	13	2639.0	9	0.09377	14	2655.3

above). In figure 3(b) analogous results for the model KL-TV are reported; we show the *error* images obtained by the inexact iterative procedure with the minimum reconstruction error (iteration $k = 4$, $\rho_4 = 0.09319$) and by AEM with $\beta_{opt} = 0.09$ (relative reconstruction error equal to 0.09033, with 1728 iterations). For the images obtained by the inexact method, the error on the background is more uniform and we observe an enhancement of the edges.

To highlight the features of the obtained reconstructions, in figure 4 we show the contour plots of the original image *micro* (levels 1, 20, 40, 60, 67), the restored images obtained by solving the KL-HS model with SGP and with the inexact iterative method at the iteration $k = 4$ (minimum reconstruction error). Analogous results are obtained for the model KL-TV. We observe that the inexact procedure provides an improvement of the contrast.

Figure 5 shows the *log-error* images related to *H-spacecraft*. For the model KL-HS, $k = 8$ is the iteration of the inexact method with minimum value of $D^{\tilde{p}^{(k)}}(x^*, \tilde{x}^{(k)})$. For the model KL-TV, $k = 16$ is the iteration of the inexact method with minimum value of $\Delta^{\tilde{p}^{(k)}}_{\epsilon_k}(x^*, \tilde{x}^{(k)})$; in this case, from the 17-th outer iteration onwards, the prefixed maximum number of iterations (15000) is reached by the inner solver, but at the iteration $k = 16$, ρ_k is less than the reconstruction error obtained by AEM with β_{opt} (0.2663). Also for this test problem, we observe from the *log-error* maps a better attenuation of the background noise with respect to the images obtained with β_{opt}.

Table 6 enables us to investigate how the estimation of the regularization parameter affects the efficiency of the iterative scheme. For test problem *micro*, we report the number of outer iterations at which we obtain the minimum relative reconstruction error for both models and the minimum value of $D^{\tilde{p}^{(k)}} f_1(x^*, \tilde{x}^{(k)})$ or of $\Delta^{\tilde{p}^{(k)}}_{\epsilon_k} f_1(x^*, \tilde{x}^{(k)})$ when the inexact method is applied with different values of β. In particular we observe that the number of outer iterations needed to obtain the restored image increases with decreasing the accuracy of the parameter estimate.

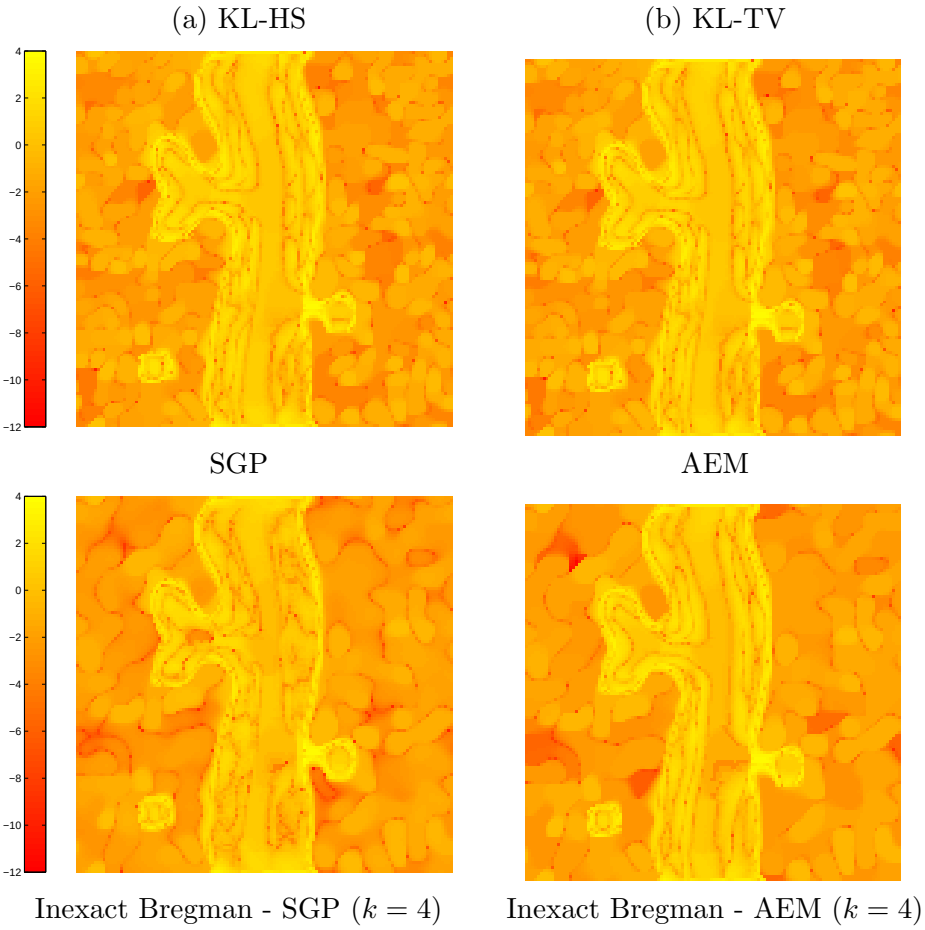


Figure 3. Test problem *micro*: *log-error* images for the model KL-HS (left-panels) and KL-TV (right-panels). In the upper panels SGP and AEM are used with an optimal regularization parameter, while the lower panels show the results obtained by the inexact iterative method at the iteration with the minimum reconstruction error. The level scale is the same for all images.

4.3. Behavior of the inexact procedure with overestimated regularization parameter: quadratic regularization

In order to evaluate the behavior of the inexact method for a different model, we consider the deblurring problem *NGB 7027*, an example of a diffuse astronomical object, that can be restored by miminizing the combination of KL with the regularization term (43). A satisfactory reconstruction can be obtained by SGP with $\beta_{opt} = 9 \cdot 10^{-9}$; the relative reconstruction error is equal to 0.079339 with 263 iterations ($tol_{SGP} = 10^{-7}$). In table 7, we report the minimum relative reconstruction error in the first 5 iterations of the inexact method with $\beta = 10\beta_{opt}$, that, for this model, is an inexact proximal point. The numerical results are consistent with the considerations for the previous models: few iterations enable to obtain a sensible restored image even in the case of an overestimated regularization parameter. The restored images obtained at the fourth ant fifth iteration are very similar to the one obtained by SGP with β_{opt}.

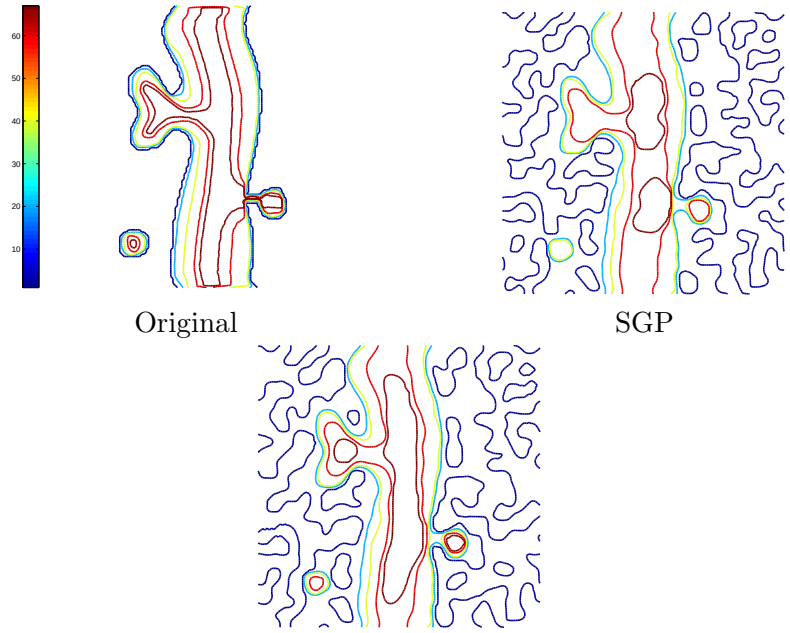


Figure 4. Test problem *micro*: contour plots of the original image (upper-left panel) and restored images for the model KL-HS obtained by SGP (upper-right panel) and by the inexact iterative method for $k = 4$ (lower panel). The level scale, given by (1, 20, 40, 60, 67), is equal for all plots.

Table 7. Inexact Bregman method for *NGB 7027*; the model is the combination of KL function with (43) as regularization term. Here $\alpha = 1.5$ and $tol_{SGP} = 10^{-4}$ at the first outer iteration.

k	$D^{\tilde{p}^{(k)}} f_1(x^*, \tilde{x}^{(k)})$	ρ_k	it
1	3959 10^7	0.1133	47
2	2475 10^7	0.08951	71
3	2001 10^7	0.08026	84
4	1928 10^7	**0.07796**	108
5	1996 10^7	0.07819	122

In table 8, we report the iterations that correspond to the minimum relative reconstruction error ρ_k and the minimum $D^{\tilde{p}^{(k)}} f_1(x^*, \tilde{x}^{(k)})$ when the inexact iterative method is applied with different values of β. Again, we observe the dependence of the number of iterations required to obtain the minimum errors on the accuracy of the β estimate.

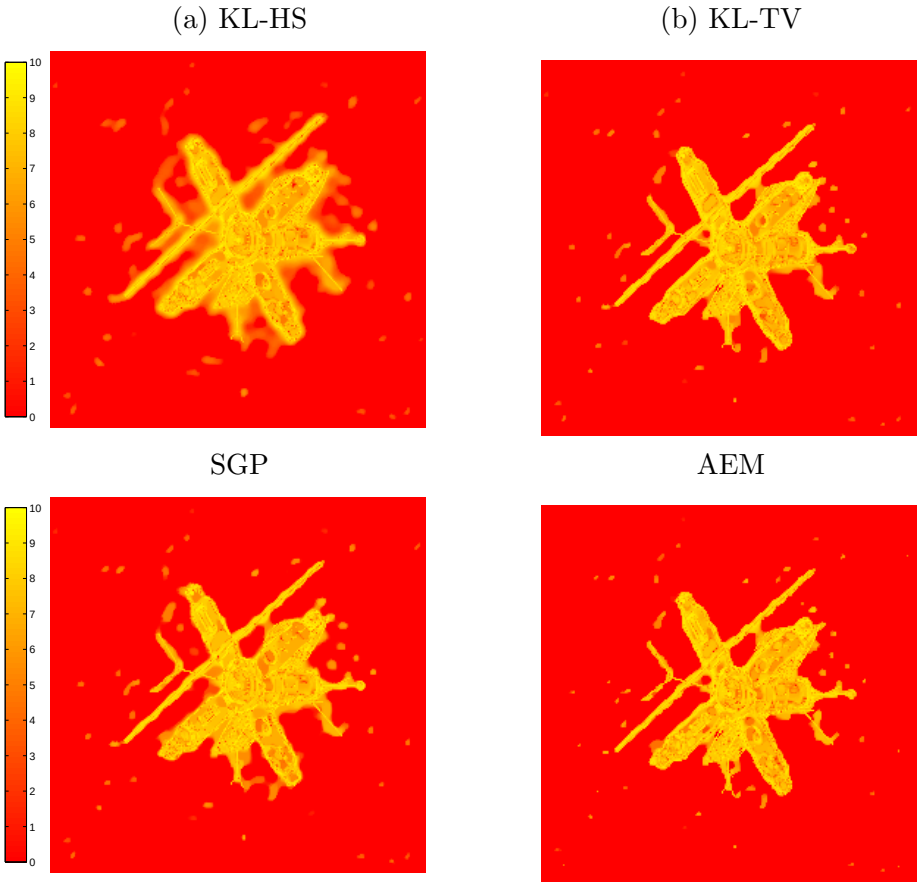


Figure 5. Test problem *H-spacecraft*: *log-error* images for the model KL-HS (left-panels) and KL-TV (right-panels). In the upper panels SGP and AEM are used with an optimal regularization parameter, while the lower panels show the results obtained by the inexact iterative method at the iteration with the minimum value of $D^{\tilde{p}^{(k)}} f_1(x^*, \tilde{x}^{(k)})$ or $\Delta^{\tilde{p}^{(k)}}_{\epsilon_k} f_1(x^*, \tilde{x}^{(k)})$.

Table 8. Test problem *NGB 7027*: minimum errors (ρ_k and $D^{\tilde{p}^{(k)}} f_1(x^*, \tilde{x}^{(k)})$) versus iterations of the inexact iterative method with different values of β; the model is the combination of KL function with the quadratic regularization.

β	k	ρ_k	k	$D^{\tilde{p}^{(k)}} f_1(x^*, \tilde{x}^{(k)})$
$10\beta_{opt}$	4	0.077963	4	1928 10^7
$15\beta_{opt}$	6	0.079085	5	2042 10^7
$20\beta_{opt}$	7	0.080287	6	2110 10^7

4.4. Behavior of the inexact procedure with overestimated regularization parameter: denoising

To evaluate the performance of the inexact procedure for a denoising problem, we refer to the LCR phantom, a piecewise constant object, frequently used in other papers (see, for example, [20]). We consider both the models KL-HS and KL-TV. The model KL-HS can be solved by SGP with an experimentally tuned optimal regularization parameter $\beta_{opt} = 0.575$ [26]; in this case the relative reconstruction error is equal to 0.04231 and the number of iterations is 1034 ($\delta = 10^{-3}$, $tol_{SGP} = 10^{-8}$). For the KL-TV model, we obtain by AEM with the same value of β_{opt} that the reconstruction error is equal to 0.04477 and the number of iterations is 2536 ($tol_{AEM} = 10^{-6}$).
In table 9 we report the results obtained by solving both models by the inexact method: the relative reconstruction error, the exact or inexact Bregman distance and the number of inner iterations at each k–th outer iteration. Also for this problem, in both cases, β is equal to $10\beta_{opt}$. In the first outer iteration, for the KL-HS model, we set $tol_{SGP} = 10^{-5}$, while for the KL-TV model, $tol_{AEM} = 10^{-4}$. Also for this test problem, we observe a semi-convergence behavior of the inexact method. For the model KL-HS, the minimum relative reconstruction error is obtained at the iteration $k = 6$, while for the KL-TV model, we obtain the minimum reconstruction error at $k = 5$. The Bregman distance and the inexact Bregman distance reach the minimum value at the same outer iterate ($k = 6$).
In figure 6 (a), we show the superposition of the line–outs from row number 128 for the

Table 9. Inexact iterative method for the denoising problem related to LCR phantom

	KL-HS			KL-TV		
k	$D^{\tilde{p}^{(k)}} f_1(x^*, \tilde{x}^{(k)})$	ρ_k	it	$\Delta^{\tilde{p}^{(k)}}_{\epsilon_k} f_1(x^*, \tilde{x}^{(k)})$	ρ_k	it
1	2685.2	0.15605	1140	2788.6	0.19737	758
2	1569.1	0.11313	2104	1691.3	0.04855	1300
3	1202.4	0.05435	3688	1285.6	0.03957	1507
4	1081.7	0.04038	2117	1167.4	0.03738	1867
5	1017.1	0.03666	1223	1102.6	**0.03654**	2264
6	998.5	**0.03640**	361	1079.5	0.03670	2657
7	1023.2	0.03791	648	1089.3	0.03848	2988
8	1124.5	0.04051	2261	1129.5	0.04122	3048

KL-HS reconstructions obtained by SGP and by the inexact method at the iteration $k = 6$ (minimum reconstruction error). The solid line is the row number 128 of the original image. Figure 6(b) shows the superposition of the line–outs from row number 128 for the KL-TV reconstructions obtained by AEM and by the inexact method at the iteration $k = 5$. We observe that for both models, the reconstruction obtained by the inexact method is able to reach the level 40 in the central pixels of the original image. This level is underestimated when we solve the models KL-HS and KL-TV with the

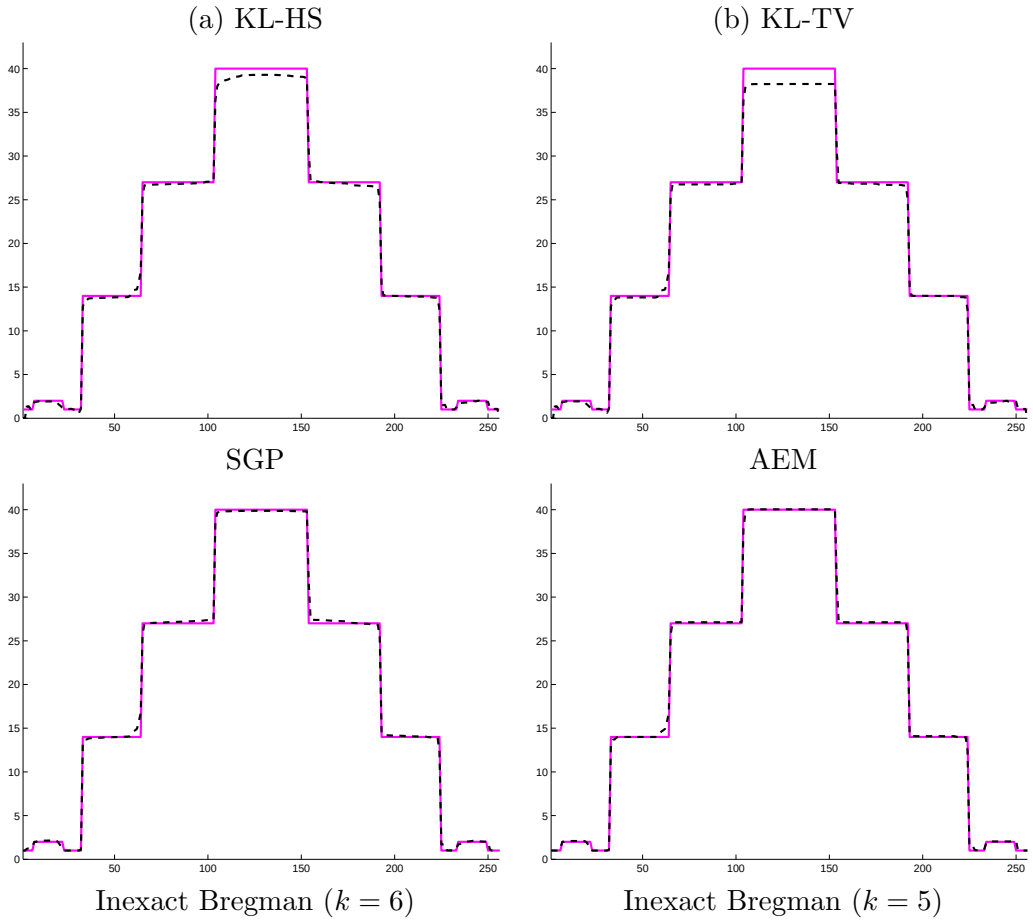


Figure 6. Test problem LCR: superposition of the line–outs from row number 128 for the KL-HS (left-panels) and KL-TV (right-panels) restored images. In the upper panels SGP and AEM are used with an optimal regularization parameter (dashed lines), while the lower panels show the images obtained by inexact iterative method at the iteration with the minimum reconstruction error (dashed lines). The solid line is the row number 128 of the original image.

optimal regularization parameter.

Figure 7 (a) shows the contour plots of the restored images for the model KL-HS obtained by SGP and by the inexact iterative method at the iteration $k = 6$ while figure 7 (b) shows the results obtained for the model KL-TV by AEM and by the inexact iterative method at the iteration $k = 5$. For the images obtained by SGP and AEM, the levels are $[2, 14, 27, 38]$, while for the results of the inexact iterative method, they are $[2, 14, 27, 40]$, since in this case the level 40 is reached. Figure 7 as well as the previous figure 6 show that the inexact iterative method provides a contrast enhancement of the restored images. Furthermore, the contour plots obtained by the inexact method show a better reconstruction, in particular of the frame around the circles and of the highest circle.

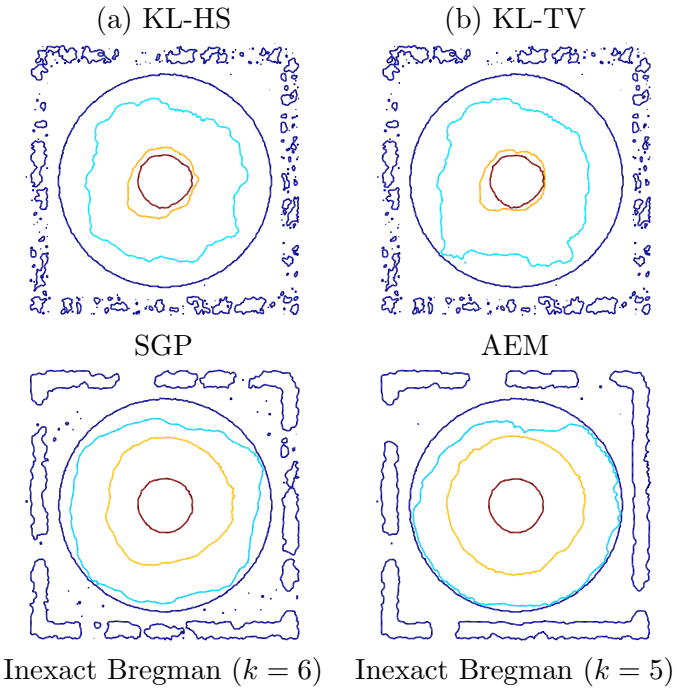


Figure 7. Test problem LCR: contour plots of the KL-HS (left-panels) and KL-TV (right-panels) restored images. In the upper panels SGP and AEM are used with an optimal regularization parameter (levels $[2, 14, 27, 38]$), while the lower panels show the contour plots of the images obtained by inexact iterative method at the iteration with the minimum reconstruction error (levels $[2, 14, 27, 40]$).

5. Conclusions

In this paper we investigate the iterative regularization method based on the Bregman iteration, already proposed in [5] for Gaussian data and in [6] for Poisson data. In a discrete setting, we have resumed the convergence analysis under general hypotheses and we have described an inexact version of the method that enables to use iterative inner solvers, devising an inner stopping criterion that assures the convergence of the scheme. Then we discuss the application of the inexact scheme to image reconstruction problems from data corrupted by Poisson noise (denoising and deblurring). The numerical experiments show that the inexact version appears promising from the point of view of the efficiency: it allows to exploit iterative schemes specialized to minimize KL function combined with a differentiable or a non differentiable regularization term, monitoring how much accurately the inner solution has to be computed to preserve the convergence. Furthermore the iterative scheme is an effective tool for image restoration when only an overestimation of the regularization parameter is known. Obviously the number of iterations needed to obtain the restored image increases with decreasing the accuracy of the parameter estimate. We observe also that, above all in denoising case, the iterative scheme seems to provide an enhancement of the restored images. Future work will

involve the investigation of other techniques for dealing with the incorrect solution of the inner subproblems, in order to further relax the inner accuracy and to obtain a greater efficiency. A crucial point for future investigations is to devise a suitable outer stopping criterion for the scheme.

Appendix A.

We report the proof of Proposition 1.
Proof.

(a) From the optimality condition for the minimizer $x^{(k+1)}$ of $Q_k(x, p^{(k)})$ and in according to Theorem 23.8 in [15], we have $0 \in \partial f_1(x^{(k+1)}) - p^{(k)} + \frac{1}{\beta}\partial f_0(x^{(k+1)})$. Then (7) follows.

(b) From $Q_{k-1}(x^{(k-1)}, p^{(k-1)}) = \frac{1}{\beta} f_0(x^{(k-1)})$ and $D^{p^{(k-1)}} f_1(x^{(k)}, x^{(k-1)}) \geq 0$, since $x^{(k)}$ is a minimizer of $Q_{k-1}(x, p^{(k-1)})$, we have $\frac{1}{\beta} f_0(x^{(k)}) \leq Q_{k-1}(x^{(k)}, p^{(k-1)}) \leq Q_{k-1}(x^{(k-1)}, p^{(k-1)})$ and (8) holds.

(c) By direct algebra, the following identity holds:

$$D^{p^{(k)}} f_1(x, x^{(k)}) - D^{p^{(k-1)}} f_1(x, x^{(k-1)}) + D^{p^{(k-1)}} f_1(x^{(k)}, x^{(k-1)}) = \\ = < (x^{(k)} - x), (p^{(k)} - p^{(k-1)}) >$$

Using (7), since $p^{(k)} - p^{(k-1)} = -\frac{1}{\beta} q^{(k)} \in \frac{1}{\beta}\partial f_0(x^{(k)})$, from the convexity of f_0, we have (9).

(d) If $\widehat{x}$ is a minimizer of f_0, from (9) with $x = \widehat{x}$, since $D^{p^{(k-1)}} f_1(x^{(k)}, x^{(k-1)}) \geq 0$, we obtain

$$D^{p^{(k)}} f_1(\widehat{x}, x^{(k)}) + \frac{1}{\beta}(f_0(x^{(k)}) - f_0(\widehat{x})) \leq D^{p^{(k-1)}} f_1(x^*, x^{(k-1)})$$

Since $f_0(x^{(k)}) - f_0(\widehat{x}) \geq 0$, the inequality (10) holds. Furthermore, summing up the inequalities (9) computed at $\widehat{x}$ related to the first k steps, we have:

$$D^{p^{(k)}} f_1(\widehat{x}, x^{(k)}) + \sum_{i=1}^{k} [D^{p^{(i-1)}} f_1(x^{(i)}, x^{(i-1)}) + \frac{1}{\beta}(f_0(x^{(i)}) - f_0(\widehat{x}))] \leq \\ \leq f_1(\widehat{x}) - f_1(x^{(0)}) \tag{A.1}$$

Since $D^{p^{(i-1)}} f_1(x^{(i)}, x^{(i-1)}) \geq 0$ for any i and $D^{p^{(k)}} f_1(\widehat{x}, x^{(k)}) \geq 0$, from the monotonicity of the sequence $f_0(x^{(i)})$, we have

$$k\frac{1}{\beta}[f_0(x^{(k)}) - f_0(\widehat{x})] \leq f_1(\widehat{x}) - f_1(x^{(0)})$$

and then (11) follows.
Since the sequence $\{x^{(k)}\}$ is bounded, there exists a subsequence of $\{x^{(k)}\}$ convergent to a limit point $\tilde{x}$ and, from (11), we have for $k \to \infty$ that $f_0(\tilde{x}) \leq f_0(\widehat{x})$. Then, $\tilde{x}$ is a minimizer of $f_0(x)$. If $\widehat{x}$ is the unique minimizer of $f_0(x)$, then $x^{(k)} \to \widehat{x}$ as $k \to \infty$.

□

Appendix B.

To analyze the behavior of the inexact Bregman iteration with respect to the exact scheme, we consider a 1D example where, under suitable hypothesis, a closed formula gives the solution of the first two inner subproblems. This is the case of a denoising problem, when it is solved by the discrete TV regularization. By Theorem 1 in [41], the exact solution can be easily computed. For completeness, we report this theorem (in a discrete setting).

Proposition 6 *[41]. Suppose the vector $f \in \mathbb{R}^n$ is defined such that $\frac{\sum_{i=1}^{n_1} f_i}{n_1} = \varphi_1$ and $\frac{\sum_{i=n_1+1}^{n} f_i}{n-n_1} = \varphi_2$, with $\varphi_1 > \varphi_2$ and $\max_{i=n_1+1,n} f_i \leq \min_{i=1,n_1} f_i$. If we assume that*

$$\max_{i=n_1+1,n} f_i \leq \varphi_2 + \frac{\beta}{n-n_1} \leq \varphi_1 - \frac{\beta}{n_1} \leq \min_{i=1,n_1} f_i \tag{B.1}$$

then the unique minimizer of the problem

$$\min_x \frac{1}{2}\|x-f\|^2 + \beta \sum_{i=1}^{n-1} |x_{i+1} - x_i|$$

is given by

$$\widehat{x_i} = \begin{cases} \varphi_1 - \frac{\beta}{n_1} & i = 1, ..., n_1 \\ \varphi_2 + \frac{\beta}{n-n_1} & i = n_1+1, ..., n \end{cases}$$

Let g^* be a vector of $\mathbb{R}^n$, defined as

$$g_i{}^* = \begin{cases} \gamma_1{}^* & i = 1, ..., n_1 \\ \gamma_2{}^* & i = n_1+1, ..., n \end{cases}$$

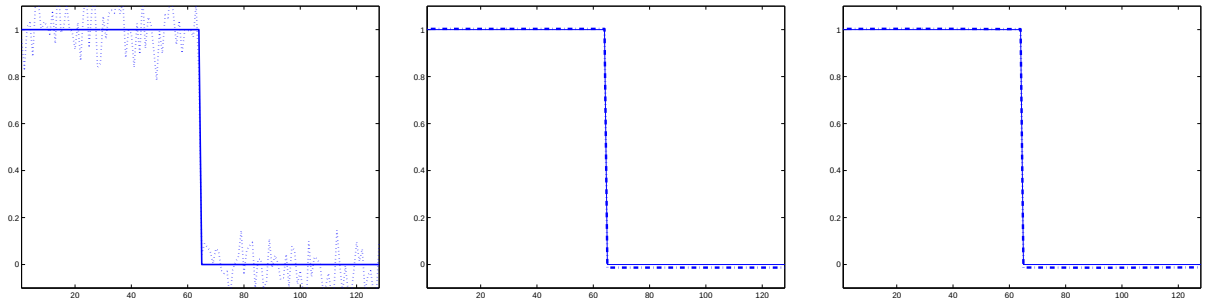
where we assume $\gamma_1{}^* > \gamma_2{}^*$. We denote by n_2 the difference $n - n_1$. By perturbing the data with Gaussian noise with standard deviation σ, we obtain a vector g, such that $g = g^* + s$, where s is the Gaussian noise. The noise is such that, if we denote by γ_1 and γ_2 the following values

$$\frac{\sum_{i=1}^{n_1} g_i}{n_1} = \gamma_1 \qquad \frac{\sum_{i=n_1+1}^{n} g_i}{n_2} = \gamma_2$$

we have $\gamma_1 > \gamma_2$.

In figure B1 (a) we show the original vector g^* (solid line) and the noisy vector g (dotted line) for $n = 128$, $n_1 = n_2 = 64$, $\gamma_1{}^* = 1$, $\gamma_2{}^* = 0$, $\gamma_1 = 1.0030$, $\gamma_2 = -0.0129$, $\frac{1}{2}\|g^* - g\|^2 = 0.5052$, $\sigma = 0.10$.
In order to solve the denoising problem of recovering an estimate of the original object g^* from the noisy data g, we consider the following variational problem (discrete approximation of the ROF model):

$$\min_x \frac{1}{2}\|x-g\|^2 + \beta \sum_{i=1}^{n-1} |x_{i+1} - x_i| \tag{B.2}$$



(a) Original and noisy vector (b) $x^{(2)}$ with exact procedure (c) $\tilde{x}^{(2)}$ with inexact procedure

Figure B1. Test problem 1D: (a) plot of the original vector (solid line) and noisy vector (dotted line); (b) plot of the exact iterate $x^{(2)}$ (dashed line) of the Bregman procedure with respect to the original vector (solid line); (c) plot of the iterate $\tilde{x}^{(2)}$ of the inexact procedure combined with AEM (dashed line) with respect to the original vector (solid line).

where $f_0(x) = \frac{1}{2}\|x-g\|^2$, $f_1(x) = \sum_{i=1}^{n-1} |x_{i+1}-x_i|$ and β is the regularization parameter. The exact Bregman iteration for the problem (B.2) consists in solving a sequence of subproblems as

$$\min_x \frac{1}{2}\|x - (g + \beta p^{(k)})\|^2 + \beta \sum_{i=1}^{n-1} |x_{i+1} - x_i| \quad k = 0, 1, .. \tag{B.3}$$

with $p^{(0)} = 0$. If the parameter β satisfies the following conditions

$$2\max\{n_2(\max_{i=n_1+1,n} g_i - \gamma_2), n_1(\gamma_1 - \min_{i=1,n_1} g_i)\} \leq \beta \leq \frac{(\gamma_1 - \gamma_2)n_1 n_2}{n} \tag{B.4}$$

the first two step of the Bregman iteration can be computed by a closed formula. At the first iteration (the initial subproblem coincides with (B.2)), we have

$$x_i{}^{(1)} = \begin{cases} \gamma_1 - \frac{\beta}{n_1} & i = 1, ..., n_1 \\ \gamma_2 + \frac{\beta}{n_2} & i = n_1 + 1, ..., n \end{cases}$$

$$p_i{}^{(1)} = p_i{}^{(0)} - \frac{1}{\beta}(x_i{}^{(0)} - g_i) \begin{cases} -\frac{1}{\beta}(\gamma_1 - \frac{\beta}{n_1} - g_i) & i = 1, ..., n_1 \\ -\frac{1}{\beta}(\gamma_2 + \frac{\beta}{n_2} - g_i) & i = n_1 + 1, ..., n \end{cases}$$

Furthermore, we have

$$f_0(x^{(1)}) = \frac{1}{2}\left(\sum_{i=1}^{n_1}(\gamma_1 - g_i)^2 + \sum_{i=n_1+1}^{n}(\gamma_2 - g_i)^2 + \frac{\beta^2}{n_1} + \frac{\beta^2}{n_2}\right)$$

$$D^{p^{(1)}} f_1(g^*, x^{(1)}) = 0$$

Now, at the second iteration, since $g^{(1)} = g + \beta p^{(1)}$ is such that

$$\frac{\sum_{i=1}^{n_1} g_i{}^{(1)}}{n_1} = \gamma_1 + \frac{\beta}{n_1} \qquad \frac{\sum_{i=n_1+1}^{n} g_i{}^{(1)}}{n - n_1} = \gamma_2 - \frac{\beta}{n_2}$$

thanks to (B.4) and $\gamma_1 > \gamma_2$, $g^{(1)}$ satisfies the assumption B.1 of Proposition 6. Then the second exact iterate is given by

$$x_i{}^{(2)} = \begin{cases} \gamma_1 & i = 1, ..., n_1 \\ \gamma_2 & i = n_1 + 1, ..., n \end{cases} \tag{B.5}$$

Furthermore, we have

$$\begin{aligned} f_0(x^{(2)}) \qquad &= \frac{1}{2}\left(\sum_{i=1}^{n_1}(\gamma_1 - g_i)^2 + \sum_{i=n_1+1}^{n}(\gamma_2 - g_i)^2\right) \\ D^{p^{(1)}} f_1(g^*, x^{(1)}) &= 0 \end{aligned} \tag{B.6}$$

Then $x^{(2)}$ can be considered an approximation of the original vector.
Figure B1 (b) shows the iterate $x^{(2)}$. In this case a value of β satisfying (B.4) is 29 which is an overestimate of an optimal value of the regularization parameter. Table B1 shows the results obtained for this test problem in the following three cases:

- exact scheme with the closed formula
- exact scheme with an inner solver
- inexact scheme.

Here $\rho_k = \frac{\|x^{(k)} - g^*\|}{\|g^*\|}$, *it* denotes the number of iterations of the inner solver for each k–th outer iteration.
As inner solver for the subproblems (B.3), we use AEM and the primal-dual Algorithm 2 in [31], denoted by CP in table B1. For the exact version, the stopping rule of both inner solvers is based on the standard relative difference in Euclidean norm between two successive primal-dual iterates, that is $\left\| \begin{pmatrix} x^{i+1} \\ y^{i+1} \end{pmatrix} - \begin{pmatrix} x^{i} \\ y^{i} \end{pmatrix} \right\| / \left\| \begin{pmatrix} x^{i+1} \\ y^{i+1} \end{pmatrix} \right\| \leq tol$, where (x^i, y^i) denotes the i-th primal-dual iterate of the inner solver. For the inexact version, in the first iteration, the standard criterion is used; then in the subsequent iterations, both the inner solvers are stopped when the conditions (37)-(38) hold, with $c = \|\eta^{(1)}\|$ and $d = \epsilon_1$. We observe that when an inner iterative method is used in the exact scheme, the computed Bregman distances can assume negative values, since the update rule determines an approximate subgradient of f_1 at the current iterate. Furthermore, in the inexact version, two outer iterations are sufficient to obtain results similar to the ones related to the exact version. For this particular test problem, the inner solver AEM appears more efficient than the CP method.

Acknowledgments

This research is supported by the PRIN2008 Project of the Italian Ministry of University and Research *optimizAtion Methods and Software for Inverse PRoblems*, grant 2008T5KA4L.

Table B1. Test problem 1D: results of different version of the Bregman iteration with $\beta = 29$.

exact version with closed formulas									
k	ρ_k	it	$f_0(x^{(k)})$	$D^{p^{(k)}} f_1(x^*, x^{(k)})$					
1	0.6286		13.6402	0					
2	0.01325		0.49955	0					
k	ρ_k	it	$f_0(x^{(k)})$	$D^{p^{(k)}} f_1(x^*, x^{(k)})$	k	ρ_k	it	$f_0(\tilde{x}^{(k)})$	$\Delta_{\epsilon_k}^{p^{(k)}} f_1(x^*, \tilde{x}^{(k)})$
exact version with AEM as inner solver, $tol = 10^{-6}$					inexact version with AEM and $tol = 10^{-3}$, $\alpha = 1.5$				
1	0.6296	830	13.64	$-8.1\ 10^{-5}$	1	0.6026	239	12.59	$5.3\ 10^{-16}$
2	0.01325	845	0.500	$-1.1\ 10^{-4}$	2	0.01326	623	0.500	$3.5\ 10^{-16}$
3	0.01923	816	0.484	$-1.7\ 10^{-4}$	3	0.02350	697	0.481	$7.0\ 10^{-17}$
exact version with CP as inner solver, $tol = 10^{-8}$					inexact version with CP and $tol = 10^{-6}$, $\alpha = 1.5$				
1	0.6289	61742	13.6126	$9.33\ 10^{-4}$	1	0.6277	22392	13.56	$2.0\ 10^{-10}$
2	0.01304	52678	0.49956	$-1.01\ 10^{-5}$	2	0.01379	34950	0.4996	$1.0\ 10^{-16}$
3	0.02005	46246	0.48373	$-1.61\ 10^{-5}$	3	0.02013	38485	0.4839	$8.0\ 10^{-16}$

References

[1] Shepp L A and Vardi Y. Maximum likelihood reconstruction for emission tomography. *Trans. Med. Imaging*, MI–1:113–122, 1982.

[2] Geman S and Geman D. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[3] Rudin L, Osher S, and Fatemi E. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[4] Bregman L M. The relaxation method of finding the common points of convex sets and its applications to the solution of problems in convex optimization. *USSR Computational Mathematics and Mathematical Physics*, 7:200–217, 1967.

[5] Osher S, Burger M, Goldfarb D, Xu J, and Yin W. An iterative regularization method for total variation-based image restoration. *SIAM Journal on Multiscale Modeling and Simulation*, 4(2):460–489, 2005.

[6] Brune C, Sawatzky A, and Burger M. Primal and dual Bregman methods with application to optical nanoscopy. *Int. J. Comput. Vis.*, 92(2):211–229, 2010.

[7] Bertero M, Boccacci P, Talenti G, Zanella R, and Zanni L. A discrepancy principle for Poisson data. *Inverse Problems*, 26:10500, 2010.

[8] J M Bardsley and J Goldes. Regularization parameter selection methods for ill-posed Poisson maximum likelihood estimation. *Inverse Problems*, 25, 2009.

[9] Staglianò A., Boccacci P, and Bertero M. Analysis of an approximate model for Poisson data reconstruction and a related discrepancy principle. *Inverse Problems*, 27:125003, 2011.

[10] Carlavan M and L Blanc-Féraud. Two constrained formulations for deblurring poisson noisy images. *Proc. IEEE International Conference on Image Processing (ICIP), Brussels, Belgium*, 2011.

[11] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer, Dordrecht, 1996.

[12] Yin W, Osher S, Goldfarb D, and Darbon J. Bregman iterative algorithms for l_1–minimization with applications to compressed sensing. *Siam J. Imaging Sciences*, 1(1):143–168, 2008.

[13] Brune C, Sawatzky A, and Burger M. *Scale Space and Variationale Methods in Computer Vision*, volume 5567 of *LNCS*, chapter Bregman–TV–EM methods with application to optical nanoscopy, pages 235–246. Springer, 2009.

[14] Eckstein J. Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming. *Mathematics of Operations Research*, 18(1):202–226, 1993.

[15] R T Rockafellar. Monotone operators and the proximal point algorithms. *SIAM J. Control*

Optimization, 14(5):877–898, 1976.

[16] Salzo S and Villa S. Inexact and accelerated proximal point algorithms. *www.optimization-online.org/DB.HTML/2011/08/3128.html*.

[17] Ng M K, Wang F, and Yuan X. Inexact alternating direction methods for image recovery. *SIAM J. Sci. Comput.*, 33(4):1643–1668, 2011.

[18] K C Kiwiel. Proximal minimization methods with generalized Bregman functions. *SIAM J Control Optim*, 35(4):1142–1168, 1997.

[19] M. Nikolova. A variational approach to remove outliers and impulse noise. *Journal of Mathematical Imaging and Vision*, 20:99–120, 2004.

[20] R Zanella, P Boccacci, L Zanni, and M Bertero. Efficient gradient projection methods for edge-preserving removal of Poisson noise. *Inverse Problems*, 25, 2009.

[21] J M Bardsley and A Luttman. Total Variation-Penalized Poisson Likelihood Estimation for Ill-Posed Problems. *Adv Comput Math*, 31(1–3):35–59, 2009.

[22] S Bonettini, G Landi, E Loli Piccolomini, and L Zanni. On scaling techniques for gradient projection–type methods in astronomical image deblurring. *in press on International Journal on Computer Mathematics*, DOI:10.1080/00207160.2012.716513, 2012.

[23] Brune C, Sawatzky A, Kösters T, Wübbeling F, and Burger M. Forward-backward EM-TV methods for inverse problems with Poisson noise. 2011.

[24] Setzer S, Steidl G, and Teuber T. Deblurring Poissonian images by split Bregman techniques. *J. Vis. Commun. Image R.*, 21:193–199, 2010.

[25] Figueiredo M A T and Bioucas-Dias J M. Restoration of Poissonian images using alternating direction optimization. *IEEE Transactions on Image Processing*, 2010.

[26] S Bonettini and V Ruggiero. An alternating extragradient method for total variation based image restoration from Poisson data. *Inverse Problems*, 27:095001, 2011.

[27] S Bonettini and V Ruggiero. On the convergence of primal–dual hybrid gradient algorithms for total variation image restoration. *Journal of Mathematical Imaging and Vision*, 44(3):236–253, 2012.

[28] D G Luenberger and Y Ye. *Linear and Nonlinear Programming*. Springer, third edition edition, 2008.

[29] Goldstein T and Osher S. The split Bregman algorithm for L1 regularized problems. *SIAM Journal on Imaging Sciences*, 2(2):323–343, 2009.

[30] R T Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, NJ, 1970.

[31] A Chambolle and T Pock. A first–order primal–dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40:120–145, 2011.

[32] M Bertero, P Boccacci, G Desiderà, and G Vicidomini. Image deblurring with Poisson data: from cells to galaxies. *Inverse Problems*, 25:123006, 2009.

[33] Charbonnier P, Blanc-Féraud L, Aubert G, and Barlaud A. Deterministic edge–preserving regularization in computed imaging. *IEEE Trans. Image Processing*, 6, 1997.

[34] S Bonettini and V Ruggiero. On the uniqueness of the solution of image reconstruction problems with Poisson data. In T.E. Simos et al., editor, *Proceedings of ICNAAM 2010*, volume 1281 of *AIP conference proceedings*, pages 1803–1806. AIP, 2010.

[35] Bonettini S., Zanella R., and Zanni L. A scaled gradient projection method for constrained image deblurring. *Inverse Problems*, 25:015002, 2009.

[36] B Zhan, L Gao, and Y H Dai. Gradient methods with adaptive step-sizes. *Computational Optimization and Applications*, 35:69–86, 2006.

[37] Frassoldati G, Zanni L, and Zanghirati G. New adaptive stepsize selections in gradient methods. *Journal of Industrial and Management Optimization*, 2008.

[38] Willett R M and Nowak R D. Platelets: A multiscale approach for recovering edges and surfaces in photon limited medical imaging. *IEEE Transactions on Medical Imaging*, 22:332–350, 2003.

[39] T Le, R Chartrand, and T J Asaki. A variational approach to reconstructing images corrupted by Poisson noise. *J. Math. Imaging Vis.*, 27:257–263, 2007.

[40] Chambolle A. Total variation minimization and a class of binary mrf models. *EMMCVPR 05. Lecture Notes in Computer Sciences*, 3757:136–152, 2005.

[41] D. M. Strong and T. F. Chan. Exact solutions to Total Variation regularization problems. CAM Reports 96-41, UCLA Center for Applied Math., 1996.