



UNIVERSITÀ DEGLI STUDI DI FERRARA

DOTTORATO DI RICERCA IN  
MATEMATICA E INFORMATICA

CICLO XXVII

COORDINATORE PROF. MASSIMILIANO MELLA

**Optimization Methods  
for Image Regularization  
from Poisson Data**

Settore Scientifico Disciplinare MAT/08

**Dottorando**

Dott. Benfenati Alessandro

**Tutore**

Prof. Ruggiero Valeria

Anni 2012-2014



# Contents

<b>Introduction</b>	<b>v</b>
<b>Notations</b>	<b>ix</b>
<b>List of Algorithms</b>	<b>1</b>
<b>1 Generalities on Image Restoration Problems</b>	<b>3</b>
1.1 Basics on the Mathematical Model . . . . .	3
1.2 Point Spread Function . . . . .	4
1.3 Noise . . . . .	7
1.3.1 Gaussian Noise . . . . .	7
1.3.2 Poisson Noise . . . . .	8
1.4 Data Restoration . . . . .	8
1.4.1 Maximum Likelihood Estimation . . . . .	10
1.4.2 Bayesian Approach . . . . .	12
1.5 Constraints . . . . .	12
1.6 Regularization Functionals . . . . .	13
1.6.1 Tikhonov Regularization . . . . .	13
1.6.2 Edge-Preserving Regularization . . . . .	14
1.7 The regularization problem for Poisson data . . . . .	15
<b>2 Optimization Methods</b>	<b>17</b>
2.1 SGP . . . . .	17
2.1.1 Scaling Matrix . . . . .	22
2.1.2 Steplength Selection . . . . .	23
2.2 EM-TV . . . . .	23
2.3 AEM . . . . .	24
2.4 PIDAL & PIDSplit+ . . . . .	29
2.5 Chambolle & Pock's algorithm . . . . .	31
2.6 PDHG with variable metric . . . . .	37
2.6.1 Scaled $\varepsilon$ -subgradient level algorithm . . . . .	43
2.6.2 Scaled Primal-Dual Hybrid Gradient Method . . . . .	47
2.6.3 Application to image restoration . . . . .	49
2.7 Numerical Comparison . . . . .	53
<b>3 Numerical Methods for Parameter Estimation in Poisson Data Inversion</b>	<b>59</b>
3.1 Two Estimation Models . . . . .	59

---

3.1.1	Relations Between Penalized and Constrained Convex Problems	60
3.2	Existence and uniqueness of the solution of the Discrepancy equation . .	65
3.3	Numerical Methods . . . . .	66
3.3.1	Model 1, <i>Crossing Model</i> . . . . .	66
3.3.2	Model 2, <i>Constrained Model</i> . . . . .	67
<b>4</b>	<b>Bregman Procedure</b>	<b>71</b>
4.1	Bregman distance: Definition and Properties . . . . .	71
4.2	Bregman Procedure . . . . .	72
4.2.1	The procedure for image restoration problems . . . . .	76
4.3	Inexact Bregman Procedure . . . . .	81
4.3.1	The inexact procedure for image restoration problems . . . . .	86
4.3.2	A simple 2-D example . . . . .	87
4.3.3	A simple 1-D example . . . . .	88
<b>5</b>	<b>Numerical Experiments and Applications</b>	<b>95</b>
5.1	Exact versus inexact procedure . . . . .	95
5.2	Denoising . . . . .	99
5.3	Comparison between the methods for $\beta$ estimation and the inexact Breg- man procedure . . . . .	102
5.4	Application to High Dynamic Range images in Astronomy . . . . .	108
5.4.1	Young Stellar Objects . . . . .	111
5.4.2	Binary Stars . . . . .	117
5.5	Conclusion . . . . .	121
<b>A</b>	<b>Convex Analysis</b>	<b>123</b>
A.1	Definitions and general results . . . . .	123
A.2	$\varepsilon$ -Subgradients . . . . .	125

# Ringraziamenti

Questa tesi è il frutto del lavoro svolto nei tre anni di dottorato. Non sarebbe mai stato possibile per me arrivare fin qui se non fosse stato per i miei genitori, che mi han sempre sostenuto, lasciandomi piena libertà nelle mie scelte e consigliandomi quando i miei sbagli fossero evidenti. Ma il ruolo principale l'ha avuto mia sorella, che ha plasmato l'uomo che sono oggi e contribuirà negli anni a venire nel rendermi ancora migliore.

In questi tre anni sono stato fortunato nel ritrovarmi in un gruppo di ricerca molto attivo, preparato e soprattutto umano. Valeria mi ha insegnato cosa vuol dire fare ricerca in Matematica, porsi le domande giuste, andare fino in fondo alle questioni e risolvere i problemi nel migliore dei modi possibili. Spero di aver acquisito anche solo un minimo dei suoi insegnamenti. I miei ringraziamenti vanno anche a Riccardo, sempre pronto a fornire consigli e suggerimenti su ogni mio dubbio informatico, e a Gaetano, che con il suo vocione mi ha dato sempre una grossa mano. Grazie anche a Silvia, per tutti gli spunti e i lavori svolti insieme, e a Marco, che con le sue osservazioni spinge sempre a dare un po' di più. Grazie anche a Luca, che assieme a Valeria ha dato vita al progetto SPINNER che mi ha permesso di poter svolgere il lavoro che mi appassiona. Anastasia e Federica ci sono sempre state nei miei momenti difficili, con parole di conforto e di incoraggiamento.

Un doveroso ringraziamento va a Mario Bertero, le cui osservazioni, suggerimenti e aiuti sono stati di fondamentale importanza nel lavoro di ricerca di questi anni. Patrizia Boccacci, Andrea La Camera e Marcell Carbillet mi han introdotto nel campo dell'imaging astronomico, un'area di ricerca che mi ha sempre affascinato. Grazie anche ad Antonin Chambolle, direttore del CMAP all'Ecole Polytechnique, presso il quale ho svolto un periodo di studi: grazie a lui ho acquisito un altro punto di vista su come fare ricerca.

Andando indietro nel tempo, mi pare doveroso ringraziare le mie professoressa del liceo Scientifico, che mi han fatto innamorare della Matematica, ed Elena Loli Piccolomini, la quale, nelle sue lezioni, mi ha fatto invaghiare dell'Analisi Numerica.

Questo percorso non sarebbe stato ciò che si è rivelato senza gli amici e le persone incontrate in questi anni. Grazie a Cosimo, un buon amico, a Giulio, sempre presente e con l'incredibile capacità di portare calma e serenità nella vita delle persone. Fabrizio, Simone e Sergio, con cui condivido il percorso irto di difficoltà nel campo della ricerca scientifica. Piero ed Elisa, che mi comunicano un senso di stabilità raro. Grazie anche alla mia vecchia squadra di nuoto, dove ho imparato che non c'è ricompensa senza sacrificio.

Infine ringrazio tutte le altre persone non nominate incontrate nei miei viaggi, nelle conferenze, che mi han dato qualcosa in più da aggiungere al mio bagaglio personale.

Ma soprattutto grazie a te, Elena.



# Introduction

*Inverse Problems* is a research area in very different domains of Applied Science. In particular they are basic in problems of image restoration [11] which are relevant in medicine, microscopy, seismology, astronomy etc. These problems are *ill-posed* in the sense of Hadamard, i.e. the solution might lack in existence, uniqueness or continuous dependence on data. In the past several approaches have been adopted in order to overcome the ill-posedness: they consist in the so-called regularization theory [95, 47, 11] and in Bayesian methods [53]. In all cases the solution of an inverse problem can lead to the minimization of a suitable functional containing, in general, both a *discrepancy* term, measuring the distance between computed and measured data, and a *penalty* or *regularization* term, expressing the required characteristics on the solution. The trade-off between the two is balanced by a parameter, namely *beta*, called *regularization parameter*. Additional constraints can also be considered (e.g., the non-negativity of the solution). Since in real applications the data we are given are discrete, one has to translate the continuous model into a discrete one. The ill-posedness of the problems results in ill-conditioning of the introduced numerical approximations.

In the most part of practical applications, the problems involve a large amount of data; hence the main aim is to solve large-scale nonlinear problems. The variational formulation, i.e. the minimization of opportune functionals, depends on the application, on the type of data and on the mathematical formulation one is considering. The choice of the discrepancy functional depends on the statistical properties of the noise affecting the data, while the penalty functional is chosen in order to take into account a-priori information, with the aim to preserve some features on the solution, such as sharp edges, high contrast regions, diffuse component and so on.

For what regards the noise, in practical applications, Gaussian noise and Poisson noise are predominant. The former one is additive and it is the starting point of the well-known Tikhonov regularization theory for the solution of inverse and ill-posed problems [95]. Poisson noise is also known as *photon noise* and it is due to errors arising in counting the incoming photons. In this case the Maximum Likelihood approach [88] leads to the minimization of a non-quadratic functional, known as Csiszar I-divergence or Kullback-Leibler functional. In recent years (see, for example, [69, 13, 33, 49, 101] and the references therein) it has become of increasing interest in astronomy, microscopy and in tomography: thus in this work we will focus our discussion on Poisson noise.

The parameter  $\beta$  has a huge influence on the computed solution: the optimal value of this parameter is obviously the one which permits to obtain a reconstruction the closest possible to the original signal. This value is not known. In the past years, several techniques have been developed; for Gaussian noise several methods have been studied and proposed: Morozov's discrepancy principle [75], GCV [97], L-curve method [58, 59, 61]. The already cited growing interest in Poisson noise gave birth to studies

regarding the choice of  $\beta$  even in this case [13, 33, 87].

This thesis mainly focuses on regularization techniques for Poisson data and, in particular, on Bregman iterative procedure in a image restoration framework. The basic idea of this procedure lies in the paper by Bregman [25]: it consists in substituting the regularization functional with its Bregman distance from the previous iterate. Osher et al. [79] discussed the semi convergence property of the method in the framework of the inverse problems; this regularizing procedure was first implemented for Gaussian noise, and then developed for Poisson noise in [28, 29]. At each step, a minimization problem has to be solved: when an explicit solution of this problem is not available, one has to use an iterative solver in order to obtain such a solution. In this case the procedure can become high expensive. On the other hand, as it will be deepened throughout the work, the Bregman procedure permits to employ an *overestimation* of the regularization parameter: the obtained reconstructions are satisfying with respect to the ones obtained by using the optimal parameter and, in some cases, the formers are more appealing than the latters. Furthermore, this overestimation can allow to gain a contrast enhancement, which is of fundamental importance in some practical problems arising in Astronomical imaging framework.

The original contribution of this thesis regarding the Bregman procedure consists in the introduction of an inexact scheme, which allows to control the *inexactness* with which one solves the subproblems with an iterative solver. This scheme preserves the main convergence property of the original procedure and, at the same time, it enables us to have satisfactory results with a lower computational cost.

As further contribution, we analyse up to the level of software implementation two models that permit to estimate the optimal parameter  $\beta$ . The first one, named *Crossing Model* [13, 102], consists in finding the root of a function, called the *discrepancy function*. The second one, called *Constrained Model*, is based on a constrained minimization problem [92]. These two models are compared with the inexact Bregman procedure.

We will show that the three approaches, –inexact Bregman procedure, Constrained Model and Crossing Model– in some cases are equivalent, i.e. one can obtain similar results. Unfortunately, for High Dynamic Range imaging problems arising in an Astronomical framework, the Crossing Model and the Constrained one are not able to give the expected results, due to the high difference of intensity in the image. On the contrary, the inexact Bregman, thanks to the property of enhancing the contrasts, allows to improve the reconstruction of the diffuse component, providing information of interest to astronomers. We describe some applications on simulated data provided by Dr. Andrea La Camera (DIBRIS, University of Genova) and Marcel Carbillet (University of Nice-Sophia Antipolis).

The thesis is organized as follows.

In Chapter 1, we give the generalities of image restoration problems, providing some details on the image formation process and on the noise arising during the data acquisition; we also introduce the notation which will be used in the following. Moreover, we resume the basics of the statistical framework lying under the whole work, focusing particularly on the Bayesian approach and on the regularization functionals which will be employed.

In Chapter 2, we give a survey of the optimization methods used in numerical experiments and employed as inner solvers in Constrained Model, in Crossing Model and in the Bregman procedure. Some of them are state-of-the-art methods for Poisson

---

data, such as the EM-TV method or the ADMM approach, while others have been developed and widely used in the recent years (e.g., Chambolle & Pock Algorithm 1 [37] or the SGP method [22]). An original contribution concerns the introduction of a *variable scaling metric* in the class of  $\varepsilon$ -subgradient methods, such as the Primal Dual Hybrid Algorithm [21]. These methods are suitable for problems with a non differentiable regularization term. They have the feature to provide information on the inaccuracy of the computed subgradient. Then, they are well suited as inner solvers of an inexact Bregman procedure. The introduction of a variable metric allows to obtain an acceleration for this class of solvers. We give the convergence proof and furthermore we present a section of experiments to numerically evaluate this technique.

In Chapter 3 Crossing and Constrained models [102] are presented. We prove the equivalence of these models, i.e. in a theoretical framework they provide us with the same estimation of  $\beta$  and the same reconstruction; moreover, extending existing results on smooth regularization functions, we prove the existence and the uniqueness of the solution of Crossing Model for seminorm regularization. Furthermore, we describe the numerical methods which allow us to exploit these models. For the first model, the method used is based on two successive steps: the first consists in finding the interval in which the root lies, the second is a secant phase to actually compute this root. For Constrained Model, we analyze the Alternate Direction of Multiplier Method (ADMM), particularly addressed for the image restoration problem [23, 49]; following [98, 63], an adaptive procedure for the estimation of the ADMM parameter is investigated.

In Chapter 4 we present the Bregman procedure. We recall the original Bregman procedure, with particular attention to the image restoration problem, and then we introduce the *inexact* procedure. Moreover, we explain the regularization behaviour of the exact and inexact Bregman procedure in the image restoration framework, also with the help of two meaningful examples. The results of this chapter have been published in [9].

Chapter 5 is devoted to numerically evaluate the procedures described in Chapters 3 and 4. The first set of experiments have the aim to compare the behaviour of the exact and the inexact Bregman procedure, showing that the latter one provide us with reliable reconstructions with lower computational cost. The second set of experiments concerns a denoising problem, in which the inexact Bregman procedure seems to provide us with better reconstructions than the ones obtained by employing the optimal value for  $\beta$ . These first two sections represent a summary of the experiments reported in [9]. The third part of this chapter is devoted to compare Crossing and Constrained models with the Bregman procedure: we observe that in some cases the three models provide comparable results in terms of accuracy, while in others, in case of low counts images, the Bregman procedure enables us to obtain reconstructed images that the other models are not able to achieve. The last part regards the High Dynamic Range images, in an Astronomical framework: this type of images are characterized by the presence of point sources of very high intensity with respect to the rest of the image. In order to overcome the difficulties arising from this huge contrast, we adopt a superposition model, distinguishing the point source component from the diffuse one. We show that the inexact Bregman procedure enable us to restore the HDR images with a good contrast enhancement, both in case of Young Stellar Objects and of binary stars surrounded by extrastellar materials. This part is an extension of the results in [8].

In Appendix A one can find some useful results in Convex Analysis; such results are widely used throughout the thesis. One can refer to the seminal works [84, 78] to a deeper insight in this field.

# Notations

- ▷  $\mathbb{R}^+$  is the set of the nonnegative real numbers:  $\mathbb{R}^+ = \{x \in \mathbb{R} | x \geq 0\}$ .
- ▷ Lower case bold letters (e.g.  $\mathbf{x}$ ) will denote vectors, upper case letters (e.g.  $\mathbf{A}$  or  $L$ ) will denote matrices.
- ▷  $\mathbf{1}$  and  $\mathbf{0}$  denote a vector with all entries equal to one and to zero, respectively.
- ▷ If  $\mathbf{x}, \mathbf{y}$  are two vectors belonging to  $\mathbb{R}^n$ ,  $\frac{\mathbf{x}}{\mathbf{y}}$  denotes the component-wise division, provided that  $\mathbf{y}_i \neq 0$  for each  $i$ .
- ▷ If  $\mathbf{x}, \mathbf{y}$  are two vectors belonging to  $\mathbb{R}^n$ ,  $\mathbf{x} \cdot \mathbf{y}$  denotes the component-wise product.
- ▷ If  $\mathbf{x}, \mathbf{y}$  are two vectors belonging to  $\mathbb{R}^n$ ,  $\langle \mathbf{x}, \mathbf{y} \rangle$  denotes the scalar product.
- ▷ If  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{x} > 0 \Leftrightarrow \mathbf{x}_i > 0$  for each  $i$ . The same notation holds also for  $<, \geq, \leq$ .
- ▷ For elementary functions, such that log, exp, square root, the application of such functions on vectors is intended component-wise. For example, if  $\mathbf{y} \in \mathbb{R}^n$ , then  $\sqrt{\mathbf{y}} = (\sqrt{\mathbf{y}_1}, \sqrt{\mathbf{y}_2}, \dots, \sqrt{\mathbf{y}_n})^t$ .
- ▷  $\mathcal{M}_{m \times n}(\mathbb{R})$  denotes the set of the matrices with  $m$  rows and  $n$  columns with real elements; when  $m = n$ , we denote this set with  $\mathcal{M}_n(\mathbb{R})$ .
- ▷ Unless noted otherwise, the function  $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$  denotes the usual Euclidean norm.



# List of Algorithms

1	<b>SGP</b> . . . . .	20
2	<b>ABB Steplength Selection</b> . . . . .	23
3	<b>EM-TV</b> . . . . .	24
4	<b>AEM</b> . . . . .	29
5	<b>ADMM</b> . . . . .	30
6	<b>PIDAL-TV or PIDSplit+</b> . . . . .	31
7	<b>Chambolle &amp; Pock Algorithm 1</b> . . . . .	32
8	<b>SSL</b> . . . . .	44
9	<b>SPDHG</b> . . . . .	54
10	<b>Bracketing Phase of Algorithm MDF</b> . . . . .	67
11	<b>Secant Phase of Algorithm MDF</b> . . . . .	68
12	<b>Bregman Procedure, General Scheme</b> . . . . .	73
13	<b>Inexact Bregman Procedure: General Scheme</b> . . . . .	82
14	<b>Inexact Bregman Procedure</b> . . . . .	86



# Chapter 1

## Generalities on Image Restoration Problems

In many fields of Applied Science we are addressed to *find an optimal solution for problems* of various type; this task consists in finding a *minimum* or a *maximum* of a function  $f$ :

$$x_{sol} = \arg \min_x f(x) \quad \text{or} \quad x_{sol} = \arg \max_x f(x)$$

We can have some examples in Machine Learning (SVM) [96, 5], Signal Processing [27], Logistic Planning [66], Data Mining [62], Image Restoration ([11, 14]). In this work we will focus on some applications in *Image Restoration problems with Poisson data*, both in deblurring and in denoising case, keeping in mind that the algorithms and the procedures described throughout this thesis can be applied to other problems.

### 1.1 Basics on the Mathematical Model

The original object in which we are interested (image, acoustic sound, ...) can be modeled by a function  $f(\mathbf{x})$ , where  $\mathbf{x}$  is a 1D, 2D or 3D dimensional vector. The data acquisition system can be modeled by an impulse response function  $\mathcal{H}$  which simulates the effects the system produces on the registered image, namely  $g$ . The acquisition process turns out to be linear in most cases, so it can be stated as

$$g(\mathbf{x}) = \int \mathcal{H}(\mathbf{x}, \mathbf{x}') f(\mathbf{x}') d\mathbf{x}' \quad (1.1)$$

where  $\mathcal{H}$  is called Point Spread Function (PSF). Many physics systems for data acquisition can be modeled by equation (1.1).

In many systems, the function  $\mathcal{H}$  is invariant with respect to translations; hence actually  $\mathcal{H}(\mathbf{x}, \mathbf{x}')$  is a function only of the difference  $\mathbf{x} - \mathbf{x}'$  of the variables. In this case the system and the corresponding function are called *space invariants* and the equation (1.1) becomes

$$g(\mathbf{x}) = \int \mathcal{H}(\mathbf{x} - \mathbf{x}') f(\mathbf{x}') d\mathbf{x}' \quad (1.2)$$

Therefore we can write the registered data  $g$  as the *convolution product* of  $\mathcal{H}$  and  $f$ :  $g = \mathcal{H} * f$ . Hereafter, we will consider only space invariant systems.

In some applications, a nonnegative constant *background term* (to which we will refer with  $bg$ ) can be added to the final image, leading to the form

$$g = \mathcal{H} * f + bg \quad (1.3)$$

Moreover, due to the physics connected to the signal registration process, the clear data  $g$  can be affected by some kind of noise, resulting in a blurred and noisy data  $gn$ . This induces to see the registered data as a realization of a random variable. It will become clearer in the forthcoming sections.

Throughout this work we will have to deal mainly with digital signals, thus we need to use a discrete model instead of a continuous one. The original signal, once a function, is treated now as a vector  $\mathbf{f}$  belonging to  $\mathbb{R}^N$ . In view of (1.3), the continuous convolution product  $\mathcal{H} * f$  can be restated in a discrete framework as the matrix-vector product  $\mathbf{H}\mathbf{f}$ , where  $\mathbf{H} \in \mathcal{M}_{M \times N}(\mathbb{R})$ . Thus a linear discrete model for the signal registration process is given by

$$\mathbf{g} = \mathbf{H}\mathbf{f} + bg\mathbf{1} \quad (1.4)$$

where  $\mathbf{g} \in \mathbb{R}^M$ . Moreover, due to the presence of noise, the detected data is a vector  $\mathbf{gn} \in \mathbb{R}^M$ , instead of  $\mathbf{g}$ ; for example, for additive noise we have  $\mathbf{gn} = \mathbf{H}\mathbf{f} + bg\mathbf{1} + v = \mathbf{g} + v$ , where the vector  $v$  represents the noise contribution. In Section 1.3, we discuss this issue.

We are giving more details about the image registration process.

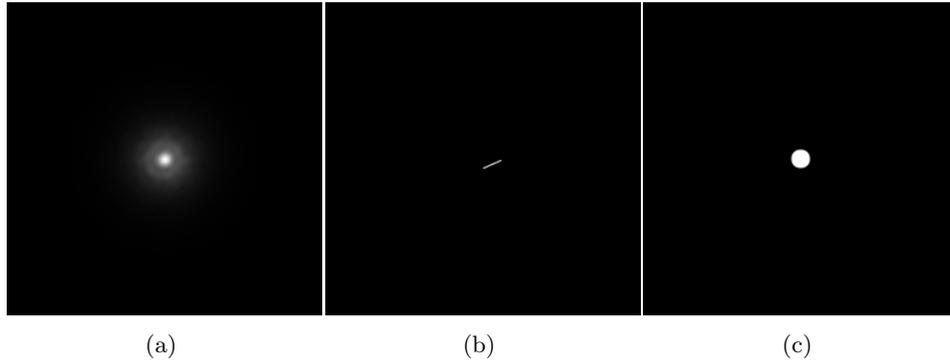
- ▷ A digital 2D image  $\mathbf{F}$  is a matrix belonging to  $\mathcal{M}_{m \times n}(\mathbb{R})$ , where in general  $n \neq m$ ; each element of the image is called pixel (picture element). For computational purposes, a rectangular image is vectorized, i.e. it becomes a vector belonging to  $\mathbb{R}^N$ , where  $N = mn$ :  $\mathbf{f} = \mathbf{vec}(\mathbf{F})$ . This vector contains the data of the image ordered by columns. The action of the  $\mathbf{vec}$  operator is clarified below.

$$\mathbf{F} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \text{ then } \mathbf{vec}(\mathbf{F}) = \begin{pmatrix} 1 \\ 4 \\ 2 \\ 5 \\ 3 \\ 6 \end{pmatrix}$$

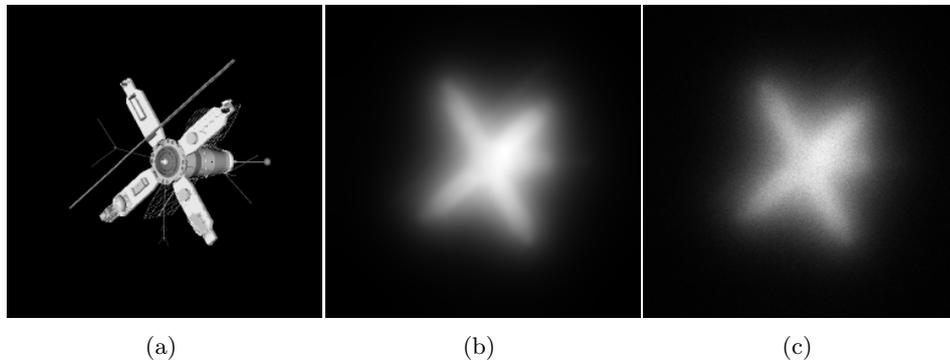
- ▷ Starting from a discretized PSF and taking into account the boundary conditions, it is possible to obtain the matrix  $\mathbf{H}$  (see [60]). Some examples of PSF can be found in Figure (1.1). We distinguish two cases: when  $\mathbf{H} = \mathbf{I}$  (where  $\mathbf{I}$  is the identity matrix) the data are perturbed only by noise, hence the problem we are facing is a *denoising* one, while when  $\mathbf{H} \neq \mathbf{I}$ , the restoration problem is called *deblurring* problem.

## 1.2 Point Spread Function

The matrix  $\mathbf{H}$  can be obtained by the *Point Spread Function* (PSF), which is the impulsive response of the acquisition system, i.e., it is the registered data coming from a point source. It models how the physics of the image recording system perturbs the



**Figure 1.1:** *Examples of Point Spread Functions in imaging. 1.1(a): telescope; 1.1(b): motion blur; 1.1(c): out-of-focus blur.*



**Figure 1.2:** *Example of the effect induced by a PSF: on the left the original image, in the middle the blurred image using the PSF used is 1.1(a). On the right blurred image corrupted by Poisson noise.*

original signal.

When the PSF satisfies the space-invariance property, we can use the model stated in (1.4), where the matrix-vector product  $\mathbf{H}\mathbf{f}$  is a discrete convolution product. In a discrete framework one must deal with conditions at the boundary of the image. The main settings of these conditions are *periodic boundary conditions*, *zero boundary conditions*, *reflexive boundary conditions*: they depend on the the type of problem or data we are treating.

Focusing on the imaging framework, the introduction of the boundary conditions induces particular structures on the imaging matrix  $\mathbf{H}$ : these structures incorporates all the information needed for the computation of the convolution product. We recall some useful definitions.

**Definition 1.1.** *Let  $Q$  be a matrix belonging to  $\mathcal{M}_{p \times q}(\mathbb{R})$ .*

- *$Q$  is a Toeplitz matrix when its entries are constant on each diagonal;*
- *$Q$  is a circulant matrix when it is a Toeplitz matrix and each row (column) is a periodic shift of its previous row (column);*

- $Q$  is a Hankel matrix when its entries are constant on each antidiagonal.

This concepts are easily expandable to block matrices; hence, following [60], we have:

- ▷ a BCCB matrix is a block circulant with circulant blocks;
- ▷ a BHHB matrix is a block Hankel matrix with Hankel blocks;
- ▷ a BTHB matrix is a block Toeplitz matrix with Hankel blocks;
- ▷ a BHTB matrix is a block Hankel matrix with Toeplitz blocks.
- ▷ a BTTB matrix is a block Toeplitz matrix with Toeplitz blocks.

Thus, we can classify the matrix  $\mathbf{H}$  depending on the boundary conditions employed in the a 2D imaging problem:

- ◇ zero boundary conditions:  $\mathbf{H}$  is a BTTB matrix;
- ◇ periodic boundary conditions:  $\mathbf{H}$  is a BCCB matrix;
- ◇ reflexive boundary conditions:  $\mathbf{H}$  is a sum of BCCB, BHHB, BTHB and BHTB matrices.

For each type of imaging matrices, the computation of the discrete matrix–vector product  $\mathbf{H}\mathbf{f}$  can be efficiently obtained by discrete transforms.

Hereafter, we will focus on periodic boundary conditions; hence the appropriate transform to be used is the Discrete Fourier Transform ( $\mathbf{DFT}$ ) and its inverse ( $\mathbf{IDFT}$ ). In this way, using the convolution theorem, the matrix–vector products  $\mathbf{H}\mathbf{f}$  and  $\mathbf{H}^t\mathbf{f}$  can be obtained in the following way

$$\begin{aligned}\mathbf{H}\mathbf{f} &= \mathbf{IDFT}[\mathbf{DFT}[h] \cdot \mathbf{DFT}[\mathbf{f}]] \\ \mathbf{H}^t\mathbf{f} &= \mathbf{IDFT}[\overline{\mathbf{DFT}[h]} \cdot \mathbf{DFT}[\mathbf{f}]]\end{aligned}$$

where  $h$  denotes the first column of  $\mathbf{H}$  and  $\bar{\alpha}$  denotes the complex conjugate of any entry of the vector  $\alpha$ .

For deblurring problems, when the imaging matrix is approximated by the cyclic convolution of the object with a periodic and nonnegative PSF, standard assumptions on the matrix  $\mathbf{H}$  are given by the following conditions:

$$\mathbf{H}_{ij} \geq 0 \quad \forall i, j, \quad \mathbf{H}^t\mathbf{1} = \mathbf{1}, \quad \mathbf{H}\mathbf{1} > \mathbf{0} \quad (1.5)$$

Moreover, if  $\mathbf{H}$  is a square matrix, a further classical condition is  $\mathbf{H}\mathbf{1} = \mathbf{1}$ . These hypothesis on  $\mathbf{H}$  are commonly used in astronomy and microscopy.

For a very detailed discussion about the imaging framework, the computational aspects of the boundary conditions and the discrete model, see [11, 12, 97, 60].

### 1.3 Noise

Following [14, 11], we recall the mathematical model of data acquisition in a statistical framework. As a consequence of the noise introduced by the detection system, the data can be viewed as a realizations of random variables. Indeed, the noise is a random process so that the detected value  $\mathbf{gn}_i$  at each  $i$ -th pixel is a realization of a random variable  $Y_i$ ; therefore, a modeling of the detection system requires a modeling of the noise, i.e. a model of its probability density. This density depends on the original object  $\mathbf{f}^*$  and therefore we denote it as

$$P_Y(\mathbf{gn}|\mathbf{f}^*)$$

where  $Y$  is a multivalued random variable whose components  $Y_i$  are random variables. In order to be able to treat such a function, the following assumptions are in general accepted as reasonable ones.

**Assumptions 1.1.** *The random variables  $Y_i$  and  $Y_j$  are pairwise statistically independent and identically distributed (i.i.d.) for each  $i, j$ . In this way, we can write*

$$P_Y(\mathbf{gn}|\mathbf{f}^*) = \prod_i p_{Y_i}(\mathbf{gn}_i|\mathbf{f}^*) \quad (1.6)$$

**Assumptions 1.2.** *The expected value of  $Y$  is*

$$E[Y] = \mathbf{Hf}^* + b\mathbf{g}\mathbf{1}$$

#### 1.3.1 Gaussian Noise

As already specified, in case of Gaussian or additive noise, the linear discrete model (1.4) can be rewritten as

$$\mathbf{gn} = \mathbf{Hf}^* + b\mathbf{g}\mathbf{1} + v$$

where  $v$  is a realization of the multivalued random variable  $\Upsilon$ . Each component  $\Upsilon_i$  is a random variable with Gaussian distribution with zero mean and standard deviation  $\sigma$ .

$$p_{\Upsilon_i}(v_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{v_i^2}{2\sigma^2}\right)$$

Thus, for the multivalued random variable, we have

$$P_{\Upsilon}(v) = \frac{1}{(2\pi\sigma)^{\frac{N}{2}}} \exp\left(-\frac{\|v\|^2}{2\sigma^2}\right)$$

Then, following [11], a suitable model for  $P_Y(\mathbf{gn}|\mathbf{f}^*)$  which takes into account the previous assumptions, is given by

$$P_Y(\mathbf{gn}|\mathbf{f}^*) = P_{\Upsilon}(\mathbf{Hf}^* + b\mathbf{g}\mathbf{1} - \mathbf{gn}) = C_N \exp\left(-\frac{\|\mathbf{Hf}^* + b\mathbf{g}\mathbf{1} - \mathbf{gn}\|^2}{2\sigma^2}\right) \quad (1.7)$$

In this way,  $Y$  is a multivalued random variable with Gaussian distribution, whose expected value is exactly  $\mathbf{Hf}^* + b\mathbf{g}\mathbf{1}$ . We point out that the Gaussian noise affecting the  $i$ -th element is independent from the intensity value of the  $i$ -th component of the data.

### 1.3.2 Poisson Noise

When the image intensity is measured via the counting of incident particles (e.g. photons,  $\gamma$ -rays, etc), the fluctuations in the emission counting process can be described by modeling the data as realizations of Poisson random variables: each component of the data follows the Poisson distribution

$$p_{Y_i}(\mathbf{gn}_i) = e^{-\lambda_i} \frac{\lambda_i^{\mathbf{gn}_i}}{(\mathbf{gn}_i)!} \quad \Rightarrow \quad P_Y(\mathbf{gn}) = \prod_i e^{-\lambda_i} \frac{\lambda_i^{\mathbf{gn}_i}}{(\mathbf{gn}_i)!}$$

due to the assumption of independence of the data values. Since  $E[Y] = \mathbf{Hf}^* + bg\mathbf{1}$ , we assume  $\lambda = \mathbf{Hf}^* + bg\mathbf{1}$ , so we can write the distribution probability for the multivalued random variable  $Y$  as

$$P_Y(\mathbf{gn}|\mathbf{f}^*) = \prod_i \frac{(\mathbf{Hf}^* + bg\mathbf{1})_i^{\mathbf{gn}_i} e^{-(\mathbf{Hf}^* + bg\mathbf{1})_i}}{\mathbf{gn}_i!} \quad (1.8)$$

The influence of the noise is highly dependent on the value of the exact data: for low counts the registered signal has an high level of noise, while for high counts the signal is less affected by noise. This behaviour is depicted in Figure (1.3).

## 1.4 Data Restoration

The goal of image restoration is to **restore the original data  $\mathbf{f}^*$  from the blurred and noisy data  $\mathbf{gn}$  when the structure of  $\mathbf{H}$  is known**<sup>1</sup>. This is a problem belonging to a wider class of problems, called **inverse problems**.

A naive approach to solve this problem consists in solving the equation (1.4) for  $\mathbf{f}$ , obtaining the following value for the original object  $\mathbf{f}^*$ :

$$\mathbf{H}^{-1}(\mathbf{gn} - bg\mathbf{1})$$

An example of what we could obtain with this procedure applied on the image in Figure 1.2(c) is shown in Figure 1.4.

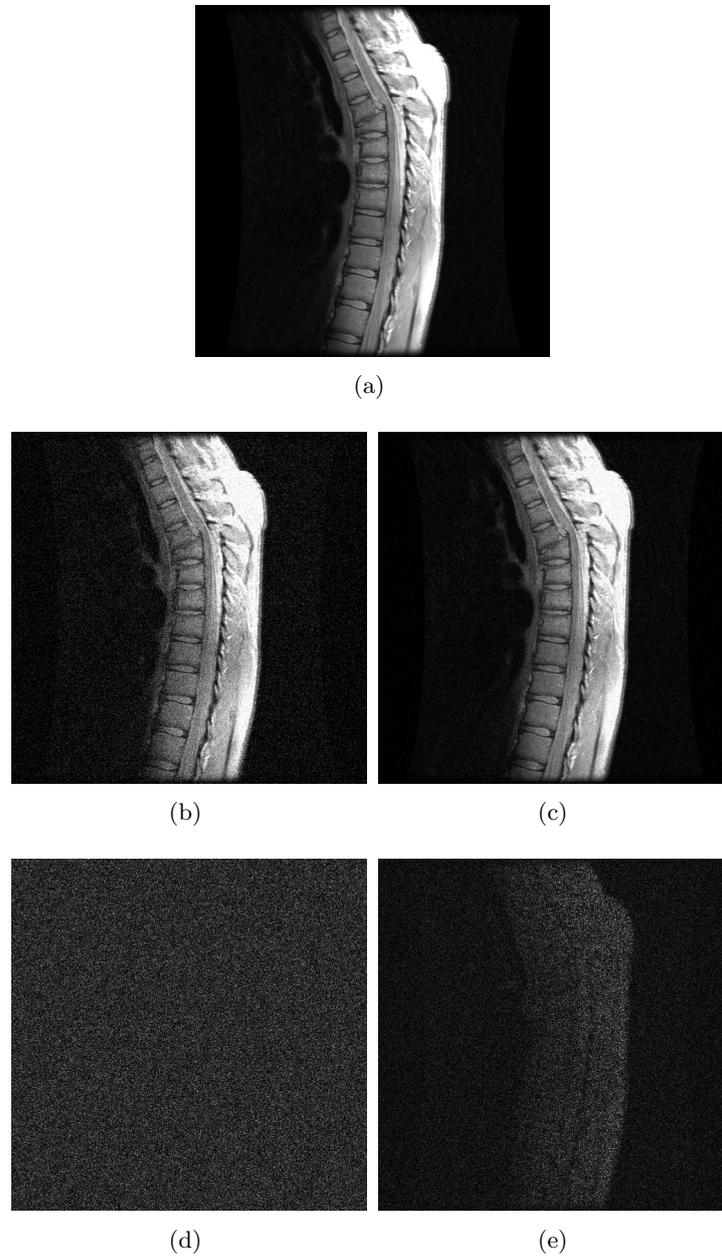
The direct inversion is not suitable for this kind of problems since in most cases they are *ill-posed* in the sense of Hadamard [57].

For discrete data, we have to consider the ill-conditioning of the problem, expressed by the condition number of the imaging matrix  $\mathbf{H}$  which can be singular or nearly singular. Due to the presence of statistical noise, the direct inversion is not viable.

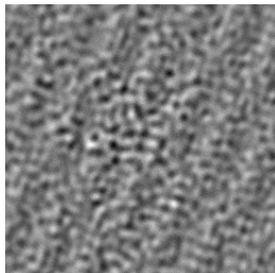
Since it is physically impossible to recover the exact original data  $\mathbf{f}^*$ , our aim is to find an *approximation*  $\mathbf{f}$  of it, taking into account all the available information on  $\mathbf{H}$  and on the noise affecting the image.

In the early 80's, in [88] the *Maximum Likelihood Estimation* approach is proposed; successively the *Bayesian Approach*, described in [53], has attracted more and more attention due to its extension of the previous model. In the following we present the basic ideas of these two approaches.

<sup>1</sup>When the operator  $\mathbf{H}$  is not fully known, we have to face a problem of *blind deconvolution*: see [44, 17].



**Figure 1.3:** *Example of noise affecting data. In 1.3(b) Gaussian noise is present on the original image 1.3(a); in 1.3(c) Poissonian noise had influenced the data acquisition; in 1.3(d) and 1.3(e) there are the differences between the original image and the data affected by Gaussian and Poisson noise, respectively. As pointed out, the Gaussian noise is uniform on the entire image, while on the contrary the other type of noise is highly dependent on the pixels' values.*



**Figure 1.4:** Attempt to reconstruct the image 1.2(a) from 1.2(c) by a naive approach.

### 1.4.1 Maximum Likelihood Estimation

The Maximum Likelihood Estimation approach consists in considering (1.6) as the likelihood function. The *ML estimator* is a vector  $\mathbf{f}_{ML}$  which maximize  $P_Y(\mathbf{gn}|\mathbf{f})$ . Since the likelihood function is the product of a very large number of factors it is convenient to take its negative logarithm: then we are led to solve a minimization problem instead of a maximization one:

$$\mathbf{f}_{ML} = \arg \max_{\mathbf{f}} P_Y(\mathbf{gn}|\mathbf{f}) = \arg \min_{\mathbf{f}} (-\log(P_Y(\mathbf{gn}|\mathbf{f})))$$

Within this approach, the ML estimator is exactly the searched approximation  $\tilde{\mathbf{f}}$ :

$$\tilde{\mathbf{f}} \equiv \mathbf{f}_{ML} = \arg \min_{\mathbf{f}} \varphi_0(\mathbf{Hf} + b\mathbf{g}\mathbf{1}; \mathbf{gn}) \quad (1.9)$$

where  $\varphi_0(\mathbf{Hf} + b\mathbf{g}\mathbf{1}; \mathbf{gn}) = -\log(P_Y(\mathbf{gn}|\mathbf{f}))$ .

**Definition 1.2.** The function  $\varphi_0 : \mathbf{f} \mapsto \varphi_0(\mathbf{Hf} + b\mathbf{g}\mathbf{1}; \mathbf{gn})$  is called *fidelity functional* or *fit-to-data functional*: it measures the discrepancy between the given data  $\mathbf{gn}$  and its argument  $\mathbf{f}$ .

Considering the data  $\mathbf{gn}$  affected by different type of noise, we obtain different functionals  $\varphi_0$ .

**Gaussian noise:** considering (1.7), neglecting some irrelevant constants, the problem (1.9) could be stated as

$$\mathbf{f}_{ML} = \arg \min_{\mathbf{f}} \frac{1}{2} \|\mathbf{Hf} + b\mathbf{g}\mathbf{1} - \mathbf{gn}\|^2 \quad (1.10)$$

i.e. it is the classical *Least Squares* minimization problem. It is worth noticing that this functional is convex, and strictly convex if and only if  $\mathbf{Hf} = 0$  when  $\mathbf{f} = 0$ . Moreover, it has always a global minimum.

**Poisson noise:** from (1.8), using Stirling's formula to approximate the factorial and neglecting again some constants, (1.9) takes the form

$$\mathbf{f}_{ML} = \arg \min_{\mathbf{f}} \mathbf{KL}(\mathbf{Hf} + b\mathbf{g}\mathbf{1}; \mathbf{gn}) \quad (1.11)$$

where

$$\mathbf{KL}(\mathbf{x}; \mathbf{y}) = \sum_{x_i > 0} y_i \log \left( \frac{y_i}{x_i} \right) + x_i - y_i \quad (1.12)$$

is the Kullback–Leibler functional, with  $0 \log(0) = 0$ ; this functional measures the distance between two nonnegative probability distributions. Since in our case the data come from Poisson random variables, it seems the right functional to be employed. We report in our notations two useful results about the properties of the Kullback–Leibler functional.

**Proposition 1.1.** [50, Prop. 2] *For any  $\mathbf{y} \geq 0$  we have  $\text{dom}(\mathbf{KL}(\cdot; \mathbf{y})) = \{\mathbf{x} > 0 \mid \mathbf{KL}(\mathbf{x}; \mathbf{y}) \leq \infty\} \neq \emptyset$ ; furthermore, the function  $\mathbf{KL}(\cdot; \mathbf{y})$  is lower semicontinuous (l.s.c.), coercive and convex on its domain. If  $\mathbf{y} > 0$ ,  $\mathbf{KL}(\cdot; \mathbf{y})$  is also strictly convex. If  $\mathbf{H}$  is a linear operator such that  $\{\mathbf{x} \mid \mathbf{H}\mathbf{x} > \mathbf{0}\} \neq \emptyset$ , we have that the function  $\mathbf{KL} \circ \mathbf{H}$*

$$\mathbf{KL} \circ \mathbf{H}(\mathbf{x}; \mathbf{y}) := \mathbf{KL}(\mathbf{H}\mathbf{x}; \mathbf{y})$$

*is proper, l.s.c. and convex and moreover if  $\mathbf{H}$  is injective, then  $\mathbf{KL} \circ \mathbf{H}$  is coercive. If  $\mathbf{H}$  is injective and  $\mathbf{y} > 0$ , then  $\mathbf{KL} \circ \mathbf{H}$  is strictly convex.*

**Remark 1.1.** Proposition 1.1 holds for  $\mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n})$  with  $b\mathbf{g} > 0$ . We observe that the domain of  $\mathbf{KL} \circ \mathbf{H}$  is  $\{\mathbf{x} \in \mathbb{R}^N \mid \mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1} > \mathbf{0}\}$ . Under suitable assumptions,  $\text{dom}(\mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}))$  is non empty. Furthermore, for  $\mathbf{x} \in \text{dom}(\mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}))$  is twice continuously differentiable.

**Proposition 1.2.** [13] *If  $\mathbf{H}$  satisfies*

$$\mathbf{H}_{ij} \geq 0, \quad \sum_{i=1}^M \mathbf{H}_{ij} > 0 \text{ for } j = 1, \dots, N, \quad \sum_{j=1}^N \mathbf{H}_{ij} > 0 \text{ for } i = 1, \dots, M \quad (1.13)$$

*the  $\mathbf{KL} \circ \mathbf{H}$  is coercive on the nonnegative orthant.*

*Proof.* The assumptions imply that, if the null space of the matrix  $\mathbf{H}$  is not trivial, then all its elements must have at least one negative component. It follows that  $\|\mathbf{H}\mathbf{x}\|$  is coercive on the nonnegative orthant. Indeed, it can not be zero on the intersection of the nonnegative orthant with the unit sphere and therefore has a positive minimum  $\alpha$ . Since it is homogeneous of order 1, it follows  $\|\mathbf{H}\mathbf{x}\| \geq \alpha\|\mathbf{x}\|$ ,  $x \geq 0$ .  $\square$

**Remark 1.2.** When  $\mathbf{H}$  satisfies (1.13) and  $b\mathbf{g} > 0$ ,  $\mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n})$  is gradient Lipschitz–continuous on its domain.

For both types of noise the solution is difficult to compute, since this problem is ill–conditioned, as stated before. The computation of the Least Squares solution leads to search for the solution of the well–known Euler equation

$$\mathbf{H}^t \mathbf{H} \mathbf{f} = \mathbf{H}^t (\mathbf{g}\mathbf{n} - b\mathbf{g}\mathbf{1})$$

and the condition number of  $\mathbf{H}$  can be very large. On the other hand, in presence of Poisson noise, the fit–to–data functional is highly non–linear and, furthermore, the condition number of  $\mathbf{H}$  can make the computation very expensive.

Due to the ill–conditioning of the problem, the basic instability of the minimizer of the likelihood was soon recognized; in order to get a better reconstruction, the Bayesian approach has been widely investigated.

### 1.4.2 Bayesian Approach

In a complete statistical framework, one can consider also the true signal as the realization of a multi-valued random variable  $F$ ; hence we can take into account the probability density function  $P_F(\mathbf{f})$  as *a priori* information.  $P_F(\mathbf{f})$  is the so-called prior probability distribution, or simply the *prior*. Introducing the marginal probability  $P_Y(\mathbf{gn})$ , we are led, via the Bayes theorem, to the conditional probability of  $F$  for a given value  $\mathbf{gn}$  of  $Y$ :

$$P_F(\mathbf{f}|\mathbf{gn}) = \frac{P_Y(\mathbf{gn}|\mathbf{f})P_F(\mathbf{f})}{P_Y(\mathbf{gn})} \quad (1.14)$$

In this way, some useful information (such as sharp edges, high contrasted values) on the object can be incorporated in the *a priori* probability  $P_F(\mathbf{f})$ ; a very common and suitable choice for this function is given by a Gibbs distribution:

$$P_F(\mathbf{f}) = C_N \exp(-\beta\varphi_1(\mathbf{f})) \quad (1.15)$$

with  $C_N \in \mathbb{R}$ ,  $\beta \in \mathbb{R}^+$ ; here  $\varphi_1$  includes all the available information on  $\mathbf{f}^*$ . A reasonable request for  $\varphi_1$  is to be convex.

In the Bayesian approach, the required approximation  $\tilde{\mathbf{f}}$  is computed as a *maximum a posteriori* (MAP) estimator of  $P_F(\mathbf{f}|\mathbf{gn})$ . Again we are led to search for the maximum of a function: for computational purposes we take again the neglog of  $P_F(\mathbf{f}|\mathbf{gn})$ , which means to compute

$$\tilde{\mathbf{f}} = \arg \min_{\mathbf{f}} \varphi_\beta(\mathbf{f}) \equiv \varphi_0(\mathbf{H}\mathbf{f} + b\mathbf{g}\mathbf{1}; \mathbf{gn}) + \beta\varphi_1(\mathbf{f}) \quad (1.16)$$

We have neglected some constants as  $C_N$  and  $P_Y(\mathbf{gn})$ , since they do not affect the minimization procedure.

The function  $\varphi_1$  is called **regularization functional**: it has the role to take into account some properties of the desired solution and to control the influence of the noise. The parameter  $\beta$  is a real, positive number which balances the trade-off between  $\varphi_0$  and  $\varphi_1$ : in a statistical framework it has the name of **hyper parameter**, while in regularization theory it is called **regularization parameter**.

In presence of Gaussian or Poisson noise the  $\varphi_0$  function is differentiable, coercive and convex, so if one chooses for  $\varphi_1$  a convex function, the function  $\varphi_\beta := \varphi_0 + \beta\varphi_1$  has global minimizers for positive value of  $\beta$ . It is worthing to notice that again this formulation does not assure us the computation of a good reconstruction: the parameter  $\beta$  has a huge influence on the computation of  $\tilde{\mathbf{f}}$ .

## 1.5 Constraints

In real applications, the physics lying under the data acquisition process can require some characteristics or *constraints* on the approximated solution  $\tilde{\mathbf{f}}$ . Hence we are led to solve a constrained minimization problem; the requested characteristics can be modeled by searching the solution in a convex set  $\mathcal{C} \subset \mathbb{R}^N$ :

$$\tilde{\mathbf{f}} = \arg \min_{\mathbf{f} \in \mathcal{C}} \varphi_\beta(\mathbf{f})$$

Some typical examples of constraint set are the following.

- $\mathcal{C}$  is the positive orthant: the requested solution must have only nonnegative elements;
- the flux of the data remains constant after the acquisition process:

$$\sum_i \mathbf{f}_i = c, \quad c \in \mathbb{R}$$

- $\mathcal{C}$  is a box, i.e., the elements of the computed solution satisfy box constraints:

$$a_i \leq \mathbf{f}_i \leq b_i, \quad a_i, b_i \in \mathbb{R} \quad i = 1, \dots, N,$$

- $\mathcal{C}$  is a convex, bounded subset of  $\mathbb{R}^N$ .

**Remark 1.3.** An interesting point of view of the minimization procedure can be the following: one can consider the following minimization problem

$$\arg \min \varphi_1(\mathbf{f}) \quad \text{subject to } \varphi_0(\mathbf{H}\mathbf{f} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) \leq \eta$$

i.e. searching a regularized solution within the elements that keep bounded the value of  $\varphi_0$ . Adopting a penalty approach, one can obtain the same model described before, just multiplying both members by  $\beta$ :

$$\arg \min_{\mathbf{f}} \varphi_1(\mathbf{f}) + \frac{1}{\beta} \varphi_0(\mathbf{H}\mathbf{f} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n})$$

The relations between the two formulations have been widely investigated.

## 1.6 Regularization Functionals

In this section we report the forms of the regularization function  $\varphi_1$  used in this work.

### 1.6.1 Tikhonov Regularization

Let  $L$  be a linear operator. The choice

$$\varphi_1(\mathbf{f}) = \frac{1}{2} \|L\mathbf{f}\|^2$$

is known as *Tikhonov regularization* ([93, 94, 95]). The starting point of this kind of regularization is the Euler Equation  $\mathbf{H}^t \mathbf{H}\mathbf{f} = \mathbf{H}^t(\mathbf{g}\mathbf{n} - b\mathbf{g}\mathbf{1})$ , rising from the least squares formulation.

The aim of this kind of regularization is to emphasize the features of smooth objects. Current choices for  $L$  can be:

- $L = \mathbf{I}$ ;
- $L = \nabla$ , where  $\nabla$  is the discretization of the gradient operator;
- $L = \nabla^2$ , where  $\nabla^2$  is the discretization of the Laplacian operator;

More precisely, in the latter cases, in presence of a 2D square image of  $N$  pixels, where  $N = m \times n$ , we have that

$$\nabla = \begin{pmatrix} \nabla_1 \\ \vdots \\ \nabla_N \end{pmatrix}$$

where  $\nabla_i$  is a  $2 \times N$  matrix such that

$$\nabla_i \mathbf{f} = \begin{pmatrix} \mathbf{f}_{i+1} - \mathbf{f}_i \\ \mathbf{f}_{i+m} - \mathbf{f}_i \end{pmatrix}$$

Here forward difference formulae at the  $i$ -th pixel are used. Then the discrete formulation is

$$\frac{1}{2} \|\nabla \mathbf{f}\|^2 = \frac{1}{2} \sum_{i=1}^N \|\nabla_i \mathbf{f}\|^2$$

A similar argument holds for  $\nabla^2$ .

It has to be pointed out that this type of regularization functionals are convex.

### 1.6.2 Edge-Preserving Regularization

In order to preserve sharp discontinuities in the solution, a suitable choice for  $\varphi_1$  is the Total Variation functional [86]. The discrete version for the Total Variation of a digital signal  $\mathbf{f}$  can be written as

$$\mathbf{TV}(\mathbf{f}) = \sum_{i=1}^N \|\nabla_i \mathbf{f}\| \quad (1.17)$$

where  $\nabla_i$  is the discretization of the gradient operator at the  $i$ -th pixel, as in the Tikhonov case.

Using the notation in [97], we can express the discrete  $\mathbf{TV}$  also as

$$\mathbf{TV}(f) = \sum_{i=1}^N \psi \left( (\mathbf{f}_{i+1} - \mathbf{f}_i)^2 + (\mathbf{f}_{i+m} - \mathbf{f}_i)^2 \right)$$

with  $\psi(t) = \sqrt{t}$ .

We have to notice that, also in this case,  $\varphi_1$  is a convex function. On the other hand, the Total Variation functional is not differentiable for  $f$  such that  $\nabla_i f = 0$  for some  $i \in \{1, \dots, N\}$ ; hence, in order to remove the singularity of  $\psi$  at the origin, we can introduce a generalization for this kind of function, using an approximation  $\psi_\delta$  of  $\psi$ . Some common choice for  $\psi_\delta$  can be [97]

$$\psi_\delta(t) = \sqrt{t + \delta^2} - \delta \quad (1.18)$$

$$\psi_\delta(t) = \begin{cases} \frac{t}{\delta} & t \leq \delta^2 \\ 2\sqrt{t} - \delta & t > \delta^2 \end{cases} \quad (1.19)$$

with  $\delta > 0$ . The only requirement is that  $\psi'(t) > 0$  for  $t \geq 0$ . In the first case, for  $\delta$  equal to zero, we obtain the classical Total Variation. The function (1.18) has been

considered in a general approach to edge preserving regularization [39]. In such a case it is called Hypersurface Potential [38]

$$\mathbf{HS}(\mathbf{f}) = \sum_{i=1}^N \psi_{\delta} ((f_{i+1} - f_i)^2 + (f_{i+m} - f_i)^2) = \sum_{i=1}^N \sqrt{\|\nabla_i \mathbf{f}\|^2 + \delta^2} - \delta \quad (1.20)$$

and  $\delta$  is considered as a thresholding parameter which tunes the value of the gradient above which a discontinuity is detected.

Edge-preserving image reconstruction can be obtained also by Markov Random Field (MRF) regularization [53, 54]. The methods which allow to deal with the  $\mathbf{HS}$  potential can be efficiently used when  $\varphi_1$  is a MRF function.

## 1.7 The regularization problem for Poisson data

When  $\mathbf{H}$  satisfies the classical assumptions (1.5), as pointed out in Proposition 1.2, the Kullback–Leibler function is coercive and convex; for  $x \in \text{dom}(\mathbf{KL} \circ \mathbf{H})$ , the Hessian takes the form

$$\nabla^2 \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) = \mathbf{H}^t \text{diag} \left( \frac{\mathbf{g}\mathbf{n}}{(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1})^2} \right) \mathbf{H}$$

The kernel of  $\nabla^2 \mathbf{KL}$  contains the kernel of  $\mathbf{H}$  and, under the previous assumptions on the imaging matrix, we have that  $\mathbf{1} \notin \text{Ker}(\nabla^2 \mathbf{KL})$ . With these preliminary results, the subsequent Proposition assure us the existence and the uniqueness of the solution of the variational problem

$$\min_{\mathbf{x} \geq 0} \varphi_{\beta} \equiv \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta \varphi_1(\mathbf{x}) \quad (1.21)$$

where the fidelity function is the Kullback–Leibler and the regularization is given by a Tikhonov–type or edge-preserving function.

We assume  $b\mathbf{g} > 0$  in order to guarantee the existence of the Hessian of  $\mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n})$  for any  $\mathbf{x} \geq 0$ . This does not give any restriction since one can set  $b\mathbf{g}$  as a very small positive value.

**Proposition 1.3.** *Let  $\mathbf{H}$  satisfies (1.5) in (1.21).*

- (a) *If  $\varphi_1(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2$ , then the solution of (1.21) exists and it is unique.*
- (b) *If  $\varphi_1(\mathbf{x}) = \frac{1}{2}\|L\mathbf{x}\|^2$  and  $\text{Ker}(L) = \{\alpha\mathbf{1} | \alpha \in \mathbb{R}\}$ , then the solution of (1.21) exists and it is unique.*
- (c) *If  $\varphi_1(\mathbf{x}) = \mathbf{HS}(\mathbf{x})$  ( $\delta > 0$ ), then the solution of (1.21) exists and it is unique.*
- (d) *If  $\varphi_1(\mathbf{x}) = \mathbf{TV}(\mathbf{x})$ , then the solution of (1.21) exists; if  $\text{Ker}(\mathbf{H}) = \{\mathbf{0}\}$  and  $\mathbf{g}\mathbf{n} > 0$  the solution is unique.*

*Proof.* In cases (a)–(c) the Hessian of  $\varphi_{\beta}$  is the sum of the Hessian of the  $\mathbf{KL}$  divergence and  $\nabla^2 \varphi_1$ ; we have to show that the intersection  $\text{Ker}(\nabla^2(\mathbf{KL})) \cap \text{Ker}(\nabla^2(\varphi_1)) = \{\mathbf{0}\}$  in order to have a positive definite Hessian matrix.

- (a) Since  $\nabla^2 \varphi_1 = \mathbf{I}$ ,  $\text{Ker}(\nabla^2 \varphi_1) = \{\mathbf{0}\}$ , the intersection is the trivial vector space and the conclusion holds.

(b) Since  $\text{Ker}(\nabla^2\varphi_1) = \text{Ker}(L^tL) = \text{Ker}(L) = \{\alpha\mathbf{1}|\alpha \in \mathbb{R}\}$ , the intersection  $\text{Ker}(\nabla^2(\mathbf{KL})) \cap \text{Ker}(\nabla^2(\varphi_1))$  is just the null vector:  $\varphi_\beta$  is thus strictly convex.

(c) From Definition (1.20), we have  $\mathbf{HS}(\mathbf{x}) = \sum_{i=1}^N \left( \sqrt{D_i} - \delta \right)$ , where  $D_i = \|\nabla_i\mathbf{x}\|^2 + \delta^2$ .

We define the block matrices

$$E(\mathbf{x}) = \text{diag}(D_i^{1/2}\mathbf{I}_2), \quad F(\mathbf{x}) = \text{diag}\left(\mathbf{I}_2 - \frac{1}{D_i}\nabla_i\mathbf{x}\mathbf{x}^t\nabla_i^t\right)$$

for  $i = 1, \dots, N$ . The matrices  $E, F \in \mathcal{M}_{2N}(\mathbb{R})$  are block diagonal matrices. Then, we can write the gradient and the Hessian of  $\varphi_1$  in terms of  $E$  and  $F$ :

$$\nabla\varphi_1(\mathbf{x}) = \nabla^t E(\mathbf{x})^{-1}\nabla\mathbf{x}, \quad \nabla^2\varphi_1(\mathbf{x}) = \nabla^t E(\mathbf{x})^{-1}F(\mathbf{x})\nabla$$

In particular,  $E(\mathbf{x})$  is diagonal with positive diagonal entries and the  $i$ -th diagonal block of  $F(\mathbf{x})$  is the difference between the identity matrix and a dyadic product: from the definition of  $D_i$  it follows that  $\frac{1}{D_i}\mathbf{x}^t\nabla_i^t\nabla_i\mathbf{x} = \frac{\|\nabla_i\mathbf{x}\|^2}{D_i} < 1$ . As a consequence, from the Sherman–Morrison theorem, it follows that all the diagonal blocks of  $F(\mathbf{x})$  (and hence  $F(\mathbf{x})$  itself) are nonsingular. Due to the last considerations, we are going to prove that the null space of the Hessian of  $\varphi_1$  is the set of the minimum points of  $\varphi_1$ .

In fact, we have that  $E(\mathbf{x})^{-1}F(\mathbf{x})$  is nonsingular for all  $\mathbf{x} \in \mathbb{R}^N$ : this leads to  $\text{Ker}(\nabla^2(\varphi_1)) = \text{Ker}(\nabla)$ . Conversely, the minimum points of the convex function  $\varphi_1$  satisfy  $\nabla(\varphi_1) = 0$ , that is  $\nabla^t E(\mathbf{x})^{-1}\nabla\mathbf{x} = 0$ . Hence the set of the minimum points of  $\varphi_1$  is the subspace  $\text{Ker}(\nabla)$ .

Since the minimum points of the  $\mathbf{HS}$  functional are all the points  $\mathbf{x}_i = \alpha, \alpha \in \mathbb{R}$  for all  $i$ , we thus have  $\text{Ker}(\nabla^2\mathbf{HS}) = \{\alpha\mathbf{1}|\alpha \in \mathbb{R}\}$ , then the intersection between the null space of the Hessians of  $\varphi_0$  and  $\varphi_1$  is just the zero vector:  $\varphi_\beta$  is thus strictly convex.

(d) In this case, see [50, Proposition 3].

□

The assumption  $\mathbf{gn} > 0$  is not very restrictive, since in real applications  $\mathbf{gn}$  contains also the background emission. When this does not occur, it is sufficient to set the zero value in the given data to very small values, with no side effects.

Furthermore, the general assumption  $\text{Ker}(L) = \{\alpha\mathbf{1}|\alpha \in \mathbb{R}\}$  holds for  $L = \nabla$  or  $L = \nabla^2$  when classic boundary conditions, such as periodic or reflexive ones, are imposed on the data; for the identity operator or for zero boundary conditions the kernel is just the trivial vector space ( $\alpha = 0$ ).

## Chapter 2

# Optimization Methods

In the previous chapter we have explained the way we are led to solve minimization problems in image restoration framework. The aim of this chapter is to give a survey on the optimization algorithms that are the state of the art in imaging applications. We mainly focus on the case of data corrupted by Poisson noise and on the methods that will be used in the subsequent chapters. Some algorithms (SGP) allow us to solve the primal formulation of the minimization of a convex function while others are employed for primal–dual formulation (AEM, PDHG, ...), taking into account also the non differentiable case. Furthermore, as original contribution, we introduce a new scaling technique with the aim of accelerating the class of the  $\varepsilon$ -subgradient methods. The problem to be faced is

$$\min_{\mathbf{x} \in \Omega} \varphi(\mathbf{x}) \quad (2.1)$$

where  $\varphi$  is a convex real function and  $\Omega$  is a closed convex set of  $\mathbb{R}^N$ .

### 2.1 SGP

The Scaled Gradient Projection (SGP) algorithm [22] is an optimization method belonging to the class of gradient projection methods, whose aim is to solve the problem (2.1) where  $\varphi$  is continuously *differentiable*. Focusing on image restoration problems, the choice of  $\varphi_\beta$  (1.16) as  $\varphi$  in (2.1) is a special case of this formulation for differentiable  $\varphi_1$  (Tikhonov or **HS** regularization, for example).

Before describing the algorithm and its properties, we recall some definitions and features of projection operators which become useful in the forthcoming sections.

**Definition 2.1.** Let  $D$  a symmetric positive definite matrix,  $D \in \mathcal{M}_N(\mathbb{R})$ ; the function

$$\begin{aligned} \|\cdot\|_D &: \mathbb{R}^N \rightarrow \mathbb{R} \\ \|\cdot\|_D &: \mathbf{x} \mapsto \|\mathbf{x}\|_D = \sqrt{\mathbf{x}^t D \mathbf{x}} \end{aligned}$$

is the norm associated to the matrix  $D$ . The projector operator  $P_{\Omega, D}$  on the closed set  $\Omega \subset \mathbb{R}^n$  is defined as

$$P_{\Omega, D}(\mathbf{x}) = \arg \min_{\mathbf{y} \in \Omega} \|\mathbf{y} - \mathbf{x}\|_D^2 \quad (2.2)$$

When  $D = \mathbf{I}$ , we have the standard Euclidean projection and we write  $P_{\Omega, \mathbf{I}} = P_\Omega$ . In the following lemma we state some properties of  $P_{\Omega, D}$ .

**Lemma 2.1.** *Let  $D$  be a symmetric positive definite matrix. The following relations hold.*

(a) [16, Proposition 3.7] *For any  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{y} \in \Omega$  we have*

$$(\mathbf{P}_{\Omega, D^{-1}}(\mathbf{x}) - \mathbf{x})^t D^{-1}(\mathbf{y} - \mathbf{P}_{\Omega, D^{-1}}(\mathbf{x})) \geq 0 \quad (2.3)$$

$$(\mathbf{P}_{\Omega, D^{-1}}(\mathbf{x}) - \mathbf{x})^t D^{-1}(\mathbf{y} - \mathbf{x}) \geq \|\mathbf{x} - \mathbf{P}_{\Omega, D^{-1}}(\mathbf{x})\|_{D^{-1}}^2 \quad (2.4)$$

(b) *For any  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  we have*

$$\|\mathbf{P}_{\Omega, D^{-1}}(\mathbf{x}) - \mathbf{P}_{\Omega, D^{-1}}(\mathbf{y})\|_{D^{-1}} \leq \|\mathbf{x} - \mathbf{y}\|_{D^{-1}}$$

(c) [22, Lemma 2.1] *Let  $L$  a positive number such that  $\|D\| = \lambda_{\max}(D) \leq L$  and  $\|D^{-1}\| = \frac{1}{\lambda_{\min}(D)} \leq L$ , where  $\lambda_{\max}$  and  $\lambda_{\min}$  are the maximum and the minimum eigenvalue of  $D$ , respectively; then*

$$\|\mathbf{P}_{\Omega, D^{-1}}(\mathbf{x}) - \mathbf{P}_{\Omega, D^{-1}}(\mathbf{y})\| \leq L\|\mathbf{x} - \mathbf{y}\|$$

*Proof.* The optimality conditions of the minimum problem (2.2) yield part (a). Part (b) is a consequence of the first inequality in (a) and of the Cauchy–Schwartz inequality. For part (c), thanks to (b), we have

$$\begin{aligned} \|\mathbf{P}_{\Omega, D^{-1}}(\mathbf{x}) - \mathbf{P}_{\Omega, D^{-1}}(\mathbf{y})\|_{D^{-1}} &\leq \|\mathbf{x} - \mathbf{y}\|_{D^{-1}} \\ &\leq \sqrt{\lambda_{\max}(D^{-1})}\|\mathbf{x} - \mathbf{y}\| \\ &= \sqrt{\frac{1}{\lambda_{\min}(D)}}\|\mathbf{x} - \mathbf{y}\| \\ &\leq \sqrt{L}\|\mathbf{x} - \mathbf{y}\| \end{aligned}$$

and

$$\begin{aligned} \|\mathbf{P}_{\Omega, D^{-1}}(\mathbf{x}) - \mathbf{P}_{\Omega, D^{-1}}(\mathbf{y})\|_{D^{-1}} &\geq \sqrt{\lambda_{\min}(D^{-1})}\|\mathbf{P}_{\Omega, D^{-1}}(\mathbf{x}) - \mathbf{P}_{\Omega, D^{-1}}(\mathbf{y})\| \\ &= \sqrt{\frac{1}{\lambda_{\max}(D)}}\|\mathbf{P}_{\Omega, D^{-1}}(\mathbf{x}) - \mathbf{P}_{\Omega, D^{-1}}(\mathbf{y})\| \\ &\geq \sqrt{\frac{1}{L}}\|\mathbf{P}_{\Omega, D^{-1}}(\mathbf{x}) - \mathbf{P}_{\Omega, D^{-1}}(\mathbf{y})\| \end{aligned}$$

and then we have the result.  $\square$

In the past years, several algorithms have been developed with the aim to perform in a easy and non-expensive way (in terms of computational cost) the projection even in case of sets of particular structures ([45, 26, 72]).

The subsequent lemma characterizes the action of the projection operator on stationary points for the problem (2.1).

**Lemma 2.2.** *A vector  $\mathbf{x}^*$  is a stationary point for (2.1) if and only if*

$$\mathbf{x}^* = P_{\Omega, D^{-1}}(\mathbf{x}^* - \alpha D \nabla \varphi(\mathbf{x}^*))$$

for any positive scalar  $\alpha$  and any symmetric definite positive matrix  $D$ .

*Proof.* Let  $\alpha \in \mathbb{R}^+$  and  $D$  be a symmetric positive definite matrix. Assume that  $\mathbf{x}^* = P_{\Omega, D^{-1}}(\mathbf{x}^* - \alpha D \nabla \varphi(\mathbf{x}^*))$ ; from the definition of the projection operator we have

$$\begin{aligned} (\mathbf{x}^* - \mathbf{x} + \alpha D \nabla \varphi(\mathbf{x}^*))^t D^{-1}(\mathbf{x}^* - \mathbf{x}) &\leq 0 \\ \alpha (D \nabla \varphi(\mathbf{x}^*))^t D^{-1}(\mathbf{x}^* - \mathbf{x}) &\leq 0 \\ \nabla \varphi(\mathbf{x}^*)^t (\mathbf{x}^* - \mathbf{x}) &\leq 0 \end{aligned}$$

$\forall \mathbf{x} \in \Omega$ . The last one is exactly the stationarity condition for (2.1). On the other hand, let us assume that  $\mathbf{x}^* \in \Omega$  is a stationary point for (2.1) and suppose  $\bar{\mathbf{x}} = P_{\Omega, D^{-1}}(\mathbf{x}^* - \alpha D \nabla \varphi(\mathbf{x}^*))$ ,  $\bar{\mathbf{x}} \neq \mathbf{x}^*$ . Then, from (2.3), we have

$$(\bar{\mathbf{x}} - \mathbf{x}^* + \alpha D \nabla \varphi(\mathbf{x}^*))^t D^{-1}(\bar{\mathbf{x}} - \mathbf{x}^*) \leq 0$$

that is

$$\|\bar{\mathbf{x}} - \mathbf{x}^*\|_{D^{-1}}^2 + \alpha \nabla \varphi(\mathbf{x}^*)^t (\bar{\mathbf{x}} - \mathbf{x}^*) \leq 0$$

which yields

$$\nabla \varphi(\mathbf{x}^*)^t (\bar{\mathbf{x}} - \mathbf{x}^*) \leq -\frac{\|\bar{\mathbf{x}} - \mathbf{x}^*\|_{D^{-1}}^2}{\alpha} < 0$$

which gives a contradiction with the stationarity assumptions on  $\mathbf{x}^*$ .  $\square$

When  $\tilde{\mathbf{x}}$  is not a stationary point,  $P_{\Omega, D^{-1}}(\tilde{\mathbf{x}} - \alpha D \nabla \varphi(\tilde{\mathbf{x}})) - \tilde{\mathbf{x}}$  can be exploited as a descent direction of  $\varphi$  at  $\tilde{\mathbf{x}}$ . This is the basic idea for the SGP method. The main steps of SGP are reported as Algorithm 1.

The direction  $\mathbf{d}^k$  in STEP 3 is a descent direction for the function  $\varphi$  at  $\mathbf{x}^k$ . Moreover, if the succession  $\{\mathbf{x}^k\}$  is bounded, then  $\{\mathbf{d}^k\}$  is bounded too [22, Lemma 2.4]. Furthermore, using the following theoretical results we can state a convergence theorem for the SGP method.

**Lemma 2.3.** *Assume that the subsequence  $\{\mathbf{x}^k\}_{k \in K}$ ,  $K \subset \mathbb{N}$  is converging to a point  $\mathbf{x}^* \in \Omega$ . Then,  $\mathbf{x}^*$  is a stationary point of (2.1) if and only if*

$$\lim_{k \in K} \nabla \varphi(\mathbf{x}^k)^t \mathbf{d}^k = 0.$$

**Lemma 2.4.** *Let  $\mathbf{x}^* \in \Omega$  be an accumulation point of the sequence  $\{\mathbf{x}^k\}$  such that  $\lim_{k \in K} \mathbf{x}^k = \mathbf{x}^*$ , for some  $K \subset \mathbb{N}$ . If  $\mathbf{x}^*$  is a stationary point of (2.1), then  $\mathbf{x}^*$  is an accumulation point also for the sequence  $\{\mathbf{x}^{k+r}\}_{k \in K}$  for any  $r \in \mathbb{N}$ . Furthermore,*

$$\lim_{k \in K} \|\mathbf{d}^{k+r}\| = 0, \quad \forall r \in \mathbb{N}.$$

The proofs of the previous lemmas can be found in [22]; thanks to these results, the following theorem holds. For sake of completeness we report the proof in our notations.

**Algorithm 1: SGP**


---

Choose the starting point  $\mathbf{x}^0 \in \Omega$ , set the parameters  $\gamma, \theta \in (0, 1)$ ,  $0 < \alpha_{\min} < \alpha_{\max}$  and fix a positive integer  $M$ .

FOR  $k = 0, 1, 2, \dots$  DO THE FOLLOWING STEPS:

STEP 1. Choose the parameter  $\alpha_k \in [\alpha_{\min}, \alpha_{\max}]$  and the scaling matrix  $D_k$ ;

STEP 2. Projection:  $\mathbf{y}^k = P_{\Omega, D^{-1}}(\mathbf{x}^k - \alpha_k D_k \nabla \varphi(\mathbf{x}^k))$ ; if  $\mathbf{y}^k = \mathbf{x}^k$ , then stop, declaring  $\mathbf{x}^k$  is a stationary point;

STEP 3. Descent direction:  $\mathbf{d}^k = \mathbf{y}^k - \mathbf{x}^k$ ;

STEP 4. Set  $\lambda_k = 1$  and  $\varphi_{\max} = \max_{0 \leq j \leq \min\{k, M-1\}} \varphi(\mathbf{x}^{k-j})$

STEP 5. Backtracking loop:  
 IF  $\varphi(\mathbf{x}^k + \lambda_k \mathbf{d}^k) \leq \varphi_{\max} + \gamma \lambda_k \nabla \varphi(\mathbf{x}^k)^t \mathbf{d}^k$   
 go to STEP 6;  
 ELSE  
 set  $\lambda_k = \theta \lambda_k$  and go to STEP 5;

STEP 6. Set  $\mathbf{x}^{k+1} = \mathbf{x}^k + \lambda_k \mathbf{d}^k$

---

**Proposition 2.1.** [22, Theorem 2.1] *Assume that the level set  $\Omega_0 = \{\mathbf{x} \in \Omega \mid \varphi(\mathbf{x}) \leq \varphi(\mathbf{x}^0)\}$  is bounded. Every accumulation point of the sequence  $\{\mathbf{x}^k\}$  generated by the SGP algorithm is a stationary point for  $\varphi$  in (2.1).*

*Proof.* Since every iterate  $\mathbf{x}^k$  lies in  $\Omega_0$ , the sequence  $\{\mathbf{x}^k\}$  is bounded and has at least one accumulation point. Let  $\mathbf{x}^* \in \Omega$  be such that  $\lim_{k \in K} \mathbf{x}^k = \mathbf{x}^*$  for a set of indexes  $K \subset \mathbb{N}$ . Let us consider separately the two cases

- a.  $\inf_{k \in K} \lambda_k = 0$ ;
- b.  $\inf_{k \in K} \lambda_k = \rho > 0$ .

Case a.

Let  $K_1 \subset K$  be a set of indexes such that  $\lim_{k \in K_1} \lambda_k = 0$ . This implies that, for  $k \in K_1$ ,  $k$  sufficiently large, the backtracking rule fails to be satisfied at least once. Thus, at the penultimate step of the backtracking loop, we have

$$\varphi\left(\mathbf{x}^k + \frac{\lambda_k}{\theta} \mathbf{d}^k\right) > \varphi(\mathbf{x}^k) + \gamma \frac{\lambda_k}{\theta} \nabla \varphi(\mathbf{x}^k)^t \mathbf{d}^k,$$

hence

$$\frac{\varphi\left(\mathbf{x}^k + \frac{\lambda_k}{\theta} \mathbf{d}^k\right) - \varphi(\mathbf{x}^k)}{\frac{\lambda_k}{\theta}} > \gamma \nabla \varphi(\mathbf{x}^k)^t \mathbf{d}^k. \quad (2.5)$$

By the mean value theorem, we have that there exists a scalar  $t_k \in \left[0, \frac{\lambda_k}{\theta}\right]$  such that the left hand side of (2.5) is equal to  $\nabla \varphi(\mathbf{x}^k + t_k \mathbf{d}^k)^t \mathbf{d}^k$ . Thus, the inequality (2.5) becomes

$$\nabla \varphi(\mathbf{x}^k + t_k \mathbf{d}^k)^t \mathbf{d}^k > \gamma \nabla \varphi(\mathbf{x}^k)^t \mathbf{d}^k. \quad (2.6)$$

Since  $\alpha_k$  and  $D_k$  are bounded, it is possible to find a set of indexes  $K_2 \subset K_1$  such that  $\lim_{k \in K_2} \alpha_k = \alpha_*$  and  $\lim_{k \in K_2} D_k = D_*$ . Thus the sequence  $\{\mathbf{d}^k\}_{k \in K_2}$  converges to the vector

$$\mathbf{d}^* = (P_{\Omega, D_*^{-1}}(\mathbf{x}^* - \alpha_* D_* \nabla \varphi(\mathbf{x}^*)) - \mathbf{x}^*)$$

and, furthermore,  $t_k \mathbf{d}^k \rightarrow 0$  when  $k$  diverges,  $k \in K_2$ . Taking limits in (2.6) as  $k \rightarrow \infty$ ,  $k \in K_2$ , we obtain

$$(1 - \gamma) \nabla \varphi(\mathbf{x}^*)^t \mathbf{d}^* \geq 0.$$

Since  $(1 - \gamma) > 0$  and  $\nabla \varphi(\mathbf{x}^k)^t \mathbf{d}^k < 0$  for all  $k$ , then we necessarily have  $\lim_{k \in K_2} \nabla \varphi(\mathbf{x}^k)^t \mathbf{d}^k =$

$\nabla \varphi(\mathbf{x}^*)^t \mathbf{d}^* = 0$ . Then, by Lemma 2.3, we conclude that  $\mathbf{x}^*$  is a stationary point.

Case b.

Let us define the point  $\mathbf{x}^{\ell(k)}$  as the point such that

$$\varphi(\mathbf{x}^{\ell(k)}) = \varphi_{\max} = \max_{0 \leq j \leq \min(k, M-1)} \varphi(\mathbf{x}^{k-j}).$$

Then, for  $k > M - 1$ ,  $k \in \mathbb{N}$ , the following condition holds:

$$\varphi(\mathbf{x}^{\ell(k)}) \leq \varphi(\mathbf{x}^{\ell(k)-1}) + \gamma \lambda_{\ell(k)-1} \nabla \varphi(\mathbf{x}^{\ell(k)-1})^t \mathbf{d}^{\ell(k)-1}. \quad (2.7)$$

Since the iterates  $\mathbf{x}^k$ ,  $k \in \mathbb{N}$  belong to a bounded set, the monotone non-increasing sequence  $\{\varphi(\mathbf{x}^{\ell(k)})\}$  admits a finite limit  $\mathcal{L} \in \mathbb{R}$  for  $k \in K$ . Let  $K_3 \subset K$  be a set of indexes such that  $\lim_{k \in K_3} \lambda_{\ell(k)-1} = \rho_1 \geq \rho > 0$  and  $\lim_{k \in K_3} \nabla \varphi(\mathbf{x}^{\ell(k)-1})^t \mathbf{d}^{\ell(k)-1}$  exists (recall that, from [22, Lemma 2.4], the sequence  $\{\mathbf{d}^k\}_{k \in \mathbb{N}}$  is bounded); taking limits on (2.7) for  $k \in K_3$  we obtain

$$\mathcal{L} \leq \mathcal{L} + \gamma \rho_1 \lim_{k \in K_3} \nabla \varphi(\mathbf{x}^{\ell(k)-1})^t \mathbf{d}^{\ell(k)-1},$$

that is

$$\lim_{k \in K_3} \nabla \varphi(\mathbf{x}^{\ell(k)-1})^t \mathbf{d}^{\ell(k)-1} \geq 0.$$

Recalling that  $\nabla \varphi(\mathbf{x}^k)^t \mathbf{d}^k < 0$  for any  $k$ , the previous inequality implies that

$$\lim_{k \in K_3} \nabla \varphi(\mathbf{x}^{\ell(k)-1})^t \mathbf{d}^{\ell(k)-1} = 0. \quad (2.8)$$

Then, by the Lemma 2.3, (2.8) implies that every accumulation point of the sequence  $\{\mathbf{x}^{\ell(k)-1}\}_{k \in K_3}$  is a stationary point of (2.1). Let us prove that the point  $\mathbf{x}^*$  is an accumulation point of  $\{\mathbf{x}^{\ell(k)-1}\}_{k \in K_3}$ .

The definition of  $\mathbf{x}^{\ell(k)}$  implies that  $k - M + 1 \leq \ell(k) \leq k$ . Thus we can write

$$\|\mathbf{x}^k - \mathbf{x}^{\ell(k)-1}\| \leq \sum_{j=0}^{k-\ell(k)} \lambda_{\ell(k)-1+j} \|\mathbf{d}^{\ell(k)-1+j}\|, \quad k \in K. \quad (2.9)$$

Let  $K_4 \subset K_3$  be a subset of indexes such that the sequence  $\{\mathbf{x}^{\ell(k)-1}\}_{k \in K_4}$  converges to an accumulation point  $\bar{\mathbf{x}} \in \Omega$ . Recalling that, from (2.8) and Lemma 2.3,  $\bar{\mathbf{x}}$  is a stationary point of (2.1), we can apply Lemma 2.4 to obtain that  $\lim_{k \in K_4} \|\mathbf{d}^{\ell(k)-1+j}\| = 0$

for any  $j \in \mathbb{N}$ . By using (2.9) we conclude that

$$\lim_{k \in K_4} \|\mathbf{x}^k - \mathbf{x}^{\ell(k)-1}\| = 0. \quad (2.10)$$

Since

$$\|\mathbf{x}^* - \mathbf{x}^{\ell(k)-1}\| \leq \|\mathbf{x}^k - \mathbf{x}^{\ell(k)-1}\| + \|\mathbf{x}^k - \mathbf{x}^*\|$$

and  $\lim_{k \in K} \mathbf{x}^k = \mathbf{x}^*$ , then (2.10) implies that  $\mathbf{x}^*$  is an accumulation point also for the sequence  $\{\mathbf{x}^{\ell(k)-1}\}_{k \in K_3}$ . Hence, we conclude that  $\mathbf{x}^*$  is a stationary point of (2.1).  $\square$

We remark that SGP method is working for *any symmetric definite positive matrix*  $D_k$  and for *any positive scalar*  $\alpha_k$ , consequently the setting of these two parameters can be done dependently on the model which is considered in the application.

In the following we describe the choices made for image restoration problems with data corrupted by Poisson noise:

$$\varphi(\mathbf{x}) = \varphi_\beta(\mathbf{x}) \equiv \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta\varphi_1(\mathbf{x})$$

### 2.1.1 Scaling Matrix

The introduction of a variable scaling matrix  $D_k$  has two main goals: reduce computational costs and improve the rate of convergence. A classical way to perform these tasks is to choose a diagonal matrix  $D_k = \text{diag}(d_1^k, d_2^k, \dots, d_N^k)$  such that  $D_k$  approximates the inverse of the Hessian of  $\varphi_\beta$ :

$$d_i^k = \left( \frac{\partial^2 \varphi_\beta(\mathbf{x}^k)}{(\partial \mathbf{x}_i)^2} \right)^{-1}, \quad i = 1, \dots, N$$

Since the computation of the Hessian could be very expansive, a suitable choice is to take an approximation of it.

When  $\beta = 0$  (i.e. no regularization term), a classical choice is :

$$d_i^k = \min \left\{ L, \max \left\{ \frac{1}{L}, x_i^k \right\} \right\}, \quad i = 1, \dots, N$$

with  $L > 0$ . For a theoretical discussion, see [22].

In presence of a regularization term, following [67], a strategy for a convenient choice for  $D_k$  is based on a splitting of the gradient of  $\varphi_\beta$ . We consider a splitting of the gradient of  $\varphi_1$ :

$$-\nabla \varphi_1(\mathbf{x}) = U(\mathbf{x}) - V(\mathbf{x})$$

with  $U(x) \geq 0, V(x) > 0$  for any  $x$ . Since  $\varphi_0(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) \equiv \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n})$ , its gradient is  $\mathbf{H}^t \left( \mathbf{1} - \frac{\mathbf{g}\mathbf{n}}{\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}} \right)$ . Consequently a splitting of the complete gradient of the function  $\varphi_\beta$  is

$$-\nabla \varphi_\beta(\mathbf{x}; \mathbf{y}) = \left( \mathbf{H}^t \frac{\mathbf{g}\mathbf{n}}{\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}} + \beta U(\mathbf{x}) \right) - (\mathbf{H}^t \mathbf{1} + \beta V(\mathbf{x}))$$

A suitable choice in this case is the following [101]

$$d_i^k = \min \left\{ L, \max \left\{ \frac{1}{L}, \frac{\mathbf{x}_i^k}{\mathbf{H}^t \mathbf{1} + \beta V(\mathbf{x}_i^k)} \right\} \right\}, \quad i = 1, \dots, N$$

### 2.1.2 Steplength Selection

The steplength  $\alpha$  has a huge impact on the convergence rate; hence the selection of this parameter had become of great interest in the past years. As described in [22], SGP method exploits the properties of Barzilai–Borwein rules [6]:

$$\alpha_k^1 = \frac{\mathbf{s}^{k-1} D_k^{-1} D_k^{-1} \mathbf{s}^{k-1}}{(\mathbf{s}^{k-1})^t D_k^{-1} \mathbf{z}^{k-1}}, \quad \alpha_k^2 = \frac{(\mathbf{s}^{k-1})^t D_k \mathbf{z}^{k-1}}{(\mathbf{z}^{k-1})^t D_k D_k \mathbf{z}^{k-1}} \quad (2.11)$$

where  $\mathbf{s}^k = \mathbf{x}^k - \mathbf{x}^{k-1}$  and  $\mathbf{z}^k = \nabla \varphi_\beta(\mathbf{x}^k) - \nabla \varphi_\beta(\mathbf{x}^{k-1})$ . Using the original ideas in [6], we recall the algorithm described in [22] to choose in a suitable way the steplength  $\alpha$  in our problem setting.

---

#### Algorithm 2: ABB Steplength Selection

---

```

IF  $k = 0$ 
  set  $\alpha_0 \in [\alpha_{\min}, \alpha_{\max}]$ ,  $\tau_1 \in (0, 1)$  and a nonnegative integer  $M$ .
ELSE
  IF  $(\mathbf{s}^{k-1})^t D_k^{-1} \mathbf{z}^{k-1} \leq 0$ 
     $\alpha_k^1 = \alpha_{\max}$ 
  ELSE
     $\alpha_k^1 = \max \left\{ \alpha_{\min}, \min \left\{ \frac{(\mathbf{s}^{k-1})^t D_k^{-1} D_k^{-1} \mathbf{s}^{k-1}}{(\mathbf{s}^{k-1})^t D_k^{-1} \mathbf{z}^{k-1}}, \alpha_{\max} \right\} \right\}$ 
  IF  $(\mathbf{s}^{k-1})^t D_k \mathbf{z}^{k-1} \leq 0$ 
     $\alpha_k^2 = \alpha_{\max}$ 
  ELSE
     $\alpha_k^2 = \max \left\{ \alpha_{\min}, \min \left\{ \frac{(\mathbf{s}^{k-1})^t D_k \mathbf{z}^{k-1}}{(\mathbf{z}^{k-1})^t D_k D_k \mathbf{z}^{k-1}}, \alpha_{\max} \right\} \right\}$ 
  IF  $\frac{\alpha_k^2}{\alpha_k^1} \leq \tau_k$ 
     $\alpha_k = \min \{ \alpha_j^2, j = \max \{ 1 - M \}, \dots, k \}$ ,  $\tau_{k+1} = 0.9\tau_k$ 
  ELSE
     $\alpha_k = \alpha_k^1$ ,  $\tau_{k+1} = 1.1\tau_k$ 

```

---

Algorithm 2 describes an updating strategy based on an alternating rule for the steplength [52] which is very efficient in applications.

Another choice for the steplength updating is the strategy recently proposed in [81] (see [51] for unconstrained optimization). The principal aim of this strategy is to acquire second order information by considering a small number (namely  $m$ ) of gradients computed in the previous iterations. These gradients enable to determine approximations of eigenvalues of the Hessian matrix, named Ritz values. The reciprocal of these values are used as steplengths. We refer to [51, 81] for further details. In [81] it is shown that the SGP version exploiting this Ritz-like steplength selection strategy shows a better efficiency compared to the alternating Barzilai–Borwein rules.

## 2.2 EM-TV

In this section, we recall the well-known EM-TV method [30]. Such a method allows to compute the approximate solution of a deblurring problem by Total Variation

regularization from data corrupted by Poisson noise:

$$\min_{\mathbf{x} \geq 0} \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta \mathbf{TV}(\mathbf{x})$$

The general procedure is described in Algorithm 3.

---

**Algorithm 3: EM-TV**

---

Choose  $\beta$  and  $\mathbf{x}^0$ .

FOR  $k = 0, 1, 2, \dots$  DO THE FOLLOWING STEPS:

STEP 1.  $\mathbf{x}^{k+\frac{1}{2}} = \mathbf{x}^k \mathbf{H}^T \frac{\mathbf{g}\mathbf{n}}{\mathbf{H}\mathbf{x}^k + b\mathbf{g}\mathbf{1}}$

STEP 2.  $\mathbf{x}^{k+1} = \arg \min_{\mathbf{x}} \left\{ \frac{1}{2} \left\| \frac{\mathbf{x} - \mathbf{x}^{k+\frac{1}{2}}}{\sqrt{\mathbf{x}^k}} \right\|^2 + \beta \mathbf{TV}(\mathbf{x}) \right\}$

---

The first step (the computation of  $\mathbf{x}^{k+\frac{1}{2}}$ ) is the EM [53] step (known in astronomic field as Lucy–Richardson algorithm [70, 83]), while the second step consists in the solution of a weighted ROF model [86]. Since an iterative solver is used for finding the solution of the TV step, the whole procedure consists in two innested cycles: this fact can be quite expansive in term of computational cost. Thus the effectiveness of EM-TV is strongly dependent on the performance of the inner solver for  $\mathbf{x}^{k+1}$ . A well-known method for the solution of the ROF model is developed in [3]; for other solvers, see [35, 104].

## 2.3 AEM

The aim of extragradient methods is to compute the saddle point of a function  $F$ :

$$\min_{\mathbf{x} \in \mathcal{X}} \max_{\mathbf{y} \in \mathcal{Y}} F(\mathbf{x}, \mathbf{y})$$

where  $\mathcal{X}$  and  $\mathcal{Y}$  are two closed and convex sets and  $F$  is a smooth convex–concave function on  $\mathcal{D} = \mathcal{X} \times \mathcal{Y}$ . Starting from  $\mathbf{x}^0 \in \mathcal{X}$ ,  $\mathbf{y}^0 \in \mathcal{Y}$ , a sequence of iterates can be generated by the following formulae:

$$\bar{\mathbf{y}}_k = P_{\mathcal{Y}} \left( \mathbf{y}^k + \alpha_k \nabla_{\mathbf{y}} F(\mathbf{x}^k, \mathbf{y}^k) \right) \quad (2.12)$$

$$\mathbf{x}^{k+1} = P_{\mathcal{X}} \left( \mathbf{x}^k - \alpha_k \nabla_{\mathbf{x}} F(\mathbf{x}^k, \bar{\mathbf{y}}^k) \right) \quad (2.13)$$

$$\mathbf{y}^{k+1} = P_{\mathcal{Y}} \left( \mathbf{y}^k + \alpha_k \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \mathbf{y}^k) \right) \quad (2.14)$$

$$(2.15)$$

The steplength  $\alpha_k$  is adaptively chosen; indeed, we define

$$\begin{aligned} A_k &= \frac{\|\nabla_{\mathbf{x}}F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k) - \nabla_{\mathbf{x}}F(\mathbf{x}^k, \bar{\mathbf{y}}^k)\|}{\|\mathbf{x}^{k+1} - \mathbf{x}^k\|} \\ B_k &= \frac{\|\nabla_{\mathbf{y}}F(\mathbf{x}^{k+1}, \mathbf{y}^k) - \nabla_{\mathbf{y}}F(\mathbf{x}^k, \mathbf{y}^k)\|}{\|\mathbf{x}^{k+1} - \mathbf{x}^k\|} \\ C_k &= \frac{\|\nabla_{\mathbf{y}}F(\mathbf{x}^{k+1}, \mathbf{y}^k) - \nabla_{\mathbf{y}}F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k)\|}{\|\mathbf{y}^k - \bar{\mathbf{y}}^k\|} \end{aligned} \quad (2.16)$$

The convergence to a saddle point of  $F$  is assured when  $\alpha_k$  is chosen in a bounded interval  $[\alpha_{\min}, \alpha_{\max}]$ ,  $0 < \alpha_{\min} < \alpha_{\max}$  so that the following conditions hold

$$\begin{cases} 1 - 2\alpha_k A_k - 2\alpha_k^2 B_k^2 \geq \varepsilon \\ 1 - 2\alpha_k C_k \geq \varepsilon \end{cases} \quad (2.17)$$

where  $\varepsilon$  is a small fixed value in  $(0, 1)$ . The following lemma allows to prove the convergence property in Proposition 2.2.

**Lemma 2.5.** [73] *Let  $\Omega \subset \mathbb{R}^N$  be a non empty, closed convex set,  $\mathbf{w}, \mathbf{z} \in \mathbb{R}^n$  and  $\mathbf{u} \in \Omega$ . Then*

$$\|\mathbf{P}_{\Omega}(\mathbf{z}) - \mathbf{u}\|^2 \leq \|\mathbf{z} - \mathbf{u}\|^2 - \|\mathbf{P}_{\Omega}(\mathbf{z}) - \mathbf{z}\|^2 \quad (2.18)$$

**Proposition 2.2.** [20, Theorem 1] *Assume that  $F$  is convex with respect to  $\mathbf{x}$  and concave with respect to  $\mathbf{y}$  in the domain  $\mathcal{D}$  and that there exists a saddle point of  $F$  in  $\mathcal{D}$ . Let  $\{\mathbf{x}^k, \mathbf{y}^k\}$  be the sequence generated by (2.12), with  $\alpha_k \in [\alpha_{\min}, \alpha_{\max}]$  and satisfying (2.17). Then  $\{\mathbf{x}^k, \mathbf{y}^k\}$  converges to a saddle point of  $F$  in  $\mathcal{D}$ .*

*Proof.* Let  $(\mathbf{x}^*, \mathbf{y}^*) \in \mathcal{D}$  a saddle point of  $F$ . By applying (2.18) to (2.14), setting  $\mathbf{z} = \mathbf{y}^k + \alpha_k \nabla_{\mathbf{y}}F(\mathbf{x}^{k+1}, \mathbf{y}^k)$  and  $\mathbf{u} = \mathbf{y}^*$  we obtain

$$\begin{aligned} \|\mathbf{y}^{k+1} - \mathbf{y}^*\|^2 &\leq \\ &\leq \|\mathbf{y}^k + \alpha_k \nabla_{\mathbf{y}}F(\mathbf{x}^{k+1}, \mathbf{y}^k) - \mathbf{y}^*\|^2 - \|\mathbf{y}^{k+1} - \mathbf{y}^k - \alpha_k \nabla_{\mathbf{y}}F(\mathbf{x}^{k+1}, \mathbf{y}^k)\|^2 \\ &= \|\mathbf{y}^k - \mathbf{y}^*\|^2 + 2\alpha_k \langle \nabla_{\mathbf{y}}F(\mathbf{x}^{k+1}, \mathbf{y}^k), \mathbf{y}^k - \mathbf{y}^* \rangle + \\ &\quad - \|\mathbf{y}^{k+1} - \mathbf{y}^k\|^2 + 2\alpha_k \langle \nabla_{\mathbf{y}}F(\mathbf{x}^{k+1}, \mathbf{y}^k), \mathbf{y}^{k+1} - \mathbf{y}^k \rangle \\ &= \|\mathbf{y}^k - \mathbf{y}^*\|^2 - \|\mathbf{y}^{k+1} - \mathbf{y}^k\|^2 + 2\alpha_k \langle \nabla_{\mathbf{y}}F(\mathbf{x}^{k+1}, \mathbf{y}^k), \mathbf{y}^{k+1} - \mathbf{y}^* \rangle \end{aligned} \quad (2.19)$$

Similarly, from (2.12) and (2.18) with  $\mathbf{z} = \mathbf{y}^k + \alpha_k \nabla_{\mathbf{y}}F(\mathbf{x}^k, \mathbf{y}^k)$  and  $\mathbf{u} = \mathbf{y}^{k+1}$  we have

$$\begin{aligned} \|\bar{\mathbf{y}}^k - \mathbf{y}^{k+1}\|^2 &\leq \\ &\leq \|\mathbf{y}^k + \alpha_k \nabla_{\mathbf{y}}F(\mathbf{x}^k, \mathbf{y}^k) - \mathbf{y}^{k+1}\|^2 - \|\bar{\mathbf{y}}^k - \mathbf{y}^k - \alpha_k \nabla_{\mathbf{y}}F(\mathbf{x}^k, \mathbf{y}^k)\|^2 \\ &= \|\mathbf{y}^k - \mathbf{y}^{k+1}\|^2 + 2\alpha_k \langle \nabla_{\mathbf{y}}F(\mathbf{x}^k, \mathbf{y}^k), \mathbf{y}^k - \mathbf{y}^{k+1} \rangle + \\ &\quad - \|\bar{\mathbf{y}}^k - \mathbf{y}^k\|^2 + 2\alpha_k \langle \nabla_{\mathbf{y}}F(\mathbf{x}^k, \mathbf{y}^k), \bar{\mathbf{y}}^k - \mathbf{y}^k \rangle \\ &= \|\mathbf{y}^k - \mathbf{y}^{k+1}\|^2 - \|\bar{\mathbf{y}}^k - \mathbf{y}^k\|^2 + 2\alpha_k \langle \nabla_{\mathbf{y}}F(\mathbf{x}^k, \mathbf{y}^k), \bar{\mathbf{y}}^k - \mathbf{y}^{k+1} \rangle \end{aligned}$$

Adding and subtracting  $2\alpha_k \langle \nabla_{\mathbf{y}}F(\mathbf{x}^{k+1}, \mathbf{y}^k), \bar{\mathbf{y}}^k - \mathbf{y}^{k+1} \rangle$  to the right-hand side of the previous inequality, one obtains

$$\begin{aligned} \|\bar{\mathbf{y}}^k - \mathbf{y}^{k+1}\|^2 &\leq \\ &\leq \|\mathbf{y}^k - \mathbf{y}^{k+1}\|^2 - \|\bar{\mathbf{y}}^k - \mathbf{y}^k\|^2 + 2\alpha_k \langle \nabla_{\mathbf{y}}F(\mathbf{x}^{k+1}, \mathbf{y}^k), \bar{\mathbf{y}}^k - \mathbf{y}^{k+1} \rangle \\ &\quad + 2\alpha_k \langle \nabla_{\mathbf{y}}F(\mathbf{x}^k, \mathbf{y}^k) - \nabla_{\mathbf{y}}F(\mathbf{x}^{k+1}, \mathbf{y}^k), \bar{\mathbf{y}}^k - \mathbf{y}^{k+1} \rangle \end{aligned} \quad (2.20)$$

Summing (2.19) and (2.20) yields

$$\begin{aligned}
& \|\mathbf{y}^{k+1} - \mathbf{y}^*\|^2 \leq \\
& \leq \|\mathbf{y}^k - \mathbf{y}^*\|^2 - \|\bar{\mathbf{y}}^k - \mathbf{y}^{k+1}\|^2 - \|\bar{\mathbf{y}}^k - \mathbf{y}^k\|^2 + \\
& \quad + 2\alpha_k \langle \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \mathbf{y}^k), \bar{\mathbf{y}}^k - \mathbf{y}^* \rangle + \\
& \quad + 2\alpha_k \langle \nabla_{\mathbf{y}} F(\mathbf{x}^k, \mathbf{y}^k) - \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \mathbf{y}^k), \bar{\mathbf{y}}^k - \mathbf{y}^{k+1} \rangle
\end{aligned} \tag{2.21}$$

Since  $F$  is concave with respect to the variable  $\mathbf{y}$ , we have

$$\langle \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \mathbf{y}^k), \mathbf{y}^* - \mathbf{y}^k \rangle \geq F(\mathbf{x}^{k+1}, \mathbf{y}^*) - F(\mathbf{x}^{k+1}, \mathbf{y}^k)$$

Then, adding and subtracting  $2\alpha_k \langle \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \mathbf{y}^k), \mathbf{y}^k \rangle$ , we have in (2.21)

$$\begin{aligned}
& \|\mathbf{y}^{k+1} - \mathbf{y}^*\|^2 \leq \\
& \leq \|\mathbf{y}^k - \mathbf{y}^*\|^2 - \|\bar{\mathbf{y}}^k - \mathbf{y}^{k+1}\|^2 - \|\bar{\mathbf{y}}^k - \mathbf{y}^k\|^2 + \\
& \quad + 2\alpha_k \langle \nabla_{\mathbf{y}} F(\mathbf{x}^k, \mathbf{y}^k) - \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \mathbf{y}^k), \bar{\mathbf{y}}^k - \mathbf{y}^{k+1} \rangle + \\
& \quad + 2\alpha_k \langle \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \mathbf{y}^k), \bar{\mathbf{y}}^k - \mathbf{y}^k \rangle + 2\alpha_k \{F(\mathbf{x}^{k+1}, \mathbf{y}^k) - F(\mathbf{x}^{k+1}, \mathbf{y}^*)\}
\end{aligned} \tag{2.22}$$

Let now consider the definition of  $\mathbf{x}^{k+1}$  in (2.13). Invoking again (2.18) with  $\mathbf{z} = \mathbf{x}^k + \alpha_k \nabla_{\mathbf{y}} F(\mathbf{x}^k, \bar{\mathbf{y}}^k)$  and  $\mathbf{u} = \mathbf{x}^*$  we can write

$$\begin{aligned}
& \|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 \leq \\
& \leq \|\mathbf{x}^k - \alpha_k \nabla_{\mathbf{x}} F(\mathbf{x}^k, \bar{\mathbf{y}}^k) - \mathbf{x}^*\|^2 - \|\mathbf{x}^{k+1} - \mathbf{x}^k + \alpha_k \nabla_{\mathbf{x}} F(\mathbf{x}^k, \bar{\mathbf{y}}^k)\|^2 \\
& = \|\mathbf{x}^k - \mathbf{x}^*\|^2 - \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 + \\
& \quad + 2\alpha_k \langle \nabla_{\mathbf{x}} F(\mathbf{x}^k, \bar{\mathbf{y}}^k), \mathbf{x}^* - \mathbf{x}^k \rangle - 2\alpha_k \langle \nabla_{\mathbf{x}} F(\mathbf{x}^k, \bar{\mathbf{y}}^k), \mathbf{x}^{k+1} - \mathbf{x}^k \rangle
\end{aligned}$$

By the convexity of  $F$  with respect to the variable  $\mathbf{x}$  we have

$$\langle \nabla_{\mathbf{x}} F(\mathbf{x}^k, \bar{\mathbf{y}}^k), \mathbf{x}^* - \mathbf{x}^k \rangle \leq F(\mathbf{x}^*, \bar{\mathbf{y}}^k) - F(\mathbf{x}^k, \bar{\mathbf{y}}^k)$$

thus we can write

$$\begin{aligned}
\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 & \leq \|\mathbf{x}^k - \mathbf{x}^*\|^2 - \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 + \\
& \quad + 2\alpha_k \{F(\mathbf{x}^*, \bar{\mathbf{y}}^k) - F(\mathbf{x}^k, \bar{\mathbf{y}}^k) - \langle \nabla_{\mathbf{x}} F(\mathbf{x}^k, \bar{\mathbf{y}}^k), \mathbf{x}^{k+1} - \mathbf{x}^k \rangle\}
\end{aligned}$$

Summing and subtracting  $2\alpha_k F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k)$  to the right hand side in the last inequality and observing that the convexity implies also that

$$F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k) - F(\mathbf{x}^k, \bar{\mathbf{y}}^k) \leq \langle \nabla_{\mathbf{x}} F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k), \mathbf{x}^{k+1} - \mathbf{x}^k \rangle$$

we obtain

$$\begin{aligned}
\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 & \leq \|\mathbf{x}^k - \mathbf{x}^*\|^2 - \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 + \\
& \quad + 2\alpha_k \{F(\mathbf{x}^*, \bar{\mathbf{y}}^k) - F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k) + \\
& \quad + \langle \nabla_{\mathbf{x}} F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k) - \nabla_{\mathbf{x}} F(\mathbf{x}^k, \bar{\mathbf{y}}^k), \mathbf{x}^{k+1} - \mathbf{x}^k \rangle\}
\end{aligned}$$

Now we recall that the saddle point  $(\mathbf{x}^*, \mathbf{y}^*)$  satisfies

$$F(\mathbf{x}^*, \bar{\mathbf{y}}^k) \leq F(\mathbf{x}^*, \mathbf{y}^*) \leq F(\mathbf{x}^{k+1}, \mathbf{y}^*)$$

which leads to the following inequality

$$\begin{aligned} \|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 &\leq \\ &\leq \|\mathbf{x}^k - \mathbf{x}^*\|^2 - \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 + 2\alpha_k \{F(\mathbf{x}^{k+1}, \mathbf{y}^*) - F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k)\} + \\ &\quad + 2\alpha_k \langle \nabla_{\mathbf{x}} F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k) - \nabla_{\mathbf{x}} F(\mathbf{x}^k, \bar{\mathbf{y}}^k), \mathbf{x}^{k+1} - \mathbf{x}^k \rangle \end{aligned} \quad (2.23)$$

Summing the inequalities (2.22) and (2.23) and adding and subtracting  $2\alpha_k \langle \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k), \bar{\mathbf{y}}^k - \mathbf{y}^k \rangle$  yields

$$\begin{aligned} &\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 + \|\mathbf{y}^{k+1} - \mathbf{y}^*\|^2 \leq \\ &\leq \|\mathbf{x}^k - \mathbf{x}^*\|^2 + \|\mathbf{y}^k - \mathbf{y}^*\|^2 - \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 - \|\bar{\mathbf{y}}^k - \mathbf{y}^{k+1}\|^2 + \\ &\quad - \|\bar{\mathbf{y}}^k - \mathbf{y}^k\|^2 + \\ &\quad + 2\alpha_k \langle \nabla_{\mathbf{x}} F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k) - \nabla_{\mathbf{x}} F(\mathbf{x}^k, \bar{\mathbf{y}}^k), \mathbf{x}^{k+1} - \mathbf{x}^k \rangle + \\ &\quad + 2\alpha_k \langle \nabla_{\mathbf{y}} F(\mathbf{x}^k, \mathbf{y}^k) - \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \mathbf{y}^k), \bar{\mathbf{y}}^k - \mathbf{y}^{k+1} \rangle + \\ &\quad + 2\alpha_k \{F(\mathbf{x}^{k+1}, \mathbf{y}^k) - F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k) + \langle \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k), \bar{\mathbf{y}}^k - \mathbf{y}^k \rangle\} + \\ &\quad + 2\alpha_k \langle \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \mathbf{y}^k) - \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k), \bar{\mathbf{y}}^k - \mathbf{y}^k \rangle \\ &\leq \|\mathbf{x}^k - \mathbf{x}^*\|^2 + \|\mathbf{y}^k - \mathbf{y}^*\|^2 - \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 - \|\bar{\mathbf{y}}^k - \mathbf{y}^{k+1}\|^2 + \\ &\quad - \|\bar{\mathbf{y}}^k - \mathbf{y}^k\|^2 + \\ &\quad + 2\alpha_k \langle \nabla_{\mathbf{x}} F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k) - \nabla_{\mathbf{x}} F(\mathbf{x}^k, \bar{\mathbf{y}}^k), \mathbf{x}^{k+1} - \mathbf{x}^k \rangle + \\ &\quad + 2\alpha_k \langle \nabla_{\mathbf{y}} F(\mathbf{x}^k, \mathbf{y}^k) - \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \mathbf{y}^k), \bar{\mathbf{y}}^k - \mathbf{y}^{k+1} \rangle + \\ &\quad + 2\alpha_k \langle \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \mathbf{y}^k) - \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k), \bar{\mathbf{y}}^k - \mathbf{y}^k \rangle \end{aligned}$$

where the last inequality follows from the concavity of  $F$  with respect to  $\mathbf{y}$ . By the Cauchy-Schwartz inequality, we obtain

$$\begin{aligned} &\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 + \|\mathbf{y}^{k+1} - \mathbf{y}^*\|^2 \leq \\ &\leq \|\mathbf{x}^k - \mathbf{x}^*\|^2 + \|\mathbf{y}^k - \mathbf{y}^*\|^2 - \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 - \|\bar{\mathbf{y}}^k - \mathbf{y}^{k+1}\|^2 + \\ &\quad - \|\bar{\mathbf{y}}^k - \mathbf{y}^k\|^2 + \\ &\quad + 2\alpha_k \|\nabla_{\mathbf{x}} F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k) - \nabla_{\mathbf{x}} F(\mathbf{x}^k, \bar{\mathbf{y}}^k)\| \|\mathbf{x}^{k+1} - \mathbf{x}^k\| + \\ &\quad + 2\alpha_k \|\nabla_{\mathbf{y}} F(\mathbf{x}^k, \mathbf{y}^k) - \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \mathbf{y}^k)\| \|\bar{\mathbf{y}}^k - \mathbf{y}^{k+1}\| + \\ &\quad + 2\alpha_k \|\nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \mathbf{y}^k) - \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \bar{\mathbf{y}}^k)\| \|\bar{\mathbf{y}}^k - \mathbf{y}^k\| \end{aligned} \quad (2.24)$$

Now we recall that, since the projection operator is non-expansive, we can write

$$\begin{aligned} \|\mathbf{y}^{k+1} - \bar{\mathbf{y}}^k\| &= \|\text{Py}(\mathbf{y}^k + \alpha_k \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \mathbf{y}^k)) - \text{Py}(\mathbf{y}^k + \alpha_k \nabla_{\mathbf{y}} F(\mathbf{x}^k, \mathbf{y}^k))\| \\ &\leq \alpha_k \|\nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \mathbf{y}^k) - \nabla_{\mathbf{y}} F(\mathbf{x}^k, \mathbf{y}^k)\| \end{aligned} \quad (2.25)$$

Using the inequality (2.25) in (2.24) and recalling the definitions (2.17), we obtain

$$\begin{aligned}
& \|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 + \|\mathbf{y}^{k+1} - \mathbf{y}^*\|^2 \\
& \leq \|\mathbf{x}^k - \mathbf{x}^*\|^2 + \|\mathbf{y}^k - \mathbf{y}^*\|^2 - \|\bar{\mathbf{y}}^k - \mathbf{y}^{k+1}\|^2 + \\
& \quad - (1 - 2\alpha_k C_k) \|\bar{\mathbf{y}}^k - \mathbf{y}^k\|^2 + \\
& \quad - (1 - 2\alpha_k A_k - 2\alpha_k^2 B_k^2) \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2
\end{aligned} \tag{2.26}$$

By the hypothesis (2.17), the coefficients  $(1 - 2\alpha_k A_k - 2\alpha_k^2 B_k^2)$  and  $(1 - 2\alpha_k C_k)$  are strictly positive and bounded away from zero. Thus, we must have that

$$\begin{aligned}
\lim_k \|\bar{\mathbf{y}}^k - \mathbf{y}^{k+1}\| &= 0 \\
\lim_k \|\bar{\mathbf{y}}^k - \mathbf{y}^k\| &= 0 \\
\lim_k \|\mathbf{x}^{k+1} - \mathbf{x}^k\| &= 0
\end{aligned}$$

Then there exists a point  $(\bar{\mathbf{x}}, \bar{\mathbf{y}}) \in \mathcal{D}$  such that  $\mathbf{x}^k$  converges to  $\bar{\mathbf{x}}$  and  $\mathbf{y}^k, \bar{\mathbf{y}}^k$  converges to  $\bar{\mathbf{y}}$ . Consider now a subsequence  $\{\alpha_{k_j}\}_j$  such that  $\lim_j \alpha_{k_j} = \bar{\alpha} > 0$ . Taking the limit for  $j \rightarrow \infty$ , by continuity of the projection operator and by the definition of the sequences  $\mathbf{x}^k, \mathbf{y}^k, \bar{\mathbf{y}}^k$ , we have

$$\begin{aligned}
\bar{\mathbf{y}} &= P_{\mathcal{Y}}(\bar{\mathbf{y}} + \bar{\alpha} \nabla_{\mathbf{y}} F(\bar{\mathbf{x}}, \bar{\mathbf{y}})) \\
\bar{\mathbf{x}} &= P_{\mathcal{X}}(\bar{\mathbf{x}} - \bar{\alpha} \nabla_{\mathbf{x}} F(\bar{\mathbf{x}}, \bar{\mathbf{y}}))
\end{aligned}$$

Thus, we can conclude that  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$  is a saddle point of  $F$ .  $\square$

The well-posedness of the method and the convergence property is assured when a steplength parameter satisfying (2.17) exists; the scheme of AEM method is depicted in Algorithm 4.

Now we focus on image restoration framework, considering the problems related to KL-TV and KL-HS. The primal-dual formulation of

$$\min_{\mathbf{x} \in \mathcal{X}} \varphi_{\beta} \equiv \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta\varphi_1(\mathbf{x})$$

for  $\varphi_1(\mathbf{x})$  given by (1.17) or (1.20) is the saddle point problem

$$\min_{\mathbf{x} \in \mathcal{X}} \max_{\mathbf{y} \in \mathcal{Y}} F(\mathbf{x}, \mathbf{y}) \equiv \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta\mathbf{y}^t M\mathbf{x}$$

where  $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^N : \mathbf{x} \geq \eta, \eta \geq 0\}$ ; for TV regularization,  $M\mathbf{x}$  and  $\mathcal{Y}$  are given by

$$\begin{aligned}
M\mathbf{x} &= \nabla\mathbf{x}, \quad \nabla = (\nabla_1^t, \dots, \nabla_N^t)^t \\
\mathcal{Y} &= \left\{ \mathbf{y} \in \mathbb{R}^{2N} \mid \sqrt{y_{2i+1}^2 + y_{2i}^2} \leq 1, i = 1, \dots, N \right\}
\end{aligned} \tag{2.27}$$

while for Hypersurface regularization we have

$$\begin{aligned}
M\mathbf{x} &= \begin{pmatrix} \nabla\mathbf{x} \\ \delta\mathbf{1} \end{pmatrix} \\
\mathcal{Y} &= \left\{ \mathbf{y} \in \mathbb{R}^{3N} \mid \sqrt{y_{2i+1}^2 + y_{2i}^2 + y_{2N+i}^2} \leq 1, i = 1, \dots, N \right\}
\end{aligned} \tag{2.28}$$

Under the assumption that  $\mathbf{1} \notin \text{Ker}(\mathbf{H})$ ,  $\varphi_{\beta}$  is coercive (see Proposition 1.2). Thus we can restrict the variable  $\mathbf{x}$  in a bounded subset of  $\mathcal{X}$  and invoke the min-max theorem [84] to ensure the existence of a saddle point for  $F(\mathbf{x}, \mathbf{y})$ .

**Algorithm 4: AEM**

Choose the starting point  $(\mathbf{x}^0, \mathbf{y}^0) \in \mathcal{D}$ , set the parameters  $\varepsilon, \theta \in (0, 1)$   $\alpha_{\max} > 0$ .

FOR  $k = 0, 1, 2, \dots$  DO THE FOLLOWING STEPS

STEP 1. Choose  $\alpha < \alpha_{\max}$ ;

STEP 2. Compute tentative points:

$$\begin{aligned}\mathbf{y}^+ &= P_{\mathbf{y}}(\mathbf{y}^k + \alpha \nabla_{\mathbf{y}} F(\mathbf{x}^k, \mathbf{y}^k)) \\ \mathbf{x}^+ &= P_{\mathbf{x}}(\mathbf{x}^k - \alpha \nabla_{\mathbf{x}} F(\mathbf{x}^k, \mathbf{y}^+))\end{aligned}$$

$$A = \frac{\|\nabla_{\mathbf{x}} F(\mathbf{x}^+, \mathbf{y}^+) - \nabla_{\mathbf{x}} F(\mathbf{x}^k, \mathbf{y}^k)\|}{\|\mathbf{x}^+ - \mathbf{x}^k\|}$$

$$B = \frac{\|\nabla_{\mathbf{y}} F(\mathbf{x}^+, \mathbf{y}^k) - \nabla_{\mathbf{y}} F(\mathbf{x}^k, \mathbf{y}^k)\|}{\|\mathbf{x}^+ - \mathbf{x}^k\|}$$

$$C_k = \frac{\|\nabla_{\mathbf{y}} F(\mathbf{x}^+, \mathbf{y}^k) - \nabla_{\mathbf{y}} F(\mathbf{x}^+, \mathbf{y}^+)\|}{\|\mathbf{y}^k - \mathbf{y}^+\|}$$

$$\bar{\alpha} = \begin{cases} \min \left\{ \frac{\sqrt{A^2 + 2B^2(1-\varepsilon)} - A}{2B^2}, \frac{1-\varepsilon}{2C} \right\} & B > 0, C > 0 \\ \min \left\{ \frac{1-\varepsilon}{2A}, \frac{1-\varepsilon}{2C} \right\} & A > 0, C > 0, B = 0 \\ \frac{1-\varepsilon}{2C} & A = 0, C > 0, B = 0 \\ \frac{1-\varepsilon}{2A} & A > 0, C = 0, B = 0 \\ \alpha & \text{otherwise} \end{cases}$$

STEP 3. Check convergence condition:

IF  $\alpha \leq \bar{\alpha}$

$\alpha_k = \alpha$ ;

$\bar{\mathbf{y}}^k = \mathbf{y}^+$ ;

$\mathbf{x}^{k+1} = \mathbf{x}^+$

ELSE

$\alpha = \min\{\bar{\alpha}, \theta\alpha\}$  and go to STEP 2;

STEP 4. Set  $\mathbf{y}^{k+1} = P_{\mathbf{y}}(\mathbf{y}^k + \alpha_k \nabla_{\mathbf{y}} F(\mathbf{x}^{k+1}, \mathbf{y}^k))$

**2.4 PIDAL & PIDSplit+**

Recently in [50, 87] a suitable version of the Alternating Direction Multiplier Method (ADMM) has been proposed for deblurring images corrupted by Poisson noise. The ADMM is designed for solving the minimization problem

$$\min_{\mathbf{x}} f_1(\mathbf{x}) + f_2(M\mathbf{x}) \quad (2.29)$$

being  $M$  a linear operator and  $\mathbf{x} \in \mathbb{R}^N$ . By introducing the constraint  $M\mathbf{x} = \mathbf{w}$ , (2.29) can be formulated as the constrained problem

$$\min_{\mathbf{x}, \mathbf{w}} f_1(\mathbf{x}) + f_2(\mathbf{w}) \text{ subject to } M\mathbf{x} = \mathbf{w} \quad (2.30)$$

or equivalently as

$$\min_{\mathbf{x}, \mathbf{w}} f_1(\mathbf{x}) + f_2(\mathbf{w}) + \frac{1}{2\gamma} \|M\mathbf{x} - \mathbf{w}\|^2 \text{ s.t. } M\mathbf{x} = \mathbf{w} \quad (2.31)$$

The augmented Lagrangian of this last formulation is

$$\max_{\mathbf{p}} \min_{\mathbf{x}, \mathbf{w}} f_1(\mathbf{x}) + f_2(\mathbf{w}) + \mathbf{p}^t(M\mathbf{x} - \mathbf{w}) + \frac{1}{2\gamma} \|M\mathbf{x} - \mathbf{w}\|^2 \quad (2.32)$$

where  $\mathbf{p}$  is the Lagrange multiplier of the equality constraint. Redefining  $\mathbf{p} \rightarrow \gamma\mathbf{p}$ , the basic idea of the iterative procedure is given in Algorithm 5 [15].

---

**Algorithm 5: ADMM**


---

Choose  $\gamma > 0$ , select  $\mathbf{x}^0$ ,  $\mathbf{w}^0 = M\mathbf{x}^0$  and  $\mathbf{p}^0 = 0$

FOR  $k = 0, 1, 2, \dots$  DO THE FOLLOWING STEPS

STEP 1.  $\mathbf{x}^{k+1} \in \arg \min_{\mathbf{x}} f_1(\mathbf{x}) + \frac{1}{2\gamma} \|M\mathbf{x} - \mathbf{w}^k + \mathbf{p}^k\|_2^2$

STEP 2.  $\mathbf{w}^{k+1} \in \arg \min_{\mathbf{w}} f_2(\mathbf{w}) + \frac{1}{2\gamma} \|M\mathbf{x}^{k+1} - \mathbf{w} + \mathbf{p}^k\|_2^2$

STEP 3.  $\mathbf{p}^{k+1} = \mathbf{p}^k + (M\mathbf{x}^{k+1} - \mathbf{w}^{k+1})$

---

In [87, 50], the function  $\varphi_\beta(\mathbf{x}; \mathbf{g}\mathbf{n}) = \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta\mathbf{TV}(\mathbf{x})$  is treated considering

$$f_1(\mathbf{x}) = \langle 0, \mathbf{x} \rangle, \quad f_2(M\mathbf{x}) = \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta\mathbf{TV}(\mathbf{x}) + i_{\mathbf{x} \geq 0}(\mathbf{x})$$

where  $i_{\mathbf{x} \geq 0}$  is the indicator function of the nonnegative orthant and  $M$  is given by

$$M\mathbf{x} = \begin{pmatrix} \mathbf{H} \\ \nabla \\ \mathbf{I} \end{pmatrix} \mathbf{x} = \begin{pmatrix} \mathbf{w}_{(1)} - b\mathbf{g}\mathbf{1} \\ \mathbf{w}_{(2)} \\ \mathbf{w}_{(3)} \end{pmatrix}$$

The Algorithm 5 formulated for the above problem is the so-called PIDSplit+ [87], also called PIDAL-TV in [50]:

It's worth to stress that the computation of  $\mathbf{x}^{k+1}$  and  $\mathbf{w}^{k+1}$  are easier than they appear; in fact, we have

$$\mathbf{x}^{k+1} = (\mathbf{I} + \mathbf{H}^t\mathbf{H} + \nabla^t\nabla)^{-1} \left( \mathbf{H}^t(\mathbf{w}_{(1)}^k - b\mathbf{g}\mathbf{1} - \mathbf{p}_{(1)}^k) + \nabla^t(\mathbf{w}_{(2)}^k - \mathbf{p}_{(2)}^k) + (\mathbf{w}_{(3)}^k - \mathbf{p}_{(3)}^k) \right)$$

thus  $\mathbf{x}^{k+1}$  is available by solving a linear system of equations. Moreover,

$$\begin{aligned} \mathbf{w}_{(1)}^{k+1} &= \frac{1}{2} \left( \mathbf{p}_{(1)}^k + \mathbf{H}\mathbf{x}^{k+1} + (b\mathbf{g} - \gamma)\mathbf{1} \right) \\ &\quad + \frac{1}{2} \left( \sqrt{\left( \mathbf{p}_{(1)}^k + \mathbf{H}\mathbf{x}^{k+1} + (b\mathbf{g} - \gamma)\mathbf{1} \right)^2 + 4\gamma\mathbf{g}\mathbf{n}} \right) \\ \mathbf{w}_{(2)}^{k+1} &= \text{shrink}_{\gamma\beta} \left( \mathbf{p}_{(2)}^k + \nabla\mathbf{x}^{k+1} \right) \\ \mathbf{w}_{(3)}^{k+1} &= \max \left\{ 0, \mathbf{p}_{(3)}^k + \mathbf{x}^{k+1} \right\} \end{aligned}$$

where  $\text{shrink}_\rho$  is the shrinkage operator of parameter  $\rho$ :

$$\text{shrink}_\rho(\mathbf{q}) = \begin{cases} \mathbf{q}_i - \rho \frac{\mathbf{q}_i}{|\mathbf{q}_i|} & |\mathbf{q}_i| \geq \rho, \quad i = 1, \dots, N \\ 0 & \text{otherwise} \end{cases}$$

**Algorithm 6: PIDAL-TV or PIDSplit+**

Choose  $\gamma, \beta > 0$ ,  $\mathbf{x}^0 = \mathbf{gn}$ ;  $\mathbf{w}_{(1)}^0 = \mathbf{H}\mathbf{x}^0 - b\mathbf{g}\mathbf{1}$ ;  $\mathbf{w}_{(2)}^0 = \Delta\mathbf{x}^0$ ;  $\mathbf{w}_{(3)}^0 = \mathbf{x}^0$ ;  $\mathbf{p}^0 = \mathbf{0}$ .  
 FOR  $k = 0, 1, 2, \dots$  DO THE FOLLOWING STEPS

$$\text{STEP 1. } \mathbf{x}^{k+1} = \arg \min_{\mathbf{x}} \frac{1}{2\gamma} \|(\mathbf{H}^t; \nabla^t; \mathbf{I})^t \mathbf{x} - \mathbf{w}^k - b\mathbf{g}\mathbf{1} + \mathbf{p}^k\|^2$$

$$\text{STEP 2. } \mathbf{w}_{(1)}^{k+1} = \arg \min_{\mathbf{w}_{(1)}} \mathbf{KL}(\mathbf{w}_{(1)}; \mathbf{gn}) + \frac{1}{2\gamma} \|\mathbf{H}\mathbf{x}^{k+1} + b\mathbf{g}\mathbf{1} - \mathbf{w}_{(1)} + \mathbf{p}_{(1)}^k\|^2$$

$$\text{STEP 3. } \mathbf{w}_{(2)}^{k+1} = \arg \min_{\mathbf{w}_{(2)}} \beta \sum_i \|(\mathbf{w}_{(2)})_i\| + \frac{1}{2\gamma} \|\nabla\mathbf{x}^{k+1} - \mathbf{w}_{(2)} + \mathbf{p}_{(2)}^k\|^2$$

$$\text{STEP 4. } \mathbf{w}_{(3)}^{k+1} = \arg \min_{\mathbf{w}_{(3)}} i_{\mathbf{x} \geq 0}(\mathbf{x}^{k+1}) + \frac{1}{2\gamma} \|\mathbf{x}^{k+1} - \mathbf{w}_{(3)} + \mathbf{p}_{(3)}^k\|^2$$

$$\text{STEP 5. } \mathbf{p}_{(1)}^{k+1} = \mathbf{p}_{(1)}^k + (\mathbf{H}\mathbf{x}^{k+1} + b\mathbf{g}\mathbf{1} - \mathbf{w}_{(1)}^{k+1})$$

$$\text{STEP 6. } \mathbf{p}_{(2)}^{k+1} = \mathbf{p}_{(2)}^k + (\nabla\mathbf{x}^{k+1} - \mathbf{w}_{(2)}^{k+1})$$

$$\text{STEP 7. } \mathbf{p}_{(3)}^{k+1} = \mathbf{p}_{(3)}^k + (\mathbf{x}^{k+1} - \mathbf{w}_{(3)}^{k+1})$$

The greatest computational cost of this algorithm is hence the computation of  $\mathbf{x}^{k+1}$ : however, under suitable assumptions on the boundary conditions, the linear system can be solved by discrete transforms (e.g., DFT or DCT).

When the regularization term is the Hypersurface Potential (1.20) the ADMM approach can be easily adapted by using the following linear constraint:

$$\begin{pmatrix} \mathbf{H} \\ \nabla \\ \mathbf{0} \\ \mathbf{I} \end{pmatrix} \mathbf{x} = \begin{pmatrix} \mathbf{w}_{(1)} - b\mathbf{g} \\ \bar{\mathbf{w}}_{(2)} \\ \tilde{\mathbf{w}}_{(2)} - \delta \\ \mathbf{w}_{(3)} \end{pmatrix}$$

where  $\mathbf{w}_{(2)} = \begin{pmatrix} \bar{\mathbf{w}}_{(2)} \\ \tilde{\mathbf{w}}_{(2)} \end{pmatrix}$ . In Algorithm 6 the only modification concerns the computation of  $\mathbf{w}_{(2)}^{k+1}$  and  $\mathbf{p}_{(2)}^{k+1}$ :

$$\mathbf{w}_{(2)}^{k+1} = \arg \min_{\mathbf{w}_{(2)}} \beta \sum_i \psi_\delta (\|(\bar{\mathbf{w}}_{(2)})_i\|^2) + \frac{1}{2\gamma} \left\| \begin{pmatrix} \nabla \\ \mathbf{0} \end{pmatrix} \mathbf{x}^{k+1} - \mathbf{w}_{(2)} + \begin{pmatrix} \mathbf{0} \\ \delta\mathbf{1} \end{pmatrix} + \mathbf{p}_{(2)}^k \right\|^2$$

$$\mathbf{p}_{(2)}^{k+1} = \mathbf{p}_{(2)}^k + \left( \begin{pmatrix} \nabla \\ \mathbf{0} \end{pmatrix} \mathbf{x}^{k+1} - \mathbf{w}_{(2)}^{k+1} + \begin{pmatrix} \mathbf{0} \\ \delta\mathbf{1} \end{pmatrix} \right)$$

**2.5 Chambolle & Pock's algorithm**

The methods presented in [37] have the aim to solve the saddle point problem

$$\min_{\mathbf{x}} \max_{\mathbf{y}} \{f_1(\mathbf{x}) + \langle M\mathbf{x}, \mathbf{y} \rangle - f_2^*(\mathbf{y})\} \quad (2.33)$$

being  $f_1, f_2^*$  two proper, convex, l.s.c. functions;  $f_2^*$  denotes the conjugate of  $f_2$  (see Appendix A). It's worth to notice that this problem is the primal-dual formulation of

$$\min_{\mathbf{x}} f_1(\mathbf{x}) + f_2(M\mathbf{x})$$

strictly connected with the formulation presented in Section 2.4. Algorithm 7 shows the scheme of the well-known Chambolle & Pock Algorithm 1 [37].

---

**Algorithm 7: Chambolle & Pock Algorithm 1**

---

Choose  $\tau, \sigma > 0$ ,  $\theta \in [0, 1]$ ,  $(\mathbf{x}^0, \mathbf{y}^0)$  and set  $\bar{\mathbf{x}}^0 = \mathbf{x}^0$ .

FOR  $k = 0, 1, 2, \dots$  DO THE FOLLOWING STEPS

STEP 1.  $\mathbf{y}^{k+1} = (I + \sigma \partial f_2^*)^{-1}(\mathbf{y}^k + \sigma M \bar{\mathbf{x}}^k)$

STEP 2.  $\mathbf{x}^{k+1} = (I + \tau \partial f_1)^{-1}(\mathbf{x}^k - \tau M^t \mathbf{y}^{k+1})$

STEP 3.  $\bar{\mathbf{x}}^{k+1} = \mathbf{x}^{k+1} + \theta(\mathbf{x}^{k+1} - \mathbf{x}^k)$

---

In this scheme  $\partial F$  denotes the subdifferential of the function  $F$  and  $(I + \sigma \partial F)^{-1}$  is the resolvent operator of  $F$  (see Appendix A and [84]).

Let  $\mathcal{X} \times \mathcal{Y}$  be a subset of  $\text{dom}(f_1) \times \text{dom}(f_2^*)$ ; the partial primal–dual gap is given by

$$\begin{aligned} \mathcal{G}_{\mathcal{X} \times \mathcal{Y}}(\mathbf{x}, \mathbf{y}) = & \max_{\mathbf{y}' \in \mathcal{Y}} \langle \mathbf{y}', M \mathbf{x} \rangle - f_2^*(\mathbf{y}') + f_1(\mathbf{x}) \\ & - \min_{\mathbf{x}' \in \mathcal{X}} \langle \mathbf{y}, M \mathbf{x}' \rangle - f_2^*(\mathbf{y}) + f_1(\mathbf{x}') \end{aligned}$$

For sake of completeness, we recall the convergence theorem and its proof in a discrete framework for the choice  $\theta = 1$  [37].

**Proposition 2.3.** [37, Theorem 1] *Let  $L = \|M\|$  and assume there exists a saddle point  $(\mathbf{x}^*, \mathbf{y}^*)$  for the problem (2.33). Choose  $\theta = 1$ ,  $\tau\sigma L^2 < 1$  and let  $\mathbf{x}^k, \bar{\mathbf{x}}^k, \mathbf{y}^k$  defined as in the Algorithm 7. Then*

(a) *for any  $k$ ,  $\mathbf{y}^k, \mathbf{x}^k$  remains bounded:*

$$\frac{\|\mathbf{y}^k - \mathbf{y}^*\|^2}{2\sigma} + \frac{\|\mathbf{x}^k - \mathbf{x}^*\|^2}{2\tau} \leq C \left( \frac{\|\mathbf{y}^0 - \mathbf{y}^*\|^2}{2\sigma} + \frac{\|\mathbf{x}^0 - \mathbf{x}^*\|^2}{2\tau} \right)$$

where  $C \leq (1 - \tau\sigma L)^{-1}$ .

(b) *If we let  $\mathbf{x}_K = \frac{1}{K} \sum_{k=1}^K \mathbf{x}^k$  and  $\mathbf{y}_K = \frac{1}{K} \sum_{k=1}^K \mathbf{y}^k$  for any bounded  $\mathcal{X} \times \mathcal{Y} \subset \text{dom}(f_1) \times \text{dom}(f_2^*)$  the the partial primal-dual gap has the following bound*

$$\mathcal{G}_{\mathcal{X} \times \mathcal{Y}}(\mathbf{x}_K, \mathbf{y}_K) \leq \frac{1}{K} \sup_{(\mathbf{x}, \mathbf{y}) \in \mathcal{X} \times \mathcal{Y}} \frac{\|\mathbf{y}^0 - \mathbf{y}\|^2}{2\sigma} + \frac{\|\mathbf{x}^0 - \mathbf{x}\|^2}{2\tau}$$

*Moreover, the weak cluster points of  $(\mathbf{x}^k, \mathbf{y}^k)$  are saddle points for (2.33).*

(c) *There exists a saddle point  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$  such that  $\mathbf{x}^k \rightarrow \bar{\mathbf{x}}$  and  $\mathbf{y}^k \rightarrow \bar{\mathbf{y}}$ .*

*Proof.* Consider the general form

$$\begin{aligned} \mathbf{y}^{k+1} &= (I + \sigma \partial f_2^*)^{-1}(\mathbf{y}^k + \sigma M \bar{\mathbf{x}}) \\ \mathbf{x}^{k+1} &= (I + \tau \partial f_1)^{-1}(\mathbf{x}^k - \sigma M^t \bar{\mathbf{y}}) \end{aligned}$$

We have

$$\begin{aligned}\frac{\mathbf{y}^k - \mathbf{y}^{k+1}}{\sigma} + M\bar{\mathbf{x}} &\in \partial f_2^*(\mathbf{y}^{k+1}) \\ \frac{\mathbf{x}^k - \mathbf{x}^{k+1}}{\sigma} - M^t \bar{\mathbf{y}} &\in \partial f_1(\mathbf{x}^{k+1})\end{aligned}$$

From the convexity of  $f_2^*$  and  $f_1$

$$\begin{aligned}f_2^*(\mathbf{y}) &\geq f_2^*(\mathbf{y}^{k+1}) + \left\langle \frac{\mathbf{y}^k - \mathbf{y}^{k+1}}{\sigma}, \mathbf{y} - \mathbf{y}^{k+1} \right\rangle + \langle M\bar{\mathbf{x}}, \mathbf{y} - \mathbf{y}^{k+1} \rangle \\ f_1(\mathbf{x}) &\geq f_1(\mathbf{x}^{k+1}) + \left\langle \frac{\mathbf{x}^k - \mathbf{x}^{k+1}}{\sigma}, \mathbf{x} - \mathbf{x}^{k+1} \right\rangle - \langle \bar{\mathbf{y}}, M(\mathbf{x} - \mathbf{x}^{k+1}) \rangle\end{aligned}$$

From

$$\begin{aligned}\left\langle \frac{\mathbf{y}^k - \mathbf{y}^{k+1}}{\sigma}, \mathbf{y} - \mathbf{y}^{k+1} \right\rangle &= -\frac{\|\mathbf{y} - \mathbf{y}^k\|^2}{2\sigma} + \frac{\|\mathbf{y} - \mathbf{y}^{k+1}\|^2}{2\sigma} + \frac{\|\mathbf{y}^k - \mathbf{y}^{k+1}\|^2}{2\sigma} \\ \left\langle \frac{\mathbf{x}^k - \mathbf{x}^{k+1}}{\tau}, \mathbf{x} - \mathbf{x}^{k+1} \right\rangle &= -\frac{\|\mathbf{x} - \mathbf{x}^k\|^2}{2\tau} + \frac{\|\mathbf{x} - \mathbf{x}^{k+1}\|^2}{2\tau} + \frac{\|\mathbf{x}^k - \mathbf{x}^{k+1}\|^2}{2\tau}\end{aligned}$$

and summing both above inequalities, we have

$$\begin{aligned}&\frac{\|\mathbf{y} - \mathbf{y}^k\|^2}{2\sigma} + \frac{\|\mathbf{x} - \mathbf{x}^k\|^2}{2\tau} \geq \\ &\geq (\langle M\mathbf{x}^{k+1}, \mathbf{y} \rangle + f_1(\mathbf{x}^{k+1}) - f_2^*(\mathbf{y})) - (\langle Mx, \mathbf{y}^{k+1} \rangle - f_2^*(\mathbf{y}^{k+1}) + f_1(\mathbf{x})) + \\ &+ \frac{\|\mathbf{y} - \mathbf{y}^{k+1}\|^2}{2\sigma} + \frac{\|\mathbf{y}^k - \mathbf{y}^{k+1}\|^2}{2\sigma} + \frac{\|\mathbf{x} - \mathbf{x}^{k+1}\|^2}{2\tau} + \frac{\|\mathbf{x}^k - \mathbf{x}^{k+1}\|^2}{2\tau} + \\ &+ \langle M\bar{\mathbf{x}}, \mathbf{y} - \mathbf{y}^{k+1} \rangle - \langle \bar{\mathbf{y}}, M(\mathbf{x} - \mathbf{x}^{k+1}) \rangle - \langle M\mathbf{x}^{k+1}, y \rangle + \langle M\mathbf{x}, \mathbf{y}^{k+1} \rangle\end{aligned}$$

Adding and subtracting  $\langle M\mathbf{x}^{k+1}, \mathbf{y}^{k+1} \rangle$ , we obtain

$$\begin{aligned}&\frac{\|\mathbf{y} - \mathbf{y}^k\|^2}{2\sigma} + \frac{\|\mathbf{x} - \mathbf{x}^k\|^2}{2\tau} \geq \\ &\geq (\langle M\mathbf{x}^{k+1}, \mathbf{y} \rangle + f_1(\mathbf{x}^{k+1}) - f_2^*(\mathbf{y})) - (\langle M\mathbf{x}, \mathbf{y}^{k+1} \rangle - f_2^*(\mathbf{y}^{k+1}) + f_1(\mathbf{x})) + \\ &+ \frac{\|\mathbf{y} - \mathbf{y}^{k+1}\|^2}{2\sigma} + \frac{\|\mathbf{y}^k - \mathbf{y}^{k+1}\|^2}{2\sigma} + \frac{\|\mathbf{x} - \mathbf{x}^{k+1}\|^2}{2\tau} + \frac{\|\mathbf{x}^k - \mathbf{x}^{k+1}\|^2}{2\tau} + \\ &+ \langle M(\bar{\mathbf{x}} - \mathbf{x}^{k+1}), \mathbf{y} - \mathbf{y}^{k+1} \rangle + \langle (\mathbf{y}^{k+1} - \bar{\mathbf{y}}), M(\mathbf{x} - \mathbf{x}^{k+1}) \rangle\end{aligned}$$

This inequality plays an important role in proving convergence.

$$\begin{aligned}&\frac{\|\mathbf{y} - \mathbf{y}^k\|^2}{2\sigma} + \frac{\|\mathbf{x} - \mathbf{x}^k\|^2}{2\tau} \geq \\ &\geq (\langle M\mathbf{x}^{k+1}, \mathbf{y} \rangle + f_1(\mathbf{x}^{k+1}) - f_2^*(y)) - (\langle Mx, \mathbf{y}^{k+1} \rangle - f_2^*(\mathbf{y}^{k+1}) + f_1(\mathbf{x})) + \\ &+ \frac{\|\mathbf{y} - \mathbf{y}^{k+1}\|^2}{2\sigma} + \frac{\|\mathbf{y}^k - \mathbf{y}^{k+1}\|^2}{2\sigma} + \frac{\|\mathbf{x} - \mathbf{x}^{k+1}\|^2}{2\tau} + \frac{\|\mathbf{x}^k - \mathbf{x}^{k+1}\|^2}{2\tau} + \\ &+ \langle M(\bar{\mathbf{x}} - \mathbf{x}^{k+1}), \mathbf{y} - \mathbf{y}^{k+1} \rangle + \langle (\mathbf{y}^{k+1} - \bar{\mathbf{y}}), M(\mathbf{x} - \mathbf{x}^{k+1}) \rangle\end{aligned}\tag{2.34}$$

The choice  $\theta = 1$  leads to the extragradient step

$$\bar{\mathbf{x}} = 2\mathbf{x}^k - \mathbf{x}^{k-1}, \quad \bar{\mathbf{y}} = \mathbf{y}^{k+1}$$

Then we have

$$\begin{aligned} & \langle M(\bar{\mathbf{x}} - \mathbf{x}^{k+1}), \mathbf{y} - \mathbf{y}^{k+1} \rangle + \langle (\mathbf{y}^{k+1} - \bar{\mathbf{y}}), M(\mathbf{x} - \mathbf{x}^{k+1}) \rangle \\ &= \langle M(2\mathbf{x}^k - \mathbf{x}^{k-1} - \mathbf{x}^{k+1}), \mathbf{y} - \mathbf{y}^{k+1} \rangle \\ &= \langle M(\mathbf{x}^k - \mathbf{x}^{k+1}), \mathbf{y} - \mathbf{y}^{k+1} \rangle + \langle M(\mathbf{x}^k - \mathbf{x}^{k-1}), \mathbf{y} - \mathbf{y}^k + \mathbf{y}^k - \mathbf{y}^{k+1} \rangle \\ &= \langle M(\mathbf{x}^k - \mathbf{x}^{k+1}), \mathbf{y} - \mathbf{y}^{k+1} \rangle + \langle M(\mathbf{x}^k - \mathbf{x}^{k-1}), \mathbf{y} - \mathbf{y}^k \rangle + \langle M(\mathbf{x}^k - \mathbf{x}^{k-1}), \mathbf{y}^k - \mathbf{y}^{k+1} \rangle \end{aligned}$$

Using  $2ab \leq \alpha a^2 + \frac{b^2}{\alpha}$ , for any  $\alpha > 0$ , we have:

$$L\|\mathbf{x}^k - \mathbf{x}^{k-1}\| \|\mathbf{y}^k - \mathbf{y}^{k+1}\| \leq \frac{L\alpha\tau}{2\tau} \|\mathbf{x}^k - \mathbf{x}^{k-1}\|^2 + \frac{L\sigma}{2\alpha\sigma} \|\mathbf{y}^k - \mathbf{y}^{k+1}\|^2$$

where we can choose  $\alpha = \sqrt{\sigma/\tau}$ , so that  $L\alpha\tau = L\sigma/\alpha = \sqrt{\sigma\tau}L < 1$ . We obtain

$$\begin{aligned} & \frac{\|\mathbf{y} - \mathbf{y}^k\|^2}{2\sigma} + \frac{\|\mathbf{x} - \mathbf{x}^k\|^2}{2\tau} \geq \\ & \geq (\langle M\mathbf{x}^{k+1}, \mathbf{y} \rangle + f_1(\mathbf{x}^{k+1}) - f_2^*(\mathbf{y})) - (\langle M\mathbf{x}, \mathbf{y}^{k+1} \rangle - f_2^*(\mathbf{y}^{k+1}) + f_1(\mathbf{x})) + \\ & \quad + \frac{\|\mathbf{y} - \mathbf{y}^{k+1}\|^2}{2\sigma} + \frac{\|\mathbf{x} - \mathbf{x}^{k+1}\|^2}{2\tau} + \\ & \quad + (1 - \sqrt{\sigma\tau}L) \frac{\|\mathbf{y}^k - \mathbf{y}^{k+1}\|^2}{2\sigma} + \frac{\|\mathbf{x}^k - \mathbf{x}^{k+1}\|^2}{2\tau} - \sqrt{\sigma\tau}L \frac{\|\mathbf{x}^k - \mathbf{x}^{k-1}\|^2}{2\tau} + \\ & \quad + \langle M(\mathbf{x}^k - \mathbf{x}^{k+1}), \mathbf{y} - \mathbf{y}^{k+1} \rangle + \langle M(\mathbf{x}^k - \mathbf{x}^{k-1}), \mathbf{y} - \mathbf{y}^k \rangle \end{aligned} \quad (2.35)$$

Summing up from 0 to  $K-1$ , we have

$$\begin{aligned} & \langle M(\mathbf{x}^K - \mathbf{x}^{K-1}), \mathbf{y}^K - \mathbf{y} \rangle + \\ & \quad + \sum_0^{K-1} (\langle M\mathbf{x}^{k+1}, \mathbf{y} \rangle + f_1(\mathbf{x}^{k+1}) - f_2^*(\mathbf{y})) - (\langle M\mathbf{x}, \mathbf{y}^{k+1} \rangle - f_2^*(\mathbf{y}^{k+1}) + f_1(\mathbf{x})) + \\ & \quad + \frac{\|\mathbf{y} - \mathbf{y}^K\|^2}{2\sigma} + \frac{\|\mathbf{x} - \mathbf{x}^K\|^2}{2\tau} + \\ & \quad + (1 - \sqrt{\sigma\tau}L) \sum_1^K \frac{\|\mathbf{y}^{k-1} - \mathbf{y}^k\|^2}{2\sigma} + (1 - \sqrt{\sigma\tau}L) \sum_1^{K-1} \frac{\|\mathbf{x}^k - \mathbf{x}^{k-1}\|^2}{2\tau} + \\ & \quad + \frac{\|\mathbf{x}^K - \mathbf{x}^{K-1}\|^2}{2\tau} \leq \\ & \leq \frac{\|\mathbf{y} - \mathbf{y}^0\|^2}{2\sigma} + \frac{\|\mathbf{x} - \mathbf{x}^0\|^2}{2\tau} \end{aligned}$$

Since

$$\langle M(\mathbf{x}^K - \mathbf{x}^{K-1}), \mathbf{y}^K - \mathbf{y} \rangle \geq -\frac{\|\mathbf{x}^K - \mathbf{x}^{K-1}\|^2}{2\tau} - \tau\sigma L^2 \frac{\|\mathbf{y} - \mathbf{y}^K\|^2}{2\sigma}$$

we finally obtain

$$\begin{aligned}
& \sum_0^{K-1} (\langle M\mathbf{x}^{k+1}, \mathbf{y} \rangle + f_1(\mathbf{x}^{k+1}) - f_2^*(\mathbf{y})) - (\langle M\mathbf{x}, \mathbf{y}^{k+1} \rangle - f_2^*(\mathbf{y}^{k+1}) + f_1(\mathbf{x})) + \\
& + (1 - \sigma\tau L^2) \frac{\|\mathbf{y} - \mathbf{y}^K\|^2}{2\sigma} + \frac{\|\mathbf{x} - \mathbf{x}^K\|^2}{2\tau} + \\
& + (1 - \sqrt{\sigma\tau}L) \sum_1^K \frac{\|\mathbf{y}^{k-1} - \mathbf{y}^k\|^2}{2\sigma} + (1 - \sqrt{\sigma\tau}L) \sum_1^{K-1} \frac{\|\mathbf{x}^k - \mathbf{x}^{k-1}\|^2}{2\tau} \leq \\
& \leq \frac{\|\mathbf{y} - \mathbf{y}^0\|^2}{2\sigma} + \frac{\|\mathbf{x} - \mathbf{x}^0\|^2}{2\tau}
\end{aligned}$$

If we choose as  $(\mathbf{x}, \mathbf{y})$  a saddle point  $(\mathbf{x}^*, \mathbf{y}^*)$ , it follows that the first summation is nonnegative and part (a) follows (the sequences are bounded):

$$\begin{aligned}
& (1 - \sigma\tau L^2) \left( \frac{\|\mathbf{y}^* - \mathbf{y}^K\|^2}{2\sigma} + \frac{\|\mathbf{x}^* - \mathbf{x}^K\|^2}{2\tau} \right) \leq \\
& \leq \frac{\|\mathbf{y}^* - \mathbf{y}^0\|^2}{2\sigma} + \frac{\|\mathbf{x}^* - \mathbf{x}^0\|^2}{2\tau}
\end{aligned}$$

In  $\mathcal{X} \times \mathcal{Y}$ , the partial primal–dual gap behaves as  $\mathcal{O}(1/K)$ . Let

$$\mathbf{x}_K = \frac{1}{K} \sum_{i=1}^K \mathbf{x}^i, \quad \mathbf{y}_K = \frac{1}{K} \sum_{i=1}^K \mathbf{y}^i$$

From the convexity of  $f_1$  and  $f_2^*$  it follows

$$\begin{aligned}
& (\langle M\mathbf{x}_K, \mathbf{y} \rangle + f_1(\mathbf{x}_K) - f_2^*(\mathbf{y})) - (\langle M\mathbf{x}, \mathbf{y}_K \rangle - f_2^*(\mathbf{y}_K) + f_1(\mathbf{x})) \\
& \leq \frac{1}{K} \left( \frac{\|\mathbf{y} - \mathbf{y}^0\|^2}{2\sigma} + \frac{\|\mathbf{x} - \mathbf{x}^0\|^2}{2\tau} \right)
\end{aligned}$$

Since  $(\mathbf{x}_K, \mathbf{y}_K)$  is a bounded sequence, there exists a cluster point  $(\mathbf{x}^*, \mathbf{y}^*)$ ; since  $f_1$  and  $f_2^*$  are convex and l.s.c., it follows that for  $K \rightarrow \infty$  we have

$$(\langle M\mathbf{x}^*, \mathbf{y} \rangle + f_1(\mathbf{x}^*) - f_2^*(\mathbf{y})) - (\langle M\mathbf{x}, \mathbf{y}^* \rangle - f_2^*(\mathbf{y}^*) + f_1(\mathbf{x})) \leq 0$$

Then  $(\mathbf{x}^*, \mathbf{y}^*)$  is a saddle-point. Part (b) is hence proved.

For part (c), we observe that  $(\mathbf{x}^k, \mathbf{y}^k)$  is bounded (first part), so that a subsequence  $(\mathbf{x}^{k_n}, \mathbf{y}^{k_n})$  converges to some limit  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$  (strongly, since  $\mathcal{X}$  and  $\mathcal{Y}$  are finite-dimensional space). From the above inequalities,  $\mathbf{x}^k - \mathbf{x}^{k-1} \rightarrow 0$  and  $\mathbf{y}^k - \mathbf{y}^{k-1} \rightarrow 0$ ; in particular,  $(\mathbf{x}^{k_n-1}, \mathbf{y}^{k_n-1})$  converges to  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ ; then  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$  is a fixed point of the iteration, that is a saddle-point of the problem. If we consider (2.35) with  $(\mathbf{x}, \mathbf{y}) = (\bar{\mathbf{x}}, \bar{\mathbf{y}})$  and we sum from  $k_n$  to  $K - 1$ , we have

$$\begin{aligned}
& \frac{\|\bar{\mathbf{y}} - \mathbf{y}^K\|^2}{2\sigma} + \frac{\|\bar{\mathbf{x}} - \mathbf{x}^K\|^2}{2\tau} + \\
& + (1 - \sqrt{\sigma\tau}L) \sum_{k_k+1}^K \frac{\|\mathbf{y}^{k-1} - \mathbf{y}^k\|^2}{2\sigma} + (1 - \sqrt{\sigma\tau}L) \sum_{k_n}^{K-1} \frac{\|\mathbf{x}^k - \mathbf{x}^{k-1}\|^2}{2\tau} + \\
& + \frac{\|\mathbf{x}^K - \mathbf{x}^{K-1}\|^2}{2\tau} - \frac{\|\mathbf{x}^{k_n} - \mathbf{x}^{k_n-1}\|^2}{2\tau} \leq \\
& \leq \frac{\|\bar{\mathbf{y}} - \mathbf{y}^{k_n}\|^2}{2\sigma} + \frac{\|\bar{\mathbf{x}} - \mathbf{x}^{k_n}\|^2}{2\tau}
\end{aligned}$$

which leads, for  $K \rightarrow \infty$ , to

$$\frac{\|\bar{\mathbf{y}} - \mathbf{y}^K\|^2}{2\sigma} + \frac{\|\bar{\mathbf{x}} - \mathbf{x}^K\|^2}{2\tau} \leq \frac{\|\bar{\mathbf{y}} - \mathbf{y}^{k_n}\|^2}{2\sigma} + \frac{\|\bar{\mathbf{x}} - \mathbf{x}^{k_n}\|^2}{2\tau}$$

i.e.  $(\mathbf{x}^K, \mathbf{y}^K) \rightarrow (\bar{\mathbf{x}}, \bar{\mathbf{y}})$  for  $K \rightarrow \infty$ .  $\square$

This method is an extrapolated version of the Arrow–Hurwicz algorithm [2]. Furthermore, under more strictly assumption on  $f_1$  and  $f_2^*$ , it is possible to develop accelerated schemes.

Algorithm 7 can be employed in image restoration framework for Poisson data. For denoising problems, where  $\mathbf{H} = \mathbf{I}$ , we can set  $f_1(\mathbf{x}) = \mathbf{KL}(\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + i_{\mathbf{x} \geq 0}$ , and  $f_2(M\mathbf{x}) = \beta\mathbf{TV}(\mathbf{x})$  or  $f_2(M\mathbf{x}) = \beta\mathbf{HS}(\mathbf{x})$  and  $M$  is given by (2.27) or (2.28), respectively.

Following [92, 103], for deblurring problems when a  $\mathbf{TV}$  regularization is required, we can set  $f_1(\mathbf{x}) = i_{\mathbf{x} \geq 0}$ ,  $f_2(\mathbf{w}) = \mathbf{KL}(\mathbf{w}_{(1)}; \mathbf{g}\mathbf{n}) + \beta \sum_i \|(\mathbf{w}_{(2)})_i\|$ , where  $\mathbf{w} = \begin{pmatrix} \mathbf{w}_{(1)} \\ \mathbf{w}_{(2)} \end{pmatrix}$  and the following constraint holds:

$$M\mathbf{x} = \begin{pmatrix} \mathbf{H} \\ \nabla \end{pmatrix} \mathbf{x} = \mathbf{w} - \begin{pmatrix} b\mathbf{g}\mathbf{1} \\ \mathbf{0} \end{pmatrix} \quad (2.36)$$

We start the Chambolle & Pock method from the  $\mathbf{x}$  iterate; then the extrapolation STEP 3 of Algorithm 7 is applied on the dual variable; the computation of  $\mathbf{y}^{k+1}$  requires to solve the problem

$$\min_{\mathbf{y}} f_2^*(\mathbf{y}) + \frac{1}{2\sigma} \left\| \mathbf{y} - \left[ \mathbf{y}^k + \sigma \begin{pmatrix} \mathbf{H} \\ \nabla \end{pmatrix} \mathbf{x}^{k+1} \right] \right\|^2$$

which can be formulated as [103]

$$\min_{\mathbf{y}} \max_{\mathbf{w}} \langle \mathbf{y}, \mathbf{w} \rangle - f_2(\mathbf{w}) + \frac{1}{2\sigma} \left\| \mathbf{y} - \left[ \mathbf{y}^k + \sigma \begin{pmatrix} \mathbf{H} \\ \nabla \end{pmatrix} \mathbf{x}^{k+1} \right] \right\|^2$$

Then we have

$$\mathbf{y}^{k+1} = \mathbf{y}^k + \sigma(M\mathbf{x}^{k+1} - \bar{\mathbf{w}})$$

where

$$\bar{\mathbf{w}} = \arg \min_{\mathbf{w}} f_2(\mathbf{w}) + \frac{\sigma}{2} \left\| \mathbf{w} - \frac{1}{\sigma} \left( \mathbf{y}^k - \sigma M\mathbf{x}^{k+1} \right) \right\|^2$$

Finally, starting from the  $\mathbf{x}$  iterate and making the extrapolation step on the dual variable, each step of the algorithm takes the following form:

$$\begin{aligned} \mathbf{x}^{k+1} &= \max\{0, \mathbf{x}^k - \tau(\mathbf{H}^t(2\mathbf{y}_{(1)}^k - \mathbf{y}_{(1)}^{k-1}) + \nabla^t(2\mathbf{y}_{(2)}^k - \mathbf{y}_{(2)}^{k-1}))\} \\ \bar{\mathbf{w}}_{(1)}^{k+1} &= \arg \min_{\mathbf{w}_{(1)}} \mathbf{KL}(\mathbf{w}_{(1)}; \mathbf{g}\mathbf{n}) + \frac{\sigma}{2} \left\| \mathbf{w}_{(1)} - \frac{1}{\sigma} \left( \mathbf{y}_{(1)}^k + \sigma \left( \mathbf{H}\mathbf{x}^{k+1} + b\mathbf{g}\mathbf{1} \right) \right) \right\|^2 \\ \bar{\mathbf{w}}_{(2)}^{k+1} &= \arg \min_{\mathbf{w}_{(2)}} \beta \sum_i \|(\bar{\mathbf{w}}_{(2)})_i\| + \frac{\sigma}{2} \left\| \bar{\mathbf{w}}_{(2)} - \frac{1}{\sigma} \left( \mathbf{y}_{(2)}^k + \sigma \nabla \mathbf{x}^{k+1} \right) \right\|^2 \\ \mathbf{y}_{(1)}^{k+1} &= \mathbf{y}_{(1)}^k + \sigma(\mathbf{H}\bar{\mathbf{x}}^{k+1} + b\mathbf{g}\mathbf{1} - \bar{\mathbf{w}}_{(1)}^{k+1}) \\ \mathbf{y}_{(2)}^{k+1} &= \mathbf{y}_{(2)}^k + \sigma(\nabla \bar{\mathbf{x}}^{k+1} - \bar{\mathbf{w}}_{(2)}^{k+1}) \end{aligned} \quad (2.37)$$

When the regularization term is the **HS** functional, this approach can be easily extended by considering the following constraint instead of (2.36):

$$\begin{pmatrix} \mathbf{H} \\ \nabla \\ \mathbf{0} \end{pmatrix} \mathbf{x} = \begin{pmatrix} \mathbf{w}_{(1)} - b\mathbf{g}\mathbf{1} \\ \bar{\mathbf{w}}_{(2)} \\ \tilde{\mathbf{w}}_{(2)} - \delta\mathbf{1} \end{pmatrix}$$

which leads to

$$\begin{aligned} \mathbf{x}^{k+1} &= \max\{0, \mathbf{x}^k - \tau(\mathbf{H}^t(2\mathbf{y}_{(1)}^k - \mathbf{y}_{(1)}^{k-1}) + \nabla^t(2\mathbf{y}_{(2)}^k - \mathbf{y}_{(2)}^{k-1}))\} \\ \bar{\mathbf{w}}_{(1)}^{k+1} &= \arg \min_{\mathbf{w}_{(1)}} \mathbf{KL}(\mathbf{w}_{(1)}; \mathbf{g}\mathbf{n}) + \frac{\sigma}{2} \left\| \mathbf{w}_{(1)} - \frac{1}{\sigma} \left( \mathbf{y}_{(1)}^k + \sigma \left( \mathbf{H}\mathbf{x}^{k+1} + b\mathbf{g}\mathbf{1} \right) \right) \right\|^2 \\ \mathbf{w}_{(2)}^{k+1} &= \arg \min_{\mathbf{w}_{(2)}} \beta \sum \left\| (\mathbf{w}_{(2)})_i \right\| + \frac{\sigma}{2} \left\| \mathbf{w}_{(2)} - \frac{1}{\sigma} \left( \mathbf{y}_{(2)}^k + \sigma \left[ \begin{pmatrix} \nabla \\ \mathbf{0} \end{pmatrix} \mathbf{x}^{k+1} + \begin{pmatrix} \mathbf{0} \\ \delta\mathbf{1} \end{pmatrix} \right] \right) \right\|^2 \\ \mathbf{y}_{(1)}^{k+1} &= \mathbf{y}_{(1)}^k + \sigma(\mathbf{H}\bar{\mathbf{x}}^{k+1} + b\mathbf{g}\mathbf{1} - \bar{\mathbf{w}}_{(1)}^{k+1}) \\ \mathbf{y}_{(2)}^{k+1} &= \mathbf{y}_{(2)}^k + \sigma(\nabla\bar{\mathbf{x}}^{k+1} - \bar{\mathbf{w}}_{(2)}^{k+1}) \\ \tilde{\mathbf{y}}_{(2)}^{k+1} &= \tilde{\mathbf{y}}_{(2)}^k + \sigma(\delta\mathbf{1} - \tilde{\mathbf{w}}_{(2)}^{k+1}) \end{aligned} \tag{2.38}$$

where  $\mathbf{w}_2 = \begin{pmatrix} \bar{\mathbf{w}}_{(2)} \\ \tilde{\mathbf{w}}_{(2)} \end{pmatrix}$  and  $\mathbf{y}_2 = \begin{pmatrix} \bar{\mathbf{y}}_{(2)} \\ \tilde{\mathbf{y}}_{(2)} \end{pmatrix}$ .

## 2.6 PDHG with variable metric

The methods we are going to present have the aim to solve the problem (2.1) where  $\varphi : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{-\infty, \infty\}$  is a convex, proper, l.s.c. function and  $\Omega$  is a convex, closed non-empty subset of  $\mathbb{R}^N$ . We are interested in the case when  $\varphi$  is not differentiable and the projection on the constraint set is easy to compute. We assume that the solution set  $\Omega^*$  is not empty.

First of all, we recall the definition of  $\varepsilon$ -subdifferential and of  $\varepsilon$ -subgradient.

**Definition 2.2.** Let  $\varphi$  a proper convex function on  $\mathbb{R}^n$ . The  $\varepsilon$ -subdifferential of  $\varphi$  at  $x \in \text{dom}(\varphi)$ , defined for  $\varepsilon \geq 0$ , is the set

$$\partial_\varepsilon \varphi(x) = \{w \in \mathbb{R}^n \mid \varphi(z) \geq \varphi(x) + \langle w, z - x \rangle - \varepsilon, \forall z \in \mathbb{R}^n\}$$

An element of  $\partial_\varepsilon \varphi(x)$  is called an  $\varepsilon$ -subgradient of  $\varphi$  at  $x$ .

When  $\varepsilon = 0$ , we have the definition of subdifferential and subgradient.

The methods we investigate belongs to the class of  $\varepsilon$ -subgradient methods, which have the general form

$$\mathbf{x}^{k+1} = P_\Omega \left( \mathbf{x}^k - \alpha_k \mathbf{u}^k \right) \tag{2.39}$$

where  $\mathbf{u}^k \in \partial_{\varepsilon_k} \varphi(\mathbf{x}^k)$  for  $\varepsilon_k \geq 0$ ,  $\alpha_k$  is the stepsize and  $P_\Omega$  is the projection on  $\Omega$ . When  $\varepsilon_k = 0$  for any  $k$ , we obtain the subgradient methods. In [80, 48] the general case in which  $\varepsilon_k > 0$  is introduced and developed; recently, under different assumptions,

convergence results are shown in [1, 65, 68, 90]. A typical assumption on the sequence  $\{\varepsilon_k\}$  is

$$\lim_{k \rightarrow \infty} \varepsilon_k = 0 \quad (2.40)$$

and, in this case, the subgradient and the  $\varepsilon$ -subgradient have very similar convergence properties.

In literature different choices for  $\alpha_k$  are discussed:

( $\mathcal{R}_1$ ) the *constant stepsize* rule  $\alpha_k = \alpha > 0$ ;

( $\mathcal{R}_2$ ) the *Polyak rule*

$$\alpha_k = c_k \frac{\varphi(\mathbf{x}^k) - \varphi^*}{\|\mathbf{u}^k\|^2}, \text{ or } \alpha_k = c_k \frac{\varphi(\mathbf{x}^k) - \varphi^*}{\max\{1, \|\mathbf{u}^k\|^2\}}, \quad c_k \in (0, 2)$$

where  $\varphi^*$  is the minimum of  $\varphi$

( $\mathcal{R}_3$ ) the *Ermoliev or diminishing, divergent series* stepsize rule, which includes any sequences  $\{\alpha_k\}$  such that

$$\alpha_k > 0, \quad \lim_{k \rightarrow \infty} \alpha_k = 0, \quad \sum_{k=0}^{\infty} \alpha_k = \infty \quad (2.41)$$

( $\mathcal{R}_4$ ) the *diminishing, divergent series, square summable* stepsize rule, which, in addition to (2.41), also requires  $\sum_{k=0}^{\infty} \alpha_k^2 < \infty$ .

( $\mathcal{R}_5$ ) the *dynamic or adaptive* stepsize rule

$$\alpha_k = \frac{\varphi(\mathbf{x}^k) - \varphi_k}{\|\mathbf{u}^k\|^2}, \text{ or } \alpha_k = \frac{\varphi(\mathbf{x}^k) - \varphi_k}{\max\{1, \|\mathbf{u}^k\|^2\}} \quad (2.42)$$

where  $\varphi_k$  is an adaptively computed estimate of  $\varphi^*$ ; several further variants of this rule, which can be considered as an approximation of ( $\mathcal{R}_2$ ) when  $\varphi^*$  is not known, depend on how  $\varphi_k$  is defined.

In literature we can find convergence results for each of these selection strategies; for the constant stepsize ( $\mathcal{R}_1$ ) in [15, 77] the sequence  $\{\varphi(\mathbf{x}^k)\}$  is proved to convergence to a suboptimal value, i.e.  $\liminf_{k \rightarrow \infty} \varphi(\mathbf{x}^k) \leq \varphi^* + C\alpha$ , being  $C > 0$ . Choosing  $\alpha_k$  as shown in ( $\mathcal{R}_3$ ) and ( $\mathcal{R}_5$ ) one can obtain  $\lim_{k \rightarrow \infty} \varphi(\mathbf{x}^k) = \varphi^*$  and  $\min_{x^* \in \Omega^*} \|\mathbf{x}^* - \mathbf{x}^k\| \rightarrow 0$  (with the assumption  $\varepsilon_k = 0$  in the latter case). Finally, the convergence of the sequence  $\{\mathbf{x}^k\}$  to a solution  $\mathbf{x}^*$  can be proved in the cases ( $\mathcal{R}_2$ ), with  $\varepsilon_k = 0$ , and ( $\mathcal{R}_4$ ) [1, 68].

The key property that the stepsize parameter has to induce on the iterates (2.39) which is exploited in the standard convergence analysis for subgradient methods is the *quasi-Féjer monotonicity* with respect to the solution set  $\Omega^*$

$$\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 \leq \|\mathbf{x}^k - \mathbf{x}^*\|^2 + \eta_k \quad \forall \mathbf{x}^* \in \Omega^*$$

for some nonnegative sequence  $\{\eta_k\}$  such that  $\sum \eta_k < \infty$  (see [7, 41]).

In the smooth case the stepsize  $\alpha_k$  has the role to ensure a sufficient decrease on the objective function, as we pointed out in section 2.1. The approaches shown (and others similar) are difficult to extend to subgradient methods, since in the latter ones  $\alpha_k$  plays a quite different role than in the smooth case.

In the same section mentioned above, we reported the results concerning the convergence properties when a scaling matrix is introduced in the iteration. We propose a similar approach, considering a scaled or a variable metric  $\varepsilon$ -subgradient method

$$\mathbf{x}^{k+1} = P_{\Omega, D_k^{-1}} \left( \mathbf{x}^k - \alpha_k D_k \mathbf{u} \right) \quad (2.43)$$

where  $D_k$  is a symmetric positive definite matrix and the stepsize  $\alpha_k$  is chosen as in  $(\mathcal{R}_4)$  or  $(\mathcal{R}_5)$ . Before giving the main convergence results regarding this particular choices, we state the following lemma.

**Lemma 2.6.** *Let  $\{L_k\}$  be a sequence of positive numbers such that  $L_k^2 = 1 + \gamma_k$ ,  $\gamma_k \geq 0$ , where  $\sum_{k=0}^{\infty} \gamma_k < \infty$ . Let  $\theta_k = \prod_{j=0}^k L_j^2$  for any  $k \geq 0$ . Then the sequence  $\{\theta_k\}$  is bounded.*

*Proof.* We want to show that there exists a constant  $M > 0$  such that  $\theta_k \leq M$  for all  $k \geq 0$ . By the monotonicity of the logarithm, this is true if and only if  $\log(\theta_k) \leq \log(M)$ . By definition of  $\theta_k$  we have

$$\log(\theta_k) = \sum_{j=0}^k \log(L_j^2) \leq \sum_{j=0}^{\infty} \log(L_j^2) \quad (2.44)$$

Thus, if the series on the right hand side of (2.44) converges, the quantities  $\theta_k$  are bounded for all  $k$ . We observe that, since  $L_k^2 = 1 + \gamma_k$ , where  $\gamma_k \rightarrow 0$  as  $k$  diverges, by the known limit  $\lim_{k \rightarrow \infty} \frac{\log(1 + \gamma_k)}{\gamma_k} = 1$ , the series  $\sum_{j=0}^{\infty} \log(L_j^2)$  and  $\sum_{j=0}^{\infty} \gamma_j$  have the same behaviour. Thus, since by hypothesis the latter one is convergent, the result follows.  $\square$

Using the previous Lemma and the properties of the projection operator stated in Lemma 2.1, we are able to prove the following theorem.

**Proposition 2.4.** *Assume that the set of the solutions  $\Omega^*$  of (2.1) is nonempty. Moreover, assume that there exists a positive constant  $\rho$  such that  $\|\mathbf{u}^k\| \leq \rho$  for any  $k$  and a sequence of positive numbers  $\{L_k\}$  such that  $\|D_k\| \leq L_k$ ,  $\|D_k^{-1}\| \leq L_k$ , with  $1 \leq L_k \leq L$  for some positive constant  $L$ , for all  $k \geq 0$ . If the following conditions holds*

$$\lim_{k \rightarrow \infty} \varepsilon_k = 0 \quad (2.45)$$

$$\sum_{k=0}^{\infty} \alpha_k = \infty \quad (2.46)$$

$$\sum_{k=0}^{\infty} \alpha_k^2 < \infty, \quad \sum_{k=0}^{\infty} \varepsilon_k \alpha_k < \infty \quad (2.47)$$

$$L_k^2 = 1 + \gamma_k, \quad \sum_{k=0}^{\infty} \gamma_k < \infty \quad (2.48)$$

then, the sequence  $\{\mathbf{x}^k\}$  generated by the iteration (2.43) converges to a solution of (2.1).

*Proof.* For all  $k$  let us define  $\mathbf{z}^k = \mathbf{x}^k - \alpha_k D_k \mathbf{u}^k$ . By Lemma 2.1 part (b), we have that

$$\begin{aligned} \|\mathbf{x}^{k+1} - \mathbf{x}^k\|_{D_k^{-1}} &= \|P_{X, D_k^{-1}}(\mathbf{z}^k) - P_{X, D_k^{-1}}(\mathbf{x}^k)\|_{D_k^{-1}} \\ &\leq \|\mathbf{z}^k - \mathbf{x}^k\|_{D_k^{-1}} = \|\alpha_k D_k \mathbf{u}^k\|_{D_k^{-1}} \\ &= \alpha_k \|\mathbf{u}^k\|_{D_k} \leq \alpha_k L_k^{\frac{1}{2}} \|\mathbf{u}^k\|_2 \\ &\leq \alpha_k L_k^{\frac{1}{2}} \rho \end{aligned} \tag{2.49}$$

Then, thanks to Lemma 2.1, for any  $\tilde{\mathbf{x}} \in \Omega^*$  we can write

$$\begin{aligned} \alpha_k^2 L_k \rho^2 &+ \|\mathbf{x}^k - \tilde{\mathbf{x}}\|_{D_k^{-1}}^2 - \|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|_{D_k^{-1}}^2 \geq \\ &\geq \|\mathbf{x}^{k+1} - \mathbf{x}^k\|_{D_k^{-1}}^2 + \|\mathbf{x}^k - \tilde{\mathbf{x}}\|_{D_k^{-1}}^2 - \|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|_{D_k^{-1}}^2 \\ &= 2(\mathbf{x}^k - \tilde{\mathbf{x}})^t D_k^{-1}(\mathbf{x}^k - \mathbf{x}^{k+1}) \\ &= 2(\mathbf{x}^k - \tilde{\mathbf{x}})^t D_k^{-1}(\mathbf{x}^k - \mathbf{z}^k) + 2(\mathbf{x}^k - \tilde{\mathbf{x}})^t D_k^{-1}(\mathbf{z}^k - \mathbf{x}^{k+1}) \\ &= 2\alpha_k (\mathbf{x}^k - \tilde{\mathbf{x}})^t \mathbf{u}^k + 2(\mathbf{x}^k - \mathbf{z}^k)^t D_k^{-1}(\mathbf{z}^k - \mathbf{x}^{k+1}) + 2(\mathbf{z}^k - \tilde{\mathbf{x}})^t D_k^{-1}(\mathbf{z}^k - \mathbf{x}^{k+1}) \\ &= 2\alpha_k (\mathbf{x}^k - \tilde{\mathbf{x}})^t \mathbf{u}^k + 2(\mathbf{x}^k - \mathbf{z}^k)^t D_k^{-1}(\mathbf{z}^k - \mathbf{x}^{k+1}) + \\ &\quad + 2(\mathbf{z}^k - \tilde{\mathbf{x}})^t D_k^{-1}(\mathbf{z}^k - P_{X, D_k^{-1}}(\mathbf{z}^k)) \\ &\geq 2\alpha_k (\mathbf{x}^k - \tilde{\mathbf{x}})^t \mathbf{u}^k + 2(\mathbf{x}^k - \mathbf{z}^k)^t D_k^{-1}(\mathbf{z}^k - \mathbf{x}^{k+1}) \\ &= 2\alpha_k (\mathbf{x}^k - \tilde{\mathbf{x}})^t \mathbf{u}^k + 2(\mathbf{x}^k - \mathbf{z}^k)^t D_k^{-1}(\mathbf{z}^k - \mathbf{x}^k) + 2(\mathbf{x}^k - \mathbf{z}^k)^t D_k^{-1}(\mathbf{x}^k - \mathbf{x}^{k+1}) \\ &= 2\alpha_k (\mathbf{x}^k - \tilde{\mathbf{x}})^t \mathbf{u}^k - 2(\mathbf{x}^k - \mathbf{z}^k)^t D_k^{-1}(\mathbf{x}^k - \mathbf{z}^k) + 2\alpha_k (\mathbf{x}^k - \mathbf{x}^{k+1})^t \mathbf{u}^k \\ &\geq 2\alpha_k (\mathbf{x}^k - \tilde{\mathbf{x}})^t \mathbf{u}^k - 2L_k \alpha_k^2 \|\mathbf{u}^k\|^2 - 2\alpha_k \|\mathbf{x}^k - \mathbf{x}^{k+1}\| \cdot \|\mathbf{u}^k\| \\ &\geq 2\alpha_k (\mathbf{x}^k - \tilde{\mathbf{x}})^t \mathbf{u}^k - 2L_k \alpha_k^2 \|\mathbf{u}^k\|^2 - 2\alpha_k^2 L_k \|\mathbf{u}^k\|^2 \\ &\geq 2\alpha_k (\mathbf{x}^k - \tilde{\mathbf{x}})^t \mathbf{u}^k - 4\alpha_k^2 L_k \rho^2 \\ &\geq 2\alpha_k (\varphi(\mathbf{x}^k) - \varphi(\tilde{\mathbf{x}}) - \varepsilon_k) - 4\alpha_k^2 L_k \rho^2 \\ &\geq -2\alpha_k \varepsilon_k - 4\alpha_k^2 L_k \rho^2 \\ &\geq -2L_k \alpha_k \varepsilon_k - 4\alpha_k^2 L_k \rho^2 \end{aligned} \tag{2.50}$$

where the first inequality follows from (2.49), while we use Lemma 2.1 part (a) in the second one, the definition of  $\mathbf{z}^k$  and the Cauchy–Schwartz inequality in the third one, the assumption  $\|D_k\| \leq L_k$ ,  $\|D_k^{-1}\| \leq L_k$  and the fact that  $\|\mathbf{x}^{k+1} - \mathbf{x}^k\| \leq \alpha_k L_k \|\mathbf{u}^k\|$  in the fourth one, the bound  $\|\mathbf{u}^k\| \leq \rho$  in the fifth one, the definition of  $\varepsilon$ -subgradient in the sixth one, the fact that  $\tilde{\mathbf{x}} \in \Omega^*$  in the seventh one and the inequality  $L_k \geq 1$  in the last one.

Upon rearranging terms, this yields

$$\|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|_{D_k^{-1}}^2 \leq \|\mathbf{x}^k - \tilde{\mathbf{x}}\|_{D_k^{-1}}^2 + 5L_k \alpha_k^2 \rho^2 + 2L_k \alpha_k \varepsilon_k \tag{2.51}$$

Since we have

$$\begin{aligned} \|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|_{D_k^{-1}}^2 &\geq \lambda_{\min}(D_k^{-1}) \|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|^2 \geq \frac{1}{L_k} \|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|^2 \\ \|\mathbf{x}^k - \tilde{\mathbf{x}}\|_{D_k^{-1}}^2 &\leq \lambda_{\max}(D_k^{-1}) \|\mathbf{x}^k - \tilde{\mathbf{x}}\|^2 \leq L_k \|\mathbf{x}^k - \tilde{\mathbf{x}}\|^2 \end{aligned} \tag{2.52}$$

we can also write

$$\begin{aligned}\|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|^2 &\leq L_k^2 \|\mathbf{x}^k - \tilde{\mathbf{x}}\|^2 + 5L_k^2 \alpha_k^2 \rho^2 + 2L_k^2 \alpha_k \varepsilon_k \\ &\leq L_k^2 \|\mathbf{x}^k - \tilde{\mathbf{x}}\|^2 + \sigma L_k^2 \alpha_k^2 + 2L_k^2 \alpha_k \varepsilon_k\end{aligned}$$

where  $\sigma = 5\rho^2$ .

By repeatedly applying the previous inequality we obtain

$$\begin{aligned}\|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|^2 &\leq L_k^2 \|\mathbf{x}^k - \tilde{\mathbf{x}}\|^2 + \sigma L_k^2 \alpha_k^2 + 2L_k^2 \alpha_k \varepsilon_k \\ &\leq L_k^2 (L_{k-1}^2 \|\mathbf{x}^{k-1} - \tilde{\mathbf{x}}\|^2 + \sigma L_{k-1}^2 \alpha_{k-1}^2 + 2L_{k-1}^2 \alpha_{k-1} \varepsilon_{k-1}) + \sigma L_k^2 \alpha_k^2 + 2L_k^2 \alpha_k \varepsilon_k \\ &\leq \theta_0^k \|\mathbf{x}^0 - \tilde{\mathbf{x}}\|^2 + \sigma \sum_{j=0}^k \theta_j^k \alpha_j^2 + 2 \sum_{j=0}^k \theta_j^k \alpha_j \varepsilon_j\end{aligned}\quad (2.53)$$

where  $\theta_j^k = \prod_{i=j}^k L_i^2$ ,  $j \leq k$ . Since  $L_i^2 \geq 1$  we have  $1 \leq \theta_j^k \leq \theta_{j-1}^k \leq \theta_0^k$ , which implies

$$\begin{aligned}\|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|^2 &\leq \theta_0^k \|\mathbf{x}^0 - \tilde{\mathbf{x}}\|^2 + \sigma \theta_0^k \sum_{j=0}^k \alpha_j^2 + 2\theta_0^k \sum_{j=0}^k \alpha_j \varepsilon_j \\ &\leq M \left( \|\mathbf{x}^0 - \tilde{\mathbf{x}}\|^2 + \sigma \sum_{j=0}^k \alpha_j^2 + 2 \sum_{j=0}^k \alpha_j \varepsilon_j \right)\end{aligned}$$

where the last inequality follows from Lemma 2.6. Thus, by conditions (2.47), the sequence  $\{\mathbf{x}^k\}$  is bounded.

In order to show that  $\{\mathbf{x}^k\}$  converges to a solution of (2.1), we now consider inequality (2.50), which, in view of (2.52), results in

$$\begin{aligned}\|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|^2 &\leq L_k^2 \|\mathbf{x}^k - \tilde{\mathbf{x}}\|^2 + 5L_k^2 \alpha_k^2 \rho^2 - 2\alpha_k L_k (\mathbf{x}^k - \tilde{\mathbf{x}})^t \mathbf{u}^k \\ &\leq L_k^2 \|\mathbf{x}^k - \tilde{\mathbf{x}}\|^2 + \sigma L_k^2 \alpha_k^2 + 2\alpha_k L_k (\tilde{\mathbf{x}} - \mathbf{x}^k)^t \mathbf{u}^k\end{aligned}$$

By repeatedly applying the previous inequality we obtain

$$\begin{aligned}\|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|^2 &\leq \theta_0^k \|\mathbf{x}^0 - \tilde{\mathbf{x}}\|^2 + \sigma \sum_{j=0}^k \theta_j^k \alpha_j^2 + \\ &\quad + 2 \sum_{j=0}^k \tilde{\theta}_j^k \alpha_j (\tilde{\mathbf{x}} - \mathbf{x}^j)^t \mathbf{u}^j\end{aligned}\quad (2.54)$$

where  $\tilde{\theta}_j^k = \theta_j^k / L_j$ . Since  $\tilde{\mathbf{x}} \in \Omega^*$  and  $\mathbf{u}^j \in \partial_{\varepsilon_k} f(\mathbf{x}^j)$ , for all  $j \geq 0$  we have

$$\varphi(\mathbf{x}^j) \geq \varphi(\tilde{\mathbf{x}}) \geq \varphi(\mathbf{x}^j) + (\tilde{\mathbf{x}} - \mathbf{x}^j)^t \mathbf{u}^j - \varepsilon_j \quad (2.55)$$

Hence,  $(\tilde{\mathbf{x}} - \mathbf{x}^j)^t \mathbf{u}^j \leq \varepsilon_j$  for all  $j \geq 0$ .

Now, we show that  $(\tilde{\mathbf{x}} - \mathbf{x}^j)^t \mathbf{u}^j \rightarrow 0$  for  $j \rightarrow \infty$ . To this end, assume by contradiction that  $(\tilde{\mathbf{x}} - \mathbf{x}^j)^t \mathbf{u}^j < -\varepsilon$  for some  $\varepsilon > 0$ . Then, by inequality (2.55) we obtain

$$\|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|^2 \leq M \|\mathbf{x}^0 - \tilde{\mathbf{x}}\|^2 + \sigma M \sum_{j=0}^k \alpha_j^2 - 2 \frac{\varepsilon}{L} \sum_{j=0}^k \alpha_j$$

where we applied the inequalities  $1 \leq \theta_j^k \leq \theta_{j-1}^k \leq \theta_0^k \leq M$ ,  $1 \leq L_k \leq L$ . Then, taking limits for  $k \rightarrow \infty$ , by assumption (2.46) we have an absurdum. Thus, there exists a subsequence  $\{\mathbf{x}^{k_i}\}$  such that  $\lim_i (\tilde{\mathbf{x}} - \mathbf{x}^{k_i})^t \mathbf{u}^{k_i} = 0$ . Then, from (2.55) and the assumption (2.47) it follows that  $\lim_i \varphi(\mathbf{x}^{k_i}) = \varphi^*$ . Since  $\{\mathbf{x}^k\}$  is bounded,  $\{\mathbf{x}^{k_i}\}$  is also bounded and, thus it has an accumulation point  $\mathbf{x}^\infty$ . By the continuity of  $\varphi(x)$  and by inequality (2.55) we can conclude that  $\mathbf{x}^\infty \in \Omega^*$ .

Now we show that the whole sequence  $\{\mathbf{x}^k\}$  converges to  $\mathbf{x}^\infty$ . Let  $\delta > 0$ ; since  $\mathbf{x}^\infty$  is an accumulation point of  $\{\mathbf{x}^k\}$  and from (2.47), there exists a positive integer  $m_\delta$  such that  $\|\mathbf{x}^\infty - \mathbf{x}^{m_\delta}\|^2 \leq \delta/(3M)$ ,  $\sum_{j=m_\delta}^{\infty} \alpha_k^2 < \delta/(3\sigma M)$  and  $\sum_{j=m_\delta}^{\infty} \alpha_k \varepsilon_k < \delta/(6M)$ . Then, for any  $k > m_\delta$ , using the same arguments as in (2.53), we obtain

$$\begin{aligned} \|\mathbf{x}^k - \mathbf{x}^\infty\|^2 &\leq \theta_{m_\delta}^{k-1} \|\mathbf{x}^{m_\delta} - \mathbf{x}^\infty\|^2 + \sigma \sum_{j=m_\delta}^{k-1} \theta_j^{k-1} \alpha_j^2 + 2 \sum_{j=m_\delta}^{k-1} \theta_j^{k-1} \alpha_j \varepsilon_j \\ &\leq M \|\mathbf{x}^{m_\delta} - \mathbf{x}^\infty\|^2 + \sigma M \sum_{j=m_\delta}^{\infty} \alpha_j^2 + 2M \sum_{j=m_\delta}^{\infty} \alpha_j \varepsilon_j \\ &\leq \delta \end{aligned}$$

Since  $\delta$  can be chosen arbitrarily small, then  $\{\mathbf{x}^k\}$  converges to  $\mathbf{x}^\infty$ .  $\square$

We discuss some issues about Proposition (2.4), relating these results with the recent literature, in particular with the papers [42, 43] and [76].

**Remark 2.1.** From equation (2.51), observing that

$$\begin{aligned} \|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|_{D_k^{-1}}^2 &\geq \lambda_{\min}(D_k^{-1}) \|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|^2 \\ &= \frac{\lambda_{\min}(D_k^{-1})}{\lambda_{\max}(D_{k+1}^{-1})} \lambda_{\max}(D_{k+1}^{-1}) \|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|^2 \\ &\geq \lambda_{\min}(D_k^{-1}) \lambda_{\min}(D_{k+1}) \|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|_{D_{k+1}^{-1}}^2 \\ &\geq \frac{1}{L_k L_{k+1}} \|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|_{D_{k+1}^{-1}}^2 \end{aligned}$$

we obtain

$$\|\mathbf{x}^{k+1} - \tilde{\mathbf{x}}\|_{D_{k+1}^{-1}}^2 \leq \zeta_k \|\mathbf{x}^k - \tilde{\mathbf{x}}\|_{D_k^{-1}}^2 + \xi \zeta_k \alpha_k^2 + 2L \zeta_k \alpha_k \varepsilon_k$$

where  $\zeta_k = \sqrt{(1 + \gamma_k)(1 + \gamma_{k+1})}$  and  $\xi = 5L\rho^2$ . By the assumptions made on  $\{\gamma_k\}$ , the sequence  $\{\zeta_k\}$  is bounded. We can also set  $\zeta_k = 1 + \eta_k$ , with  $\eta_k = \sqrt{(1 + \gamma_k)(1 + \gamma_{k+1})} - 1$ , and observe that the series  $\sum \eta_k$  and  $\sum \gamma_k$  have the same behaviour, thanks to the limit  $\lim_{z \rightarrow 0} (\sqrt{1+z} - 1)/z = 1/2$ . Then, since the assumption (2.48), we can conclude that  $\sum \eta_k$  is a convergent series.

Thus, the sequence  $\{\mathbf{x}^k\}_{k \in \mathbb{N}}$  is *quasi-Fejér monotone* with respect to  $\Omega^*$  relative to  $\{D_k^{-1}\}$ , in the sense of [42, Definition 3.1] and we could apply Proposition 3.2 in [42] (see also [43]) to obtain that  $\{\|\mathbf{x}^k - \tilde{\mathbf{x}}\|_{D_k^{-1}}\}$  converges and, thus,  $\{\mathbf{x}^k\}$  is bounded.

**Remark 2.2.** Variable metric was introduced also in [76, Chapter 5] in the context of subgradient methods for unconstrained problems (i.e.  $\Omega = \mathbb{R}^n$ ). In this case, setting  $D_k = B_k B_k^T$ , the scaling matrices are assumed to satisfy  $\|B_{k+1}^{-1} B_k\| \geq 1$  and  $\prod_{k=0}^{\infty} \|B_{k+1}^{-1} B_k\|^2 < \infty$ . Even if the second condition is verified under the assumptions of Proposition 2.4, we observe that the requirement  $\|B_{k+1}^{-1} B_k\| \geq 1$  restricts the choice of the scaling matrix, strictly connecting the metrics adopted in two successive iterates.

We introduce also a variant of the method (2.43), obtained by modifying the stepsize  $\alpha_k$ .

**Corollary 2.1.** Let  $\{\mathbf{x}^k\}$  a sequence chosen as

$$\mathbf{x}^{k+1} = P_{\Omega, D_k^{-1}} \left( \mathbf{x}^k - \frac{\alpha_k}{\max(1, \|\mathbf{u}^k\|_{D_k})} D_k \mathbf{u}^k \right) \quad (2.56)$$

such that the assumptions of Proposition 2.4 are satisfied. Then,  $\{\mathbf{x}^k\}$  converges to a solution of problem (2.1).

*Proof.* Let us define  $\bar{\alpha}_k = \frac{\alpha_k}{\max(1, \|\mathbf{u}^k\|_{D_k})}$ . From the inequalities

$$1 \leq \max(1, \|\mathbf{u}^k\|_{D_k}) \leq \max(1, L^{\frac{1}{2}} \rho)$$

it follows that

$$\begin{aligned} \sum \bar{\alpha}_k &= \sum \frac{\alpha_k}{\max(1, \|\mathbf{u}^k\|_{D_k})} \geq \frac{1}{\max(L^{\frac{1}{2}} \rho, 1)} \sum \alpha_k = \infty \\ \sum \bar{\alpha}_k^2 &= \sum \frac{\alpha_k^2}{\max(1, \|\mathbf{u}^k\|_{D_k})^2} \leq \sum \alpha_k^2 < \infty \\ \sum \bar{\alpha}_k \varepsilon_k &= \sum \frac{\alpha_k}{\max(1, \|\mathbf{u}^k\|_{D_k})} \varepsilon_k \leq \sum \alpha_k \varepsilon_k < \infty \end{aligned}$$

Then, we can invoke Proposition 2.4 to obtain the result.  $\square$

### 2.6.1 Scaled $\varepsilon$ -subgradient level algorithm

A critical point of the implementation of the methods (2.43) and (2.56) is the selection of the steplength parameter. The principal aim is to obtain good performances, but it is still an open problem since these methods are quite sensitive to the option implemented. Borrowing the ideas of [24] and [56], in this section we describe a *level algorithm* that allows to adaptively compute a stepsize  $\alpha_k$  of the form (2.42) in the iteration (2.56). In our scheme we introduce the use of the  $\varepsilon$ -subgradient of  $f$  at the current iterate (instead of the subgradient) and a variable metric. The resulting scheme is detailed in Algorithm 8, whose underlying assumption is that, for any given  $\varepsilon_k \geq 0$ , we are able to provide an element  $\mathbf{u}^k$  of the set  $\partial_{\varepsilon_k} f(\mathbf{x}^k)$  such that  $\lim_{k \rightarrow \infty} \varepsilon_k = 0$ .

In Algorithm 8,  $\varphi_k^{rec} = \min_{j=0, \dots, k} \varphi(\mathbf{x}^j)$ . The index  $l$  represents the number of times  $\varphi^{lev}$  has been updated and  $k(l)$  is the iteration at which the updating occurred. Finally,  $\sigma_k$  is the cumulative path length between two successive updating of  $\varphi^{lev}$ .

**Algorithm 8: SSL**

Choose  $B > 0$ ,  $\nu_1, \nu_2 \in (0, 1)$ ,  $\varphi_{-1}^{rec} = \infty$ ,  $\varepsilon_{-1}^{rec} = 0$ ;  $k = 0$ ,  $l = 0$ ,  $k(l) = 0$ ,  $\delta_0 > 0$ ; choose  $x^0 \in \Omega$ .

FOR  $k = 0, 1, 2, \dots$  DO THE FOLLOWING STEPS

STEP 1. Computation of  $\varphi(\mathbf{x}^k)$

STEP 2. IF  $\varphi(\mathbf{x}^k) < \varphi_{k-1}^{rec}$   
 $\varphi_k^{rec} = \varphi(\mathbf{x}^k)$

ELSE

$$\varphi_k^{rec} = \varphi_{k-1}^{rec}$$

STEP 3. IF  $\varphi(\mathbf{x}^k) < \varphi_{k(l)}^{rec} - \nu_1 \delta_l$

$$k(l+1) = k, \sigma_k = 0, \delta_{l+1} = \delta_l, l = l + 1 \text{ and go to STEP 5}$$

STEP 4. IF  $\sigma_k > B$

$$k(l+1) = k, \sigma_k = 0, \delta_{l+1} = \nu_2 \delta_l, l = l + 1$$

STEP 5. Set  $\varphi_k^{lev} = \varphi_{k(l)}^{rec} - \delta_l$ .

STEP 6. Update the stepsize and compute the new iterate

$$\alpha_k = \frac{\varphi(\mathbf{x}^k) - \varphi_k^{lev}}{\max(1, \|\mathbf{u}^k\|_{D_k})};$$

$$\mathbf{x}^{k+1} = P_{X, D_k^{-1}} \left( \mathbf{x}^k - \alpha_k D_k \frac{\mathbf{u}^k}{\max(1, \|\mathbf{u}^k\|_{D_k})} \right) \quad (2.57)$$

STEP 7.  $\sigma_{k+1} = \sigma_k + \alpha_k$  and go to Step 1.

The idea of the procedure is the following: STEPS 2–5 have the aim to compute an estimation of the optimal value  $\varphi^*$ , which is employed to create a target value for the successive iterates, which the algorithm tries to approach to; if the procedure leads the sequence  $\varphi(\mathbf{x}^k)$  close to this target value or the iterates follow a long path without approaching it, the target level is updated. In the first case, i.e. when the condition at STEP 3 is satisfied,  $\varphi_k^{lev}$  is reduced at STEP 5 by diminishing the best value obtained  $\varphi^{rec}$  by the quantity  $\delta_l$ ; in the other case, if the inequality at STEP 4 is satisfied,  $\delta_l$  is reduced and, consequently,  $\varphi_k^{lev}$  is increased.

The main difference between the proposed procedure and the square summable sequence considered in Proposition 2.4 is that the former one does not necessarily converge to zero.

In the rest of this section we prove that the sequence  $\{f(x_k)\}$ , where  $x_k$  is computed by Algorithm 8, converges to the minimum of  $f$ , using similar techniques as in [77]. Before giving the main result, we recall the following technical lemma.

**Lemma 2.7.** *Assume that the set  $\Omega^*$  of the solutions of (2.1) is nonempty. Assume that there exists a positive constant  $\rho$  such that  $\|\mathbf{u}^k\| \leq \rho$  and a sequence of positive numbers  $\{L_k\}$  such that  $\|D_k\| \leq L_k$ ,  $\|D_k^{-1}\| \leq L_k$ , with  $1 \leq L_k \leq L$  for some positive constant  $L$ , for all  $k \geq 0$ . Given  $B > 0$  and  $\{\varepsilon_k\}$  such that  $\varepsilon_k \rightarrow 0$  as  $k \rightarrow \infty$  in Algorithm 8, we have  $l \rightarrow \infty$  and  $\delta_l \rightarrow 0$  as  $l \rightarrow \infty$ .*

*Proof.* Assume by contradiction that  $l$  takes only a finite number of values, say  $l = 0, 1, \dots, \bar{l}$ . In this case, we have  $\sigma_k + \alpha_k = \sigma_{k+1} \leq B$  for all  $k \geq k(\bar{l})$ , so that we have  $\lim_{k \rightarrow \infty} \alpha_k = 0$ . But this is impossible, since for all  $k \geq k(\bar{l})$  we have

$$\alpha_k = \frac{\varphi(\mathbf{x}^k) - \varphi_k^{lev}}{\max(1, \|\mathbf{u}^k\|_{D_k})} \geq (1 - \nu_1) \frac{\delta_{\bar{l}}}{\max(1, L^{\frac{1}{2}} \rho)} > 0$$

Hence  $l \rightarrow \infty$ . Let  $\delta = \lim_{l \rightarrow \infty} \delta_l$ . If  $\delta > 0$ , then from STEPS 3 and 4 of Algorithm 8, it follows that for all  $l$  large enough we have  $\delta_l = \delta$ . Given  $\mu > 0$ , for  $l$  large enough, we have  $\mu < \nu_1 \delta$  and, consequently, we can write

$$\begin{aligned} \varphi_{k(l+1)}^{rec} - \mu &\leq \varphi_{k(l)}^{rec} - \nu_1 \delta \\ \varphi_{k(l+1)}^{rec} - \varphi_{k(l)}^{rec} &\leq -\nu_1 \delta + \mu < 0 \end{aligned}$$

implying that  $\inf_{k \geq 0} \varphi(\mathbf{x}^k) = -\infty$ ; this is a contradiction, since  $\varphi(x)$  is bounded below.  $\square$

Using the Lemma above and exploiting the results of the convergence proof of the previous Proposition, we can state the following result, which can be considered as a generalization of Proposition 2.7 in [77], which only deals with the case  $D_k = I$ ,  $\varepsilon_k = 0$  for all  $k$ .

**Proposition 2.5.** *Under the same assumptions of the previous Lemma, if  $L_k^2 = 1 + \gamma_k$ , with  $\sum_1^\infty \gamma_k < \infty$ , for SSL we have  $\tilde{\varphi} = \inf_{k \geq 0} \varphi(\mathbf{x}^k) = \varphi(\mathbf{x}^*)$ , with  $\mathbf{x}^* \in \Omega^*$ .*

*Proof.* The first part of the proof aims to show that  $\sum_j \alpha_j = \infty$  and runs as Proposition 2.7 in [77]. For sake of completeness, we report below the detailed derivation of the result.

From the previous Lemma,  $\delta_l \rightarrow 0$  as  $l \rightarrow \infty$ . Let  $S = \{l \in \{1, 2, \dots\}, \delta_l = \nu_2 \delta_{l-1}\}$ . Then, from Step 4 and 6 of Algorithm 8, we obtain

$$\sigma_k = \sigma_{k-1} + \alpha_{k-1} = \sum_{j=k(l)}^{k-1} \alpha_j$$

so that  $k(l+1) = k$  and  $l+1 \in S$  whenever  $\sum_{j=k(l)}^{k-1} \alpha_j > B$  at Step 4. Hence

$$\sum_{j=k(l-1)}^{k(l)-1} \alpha_j > B \quad \forall l \in S$$

and since the cardinality of  $S$  is infinite, we have

$$\sum_{k=k(l)}^\infty \alpha_k \geq \sum_{l \geq \bar{l}, l \in S} \sum_{j=k(l-1)}^{k(l)-1} \alpha_j > \sum_{l \geq \bar{l}, l \in S} B = \infty \quad (2.58)$$

Now in order to obtain a contradiction, assume that  $\tilde{\varphi} > \varphi(\mathbf{x}^*)$ , so that for some  $\tilde{\mathbf{y}} \in \Omega$  and some  $\eta > 0$

$$\tilde{\varphi} - \eta \geq \varphi(\tilde{\mathbf{y}}) \quad (2.59)$$

Since  $\delta_l \rightarrow 0$  and  $\varepsilon_k \rightarrow 0$ , there are a large enough  $\bar{l}$  and a large enough  $\bar{k}$  such that, for all  $l \geq \bar{l}$  and  $k(l) \geq \bar{k}$ , we have  $\delta_l < \frac{\eta}{2}$  and  $\varepsilon_k < \frac{\eta}{2}$ ; then for all  $k \geq \bar{k} = \max\{k(\bar{l}), \bar{k}\}$

$$\varphi_k^{lev} - \varepsilon_k = \varphi_{k(l)}^{rec} - \varepsilon_k - \delta_l > \tilde{\varphi} - \eta \geq \varphi(\tilde{\mathbf{y}})$$

From this inequality, by Lemma 2.1 (iii), the definition of  $\varepsilon$ -subgradient, the definition of  $\alpha_k$  we obtain

$$\begin{aligned} \|\mathbf{x}^{k+1} - \tilde{\mathbf{y}}\|_{D_k^{-1}}^2 &\leq \|\mathbf{x}^k - \tilde{\mathbf{y}}\|_{D_k^{-1}}^2 - 2 \frac{\alpha_k}{\max(1, \|\mathbf{u}^k\|_{D_k})} (\varphi(\mathbf{x}^k) - \varphi(\tilde{\mathbf{y}}) - \varepsilon_k) + \alpha_k^2 \\ &\leq \|\mathbf{x}^k - \tilde{\mathbf{y}}\|_{D_k^{-1}}^2 - 2 \frac{\alpha_k}{\max(1, \|\mathbf{u}^k\|_{D_k})} (\varphi(\mathbf{x}^k) - \varphi_k^{lev} - \varepsilon_k) + \alpha_k^2 \\ &\leq \|\mathbf{x}^k - \tilde{\mathbf{y}}\|_{D_k^{-1}}^2 - \alpha_k^2 \end{aligned}$$

In view of (2.52) with  $\tilde{\mathbf{x}} = \tilde{\mathbf{y}}$  and  $L_k \geq 1$ , we can write

$$\|\mathbf{x}^{k+1} - \tilde{\mathbf{y}}\|^2 \leq L_k^2 \|\mathbf{x}^k - \tilde{\mathbf{y}}\|^2 - \alpha_k^2 \quad (2.60)$$

By repeatedly applying the previous inequality we obtain

$$\|\mathbf{x}^{k+1} - \tilde{\mathbf{y}}\|^2 \leq \theta_k^k \|\mathbf{x}^{\bar{k}} - \tilde{\mathbf{y}}\|^2 - \sum_{j=k(\bar{l})}^k \theta_{j+1}^k \alpha_j^2$$

where  $\theta_j^k = L_j^2 \cdot \dots \cdot L_k^2$ ; since  $1 \leq \theta_j^k \leq \theta_0^k \leq M$  we have

$$\sum_{j=\bar{k}}^{\infty} \alpha_j^2 \leq M \|\mathbf{x}^{\bar{k}} - \tilde{\mathbf{y}}\|^2$$

and consequently  $\sum_{j=\bar{k}}^{\infty} \alpha_k^2 < \infty$ . Then  $\alpha_k \rightarrow 0$  as  $k \rightarrow \infty$  and, from (2.58),  $\sum_{j=\bar{k}}^{\infty} \alpha_k = \infty$ .

Now we show that  $\sum \alpha_k \varepsilon_k < \infty$ . Indeed, since  $\varepsilon_k \rightarrow 0$  as  $k \rightarrow \infty$ , there exists  $\bar{k}$  such that  $2\varepsilon_k < \eta$  for  $k \geq \bar{k}$ , where  $\eta$  is such that (2.59) holds. We consider the inequality

$$\|\mathbf{x}^{k+1} - \tilde{\mathbf{y}}\|_{D_k^{-1}}^2 \leq \|\mathbf{x}^k - \tilde{\mathbf{y}}\|_{D_k^{-1}}^2 + \alpha_k^2 - 2 \frac{\alpha_k}{\max(1, \|\mathbf{u}^k\|_{D_k})} (\mathbf{u}^k)^t (\mathbf{x}^k - \tilde{\mathbf{y}}) \quad (2.61)$$

For the convexity of  $\varphi$ , the inequality (2.59) and inequality  $2\varepsilon_k < \eta$ , we have

$$\varphi(\mathbf{x}^k) + (\mathbf{u}^k)^t (\tilde{\mathbf{y}} - \mathbf{x}^k) - \varepsilon_k \leq \varphi(\tilde{\mathbf{y}}) \leq \tilde{\varphi} - \eta \leq \varphi(\mathbf{x}^k) - 2\varepsilon_k$$

Then we have

$$(\mathbf{u}^k)^t (\tilde{\mathbf{y}} - \mathbf{x}^k) \leq -\varepsilon_k$$

Using this inequality in (2.61), we obtain

$$\|\mathbf{x}^{k+1} - \tilde{\mathbf{y}}\|_{D_k^{-1}}^2 \leq \|\mathbf{x}^k - \tilde{\mathbf{y}}\|_{D_k^{-1}}^2 + \alpha_k^2 - 2 \frac{\alpha_k \varepsilon_k}{\max(1, \|\mathbf{u}^k\|_{D_k})}$$

Using the same arguments as above, we obtain

$$\|\mathbf{x}^{k+1} - \tilde{\mathbf{y}}\|^2 \leq L_k^2 \|\mathbf{x}^k - \tilde{\mathbf{y}}\|^2 + L_k^2 \alpha_k^2 - 2 \frac{\alpha_k \varepsilon_k}{\max(1, L^{\frac{1}{2}} \rho)}$$

By repeatedly applying the previous inequality we have

$$\begin{aligned} \|\mathbf{x}^{k+1} - \tilde{\mathbf{y}}\|^2 &\leq \theta_{\bar{k}}^k \|x^{(\bar{k})} - \tilde{\mathbf{y}}\|^2 + \theta_{\bar{k}}^k \sum_{j=\bar{k}}^k \alpha_j^2 - \frac{2}{\max(1, L^{\frac{1}{2}}\rho)} \sum_{j=\bar{k}}^k \alpha_j \varepsilon_j \\ &\leq M \left( \|\mathbf{x}^{\bar{k}} - \tilde{\mathbf{y}}\|^2 + \sum_{j=\bar{k}}^k \alpha_j^2 \right) - \frac{2}{\max(1, L^{\frac{1}{2}}\rho)} \sum_{j=\bar{k}}^k \alpha_j \varepsilon_j \end{aligned}$$

Then we have

$$\sum_{j=\bar{k}}^{\infty} \alpha_j \varepsilon_j \leq \frac{M}{2} \max(1, L^{\frac{1}{2}}C) \left( \|\mathbf{x}^{\bar{k}} - \tilde{\mathbf{y}}\|^2 + \sum_{j=\bar{k}}^{\infty} \alpha_j^2 \right) < \infty$$

According to Corollary 2.1 we have  $\tilde{\varphi} = \varphi^*$  that contradicts (2.59).  $\square$

The SSL algorithm can be generalized by introducing a variable path bound  $B$  and different strategies for the updating of  $\delta_l$  (see [77] for some examples).

### 2.6.2 Scaled Primal–Dual Hybrid Gradient Method

We present here an application of the previous SSL algorithm when  $\varphi$  is the sum of two convex, proper l.s.c functions, namely  $\varphi = f_1 + f_2 \circ M$ , such that  $\text{diam}(\text{dom}(f_2^*))$  is finite and where  $M$  is a suitable matrix; the problem (2.1) becomes hence

$$\min_{\mathbf{x} \in \Omega} f_1(\mathbf{x}) + f_2(M\mathbf{x}) \quad (2.62)$$

We propose a Scaled Primal–Dual Hybrid Gradient (SPDHG) method for the solution of (2.62):

$$\mathbf{y}^{k+1} = (I + \tau_k \partial f_2^*)^{-1}(\mathbf{y}^k + \tau_k M\mathbf{x}^k) \quad (2.63)$$

$$\mathbf{x}^{k+1} = \text{P}_{\Omega, D_k^{-1}}(\mathbf{x}^k - \alpha_k D_k(\mathbf{w}^k + M^t \mathbf{y}^{k+1})) \quad (2.64)$$

where  $\mathbf{w}_k \in \partial_{\delta_k} f_1(\mathbf{x}^k)$ , for some  $\delta_k \geq 0$ ,  $\{\tau_k\}$  and  $\{\alpha_k\}$  are the dual and primal steplength sequences respectively. The results that allows us to classify this method as a  $\varepsilon$ -subgradient method is the following lemma.

**Lemma 2.8.** [21, Lemma 1] *Let  $\mathbf{y}^{k+1}$  defined as in (2.63). Then,  $\mathbf{y}^{k+1} \in \text{dom}(f_2^*)$  and, thus,  $M^t \mathbf{y}^{k+1} \in \partial_{\sigma_k}(f \circ M)(\mathbf{x}^k)$ , where  $\sigma_k = f_2(M\mathbf{x}^k) + f_2^*(\mathbf{y}^{k+1}) - (\mathbf{y}^{k+1})^t M\mathbf{x}^k$ . If there exists a positive number  $D$  such that  $\text{diam}(\text{dom}(f_2^*)) \leq D$ , then  $\sigma_k \leq (2\tau_k)^{-1} D^2$ .*

*Proof.* From (2.63) we have

$$\begin{aligned} \mathbf{y}^{k+1} &= \arg \min_{\mathbf{y}} f_2^*(\mathbf{y}) + \frac{1}{2\tau_k} \|\mathbf{y} - (\mathbf{y}^k + \tau_k M\mathbf{x}^k)\|^2 \\ &= \arg \min_{\mathbf{y}} f_2^*(\mathbf{y}) - \langle \mathbf{y}, M\mathbf{x}^k \rangle + \frac{1}{2\tau_k} \|\mathbf{y} - \mathbf{y}^k\|^2 \end{aligned}$$

Then

$$\begin{aligned}
f_2^*(\mathbf{y}^{k+1}) - \langle \mathbf{y}^{k+1}, M\mathbf{x}^k \rangle &\leq f_2^*(\mathbf{y}^{k+1}) - \langle \mathbf{y}^{k+1}, M\mathbf{x}^k \rangle + \frac{1}{2\tau_k} \|\mathbf{y}^{k+1} - \mathbf{y}^k\|^2 \\
&= \min_y f_2^*(\mathbf{y}) - \langle \mathbf{y}, M\mathbf{x}^k \rangle + \frac{1}{2\tau_k} \|\mathbf{y} - \mathbf{y}^k\|^2 \\
&\leq \min_y f_2^*(\mathbf{y}) - \langle \mathbf{y}, M\mathbf{x}^k \rangle + \frac{1}{2\tau_k} D^2 \\
&\leq -f_2(M\mathbf{x}^k) + \frac{1}{2\tau_k} D^2
\end{aligned}$$

which results in  $\sigma_k = f_2^*(\mathbf{y}^{k+1}) - \langle \mathbf{y}^{k+1}, M\mathbf{x}^k \rangle + f_2(M\mathbf{x}^k) \leq \frac{1}{2\tau_k} D^2$ .  $\square$

Then, for the additivity of the  $\varepsilon$ -subgradient (see Appendix A), we have  $\mathbf{u}^k = \mathbf{w}^k + M^t \mathbf{y}^{k+1} \in \partial_{\varepsilon_k} \varphi(\mathbf{x}^k)$ , being  $\varphi = f_1 + f_2 \circ M$ .

We develop two implementation of the SPDHG method, using the above observation and the results of the previous section; in the first case, prefixed sequences  $\{\tau_k\}, \{\alpha_k\}, \{L_k\}$  are selected: the following corollary, based on Lemma 2.8 and Proposition 2.4, states some conditions on these sequences to guarantee the convergence of  $\{\mathbf{x}^k\}$  to a point  $\mathbf{x}^* \in \Omega^*$ .

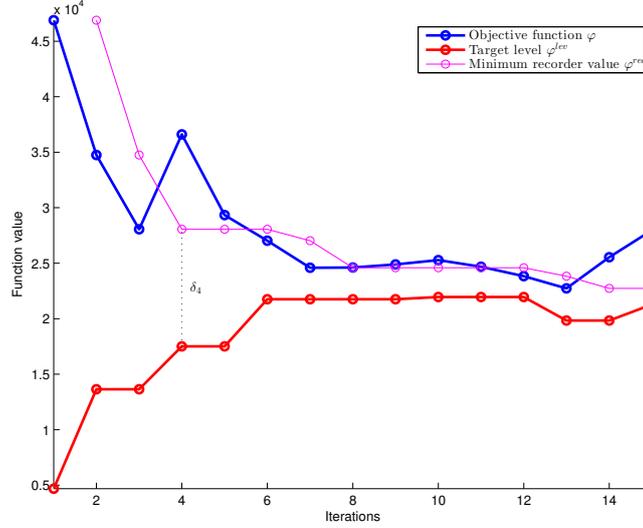
**Corollary 2.2.** *Let  $\{\mathbf{x}^k\}$  be the sequence generated by iteration (2.63)–(2.64). Assume that  $\mathbf{w}^k \in \partial_{\delta_k} f_1(\mathbf{x}^k)$  and that there exists  $\bar{\rho} > 0$  such that  $\|\mathbf{w}^k\| \leq \bar{\rho}$  for all  $k$ . Let the steplength sequences  $\{\tau_k\}, \{\alpha_k\}$  and the scaling matrix bounds  $\{L_k\}$  satisfy*

$$\alpha_k = \mathcal{O}\left(\frac{1}{k^p}\right), \quad \frac{1}{\tau_k} = \mathcal{O}(k^{-p}), \quad L_k = \sqrt{1 + \mathcal{O}\left(\frac{1}{k^q}\right)} \quad \frac{1}{2} < p \leq 1, \quad q > 1. \quad (2.65)$$

Moreover, assume that  $\delta_k$  converges to zero at least as  $\tau_k^{-1}$ . If the set of the solutions of (2.1) is nonempty and  $\text{diam}(\text{dom}(f_2^*))$  is finite, then, the sequence  $\{\mathbf{x}^k\}$  converges to a solution of (2.1).

*Proof.* As observed above, we have  $\mathbf{u}^k = \mathbf{w}^k + M^t \mathbf{y}^{k+1} \in \partial_{\varepsilon_k} \varphi(\mathbf{x}^k)$ , where  $\varepsilon_k = \delta_k + \sigma_k$ . Since  $\text{diam}(\text{dom}(f_2^*))$  is finite, we can apply Lemma 2.8 obtaining  $\sigma_k \leq (2\tau_k)^{-1} D^2$ . By the assumption (2.65) on  $\tau_k$  and on  $\delta_k$  we obtain that  $\varepsilon_k = \mathcal{O}\left(\frac{1}{k^p}\right)$  and, as a consequence,  $\alpha_k \varepsilon_k = \mathcal{O}\left(\frac{1}{k^{2p}}\right)$ . Since  $\frac{1}{2} < p \leq 1$  and  $q > 1$ , all assumptions (2.45)–(2.48) of Theorem 2.4 are satisfied and we obtain the result.  $\square$

On the other hand, we can employ the SSL procedure to dynamically compute the stepsize  $\alpha_k$ ; in this case, a sequence  $L_k = \sqrt{1 + \gamma_k}$ , with  $\sum \gamma_k < \infty$  has to be provided. For sake of simplicity we consider the case  $\delta_k = 0$ , thus  $\sigma_k = f_2(M\mathbf{x}^k) - (\mathbf{y}^{k+1})^t M\mathbf{x}^k + f_2^*(\mathbf{y}^{k+1})$  is controlled by the dual stepsize  $\tau_k$  (see Lemma 2.8). Proposition 2.5 allows to conclude that the sequence  $\{f_1(\mathbf{x}^k) + f_2(M\mathbf{x})\}$  generated by SPDHG method combined with SSL algorithm converges to the minimum of the objective function in (2.62), as stated in the following Corollary, under the same boundedness assumptions as in Corollary 2.2. In Figure 2.1 the idea beyond the entire procedure is shown.



**Figure 2.1:** Behaviour of the  $\varphi$ ,  $\varphi^{rec}$  and  $\varphi^{lev}$  in the first 15 iterations of the Algorithm 8 for the test-problem micro described in Section 2.7. The blue line refers to the objective function's values  $\varphi(\mathbf{x}^k)$ , the magenta one represents the recorded minimum values  $\varphi_k^{rec}$ ; the red line is the target value  $\varphi_k^{lev}$  at each iteration. The  $\varphi^{lev}$  values are used as an estimate of the minimum value  $\varphi^*$ ; the procedure tries, at each  $k$ -step, to lead the objective function to the estimate  $\varphi^{lev}$ . We can observe that in the first 6 iterations this estimate is too low; thus the procedure at each step computes a greater value for  $\varphi^{lev}$ . Once the objective function is near the current  $\varphi^{lev}$  value (for example at the 12-th iteration), the algorithm decreases the estimate of the minimum in order to force the minimization of  $\varphi$ . If the new estimate results to be too low (iteration 13-th and the subsequents), then the procedure again increases  $\varphi^{lev}$ .

**Corollary 2.3.** Let  $\{\mathbf{x}^k\}$  be the sequence generated by Algorithm 8. Here  $\mathbf{u}^k = \mathbf{w}^k + M^t \mathbf{y}^{k+1}$  in (2.57),  $\mathbf{w}^k \in \partial f_1(\mathbf{x}^k)$  and  $\mathbf{y}^{k+1}$  is computed as in (2.63). Assume that  $\lim_{k \rightarrow \infty} \tau_k = \infty$ ,  $L_k = \sqrt{1 + \gamma_k}$  with  $\sum \gamma_k < \infty$  and that there exists  $\rho > 0$  such that  $\|\mathbf{u}^k\| \leq \rho$ . If the set of the solutions of (2.1) is nonempty and  $\text{diam}(\text{dom}(f_2^*))$  is finite, then the sequence  $\{f_1(\mathbf{x}^k) + f_2(M\mathbf{x}^k)\}$  converges to  $f_1(\mathbf{x}^*) + f_2(M\mathbf{x}^*)$ , with  $\mathbf{x}^* \in \Omega^*$ .

*Proof.* Since  $\mathbf{w}^k \in \partial f_1(x_k)$ , by Lemma 2.8 we have  $\mathbf{u}^k \in \partial_{\varepsilon_k} f(x_k)$ , where  $\varepsilon_k = f_1(M\mathbf{x}^k) + f_1^*(\mathbf{y}^{k+1}) - \mathbf{y}^{k+1t} A x_k$ . Since  $\text{diam}(\text{dom}(f_1^*))$  is finite, we can apply the second part of Lemma 2.8 obtaining  $\varepsilon_k \leq (2\tau_k)^{-1} D^2$ , for a positive constant  $D$  such that  $\text{diam}(\text{dom}(f_1^*)) \leq D$ . Since  $\lim_{k \rightarrow \infty} \tau_k = \infty$ , we have  $\lim_{k \rightarrow \infty} \varepsilon_k = 0$  and by Theorem 2.5 we obtain the result.  $\square$

### 2.6.3 Application to image restoration

We apply the SPDHG method to deblurring problems from data corrupted by Poisson noise. We assume that the detected image is the vectorized square image of dimension  $N$ , where  $N = n^2$ . We employ the Total Variation regularization: we set  $f_1(\mathbf{x}) = \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n})$ ,  $f_2(M\mathbf{x}) = \beta \mathbf{TV}(\mathbf{x}) = \beta \sum_i \|\nabla_i \mathbf{x}\|$  where  $M\mathbf{x} = \nabla \mathbf{x}$ ,  $\nabla = (\nabla_1^t, \dots, \nabla_N^t)^t$  as in (2.27) and  $\Omega$  is the nonnegative orthant. In this case,

we have

$$\mathbf{w}^k = \nabla \mathbf{KL}(\mathbf{H}\mathbf{x}^k + bg\mathbf{1}; \mathbf{g}\mathbf{n}) = \mathbf{H}^t \left( \mathbf{1} - \frac{\mathbf{g}\mathbf{n}}{\mathbf{H}\mathbf{x}^k + bg\mathbf{1}} \right)$$

Thus  $\delta_k = 0$ . Furthermore, the resolvent operator for  $f_2$  is just the projection on the set  $\mathcal{Y} \subset \mathbb{R}^{2N}$ , defined in (2.27). In order to devise a suitable scaling matrix  $D_k$ , we use a strategy similar to the one presented in section 2.1.1. We split the  $\varepsilon$ -subgradient of the objective function in two nonnegative parts, such that

$$\mathbf{u}^k = V(\mathbf{x}^k) - U(\mathbf{x}^k) \quad V > 0, \quad U \geq 0$$

and we define the diagonal entries of the matrix  $D_k$  as the projection of  $\frac{\mathbf{x}^k}{V(\mathbf{x}^k)}$  on the set  $[L_k^{-1}, L_k]^N$ . This strategy has the advantage to agree with the non negativity constraints and strongly depends on the form of  $\mathbf{u}^k$ .

We have to find a decomposition of  $\mathbf{u}^k = \nabla \mathbf{KL}(\mathbf{H}\mathbf{x} + bg\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta M^t \mathbf{y}^{k+1}$  as the difference of two non negative terms. Thanks to the hypothesis (1.13) on the matrix  $\mathbf{H}$ , the gradient of the Kullback–Leibler can be split in

$$U_{\mathbf{KL}}(\mathbf{x}^k) = \mathbf{H}^t \mathbf{1}, \quad V_{\mathbf{KL}}(\mathbf{x}) = \mathbf{H}^t \frac{\mathbf{g}\mathbf{n}}{\mathbf{H}\mathbf{x}^k + bg\mathbf{1}}$$

It remains to compute a decomposition of the  $\beta M^t \mathbf{y}^{k+1}$  term: it could be done by writing it as a function depending on  $\mathbf{x}^k$ . Let's consider the dual variable  $\mathbf{y} \in \mathbb{R}^{2N}$  as

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_N \end{pmatrix}$$

where  $\mathbf{y}_i \equiv \begin{pmatrix} y_{2i-1} \\ y_{2i} \end{pmatrix} \in \mathbb{R}^{2 \times N}$ , for  $i = 1, \dots, N$ . The updating rule (2.63) can be written as

$$\begin{aligned} \bar{\mathbf{y}}^k &= \mathbf{y}^k + \tau_k \beta M \mathbf{x}^k \\ \mathbf{y}^{k+1} &= S_k \bar{\mathbf{y}}^k \end{aligned}$$

where  $S_k$  is a diagonal  $2N \times 2N$  matrix with the following diagonal entries

$$(S_k)_{2i-1, 2i-1} = (S_k)_{2i, 2i} = \frac{1}{\max\{1, \|\bar{\mathbf{y}}_i^k\|\}}, \quad i = 1, \dots, N \quad (2.66)$$

If the method is initialized with  $\mathbf{y}^0 = 0$ , the dual variable can be written as

$$\begin{aligned} \mathbf{y}^0 &= 0 \\ \mathbf{y}^1 &= \beta \tau_0 S_0 M \mathbf{x}^0 \\ \mathbf{y}^2 &= \beta S_1 (\tau_0 S_0 M \mathbf{x}^0 + \tau_1 M \mathbf{x}^1) \\ \mathbf{y}^3 &= \beta S_2 (\tau_0 S_1 S_0 M \mathbf{x}^0 + \tau_1 S_1 M \mathbf{x}^1 + \tau_2 M \mathbf{x}^2) \\ &\vdots \\ \mathbf{y}^{k+1} &= \beta \sum_{j=0}^k \tau_j \tilde{S}_j^k M \mathbf{x}^j \end{aligned}$$

where

$$\tilde{S}_j^k = \prod_{i=j}^k S_i$$

As a consequence, the  $\varepsilon$ -subgradient of  $f_2 \circ M$  employed in (2.64) can be expressed as

$$\beta M^t \mathbf{y}^{k+1} = \beta^2 \sum_{j=0}^k \tau_j M^t \tilde{S}_j^k M \mathbf{x}^j \quad (2.67)$$

The following lemma, which directly follows from the definition of  $M$ , indicates a possible decomposition of each term in the summation at the right hand side of (2.67) as the difference between a positive and a nonnegative term.

**Lemma 2.9.** *Every matrix–vector product of the form  $M^t S M \mathbf{x}$  where  $S$  is a  $2N \times 2N$  diagonal matrix with positive entries such that  $S_{2\ell, 2\ell} = S_{2\ell-1, 2\ell-1} = s_\ell$ ,  $\ell = 1, \dots, N$ ,  $\mathbf{x} \geq 0$ , can be decomposed as*

$$M^t S M \mathbf{x} = V_S(\mathbf{x}) - U_S(\mathbf{x})$$

where

$$\begin{aligned} (V_S(\mathbf{x}))_{i,j} &= (2s_{i,j} + s_{i,j-1} + s_{i-1,j}) \mathbf{x}_{i,j} \geq 0 \\ (U_S(\mathbf{x}))_{i,j} &= s_{i,j} (\mathbf{x}_{i+1,j} + \mathbf{x}_{i,j+1}) + s_{i,j-1} \mathbf{x}_{i,j-1} + s_{i-1,j} \mathbf{x}_{i-1,j} \geq 0 \end{aligned}$$

with the correspondence

$$s_\ell \equiv s_{i,j}, \quad j = \lfloor (\ell - 1)/n \rfloor + 1, \quad i = \ell - \lfloor (\ell - 1)/n \rfloor \cdot n$$

where  $\lfloor \cdot \rfloor$  denotes the integer quotient.

*Proof.* Since  $M = \begin{pmatrix} \nabla_1 \\ \vdots \\ \nabla_N \end{pmatrix}$ , the matrix  $M^t S M$  can be expressed in the following way:

$$M^t S M = s_1 \nabla_1^t \nabla_1 + \dots + s_N \nabla_N^t \nabla_N$$

We observe that when  $\ell$  is related to the pair  $(i, j)$ ,  $i = \ell - \lfloor (\ell - 1)/n \rfloor \cdot n$ ,  $j = \lfloor (\ell - 1)/n \rfloor + 1$ , with  $i, j \neq n$ , the matrix product  $D_\ell = \nabla_\ell^t \nabla_\ell$  is a  $N \times N$  matrix whose entries are zero except for the following components:

- on the main diagonal:  $(D_\ell)_{\ell, \ell} = 2, (D_\ell)_{\ell+n, \ell+n} = 1$ ;
- off-diagonal nonzero entries:  $(D_\ell)_{\ell, \ell+1} = -1, (D_\ell)_{\ell+1, \ell} = -1, (D_\ell)_{\ell, \ell+n} = -1, (D_\ell)_{\ell+n, \ell} = -1$ .

With a similar argument, it is possible to show that  $D_\ell$  for  $\ell$  related to a pair  $(i, N)$  or  $(N, j)$  also has positive entries only in two positions on the main diagonal and negative values in four off-diagonal positions. For example, the nonzero entries of  $D_N$  are:  $(D_N)_{n, n} = 1, (D_N)_{n, N} = -1, (D_N)_{N-n, N-n} = 1, (D_N)_{N-n, N} = -1, (D_N)_{N, N-n} = -1, (D_N)_{N, N} = 2$ .

Consequently, in view of the relation between  $\ell$  and the pair  $(i, j)$ , each  $\ell$ -th diagonal entry of  $M^t S M$  is obtained by the following terms:

- $s_{i,j} \nabla_{i,j}^t \nabla_{i,j} = 2s_{i,j}$
- $s_{i-1,j} \nabla_{i-1,j}^t \nabla_{i-1,j} = s_{i-1,j}$
- $s_{i,j-1} \nabla_{i,j-1}^t \nabla_{i,j-1} = s_{i,j-1}$

Then,  $M^t SMx$  can be written as  $V(\mathbf{x}) - U(\mathbf{x})$ , where  $V(\mathbf{x})$  is given by

$$(V_S(\mathbf{x}))_{i,j} = (2s_{i,j} + s_{i-1,j} + s_{i,j-1})\mathbf{x}_{i,j}$$

while  $(U(\mathbf{x}))_{i,j}$  is obtained by the negative contribution of  $D_{i,j}$ ,  $D_{i,j-1}$  and  $D_{i-1,j}$ :

$$(U(\mathbf{x}))_{i,j} = s_{i,j}(\mathbf{x}_{i+1,j} + \mathbf{x}_{i,j+1}) + s_{i-1,j}\mathbf{x}_{i-1,j} + s_{i,j-1}\mathbf{x}_{i,j-1}$$

□

Hence, the subgradient  $\mathbf{u}^k$  can be splitted out in two parts  $U(\mathbf{x}^k) \geq 0, V(\mathbf{x}^k) > 0$ , with

$$V(\mathbf{x}^k) = \mathbf{H}^t \mathbf{1} + \beta^2 \sum_{j=0}^k \tau_j V_{\bar{S}_j^k}(\mathbf{x}_j) \quad (2.68)$$

It is worth to notice that since the matrix  $D_k$  is diagonal and the constraint set  $\Omega$  is the nonnegative orthant, the projection  $P_{\Omega, D_k^{-1}}$  is simply the usual Euclidean projection  $P_\Omega$ .

**Remark 2.3.** Introducing the following auxiliary vectors:

$$\mathbf{p}_{i,j}^k = (\mathbf{p}_{i,j}^{k-1} + \beta^2 \tau_k \mathbf{x}_{i,j}^k) s_{i,j}^k \quad (2.69)$$

$$\mathbf{q}_{i,j}^k = (\mathbf{q}_{i,j}^{k-1} + \beta^2 \tau_k \mathbf{x}_{i,j}^k) s_{i-1,j}^k \quad (2.70)$$

$$\mathbf{r}_{i,j}^k = (\mathbf{r}_{i,j}^{k-1} + \beta^2 \tau_k \mathbf{x}_{i,j}^k) s_{i,j-1}^k \quad (2.71)$$

for  $i, j = 1, \dots, N$

it can be shown that the computation of the sum in the second term of (2.68) can be written as

$$\beta^2 \sum_{j=0}^k \tau_j V_{\bar{S}_j^k}(\mathbf{x}_j) = 2\mathbf{p}^k + \mathbf{q}^k + \mathbf{r}^k \quad (2.72)$$

For sake of simplicity, we limit ourselves to show that this is true for  $k = 0, 1$ . Indeed, from Lemma 2.9 and from (2.69)–(2.71) we have

$$\begin{aligned} V^R(\mathbf{x}^0)_{i,j} &= \beta^2 \tau_0 (V_{S_0} \mathbf{x}^0)_{i,j} \\ &= \beta^2 \tau_0 (2s_{i,j}^0 + s_{i-1,j}^0 + s_{i,j-1}^0) \mathbf{x}_{i,j}^0 \\ &= 2\mathbf{p}_{i,j}^0 + \mathbf{q}_{i,j}^0 + \mathbf{r}_{i,j}^0 \end{aligned}$$

$$\begin{aligned} V^R(\mathbf{x}^1)_{i,j} &= \beta^2 \tau_0 (V_{S_0 S_1} \mathbf{x}^0)_{i,j} + \beta^2 \tau_1 (V_{S_1} \mathbf{x}^1)_{i,j} \\ &= \beta^2 \tau_0 (2s_{i,j}^0 s_{i,j}^1 + s_{i-1,j}^0 s_{i-1,j}^1 + s_{i,j-1}^0 s_{i,j-1}^1) \mathbf{x}_{i,j}^0 + \\ &\quad + \beta^2 \tau_1 (2s_{i,j}^1 + s_{i-1,j}^1 + s_{i,j-1}^1) \mathbf{x}_{i,j}^1 \\ &= 2(\beta^2 \tau_0 s_{i,j}^0 \mathbf{x}_{i,j}^0 + \beta^2 \tau_1 \mathbf{x}_{i,j}^1) s_{i,j}^1 + \\ &\quad + (\beta^2 \tau_0 s_{i-1,j}^0 \mathbf{x}_{i,j}^0 + \beta^2 \tau_1 \mathbf{x}_{i,j}^1) s_{i-1,j}^1 + \\ &\quad + (\beta^2 \tau_0 s_{i,j-1}^0 \mathbf{x}_{i,j}^0 + \beta^2 \tau_1 \mathbf{x}_{i,j}^1) s_{i,j-1}^1 \\ &= 2(\mathbf{p}_{i,j}^0 + \beta^2 \tau_1 \mathbf{x}_{i,j}^1) s_{i,j}^1 + (\mathbf{q}_{i,j}^0 + \beta^2 \tau_1 \mathbf{x}_{i,j}^1) s_{i-1,j}^1 + (\mathbf{r}_{i,j}^0 + \beta^2 \tau_1 \mathbf{x}_{i,j}^1) s_{i,j-1}^1 \\ &= 2\mathbf{p}_{i,j}^1 + \mathbf{q}_{i,j}^1 + \mathbf{r}_{i,j}^1 \end{aligned}$$

The entire procedure for the SPDHG algorithm is shown in Algorithm 9. The stepsize selection strategies presented above could be employed in SPDHG: the choice of prefixed sequences requires that the assumptions of Corollary 2.2 are satisfied.

---

**Algorithm 9: SPDHG**


---

Choose the starting point  $\mathbf{x}^0 \in X$  and set  $\mathbf{y}_0 = 0$ ,  $\mathbf{p}^{(-1)} = \mathbf{q}^{(-1)} = \mathbf{r}^{(-1)} = 0$ . Choose the sequences  $\{\alpha_k\}$ ,  $\{\tau_k\}$ ,  $\{\gamma_k\}$ .

FOR  $k = 0, 1, 2, \dots$  DO THE FOLLOWING STEPS

STEP 1. Compute  $\bar{\mathbf{y}}^k = \mathbf{y}^k + \beta\tau_k M\mathbf{x}^k$ ;

STEP 2. Compute  $s_\ell^k = \frac{1}{\max\{1, \|\bar{\mathbf{y}}^k_\ell\|\}}$ ,  $\ell = 1, \dots, n$  and define  $S_k$  as in (2.66);

STEP 3. Dual update:  $\mathbf{y}^{k+1} = S_k \bar{\mathbf{y}}^k$ ;

STEP 4. Auxiliary vectors update for the decomposition:

$$\begin{aligned} \mathbf{p}_{i,j}^k &= (\mathbf{p}_{i,j}^{k-1} + \beta^2 \tau_k \mathbf{x}_{i,j}^k) s_{i,j}^k \\ \mathbf{q}_{i,j}^k &= (\mathbf{q}_{i,j}^{k-1} + \beta^2 \tau_k \mathbf{x}_{i,j}^k) s_{i-1,j}^k \\ \mathbf{r}_{i,j}^k &= (\mathbf{r}_{i,j}^{k-1} + \beta^2 \tau_k \mathbf{x}_{i,j}^k) s_{i,j-1}^k \\ &\text{for } i, j = 1, \dots, N \end{aligned}$$

STEP 5. Compute the positive part of the decomposition:

$$V(\mathbf{x}^k) = \mathbf{H}^t \mathbf{1} + (2\mathbf{p}^k + \mathbf{q}^k + \mathbf{r}^k)$$

STEP 6. Compute the scaling matrix:

$$\begin{aligned} L_k &= \sqrt{1 + \gamma_k} \\ (D_k)_{\ell,\ell} &= \min \left\{ L_k, \max \left\{ L_k^{-1}, \frac{x_\ell^k}{V(\mathbf{x}^k)_\ell} \right\} \right\} \end{aligned}$$

STEP 7. Primal update:  $\mathbf{x}^{k+1} = P_{\geq 0}(\mathbf{x}^k - \alpha_k D_k (\nabla f_1(\mathbf{x}^k) + \beta M^t \mathbf{y}^{k+1}))$ .

---

## 2.7 Numerical Comparison

In this section we compare the numerical behavior of some of the algorithms presented above for the solution of the deblurring problem

$$\arg \min_{\mathbf{x} \geq 0} \varphi_\beta(\mathbf{x}) \equiv \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta \mathbf{TV}(\mathbf{x})$$

The regularization parameter has been empirically chosen by solving the minimization problem for different values of  $\beta$  with very high accuracy and selecting the one corresponding to the minimum relative  $\ell_2$  distance from the original image  $\mathbf{x}^*$ . The parameter found is denoted by  $\beta_{\text{opt}}$ .

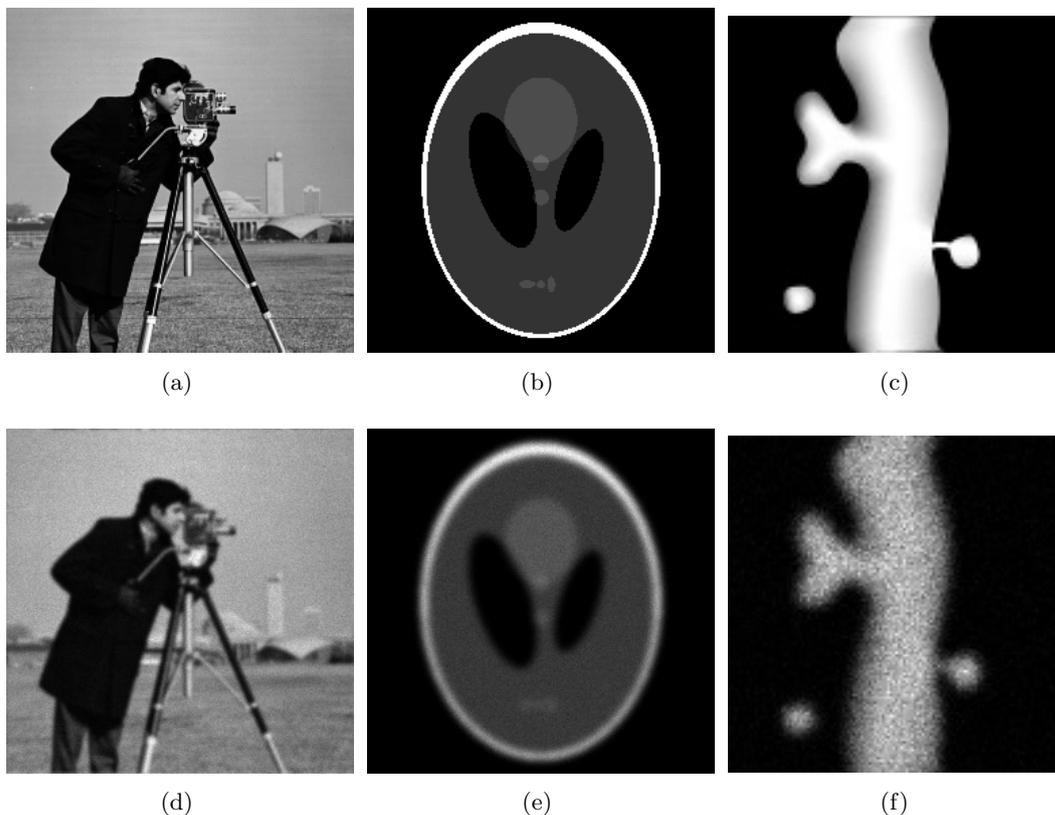
The tested algorithms are

- **PDHG**[21] - Algorithm 9 with the choice  $D_k = \mathbf{I}$ ;  $\alpha_k$  is *a priori* sequence such that

$$\sum_k \alpha_k = \infty, \quad \sum_k \alpha_k^2 < \infty$$

i.e. is a diminishing, divergent series and square summable sequence.

- **SPDHG** - Algorithm 9;  $\alpha_k$  is an *a priori* diminishing, divergent series and square summable sequence, as in the previous case.
- **SL** - Algorithms 8 with  $D_k = I$  and  $\mathbf{u}^k = \nabla \text{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta M^t \mathbf{y}^{k+1}$ , where  $\mathbf{y}^{k+1}$  is given in (2.63).
- **SSL** - Algorithms 8; in this case,  $D_k$  is given by Algorithm 9.
- **CP** - Algorithm 7; (Chambolle & Pock method tailored for image deblurring problem (2.38). In our setting  $\tau\sigma = 0.1$ ,  $\theta = 1$ ; here 0.1 is an estimate of  $1/\|M^t M\|$ , where  $M$  is given in (2.36)).
- **PIDSplit+** - Algorithm 6 (PIDSplit+ method, with  $\gamma = \frac{a}{\beta}$ , with  $a \in \mathbb{R}^+$ ).



**Figure 2.2:** Test problems for numerical comparison: in the first row the original images are shown, while in the second one the perturbed ones are presented.

The data set on which the algorithms are tested is given by three different problems (see Figure 2.2):

- ▷ *cameraman*: the classical cameraman  $256 \times 256$  image; the blurring is due to a Gaussian PSF with standard deviation 1.3, truncated at the  $9 \times 9$  central pixels. The values of the exact image lie in the interval  $[7, 253]$ , while the perturbed and

noisy image  $\mathbf{gn}$  ranges between 3.5 and 249.5; the  $bg$  value is set to 0. The relative  $\ell_2$  distance of  $\mathbf{gn}$  from the exact data is 0.1290. The value of  $\beta_{\text{opt}}$  is 0.005.

- ▷ *phantom*: this  $256 \times 256$  image is the Shepp–Logan phantom generated by the Matlab function `phantom`; the original image ranges in  $[0, 1000]$ , it is blurred by a Gaussian PSF of standard deviation 3 and truncated at the central  $9 \times 9$  pixels. A constant background value of 10 is added to the blurred image; the values of  $\mathbf{gn}$  lie in  $[1, 934]$ . The relative  $\ell_2$  distance of the perturbed image from the original one is 0.4661. The value of  $\beta_{\text{opt}}$  is 0.00526.
- ▷ *micro*: the original image is the confocal microscopy phantom of size  $128 \times 128$  described in [99] multiplied by 10; the PSF is the one used in [99], truncated at the  $9 \times 9$  central pixels. Here the background value is set to 0, the original image pixels are in the range  $[10, 690]$ , while  $\mathbf{gn}$  has values in  $[1, 778]$ . The relative  $\ell_2$  relative distance between the original and perturbed images is 0.1442. The computed  $\beta_{\text{opt}}$  is 0.0477.

	$k$	$e^k$	$E^k$	time				
phantom								
PDHG	*3000	0.1775	0.3140	68.72				
SPDHG	847	0.0157	0.2331	21.54				
SL	*3000	0.2745	0.3822	75.63				
SSL	1725	0.0025	0.2323	53.05				
CP	405	0.0223	0.2361	14.15				
PIDSPLIT+	351	0.0154	0.2370	18.47				

	$k$	$e^k$	$E^k$	time		$k$	$e^k$	$E^k$	time
cameraman						micro			
PDHG	1462	0.0032	0.0872	34.81		2141	0.0289	0.0750	12.43
SPDHG	914	0.0007	0.0875	24.01		550	0.0094	0.0720	3.66
SL	327	0.0257	0.0880	7.52		1549	0.0696	0.0979	9.35
SSL	316	0.0021	0.0872	8.23		916	0.0130	0.0715	6.25
CP	280	0.0038	0.0872	9.33		1926	0.0283	0.0748	17.87
PIDSPLIT+	199	0.0023	0.0873	10.56		281	0.0112	0.0720	3.54

**Table 2.1:** Relative reconstruction error corresponding to the stopping criterion (2.73) with  $\text{tol} = 10^{-6}$ . The asterisk denotes that the condition (2.73) was not met in the first 3000 iterations.

In order to evaluate the behaviour of the procedures, we have computed the minimum  $\tilde{\varphi}$  of  $\varphi_{\beta_{\text{opt}}}$  and the corresponding minimum argument  $\tilde{\mathbf{x}}$  by running 100000 iterations of PIDSplit+ algorithm. Then, we evaluate the progress toward this solution at each iteration in terms of the  $\ell_2$  relative error from the minimum point and the relative difference from the optimal value

$$e^k = \frac{\|\mathbf{x}^k - \tilde{\mathbf{x}}\|}{\|\tilde{\mathbf{x}}\|} \quad f^k = \frac{\|\varphi(\mathbf{x}^k) - \tilde{\varphi}\|}{\|\tilde{\varphi}\|}$$

Following the assumptions of Corollaries 2.2 and 2.3, we choose the sequences of

		PDHG						
		$\tau_k$	$\alpha_k$	$\gamma_k$				
cameraman		$0.9 + 10^{-2}k$	$(0.04 + 10^{-5}k)^{-1}$	-				
micro		$0.9 + 10^{-3}k$	$(0.04 + 10^{-4}k)^{-1}$	-				
phantom		$0.9 + 10^{-3}k$	$(0.2 + 10^{-5}k)^{-1}$	-				
		SPDHG						
		$\tau_k$	$\alpha_k$	$\gamma_k$				
cameraman		$0.5 + 5 \cdot 10^{-3}k$	$(0.5 + 10^{-5}k)^{-1}$	$10^{13}k^{-2}$				
micro		$0.4 + 10^{-5}k$	$(0.4 + 10^{-5}k)^{-1}$	$10^{13}k^{-2}$				
phantom		$0.5 + 10^{-4}k$	$(0.5 + 10^{-5}k)^{-1}$	$10^{13}k^{-2}$				
		SL	SSL	PIDSplit+	CP			
		$\tau_k$	$\gamma_k$	$a$	$\sigma$	$\tau$		
cameraman		$0.5 + 5 \cdot 10^{-2}k$	-	$0.7 + 5 \cdot 10^{-3}k$	$10^{13}k^{-2}$	5	100	0.001
micro		$0.9 + 10^{-3}k$	-	$0.4 + 10^{-5}k$	$10^{13}k^{-2}$	50	10	0.01
phantom		$0.9 + 10^{-3}k$	-	$0.9 + 10^{-4}k$	$10^{13}k^{-2}$	50	1000	0.0001

**Table 2.2:** Parameter settings.

parameters as follows

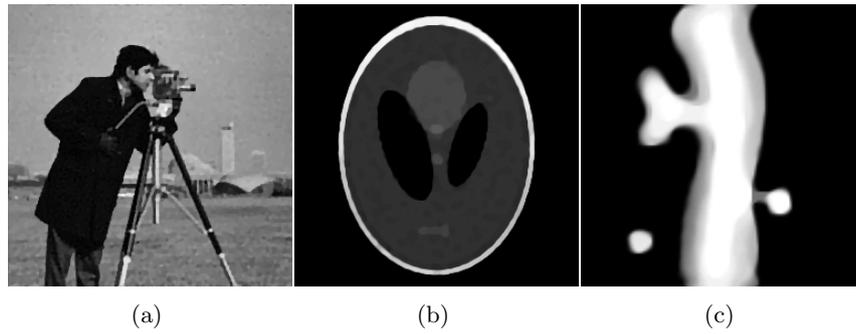
$$\tau_k = t_1 + t_2k, \quad \alpha_k = \frac{1}{t_3 + t_4k}, \quad \gamma_k = \frac{t_5}{k^{1+t_6}}$$

In order to illustrate the effectiveness of the methods, the values  $t_i$  have been manually optimized for each test problem to obtain a faster decrease of  $e^k$ . The same manual procedure has been adopted also for the parameters of CP and PIDSplit+ (see Table 2.2). Moreover, for the initialization of both SL and SSL, we adopt the rule  $\delta_0 = 0.9f(\mathbf{x}^0)$ , while the other parameters are  $\nu_1 = \nu_2 = 0.5$ ,  $B = 0.9\|\mathbf{u}^0\| \|D_0\|_{\infty}^{\frac{1}{2}}$ .

In the first numerical experiment we force the algorithms to perform 3000 iterations, in order to compare the convergence rate of the methods; the results obtained are depicted in Figure 2.4. From our numerical experience, it seems that introducing the scaling can give an acceleration towards the solution, with both steplength  $\alpha_k$  strategy choices. Moreover, we observe that the best results are reached by setting the initial values for the scaling matrix very large (see Table 2.2). It is worth noticing that the adaptive choice for  $\alpha_k$  combined with the proposed scaling strategy seems to work well, reaching results which are close to the ones obtained by tuning these parameters in the optimal way. Indeed, SPDHG and its non scaled version PDHG are very sensitive to the choice of  $\tau_k$  and  $\alpha_k$ : depending strongly their performances on these parameters, it is quite difficult to devise a general strategy. Moreover, it is worth mentioning that also the effectiveness of PIDSplit+ and CP depends on the choice of the relative parameters.

In the second experiment in order to show the behaviour of the proposed methods in image restoration framework, we employ the following stopping criterion:

$$\frac{|\varphi_{\beta}(\mathbf{x}^{k+1}) - \varphi_{\beta}(\mathbf{x}^k)|}{|\varphi_{\beta}(\mathbf{x}^{k+1})|} \leq tol \quad \text{and} \quad \sum_{i=k-4}^k \frac{|\varphi_{\beta}(\mathbf{x}^{i+1}) - \varphi_{\beta}(\mathbf{x}^i)|}{|\varphi_{\beta}(\mathbf{x}^{i+1})|} \leq 10tol \quad (2.73)$$

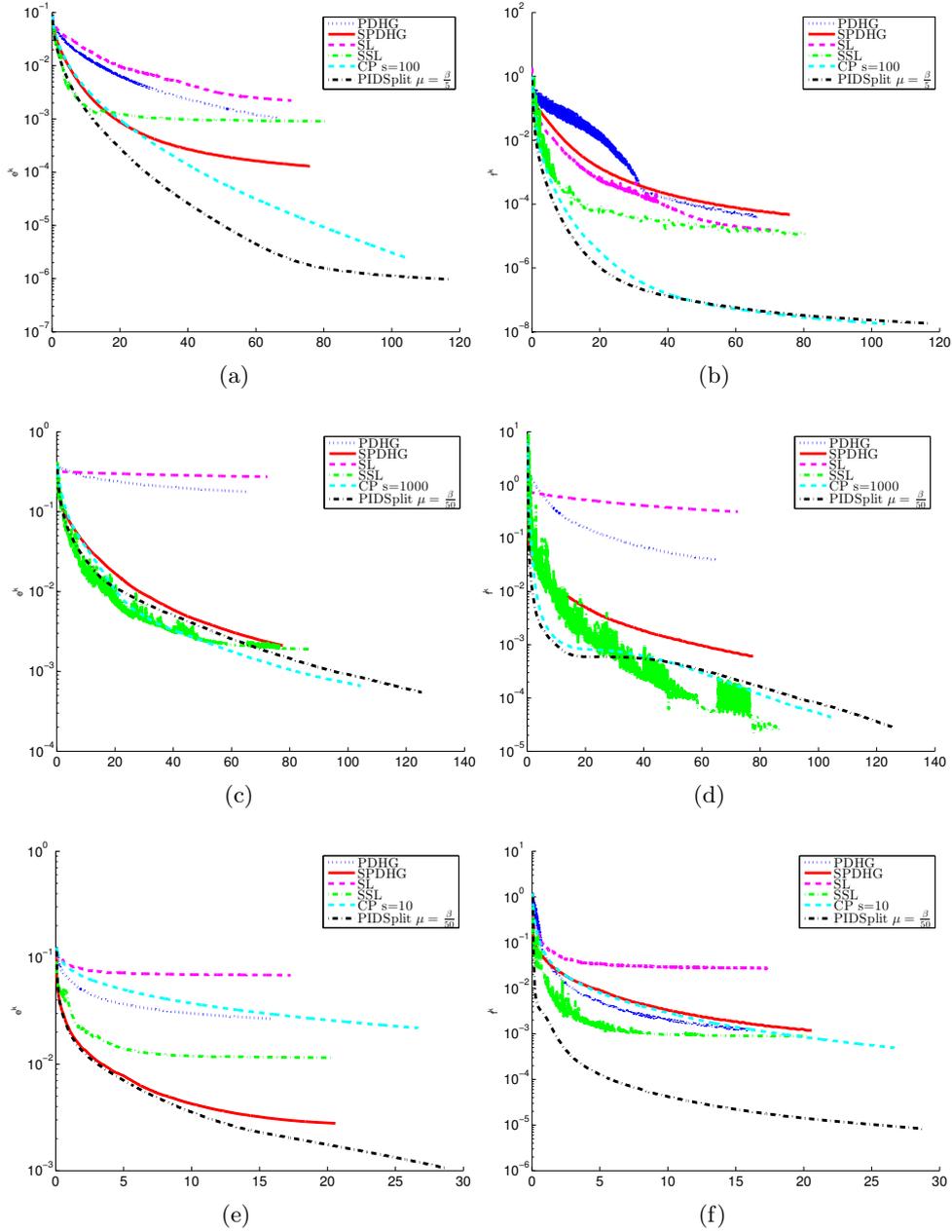


**Figure 2.3:** Reconstructed images obtained with SPDHG method.

with  $tol = 10^{-6}$  and  $\beta = \beta_{\text{opt}}$ . The numerical results are presented in Table 2.1, where we show the relative reconstruction error  $E^k$ :

$$E^k = \frac{\|\mathbf{x}^k - \mathbf{x}^*\|}{\|\mathbf{x}^*\|}$$

where  $\mathbf{x}^*$  is the original image. In Figure 2.3 the reconstructed images obtained with SPDHG are shown. These results are coherent with those in [18], where scaled gradient methods show to be more effective in providing a good reconstruction than nonscaled ones.



**Figure 2.4:** First numerical experiment. In the first row the results concerning the cameraman test problem are shown, while in the second one the plots refer to the phantom problem and, finally, in the third one the plots regard the micro test problem. In the first column, for each problem, the behaviour of the relative minimization error is depicted, while in the second one the plots are related to the behaviour of the relative distance from the minimum function value.

## Chapter 3

# Numerical Methods for Parameter Estimation in Poisson Data Inversion

The parameter  $\beta$  that appears in the variational problem

$$\min_{\mathbf{x}} \varphi_{\beta}(\mathbf{x}) \equiv \varphi_0(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta\varphi_1(\mathbf{x})$$

has a crucial role. As already shown in the previous chapters, we define the *optimal parameter* as the value for  $\beta$  which gives the minimum reconstruction error. This value is not known, hence we need some methods in order to *estimate* this value: in this chapter we propose two different methodologies which have exactly this aim for data corrupted by Poisson noise. The first one is based on solving a *discrepancy equation* [13], while the other one is based on a constrained approach. We prove the equivalence of these models, i.e. we show that they theoretically provide us with the same estimate for  $\beta$ ; moreover, we propose a numerical method based on the algorithm by Dai and Fletcher [45] to solve the discrepancy equation and a particular implementation of the ADMM method to solve the constrained model.

### 3.1 Two Estimation Models

We consider the linear model

$$\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1} = \mathbf{g}\mathbf{n}$$

where the assumptions (1.5) hold. As already shown, in order to recover the original signal  $\mathbf{x}^*$  we have to solve

$$\min_{\mathbf{x}} \varphi_{\beta}(\mathbf{x}) \equiv \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta\varphi_1(\mathbf{x}) \quad (3.1)$$

As shown in [101], the expected value for  $\mathbf{KL}(\mathbf{H}\mathbf{x}^* + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n})$  is  $N/2$ , being  $N$  the dimension of  $\mathbf{x}$ . Therefore, if we define the *normalized discrepancy function*

$$\mathcal{D}_{\mathbf{H}}(\mathbf{x}; \mathbf{y}) = \frac{2}{N}\mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) \quad (3.2)$$

and we denote as  $\mathbf{x}_{\beta}$  the solution of (3.1), we can introduce the following model for the selection of  $\beta$ , equivalent to require that the discrepancy corresponding to the selected minimizer is close to that of the true source image.

**Model 1.** Select the value of  $\beta$  such that

$$\mathcal{D}_{\mathbf{H}}(\mathbf{x}_\beta; \mathbf{g}\mathbf{n}) = \eta \quad (3.3)$$

where  $\mathbf{x}_\beta \geq 0$  is the minimizer of (3.1) and  $\eta$  is a given number close to one. If  $\bar{\beta}$  is a solution of the equation (3.3), then  $\bar{\mathbf{x}} = \mathbf{x}_{\bar{\beta}}$  is called solution of Model 1. We will also denote Model 1 as the *Crossing Model*

It has been remarked by several authors [33, 34, 92] that the approach of Model 1 can lead to highly time-consuming computations since it can require the solution of several minimization problems; therefore a *Constrained Model* has been proposed and formulated as follows.

**Model 2.** For a given  $\eta$ , solve the problem

$$\min_{\mathbf{x} \geq 0} \varphi_1(\mathbf{x}), \quad \text{subject to } \mathcal{D}_{\mathbf{H}}(\mathbf{x}; \mathbf{g}\mathbf{n}) \leq \eta. \quad (3.4)$$

Any solution  $\bar{\mathbf{x}}$  of this problem is called solution of Model 2. It is obvious that in such a case only one minimization is required.

We show the relations between the two models: using and extending results proved in [92], we prove that the two models have the same solutions for a suitable class of potential functions. Moreover we report and extend results, proved in [13], concerning Model 1.

In the following we denote by  $\text{lev}_\alpha f(t)$  the level set of a function  $f$ , i.e.  $\{t | f(t) \leq \alpha\}$ .

### 3.1.1 Relations Between Penalized and Constrained Convex Problems

For a unified treatment of the cases we are investigating it is convenient to set

$$\varphi_1(\mathbf{x}) = R(L\mathbf{x}), \quad (3.5)$$

where  $R(t)$  is a convex and nonnegative function and  $L$  is a suitable matrix. We recall that any seminorm on  $\mathbb{R}^n$  can be written as  $\|L \cdot\|$ , and therefore it can be referred to by the notation (3.5).

In order to have notations close to those used in [92], we set  $\tau = \eta \frac{N}{2}$  and  $\lambda = \frac{1}{\bar{\beta}}$ . Moreover we introduce the quantities

$$\begin{aligned} \tau_0 &= \min_{\mathbf{x} \geq 0} \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}), \\ \tau_L &= \min_{\substack{\mathbf{x} \geq 0 \\ \mathbf{x} \in \text{Ker}(L)}} \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}), \end{aligned} \quad (3.6)$$

where  $\text{Ker}(L)$  is the null space of the matrix  $L$  in (3.5).

We consider the following problems, strictly connected to the Models (3.3) and (3.4) respectively:

$$\min_{\mathbf{x} \geq 0} R(L\mathbf{x}) + \lambda \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}), \quad \lambda \geq 0, \quad (3.7)$$

and

$$\min_{\mathbf{x} \geq 0} R(L\mathbf{x}), \quad \text{subject to } \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) \leq \tau, \quad (3.8)$$

with  $\tau \geq \tau_0$ . The regularization functionals considered in the previous chapters, such as the Tikhonov regularization, the Total Variation and the Hypersurface potential, belong to the class of functional of type (3.5); in all the three cases, the function  $R$  satisfies a set of general assumptions, applicable to other regularization functions such as, for example,  $\|L\mathbf{x}\|_1$  or Markov Random Field regularizations. We denote as **Assumption REG** the following assumptions, also satisfied by any seminorm:

- $R : \mathbb{R}^p \rightarrow \mathbb{R}$  is a proper, convex, continuous function, such that  $R(t) \geq 0$  and  $R(t) = 0 \Leftrightarrow t = 0$ ;
- the level sets  $\text{lev}_\alpha R(t)$  of  $R(t)$  are nonempty and bounded for  $\alpha \geq 0$ ;
- $L \in \mathbb{R}^{p \times N}$ ,  $p \geq N$ .

Under **Assumption REG**, it is evident that  $\text{Ker}(L)$  is the set of minimum points of  $R(L\mathbf{x})$ .

Theoretical results about the relations between the convex problems (3.7) and (3.8) can be found in [64]; we point out that, when the minimizer of a general penalized problem  $\lambda F(\mathbf{x}) + G(\mathbf{x})$  is not unique,  $F(\mathbf{x})$  can assume different values at the minimum points; furthermore, it may happen that the multiplier of the constrained problem  $\min_{\mathbf{x} \in \text{lev}_\tau(F(\mathbf{x}))} G(\mathbf{x})$  is not unique [40].

In [92], the authors specialize these general results to the problems (3.7) and (3.8), devising the conditions that assure their equivalence when  $R(L\mathbf{x})$  is a seminorm and  $\mathbf{gn} > 0$ . The discussion can be easily extended to a differentiable  $R(L\mathbf{x})$  satisfying the **Assumption REG** (as, for example, Tikhonov-like regularization or Hypersurface potential). For sake of completeness, we report these results, requiring only **Assumption REG** for  $R(L\mathbf{x})$ .

**Proposition 3.1.** [92, Th. 3.2] *Let  $\mathbf{H} \in \mathbb{R}^{M \times N}$  be such that*

$$\mathcal{K} = \{\mathbf{x} \geq 0 : \mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1} > 0\} \neq \emptyset \quad (3.9)$$

and

$$\{\mathbf{x} \geq 0\} \not\subseteq \text{Ker}(\mathbf{H}). \quad (3.10)$$

We assume  $L \in \mathbb{R}^{p \times N}$  with  $\text{Ker}(L) \cap \text{Ker}(\mathbf{H}) = \{\mathbf{0}\}$  and  $R(L\mathbf{x})$  satisfying the **Assumption REG**. Then we have

i) the problems (3.7) with  $\lambda \geq 0$  and (3.8) with  $\tau \geq \tau_0$  have a solution;

ii) under the assumption  $\mathbf{gn} > 0$ , if  $\hat{\mathbf{x}}$  and  $\tilde{\mathbf{x}}$  are solution of (3.7) for a given  $\lambda > 0$ , then  $R(L\hat{\mathbf{x}}) = R(L\tilde{\mathbf{x}})$  and  $\mathbf{H}\hat{\mathbf{x}} = \mathbf{H}\tilde{\mathbf{x}}$ , i.e.  $\mathbf{KL}(\mathbf{H}\hat{\mathbf{x}} + b\mathbf{g}\mathbf{1}; \mathbf{gn}) = \mathbf{KL}(\mathbf{H}\tilde{\mathbf{x}} + b\mathbf{g}\mathbf{1}; \mathbf{gn})$ ;

iii) let

$$\arg \min_{\mathbf{x} \geq 0} \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{gn}) \cap \text{Ker}(L) = \emptyset$$

and  $\tau_0 < \tau < \tau_L$ ; under the assumption  $\mathbf{gn} > 0$ , if  $\hat{\mathbf{x}}$  and  $\tilde{\mathbf{x}}$  are solution of (3.8) for a given  $\tau$ , then  $R(L\hat{\mathbf{x}}) = R(L\tilde{\mathbf{x}})$  and  $\mathbf{H}\hat{\mathbf{x}} = \mathbf{H}\tilde{\mathbf{x}}$ , with  $\mathbf{KL}(\mathbf{H}\hat{\mathbf{x}} + b\mathbf{g}\mathbf{1}; \mathbf{gn}) = \mathbf{KL}(\mathbf{H}\tilde{\mathbf{x}} + b\mathbf{g}\mathbf{1}; \mathbf{gn}) = \tau$ .

In presence of nonzero background, the assumption (3.10) allows to exclude the trivial case when the nonnegative orthant is a subset of  $\text{Ker}(\mathbf{H})$ . Obviously, when  $\mathbf{H}$  satisfies the conditions (1.5), the assumptions (3.9) and (3.10) hold. Under the assumptions of Proposition 3.1, using the same arguments of the proof of Lemma 3.3 in [92], it is possible to state that each solution  $\bar{\mathbf{x}}$  of (3.7) does not belong to  $\text{Ker}(L)$  and the following relation with  $\bar{\lambda}$  holds :

$$\bar{\lambda} = \frac{\langle \mathbf{p}_1, L\bar{\mathbf{x}} \rangle}{\left\langle \frac{\mathbf{gn}}{\mathbf{H}\bar{\mathbf{x}} + bg\mathbf{1}} - \mathbf{1}, \mathbf{H}\bar{\mathbf{x}} \right\rangle} \quad (3.11)$$

where  $\mathbf{p}_1 \in \partial R(L\bar{\mathbf{x}})$ ,

When  $R(L\mathbf{x})$  is a seminorm, i.e.  $R(L\mathbf{x}) = \|L\mathbf{x}\|$ , since  $\bar{\mathbf{x}} \notin \text{Ker}(L)$  and  $\|\mathbf{p}_1\|_* = 1$  (where  $\|\cdot\|_*$  is the dual norm of  $\|\cdot\|$ ), we have  $\langle \mathbf{p}_1, L\bar{\mathbf{x}} \rangle = \|L\bar{\mathbf{x}}\| > 0$  and, consequently,

$$\bar{\lambda} = \frac{R(L\bar{\mathbf{x}})}{\left\langle \frac{\mathbf{gn}}{\mathbf{H}\bar{\mathbf{x}} + bg\mathbf{1}} - \mathbf{1}, \mathbf{H}\bar{\mathbf{x}} \right\rangle} \quad (3.12)$$

Then, for part ii) of Proposition 3.1, in the case of the seminorm regularization,  $\bar{\lambda}$  is uniquely determined.

When  $R(L\bar{\mathbf{x}})$  is differentiable,  $\mathbf{p}_1 = \nabla_t R(t)|_{t=L\bar{\mathbf{x}}}$  and the equation (3.11) can be written as:

$$\bar{\lambda} = \frac{\langle \nabla_t R(t)|_{t=L\bar{\mathbf{x}}}, L\bar{\mathbf{x}} \rangle}{\left\langle \frac{\mathbf{gn}}{\mathbf{H}\bar{\mathbf{x}} + bg\mathbf{1}} - \mathbf{1}, \mathbf{H}\bar{\mathbf{x}} \right\rangle} \quad (3.13)$$

Then, as in the previous case, for differentiable  $R(t)$  satisfying the **Assumption REG**,  $\bar{\lambda}$  is uniquely determined.

In particular, for Tikhonov-like regularization,

$$\langle \nabla_t R(t)|_{t=L\bar{\mathbf{x}}}, L\bar{\mathbf{x}} \rangle = \langle \bar{\mathbf{x}}, \nabla_{\mathbf{x}} R(L\bar{\mathbf{x}}) \rangle = \|L\bar{\mathbf{x}}\|^2$$

For HS regularization, taking into account the expression of the gradient [19], we have

$$\langle \nabla_t R(t)|_{t=L\bar{\mathbf{x}}}, L\bar{\mathbf{x}} \rangle = \sum_{i=1}^N \frac{\bar{\mathbf{x}}^t L_i^t L_i \bar{\mathbf{x}}}{\sqrt{\|L_i \bar{\mathbf{x}}\|^2 + \delta^2}}$$

Using part iii) of Proposition 3.1 and the above relation between the value of  $\bar{\lambda}$  and the corresponding solutions of problem (3.7) the following theorem states the relationship between problems (3.7) and (3.8).

**Proposition 3.2.** [92, Theorem 3.4] *Let  $\mathbf{H} \in \mathbb{R}^{M \times N}$  be such that (3.9)-(3.10) hold and  $L \in \mathbb{R}^{p \times N}$  with  $\text{Ker}(L) \cap \text{Ker}(\mathbf{H}) = \{\mathbf{0}\}$ . Let*

$$\arg \min_{\mathbf{x} \geq 0} \mathbf{KL}(\mathbf{H}\mathbf{x} + bg\mathbf{1}; \mathbf{gn}) \cap \text{Ker}(L) = \emptyset$$

and  $\tau_0 < \tau < \tau_L$ . When  $R(L\mathbf{x})$  is a seminorm regularization or a differentiable function satisfying the **Assumption REG**, if  $\bar{\mathbf{x}}$  is a solution of (3.8), then there exists a unique  $\bar{\lambda} > 0$  such that  $\bar{\mathbf{x}}$  is a solution of (3.7).

In conclusion Proposition 3.2, combined with part ii) and part iii) of Proposition 3.1, states that Model 2 has solutions and that they are all the solutions of Model 1.

For suitable differentiable regularizations, the previous results can be established under assumptions weaker than those of the previous Proposition 3.2: more precisely one can remove the assumption that all the values of  $\mathbf{gn}$  are strictly positive. Before introducing these results, we recall the conditions for the existence and uniqueness of the solution of the problem (3.7) for  $\lambda > 0$  in some special cases and we extend part iii) of Proposition 3.1.

Let  $I_1 = \{i \mid \mathbf{gn}_i > 0\}$  and  $I_2 = \{1, 2, \dots, M\} - I_1$ . The cardinality of  $I_1$  is denoted by  $M_1$ . We denote by  $\mathbf{gn}_{I_1}$  the vector of nonzero entries of  $\mathbf{gn}$  and by  $\mathbf{H}_{I_1}$  and  $\mathbf{H}_{I_2}$  the submatrices of  $\mathbf{H}$  given by the rows with indexes in  $I_1$  and  $I_2$ , respectively. The following Lemma is a generalization of Proposition (1.3) cases (a)–(c).

**Lemma 3.1.** *Let  $\mathbf{H} \in \mathbb{R}^{M \times N}$  be such that the assumptions (3.9)–(3.10) hold and  $L \in \mathbb{R}^{p \times N}$ . Let  $\text{Ker}(L) \cap \text{Ker}(\mathbf{H}_{I_1}) = \{\mathbf{0}\}$ ,  $I_1 \neq \emptyset$ . If  $\text{R}(L\mathbf{x})$  satisfies the **Assumption REG** and  $\text{R}(t)$  is a differentiable function with positive definite Hessian, then the problem (3.7) for  $\lambda > 0$  has a unique solution.*

*Proof.* Under the assumption

$$\text{Ker}(L) \cap \text{Ker}(\mathbf{H}_{I_1}) = \{\mathbf{0}\}$$

$I_1 \neq \emptyset$  and  $\lambda > 0$ , we prove that the objective function of the problem (3.7) is strictly convex by showing that the intersection between the null spaces of the Hessian of  $\mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{gn})$  and  $\text{R}(Lx)$ , i.e. the null space of the Hessian of the objective function, is trivial.

Indeed we have that  $\text{Ker}(\nabla^2 \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{gn})) = \text{Ker}(\mathbf{H}_{I_1})$ . On the other hand, the Hessian matrix of  $\text{R}(Lx)$  is given by  $L^t \nabla^2 \text{R}(Lx) L$ . Since  $\nabla^2 \text{R}(t)$  is positive definite for any  $t$ ,  $\text{Ker}(\nabla^2 \text{R}(Lx)) = \text{Ker}(L) = \{\mathbf{x} = \arg \min_{\mathbf{x}} \text{R}(Lx)\}$ . As a consequence of the assumption  $\text{Ker}(L) \cap \text{Ker}(\mathbf{H}_{I_1}) = \{\mathbf{0}\}$ , the objective function is strictly convex and the minimum point is unique.  $\square$

Lemma 3.1 holds for Tikhonov-like regularization, HS potential and Markov random field functions.

Under the assumptions of Lemma 3.1, the following relation between  $\lambda$  and the unique solution  $\hat{\mathbf{x}}$  of the problem (3.7) holds:

$$\lambda = \frac{\langle \nabla_t \text{R}(t)|_{t=L\hat{\mathbf{x}}}, L\hat{\mathbf{x}} \rangle}{\left\langle \frac{\mathbf{gn}_{I_1}}{\mathbf{H}_{I_1}\hat{\mathbf{x}} + b\mathbf{g}\mathbf{1}}, \mathbf{H}_{I_1}\hat{\mathbf{x}} \right\rangle - \langle \mathbf{1}, \mathbf{H}\hat{\mathbf{x}} \rangle}, \quad (3.14)$$

Furthermore, part iii) of Proposition 3.1 can be restated under more general assumptions.

**Proposition 3.3.** *Let  $\mathbf{H} \in \mathbb{R}^{M \times N}$  be such that the assumptions (3.9)–(3.10) hold,  $L \in \mathbb{R}^{p \times N}$  and let  $\text{R}(Lx)$  satisfy **Assumption REG**. Let*

$$\arg \min_{\mathbf{x} \geq 0} \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{gn}) \cap \text{Ker}(L) = \emptyset$$

and  $\tau_0 < \tau < \tau_L$ ; under the assumption

$$\text{Ker}(L) \cap \text{Ker}(\mathbf{H}_{I_1}) = \{\mathbf{0}\}, \quad I_1 \neq \emptyset$$

if  $\hat{\mathbf{x}}$  and  $\tilde{\mathbf{x}}$  are solution of (3.8), then  $R(L\hat{\mathbf{x}}) = R(L\tilde{\mathbf{x}})$  and

$$\mathbf{KL}(\mathbf{H}\hat{\mathbf{x}} + bg\mathbf{1}; \mathbf{g}\mathbf{n}) = \mathbf{KL}(\mathbf{H}\tilde{\mathbf{x}} + bg\mathbf{1}; \mathbf{g}\mathbf{n}) = \tau$$

with  $\mathbf{H}_{I_1}\hat{\mathbf{x}} = \mathbf{H}_{I_1}\tilde{\mathbf{x}}$  and  $\langle \mathbf{1}_{M-M_1}, \mathbf{H}_{I_2}\tilde{\mathbf{x}} \rangle = \langle \mathbf{1}_{M-M_1}, \mathbf{H}_{I_2}\hat{\mathbf{x}} \rangle$ .

*Proof.* When  $\hat{\mathbf{x}}$  and  $\tilde{\mathbf{x}}$  are two solutions of problem (3.8), it is evident that  $R(L\hat{\mathbf{x}}) = R(L\tilde{\mathbf{x}})$ . We have to prove that  $\mathbf{KL}(\mathbf{H}\hat{\mathbf{x}} + bg\mathbf{1}; \mathbf{g}\mathbf{n}) = \tau$  for any solution  $\hat{\mathbf{x}}$  of the problem. Since  $\hat{\mathbf{x}}$  is a solution of (3.8),  $\text{lev}_\tau(H) \cap i_{x \geq 0} \neq \emptyset$  and  $R(Lx)$  is continuous on its domain  $\mathbb{R}^n$ , there exists  $\mathbf{w} \in L^t \partial R(L\hat{\mathbf{x}})$  (or  $\mathbf{w} = L^t \nabla R(L\hat{\mathbf{x}})$  if  $R(Lx)$  is differentiable) and a nonnegative scalar  $\mu$  [84, §28] such that

$$\begin{aligned} (\mathbf{w} + \mu \nabla \mathbf{KL}(\mathbf{H}\hat{\mathbf{x}} + bg\mathbf{1}; \mathbf{g}\mathbf{n}))\hat{\mathbf{x}} &= 0 \\ \mathbf{w} + \mu \nabla \mathbf{KL}(\mathbf{H}\hat{\mathbf{x}} + bg\mathbf{1}; \mathbf{g}\mathbf{n}) &\geq 0 \\ \hat{\mathbf{x}} &\geq 0 \\ \mu(\mathbf{KL}(\mathbf{H}\hat{\mathbf{x}} + bg\mathbf{1}; \mathbf{g}\mathbf{n}) - \tau) &= 0 \\ \mathbf{KL}(\mathbf{H}\hat{\mathbf{x}} + bg\mathbf{1}; \mathbf{g}\mathbf{n}) &\leq \tau \\ \mu &\geq 0 \end{aligned} \tag{3.15}$$

Let  $\mathbf{w} = L^t \mathbf{p}$  for a suitable  $\mathbf{p} \in \partial R(t)|_{t=L\hat{\mathbf{x}}}$  (or  $\mathbf{p} = \nabla_t R(t)|_{t=L\hat{\mathbf{x}}}$ ). Since  $\tau < \tau_L$ ,  $\hat{\mathbf{x}} \notin \text{Ker}(L)$ , and, consequently,  $\mathbf{w}$  is different from  $\mathbf{0}$  (indeed  $\arg \min R(Lx) = \text{Ker}(L)$ ). Then, if  $\mathbf{KL}(\mathbf{H}\hat{\mathbf{x}} + bg\mathbf{1}; \mathbf{g}\mathbf{n}) < \tau$ , since  $\mu(\mathbf{KL}(\mathbf{H}\hat{\mathbf{x}} + bg\mathbf{1}; \mathbf{g}\mathbf{n}) - \tau) = 0$ , it follows  $\mu = 0$ ; therefore  $\langle \mathbf{w}, \hat{\mathbf{x}} \rangle = 0$  and the problem is reduced to the minimum of  $R(L\mathbf{x})$  on the nonnegative orthant. Then  $\mathbf{x} = \mathbf{0}$  is a solution as well as  $\hat{\mathbf{x}}$ . But  $\mathbf{0} \in \text{Ker}(L)$  and  $\mathbf{0} \notin \text{lev}_\tau \mathbf{KL}(\mathbf{H}\mathbf{x} + bg\mathbf{1}; \mathbf{g}\mathbf{n})$ . Then  $\mathbf{KL}(\mathbf{H}\hat{\mathbf{x}} + bg\mathbf{1}; \mathbf{g}\mathbf{n}) = \tau$ .

Now, we prove that, if  $\hat{\mathbf{x}}$  and  $\tilde{\mathbf{x}}$  are solutions of (3.8), then  $\mathbf{H}_{I_1}\hat{\mathbf{x}} = \mathbf{H}_{I_1}\tilde{\mathbf{x}}$ ; we write  $\hat{\mathbf{x}} = \hat{\mathbf{x}}_1 + \hat{\mathbf{x}}_2$  and  $\tilde{\mathbf{x}} = \tilde{\mathbf{x}}_1 + \tilde{\mathbf{x}}_2$  with  $\hat{\mathbf{x}}_1, \tilde{\mathbf{x}}_1 \in \mathcal{R}(\mathbf{H}_{I_1}^t)$ ,  $\hat{\mathbf{x}}_1 \neq \tilde{\mathbf{x}}_1$  and  $\hat{\mathbf{x}}_2, \tilde{\mathbf{x}}_2 \in \text{Ker}(\mathbf{H}_{I_1})$ . Let  $\mathbf{x} = \mu\hat{\mathbf{x}} + (1 - \mu)\tilde{\mathbf{x}}$ , where  $\mu \in (0, 1)$  is chosen so that  $\mathbf{x} \geq 0$ . Then

$$\begin{aligned} \mathbf{KL}(\mathbf{H}\mathbf{x} + bg\mathbf{1}; \mathbf{g}\mathbf{n}) &= \mathbf{KL}(\mathbf{H}_{I_1}(\mu\hat{\mathbf{x}} + (1 - \mu)\tilde{\mathbf{x}}) + bg\mathbf{1}; \mathbf{g}\mathbf{n}_{I_1}) + \\ &\quad + \langle \mathbf{1}_{M-M_1}, \mathbf{H}_{I_2}(\mu\hat{\mathbf{x}} + (1 - \mu)\tilde{\mathbf{x}}) + bg\mathbf{1} \rangle = \\ &= \mathbf{KL}(\mathbf{H}_{I_1}(\mu\hat{\mathbf{x}}_1 + (1 - \mu)\tilde{\mathbf{x}}_1) + bg\mathbf{1}; \mathbf{g}\mathbf{n}_{I_1}) + \\ &\quad + \mu \langle \mathbf{1}_{M-M_1}, \mathbf{H}_{I_2}\hat{\mathbf{x}} + bg\mathbf{1} \rangle + \\ &\quad + (1 - \mu) \langle \mathbf{1}_{M-M_1}, \mathbf{H}_{I_2}\tilde{\mathbf{x}} + bg\mathbf{1} \rangle < \\ &< \mu(\mathbf{KL}(\mathbf{H}_{I_1}\hat{\mathbf{x}}_1 + bg\mathbf{1}; \mathbf{g}\mathbf{n}_{I_1}) + \langle \mathbf{1}_{M-M_1}, \mathbf{H}_{I_2}\hat{\mathbf{x}} \rangle) + \\ &\quad + (1 - \mu)(\mathbf{KL}(\mathbf{H}_{I_1}\tilde{\mathbf{x}}_1 + bg\mathbf{1}; \mathbf{g}\mathbf{n}_{I_1}) + \\ &\quad + \langle \mathbf{1}_{M-M_1}, \mathbf{H}_{I_2}\tilde{\mathbf{x}} + bg\mathbf{1} \rangle) \\ &= \mu\tau + (1 - \mu)\tau = \tau \end{aligned} \tag{3.16}$$

where the strict inequality follows from the strict convexity of  $\mathbf{KL}(\mathbf{H}_{I_1}\mathbf{x} + bg\mathbf{1}; \mathbf{g}\mathbf{n}_{I_1})$  on  $\mathcal{R}(\mathbf{H}_{I_1}^t)$ . Then we have  $\mathbf{KL}(\mathbf{H}\mathbf{x} + bg\mathbf{1}; \mathbf{g}\mathbf{n}) < \tau$ . On the other hand, we obtain

$$\begin{aligned} R(L\mathbf{x}) &= R(\mu L\hat{\mathbf{x}} + (1 - \mu)L\tilde{\mathbf{x}}) \leq \\ &\leq \mu R(L\hat{\mathbf{x}}) + (1 - \mu)R(L\tilde{\mathbf{x}}) = \\ &= R(L\tilde{\mathbf{x}}) = R(L\hat{\mathbf{x}}) \end{aligned} \tag{3.17}$$

so that  $\mathbf{x}$  should be a minimizer of (3.8). This is impossible since any minimizer has to fulfill  $\mathbf{KL}(\mathbf{H}\mathbf{x} + bg\mathbf{1}; \mathbf{g}\mathbf{n}) = \tau$ .

Finally, since  $\mathbf{KL}(\mathbf{H}\hat{\mathbf{x}} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) = \mathbf{KL}(\mathbf{H}\tilde{\mathbf{x}} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) = \tau$  and  $\mathbf{H}_{I_1}\hat{\mathbf{x}} = \mathbf{H}_{I_1}\tilde{\mathbf{x}}$ , it follows

$$\langle \mathbf{1}_{M-M_1}, \mathbf{H}_{I_2}\tilde{\mathbf{x}} \rangle = \langle \mathbf{1}_{M-M_1}, \mathbf{H}_{I_2}\hat{\mathbf{x}} \rangle \quad (3.18)$$

□

Following the same arguments used in [92] for proving Proposition 3.2, Proposition 3.3 enables to extend the results of Proposition 3.2 to the case  $\mathbf{g}\mathbf{n} \geq 0$  for differentiable regularization functions satisfying Lemma 3.1.

**Proposition 3.4.** *Let  $\mathbf{H} \in \mathbb{R}^{M \times N}$  be such that the assumptions (3.9)-(3.10) hold and  $L \in \mathbb{R}^{p \times N}$ . Let*

$$\text{Ker}(L) \cap \text{Ker}(\mathbf{H}_{I_1}) = \{\mathbf{0}\}$$

$I_1 \neq \emptyset$ . Let  $\arg \min_{\mathbf{x} \geq 0} \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) \cap \text{Ker}(L) = \emptyset$ , and  $\tau_0 < \tau < \tau_L$ . Assume that  $\text{R}(L\mathbf{x})$  satisfies **Assumption REG** and  $\text{R}(t)$  is a differentiable function with positive definite Hessian. If  $\bar{\mathbf{x}}$  is a solution of (3.8), then there exists a unique  $\bar{\lambda} > 0$  such that  $\bar{\mathbf{x}}$  is a solution of (3.7).

## 3.2 Existence and uniqueness of the solution of the Discrepancy equation

In [13], conditions for the existence and uniqueness of the solution of the discrepancy equation (3.3) are devised for denoising and deblurring of images corrupted by Poisson noise, when  $\varphi_1(\mathbf{x}) \equiv \text{R}(L\mathbf{x})$  is the Tikhonov-like regularization or the HS potential.

Using the previous results, the discrepancy equation can be used also when  $\text{R}(L\mathbf{x})$  is a seminorm regularization. We recall that  $\mathbf{x}_\beta$  denotes a solution of the penalized problem (3.1). The following proposition states the conditions assuring the well-definiteness of the function  $\mathcal{D}_{\mathbf{H}}(\mathbf{x}; \mathbf{g}\mathbf{n})$  for  $\mathbf{g}\mathbf{n} > 0$ .

**Proposition 3.5.** *Let  $\mathbf{H} \in \mathbb{R}^{M \times N}$  be such that (3.9)-(3.10) hold and  $L \in \mathbb{R}^{p \times N}$  with  $\text{Ker}(L) \cap \text{Ker}(H) = \{\mathbf{0}\}$ . When  $\mathbf{g}\mathbf{n} > 0$  and  $\text{R}(L\mathbf{x})$  satisfies the **Assumption REG**,  $\mathcal{D}_{\mathbf{H}}(\mathbf{x}_\beta; \mathbf{g}\mathbf{n})$  is well defined for  $\beta > 0$ .*

*Proof.* Given  $\beta > 0$ , Proposition 3.1 assures that the problem (3.7) with  $\lambda = \frac{1}{\beta}$  has at least a solution and, if we consider two different solutions of the problem  $\hat{\mathbf{x}}$  and  $\tilde{\mathbf{x}}$ , we have  $\mathbf{KL}(\mathbf{H}\hat{\mathbf{x}} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) = \mathbf{KL}(\mathbf{H}\tilde{\mathbf{x}} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n})$ . Therefore for all  $\beta > 0$ ,  $\mathcal{D}_{\mathbf{H}}(\mathbf{x}_\beta; \mathbf{g}\mathbf{n})$  is well defined. □

Finally, Proposition 3.6 establishes the uniqueness of the solution of Model 1 for seminorm regularization (i) and restates the results already obtained in [13] (ii).

**Proposition 3.6.** *Let  $\mathbf{H} \in \mathbb{R}^{N \times M}$  be such that (3.9)-(3.10) hold and  $L \in \mathbb{R}^{p \times N}$  with  $\text{Ker}(L) \cap \text{Ker}(H) = \{\mathbf{0}\}$ . Assume  $\tau = \eta \frac{N}{2}$  such that  $\tau_0 < \tau < \tau_L$ , i.e.*

$$\arg \min_{\mathbf{x} \geq 0} \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) \cap \text{Ker}(L) = \emptyset$$

Thus we have:

- i) under the assumption  $\mathbf{g}\mathbf{n} > 0$ , the solution  $\bar{\beta}$  of the discrepancy equation combined with a seminorm regularization exists and is unique;

ii) let  $R(L\mathbf{x})$  be a differentiable regularization term satisfying **Assumption REG**, such that the Hessian of  $R(t)$  is positive definite; under the assumption  $\text{Ker}(L) \cap \text{Ker}(\mathbf{H}_{I_1}) = \{\mathbf{0}\}$ ,  $I_1 \neq \emptyset$ , the solution  $\bar{\beta}$  of the discrepancy equation combined with the regularization term  $R(L\mathbf{x})$  exists and is unique; furthermore the vector  $\mathbf{x}_{\bar{\beta}}$  corresponding to the solution is unique.

*Proof.* i) We consider the problem (3.8) with  $\tau = \eta \frac{N}{2}$  and  $R(L\mathbf{x})$  given by a seminorm. For part iii) of Proposition 3.1, the solution  $\bar{\mathbf{x}}$  of this problem exists and  $\mathbf{KL}(\mathbf{H}\bar{\mathbf{x}} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) = \tau$ . For Proposition 3.2, there exists a unique  $\bar{\lambda} > 0$  such that  $\bar{\mathbf{x}}$  is a solution of (3.7) with  $\bar{\lambda} = 1/\bar{\beta}$  and this value does not depend on  $\bar{\mathbf{x}}$ . Then  $\bar{\beta} = 1/\bar{\lambda}$  is the unique solution of the discrepancy equation combined with the seminorm regularization.

ii) We consider the problem (3.8) with  $\tau = \eta \frac{N}{2}$  and  $R(L\mathbf{x})$  satisfying **Assumption REG**. For Proposition 3.3, the solution  $\bar{\mathbf{x}}$  of this problem exists and  $\mathbf{KL}(\mathbf{H}\bar{\mathbf{x}} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) = \tau$ . Since  $R(t)$  is a differentiable function with positive definite Hessian, Proposition 3.4 assures that there exists a unique  $\bar{\lambda} > 0$  such that  $\bar{\mathbf{x}}$  is a solution of (3.7) with  $\bar{\lambda} = 1/\bar{\beta}$  and this value does not depend on  $\bar{\mathbf{x}}$ . Then  $\bar{\beta} = 1/\bar{\lambda}$  is the unique solution of the discrepancy equation. Furthermore, from Lemma 3.1, the solution  $\mathbf{x}_{\bar{\beta}}$  of (3.7) for a given  $\bar{\beta}$  is unique.  $\square$

In [13], under the assumption of part ii) of the previous Proposition, for differentiable function with positive definite Hessian  $R(t)$  and under the assumption (1.5) on  $\mathbf{H}$ , the authors give the conditions that assure  $\tau_0 < \eta \frac{N}{2} < \tau_L$  for HS and Tikhonov-like regularization. For example, for image deconvolution ( $N = M$  and  $\mathbf{H}\mathbf{1} = \mathbf{1}$ ), the condition  $\frac{M}{2} < \tau_L$  is satisfied if

$$\frac{1}{N} \sum_{i \in I_1} \mathbf{g}\mathbf{n}_i \log(\mathbf{g}\mathbf{n}_i) > \frac{1}{2} + \overline{\mathbf{g}\mathbf{n}} \log(\overline{\mathbf{g}\mathbf{n}}), \quad (3.19)$$

where  $\overline{\mathbf{g}\mathbf{n}} = \frac{1}{N} \sum_{i \in I_1} \mathbf{g}\mathbf{n}_i$ .

### 3.3 Numerical Methods

In this section, we introduce the numerical methods enabling us to exploit the two considered models. For the *Crossing Model* we propose an approach based on the Dai-Fletcher algorithm [45], while for the *Constrained* one we use the ADMM method described in Section 2.4.

#### 3.3.1 Model 1, *Crossing Model*

The solution  $\bar{\beta}$  of the discrepancy equation (3.3) can be approximated by solving the root finding problem

$$\mathcal{F}(\beta) = \mathcal{D}_{\mathbf{H}}(\mathbf{x}_{\beta}; \mathbf{g}\mathbf{n}) - \eta = 0, \quad (3.20)$$

where  $\mathcal{F}(\beta)$  is a continuous increasing function. We solve this problem by a specialized version of the algorithm proposed in [45], called Modified Dai-Fletcher (MDF) method. The root finding solver [45] consists in two phases: a bracketing phase to determine

the extremes  $0 < \beta_l < \beta_u$  of an interval containing the root and a secant phase for the root approximation. The implementation of the bracketing phase has been changed, according to the special regularization framework we are considering. In particular, we have taken into account that the evaluation of the discrepancy equation in  $\beta$  is generally less expensive for large values of  $\beta$ , due to the special form of the penalized problem providing  $\mathbf{x}_\beta$ . For this reason, when we start from  $\beta$  such that  $\mathcal{F}(\beta) < 0$ , we look for  $\beta_l$  and  $\beta_u$  by means of a secant-like approach, allowing steps in the interval  $[d\beta, 10d\beta]$ , where  $d\beta$  denotes the previous step; on the other hand, when we move from  $\beta$  such that  $\mathcal{F}(\beta) > 0$ , we reduce the tentative  $\beta$  by a constant factor  $\omega \in (0, 1)$ , for preserving the positivity of  $\beta$  and for avoiding the evaluation of the discrepancy function in too small values. The bracketing phase is sketched as in Algorithm 10.

---

**Algorithm 10: Bracketing Phase of Algorithm MDF**


---

```

Choose an initial value  $\beta > 0$ , an initial step  $d\beta > 0$  and  $\omega \in (0, 1)$ .
IF  $\mathcal{F}(\beta) < 0$ 
     $\beta_l = \beta, \quad \beta = \beta + d\beta$ 
    WHILE  $\mathcal{F}(\beta) < 0$ 
         $s = \min(1, \max((\mathcal{F}(\beta_l) - \mathcal{F}(\beta)) / \mathcal{F}(\beta), 0.1))$ ,
         $\beta_l = \beta, \quad d\beta = d\beta/s, \quad \beta = \beta + d\beta$ ,
     $\beta_u = \beta$ 
ELSE
     $\beta_u = \beta, \quad \beta = \beta_u\omega$ 
    WHILE  $\mathcal{F}(\beta) > 0$ 
         $\beta_u = \beta, \quad \beta = \beta_u\omega$ 
     $\beta_l = \beta$ 
ENDIF  $\beta_l, \beta_u$  are such that  $\bar{\beta} \in [\beta_l, \beta_u]$ .

```

---

The second phase of the MDF method is essentially the secant-based strategy described in [45]: starting from the values  $\beta_l$  and  $\beta_u$  provided by the bracketing phase, the root  $\bar{\beta}$  is approximated by exploiting standard secant steps or modified steps designed to accelerate the convergence remote from the solution. For completeness, we report the main steps of the MDF secant phase in Algorithm 11. The stopping rule used in the MDF secant phase is

$$\begin{aligned}
 |\mathcal{F}(\beta_k)| &\leq \varepsilon_1 \quad \text{or} \\
 (|\beta_k - \beta_{k-1}| &\leq \varepsilon_2\beta_k \quad \text{and} \quad |\mathcal{F}(\beta_k)| \leq 10\varepsilon_1),
 \end{aligned} \tag{3.21}$$

where  $\beta_k$  denotes the value of  $\beta$  at the  $k$ -th iteration of Algorithm 11 and  $\varepsilon_1, \varepsilon_2$  are two small positive constants. From a practical point of view, the MDF performance strictly depends on the effectiveness of the minimization method used for obtaining the solution  $\mathbf{x}_\beta$  of the penalized problem (3.1). In our experience, two suitable solvers for (3.1) are the Scaled Gradient Projection method, for the case of differentiable objective functional, and the PIDSplit+ algorithm.

### 3.3.2 Model 2, *Constrained Model*

In [92], the authors propose to estimate the regularization parameter by solving the constrained problem (3.8) with the well-known Alternating Direction Method of Mul-

**Algorithm 11: Secant Phase of Algorithm MDF**


---

Set  $\beta_l, \beta_u$  such that  $\bar{\beta} \in [\beta_l, \beta_u]$ .  
 $s = (\mathcal{F}(\beta_u) - \mathcal{F}(\beta_l)) / \mathcal{F}(\beta_u)$ ,  $d\beta = (\beta_u - \beta_l) / s$ ,  $\beta = \beta_u - d\beta$   
WHILE (stopping rule is not satisfied)  
  IF  $\mathcal{F}(\beta) > 0$   
    IF  $s \leq 2$   
       $\beta_u = \beta$ ,  $s = (\mathcal{F}(\beta_u) - \mathcal{F}(\beta_l)) / \mathcal{F}(\beta_u)$   
       $d\beta = (\beta_u - \beta_l) / s$ ,  $\beta = \beta_u - d\beta$   
    ELSE  
       $s = \max((\mathcal{F}(\beta_u) - \mathcal{F}(\beta)) / \mathcal{F}(\beta), 0.1)$   
       $d\beta = (\beta_u - \beta) / s$ ,  $\beta_u = \beta$   
       $\beta = \max(\beta - d\beta, 0.75\beta_l + 0.25\beta)$   
       $s = (\beta_u - \beta_l) / (\beta_u - \beta)$   
  ELSE  
    IF  $s \geq 2$   
       $\beta_l = \beta$ ,  $s = (\mathcal{F}(\beta_u) - \mathcal{F}(\beta_l)) / \mathcal{F}(\beta_u)$   
       $d\beta = (\beta_u - \beta_l) / s$ ,  $\beta = \beta_u - d\beta$   
    ELSE  
       $s = \max((\mathcal{F}(\beta_l) - \mathcal{F}(\beta)) / \mathcal{F}(\beta), 0.1)$   
       $d\beta = (\beta - \beta_l) / s$ ,  $\beta_l = \beta$   
       $\beta = \min(\beta + d\beta, 0.75\beta_u + 0.25\beta)$   
       $s = (\beta_u - \beta_l) / (\beta_u - \beta)$   
 $\beta$  is an approximation of  $\bar{\beta}$

---

multipliers (ADMM). The problem (3.8) is equivalent to

$$\begin{aligned} \min_{\mathbf{x}} f_2(\mathbf{w}) + \frac{1}{2\gamma} \|M\mathbf{x} - \mathbf{w}\|^2, \\ \text{subject to } M\mathbf{x} = \mathbf{w}, \end{aligned} \quad (3.22)$$

where  $\gamma$  is a positive parameter,  $f_2(\mathbf{w}) = i_{\text{ev}, \tau} \mathbf{KL}(\mathbf{H}\mathbf{w}_{(1)} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + i_{\mathbf{w}_{(3)} \geq 0} + \mathbf{R}(\mathbf{w}_{(2)})$ , with similar notations introduced in Section 2.4 and with the linear constraints depending on the choice of the regularization function. As already explained in Section 2.4, the basic idea of ADMM method is to compute a saddle-point of the Augmented Lagrangian of the problem (3.22), decoupling the minimization step of the Augmented Lagrangian method in a sequence of minimizations with respect to the different sets of variables  $\mathbf{x}$  and  $\mathbf{w}_{(k)}$ . The application of ADMM for the constrained problem (3.8) is analyzed in [92] for a seminorm regularization: the authors describe the algorithm in detail and prove the convergence of the sequence of multipliers obtained for the inner subproblem in  $\mathbf{w}_{(1)}$  to the multiplier  $\lambda$  of  $\mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}) \leq \tau$ . In view of the Propositions of the previous section, these results can be easily extended to suitable differentiable regularization functions. It is well known that the performance of ADMM can be strongly dependent on the choice of the parameter  $\gamma$ . In order to reduce this drawback, in [63, 98] strategies for adaptively adjusting  $\gamma$  at any iteration have been proposed. The convergence of this adaptive version of ADMM is proved, assuming that  $\gamma$  becomes fixed after a finite number of iterations. Then the choice of the updating strategy is based on the numerical effectiveness. The procedure proposed in [63] is based on the

control of the  $\ell_2$  norm of two vectors, the so-called primal and dual residual:

$$\begin{aligned}\mathbf{r}^i &= M\mathbf{x}^i - \mathbf{w}^i, \\ \mathbf{s}^i &= \frac{1}{\gamma}M^t\sigma^i,\end{aligned}\tag{3.23}$$

with  $\sigma^i = \mathbf{w}^i - \mathbf{w}^{i-1}$ . Indeed, the vectors  $\mathbf{r}^i$  and  $\sigma^i$  are crucial components of the upper bound of the absolute error between the objective function at the current iterate and its minimum value. We use the following updating procedure:

$$\gamma^{i+1} = \begin{cases} \frac{\alpha}{\gamma^i} & \text{if } \|\mathbf{r}^i\| > \mu\|\sigma^i\| \text{ and } i \leq k_{max} \\ \alpha\gamma^i & \text{if } \|\sigma^i\| > \mu\|\mathbf{r}^i\| \text{ and } i \leq k_{max} \\ \gamma^i & \text{otherwise} \end{cases}\tag{3.24}$$

where  $\alpha$  and  $\mu$  are positive values greater than 1 and  $\gamma^0$  is a prefixed positive value. From the practical point of view, this adaptive version of ADMM appears less dependent on the parameter settings than the standard ADMM approach.

Finally, we mention that other methods have been adapted in [92] for the numerical solution of the constrained problem, as Arrow–Hurwitz method [2] or its extrapolated version [37]. As in the case of ADMM, the performance of these algorithms depends on a suitable selection of two prefixed parameters.

In [102] it is reported a wide computational study about the efficiency of the numerical procedures involved in the two approaches. For each model, state of the art computational techniques have been used. Two different regularizations have been considered: a differentiable one (HS) and a non-differentiable one (TV). In the first case we show that MDF, combined with an efficient solver for the inner optimization problems, enables Model 1 to be less time demanding than Model 2. On the other hand, in the case of TV regularization, Model 1 can be competitive with Model 2 but the latter is definitely more efficient in some cases. These results are in according to those reported in Section 5.3.

It should be stressed, however, that for images with low counts or for special applications (such as High Dynamic Range images in Astronomy), the described models do not always achieve significant results.



## Chapter 4

# Bregman Procedure

In this chapter we show the main features of the Bregman Procedure, whose original idea lies in [25]. In [79] the idea was applied to restoration problems in presence of Gaussian noise; recently in [28, 29] this method was extended to Poisson data. In this chapter, we recall the results previously mentioned. Moreover, we introduce an *inexact procedure*, employing the  $\varepsilon$ -subgradients instead of the subgradients which allows to control the level of *inexactness* with which we solve the subproblems of the procedure and we preserve its convergence properties. Moreover, we will show that we can employ an *overestimation* of the regularization parameter  $\beta$  in the model  $\varphi_0 + \beta\varphi_1$  in the image restoration framework: such an estimation will induce a contrast enhancement in the reconstructed images, as already observed in [79].

### 4.1 Bregman distance: Definition and Properties

The Bregman distance for convex functions is defined below.

**Definition 4.1.** *Let  $f$  a convex function with a non-empty subdifferential  $\partial f(\mathbf{y})$ , let  $\mathbf{x}, \mathbf{y} \in \text{dom}(f)$ . The Bregman distance of  $f$  from  $\mathbf{x}$  to  $\mathbf{y}$  is*

$$D_f^{\mathbf{p}}(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) - f(\mathbf{y}) - \langle \mathbf{p}, \mathbf{x} - \mathbf{y} \rangle \quad (4.1)$$

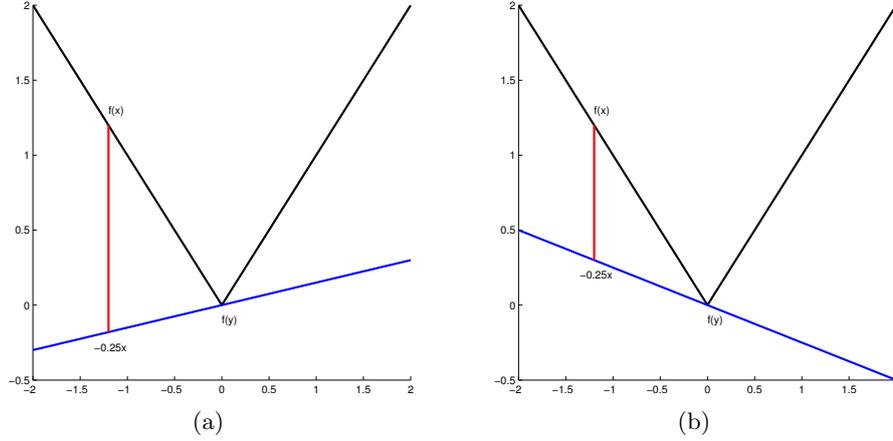
where  $\mathbf{p}$  is a subgradient of  $f$  at  $\mathbf{y}$ :  $\mathbf{p} \in \partial f(\mathbf{y})$ .

Although its name "distance", the function  $D_f^{\mathbf{p}}$  is not a distance in the usual sense because it is not symmetric; one can consider it as a measure of the difference between the function  $f$  and its first order approximation (see Figure 4.1).

**Example 4.1.** Let us consider the differentiable function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f : \mathbf{x} \mapsto \frac{1}{2}\|M\mathbf{x}\|^2$ , with  $M \in \mathcal{M}_n(\mathbb{R})$ . We have  $\partial f(\mathbf{x}) = \nabla f(\mathbf{x}) = M^t M\mathbf{x}$ . Considering its Bregman distance between  $\mathbf{x}, \mathbf{y}$  we obtain

$$\begin{aligned} D_f^{\mathbf{p}}(\mathbf{x}, \mathbf{y}) &= \frac{1}{2}\|M\mathbf{x}\|^2 - \frac{1}{2}\|M\mathbf{y}\|^2 - \langle M^t M\mathbf{y}, \mathbf{x} - \mathbf{y} \rangle \\ &= \frac{1}{2}\|M\mathbf{x}\|^2 + \frac{1}{2}\|M\mathbf{y}\|^2 - \langle M\mathbf{y}, M\mathbf{x} \rangle \\ &= \frac{1}{2}\|M(\mathbf{x} - \mathbf{y})\|^2 = \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|_{M^t M}^2 \end{aligned}$$

Hence the Bregman distance in this case is equivalent to the squared  $\ell_2$ -distance of  $\mathbf{x}$  from  $\mathbf{y}$  with respect to the linear operator  $M^T M$ .



**Figure 4.1:** Bregman distance. In (a) the function  $f(z) = |z|$  is shown; when  $\mathbf{y} = 0$ ,  $\mathbf{x} = -1.2$  and  $p = 0.15$  the length of the red line is the Bregman distance  $D_{|\mathbf{x}|}^{0.15}(-1.2, 0)$ . In the right figure we have taken a different subgradient,  $p = -\frac{1}{4}$ : hence it is evident that the Bregman distance depends strongly on the choice of the subgradient.

**Example 4.2.** Let consider the function  $f : (0, +\infty) \rightarrow \mathbb{R}$ ,  $f : \mathbf{x} \mapsto \mathbf{x} \log(\mathbf{x})$ . Then we have

$$\begin{aligned} D_f^p(\mathbf{x}, \mathbf{y}) &= \mathbf{x} \log(\mathbf{x}) - \mathbf{y} \log(\mathbf{y}) - \langle 1 + \log(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \\ &= \mathbf{x} \log(\mathbf{x}) - \mathbf{y} \log(\mathbf{y}) - \mathbf{x} + \mathbf{y} - \mathbf{x} \log(\mathbf{y}) + \mathbf{y} \log(\mathbf{y}) \\ &= \mathbf{x} \log\left(\frac{\mathbf{x}}{\mathbf{y}}\right) - \mathbf{x} + \mathbf{y} \end{aligned}$$

i.e. we obtain that the Bregman distance of the function  $f$  related to  $\mathbf{x}$  and  $\mathbf{y}$  is the Kullback–Leibler functional of  $\mathbf{x}$  and  $\mathbf{y}$ .

The basic properties of Bregman distance depend on the convexity of  $f$ :  $D_f^p(\mathbf{x}, \mathbf{y}) \geq 0$  for any  $\mathbf{p} \in \partial f(\mathbf{y})$  and  $D_f^p(\mathbf{x}, \mathbf{x}) = 0$ ; moreover, we have that  $D_f^p(\mathbf{x}, \mathbf{y}) \geq D_f^p(\mathbf{w}, \mathbf{y})$  for any  $\mathbf{w} \in \{(1-t)\mathbf{x} + t\mathbf{y} | t \in [0, 1]\}$ . If  $f$  is strongly convex,  $D_f^p(\mathbf{x}, \mathbf{y}) > 0$  for  $\mathbf{x} \neq \mathbf{y}$  and for any  $\mathbf{p} \in \partial f(\mathbf{y})$ . Furthermore, no triangle inequality holds for the Bregman distance.

## 4.2 Bregman Procedure

In this section we show the Bregman Procedure for a general problem of convex optimization. Then we describe its regularization behaviour in image restoration problems. Since we are treating discrete data, the whole discussion will be done in a discrete framework.

The original problem considered in [25] can be stated as

$$\min_{\mathbf{x}} f_2(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{H}\mathbf{x} = \mathbf{c} \quad (4.2)$$

where  $f_2 : \mathbb{R}^N \rightarrow \mathbb{R}$  is a proper, convex, closed and nonnegative function,  $\mathbf{H} \in \mathbb{R}^{M \times N}$ . Assuming that the set  $\{\mathbf{x} \in \text{dom}(f_2) | \mathbf{H}\mathbf{x} = \mathbf{c}\}$  is not empty, the previous problem can

be replaced by a sequence of unconstrained subproblems

$$\min_{\mathbf{x}} q_k(\mathbf{x}) \equiv f_2(\mathbf{x}) + \frac{1}{\beta_k} f_1(\mathbf{x}) \quad (4.3)$$

where  $\beta_k > 0$  for any  $k$  and  $f_1$  is a penalty function for  $\mathbf{H}\mathbf{x} = \mathbf{c}$ , i.e.  $f_1(\mathbf{x}) \geq 0$  for any  $\mathbf{x}$  and  $f_1(\mathbf{x}) = 0$  if and only if  $\mathbf{H}\mathbf{x} = \mathbf{c}$ ;  $f_1$  is chosen as a coercive and convex function in order to be able to solve any  $k$ -th subproblem. A well-known result [71] states that when the subproblems (4.3) have a solution, namely  $\mathbf{x}^k$ , and when the sequence  $\beta_k \rightarrow 0$  as  $k \rightarrow \infty$  any limit point of the sequence  $\{\mathbf{x}^k\}$  is a solution of (4.2). The difficulties arise when  $\beta_k$  assumes very small values.

The procedure based on the Bregman Iteration [25] sets  $\beta_k = \beta$  for any  $k$  while  $f_2$  is replaced with its Bregman Distance at the current iterate. The procedure then modifies the formulation (4.3) by solving a sequence of minimization subproblems as shown in Algorithm 12.

---

**Algorithm 12: Bregman Procedure, General Scheme**

---

Given  $\mathbf{x}^0$  s.t  $\mathbf{p}^0 = \mathbf{0} \in \partial f_2(\mathbf{x}^0)$

FOR  $k = 0, 1, 2, \dots$  DO THE FOLLOWING STEPS

$$\mathbf{x}^{k+1} = \arg \min_x \bar{Q}_k(\mathbf{x}, \mathbf{p}^k) \quad (4.4)$$

where

$$\bar{Q}_k(\mathbf{x}, \mathbf{p}^k) = D_{f_2}^{\mathbf{p}^k}(\mathbf{x}, \mathbf{x}^k) + \frac{1}{\beta} f_1(\mathbf{x})$$

and  $\mathbf{p}^k \in \partial f_2(\mathbf{x}^k)$ .

---

As already observed in [46], this method can be viewed as a generalization of the proximal point algorithm, in which the  $\ell_2$  norm is replaced by the Bregman Distance. The updating rule is given by

$$\mathbf{x}^{k+1} = \left( \partial f_2 + \frac{1}{\beta} \partial f_1 \right)^{-1} \left( \partial f_2(\mathbf{x}^k) \right)$$

In [79] the choice for  $f_1(\mathbf{x}) = \frac{1}{2} \|\mathbf{H}\mathbf{x} - \mathbf{c}\|^2$  has been investigated, showing some features of the Bregman iteration; moreover, in [29, 87] the results are extended to the Poisson case. For sake of completeness, we summarize these features below in a discrete framework.

**Proposition 4.1.** *Let  $f_1$  and  $f_2$  be nonnegative, proper, closed and convex functions, with  $\text{dom}(f_1) \subseteq \text{dom}(f_2)$  and the relative interiors of  $f_1$  and  $f_2$  have at least a point in common. We assume that, for any  $k$ , there exists a minimizer  $\mathbf{x}^{k+1}$  of the subproblem (4.4); then, the following conditions hold:*

(a) *there exists a subgradient  $\mathbf{p}^{k+1} \in \partial f_2(\mathbf{x}^{k+1})$  and  $\mathbf{q}^{k+1} \in \partial f_1(\mathbf{x}^{k+1})$  such that*

$$\mathbf{p}^{k+1} = \mathbf{p}^k - \frac{1}{\beta} \mathbf{q}^{k+1} \quad (4.5)$$

(b) *the sequence  $f_1(\mathbf{x}^k)$  is monotonically non increasing and we have*

$$f_1(\mathbf{x}^k) \leq f_1(\mathbf{x}^k) + \beta D_{f_2}^{\mathbf{p}^{k-1}}(\mathbf{x}^k, \mathbf{x}^{k-1}) \leq f_1(\mathbf{x}^{k-1}) \quad (4.6)$$

(c) if there exists  $\mathbf{x}$  such that  $f_2(\mathbf{x}) < \infty$ , we have

$$\begin{aligned} \beta \left( D_{f_2}^{\mathbf{p}^k}(\mathbf{x}, \mathbf{x}^k) + D_{f_2}^{\mathbf{p}^{k-1}}(\mathbf{x}^k, \mathbf{x}^{k-1}) \right) + f_1(\mathbf{x}^k) &\leq \\ &\leq f_1(\mathbf{x}) + \beta D_{f_2}^{\mathbf{p}^{k-1}}(\mathbf{x}, \mathbf{x}^{k-1}) \end{aligned} \quad (4.7)$$

(d) if  $\tilde{\mathbf{x}}$  is a minimizer of  $f_1$  such that  $f_2(\tilde{\mathbf{x}}) < \infty$ , we have that

$$D_{f_2}^{\mathbf{p}^k}(\tilde{\mathbf{x}}, \mathbf{x}^k) \leq D_{f_2}^{\mathbf{p}^{k-1}}(\tilde{\mathbf{x}}, \mathbf{x}^{k-1}) \quad (4.8)$$

and

$$f_1(\mathbf{x}^k) \leq f_1(\tilde{\mathbf{x}}) + \beta \frac{f_2(\tilde{\mathbf{x}}) - f_2(\mathbf{x}^0)}{k} \quad (4.9)$$

Moreover, if the level subsets of  $f_1$  are bounded, a limit point of the sequence  $\{\mathbf{x}^k\}$  is a minimizer of  $f_1(x)$ ; if  $\tilde{\mathbf{x}}$  is the unique minimizer of  $f_1(x)$ , then  $\mathbf{x}^k \rightarrow \tilde{\mathbf{x}}$  as  $k \rightarrow \infty$ .

*Proof.*

- (a) From the optimality condition for the minimizer  $\mathbf{x}^{k+1}$  of  $\bar{Q}_k(\mathbf{x}, \mathbf{p}^k)$  and in accordance to Proposition 23.8 in [85], we have  $\mathbf{0} \in \partial f_2(\mathbf{x}^{k+1}) - \mathbf{p}^k + \frac{1}{\beta} \partial f_1(\mathbf{x}^{k+1})$ . Then (4.5) follows.
- (b) From  $\bar{Q}_{k-1}(\mathbf{x}^{k-1}, \mathbf{p}^{k-1}) = \frac{1}{\beta} f_1(\mathbf{x}^{k-1})$  and  $D_{f_2}^{\mathbf{p}^{k-1}}(\mathbf{x}^k, \mathbf{x}^{k-1}) \geq 0$ , since  $\mathbf{x}^k$  is a minimizer of  $\bar{Q}_{k-1}(\mathbf{x}, \mathbf{p}^{k-1})$ , we have  $\frac{1}{\beta} f_1(\mathbf{x}^k) \leq \bar{Q}_{k-1}(\mathbf{x}^k, \mathbf{p}^{k-1}) \leq \bar{Q}_{k-1}(\mathbf{x}^{k-1}, \mathbf{p}^{k-1})$  and (4.6) holds.
- (c) By direct algebra, the following identity holds:

$$\begin{aligned} D_{f_2}^{\mathbf{p}^k}(\mathbf{x}, \mathbf{x}^k) - D_{f_2}^{\mathbf{p}^{k-1}}(\mathbf{x}, \mathbf{x}^{k-1}) + D_{f_2}^{\mathbf{p}^{k-1}}(\mathbf{x}^k, \mathbf{x}^{k-1}) &= \\ &= \langle \mathbf{x}^k - \mathbf{x}, \mathbf{p}^k - \mathbf{p}^{k-1} \rangle > \end{aligned}$$

Using (4.5), since  $\mathbf{p}^k - \mathbf{p}^{k-1} = -\frac{1}{\beta} \mathbf{q}^k \in \frac{1}{\beta} \partial f_1(\mathbf{x}^k)$ , from the convexity of  $f_1$ , we have (4.7).

- (d) If  $\tilde{\mathbf{x}}$  is a minimizer of  $f_1$ , from (4.7) with  $\mathbf{x} = \tilde{\mathbf{x}}$ , since  $D_{f_2}^{\mathbf{p}^{k-1}}(\mathbf{x}^k, \mathbf{x}^{k-1}) \geq 0$ , we obtain

$$D_{f_2}^{\mathbf{p}^k}(\tilde{\mathbf{x}}, \mathbf{x}^k) + \frac{1}{\beta} (f_1(\mathbf{x}^k) - f_1(\tilde{\mathbf{x}})) \leq D_{f_2}^{\mathbf{p}^{k-1}}(\tilde{\mathbf{x}}, \mathbf{x}^{k-1})$$

Since  $f_1(\mathbf{x}^k) - f_1(\tilde{\mathbf{x}}) \geq 0$ , the inequality (4.8) holds. Furthermore, summing up the inequalities (4.7) computed at  $\tilde{\mathbf{x}}$  related to the first  $k$  steps, we have:

$$\begin{aligned} D_{f_2}^{\mathbf{p}^k}(\tilde{\mathbf{x}}, \mathbf{x}^k) + \sum_{i=1}^k \left[ D_{f_2}^{\mathbf{p}^{i-1}}(\mathbf{x}^i, \mathbf{x}^{i-1}) + \frac{1}{\beta} (f_1(\mathbf{x}^i) - f_1(\tilde{\mathbf{x}})) \right] &\leq \\ &\leq f_2(\tilde{\mathbf{x}}) - f_2(\mathbf{x}^0) \end{aligned} \quad (4.10)$$

Since  $D_{f_2}^{p^{i-1}}(\mathbf{x}^i, \mathbf{x}^{i-1}) \geq 0$  for any  $i$  and  $D_{f_2}^{p^k}(\tilde{\mathbf{x}}, \mathbf{x}^k) \geq 0$ , from the monotonicity of the sequence  $f_1(\mathbf{x}^i)$ , we have

$$\frac{k}{\beta}[f_1(\mathbf{x}^k) - f_1(\tilde{\mathbf{x}})] \leq f_2(\tilde{\mathbf{x}}) - f_2(\mathbf{x}^0)$$

and then (4.9) follows.

Since the sequence  $\{\mathbf{x}^k\}$  is bounded, there exists a subsequence of  $\{\mathbf{x}^k\}$  convergent to a limit point  $\bar{\mathbf{x}}$  and, from (4.9), we have for  $k \rightarrow \infty$  that  $f_1(\bar{\mathbf{x}}) \leq f_1(\tilde{\mathbf{x}})$ . Then,  $\bar{\mathbf{x}}$  is a minimizer of  $f_1$ . If  $\tilde{\mathbf{x}}$  is the unique minimizer of  $f_1$ , then  $\mathbf{x}^k \rightarrow \tilde{\mathbf{x}}$  as  $k \rightarrow \infty$ .  $\square$

**Remark 4.1.** Under suitable hypothesis, Proposition 4.1 guarantees the convergence of the sequence of the minimizers of the subproblems (4.4) to a solution  $\tilde{\mathbf{x}}$  of  $f_1(\mathbf{x}) = 0$  (i.e. to a solution of  $\mathbf{H}\mathbf{x} = \mathbf{c}$ ) while the sequence  $\{D_{f_2}^{p^k}(\tilde{\mathbf{x}}, \mathbf{x}^k)\}$  is decreasing. This convergence result is strongly dependent on the decreasing behaviour of the sequence  $\{f_1(\mathbf{x}^k)\}$ , which follows directly from the nonnegativity of the Bregman Distance.

In a numerical framework, one can obtain this decreasing behaviour when (4.5) is able to give an *exact* subgradient at the current iterate: when a closed formula for the minimizer of  $\bar{Q}_k(\mathbf{x}, \mathbf{p}^k)$  is not available, we have to use a numerical approach. If  $\mathbf{x}^k$  is a raw minimizer of  $\bar{Q}_{k-1}$ , it could happen that  $\mathbf{p}^k \notin \partial f_2(\mathbf{x}^k)$ , leading to  $D_{f_2}^{p^k} < 0$ ; at the successive step, we may have

$$\bar{Q}_k(\mathbf{x}^{k+1}, \mathbf{1p}^k) = \frac{1}{\beta}f_1(\mathbf{x}^{k+1}) + D_{f_2}^{p^k}(\mathbf{x}^{k+1}, \mathbf{x}^k) < \frac{1}{\beta}f_1(\mathbf{x}^{k+1})$$

Consequently the sequence  $\{f_1(\mathbf{x}^k)\}$  may have a non-monotone behaviour.

We can obtain an exact subgradient when an *explicit* solution of each inner problem is available; when it is not possible to obtain such a solution, the procedure has to solve each minimization subproblem by a numerical iterative solver with a very high degree of accuracy. In the latter case, the computational cost would be very high.

**Remark 4.2.** We point out that the updating rule for the subgradient can be restate by a variable change, in order to redefine  $\mathbf{x}^{k+1}$ ; in fact, provided  $\mathbf{y} = \mathbf{H}\mathbf{x}$  we can consider  $f_1$  as a function of  $\mathbf{y}$ : hence, we obtain  $\mathbf{q}^{k+1} = \mathbf{H}^t \mathbf{u}^{k+1}$ , having  $\mathbf{u}^{k+1} \in \partial_y f_1(\mathbf{H}\mathbf{x}^{k+1})$ . In the setup of the procedure,  $\mathbf{p}^0 = \mathbf{0}$ ; hence, considering  $\mathbf{v}^0 = \mathbf{0}$ , a direct computation leads to the following rule:

$$\mathbf{p}^{k+1} = \mathbf{H}^t \mathbf{v}^{k+1} = \mathbf{H}^t \left( \mathbf{v}^k - \frac{1}{\beta} \mathbf{u}^{k+1} \right)$$

Then, the formula (4.5) can be substituted by

$$\mathbf{v}^{k+1} = \mathbf{v}^k - \frac{1}{\beta} \mathbf{u}^{k+1} \quad (4.11)$$

In this way, the definition in (4.4) of  $\mathbf{x}^{k+1}$  can be rewritten as

$$\mathbf{x}^{k+1} = \arg \min_{\mathbf{x}} \left( f_2(\mathbf{x}) - \langle \mathbf{v}^k, \mathbf{H}\mathbf{x} \rangle + \frac{1}{\beta} f_1(\mathbf{x}) \right) \quad (4.12)$$

In Proposition 4.2 we are able to show that a solution  $\mathbf{x}^\kappa$  of  $f_1(\mathbf{x}) = 0$ , obtained thanks to (4.11) and (4.12), is a solution of the original constrained problem (4.2), or its equivalent formulation

$$\min_{\mathbf{x}} f_2(\mathbf{x}) \quad \text{s.t.} \quad f_1(\mathbf{x}) = 0 \quad (4.13)$$

**Proposition 4.2.** *Let  $f_1$  be a convex function such that  $f_1(\mathbf{x}) = 0$  if and only if  $\mathbf{H}\mathbf{x} = \mathbf{c}$ . Suppose that some iterate  $\mathbf{x}^\kappa$  of the Bregman procedure satisfies  $f_1(\mathbf{x}^\kappa) = 0$ . Then  $\mathbf{x}^\kappa$  is a solution of the constrained problem (4.2) (or (4.13)).*

*Proof.* Let  $\mathbf{x}^\kappa$  be such that  $f_1(\mathbf{x}^\kappa) = 0$  and

$$\mathbf{x}^\kappa = \arg \min_{\mathbf{x}} \left( f_2(\mathbf{x}) - \langle \mathbf{v}^\kappa, \mathbf{H}\mathbf{x} \rangle + \frac{1}{\beta} f_1(\mathbf{x}) \right) \quad (4.14)$$

for a suitable  $\mathbf{v}^\kappa$ . Let  $\tilde{\mathbf{x}}$  be a solution of the problem (4.2). Then  $f_1(\tilde{\mathbf{x}}) = 0$  and, for the hypothesis on  $f_1$ ,

$$\mathbf{H}\tilde{\mathbf{x}} = \mathbf{c} = \mathbf{H}\mathbf{x}^\kappa \quad (4.15)$$

Since  $\mathbf{x}^\kappa$  satisfies (4.14), we have

$$f_2(\mathbf{x}^\kappa) - \langle \mathbf{v}^\kappa, \mathbf{H}\mathbf{x}^\kappa \rangle + \frac{1}{\beta} f_1(\mathbf{x}^\kappa) \leq f_2(\tilde{\mathbf{x}}) - \langle \mathbf{v}^\kappa, \mathbf{H}\tilde{\mathbf{x}} \rangle + \frac{1}{\beta} f_1(\tilde{\mathbf{x}}) \quad (4.16)$$

Using (4.15) in (4.16) and taking into account that  $f_1(\mathbf{x}^\kappa) = f_1(\tilde{\mathbf{x}}) = 0$ , we have that

$$f_2(\mathbf{x}^\kappa) \leq f_2(\tilde{\mathbf{x}})$$

Because  $\tilde{\mathbf{x}}$  is a solution of the original optimization problem, this last inequality is an equality, showing that  $\mathbf{x}^\kappa$  solves (4.2).  $\square$

#### 4.2.1 The procedure for image restoration problems

In image restoration framework the Bregman procedure has a huge importance: in fact, employing an early stopping criterion on the external cycle, it can be used as a regularization technique. To deepen this aspect, let us consider the problem

$$\min_{\mathbf{x}} \varphi_0(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta\varphi_1(\mathbf{x})$$

where  $\varphi_0$  is the data fidelity function (e.g. the Kullback–Leibler) and  $\varphi_1$  a regularization term. In this case  $\varphi_0$  plays the role of penalty term for  $\mathbf{H}\mathbf{x} = \mathbf{g}\mathbf{n} - b\mathbf{g}\mathbf{1}$  in formulation (4.2), with  $\mathbf{c} = \mathbf{g}\mathbf{n} - b\mathbf{g}\mathbf{1}$ . Thanks to Proposition 4.1 we can state that a limit point  $\{\mathbf{x}^k\}$  is a solution  $\tilde{\mathbf{x}}$  for  $\varphi_0(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) = 0$ , but actually we are interested in finding a solution  $\mathbf{x}^*$  of  $\varphi_0(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}) = 0$ , where  $\mathbf{g}$  is the noise-free data (here  $\mathbf{c} = \mathbf{g} - b\mathbf{g}\mathbf{1}$ ). In this framework, the Bregman procedure has the typical semiconvergence behaviour of the iterative methods for inverse problems, as described in [47]: the sequence  $\{\mathbf{x}^k\}$  first approaches the requested solution  $\mathbf{x}^*$ , and then runs away, converging toward the unrequested solution  $\tilde{\mathbf{x}}$  [79].

When an estimation  $\gamma$  of the noise level is available, e.g.  $\varphi_0(\mathbf{H}\mathbf{x}^* + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) \leq \gamma$ , from (4.7) we can state that for  $\mathbf{x} = \mathbf{x}^*$  and while  $\varphi_0(\mathbf{H}\mathbf{x}^k + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) \geq \gamma$  we can show that the Bregman Distance of the iterates from the desired solution  $\mathbf{x}^*$  decreases:

$$D_{\varphi_1}^{\mathbf{p}^k}(\mathbf{x}^*, \mathbf{x}^k) \leq D_{\varphi_1}^{\mathbf{p}^{k-1}}(\mathbf{x}^*, \mathbf{x}^{k-1}) \quad (4.17)$$

Indeed, from (4.7) with  $\mathbf{x} = \mathbf{x}^*$ ,  $f_1 = \varphi_0$ ,  $f_2 = \varphi_1$  we can write

$$\begin{aligned} & \beta \left( D_{\varphi_1}^{\mathbf{P}^k}(\mathbf{x}^*, \mathbf{x}^k) + D_{\varphi_1}^{\mathbf{P}^{k-1}}(\mathbf{x}^k, \mathbf{x}^{k-1}) \right) + \varphi_0(\mathbf{H}\mathbf{x}^k + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) \leq \\ & \leq \varphi_0(\mathbf{H}\mathbf{x}^* + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta D_{\varphi_1}^{\mathbf{P}^{k-1}}(\mathbf{x}^*, \mathbf{x}^{k-1}) \end{aligned}$$

In the first term, since the iterates stand above the noise level and the Bregman distance is nonnegative, we have

$$\beta \left( D_{\varphi_1}^{\mathbf{P}^k}(\mathbf{x}^*, \mathbf{x}^k) + D_{\varphi_1}^{\mathbf{P}^{k-1}}(\mathbf{x}^k, \mathbf{x}^{k-1}) \right) + \varphi_0(\mathbf{H}\mathbf{x}^k + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) \geq \beta D_{\varphi_1}^{\mathbf{P}^k}(\mathbf{x}^*, \mathbf{x}^k) + \gamma$$

while the second term satisfies

$$\varphi_0(\mathbf{H}\mathbf{x}^* + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta D_{\varphi_1}^{\mathbf{P}^{k-1}}(\mathbf{x}^*, \mathbf{x}^{k-1}) \leq \gamma + \beta D_{\varphi_1}^{\mathbf{P}^{k-1}}(\mathbf{x}^*, \mathbf{x}^{k-1})$$

leading finally to (4.17).

The same argument is presented in [79]. Furthermore, thanks to (4.6), we can devise a stopping criterion for the iterative procedure: the method terminates at the iteration  $k^*$  defined as

$$k^* = \max\{k \mid \varphi_0(\mathbf{H}\mathbf{x}^k + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) \geq \gamma\}$$

In presence of Gaussian noise this strategy is the Morozov discrepancy principle, which seems to work quite well. In presence of Poisson data, the Bregman iteration could be stopped when the Kullback–Leibler divergence between  $\mathbf{H}\mathbf{x}^k + b\mathbf{g}\mathbf{1}$  and the given data  $\mathbf{g}\mathbf{n}$  reaches the noise level, namely  $\mathbf{KL}(\mathbf{H}\mathbf{x}^k + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) \sim \gamma$ . For an estimate for the Poisson noise level, one can employ the strategy developed in the previous chapter and in [13].

To better explain the effect of the Bregman iteration in the image reconstruction framework, in the following we show the behaviour of the first two steps of the Bregman iteration for data affected by Gaussian noise, where explicit formulae for the minimizers of the subproblems are available.

**Remark 4.3.** [10] In presence of Gaussian noise, a common choice for the objective functional is the least square functional combined with Tikhonov regularization. We apply the Bregman method to this particular case, obtaining a closed form for the  $k$ -th solution and showing the effectiveness of the method in reconstructing the original signal.

In this framework the method consists in solving at each  $k$  step

$$\mathbf{x}^{k+1} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{H}\mathbf{x} - \mathbf{g}\mathbf{n}\|^2 + \frac{\beta}{2} \|\mathbf{x} - \mathbf{x}^k\|^2 \quad (4.18)$$

The explicit formula for the minimizer is available:

$$\mathbf{x}^{k+1} = (\mathbf{H}^t \mathbf{H} + \beta \mathbf{I})^{-1} (\mathbf{H}^t \mathbf{g}\mathbf{n} + \beta \mathbf{x}^k) \quad (4.19)$$

Looking at the first iterations, we have

$$\begin{aligned} \mathbf{x}^0 &= \mathbf{0} \\ \mathbf{x}^1 &= (\mathbf{H}^t \mathbf{H} + \beta \mathbf{I})^{-1} \mathbf{H}^t \mathbf{g}\mathbf{n} \equiv \mathbf{x}_\beta \\ \mathbf{x}^2 &= \mathbf{x}_\beta + \beta (\mathbf{H}^t \mathbf{H} + \beta \mathbf{I})^{-1} \mathbf{x}_\beta \\ &\dots \end{aligned}$$

where  $\mathbf{x}_\beta$  denotes the minimizer of the first subproblem ( $\mathbf{x}^0 = \mathbf{0}$ ). By induction

$$\mathbf{x}^k = \mathbf{x}_\beta + \beta (\mathbf{H}^t \mathbf{H} + \beta \mathbf{I})^{-1} \mathbf{x}_\beta + \cdots + \beta^{k-1} (\mathbf{H}^t \mathbf{H} + \beta \mathbf{I})^{-(k-1)} \mathbf{x}_\beta$$

We introduce the Singular Value Decomposition (SVD) of the matrix  $\mathbf{H}$ :

$$\mathbf{H} = U \Sigma V^t$$

where  $U$  and  $V$  are orthogonal matrices and  $\Sigma$  is a diagonal matrix; its positive diagonal entries are the singular values of  $\mathbf{H}$ . Using the SVD of  $\mathbf{H}$ , one can compute the filter corresponding to this linear solution; recalling  $\mathbf{H} \mathbf{v}_j = \sigma_j \mathbf{u}_j$ ,  $\mathbf{H}^t \mathbf{u}_j = \sigma_j \mathbf{v}_j$ , direct computation leads to

$$\begin{aligned} \langle \mathbf{x}^k, \mathbf{v}_j \rangle &= \left\{ \sum_{i=0}^{k-1} \left( \frac{\beta}{\beta + \sigma_j^2} \right)^i \right\} \langle \mathbf{x}_\beta, \mathbf{v}_j \rangle \\ &= \frac{\beta + \sigma_j^2}{\sigma_j^2} \left\{ 1 - \left( \frac{\beta}{\beta + \sigma_j^2} \right)^k \right\} \langle \mathbf{x}_\beta, \mathbf{v}_j \rangle \\ &= \left\{ 1 - \left( \frac{\beta}{\beta + \sigma_j^2} \right)^k \right\} \frac{1}{\sigma_j} \langle \mathbf{g} \mathbf{n}, \mathbf{u}_j \rangle \end{aligned}$$

where in the last equality we have used the definition of  $\mathbf{x}_\beta$  and the well-known properties of the SVD decomposition. Defining  $B_j^k(\beta)$  as

$$B_j^k(\beta) = 1 - \left( \frac{\beta}{\beta + \sigma_j^2} \right)^k$$

we have proved the following result.

**Proposition 4.3.** *The solution of the  $k$ -th Bregman iteration can be expressed as*

$$\mathbf{x}^k = \sum_j B_j^k(\beta) \frac{1}{\sigma_j} \langle \mathbf{g} \mathbf{n}, \mathbf{u}_j \rangle \mathbf{v}_j \quad (4.20)$$

*in terms of SVD of the matrix  $\mathbf{H}$ .*

The limit point of the sequence  $\{\mathbf{x}^k\}$ , i.e.  $\lim_{k \rightarrow \infty} \mathbf{x}^k = \mathbf{x}^\dagger$ , is the generalized solution of the linear problem.

In order to prove the semiconvergence behaviour of the procedure, it is sufficient to prove that the reconstruction error obtained at the second iteration is smaller than the one provided by the first iteration. Recalling the statistical framework (see Chapter 1), the data is  $\mathbf{g} \mathbf{n} = \mathbf{H} \mathbf{x}^* + v$  where the vectors  $\mathbf{x}^*$  and  $v$  are realizations of Gaussian multi-valued random variables with entries that are i.i.d. variables with zero mean and variance  $S$  and  $s$  respectively. Then we can write

$$E[\mathbf{x}_i^*, v_j] = 0, \quad E[\mathbf{x}_i^*, \mathbf{x}_j^*] = S \delta_{ij}, \quad E[v_i, v_j] = s \delta_{ij}$$

where  $E[\cdot]$  is the expected value. Hence, the relation  $\langle \mathbf{gn}, \mathbf{u}_j \rangle = \langle \mathbf{x}^*, \mathbf{H}^t \mathbf{u}_j \rangle + \langle v, \mathbf{u}_j \rangle = \sigma_j \langle \mathbf{x}^*, \mathbf{v}_j \rangle + \langle v, \mathbf{u}_j \rangle$  implies

$$\langle \mathbf{x}^* - \mathbf{x}^k, \mathbf{v}_j \rangle = \left( \frac{\beta}{\beta + \sigma_j^2} \right)^k \langle \mathbf{x}^*, \mathbf{v}_j \rangle - \frac{1}{\sigma_j} B_j^k(\beta) \langle v, \mathbf{u}_j \rangle$$

The expected value of the square of this quantity is

$$E [ |\langle \mathbf{x}^* - \mathbf{x}^k, \mathbf{v}_j \rangle|^2 ] = S^2 \left( \frac{\beta}{\beta + \sigma_j^2} \right)^{2k} + \frac{s^2}{\sigma_j^2} |B_j^k(\beta)|^2$$

which allows us to compute the reconstruction error  $\rho_k$  at the  $k$ -th Bregman iteration:

$$\begin{aligned} \rho_k &= E[\|\mathbf{x}^* - \mathbf{x}^k\|^2] \\ &= S^2 \sum_j \left\{ \left( \frac{\beta}{\beta + \sigma_j^2} \right)^{2k} + \frac{s^2}{S^2 \sigma_j^2} \left[ 1 - \left( \frac{\beta}{\sigma_j^2 + \beta} \right)^k \right]^2 \right\} \end{aligned} \quad (4.21)$$

For  $k = 1$ ,  $\rho_k$  is the reconstruction error obtained solving the problem (4.18), while  $\rho_2$  is associated to the first effective Bregman iteration. The semiconvergence behaviour is then proved if we prove that  $\rho_2 < \rho_1$ . Setting

$$\alpha = \left( \frac{s}{S} \right)^2, \quad P_j(t) = t^2 + \frac{\alpha}{\sigma_j^2} (1 - t)^2$$

the reconstruction error assumes the following form

$$\rho_k = S^2 \sum_j P_j \left[ \left( \frac{\beta}{\beta + \sigma_j^2} \right)^k \right] \quad (4.22)$$

The following expression

$$P_j(t) = \left( 1 + \frac{\alpha}{\sigma_j^2} \right) \left\{ \left( t - \frac{\alpha}{\alpha + \sigma_j^2} \right)^2 + \left( \alpha \frac{\sigma_j}{\alpha^2 + \sigma_j^2} \right)^2 \right\}$$

shows that for a given  $j$ ,  $P_j(t)$  is strictly positive and has the minimum point

$$t_j = \frac{\alpha}{\alpha + \sigma_j^2}$$

i.e. it is a parabola symmetric with respect to this point. Since  $\rho_1$  and  $\rho_2$  are the values of this polynomial at the points

$$t_{j,1} = \frac{\beta}{\beta + \sigma_j^2}, \quad t_{j,2} = \left( \frac{\beta}{\beta + \sigma_j^2} \right)^2$$

we can have  $\rho_2 < \rho_1$  if it is possible to find values of  $\beta$  such that the following inequalities hold true for any  $j$

$$t_{j,1} > t_{j,2} > 2t_j - t_{j,1} \quad (4.23)$$

The first inequality is trivially satisfied and therefore the second is the crucial one: it implies that, for any  $j$ , the value of the polynomial  $P_j(t)$  in  $t_{j,2}$  is smaller than the value of the same polynomial in  $t_{j,1}$ .

**Proposition 4.4.** *The second inequality of (4.23) is satisfied for any  $j$  if*

$$\beta > 2\alpha + \mathcal{O}\left(\frac{\alpha^2}{\sigma_1^2}\right) \quad (4.24)$$

where  $\sigma_1$  is the largest singular value of the imaging matrix  $\mathbf{H}$ .

*Proof.* If we remark that  $t_{j,2} = t_{j,2}^2$  the inequality becomes  $t_{j,1}^2 + t_{j,1} > 2t_j$  and, for positive  $t_{j,1}$ , it is equivalent to

$$t_{j,1} > \sqrt{2t_j + \frac{1}{4}} - \frac{1}{2}$$

Thanks to the dependence of  $t_{j,1}$  on  $\beta$  this inequality implies

$$\left(\frac{3}{2} - \sqrt{2t_j + \frac{1}{4}}\right)\beta > \left(\sqrt{2t_j + \frac{1}{4}} - \frac{1}{2}\right)\sigma_j^2$$

so that, multiplying both members by  $(3/2 + \sqrt{2t_j + 1/4})$  and remarking that  $t_j > 1$ , after some algebra, we obtain

$$\beta > \frac{1}{2} \frac{2t_j + \sqrt{2t_j + \frac{1}{4}} - \frac{1}{2}}{1 - t_j} \sigma_j^2$$

Finally, if we write  $t_j$  in terms of  $\alpha$ , an elementary computation provides

$$\beta > \frac{1}{4} \left(3\alpha - \sigma_j^2 + \sqrt{9\alpha^2 + 10\alpha\sigma_j^2 + \sigma_j^4}\right)$$

If we remark that the function

$$\phi(t) = \sqrt{t^2 + 10\alpha t + 9\alpha^2} - t$$

is an increasing function of  $t$ , then we can conclude that  $\beta$  satisfies the inequality for any  $j$  if it satisfies the inequality in the case  $j = 1$  corresponding to the largest singular value of  $\mathbf{H}$ . Therefore the required condition on  $\beta$  is

$$\beta > \frac{1}{4} \left(3\alpha - \sigma_1^2 + \sqrt{9\alpha^2 + 10\alpha\sigma_1^2 + \sigma_1^4}\right)$$

If we remark that  $\alpha$  is much smaller than  $\sigma_1$ , the Proposition follows from  $\sqrt{1+t} = 1 + t/2 + \mathcal{O}(t^2)$ .  $\square$

The result implies that, if we do not know the optimal value of the regularization parameter (i.e.  $\alpha$  in these computations) but only an upper bound  $\beta$ , so that the corresponding  $\mathbf{x}_\beta$  is an over-smoothed solution of the problem, then the second Bregman iteration improves the reconstruction. The reconstruction may be further improved by the subsequent Bregman iterations, before degrading as an effect of noise propagation. Since, for any  $k$ , the reconstruction errors are given by the values of the same polynomial computed in points given by increasing powers of  $t_{j,1}$ , an analysis of this situation does not look impossible, even if much more involved.

### 4.3 Inexact Bregman Procedure

When a closed formula for the solution of inner minimization subproblem (4.4) is unavailable, at any step we 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 order to avoid this computational cost, we propose a strategy whose aim is to preserve the convergence property and the features of the whole procedure, taking into account the inexactness of the computed solutions of the inner subproblems. The crucial point of the proposed scheme is devising a suitable stopping criterion for the inner solver of subproblems (4.4).

To explain this proposed scheme, we recall the notion of  $\varepsilon$ -subgradient (see Definition 2.2 and Appendix A); moreover, being  $f$  a proper, convex function on  $\mathbb{R}^n$ , if  $\varepsilon_1 > \varepsilon_2 > 0$  then

$$\partial_{\varepsilon_1} f(\mathbf{x}) \supseteq \partial_{\varepsilon_2} f(\mathbf{x}) \supseteq \partial f(\mathbf{x}) \quad (4.25)$$

Recalling also the definition of the conjugate of a convex function (see Appendix A, definition A.5), we report an useful result in Proposition 4.5: this result will turn out to be helpful in the computational framework.

**Proposition 4.5.** *Let  $f$  a proper, convex and l.s.c. function. Then for every  $\mathbf{x} \in \text{dom}(f)$  and  $\xi \in \text{dom}(f^*)$ , we have  $\xi \in \partial_\varepsilon f(\mathbf{x})$  with  $\varepsilon = f(\mathbf{x}) - \langle \xi, \mathbf{x} \rangle + f^*(\xi)$ .*

*Proof.* Let  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$  and  $\xi \in \text{dom}(f^*)$ . Then, we can write

$$\begin{aligned} f(\mathbf{x}) + \langle \xi, \mathbf{y} - \mathbf{x} \rangle &= f(\mathbf{x}) - (\langle \xi, \mathbf{x} \rangle - f^*(\xi)) + \langle \xi, \mathbf{y} \rangle - f^*(\xi) \\ &\leq f(\mathbf{x}) - (\langle \xi, \mathbf{x} \rangle - f^*(\xi)) + \sup_{\xi} \langle \xi, \mathbf{y} \rangle - f^*(\xi) \\ &= \underbrace{f(\mathbf{x}) - (\langle \xi, \mathbf{x} \rangle - f^*(\xi))}_{=\varepsilon} + f(\mathbf{y}) \end{aligned}$$

Since  $f(x) = \sup_x \langle y, x \rangle - f^*(y)$ , then  $\varepsilon \geq 0$ ; thus, Definition 2.2 is fulfilled.  $\square$

The computation of the Bregman distance  $D_f^{\mathbf{p}}(\mathbf{x}, \mathbf{y})$  requires the subgradient  $\mathbf{p}$ . When an  $\varepsilon$ -subgradient  $\xi$  of  $f$  is available, we introduce

$$\Delta_f^\xi(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) - f(\mathbf{y}) - \langle \xi, \mathbf{x} - \mathbf{y} \rangle + \varepsilon \quad (4.26)$$

We will refer to it as the *inexact Bregman distance*. We have that for any  $\mathbf{y}, \mathbf{x} \in \mathbb{R}^n$ ,  $\Delta_f^\xi(\mathbf{x}, \mathbf{y}) \geq 0$ . For  $\varepsilon = 0$  (i.e.  $\xi \in \partial f(\mathbf{x})$ ) we obtain the classical Bregman distance:  $\Delta_f^\xi = D_f^\xi$ .

The scheme shown in Algorithm 12 can be modified by considering the inexact distance instead of the exact one. Provided that  $\xi^0 = \mathbf{0} \in \partial f_2(\mathbf{x}^0)$  ( $\varepsilon_0 = 0$ ), at any step  $k \geq 0$  we have to find the minimizer of the function  $Q_k(\mathbf{x}, \xi^k)$ , where  $\xi^k$  is an  $\varepsilon_k$ -subgradient of  $f_2$  in  $\mathbf{x}^k$ . The new scheme is shown in Algorithm 13.

We suppose that every  $k$ -th subproblem (4.27) is solved by employing an iterative method, which enables us to obtain an approximate solution  $\mathbf{x}^{k+1}$ , a suitable subgradient  $\mathbf{q}^{k+1} \in \partial f_1(\mathbf{x}^{k+1})$  and an  $\varepsilon_{k+1}$ -subgradient  $\xi^{k+1}$  at  $\mathbf{x}^{k+1}$ , for any  $k \geq 0$ . Since we use an iterative method to solve the inner subproblems, the whole procedure consists in two nested cycle: we point out that we use the index  $k$  for indicating the

---

**Algorithm 13: Inexact Bregman Procedure: General Scheme**


---

Given  $\mathbf{x}^0$  s.t.  $\mathbf{0} = \xi^0 \in \partial f_2(\mathbf{x}^0)$ ,

FOR  $k = 0, 1, 2, \dots$  DO THE FOLLOWING STEPS

$$\mathbf{x}^{k+1} \simeq \arg \min_x Q_k(\mathbf{x}, \xi^k) \quad (4.27)$$

where

$$Q_k(\mathbf{x}, \xi^k) = \frac{1}{\beta} f_1(\mathbf{x}) + \Delta_{f_2}^{\xi^k}(\mathbf{x}, \mathbf{x}^k) \quad (4.28)$$

and  $\xi^k \in \partial_{\varepsilon_k} f_2(\mathbf{x}^k)$ .

---

outer iteration and the index  $\ell$  for the inner solver iteration.

We distinguish two cases, based on the differentiability of  $f_2$ .

**Case  $f_2$  differentiable.** We assume that the inner solver generates two sequences  $\{\mathbf{x}_\ell^k\}$  and  $\{\mathbf{q}_\ell^k\}$  such that

$$\lim_{\ell \rightarrow \infty} \mathbf{x}_\ell^k = \bar{\mathbf{x}}^k, \quad \lim_{\ell \rightarrow \infty} \mathbf{q}_\ell^k = \bar{\mathbf{q}}^k$$

where  $\bar{\mathbf{x}}^k$  is a minimizer of

$$Q_k(\mathbf{x}, \xi^k) = \frac{1}{\beta} f_1(\mathbf{x}) + \Delta_{f_2}^{\xi^k}(\mathbf{x}, \mathbf{x}^k)$$

and  $\bar{\mathbf{q}}^k \in \partial f_1(\bar{\mathbf{x}}^k)$ . As a consequence, given a  $\mu_{k+1} > 0$  for any  $k$ , there exists an index  $\mathcal{L}$  such that  $\|\eta_{\mathcal{L}}^k\| \leq \mu_{k+1}$ , where

$$\eta_\ell^k = \frac{1}{\beta} \mathbf{q}_\ell^k + \nabla f_2(\mathbf{x}_\ell^k) - \xi^k$$

Then we set the approximate solution  $\mathbf{x}^{k+1}$  of the inner subproblem as  $\mathbf{x}_{\mathcal{L}}^k$ ; the other quantities are set in the following way

$$\xi^{k+1} = \nabla f_2(\mathbf{x}^{k+1}), \quad \varepsilon_{k+1} = 0 \quad \text{and} \quad \mathbf{q}^{k+1} = \mathbf{q}_{\mathcal{L}}^k$$

**Case  $f_2$  non differentiable.** We can consider the primal–dual formulation of (4.27) given by

$$\min_{\mathbf{x}} \max_{\mathbf{y}} \Phi_k(\mathbf{x}, \mathbf{y}) \equiv \frac{1}{\beta} f_1(\mathbf{x}) + \langle \mathbf{y}, \mathbf{x} \rangle - f_2^*(\mathbf{y}) - f_2(\mathbf{x}^k) - \langle \xi^k, \mathbf{x} - \mathbf{x}^k \rangle + \varepsilon_k \quad (4.29)$$

and we can apply a suitable primal–dual method generating sequences  $\{\mathbf{x}_\ell^k\}$ ,  $\{\mathbf{y}_\ell^k\}$  convergent to a saddle point  $(\bar{\mathbf{x}}^k, \bar{\mathbf{y}}^k)$  of the convex–concave proper function  $\Phi_k(\mathbf{x}, \mathbf{y})$ . We recall that  $(\bar{\mathbf{x}}^k, \bar{\mathbf{y}}^k)$  is a saddle point of  $\Phi_k(\mathbf{x}, \mathbf{y})$  if there exist  $\bar{\mathbf{q}}^k \in \partial f_1(\bar{\mathbf{x}}^k)$  and  $\mathbf{w}^k \in \partial f_2^*(\bar{\mathbf{y}}^k)$  such that the following conditions hold:

$$\begin{aligned} \frac{1}{\beta} \bar{\mathbf{q}}^k + \bar{\mathbf{y}}^k - \xi^k &= 0 \\ \bar{\mathbf{x}}^k &= \mathbf{w}^k \end{aligned}$$

The inner solver generates three sequences  $\{\mathbf{x}_\ell^k\}$ ,  $\{\mathbf{y}_\ell^k\}$  and  $\{\mathbf{q}_\ell^k\}$ : we assume that for  $\ell \rightarrow \infty$  they converge to  $\bar{\mathbf{x}}^k$ ,  $\bar{\mathbf{y}}^k$  and  $\bar{\mathbf{q}}^k$  respectively. Thanks to that, the sequence of the dual iterates enables us to compute an  $\varepsilon$ -subgradient of  $f_2$  at the current primal iterate. In fact,  $\mathbf{y}_\ell^k \in \text{dom}(f_2^*)$  and thanks to Proposition 4.5 we have

$$\mathbf{y}_\ell^k \in \partial_{\varepsilon_{k,\ell}} f_2(\mathbf{x}_\ell^k), \quad \varepsilon_{k,\ell} = f_2(\mathbf{x}_\ell^k) - \langle \mathbf{y}_\ell^k, \mathbf{x}_\ell^k \rangle + f_2^*(\mathbf{y}_\ell^k)$$

with

$$\lim_{\ell \rightarrow \infty} \varepsilon_{k,\ell} = 0$$

Again, taking  $\mu_{k+1} > 0$  and  $\nu_{k+1} > 0$  for any  $k$ , there exists an index  $\mathcal{L}$  such that

$$\|\eta_{\mathcal{L}}^k\| \leq \mu_{k+1}, \quad \varepsilon_{k,\mathcal{L}} \leq \nu_{k+1}$$

with  $\eta_{\mathcal{L}}^k = \frac{1}{\beta} \mathbf{q}_{\mathcal{L}}^k + \mathbf{y}_{\mathcal{L}}^k - \xi^k$ . Finally, at the end of the  $k$ -th outer step we set

$$\begin{aligned} \mathbf{x}^{k+1} &= \mathbf{x}_{\mathcal{L}}^k & \mathbf{q}^{k+1} &= \mathbf{q}_{\mathcal{L}}^k \\ \xi^{k+1} &= \mathbf{y}_{\mathcal{L}}^k & \varepsilon_{k+1} &= \varepsilon_{k,\mathcal{L}} \end{aligned}$$

The following Lemma enables us to compute an  $\varepsilon$ -subgradient of  $Q_k$  at  $\bar{\mathbf{x}}^{k+1}$ .

**Lemma 4.1.** *Let  $f_1$  and  $f_2$  be nonnegative, proper, l.s.c. and convex functions, with  $\text{dom}(f_1) \subset \text{dom}(f_2)$  and the relative interiors of  $f_1$  and  $f_2$  have at least a point in common. If  $\mathbf{q}^{k+1} \in \partial f_1(\mathbf{x}^{k+1})$  and  $\xi^{k+1} \in \partial_{\varepsilon_{k+1}} f_2(\mathbf{x}^{k+1})$ , then the following vector*

$$\eta^{k+1} = \frac{1}{\beta} \mathbf{q}^{k+1} + \xi^{k+1} - \xi^k \tag{4.30}$$

is an  $\varepsilon_{k+1}$ -subgradient of  $Q_k$  at  $\mathbf{x}^{k+1}$ , that is  $\eta^{k+1} \in \partial_{\varepsilon_{k+1}} Q_k(\mathbf{x}^{k+1}, \xi^k)$ .

*Proof.* For the convexity of  $f_1(x)$  and the definition of  $\xi^{k+1}$  as  $\varepsilon_{k+1}$ -subgradient of  $f_2$  at  $\mathbf{x}^{k+1}$ , for any  $\mathbf{x} \in \mathbb{R}^n$  we have

$$\begin{aligned} & Q_k(\mathbf{x}^{k+1}, \xi^k) + \langle \eta^{k+1}, \mathbf{x} - \mathbf{x}^{k+1} \rangle \\ &= \frac{1}{\beta} f_1(\mathbf{x}^{k+1}) + f_2(\mathbf{x}^{k+1}) - f_2(\mathbf{x}^k) - \langle \xi^k, \mathbf{x}^{k+1} - \mathbf{x}^k \rangle + \varepsilon_k + \\ &+ \langle \frac{1}{\beta} \mathbf{q}^{k+1} + \xi^{k+1} - \xi^k, \mathbf{x} - \mathbf{x}^{k+1} \rangle \\ &\leq \frac{1}{\beta} f_1(\mathbf{x}) + f_2(\mathbf{x}) + \varepsilon_{k+1} - f_2(\mathbf{x}^k) - \langle \xi^k, \mathbf{x} - \mathbf{x}^k \rangle + \varepsilon_k \\ &= Q_k(\mathbf{x}, \xi^k) + \varepsilon_{k+1} \end{aligned}$$

Then  $\eta^{k+1} \in \partial_{\varepsilon_{k+1}} Q_k(\mathbf{x}^{k+1}, \xi^k)$ . □

For the sequence  $\{f_1(\mathbf{x}^{k+1})\}$  generated by the inexact scheme, the monotonicity property is replaced by

$$\begin{aligned} \frac{1}{\beta} f_1(\mathbf{x}^{k+1}) &\leq Q_k(\mathbf{x}^{k+1}, \xi^k) \\ &= \frac{1}{\beta} f_1(\mathbf{x}^{k+1}) + \Delta_{f_2}^{\xi^k}(\mathbf{x}^{k+1}, \mathbf{x}^k) \\ &\leq \frac{1}{\beta} f_1(\mathbf{x}^k) + \varepsilon_k \end{aligned} \tag{4.31}$$

where the first inequality follows from the nonnegativity of  $\Delta_{f_2}^{\xi^k}(\mathbf{x}^{k+1}, \mathbf{x}^k)$ . Obviously, when  $f_2$  is differentiable,  $\varepsilon_k = 0$  and the monotonicity property is preserved. From the last inequality in (4.31), we have

$$-\varepsilon_k \leq \Delta_{f_2}^{\xi^k}(\mathbf{x}^{k+1}, \mathbf{x}^k) - \varepsilon_k \leq \frac{1}{\beta}(f_1(\mathbf{x}^k) - f_1(\mathbf{x}^{k+1}))$$

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

$$-\sum_{i=1}^{k-1} i\varepsilon_i \leq \frac{1}{\beta} \sum_{i=1}^{k-1} f_1(\mathbf{x}^i) - \frac{k-1}{\beta} f_1(\mathbf{x}^k) \quad (4.32)$$

Furthermore, in view of (4.30), we obtain that (4.7) can be restated as follows:

$$\begin{aligned} \Delta_{f_2}^{\xi^k}(\mathbf{x}, \mathbf{x}^k) + \Delta_{f_2}^{\xi^{k-1}}(\mathbf{x}^k, \mathbf{x}^{k-1}) + \frac{1}{\beta} f_1(\mathbf{x}^k) &\leq \\ &\leq \frac{1}{\beta} f_1(\mathbf{x}) + \Delta_{f_2}^{\xi^{k-1}}(\mathbf{x}, \mathbf{x}^{k-1}) + \\ &\quad + \langle \eta^k, \mathbf{x}^k - \mathbf{x} \rangle + \varepsilon_k \end{aligned} \quad (4.33)$$

for any  $\mathbf{x}$  such that  $f_2(\mathbf{x}) < \infty$ . 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 4.6.** *Let  $f_1$  and  $f_2$  be nonnegative, proper, l.s.c. and convex functions, with  $\text{dom}(f_1) \subset \text{dom}(f_2)$  and the relative interiors of  $f_1$  and  $f_2$  have at least a point in common. We assume that, for any  $k$ , there exists a minimizer of the subproblem (4.27) and that  $\tilde{\mathbf{x}}$  is a minimizer of  $f_1(x)$  such that  $f_2(\tilde{\mathbf{x}}) < \infty$ . If for any  $k \geq 0$  the inner solver determines*

$$\mathbf{x}^{k+1}, \quad \mathbf{q}^{k+1} \in \partial f_1(\mathbf{x}^{k+1}) \quad \text{and} \quad \xi^{k+1} \in \partial_{\varepsilon_{k+1}} f_2(\mathbf{x}^{k+1})$$

so that the following conditions on  $\eta^{k+1} = \frac{1}{\beta} \mathbf{q}^{k+1} + \xi^{k+1} - \xi^k$  and  $\varepsilon_{k+1}$  hold

$$\|\eta^{k+1}\| \leq \mu_{k+1} \quad \text{and} \quad \varepsilon_{k+1} \leq \nu_{k+1} \quad (4.34)$$

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

$$\Delta_{f_2}^{\xi^k}(\tilde{\mathbf{x}}, \mathbf{x}^k) \leq \Delta_{f_2}^{\xi^{k-1}}(\tilde{\mathbf{x}}, \mathbf{x}^{k-1}) + \langle \eta^k, \mathbf{x}^k - \tilde{\mathbf{x}} \rangle + \varepsilon_k \quad (4.35)$$

and

$$f_1(\mathbf{x}^k) \leq f_1(\tilde{\mathbf{x}}) + \frac{\beta}{k} \left( f_2(\tilde{\mathbf{x}}) - f_2(\mathbf{x}^0) + \sum_{i=1}^k \langle \eta^i, \mathbf{x}^i - \tilde{\mathbf{x}} \rangle + \sum_{i=1}^k (i+1)\varepsilon_i \right) \quad (4.36)$$

Moreover, if the level subsets of  $f_1$  are bounded, a limit point of the sequence  $\{\mathbf{x}^k\}$  is a minimizer of  $f_1$ ; if  $\tilde{\mathbf{x}}$  is the unique minimizer of  $f_1(x)$ , then  $\mathbf{x}^k \rightarrow \tilde{\mathbf{x}}$  as  $k \rightarrow \infty$ .

*Proof.* In view of (4.33) with  $\mathbf{x} = \tilde{\mathbf{x}}$ , since for any  $k \geq 0$  we have that  $\Delta_{f_2}^{\xi^k}(\mathbf{x}^{k+1}, \mathbf{x}^k) \geq 0$  and  $f_1(\mathbf{x}^k) - f_1(\tilde{\mathbf{x}}) \geq 0$ , the inequality (4.35) holds.

Summing up the inequalities (4.33) computed at  $\tilde{\mathbf{x}}$  related to the first  $k$  steps, we have:

$$\begin{aligned} & \Delta_{f_2}^{\xi^k}(\tilde{\mathbf{x}}, \mathbf{x}^k) + \sum_{i=1}^k \Delta_{f_2}^{\xi^{k-1}}(\mathbf{x}^i, \mathbf{x}^{i-1}) + \frac{1}{\beta} \sum_{i=1}^k f_1(\mathbf{x}^i) \leq \\ & \leq \frac{k}{\beta} f_1(\tilde{\mathbf{x}}) + (f_2(\tilde{\mathbf{x}}) - f_2(\mathbf{x}^0)) + \sum_{i=1}^k \langle \eta^i, \mathbf{x}^i - \tilde{\mathbf{x}} \rangle + \sum_{i=1}^k \varepsilon_i \end{aligned} \quad (4.37)$$

Combining (4.32) with (4.37), we obtain

$$\begin{aligned} & \frac{\beta}{k} \left( \Delta_{f_2}^{\xi^k}(\tilde{\mathbf{x}}, \mathbf{x}^k) + \sum_{i=1}^k \Delta_{f_2}^{\xi^{i-1}}(\mathbf{x}^i, \mathbf{x}^{i-1}) \right) + f_1(\mathbf{x}^k) \leq \\ & \leq f_1(\tilde{\mathbf{x}}) + \\ & + \frac{\beta}{k} \left( f_2(\tilde{\mathbf{x}}) - f_2(\mathbf{x}^0) + \sum_{i=1}^k \langle \eta^i, \mathbf{x}^i - \tilde{\mathbf{x}} \rangle + \sum_{i=1}^{k-1} (i+1)\varepsilon_i + \varepsilon_k \right) \end{aligned} \quad (4.38)$$

Since  $\Delta_{f_2}^{\xi^{i-1}}(\mathbf{x}^i, \mathbf{x}^{i-1}) \geq 0$  for any  $i$  and  $\Delta_{f_2}^{\xi^k}(\tilde{\mathbf{x}}, \mathbf{x}^k) \geq 0$ , we have that (4.36) follows. Furthermore, if we denote by  $D$  the diameter of the level set  $\{\mathbf{x} | f_1(\mathbf{x}) \leq f_1(\mathbf{x}^0)\}$ , by applying the Cauchy–Schwarz inequality and condition (4.34) to inequality (4.36), we obtain

$$f_1(\mathbf{x}^k) \leq f_1(\tilde{\mathbf{x}}) + \beta \frac{f_2(\tilde{\mathbf{x}}) - f_2(\mathbf{x}^0)}{k} + \frac{\beta}{k} \left( D \sum_{i=1}^k \mu_i + \sum_{i=1}^k (i+1)\nu_i \right) \quad (4.39)$$

Since the sequence  $\{\mathbf{x}^k\}$  is bounded, there exists a subsequence of  $\{\mathbf{x}^k\}$  convergent to a limit point  $\bar{\mathbf{x}}$ . Thanks to the fact  $D \sum_{i=1}^{\infty} \mu_i + \sum_{i=1}^{\infty} (i+1)\nu_i < \infty$ , in view of (4.39), we have  $f_1(\bar{\mathbf{x}}) \leq f_1(\tilde{\mathbf{x}})$  for  $k \rightarrow \infty$ . Then,  $\bar{\mathbf{x}}$  is a minimizer of  $f_1(x)$ . If  $\tilde{\mathbf{x}}$  is the unique minimizer of  $f_1(x)$ , then  $\mathbf{x}^k \rightarrow \tilde{\mathbf{x}}$  as  $k \rightarrow \infty$ .  $\square$

In Algorithm 14 we employ the above result to update the general scheme for the inexact Bregman procedure. The following Corollary extends the previous Proposition for a differentiable  $f_2$ .

**Corollary 4.1.** *Let  $f_2$  be a differentiable function. Under the same hypotheses of the previous proposition, if for any  $k \geq 0$  the inner solver determines*

$$\mathbf{x}^{k+1}, \quad \mathbf{q}^{k+1} \in \partial f_1(\mathbf{x}^{k+1}) \quad \text{and} \quad \xi^{k+1} = \nabla f_2(\mathbf{x}^{k+1})$$

*so that the following condition on  $\eta^{k+1} = \frac{1}{\beta} \mathbf{q}^{k+1} + \xi^{k+1} - \xi^k$  holds*

$$\|\eta^{k+1}\| \leq \mu_{k+1} \quad (4.40)$$

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

$$D_{f_2}^{\xi^k}(\tilde{\mathbf{x}}, \mathbf{x}^k) \leq D_{f_2}^{\xi^{k-1}}(\tilde{\mathbf{x}}, \mathbf{x}^{k-1}) + \langle \eta^{(k)}, \mathbf{x}^k - \tilde{\mathbf{x}} \rangle \quad (4.41)$$

and

$$f_1(\mathbf{x}^k) \leq f_1(\tilde{\mathbf{x}}) + \frac{\beta}{k} \left( f_2(\tilde{\mathbf{x}}) - f_2(\mathbf{x}^0) + \sum_{i=1}^k \langle \eta^i, \mathbf{x}^i - \tilde{\mathbf{x}} \rangle \right) \quad (4.42)$$

Moreover, if the level subsets of  $f_1$  are bounded, a limit point of the sequence  $\{\mathbf{x}^k\}$  is a minimizer of  $f_1$ ; if  $\tilde{\mathbf{x}}$  is the unique minimizer of  $f_1$ , then  $\mathbf{x}^k \rightarrow \tilde{\mathbf{x}}$  as  $k \rightarrow \infty$ .

---

**Algorithm 14: Inexact Bregman Procedure**

---

Choose  $\mathbf{x}^0$  such that  $\xi^0 = \mathbf{0} \in \partial f_2(\mathbf{x}^0)$ ,  $\varepsilon_0 = 0$ ,  $\beta > 0$ ; choose sequences  $\{\mu_k\}$  and  $\{\nu_k\}$  such that  $\sum_i \mu_i < \infty$  and  $\sum_i i\nu_i < \infty$

FOR  $k = 0, 1, 2, \dots$  DO THE FOLLOWING STEPS:

Determine by an iterative solver an approximate minimizer  $\mathbf{x}^{k+1}$  of  $Q_k(\mathbf{x}, \xi^k)$ , i.e.

$$\mathbf{x}^{k+1} \sim \min_{\mathbf{x}} Q_k(\mathbf{x}, \xi^k)$$

and the related  $\mathbf{q}^{k+1} \in \partial f_1(\mathbf{x}^{k+1})$  and  $\xi^{k+1} \in \partial_{\varepsilon_{k+1}} f_2(\mathbf{x}^{k+1})$  so that

$$\|\eta^{k+1}\| \leq \mu_{k+1} \quad \text{and} \quad \varepsilon_{k+1} \leq \nu_{k+1}$$

where  $\eta^{(k+1)} = \frac{1}{\beta} \mathbf{q}^{k+1} + \xi^{k+1} - \xi^k$

Terminate if a stopping criterion is satisfied

---

**Remark 4.4.** From a numerical point of view, a practical choice for the stopping criterion (4.34) (and consequently for (4.40)) is

$$\|\eta^{k+1}\| \leq \frac{c}{(k+1)^\alpha} \quad (4.43)$$

$$\varepsilon_{k+1} \leq \frac{d}{(k+1)^\vartheta} \quad (4.44)$$

for  $k \geq 0$ ,  $\alpha, \vartheta, c, d \in \mathbb{R}^+$  and  $\alpha > 1$ ,  $\vartheta > 2$ . 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  $\alpha$  and  $\vartheta$  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 we set  $c = \|\eta^{(1)}\|$ ,  $d = \varepsilon_1$ ,  $\vartheta = 2.1$ ; in this way, the only parameter to set is  $\alpha$ .

### 4.3.1 The inexact procedure for image restoration problems

We apply the procedure in image restoration framework, with the aim to solve

$$\min_{\mathbf{x}} \varphi_0(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta\varphi_1(\mathbf{x})$$

As in Section 4.2.1, the discrepancy  $\varphi_0$  plays the role of the penalty function and  $\varphi_1$  is the regularization term. For sake of completeness, we restate that our aim is to find a solution, namely  $\mathbf{x}^*$ , of the equation  $\varphi_0(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}) = 0$ , where we recall that  $\mathbf{g}$

is the noise-free data (here  $\mathbf{c} = \mathbf{g} - b\mathbf{g}\mathbf{1}$ ). We know that the minimization procedure reaches instead an element  $\tilde{\mathbf{x}}$  which is a solution of  $\varphi_0(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) = 0$ , where  $\mathbf{g}\mathbf{n}$  is the noisy data; then  $\mathbf{c} = \mathbf{g}\mathbf{n} - b\mathbf{g}\mathbf{1}$ . We assume that an estimate  $\gamma$  of the noise level is known (i.e.  $\varphi_0(\mathbf{H}\mathbf{x}^* + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) \leq \gamma$ ); hence, as long as the iterates make the function  $\varphi_0$  stay above the noise level, i.e.

$$\varphi_0(\mathbf{H}\mathbf{x}^k + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) \geq \gamma$$

from (4.33) with  $\mathbf{x} = \mathbf{x}^*$ , employing the Cauchy–Schwarz inequality and thanks to the boundedness of the level set of  $\varphi_0(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n})$ , we have

$$\begin{aligned} -\nu_k - \mu_k D + \Delta_{\varphi_1}^{\xi^k}(\mathbf{x}^*, \mathbf{x}^k) + \Delta_{\varphi_1}^{\xi^{k-1}}(\mathbf{x}^k, \mathbf{x}^{k-1}) + \frac{1}{\beta} \varphi_0(\mathbf{H}\mathbf{x}^k + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) &\leq \\ &\leq \frac{\gamma}{\beta} + \Delta_{\varphi_1}^{\xi^{k-1}}(\mathbf{x}^*, \mathbf{x}^{k-1}) \end{aligned}$$

Then a sufficient condition to assure a decreasing behaviour for the *inexact* Bregman distance  $\Delta_{\varphi_1}^{\xi^{k-1}}(\mathbf{x}^*, \mathbf{x}^{k-1})$  between the iterates and the object  $\mathbf{x}^*$  is that, while  $\varphi_0(\mathbf{H}\mathbf{x}^k + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) \geq \gamma$  the following inequality holds:

$$\begin{aligned} \frac{1}{\beta}(\varphi_0(\mathbf{H}\mathbf{x}^k + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) - \gamma) + \Delta_{\varphi_1}^{\xi^{k-1}}(\mathbf{x}^k, \mathbf{x}^{k-1}) &= Q_{k-1}(\mathbf{x}^k, \xi^{k-1}) - \frac{\gamma}{\beta} \\ &\geq \mu_k D + \nu_k \end{aligned} \quad (4.45)$$

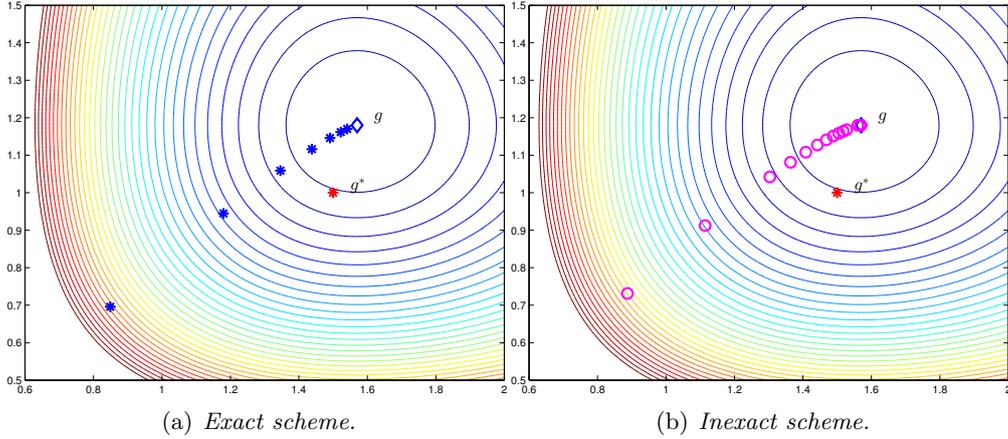
This condition depends on the *exactness* level required to the inner minimizer: the numerical experience shows that the term  $-\mu_k D - \nu_k$  is a pessimistic estimate of the absolute value of  $\langle \eta^k, \mathbf{x}^* - \mathbf{x}^k \rangle > -\varepsilon_k$ . This fact becomes clear from the numerical experiments: when  $\varphi_1$  is a differentiable function, the Bregman distance  $D_{\varphi_1}^{\mathbf{p}^k}(\mathbf{x}^*, \mathbf{x}^k)$  is a decreasing sequence until the relative reconstruction error decreases and sometimes even later. For a nondifferentiable  $\varphi_1$ , the inexact Bregman distance between the desired solution  $\mathbf{x}^*$  and the iterates  $\Delta_{\varphi_1}^{\xi^k}(\mathbf{x}^*, \mathbf{x}^k)$  is again a decreasing sequence with a very similar behaviour to the differentiable case.

Then, as for the exact Bregman method, a discrepancy criterion can provide a reasonable stopping rule. Nevertheless, the numerical experience shows that in general few iterations allow to observe the semi-convergence behaviour.

**Remark 4.5.** We had always assumed that the subproblems (4.4) and (4.27) have a solution. In presence of Poisson noise, under on the classical assumptions (1.5) on the imaging matrix  $\mathbf{H}$ , direct computation shows that the two variational subproblems related to  $Q_k(\mathbf{x}, \mathbf{p}^k)$  and  $\bar{Q}_k(\mathbf{x}, \mathbf{p}^k)$  are very similar to  $\min_{\mathbf{x} \geq 0} \varphi_\beta(\mathbf{x}) \equiv \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta\varphi_1(\mathbf{x})$ . Thus, the existence and the uniqueness of the subproblems' solution is assured by the Proposition 1.3 for standard regularization functions.

### 4.3.2 A simple 2–D example

This 2–dimensional example is devoted to clarify the regularization effect of both exact and inexact Bregman procedures when  $\varphi_0(\mathbf{x}; \mathbf{y}) = \mathbf{KL}(\mathbf{x}; \mathbf{y})$ , ( $\mathbf{H} = \mathbf{I}$  of order 2) and  $\varphi_1$  is the Tikhonov regularization. Here  $\mathbf{g}\mathbf{n} = (1.57, 1.18)^t$  is a perturbation of the exact data  $\mathbf{g}^* = (1.5, 1)^t$ . For this simple example, an explicit solution of the inner



**Figure 4.2:** Behaviour of the exact and inexact procedures. One can easily observe that the trajectories of both methods pass close to  $\mathbf{g}^*$  and then converge to the undesired solution  $\mathbf{g}$ : the problem is to devise a stopping criterion which forces the procedure to terminate itself near  $\mathbf{g}^*$ .

subproblems is available. In Figure 4.2 the behaviour of the procedures is shown. We can observe that the trajectories of both methods pass close to the desired solution  $\mathbf{x}^* = \mathbf{g}^*$  but actually the iterates converge to the perturbed data  $\mathbf{g}\mathbf{n}$ .

In Table 4.1 we report some numerical results:  $k$  is the outer iteration,  $\varphi_0(\mathbf{x}^k; \mathbf{g}\mathbf{n})$  is the Kullback–Leibler divergence from the given data,  $D_{\varphi_1}^{\mathbf{p}^k}(\mathbf{g}^*, \mathbf{x}^k)$  is the Bregman distance of iterates from the solution  $\mathbf{g}^*$  and, finally,  $\rho_k$  is the relative error on  $\mathbf{g}^*$ :

$$\rho_k = \frac{\|\mathbf{x}^k - \mathbf{g}^*\|}{\|\mathbf{g}^*\|}.$$

We assume as estimate  $\gamma$  of the noise the value  $\mathbf{KL}(\mathbf{g}\mathbf{n}; \mathbf{g}^*) = 0.016916$ : employing a discrepancy criterion for terminating the algorithm, it seems that the exact procedure has to be stopped at the 3rd iteration, where the value for  $\varphi_0$  is still above the noise level. Actually, if we let the procedure make a step further, the successive iteration provides a vector closer to the solution of the unperturbed problem. For the inexact procedure, we have the same behaviour: we should stop the procedure at the 5-th iteration, at which  $\varphi_0(\mathbf{x}^k; \mathbf{g}\mathbf{n}) > \gamma$ , but the next iterates are closer to the desired solution  $\mathbf{g}^*$ . Furthermore, the decreasing of the sequence  $D_{\varphi_1}^{\mathbf{p}^k}(\mathbf{g}^*, \mathbf{x}^k)$  is assured when the values of the fidelity function lie above the noise level, but it could happen that the sequence keep on decreasing for some further iterations.

Figure 4.2 describes the semiconvergence behaviour. Indeed the Bregman Procedure leads the iterates to the minimum of  $\varphi_0$ . The term of the exact or inexact Bregman distance in the objective function causes that the convergence curve towards the minimum of  $\varphi_0$  passes close to the real object  $\mathbf{g}^*$ . The introduction of the *inexact* distance induces a slowing down on the convergence rate of the procedure; this slowing down allows to obtain more control on the choice of the stopping iteration.

### 4.3.3 A simple 1–D example

We introduce a simple 1–dimensional example to analyze the behaviour of the inexact Bregman procedure with respect to the exact one. In this example, under suitable

$k$	$\varphi_0(\mathbf{x}^k; \mathbf{gn})$	$D_{\varphi_1}^{\mathbf{P}^k}(\mathbf{g}^*, \mathbf{x}^k)$	$\rho_k$	$k$	$\varphi_0(\mathbf{x}^k; \mathbf{gn})$	$D_{\varphi_1}^{\mathbf{P}^k}(\mathbf{g}^*, \mathbf{x}^k)$	$\rho_k$
Exact Procedure				Inexact Procedure			
1	0.383195	0.2580	0.3990	1	3.029703	1.1660	0.8470
2	0.085509	0.0528	0.1800	2	0.328989	0.2240	0.3710
3	0.024324	0.0136	0.0914	3	0.118477	0.0782	0.2190
4	0.007676	0.0087	0.0731	4	0.034295	0.0202	0.1110
5	0.002561	0.0107	0.0811	5	0.019239	0.0126	0.0879
				6	0.011214	0.0100	0.0786
				7	0.006711	0.0098	0.0775
				8	0.004093	0.0104	0.0802

**Table 4.1:** Numerical results for the 2d example.  $\varphi_1(x) = \frac{1}{2}\|x\|^2$ ,  $\mathbf{gn} = (1.57, 1.18)^t$ ,  $\mathbf{g} = (1.5, 1)^t$ .

hypothesis, a closed formula for the first two steps is available; this is the case of a denoising problem, regularized by  $\mathbf{TV}$  functional. Using Theorem 1 in [91], an explicit solution is easily computed; for sake of completeness, we report this theorem in a discrete framework.

**Proposition 4.1.** Suppose the vector  $\mathbf{f} \in \mathbb{R}^N$  is defined such that

$$\frac{1}{n_1} \sum_{i=1}^{n_1} \mathbf{f}_i = \psi_1, \quad \frac{1}{N-n_1} \sum_{i=n_1+1}^N \mathbf{f}_i = \psi_2$$

with  $\psi_1 > \psi_2$ ,  $1 < n_1 < N$  and  $\max_{i=n_1+1, \dots, N} \mathbf{f}_i \leq \min_{i=1, \dots, n_1} \mathbf{f}_i$ . If we assume that

$$\max_{i=n_1+1, \dots, N} \mathbf{f}_i \leq \psi_2 + \frac{\beta}{N-n_1} \leq \psi_1 - \frac{\beta}{n_1} \leq \min_{i=1, \dots, n_1} \mathbf{f}_i \quad (4.46)$$

then the unique minimizer to the minimization problem

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{x} - \mathbf{f}\|^2 + \beta \sum_{i=1}^{N-1} |\mathbf{x}_{i+1} - \mathbf{x}_i|$$

is given by

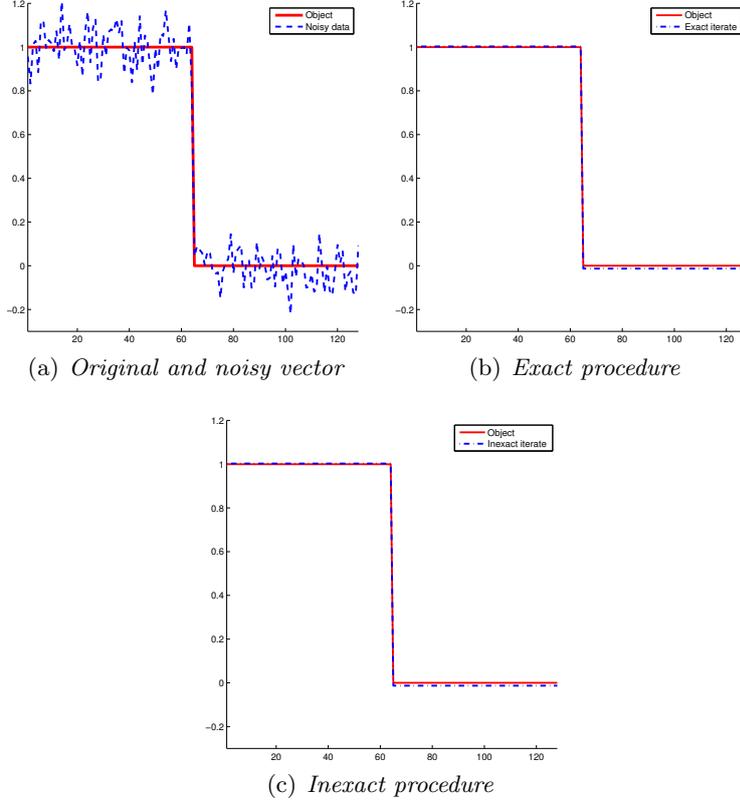
$$\widehat{\mathbf{x}}_i = \begin{cases} \psi_1 - \frac{\beta}{n_1} & i = 1, \dots, n_1 \\ \psi_2 + \frac{\beta}{N-n_1} & i = n_1 + 1, \dots, N \end{cases}$$

Let  $\mathbf{g}^*$  a vector in  $\mathbb{R}^N$  defined as

$$\mathbf{g}_i^* = \begin{cases} \gamma_1^* & i = 1, \dots, n_1 \\ \gamma_2^* & i = n_1 + 1, \dots, N \end{cases}$$

where we assume  $\gamma_1^* > \gamma_2^*$ . We denote  $n_2 = N - n_1$ . Perturbing the data with Gaussian noise with zero mean and standard deviation  $\sigma$ , we obtain a vector  $\mathbf{gn}$  such that  $\mathbf{gn} = \mathbf{g}^* + v$ , where  $v$  is the Gaussian noise. Denoting  $\gamma_1$  and  $\gamma_2$  as

$$\frac{1}{n_1} \sum_{i=1}^{n_1} \mathbf{gn}_i = \gamma_1 \quad \frac{1}{n_2} \sum_{i=n_1+1}^N \mathbf{gn}_i = \gamma_2$$



**Figure 4.3:** Test problem 1D: (a) plot of the original vector (solid line) and noisy vector (dotted line); (b) plot of the exact iterate  $\mathbf{x}^2$  of the Bregman procedure with respect to the original vector (solid line); (c) plot of the iterate  $\mathbf{x}^2$  of the inexact procedure combined with AEM (dashed line) with respect to the original vector (solid line).

we have  $\gamma_1 > \gamma_2$ . In Figure 4.3 (a) we show the original vector  $\mathbf{g}^*$  (solid line) and the noisy vector  $\mathbf{gn}$  (dotted line) for  $N = 128$ ,  $n_1 = n_2 = 64$ ,  $\gamma_1^* = 1$ ,  $\gamma_2^* = 0$ ,  $\gamma_1 = 1.0030$ ,  $\gamma_2 = -0.0129$ ,  $1/2\|\mathbf{g}^* - \mathbf{gn}\|^2 = 0.5052$ ,  $\sigma = 0.10$ .

In order to reconstruct the original signal from the noise data  $\mathbf{gn}$ , i.e. to obtain an estimate of  $\mathbf{g}^*$ , we consider the following variational problem

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{x} - \mathbf{gn}\|^2 + \beta \sum_{i=1}^{N-1} |\mathbf{x}_{i+1} - \mathbf{x}_i| \quad (4.47)$$

where  $\varphi_0(\mathbf{x}) = \frac{1}{2}\|\mathbf{x} - \mathbf{gn}\|^2$  and  $\varphi_1(\mathbf{x}) = \sum_{i=1}^{N-1} |\mathbf{x}_{i+1} - \mathbf{x}_i|$ . The exact Bregman procedure consists in solving a sequence of subproblems

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{x} - (\mathbf{gn} + \beta\mathbf{p}^k)\|^2 + \beta \sum_{i=1}^{N-1} |\mathbf{x}_{i+1} - \mathbf{x}_i|$$

provided  $\mathbf{p}^0 = \mathbf{0}$ . If the parameter  $\beta$  satisfies

$$2 \max \left\{ n_2 \left( \max_{i=n_1+1, N} \mathbf{gn}_i - \gamma_2 \right), n_1 \left( \gamma_1 - \min_{i=1, n_1} \mathbf{gn}_i \right) \right\} \leq \beta \leq \frac{(\gamma_1 - \gamma_2)n_1n_2}{N} \quad (4.48)$$

then, the first two step of the Bregman iteration can be computed by the closed formula provided in Proposition 4.1. At the first iteration we have

$$\mathbf{x}_i^1 = \begin{cases} \gamma_1 - \frac{\beta}{n_1} & i = 1, \dots, n_1 \\ \gamma_2 + \frac{\beta}{n_2} & i = n_1 + 1, \dots, N \end{cases}$$

$$\mathbf{p}_i^1 = \mathbf{p}_i^0 - \frac{1}{\beta} (\mathbf{x}_i^1 - \mathbf{g}\mathbf{n}_i) = \begin{cases} -\frac{1}{\beta}(\gamma_1 - \frac{\beta}{n_1} - \mathbf{g}\mathbf{n}_i) & i = 1, \dots, n_1 \\ -\frac{1}{\beta}(\gamma_2 + \frac{\beta}{n_2} - \mathbf{g}\mathbf{n}_i) & i = n_1 + 1, \dots, N \end{cases}$$

Hence, the fit to data value is

$$\varphi_0(\mathbf{x}^1) = \frac{1}{2} \left( \sum_{i=1}^{n_1} (\gamma_1 - \mathbf{g}\mathbf{n}_i)^2 + \sum_{i=n_1+1}^N (\gamma_2 - \mathbf{g}\mathbf{n}_i)^2 + \frac{\beta^2}{n_1} + \frac{\beta^2}{n_2} \right)$$

and moreover

$$\begin{aligned} D_{\mathbf{TV}}^{\mathbf{p}^1}(\mathbf{g}^*, \mathbf{x}^1) &= \sum_{i=1}^{N-1} |\mathbf{g}_{i+1}^* - \mathbf{g}_i^*| - \sum_{i=1}^{N-1} |\mathbf{x}_{i+1}^1 - \mathbf{x}_i^1| - \langle \mathbf{p}^1, \mathbf{g}^* - \mathbf{x}^1 \rangle \\ &= |\gamma_2^* - \gamma_1^*| - \left| \gamma_2 + \frac{\beta}{n_2} - \gamma_1 + \frac{\beta}{n_1} \right| \\ &\quad - \sum_{i=1}^{n_1} -\frac{1}{\beta} (\gamma_1 - \frac{\beta}{n_1} - \mathbf{g}\mathbf{n}_i) (\mathbf{g}_i^* - \mathbf{x}_i^1) \\ &\quad - \sum_{i=n_1+1}^{n_2} -\frac{1}{\beta} (\gamma_2 + \frac{\beta}{n_2} - \mathbf{g}\mathbf{n}_i) (\mathbf{g}_i^* - \mathbf{x}_i^1) \\ &= |\gamma_2^* - \gamma_1^*| - \left| \gamma_2 + \frac{\beta}{n_2} - \gamma_1 + \frac{\beta}{n_1} \right| \\ &\quad - \left( -\frac{n_1\gamma_1}{\beta} + 1 + \frac{n_1\gamma_1}{\beta} \right) \left( \gamma_1^* - \gamma_1 + \frac{\beta}{n_1} \right) \\ &\quad - \left( -\frac{n_2\gamma_2}{\beta} - 1 + \frac{n_2\gamma_2}{\beta} \right) \left( \gamma_2^* - \gamma_2 - \frac{\beta}{n_2} \right) \\ &= \gamma_1^* - \gamma_2^* - \gamma_1 + \gamma_2 + \frac{\beta}{n_1} + \frac{\beta}{n_2} \\ &\quad - \gamma_1^* + \gamma_1 - \frac{\beta}{n_1} + \gamma_2^* - \gamma_2 - \frac{\beta}{n_2} \\ &= 0 \end{aligned}$$

At the second step, we apply the Proposition 4.1 with  $f = \mathbf{g}\mathbf{n}^1 = \mathbf{g}\mathbf{n} + \beta\mathbf{p}^1$ ; we have

$$\frac{1}{n_1} \sum_{i=1}^{n_1} \mathbf{g}\mathbf{n}_i^1 = \gamma_1 + \frac{\beta}{n_1} \quad \frac{1}{n_2} \sum_{i=n_1+1}^N \mathbf{g}\mathbf{n}_i^1 = \gamma_2 + \frac{\beta}{n_2}$$

thanks to (4.48) and remembering  $\gamma_1 > \gamma_2$ ,  $\mathbf{g}\mathbf{n}^1$  satisfies the assumptions of Proposition 4.1; then, the second exact iterate assumes the form

$$\mathbf{x}_i^2 = \begin{cases} \gamma_1 & i = 1, \dots, n_1 \\ \gamma_2 & i = n_1 + 1, \dots, N \end{cases} \quad (4.49)$$

Furthermore, we have

$$\varphi_0(\mathbf{x}^2) = \frac{1}{2} \left( \sum_{i=1}^{n_1} (\gamma_1 - \mathbf{g}\mathbf{n}_i)^2 + \sum_{i=n_1+1}^N (\gamma_2 - \mathbf{g}\mathbf{n}_i)^2 \right)$$

and the Bregman distance is

$$D_{\mathbf{T}\mathbf{V}}^{\mathbf{p}^1}(\mathbf{g}^*, \mathbf{x}^2) = 0$$

indeed we we have taken into account that

$$\mathbf{p}^2 = \mathbf{p}^1 - \frac{1}{\beta} (\mathbf{x}^2 - \mathbf{g}\mathbf{n})$$

Then  $\mathbf{x}^2$  can be considered as an approximation of the original vector. Figure 4.3 (b)

exact version with closed formulae									
$k$	$\rho_k$	$\varphi_0$	$D_{\varphi_1}^{\mathbf{p}^k}$						
1	0.62860	13.6402	0						
2	0.01325	0.49955	0						
exact version with AEM $tol = 10^{-6}$					inexact version with AEM $tol = 10^{-3}, \alpha = 1.5$				
$k$	$\rho_k$	$it$	$\varphi_0$	$D_{\varphi_1}^{\mathbf{p}^k}$	$k$	$\rho_k$	$it$	$\varphi_0$	$\Delta_{\varphi_1}^{\xi^k}$
1	0.62960	830	13.64	$-8.1 \cdot 10^{-5}$	1	0.60260	239	12.59	$5.3 \cdot 10^{-16}$
2	0.01325	845	0.500	$-1.1 \cdot 10^{-4}$	2	0.01326	623	0.500	$3.5 \cdot 10^{-16}$
3	0.01923	816	0.484	$-1.7 \cdot 10^{-4}$	3	0.02350	697	0.481	$7.0 \cdot 10^{-17}$
exact version with CP $tol = 10^{-8}$					inexact version with CP $tol = 10^{-6}, \alpha = 1.5$				
1	0.6289	61742	13.6126	$9.33 \cdot 10^{-4}$	1	0.6277	22392	13.56	$2.0 \cdot 10^{-10}$
2	0.01304	52678	0.49956	$-1.01 \cdot 10^{-5}$	2	0.01379	34950	0.4996	$1.0 \cdot 10^{-16}$
3	0.02005	46246	0.48373	$-1.61 \cdot 10^{-5}$	3	0.02013	38485	0.4839	$8.0 \cdot 10^{-16}$

**Table 4.2:** 1D Test problem: results of different version of the Bregman iteration with  $\beta = 29$ .  $\varphi_0$  stands for  $\varphi_0(\mathbf{x}^k; \mathbf{g}\mathbf{n})$  and  $D_{\varphi_1}^{\mathbf{p}^k}$  and  $\Delta_{\varphi_1}^{\xi^k}$  for the exact and inexact Bregman distance, respectively, of  $\mathbf{g}^*$  from  $\mathbf{x}^k$

shows the exact iterate  $x^2$  related to the noisy problem in Figure 4.3 (a). In this case the value of  $\beta$  satisfying (4.48) is 29 which is an overestimate of an optimal value of the regularization parameter. Table 4.2 shows the results obtained for this test problem in three different cases:

- ▷ exact scheme with closed formulae

- ▷ exact scheme with an inner solver
- ▷ inexact scheme

Here  $\rho_k$  is the relative reconstruction error with respect to the Euclidean norm,  $it$  denotes the number of iterations of the inner solver for each  $k$ -th outer iteration.

As inner solvers we employ the algorithm of Chambolle & Pock and the AEM method (see Chapter 2 for the details); for the exact version the stopping rule of both is based on the standard relative difference of the primal–dual iterates  $(\mathbf{x}^i, \mathbf{y}^i)$  and  $(\mathbf{x}^{i+1}, \mathbf{y}^{i+1})$ , i.e.

$$\left\| \begin{pmatrix} \mathbf{x}^{i+1} \\ \mathbf{y}^{i+1} \end{pmatrix} - \begin{pmatrix} \mathbf{x}^i \\ \mathbf{y}^i \end{pmatrix} \right\| / \left\| \begin{pmatrix} \mathbf{x}^{i+1} \\ \mathbf{y}^{i+1} \end{pmatrix} \right\| \leq tol$$

For the inexact version, during the first iteration the standard stopping rule is implemented; for the successive iterations the rules (4.34) and (4.40) are used, setting  $c = \|\eta_1\|$  and  $d = \varepsilon_1$  as previously described.

We observe that when an inner iterative method is used in the exact scheme, the computed Bregman distances can assume negative values, since the updating 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.



## Chapter 5

# Numerical Experiments and Applications

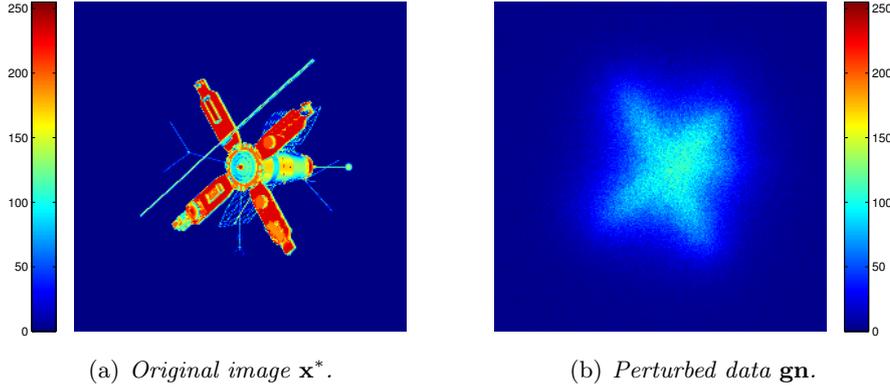
This chapter is devoted to evaluate the effectiveness of the optimization techniques described in the previous sections for the regularization of data corrupted by Poisson noise. First of all we compare the behaviour of the exact version and the inexact version of the Bregman iteration, showing that the second one is more effective in terms of computational cost. The second set of experiments is aimed to show that the inexact procedure provides also a better reconstruction in terms of contrast enhancement, specially in denoising problems. Furthermore, on some test problems, we compare the methods for the estimate of the regularization parameter  $\beta$  with the inexact Bregman procedure. These techniques provide in general high quality reconstructions, but for some cases, as for example images resulting from low counts, the Bregman iteration allows us to obtain better reconstructions. Furthermore, in the last part, we present an astronomical imaging problem, where only the Bregman procedure allows to obtain meaningful results. In this application the main issue is to reconstruct an image containing diffuse structures lying around point sources: due to their very high intensities, the action of the PSF destroys all the details around these sources. The contrast enhancement given by the Bregman procedure allows to obtain a good reconstruction of both the point sources and the diffuse component.

### 5.1 Exact versus inexact procedure

In the first set of experiments, we compare the effectiveness of the Bregman procedure for the regularization of Poisson data with respect to the inexact version, showing that the inexact version appears more promising than the exact one for what regards the efficiency.

The data on which we perform the numerical tests are presented in Figure 5.1. The original image is a  $256 \times 256$  image with sharp details, with values lying in  $[0, 255]$  and the constant background value  $bg$  is equal to 1; the PSF used simulates the one taken by a ground based telescope, and one can download it from [www.mathcs.emory.edu/nagy/RestoreTools/index.html](http://www.mathcs.emory.edu/nagy/RestoreTools/index.html): it is the one shown in Figure 1.1(a). The noise affecting the blurred image is Poisson noise. We refer to this problem as *spacecraft* problem.

In order to restore *spacecraft*, we consider the minimization of the **KL** functional combined with the **HS** regularization and with the **TV** regularization. We will denote the former one as KL-HS problem and the latter with KL-TV problem. For both models, the value of  $\beta$  is obtained by trial with several runs. We choose the value corresponding to the best reconstruction error on the original image  $\mathbf{x}^*$ . Such a value is denoted by  $\beta_{\text{opt}}$ , and it is  $1.63 \cdot 10^{-3}$ .



**Figure 5.1:** Data set for spacecraft problem. In figure (a) the original image is shown; one can observe the presence of sharp details and straight lines in the object: these information suggest to employ **HS** (or **TV**) regularization functional. In Figure (b)  $\mathbf{gn}$  is shown: the noise is predominant, and the blur generated by the PSF is very strong. Both images range in  $[0, 255]$ .

The first experiment, which will be used as benchmark, consists in solving **KL** +  $\beta_{\text{opt}}$ **HS** and **KL** +  $\beta_{\text{opt}}$ **TV** models where the minimization procedure is performed using SGP and AEM respectively (see Sections 2.1 and 2.3 for the description of the algorithms). SGP is stopped when the relative difference between two successive values of the of the objective function is less than a prefixed tolerance  $\tau_{SGP} = 10^{-7}$  and the mean of this difference over the last ten iteration is less than  $10\tau_{SGP}$ . With this setting, SGP stops after 497 iterations, giving a relative reconstruction error equal to 0.36527. For the problem with Total Variation regularization, AEM is used; the stopping criterion is satisfied when the relative difference in Euclidean norm of two successive iteration is less than  $\tau_{AEM} = 4 \cdot 10^{-5}$ . The solution is obtained after 2042 iterations with a reconstruction error of 0.36967. The results are summed up in Table 5.1.

SGP		AEM	
$\rho$	Iters	$\rho$	Iters
0.36527	497	0.36967	2042

**Table 5.1:** Results obtained by solving the problems **KL** +  $\beta_{\text{opt}}$ **HS** (SGP) and **KL** +  $\beta_{\text{opt}}$ **TV** (AEM).

We apply the exact and the inexact Bregman procedures to problems (KL-HS and KL-TV). The setup for the exact procedure consists in setting  $\beta = 10\beta_{\text{opt}}$  and

$\mathbf{x}^0 = 1/N \sum_i \mathbf{g}_i - b\mathbf{g}$  as the starting point. As stopping criteria for the iterative solvers of the inner subproblems we use the standard ones previously described, with severe tolerance  $\tau_{\text{SGP}} = 10^{-10}$  and  $\tau_{\text{AEM}} = 10^{-5}$  for SGP and AEM respectively. These severe tolerances assure that the rule (4.5) gives  $\mathbf{p}^{k+1} \in \partial\varphi_1(\mathbf{x}^{k+1})$ : if this does not occur, the Bregman Distance  $D_{\varphi_1}^{\mathbf{p}^{k+1}}(\mathbf{x}, \mathbf{x}^{k+1})$  could be negative and inequality (4.6) does not hold.

The inexact version employs the stopping criterion (4.34) for AEM and (4.40) for SGP, respectively. The sequences  $\nu_k$  and  $\mu_k$  are set up as described in Remark 4.4: the first subproblem  $Q_1$  is solved by using standard stopping rules with moderate tolerances ( $\tau_{\text{SGP}} = 10^{-7}$ ,  $\tau_{\text{AEM}} = 4 \cdot 10^{-5}$ , respectively). Then we set  $c = \|\eta_1\|$  and  $d = \varepsilon_1$ ; moreover, we fix  $\vartheta = 2.1$  and  $\alpha = 1.5$ .

	Exact iterative procedure-SGP			Inexact iterative procedure-SGP		
$k$	$\rho_k$	$it$	time	$\rho_k$	$it$	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	<u>0.3644</u>	2609	831.1	0.3697	3763	572.3
6	0.3655	3073	933.9	<u>0.3681</u>	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 5.2:** spacecraft *deblurring test problem*: exact and inexact iterative methods for the KL-HS problem ( $\delta = 0.0134$ ), using SGP as inner solver. In the exact version  $\tau_{\text{SGP}} = 10^{-10}$ ; in the inexact version, in the first outer iteration  $\tau_{\text{SGP}} = 10^{-7}$ , while in the subsequent iterations (4.40) is used as stopping rule with  $c = 5.38$  and  $\alpha = 1.5$ .  $it$  denotes the number of the iterations of the inner solver at the  $k$ -th outer step, while  $time$  denotes the execution time in seconds at the end of the current outer iteration.

In Tables 5.2 and 5.3 the different behaviours of the two procedures are shown. For each outer iteration  $k$  we report the number  $it$  of iterations of the inner method, the execution time (in seconds) at the end of the  $k$ -th outer iteration and finally the relative reconstruction error  $\rho_k = \|\mathbf{x}^k - \mathbf{x}^*\|/\|\mathbf{x}^*\|$ .

For the KL-HS problem, we can observe (Table 5.2) that the exact procedure is able to reach the minimum reconstruction error within 5 outer iterations, while the inexact one needs 6 iterations: although this difference, the inexact procedure permits to obtain a restored image in a shorter time. Indeed, the total number of iterations of the inner solver is 24809 for the exact procedure, while the inexact one needs 18713 total iterations.

The behaviour of the two procedures for the problem with **TV** regularization is similar to the previous case (Table 5.3): the exact procedure allows to obtain the restored image in 5 outer iterations, with a reconstruction error of 0.36414 and a total number of 23077 iterations of the inner solver, while the inexact Bregman procedure needs 5 outer iterations too, but the number of inner iterations is 9879 and it provides a reconstruction error of 0.36682. Comparing the results obtained with the Bregman procedure and the standard algorithms (AEM and SGP), we can observe that the reconstruction error is similar: however it is worth to notice that the  $\beta_{\text{opt}}$  actually *is not available*. In

$k$	Exact iterative procedure-AEM			Inexact iterative procedure-AEM		
	$\rho_k$	$it$	time	$\rho_k$	$it$	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	<u>0.36414</u>	3363	910.9	<u>0.36682</u>	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

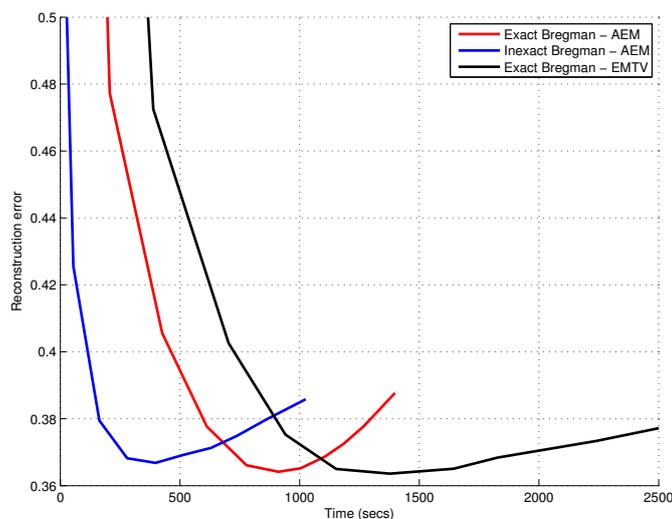
**Table 5.3:** spacecraft *deblurring test problem*: exact and inexact iterative methods for the KL-TV problem, using AEM as inner solver. In the exact version  $\tau_{\text{AEM}} = 10^{-5}$ ; in the inexact version, in the first outer iteration  $\tau_{\text{AEM}} = 5 \cdot 10^{-4}$  while in the subsequent iterations (4.34) are used as stopping rules with  $c = 19.3$ ,  $d = 65.4$ ,  $\vartheta = 2.1$  and  $\alpha = 1.5$ .  $it$  denotes the number of the iterations of the inner solver at the  $k$ -th outer step, while time denotes the execution time in seconds at the end of the current outer iteration.

the Bregman procedure the regularization parameter is *ten times* the optimal one, and, nevertheless, one can obtain an image very similar to that provided by standard methods. Moreover, this overestimation of the regularization parameter has an enhancing contrast effect as it will shown in the forthcoming sections.

In Figure 5.2 we compare the behaviour of the *relative* reconstruction error  $\rho_k$  with respect to the execution time of three different approaches for solving the KL-TV method:

- ▷ exact version of the Bregman Procedure, with AEM as inner solver;
- ▷ inexact version of the Bregman Procedure, with AEM as inner solver;
- ▷ exact version of the Bregman Procedure, with EM-TV as inner solver [29].

We point out that the Bregman Procedure combined with the EM-TV method has a complex structure, since it consists in three nested iterative methods (see section 2.2) incorporated one in each other. For more details about this implementation, see [29]. For any  $k$ -th outer iteration, we execute 2000 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 (4.6) is verified, we have determined experimentally a minimum prefixed number of inner iterations assuring an approximately correct behaviour for the KL and the objective function of the subproblems. Furthermore the inner step of EM-TV solver uses the Chambolle method [36], which is stopped when the maximum difference between two successive dual iterates is less than  $10^{-2}$ . This comparison shows that one has to choose wisely the inner solver, because the use of an expensive inner solver in terms of computational cost could make loose the advantages of the whole procedure. Moreover, we can observe that the inner stopping rules (4.34) and (4.40) improve the efficiency of the procedure without degrading its features.



**Figure 5.2:** Reconstruction error versus running time. The black line represents the exact procedure with EM-TV as inner solver, the blue line is the inexact procedure with AEM as inner solver and finally the red line represents the exact procedure with AEM as inner.

$\alpha$	$k$	$\rho_k$	$cum-it$	$\alpha$	$k$	$\rho_k$	$cum-it$
	KL-HS problem				KL-TV problem		
1.2	7	0.37039	20849	1.5	5	0.36682	9879
1.5	6	0.36809	18713	3	5	0.36721	12196
1.7	6	0.36736	22030	4	4	0.36665	16634

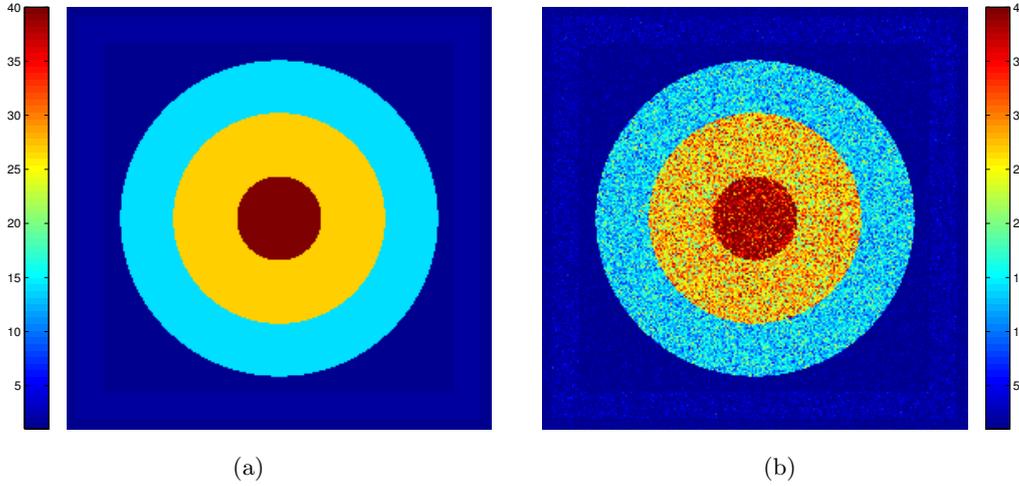
**Table 5.4:** spacecraft deblurring test problem: results obtained by inexact procedure with different values of  $\alpha$  in the inner stopping rule. For the KL-HS problem, the inner solver is SGP, while for KL-TV problem, AEM is used.

To investigate more deeply the meaning of the new stopping rules (4.34) and (4.40), we solve *spacecraft*, with different value of  $\alpha$  in the stopping rule of the inner solver; the other parameters are set as specified for the results shown in Tables 5.2 and 5.3. The parameter  $\alpha$  has the role of adjusting the accuracy needed for the solution of each subproblem; in Table 5.4 we report the results of these tests. For each one, we report the outer iteration  $k$  corresponding to the minimum reconstruction error  $\rho_k$ , the value of  $\rho_k$  and the cumulative number  $cum-it$  of the iterations of the inner solver. We can conclude that a suitable choice for  $\alpha$  is the one corresponding to a not-too-much severe accuracy after the first five to six outer iterations; in any case, the inexact procedure appears more efficient in terms of inner iterations.

## 5.2 Denoising

To evaluate the performance of the inexact iterative procedure for a denoising problem we use the data set *LCR phantom* [69], a piecewise constant object frequently used in literature (see for example [101]). It consists in concentric circles of intensities 14, 27, 40, enclosed within a square frame of intensity 2; the recorded data are affected by Pois-

son noise and the background emission term is 0. The relative error in Euclidean norm is 0.21273: see Figure 5.3 for visual inspection. In order to restore this image, we use



**Figure 5.3:** LCR phantom: on the left the original image is presented, while on the right the noisy data are shown. The relative error in Euclidean norm is 0.21273, and the background emission is 0.

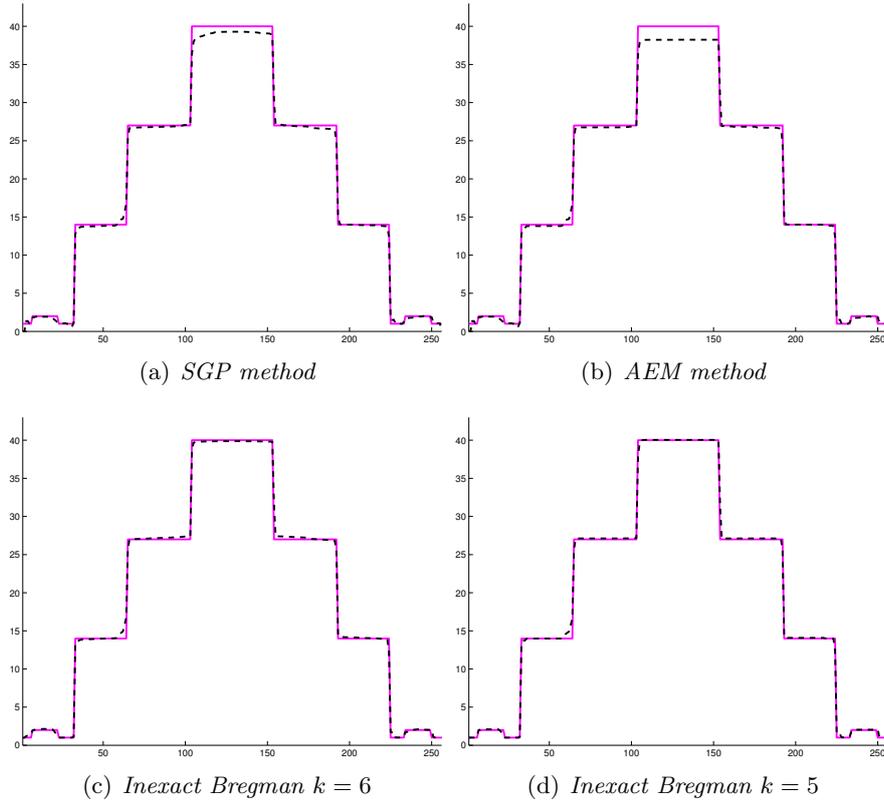
both problems KL-HS and KL-TV; the former is solved by using SGP with an optimal parameter obtained experimentally  $\beta_{\text{opt}} = 0.575$  (see [20]). In this case, the minimum  $\rho_k$  is 0.04231 and the total number of iterations is 1034 ( $\delta = 10^{-3}$ ,  $\tau_{\text{SGP}} = 10^{-8}$ ). The KL-TV problem is solved by AEM method, with the same value for  $\beta_{\text{opt}}$ , obtaining a reconstruction error of 0.04477 and 2536 total iterations ( $\tau_{\text{AEM}} = 10^{-6}$ ).

In Table 5.5 we report the results obtained by solving both problems by the inexact

$k$	KL-HS			KL-TV		
	$D_{\text{HS}}^{\text{P}^k}(x^*, \mathbf{x}^k)$	$\rho_k$	$it$	$\Delta_{\text{TV}}^{\xi^k}(x^*, \mathbf{x}^k)$	$\rho_k$	$it$
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	<u>0.03654</u>	2264
6	998.5	<u>0.03640</u>	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

**Table 5.5:** Inexact iterative method for the denoising problem related to LCR phantom

iterative method: we show the relative reconstruction error, the exact or inexact Bregman distance between the original image and the current iterate and the number of inner iterations at each  $k$ -th outer iteration. In both cases,  $\beta$  is an overestimation of the optimal parameter and it is equal to  $10\beta_{\text{opt}}$ . The stopping tolerances for the first outer iteration is  $\tau_{\text{SGP}} = 10^{-5}$  for SGP and  $\tau_{\text{AEM}} = 10^{-4}$  for AEM. We observe also for this test problem the typical semiconvergence behaviour. The minimum reconstruction

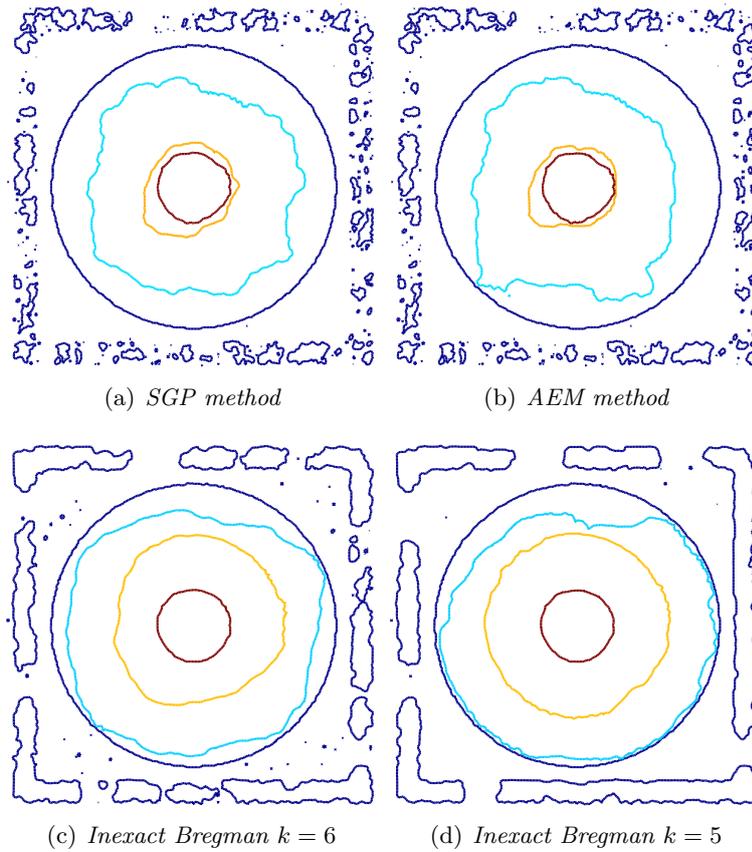


**Figure 5.4:** LCR phantom: superposition of the line-outs from row number 128 for the KL-HS (left-panels) and KL-TV (right-panels) restored images. In the upper panels SGP and AEM are used with an optimal regularization parameter (dashed lines), while the lower panels show the images obtained by inexact iterative method at the iteration with the minimum reconstruction error (dashed lines). The solid line is the row number 128 of the original image.

error is obtained at the 6-th iteration for the problem with Hypersurface regularization, while for Total Variation functional  $\rho_k$  reaches its minimum at  $k = 5$ . Moreover, the Bregman distances, both exact and inexact, reach the minimum at the same outer iteration ( $k = 6$ ).

In Figure 5.4, we show the superposition of the line-outs from row number 128 for KL-HS reconstructions obtained by SGP and by the inexact method at the iteration  $k = 6$  (minimum reconstruction error) (panels (a) and (c)). The solid line is the 128-th row of the original image. Figure 5.4(b) shows the superposition of the line-outs from row number 128 for KL-TV reconstructions obtained by AEM and in 5.4(d) by the inexact method at the iteration  $k = 5$ . We observe that for both problems, the reconstruction obtained by the inexact iterative method is able to reach the level 40 in the central pixels of the original image. This level is underestimated when we solve the KL-HS and KL-TV problems with the optimal regularization parameter.

In Figure 5.5(a) we show the contour plot of the restored image obtained by solving the KL-HS problem with SGP, while in 5.5(c) we present the contour plot of the result obtained by the inexact procedure at the 6-th outer iteration. The images on the right panels of Figure 5.5 are related to the problem involving the Total Variation function:



**Figure 5.5:** 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]).

in 5.5(b) the result obtained by the AEM method is shown while in 5.5(d) the contour plot of the images obtained by the inexact Bregman procedure at the 5-th outer iteration is presented.

The images in the first row of Figure 5.5 have been plotted setting the contour levels to (2, 14, 27, 38), while the images related to the inexact procedure are plotted using levels (2, 14, 27, 40), since in the former cases the top level of the original image is not reached. Figures 5.4 and 5.5 show that the inexact iterative method provides a contrast enhancement in 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. This contrast enhancement is discussed in [100, 28].

### 5.3 Comparison between the methods for $\beta$ estimation and the inexact Bregman procedure

The *Crossing Model* and the *Constrained Model* allow to restore the original signal by estimating the regularization parameter  $\beta$  during the computation, while the Bregman

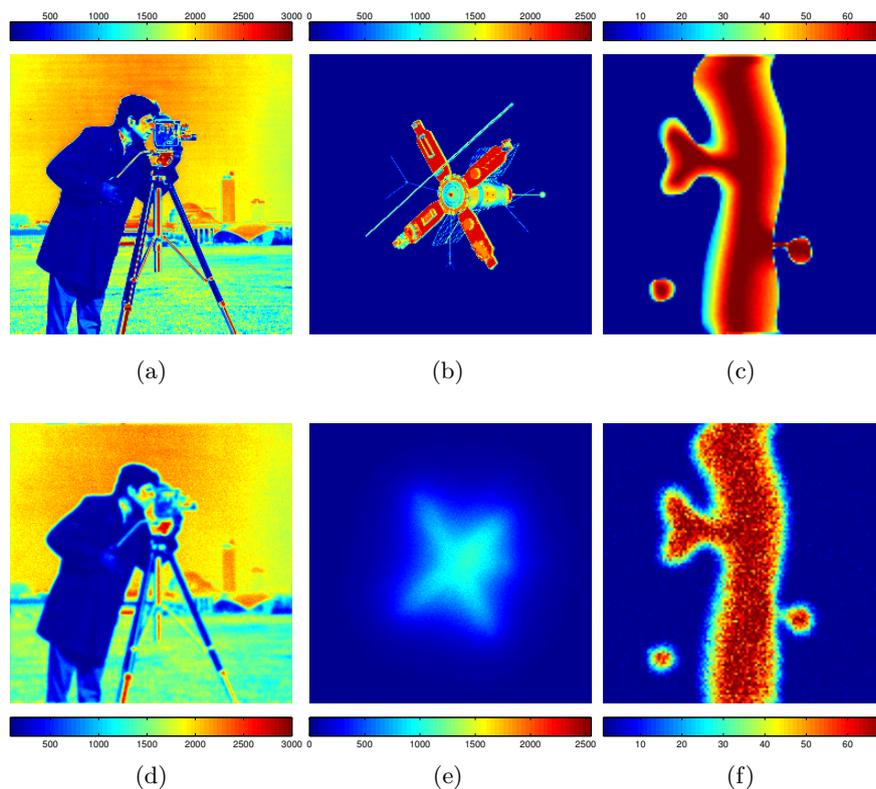
procedure requires an overestimation of such a parameter. As for the Bregman procedure, Model 1 and Model 2 need to be properly set up in order to have a satisfying result. In this section we compare the behaviour of the two estimation models and the Bregman procedure. The test problems used for this experimentation are described below.

**cameraman:** the test problem is similar to that used in Section 2.7: the original  $256 \times 256$  image has values scaled in  $[0, 3000]$ , the image has been corrupted by Gaussian blur (with standard deviation 1.3) and contaminated by Poisson noise; the background is 0 and the values of  $\mathbf{gn}$  are in the range  $[75, 2853]$ .

**Lspacecraft:** the original image is the  $256 \times 256$  image used in the experiments in Section 5.1, multiplied by 10: its values are in the range  $[0, 2550]$  and the background term  $bg$  is set to 10; the PSF is the one used in the cited experiments and presented in Figure 1.1(a). The values of the detected image are in the range  $[5, 1135]$ .

**micro:** the test problem described in Section 2.7.

**spacecraft:** the test problem described in Section 5.1.

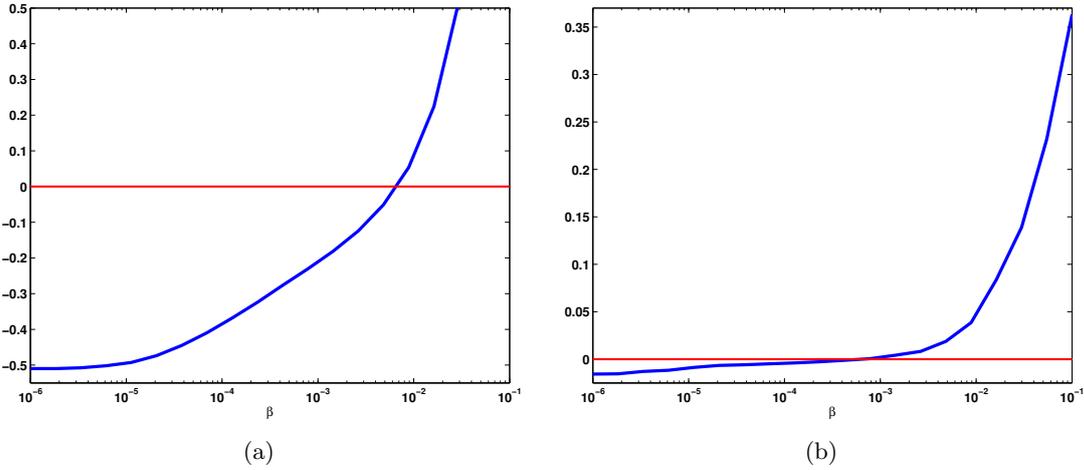


**Figure 5.6:** First row: original signals  $\mathbf{x}^*$ ; second row: perturbed data  $\mathbf{gn}$ . From left to right: cameraman, Lspacecraft and micro

Since both images are affected by Poisson noise, the data fidelity function employed is the Kullback–Leibler function; moreover, due to the characteristics of the image, we

used in a first run of tests the Hypersurface regularization and in a second one the Total Variation regularization: the two variational problems will be indicated as KL-HS and KL-TV, respectively.

**Model 1.** We solve the discrepancy equation (3.2) with MDF, the root finding solver described in Section 3.3.1. In Figure 5.7, we show the plot of  $\mathcal{D}_{\mathbf{H}}(\mathbf{x}_\beta; \mathbf{g}\mathbf{n}) - 1$  with respect to  $\beta$  for the test problems *cameraman* and *Lspacecraft*. We observe that, for the second test problem, the behaviour of this function around its zero makes very hard the localization of the crossing value. On the other hand, for *cameraman* the computation of the root of  $\mathcal{D}_{\mathbf{H}}(\mathbf{x}_\beta; \mathbf{g}\mathbf{n}) - 1$  appears easier.



**Figure 5.7:** *Left: behaviour of  $\mathcal{D}_{\mathbf{H}}(\mathbf{x}_\beta; \mathbf{g}\mathbf{n}) - 1$  for cameraman. Right: behaviour of  $\mathcal{D}_{\mathbf{H}}(\mathbf{x}_\beta; \mathbf{g}\mathbf{n}) - 1$  for Lspacecraft*

The stopping rule in MDF solver is given by (3.21):

$$\begin{aligned} |\mathcal{F}(\beta_k)| &\leq \varepsilon_1 \quad \text{or} \\ (|\beta_k - \beta_{k-1}| &\leq \varepsilon_2 \beta_k \quad \text{and} \quad |\mathcal{F}(\beta_k)| \leq 10\varepsilon_1) \end{aligned} \quad (5.1)$$

with  $\varepsilon_1 = 5 \cdot 10^{-4}$ ,  $\varepsilon_2 = 5 \cdot 10^{-3}$ . We distinguish the two regularization choices:

HS Regularization. SGP is stopped when the following criterion is satisfied

$$|\varphi_{\beta_i}(\mathbf{x}^l; \mathbf{g}\mathbf{n}) - \varphi_{\beta_i}(\mathbf{x}^{l-1}; \mathbf{g}\mathbf{n})| \leq \varepsilon_{inn} |\varphi_{\beta_i}(\mathbf{x}^l; \mathbf{g}\mathbf{n})| \quad (5.2)$$

or when a maximum number of iterations, equal to 5000, is performed without obtaining the required accuracy. We recall that  $\varphi_{\beta_i}$  denotes the objective function of the penalized problem

$$\min_{\mathbf{x} \geq 0} \varphi_{\beta_i}(\mathbf{x}; \mathbf{g}\mathbf{n}) \equiv \mathbf{KL}(\mathbf{H}\mathbf{x} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta_i \varphi_1(\mathbf{x})$$

As described in Section 3.3.1, the  $\varepsilon_{inn}$  is a mild tolerance in the Bracketing Phase and it becomes more severe in the MDF Secant Phase, up to a lower bound  $\varepsilon_{maxinn}$ , reported in the caption of the tables concerning Model 1.

TV Regularization. In this case, PIDSplit+ is used as inner solver of MDF method; the stopping criterion (5.2) is coupled with a specific rule for the alternating direction multipliers methods [23], involving current values of the primal and dual residuals (3.23):

$$\|\mathbf{r}^l\| \leq \varepsilon_{pri} \text{ and } \|\mathbf{s}^l\| \leq \varepsilon_{dual}, \quad (5.3)$$

where  $\varepsilon_{pri}$  and  $\varepsilon_{dual}$  are positive feasibility tolerances. These tolerances are defined by means of an absolute and a relative criterion:

$$\begin{aligned} \varepsilon_{pri} &= \varepsilon_a 2\sqrt{N} + \\ &\quad + \varepsilon_r \max(\|M\mathbf{x}^k\|, \|\mathbf{w}^k\|), \\ \varepsilon_{dual} &= \varepsilon_a \sqrt{N} + \varepsilon_r \|M^T \mathbf{p}^k\|, \end{aligned} \quad (5.4)$$

where  $\mathbf{p}^k$  is the current value of the multiplier of the linear constraint  $M\mathbf{x} = \mathbf{w}$ ; we fix  $\varepsilon_a = \varepsilon_r = \varepsilon_1$ . Furthermore, in order to avoid the matrix-vector operations  $M^T \sigma^k$  and  $M^T \mathbf{p}^k$  (arising in the computation of  $\mathbf{s}^k$  and  $\varepsilon_{dual}$ , respectively), we overestimate  $\|M^T \cdot\|$  with  $\sqrt{10} \|\cdot\|$ , where  $\sqrt{10}$  is an approximation of  $\|M^T\|$ . Then PIDSplit+ is stopped when one of the conditions (5.3) or (5.2) is satisfied or when a maximum number of iteration, equal to 5000, is performed without obtaining the required accuracy. The effectiveness of PIDSplit+ strongly depends on the choice of  $\gamma$ ; following [87], for the solution of the subproblem corresponding to  $\beta_i$ , the value of  $\gamma$  is set equal to  $\frac{a}{\beta_i}$  where  $a > 0$  is a fixed parameter in MDF; the selection of a suitable value for  $a$  requires a trial process. To overcome this drawback, we exploit the updating procedure (3.24), implementing an adaptive version of PIDSplit+; in this case, the initial value  $\gamma^0$  for the  $i$ -th subproblem of MDF is set equal to  $\frac{1}{\beta_i}$  while the parameters of the updating procedure are  $\mu = 5$ ,  $\alpha = 2$ ,  $k_{max} = 3500$ .

Test problem	$k$	$k_{tot}$	$\beta_k$	$\mathcal{D}_k$	$err$	Time(s)
KL-HS, SGP						
<i>cameraman</i>	8	815	$6.689 \cdot 10^{-3}$	0.9996	$8.562 \cdot 10^{-2}$	16.67
<i>Lspacecraft</i>	7	1956	$5.332 \cdot 10^{-4}$	0.9999	$3.044 \cdot 10^{-1}$	39.39
<i>micro</i>	11	1458	$3.374 \cdot 10^{-3}$	0.9997	$1.658 \cdot 10^{-1}$	9.77
<i>spacecraft</i>	41	3731	$1 \cdot 10^{-41}$	1.0494	$7.783 \cdot 10^{-1}$	285.91
KL-TV, PIDSplit+						
<i>cameraman</i>	11	600	$6.575 \cdot 10^{-3}$	0.9972	$8.514 \cdot 10^{-2}$	26.26
<i>Lspacecraft</i>	7	1744	$1.000 \cdot 10^{-5}$	0.9998	$3.5452 \cdot 10^{-1}$	71.22
<i>micro</i>	8	696	$1.6 \cdot 10^{-3}$	0.9999	$1.577 \cdot 10^{-1}$	13.76
<i>spacecraft</i>	41	5863	$1 \cdot 10^{-41}$	86.3081	$1 \cdot 10^0$	258.73

**Table 5.6:** Model 1. When MDF is solved by SGP,  $\varepsilon_{maxinn} = 5 \cdot 10^{-8}$ .  $k$  is the number of MDF iterations,  $k_{tot}$  is the total number of inner solver iterations,  $\beta_k$  is the estimate of  $\beta$ ,  $\mathcal{D}_k \equiv \mathcal{D}_{\mathbf{H}}(\mathbf{x}_{\beta_k}; \mathbf{g}_{\mathbf{n}})$ ,  $err$  is the relative reconstruction error, Time is the execution time in seconds. When the inner solver is PIDSplit+, for each  $i$ -th subproblem  $\gamma^0 = 1/\beta_i$ .

Table 5.6 shows the results obtained when the regularization parameter is estimated by using MDF with inner solver SGP with HS regularization and PIDSplit+ with TV

Test problem	$k$	$\beta_k$	$\mathcal{D}_k$	$err$	$Time(s)$
KL–HS, adaptive ADMM					
<i>cameraman</i>	451	$6.699 \cdot 10^{-3}$	0.9997	$8.535 \cdot 10^{-2}$	30.28
<i>Lspacecraft</i>	957	$8.035 \cdot 10^{-4}$	1.0050	$3.434 \cdot 10^{-1}$	64.09
<i>micro</i>	5000	$7.637 \cdot 10^{-3}$	1.0072	$1.294 \cdot 10^{-1}$	94.41
<i>spacecraft</i>	5000	$1.501 \cdot 10^{-4}$	1.0506	$5.098 \cdot 10^{-1}$	360.20
KL–TV, adaptive ADMM					
<i>cameraman</i>	441	$6.689 \cdot 10^{-3}$	0.9997	$8.534 \cdot 10^{-2}$	28.96
<i>Lspacecraft</i>	956	$8.036 \cdot 10^{-4}$	1.0050	$3.435 \cdot 10^{-1}$	60.67
<i>micro</i>	5000	$7.641 \cdot 10^{-3}$	1.0071	$1.295 \cdot 10^{-1}$	92.53
<i>spacecraft</i>	5000	$1.501 \cdot 10^{-4}$	1.0506	$5.098 \cdot 10^{-1}$	339.78

**Table 5.7:** Model 2.  $\gamma^0 = 1$  for *cameraman* and *Lspacecraft*, while  $\gamma^0 = 0.5$  for *micro* and  $\gamma^0 = 0.1$  for *spacecraft*.  $k$  is the number of ADMM iteration,  $\beta_k$  is the estimate of  $\beta$ ,  $\mathcal{D}_k \equiv \mathcal{D}_{\mathbf{H}}(\mathbf{x}_{\beta_k}; \mathbf{g}\mathbf{n})$ ,  $err$  is the relative reconstruction error,  $Time$  is the execution time in seconds.

regularization, respectively. In the table,  $k$  is the number of iterations required by MDF to satisfy the stopping criterion,  $k_{tot}$  is the total number of iterations of the inner solver,  $\beta_k$  is the obtained estimate,  $\mathcal{D}_k \equiv \mathcal{D}_{\mathbf{H}}(\mathbf{x}_{\beta_k}; \mathbf{g}\mathbf{n})$ ,  $err$  is the relative reconstruction error on the original object  $\mathbf{x}^*$ , and, finally,  $Time$  denotes the execution time in seconds.

**Model 2.** Model 2 is solved by ADMM. For stopping the method, before verifying a condition similar to (5.1), it has to be verified that the current iterate is an approximate solution of the minimization problem; indeed, it could happen that the constraint is satisfied but the solution of the problem is not yet sufficiently accurate. Therefore, when the current iterate  $\mathbf{x}^k$  satisfies a condition as (5.3), the solver switches to check that the following stopping rule is satisfied

$$|\mathcal{D}_{\mathbf{H}}(\mathbf{x}_{\beta_k}; \mathbf{g}\mathbf{n}) - \eta| \leq \varepsilon_1 \text{ or} \\ (|\beta_k - \beta_{k-1}| \leq \varepsilon_2 |\beta_k| \text{ and } |\mathcal{D}_{\mathbf{H}}(\mathbf{x}_{\beta_k}; \mathbf{g}\mathbf{n}) - \eta| \leq 10\varepsilon_1)$$

The maximum number of ADMM iterations is fixed to 5000. Furthermore, at each step of ADMM, the constrained least squares subproblem related to  $\mathbf{w}_{(1)}$  is solved by the Newton's method with a stopping tolerance of  $10^{-12}$  on the distance between two successive approximations. In our experiments, after the initial iterations of ADMM, four Newton's steps enable to compute an approximation of  $\beta^i$  within the required tolerance.

In Table 5.7 we report the results obtained by ADMM combined with the updating procedure (3.24), with  $\alpha = 2$ ,  $\mu = 10$  and  $k_{max} = 3500$ . Here the initial value of the parameter  $\gamma^0$  is set to 1 for *cameraman* and *Lspacecraft* test problems, while  $\gamma^0 = 0.5$  and  $\gamma^0 = 0.1$  for *micro* and *spacecraft*, respectively. Our numerical experience has shown that this adaptive version of ADMM allows to obtain satisfactory solutions of Model 2 without having to find a suitable value for  $\gamma$  in the non-adaptive version. Also in this case the effectiveness of the method can depend on the computational setting, for example on the value of  $\gamma^0$ , but this dependence appears less relevant than for the standard ADMM.

**Bregman Regularization.** We apply the inexact Bregman procedure to these test problems; we solve the model KL-HS by the SGP method as inner solver in the Bregman iteration, while for the model KL-TV we use AEM method for solving each  $k$ -th subproblem. In the former case, we use the stopping criterion (4.40), by setting  $\tau_{\text{SGP}} = 10^{-6}$ ,  $M = 10$ . Then  $c = \|\eta_1\|$  and  $\alpha = 1.5$ . For KL-TV model, we employ (4.34) as stopping rule: for the first subproblem the stopping criterion is based on the relative difference on two successive iterates with  $\tau_{\text{AEM}} = 10^{-5}$ , while in the subsequent iterations  $c = \|\eta_1\|$ ,  $\alpha = 1.5$  and  $\vartheta = 2.1$ .

In each test problem and for any regularization function chosen, the regularization parameter is set to ten times the optimal one, found empirically.

In Table 5.9 we present the numerical results obtained in these experiments:  $k$  is the outer iteration,  $k_{\text{tot}}$  is the total number of the iterations of the inner solver,  $\mathcal{D}_k \equiv \mathcal{D}_{\mathbf{H}}(\mathbf{x}^k; \mathbf{g}\mathbf{n})$ ,  $\text{err}$  is the relative reconstruction error and finally  $\text{Time}$  is the execution time in seconds.

KL-HS, Bregman method with SGP									
<i>cameraman</i>					<i>Lspacecraft</i>				
$k$	$k_{\text{tot}}$	$\mathcal{D}_k$	$\text{err}$	$\text{Time}$	$k$	$k_{\text{tot}}$	$\mathcal{D}_k$	$\text{err}$	$\text{Time}$
5	3131	1.0301	0.0889	88.39	7	21418	1.0001	0.3071	537.24
6	3906	0.9960	0.0873	109.14	8	26418	0.9997	0.3063	654.01
7	4578	0.9708	0.0859	127.29	9	31418	0.9993	0.3062	772.79
<i>micro</i>					<i>spacecraft</i>				
$k$	$k_{\text{tot}}$	$\mathcal{D}_k$	$\text{err}$	$\text{Time}$	$k$	$k_{\text{tot}}$	$\mathcal{D}_k$	$\text{err}$	$\text{Time}$
8	4147	1.0433	0.0827	41.43	8	24580	1.0528	0.3735	868.01
9	4615	1.0406	0.0837	45.81	9	27480	1.0527	0.3780	972.14
10	5214	1.0388	0.0848	51.81	10	32191	1.0525	0.3848	1150.91

**Table 5.8:** Bregman procedure for the test problems, HS regularization.  $k$  denotes the outer iteration,  $k_{\text{tot}}$  is the total number of iterations of the inner solver at the  $k$ -th outer iteration,  $\mathcal{D}_k \equiv \mathcal{D}_{\mathbf{H}}(\mathbf{x}^k; \mathbf{g}\mathbf{n})$ ,  $\text{err}$  is the relative reconstruction error and finally  $\text{Time}$  is the execution time in seconds. We report the iterations where  $\mathcal{D}_k$  is close to 1.

One can observe that in some test problems the three procedures are quite equivalent: e.g., the results obtained for *cameraman* problem are similar for each model used. In other cases, such as *Lspacecraft*, the inexact Bregman procedure reaches results very close to the ones by Model 1 and Model 2, but it has a higher computational cost. For problems such as *micro* and *spacecraft*, the *Constrained Model* and the *Crossing Model* provide unsatisfactory results; for example, the *spacecraft* test problem is not solved properly by Model 1 since the zero of the equation  $\mathcal{D}_{\mathbf{H}}(\mathbf{x}_\beta; \mathbf{g}\mathbf{n}) - 1 = 0$  is not reached. On the other hand, the inexact Bregman procedure, even if it requires a stronger computational effort, gives the desired result using an overestimation of  $\beta$  (see Figure 5.8).

Model 1 and Model 2 behave as in *micro* and *spacecraft* problems when one tries to reconstruct High Dynamic Range (HDR) images, while the inexact Bregman procedure enables us to obtain satisfying reconstruction and, at the same time, to gain a con-

KL-TV, Bregman method with AEM									
<i>cameraman</i>					<i>Lspacecraft</i>				
$k$	$k_{tot}$	$\mathcal{D}_k$	$err$	$Time$	$k$	$k_{tot}$	$\mathcal{D}_k$	$err$	$Time$
5	16594	1.0317	0.0886	492.06	8	40000	1.0008	0.3167	1123.63
6	20979	0.9976	0.0870	617.75	9	45000	1.0002	0.3142	1274.62
7	25979	0.9717	0.0856	765.20	10	50000	0.9998	0.3123	1427.10
<i>micro</i>					<i>spacecraft</i>				
$k$	$k_{tot}$	$\mathcal{D}_k$	$err$	$Time$	$k$	$k_{tot}$	$\mathcal{D}_k$	$err$	$Time$
8	33587	1.0434	0.0991	348.35	8	18333	1.0528	0.3748	736.02
9	41087	1.0408	0.0987	426.87	9	21567	1.0527	0.3800	865.15
10	45770	1.0390	0.0987	476.08	10	25440	1.0525	0.3858	1024.73

**Table 5.9:** Bregman procedure for the test problems, TV regularization.  $k$  denotes the outer iteration,  $k_{tot}$  is the total number of iterations of the inner solver at the  $k$ -th outer iteration,  $\mathcal{D}_k \equiv \mathcal{D}_{\mathbf{H}}(\mathbf{x}^k; \mathbf{g}\mathbf{n})$ ,  $err$  is the relative reconstruction error and finally  $Time$  is the execution time in seconds. We report the iteration where  $\mathcal{D}_k$  is close to 1.

trast enhancement. In the forthcoming section, the behaviour of the inexact Bregman procedure when HDR images are treated is described.

## 5.4 Application to High Dynamic Range images in Astronomy

A common problem in Astronomy imaging is the reconstruction of high contrast images, as the ones arising in the framework of ground-based telescopes when a faint signal is very close to bright stars. The actual problem lies in the high intensity difference between the stars themselves and the surrounding area: the action of the PSF destroys all the information on the region around the stars. Following [70, 31], we assume that the position of the stars is known: the main idea is to consider every point source as a delta function which is zero everywhere except at the known position of the source itself. The signal  $\mathbf{x}$  is therefore considered as the sum of two components:

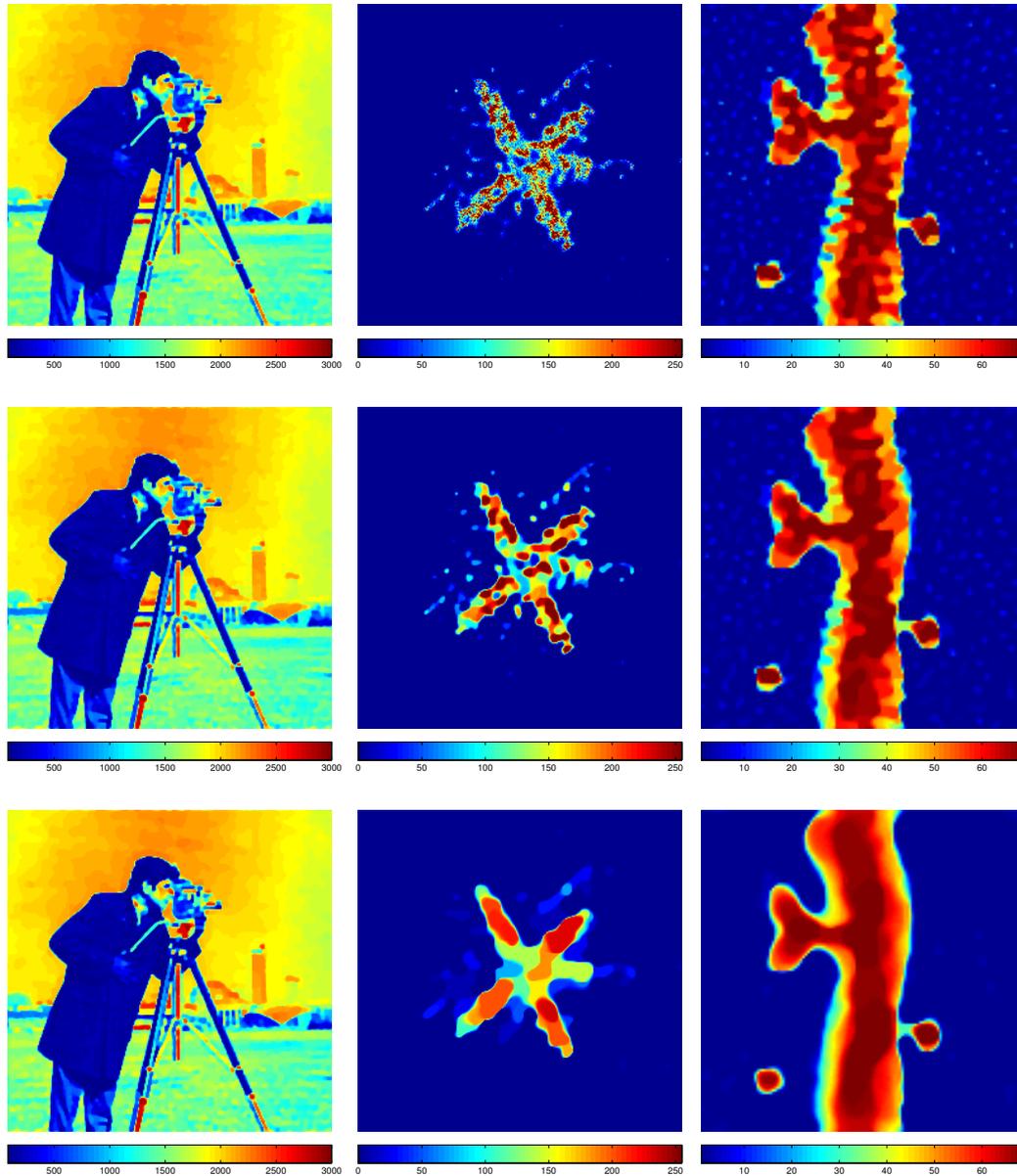
$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_d \quad (5.5)$$

where  $\mathbf{x}_p$  represents the *point sources*, while  $\mathbf{x}_d$  is the smooth component in which one is interested into; in  $\mathbf{x}_d$  the values corresponding to the sources' positions are set to zero. We can write  $\mathbf{x}_p$  as the sum of delta functions centered in the known positions multiplied for the corresponding intensities:

$$\mathbf{x}_p = \sum_{i=1}^q \bar{\mathbf{x}}_i \delta(v_i) \quad (5.6)$$

where  $q$  is the number of the stars present in the region of interest,  $\bar{\mathbf{x}}_i$  is the intensity of the  $i$ -th star and  $v_i$  is its known position. Namely,

$$\bar{\mathbf{x}}_p = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_q \end{pmatrix}$$



**Figure 5.8:** Restored images obtained by SGP method with HS regularization; from left to right, cameraman, spacecraft and micro test problems. In the first row the the results obtained for Model 1 are shown; the second row contains the images restored by Model 2. Finally, in the last row, the results obtained by solving the restoration problems by the Bregman procedure are reported. One can observe that in some problems, such as the cameraman one, the three procedures have a very similar behaviour, while in other ones (spacecraft and micro) Model 1 and Model 2 provide unsatisfactory results. On the other hand, the Bregman procedure seems to work in a proper way.

is the vector containing the intensities of the  $q$  stars.

The next step is to introduce different regularizations [74, 55, 4] for the two components. A brief discussion of the method based on this approach for Gaussian data is in [55]; in presence of Poisson noise, the same idea was proposed in [70] and in the recent years in [31, 32].

Actually, the noise affecting the given data is a mixture of Poisson noise and additive Gaussian noise: the former is due to the counting process, the latter is due to the read-out-noise (RON). In the past [89] it has been shown that one can treat the RON as a Poisson process: we adopt this approach.

Thanks to this assumption, we can employ the Kullback–Leibler functional as fit-to-data functional in our restoration problem, but we have to introduce new notations about the imaging matrix  $\mathbf{H}$  in order to take into account the decomposition (5.5).

Denoting  $\bar{\mathbf{x}} = \begin{pmatrix} \bar{\mathbf{x}}_p \\ \mathbf{x}_d \end{pmatrix}$ , we write  $\bar{\mathbf{H}} = [\mathcal{H}, \mathbf{H}]$ , where  $\mathcal{H} = [h_{v_1}, h_{v_2}, \dots, h_{v_q}]$  and  $h_j$  denotes the  $j$ -th column of  $\mathbf{H}$ . In this way the functional  $\varphi_0$  depends on  $q + N$  variables:  $\varphi_0(\bar{\mathbf{H}}\bar{\mathbf{x}} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n})$ , where we recall that  $\mathbf{g}\mathbf{n}$  is the blurred and noisy data.

Finally, we impose a regularization term only on the diffuse component: the restoration problem becomes thus

$$\min_{\mathbf{x}} \varphi_0(\bar{\mathbf{H}}\bar{\mathbf{x}} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta\varphi_1(\mathbf{x}_d) \quad (5.7)$$

with  $\varphi_1(\mathbf{x}) = 1/2\|\mathbf{L}\mathbf{x}\|^2$ . The choice of Tikhonov regularization is due to the smooth structure of the diffuse component. Assuming the conditions (1.5) hold on the imaging matrix  $\mathbf{H}$ , we prove the existence and uniqueness of the solution of (5.7) in the following Proposition, which is an extension of Proposition 1.3.

**Proposition 5.1.** *Let  $\mathbf{g}\mathbf{n} \in \mathbb{R}^M$  be positive and  $b\mathbf{g} > 0$ . Assume that the matrix  $\mathbf{H}$  satisfies (1.5) and the column rank of  $\mathcal{H} \in \mathbb{R}^{M \times q}$  is equal to  $q$ ; then we have:*

- when the null space of  $L$  is trivial, the solution of the problem (5.7) is unique for any  $\beta > 0$ ;
- when the null space of  $L$  is  $\{\alpha\mathbf{1}, \alpha \in \mathbb{R}\}$  and  $\mathbf{H}\mathbf{1} \notin \text{span}(\mathcal{H})$ , the solution of the problem (5.7) is unique for any  $\beta > 0$ .

*Proof.* In both cases, we prove that the positive semidefinite matrix  $\nabla_{\bar{\mathbf{x}}}^2\varphi_\beta(\bar{\mathbf{x}})$ , given by

$$\nabla_{\bar{\mathbf{x}}}^2\varphi_\beta(\bar{\mathbf{x}}) = \bar{\mathbf{H}}^t \text{diag} \left( \frac{\mathbf{g}\mathbf{n}}{(\bar{\mathbf{H}}\bar{\mathbf{x}} + b\mathbf{g}\mathbf{1})^2} \right) \bar{\mathbf{H}} + \beta \begin{pmatrix} 0 & 0 \\ 0 & L^t L \end{pmatrix} \quad (5.8)$$

is positive definite for  $\beta > 0$  and, consequently,  $\varphi_\beta$  is strictly convex for  $\bar{\mathbf{x}} \geq 0$ . Indeed, we can show that the intersection of the null spaces of  $\nabla_{\bar{\mathbf{x}}}^2\varphi_0$  and  $\beta\nabla_{\bar{\mathbf{x}}}^2\varphi_1$  is trivial.

- Let  $\text{Ker}(L) = \{\mathbf{0}\}$ ; since  $\text{Ker}(\nabla_{\bar{\mathbf{x}}}^2\varphi_0) = \text{Ker}(\bar{\mathbf{H}})$  and  $\text{Ker}(\nabla_{\bar{\mathbf{x}}}^2\varphi_1) = \left\{ \begin{pmatrix} \bar{\mathbf{x}}_p \\ \mathbf{0} \end{pmatrix}, \mathbf{0} \in \mathbb{R}^M \right\}$ , we have to prove that for any  $\bar{\mathbf{x}}_p \neq \mathbf{0}$

$$\bar{\mathbf{H}} \begin{pmatrix} \bar{\mathbf{x}}_p \\ \mathbf{0} \end{pmatrix} \neq \mathbf{0}$$

Indeed, for the linear independence of the columns of  $\mathcal{H}$ ,  $\mathcal{H}\bar{\mathbf{x}}_p \neq \mathbf{0}$ . Then  $\nabla_{\bar{\mathbf{x}}}^2\varphi_\beta(\bar{\mathbf{x}})$  is positive definite for  $\bar{\mathbf{x}} \geq \mathbf{0}$ .

- b. When  $\text{Ker}(L) = \{\alpha \mathbf{1}, \alpha \in \mathbb{R}\}$ ,  $\text{Ker}(\nabla_{\bar{\mathbf{x}}}^2 \varphi_1) = \left\{ \begin{pmatrix} \bar{\mathbf{x}}_p \\ \alpha \mathbf{1} \end{pmatrix}, \mathbf{1} \in \mathbb{R}^n \right\}$ . Then, we have to prove that

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

Indeed, by contradiction, if we assume

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

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

One can observe that the second part of the previous proposition still holds true also for Hypersurface regularization or Markov random field regularization: indeed for these regularization terms we have  $\text{Ker}(\nabla_{\bar{\mathbf{x}}}^2) = (\bar{\mathbf{x}}_p, \alpha \mathbf{1})^t$ .

When  $\mathbf{H} \mathbf{1} = \mathbf{1}$  holds, the assumption  $\mathbf{H} \mathbf{1} \notin \text{span}(\mathcal{H})$  is equivalent to require that  $\{\alpha \mathbf{1}, \alpha \in \mathbb{R}\}$  is not a subspace of  $\text{span}(\mathcal{H})$ ; since  $q$  is small, it is easy to verify if there exists the unique solution of  $\mathcal{H} \mathbf{x} = \mathbf{1}$ ; for example, when  $q = 1$  and  $h_{\nu_1}$  is not a constant vector,  $\mathbf{H} \mathbf{1} \notin \text{span}(\mathcal{H})$ . Moreover, when  $\mathbf{H}$  is a nonsingular matrix, there does not exist  $\mathbf{x}$  such that  $\mathcal{H} \mathbf{x} = \mathbf{1}$ ; indeed, in this case, by  $\mathbf{H} \mathbf{1} = \mathbf{1}$ , there would be contradiction with the linear independence of the columns of  $\mathbf{H}$ :

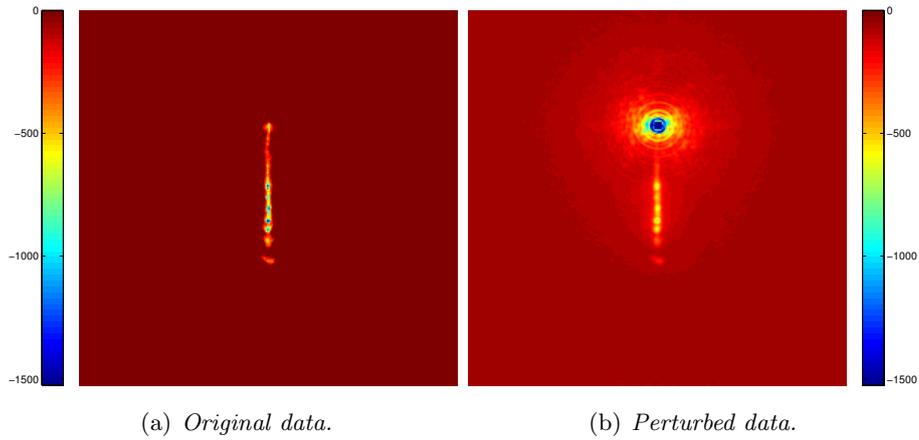
$$0 = \mathbf{1} - \mathcal{H} \mathbf{x} = \mathbf{H} \mathbf{1} - \mathcal{H} \mathbf{x} = \sum_{j \neq \nu_1, \dots, \nu_q} h_j + \sum_{j = \nu_1, \dots, \nu_q} h_j (1 - \mathbf{x}_j)$$

We show two practical case in which the model described can be applied; moreover, the Bregman procedure can be implemented for problem (5.7) in order to preserve the intensity's difference of the two terms, thanks to the ability to enhance the contrast. We remark also that the numerical experimentation shows that the techniques for the estimation of  $\beta$  do not enable us to obtain similar results.

#### 5.4.1 Young Stellar Objects

The first case we consider consists in the image of Young Stellar Objects (YSO): actually the data, kindly provided by the authors of [32], are simulated Large Binocular Telescope (LBT) infrared narrow-band observations of a star-jet system.

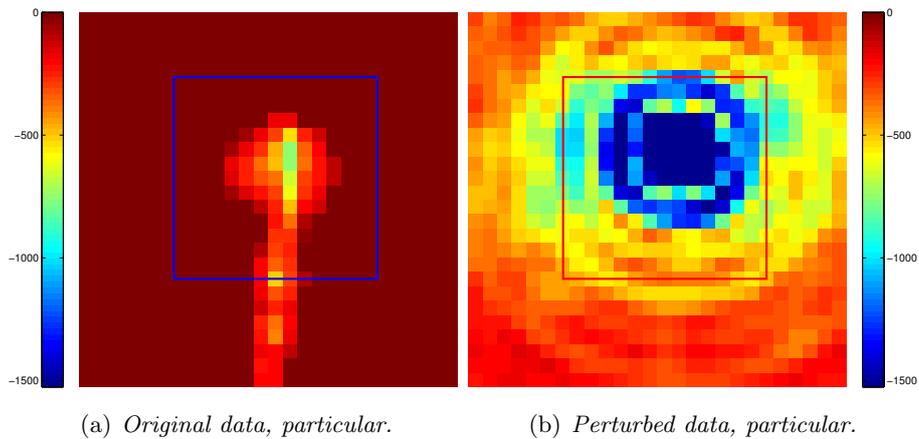
We have four test problems, in which the point source is a single star ( $q = 1$ ). The images are obtained from an optical image taken with Hubble Space Telescope of the HH34 jet (Herbig-Haro objects) [82], downscaled by a factor 3 and immersed into a  $256 \times 256$  array, surrounded by zeros. A pixel scale of 15 mas/pixel (milliarcseconds per pixel) has been assumed as typical scale of a camera attached to a single LBT dish. In the pixel (178, 129) corresponding to the source, a point-like star with variable magnitude (from 8 to 11 mag) was added in order to simulate four different objects at different contrast. For each object, a  $256 \times 256$  image has been obtained by convolving the object with an adaptive-optics corrected PSF with a Strehl ratio of 0.67. An average H-band sky brightness of 15.5mag/arcsec<sup>2</sup> has been added as background emission and,



**Figure 5.9:** Star9 test problem. On the left the original image, on the right the blurred and noisy data. The image are shown in a reverse sqrt-scale.

finally, the results are corrupted by a mixture of Poisson noise and additive Gaussian noise with  $\sigma = 10e^- / \text{pixel}$  (i.e., 10 electrons per pixel). A more precise and detailed description on the image acquisition process is available in [32].

The four test-problems are denoted by *Star8*, *Star9*, *Star10*, *Star11*, in dependence of the magnitude of the point source; in Figure 5.9 the data relative to the problem *Star9* are shown, while in Figure 5.10 a smaller region is presented. The functional to be minimized involves Tikhonov regularization  $\varphi_1(\mathbf{x}) = 1/2\|L\mathbf{x}\|^2$  since the  $\mathbf{x}_d$  component has a smooth structure.



**Figure 5.10:** Star9 test problem. A smaller region around the point source is presented. The blue square (in 5.10(a)) and the red square (in 5.10(b)) represent the region of interest with respect to we will compute  $\rho_w$ .

For each test problem, we use Model 1 described in Chapter 3 for the estimation of the regularization parameter. The equation which has to be solved is

$$\mathcal{D}_{\mathbf{H}}(\bar{\mathbf{x}}^{(\beta)}; \mathbf{g}\mathbf{n}) = \eta \quad (5.9)$$

where  $\bar{\mathbf{x}}^{(\beta)}$  is the solution of (5.7) and  $\eta \sim 1$ . We pursue this scope using the method described in section 3.3.1, which combines bisection and secant iterations. The interval in which the optimal value is found is determined by the bisection iteration and it is  $[10^{-10}, 10^{-8}]$  for all the test problems. In Table 5.10 we report the value of  $\beta$  computed for each test problem with different regularization terms. The minimization problems appearing in each step of the root-finding procedure of (5.9) are solved by SGP method. The stopping rule is the standard one on the relative difference on two successive objective function values:

$$\frac{|\varphi_{\beta}(\bar{\mathbf{x}}^k) - \varphi_{\beta}(\bar{\mathbf{x}}^{k+1})|}{|\varphi_{\beta}(\bar{\mathbf{x}}^{k+1})|} \leq \tau_{\text{SGP}} \quad (5.10)$$

coupled with the request that the mean of the last  $M$  iteration of this difference is less than  $10\tau_{\text{SGP}}$ ; in these experiments,  $\tau_{\text{SGP}} = 0.5 \cdot 10^{-9}$  and  $M = 10$ . In Table 5.10 we report also the total number *totit* of SGP iterations needed for the evaluations of  $\mathcal{D}_{\mathbf{H}}(\bar{\mathbf{x}}^{(\beta)}, \mathbf{g}\mathbf{n})$  and, in brackets, the number  $T$  of evaluations performed to obtain the numerical solution of (5.9) so that  $|\mathcal{D}_{\mathbf{H}}(\bar{\mathbf{x}}^{(\beta)}, \mathbf{g}\mathbf{n}) - 1| \leq 10^{-4}$  and the relative distance between the last two iterates is less than  $10^{-3}$ .

To evaluate the reconstructions obtained, we compute the relative error on three different objects:

- ▷ denoting by  $\bar{\mathbf{x}}^{(\beta)} = \begin{pmatrix} \bar{\mathbf{x}}_p^{(\beta)} \\ \bar{\mathbf{x}}_d^{(\beta)} \end{pmatrix}$  the reconstruction obtained with the  $\beta$  value, we indicate with  $\rho$  the relative error on the diffuse component:

$$\rho = \frac{\|\mathbf{x}_d^{(\beta)} - \mathbf{x}_d^*\|}{\|\mathbf{x}_d^*\|}$$

where  $\mathbf{x}_d^*$  is the original diffuse component;

- ▷ we indicate with  $\rho_p$  the relative reconstruction error on the value of the point source;
- ▷ the astrophysical interest for high contrast images actually lies in the region very close to the point source(s) (in astronomical measure units), in fact astronomers look for orbiting exoplanets and brown dwarfs, or they would observe circumstellar structures around them. Hence, we indicate with  $\rho_w$  the relative reconstruction error on the diffuse component in a small window around the star(s); for example, the region of interest for the images of the YSOs is the red one in Figure 5.10(a) (corresponding to the blue one in Figure 5.10(b)).

One can observe that the value of  $\beta$  found by solving (5.9) does not correspond to the reconstructed solution with the minimum reconstruction error: in general, for the intermediate iterations, the errors can be smaller than the ones presented in Table 5.10. The behaviour of the reconstruction errors could seem strange: in fact,  $\rho$ , the reconstruction error on the diffuse image, is lower than  $\rho_w$ , the reconstruction error on

$\varphi_1$	$\bar{\beta}$	totit ( $T$ )	$\rho$	$\rho_w$	$\rho_p$
<i>Star8</i>					
$L = \mathbf{I}$	$3.1 \cdot 10^{-9}$	5462(10)	0.24	0.72	$5.1 \cdot 10^{-4}$
$L = \nabla$	$1.4 \cdot 10^{-9}$	5600(13)	0.23	0.58	$2.4 \cdot 10^{-4}$
$L = \nabla^2$	$5.8 \cdot 10^{-10}$	4428(8)	0.23	0.52	$1.5 \cdot 10^{-4}$
<i>Star9</i>					
$L = \mathbf{I}$	$3.0 \cdot 10^{-9}$	6286(11)	0.23	0.69	$1.1 \cdot 10^{-3}$
$L = \nabla$	$1.4 \cdot 10^{-9}$	5828(10)	0.23	0.53	$5.4 \cdot 10^{-4}$
$L = \nabla^2$	$5.3 \cdot 10^{-10}$	4064(9)	0.22	0.47	$3.3 \cdot 10^{-4}$
<i>Star10</i>					
$L = \mathbf{I}$	$2.9 \cdot 10^{-9}$	6447(10)	0.23	0.66	$3.1 \cdot 10^{-3}$
$L = \nabla$	$1.4 \cdot 10^{-9}$	5792(20)	0.22	0.50	$1.7 \cdot 10^{-3}$
$L = \nabla^2$	$5.0 \cdot 10^{-10}$	3882(9)	0.22	0.45	$1.2 \cdot 10^{-3}$
<i>Star11</i>					
$L = \mathbf{I}$	$2.7 \cdot 10^{-9}$	5458(9)	0.23	0.64	$7.4 \cdot 10^{-3}$
$L = \nabla$	$1.3 \cdot 10^{-9}$	5664(8)	0.22	0.48	$4.5 \cdot 10^{-3}$
$L = \nabla^2$	$4.5 \cdot 10^{-10}$	4228(20)	0.22	0.44	$3.6 \cdot 10^{-3}$

**Table 5.10:** Numerical results obtained by applying the discrepancy criterion (Model 1);  $\bar{\beta}$  is the estimation of the regularization parameter, totit is the total number of inner iterations while  $T$  is the number of the MDF's steps;  $\rho$  is the relative reconstruction error on the diffuse component,  $\rho_w$  is the relative error in the small window considered and finally  $\rho_p$  is the relative reconstruction error on the intensities of the point source.

the window of interest, since the computation of the former involves a larger portion of the diffuse component. Furthermore, the reconstruction error on the point source component is very small due to the fact that the intensities of the sources are very high, hence even an approximate reconstruction is sufficient to assure a small  $\rho_p$ . We stress that the main interest, however, lies in the reconstruction of the window of interest. To overcome this drawback, we have determined empirically the optimal value for the regularization parameter, by solving the minimization problem (5.7) via the SGP method employing different values for  $\beta$ . In Table 5.11 we report the results of these experiments for the search of the optimal  $\beta$ . In order to choose this optimal value, the measure of error  $\rho$  has been considered.

Once we have determined the regularization parameter (or an overestimation of it) we can apply the inexact Bregman procedure in order to recover the high contrast image. We implement again SGP method as inner solver for the subproblems appearing in the Bregman procedure, and we employ the values found by solving the discrepancy equation as estimations for the regularization parameter  $\beta$ , since they are greater than the optimal ones found empirically. For each test problem and for each regularization functional we perform numerical tests also for values lying in the interval  $[\beta_{\text{opt}}, \bar{\beta}]$ . For the stopping criterion, the sequence  $\mu_k$  is chosen as in (4.40), setting the value for  $c$  as explained in Remark 4.4: the first iteration is stopped when the standard rule described above is satisfied with  $\tau_{\text{SGP}} = 10^{-5}$  and  $M = 1$  or  $M = 5$ , then  $c = \|\eta_1\|$  and  $\alpha = 1.5$ . Furthermore, in order to avoid a too fast decrease of  $\varphi_0$  close to the levels of interest, in SGP method the maximum steplength  $\alpha_{\text{max}}$  is set equal to 10 until  $\mu_k = c/k^\alpha > 5 \cdot 10^{-4}$ ; this happens within the first ten/twenty outer iterations; then  $\alpha_{\text{max}} = 1000$ .

$\varphi_1(\mathbf{x}_d)$	$\tilde{\beta}$	$\tilde{k}$	$\rho$	$\rho_w$	$\rho_p$	$\bar{\varphi}_0$
<i>Star8</i> , $\tau_{\text{SGP}} = 10^{-8}$ , $M = 5$						
$L = \mathbf{I}$	$1.2 \cdot 10^{-10}$	2433	0.1583	0.4342	$9.0 \cdot 10^{-5}$	30879.4
$L = \nabla$	$4.0 \cdot 10^{-11}$	2492	0.1585	0.3825	$2.6 \cdot 10^{-4}$	30882.3
$L = \nabla^2$	$1.0 \cdot 10^{-11}$	3493	0.1578	0.3149	$2.9 \cdot 10^{-4}$	30874.0
<i>Star9</i> , $\tau_{\text{SGP}} = 10^{-8}$ , $M = 5$						
$L = \mathbf{I}$	$1.6 \cdot 10^{-10}$	2942	0.1597	0.4632	$5.6 \cdot 10^{-5}$	31024.6
$L = \nabla$	$5.0 \cdot 10^{-11}$	2594	0.1548	0.3072	$4.4 \cdot 10^{-4}$	31032.4
$L = \nabla^2$	$1.0 \cdot 10^{-11}$	3172	0.1558	0.3399	$7.9 \cdot 10^{-4}$	31020.4
<i>Star10</i> , $\tau_{\text{SGP}} = 0.5 \cdot 10^{-8}$ , $M = 5$						
$L = \mathbf{I}$	$9.0 \cdot 10^{-11}$	4423	0.1598	0.4264	$3.9 \cdot 10^{-4}$	31124.4
$L = \nabla$	$5.0 \cdot 10^{-11}$	3253	0.1578	0.2760	$3.9 \cdot 10^{-4}$	31143.2
$L = \nabla^2$	$8.0 \cdot 10^{-12}$	4107	0.1548	0.3318	$1.4 \cdot 10^{-3}$	31124.9
<i>Star11</i> , $\tau_{\text{SGP}} = 0.5 \cdot 10^{-8}$ , $M = 5$						
$L = I$	$8.0 \cdot 10^{-11}$	5096	0.1578	0.4179	$1.5 \cdot 10^{-3}$	31306.7
$L = \nabla$	$5.0 \cdot 10^{-11}$	3020	0.1528	0.2603	$7.2 \cdot 10^{-4}$	31331.1
$L = \nabla^2$	$7.0 \cdot 10^{-12}$	4019	0.1525	0.3279	$3.2 \cdot 10^{-3}$	31312.9

**Table 5.11:** Numerical results obtained by applying SGP method to problem (5.7) with Tikhonov regularization; the value of the regularization parameter  $\tilde{\beta}$  corresponds to the best relative reconstruction error  $\rho$  for the diffuse component at the iteration  $\tilde{k}$ ; the minimum  $\rho$  is obtained by several trials with different values of  $\beta$ .  $\bar{\varphi}_0$  stands for  $\mathbf{KL}(\bar{\mathbf{H}}\bar{\mathbf{x}}^\kappa + b\mathbf{g}; \mathbf{gn})$ .

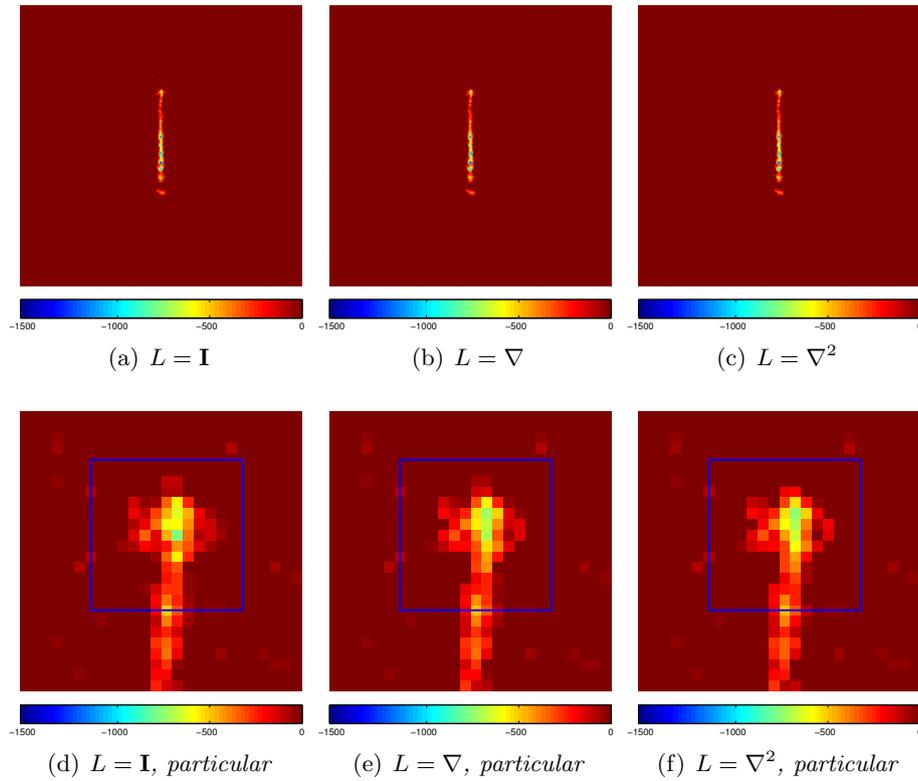
In Table 5.12 we report the results obtained;  $\kappa$  denotes the outer iteration of the inexact Bregman scheme at which we obtain the minimum relative reconstruction error  $\rho^\kappa$  of the diffuse component; we report in brackets also the total number of SGP iterations (*totit*) involved in  $\kappa$  outer iterations, the relative error  $\rho_w^\kappa$  for the square window around the star, the relative error  $\rho_p^\kappa$  for the point component and the value of the  $\mathbf{KL}$  function at the iterate  $\bar{\mathbf{x}}^\kappa$ . Table 5.12 shows that the inexact Bregman scheme enables us to obtain relative reconstruction errors at least comparable, but in the most cases better than those obtained by solving the variational problem with a suitable value of the regularization parameter. The same observation can be done about the reconstruction error  $\rho_w$  of the window around the star.

In order to visually evaluate the results obtained, we show for the problem *Star08* in Figure 5.12 (first row) the superposition of the column number 129 (from pixel 167 to 180) of the original object, of the reconstruction obtained by SGP using the optimal values for  $\beta$  presented in Table 5.11 and of the reconstruction by the Bregman Procedure. These figures highlight that the choice for  $\varphi_1$  involving a first or second order finite difference provides a contrast enhancement and a better quality in the reconstructed images, also in presence of a very bright stars (Figure 5.12, second row).

For a visual inspection of the results obtained so far, we show in Figure 5.11 the restored images for *Star09* problem with the three different types of regularization; these images are displayed in the same range of the object of Figure 5.10.

$\varphi_1$	$\beta$	$\kappa$ (totit)	$\rho^\kappa$	$\rho_w^\kappa$	$\rho_P^\kappa$	$\bar{\varphi}_0$
<i>Star8</i>						
$L = \mathbf{I}$	$3.1 \cdot 10^{-9}$	43(3603)	0.1409	0.3926	$3.469 \cdot 10^{-4}$	30856.4
$L = \mathbf{I}$	$1.1 \cdot 10^{-9}$	20 (4196)	0.1382	0.3318	$3.510 \cdot 10^{-4}$	30854.0
$L = \nabla$	$1.4 \cdot 10^{-9}$	34 (5133)	0.1579	0.3678	$1.223 \cdot 10^{-4}$	30854.2
$L = \nabla$	$0.5 \cdot 10^{-9}$	18 (4809)	0.1528	0.3495	$3.212 \cdot 10^{-4}$	30852.0
$L = \nabla^2$	$5.8 \cdot 10^{-10}$	48 (6689)	0.1590	0.3344	$1.224 \cdot 10^{-4}$	30854.9
$L = \nabla^2$	$3.8 \cdot 10^{-10}$	36 (5864)	0.1566	0.3072	$1.665 \cdot 10^{-4}$	30853.8
<i>Star9</i>						
$L = \mathbf{I}$	$3 \cdot 10^{-9}$	26 (4036)	0.1473	0.3712	$3.045 \cdot 10^{-3}$	30999.9
$L = \mathbf{I}$	$2 \cdot 10^{-9}$	20 (3590)	0.1410	0.3292	$6.126 \cdot 10^{-4}$	30999.8
$L = \nabla$	$1.4 \cdot 10^{-9}$	41 (5755)	0.1472	0.2477	$4.945 \cdot 10^{-4}$	30996.8
$L = \nabla$	$0.4 \cdot 10^{-9}$	16 (5287)	0.1535	0.3368	$4.600 \cdot 10^{-4}$	30994.6
$L = \nabla^2$	$5.3 \cdot 10^{-10}$	94 (10113)	0.1478	0.2265	$5.166 \cdot 10^{-4}$	30994.2
$L = \nabla^2$	$3.3 \cdot 10^{-10}$	63 (10071)	0.1489	0.2547	$5.479 \cdot 10^{-4}$	30994.1
<i>Star10</i>						
$L = \mathbf{I}$	$2.9 \cdot 10^{-9}$	20 (4187)	0.1579	0.4139	$3.140 \cdot 10^{-4}$	31108.6
$L = \mathbf{I}$	$1.9 \cdot 10^{-9}$	17 (3769)	0.1546	0.4289	$5.737 \cdot 10^{-4}$	31108.1
$L = \nabla$	$1.4 \cdot 10^{-9}$	39 (6682)	0.1524	0.2791	$5.297 \cdot 10^{-4}$	31107.3
$L = \nabla$	$0.4 \cdot 10^{-9}$	15 (4932)	0.1547	0.3306	$7.536 \cdot 10^{-4}$	31105.4
$L = \nabla^2$	$5.0 \cdot 10^{-10}$	77 (6704)	0.1485	0.3056	$1.448 \cdot 10^{-3}$	31105.8
$L = \nabla^2$	$4.0 \cdot 10^{-10}$	69 (9764)	0.1512	0.2684	$9.411 \cdot 10^{-4}$	31105.1
<i>Star11</i>						
$L = \mathbf{I}$	$2.7 \cdot 10^{-9}$	26 (5673)	0.1512	0.3595	$1.131 \cdot 10^{-3}$	31296.4
$L = \mathbf{I}$	$1.7 \cdot 10^{-9}$	20 (4621)	0.1442	0.3057	$5.040 \cdot 10^{-4}$	31296.9
$L = \nabla$	$1.3 \cdot 10^{-9}$	54 (9057)	0.1481	0.2409	$8.211 \cdot 10^{-4}$	31295.5
$L = \nabla$	$0.3 \cdot 10^{-9}$	13 (5553)	0.1469	0.2464	$8.877 \cdot 10^{-5}$	31296.6
$L = \nabla^2$	$4.5 \cdot 10^{-10}$	87(10782)	0.1445	0.1845	$1.251 \cdot 10^{-4}$	31295.8
$L = \nabla^2$	$2.5 \cdot 10^{-10}$	45 (10057)	0.1456	0.1856	$6.999 \cdot 10^{-5}$	31296.4

**Table 5.12:** YSO. Numerical results obtained by the inexact Bregman iteration.  $\beta$  is the value for the regularization parameter,  $\kappa$  is the external iteration at which the best result is obtained;  $\rho^\kappa$  is the relative reconstruction error on the diffuse component,  $\rho_w^\kappa$  is the reconstruction error on the window of interest and finally  $\rho_P^\kappa$  is the relative reconstruction error on the values of the point source.  $\bar{\varphi}_0$  stands for  $\mathbf{KL}(\bar{\mathbf{H}}\bar{\mathbf{x}}^\kappa + \mathbf{bg}\mathbf{1}; \mathbf{gn})$ . For each regularization functional we tested the procedure for two different values of  $\beta$ : the first one is the value found by Model 1, while the other is another overestimation of  $\beta_{\text{opt}}$ .



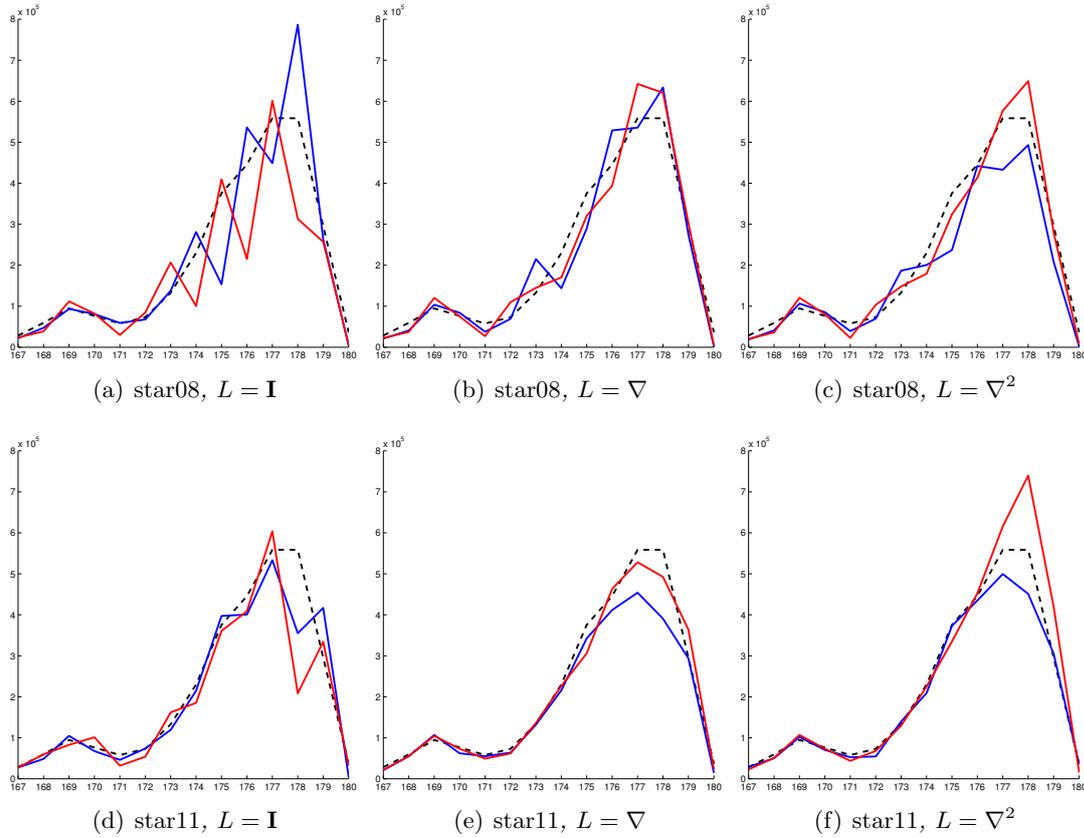
**Figure 5.11:** Restored images for Star09 problem. In the first row the whole reconstructions are displayed, while in the second one a particular related to the area near the point source is presented. The blue square emphasize the region in which we have computed the  $\rho_w$  relative reconstruction error.

From the point of view of the efficiency, the inexact Bregman scheme is not too expensive; indeed, because of the moderate tolerance required in the determination of the outer iterates, the total number of SGP iterations is two or three times greater than the one required to solve with high accuracy the variational problem (5.7) with an optimal value of the regularization parameter.

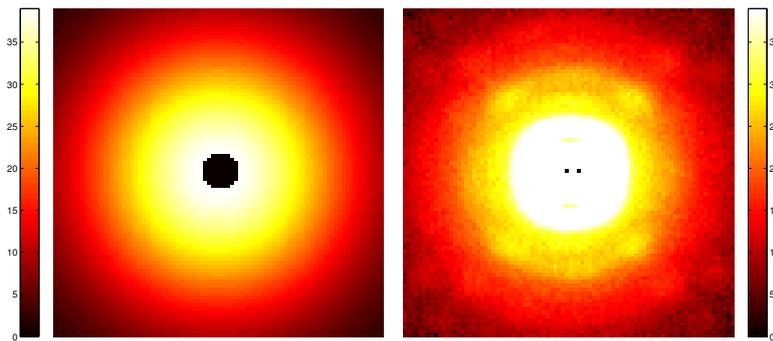
### 5.4.2 Binary Stars

The second case we are considering is the simulated image of binary stars obtained by a telescope. In Figure 5.13 the original signal and the given data  $\mathbf{gn}$  are shown: in both of them, the pixels corresponding to the stars are set to zero to make the visualization easier.

The background emission is 101.14 counts/pixel and, again, the noise affecting the image is a mixture of Poisson noise and RON with 10 counts/pixel. Actually the images shown in Figure 5.13 are a small part of the original images: since the original ones have dimension  $1024 \times 1024$ , we decided to display only the window of main interest, around the binary stars; this window will be the one on which we will compute the reconstruction error  $\rho_w$ . The window considered ranges from the 470-th pixel to the 555-th one in both dimensions.



**Figure 5.12:** Star08 (1st row) and Star11 (2nd row) problems. From left to right the lineplot of the column 129 is shown, for classical Tikhonov regularization, for  $L = \nabla$  and for  $L = \nabla^2$ , respectively. The dashed line is the original object, the red one is the reconstruction obtained by SGP with the optimal value for  $\beta$  and finally the blue line is the Bregman reconstruction.



**Figure 5.13:** Left panel: object, diffuse component. Right panel: given data, diffuse component. The images are displayed in a square root scale.

The structure of the data suggests to employ the following problem

$$\min_{\bar{\mathbf{x}} \geq 0} \mathbf{KL}(\bar{\mathbf{H}}\bar{\mathbf{x}} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \frac{\beta}{2} \|\mathbf{x}_d\|^2 \quad (5.11)$$

i.e. using a Tikhonov regularization since the diffuse component seems to have a smooth structure.

Proceeding as in the case of the Young Stellar Objects, we run the SGP method to solve problem (5.11) for several values of  $\beta$ , in order to find the value of the regularization parameter which gives the minimum  $\rho_w$ . The stopping criterion implemented is the standard one (5.10) widely used in the previous experiments, with  $\tau_{\text{SGP}} = 10^{-7}$ . The optimal value for  $\beta$  obtained is  $5 \cdot 10^{-6}$ , which gives a reconstruction error on the considered window of  $\rho_w = 0.13392$ .

In the first run of the inexact Bregman iteration we set the regularization parameter  $\beta = 10\beta_{\text{opt}}$ ; the first subproblem is stopped with the criterion (5.10) with  $\tau_{\text{SGP}} = 10^{-7}$  and  $M = 1$ . The second inner subproblem (and the successive ones) are stopped using the rule (4.34) with  $\alpha = 1.5$  and  $c = \|\eta_1\|$ .

In Table 5.13 we present the results obtained in this first test: although with a slightly

$k$	$it$	$totit$	$\rho_w$	$\bar{\varphi}_0$
1	145	145	0.20066	527342.28
2	113	258	0.14126	521991.75
3	126	384	0.13395	521692.73
4	61	445	0.13245	521627.44
5	54	499	0.13217	521595.31
6	61	560	0.13236	521576.04
7	55	615	0.13281	521562.26
8	61	676	0.13342	521551.86
9	69	745	0.13413	521543.62
10	66	811	0.13489	521536.84

**Table 5.13:** Binary stars; Bregman procedure with Tikhonov regularization. Results obtained with  $\beta = 10\beta_{\text{opt}}$ .  $k$  denotes the external iteration,  $it$  the number of iterations required by the inner solver to satisfy the convergence criterion,  $totit$  is the cumulative sum of the inner iterations,  $\rho_w$  is the relative reconstruction error of the window of major interest; finally,  $\bar{\varphi}_0 = \text{KL}(\bar{\mathbf{H}}\bar{\mathbf{x}}^k + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n})$

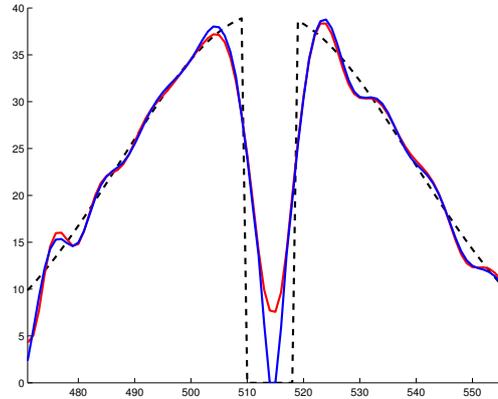
greater number of total iterations, the Bregman procedure seems to give a reconstruction error  $\rho_w$  comparable to the one obtained by tuning SGP method in an optimal way. The minimum error is reached at the 5–th external iteration; the inner solver requires a total number of 499 inner iterations versus the 178 ones of SGP method with the optimal regularization parameter.

In a second run, we employ again the Bregman procedure with the same settings as the previous ones, but we use a greater overestimate for the regularization parameter  $\beta = 10^{-4}$ .

Due to this great overestimation, in this experiment the procedure needs more external iterations to recover the image, since we are forcing an over regularization on the computed solution. Nevertheless, the numerical results seem comparable, if not better, to the reconstruction obtained via the optimal–tuned SGP.

In order to visually evaluate the reconstruction, in Figure 5.14 we plot the superposition of the 513rd line (the one containing the stars) of the images of the original signal, the reconstruction obtained via SGP with  $\beta_{\text{opt}}$  and the results of the running of the

Bregman procedure, with  $\beta = 10^{-4}$ : we do not plot the reconstruction with  $10\beta_{\text{opt}}$  since the differences are minimal. The Bregman procedure enables us to better devise



**Figure 5.14:** *Binary stars, line plots. The black dashed line represents the 513rd line of the original signal; the red one is the SGP reconstruction with optimal  $\beta$ ; the blue one is the Bregman reconstruction with  $\beta = 10^{-4}$ . The reconstructions are obtained by Tikhonov regularization. The lines are plotted in a sqrt scale.*

the hole at the center of the image: although the peaks at the top are not reached completely, the bottom of this hole is recognized.

One can observe that the Tikhonov regularization seems not to be the suitable functional to be used, since the boundaries of the hole at the centre of the image are sharp: indeed the portion reconstructed at the bottom level is too tight, even it is reached pretty well. Hence, due to the presence of the sharp edges, we can use the Hypersurface functional as regularization term, modifying (5.11) in

$$\min_{\bar{\mathbf{x}} \geq 0} \mathbf{KL}(\bar{\mathbf{H}}\bar{\mathbf{x}} + b\mathbf{g}\mathbf{1}; \mathbf{g}\mathbf{n}) + \beta\mathbf{HS}(\mathbf{x}_d) \quad (5.12)$$

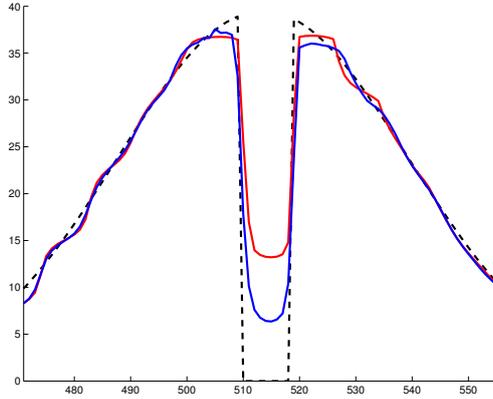
We proceed as in the Tikhonov case: we empirically determine an optimal value for the regularization parameter in (5.12), which turns out to be  $5.97 \cdot 10^{-3}$ ; solving (5.12) with this parameter ( $\tau_{\text{SGP}} = 10^{-7}$  and  $M = 1$ ) we obtain in 133 iterations a reconstruction error  $\rho_w$  of 0.101427. The numerical results confirm the fact that the Hypersurface regularization seems to be more indicated for this kind of image.

In order to evaluate the behaviour of the Bregman procedure also for this kind of regularization, we set the tolerance for the stopping criterion of the first subproblem with a mild tolerance  $\tau_{\text{SGP}} = 10^{-5}$  and  $M = 1$ ; then, for the stopping rule (4.34)  $c$  is equal to  $\|\eta_1\|$  and  $\alpha = 1.001$ . The regularization parameter is ten times the optimal one; the results are shown in Table 5.14. The obtained reconstruction error is lower than the one reached by the optimally tuned SGP method: moreover, even if the number of outer iterations seems huge (33), the total number of the iterations of the inner solvers is just 372, still comparable to the number of SGP iterations.

Proceeding as in the previous case, we force the overestimation by taking a regularization parameter  $15\beta_{\text{opt}}$ ; the setup of the procedure is the same of the case  $\beta = 10\beta_{\text{opt}}$ . One can observe that also in this case the reconstruction is still comparable with the one obtained by SGP with  $\beta_{\text{opt}}$ . In Figure 5.15 we show the superposition of the lines 513rd of the restored images. Using the Hypersurface regularization we do not reach

$k$	$it$	$totit$	$\rho_w$	$\bar{\varphi}_0$	$k$	$it$	$totit$	$\rho_w$	$\bar{\varphi}_0$
1	121	121	0.15370	524716.31	11	5	204	0.08560	522462.80
2	9	130	0.13010	523867.63	12	7	211	0.08516	522428.43
3	7	137	0.11278	523337.16	13	5	216	0.08484	522399.59
4	8	145	0.10254	523077.85	14	7	223	0.08449	522371.75
5	11	156	0.09514	522885.22	15	7	230	0.08418	522346.75
6	12	168	0.09118	522758.38	16	7	237	0.08391	522323.96
7	8	176	0.08914	522672.15	17	7	244	0.08367	522302.98
8	9	185	0.08754	522598.15	18	7	251	0.08345	522283.58
9	9	194	0.08655	522542.04	19	7	258	0.08326	522265.55
10	5	199	0.08602	522499.56	20	7	265	0.08309	522248.71

**Table 5.14:** Binary stars; Bregman procedure with  $\varphi_1 = \mathbf{HS}$  and  $\beta = 10\beta_{\text{opt}}$ .  $k$  denotes the external iteration,  $it$  the number of iterations required by the inner solver to satisfy the convergence criterion,  $totit$  is the cumulative sum of the inner iterations,  $\rho_w$  is the relative reconstruction error of the window of major interest; finally,  $\bar{\varphi}_0 = \mathbf{KL}(\overline{\mathbf{H}\mathbf{x}}^k + \mathbf{bg}\mathbf{1}; \mathbf{gn})$

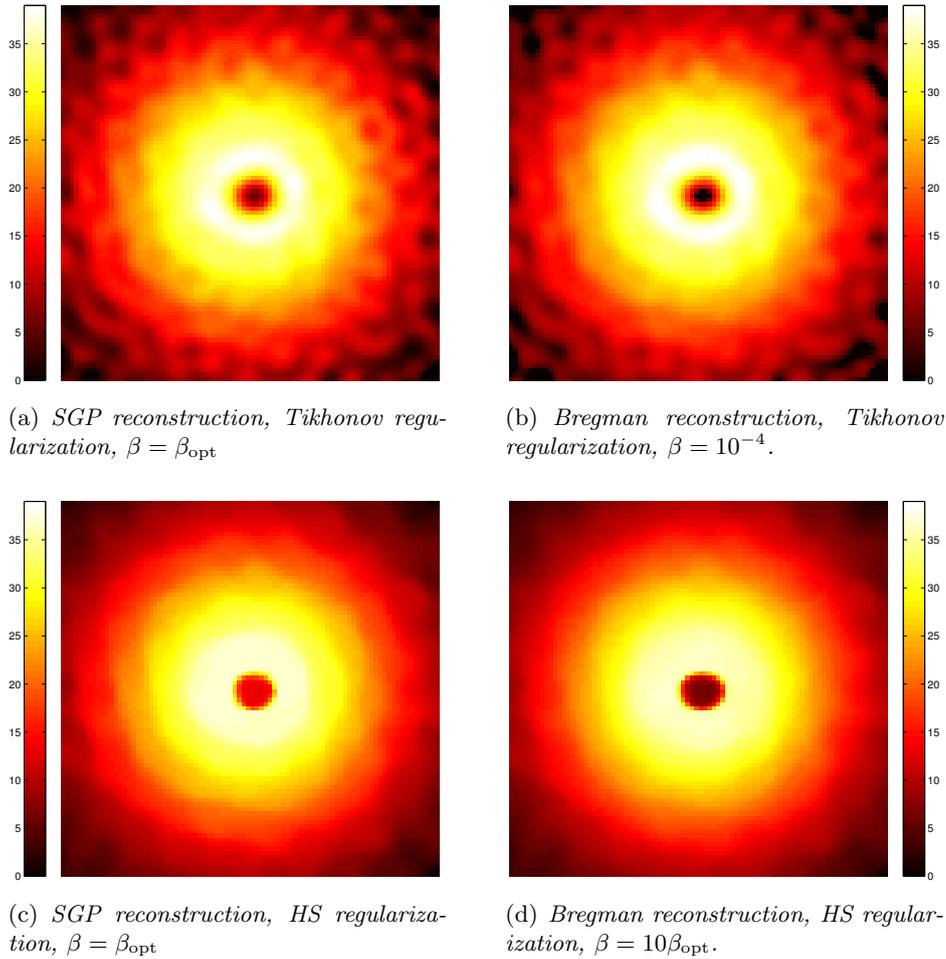


**Figure 5.15:** Binary stars, line plots. The black dashed line represents the 513rd line of the original signal; the red one is the SGP reconstruction with optimal  $\beta$ ; the blue one is the Bregman reconstruction with  $\beta = 10\beta_{\text{opt}}$ . The reconstructions are obtained by Hypersurface regularization. The lines are plotted in a sqrt scale.

the bottom of the hole neither the peaks on the boundaries, but the "flatness" is recognized in a better way, proving then a contrast enhancement. Finally, in Figure 5.16, we present the restored images obtained both with Tikhonov and the Hypersurface regularization, via the SGP method with  $\beta_{\text{opt}}$  and via the Bregman procedure.

## 5.5 Conclusion

In this chapter, we evaluate the inexact Bregman procedure, developed in the previous Chapter for image restoration in presence of Poisson noise. We show that this procedure has a lower computational cost with respect to the classical Bregman procedure. Both procedures enable us to employ an overestimation of the regularization parameter  $\beta$  appearing in the variational formulation. Although this overestimation, the results achieved are reliable with the ones obtained by using the optimal parameter. More-



**Figure 5.16:** Binary stars, restored images. In the first row the results obtained by using the Tikhonov regularization are shown, while in the second one the Hypersurface potential is used. The HS regularization induces a sort of smoothness in the diffuse component  $\mathbf{x}_d$ , even if the central hole is not fully restored.

over, above all in case of denoising, the contrast enhancement behaviour of the inexact Bregman procedure provides us with very satisfying reconstruction.

We have also compared the behaviour of Constrained and Crossing models with this inexact procedure. We recall that this two procedures allows us to estimate the optimal parameter and, at the same time, to obtain the related reconstruction. We observe that in some practical cases, such as the High Dynamic Range images in Astronomical imaging, these two approaches can not achieve good results, while the inexact Bregman procedure provides us with very good restored images, again with a gain in contrast enhancement.

# Appendix A

## Convex Analysis

This appendix is devoted to recall some basic concepts of Convex Analysis. We refer to [84, 78] for a very exhaustive and comprehensive overview.

### A.1 Definitions and general results

**Definition A.1.** Let  $f$  be a function defined on  $\mathbb{R}^n$  with real values; we denote with  $\text{dom}(f)$  the set  $\{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) < \infty\}$ . The epigraph  $\text{epi}(f)$  is the set

$$\text{epi}(f) = \{(\mathbf{x}, t) \in \text{dom}(f) \times \mathbb{R} \mid f(\mathbf{x}) \leq t\}$$

**Definition A.2.** A function  $f$  is convex if and only if its epigraph is a convex set.

From the definition above one has the following equivalent characterizations for a convex function.

- A function  $f$  is convex if and only if its domain is convex and for any  $\mathbf{x}, \mathbf{y} \in \text{dom}(f)$  and  $\alpha \in [0, 1]$  the following inequality holds

$$f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{y})$$

- A function  $f$  is convex if and only if for all  $\mathbf{x}, \mathbf{y} \in \text{dom}(f)$  and  $\beta \geq 0$  such that  $\mathbf{y} + \beta(\mathbf{y} - \mathbf{x}) \in \text{dom}(f)$ , we have

$$f(\mathbf{y} + \beta(\mathbf{y} - \mathbf{x})) \geq f(\mathbf{y}) + \beta(f(\mathbf{y}) - f(\mathbf{x})).$$

Below we state some general results regarding convex functions.

**Lemma A.1.** [Jensen inequality] For any  $\mathbf{x}_1, \dots, \mathbf{x}_m \in \text{dom}(f)$  and  $\alpha_1, \dots, \alpha_m$  such that

$$\sum_{i=1}^m \alpha_i = 1, \quad \alpha_i \geq 0, \quad i = 1, \dots, m$$

we have:

$$f\left(\sum_{i=1}^m \alpha_i \mathbf{x}_i\right) \leq \sum_{i=1}^m \alpha_i f(\mathbf{x}_i)$$

**Lemma A.2.** If  $f$  is a convex function, then all its sublevel sets

$$\text{lev}_\tau(f) = \{\mathbf{x} \in \text{dom}(f) \mid f(\mathbf{x}) \leq \tau\}$$

are either convex or empty.

**Definition A.3.** A function  $f$  is called lower semicontinuous (l.s.c.) at  $\mathbf{x}$  if and only if

$$f(\mathbf{x}) = \liminf_{\mathbf{y} \rightarrow \mathbf{x}} f(\mathbf{y}) = \lim_{\epsilon \rightarrow 0} (\inf \{f(\mathbf{y}) \mid \|\mathbf{y} - \mathbf{x}\| \geq \epsilon\})$$

$f$  is upper semicontinuous (u.s.c.) if

$$f(\mathbf{x}) = \limsup_{\mathbf{y} \rightarrow \mathbf{x}} f(\mathbf{y}) = \lim_{\epsilon \rightarrow 0} (\sup \{f(\mathbf{y}) \mid \|\mathbf{y} - \mathbf{x}\| \geq \epsilon\})$$

**Definition A.4.** A convex function  $f$  is called closed if its epigraph is a closed set.

The following result is a straightforward consequence of the previous definition.

**Proposition A.1.** If a convex function  $f$  is closed, then all its sublevel sets are either empty or closed.

**Proposition A.2.** Let  $f_1$  and  $f_2$  be two convex and closed functions,  $\beta \geq 0$ . Then the subsequent functions are all closed and convex:

- $f(\mathbf{x}) = \beta f_1(\mathbf{x})$ ,  $\text{dom}(f) = \text{dom}(f_1)$ ;
- $f(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x})$ ,  $\text{dom}(f) = \text{dom}(f_1) \cap \text{dom}(f_2)$
- $f(\mathbf{x}) = \max \{f_1(\mathbf{x}), f_2(\mathbf{x})\}$ ,  $\text{dom}(f) = \text{dom}(f_1) \cap \text{dom}(f_2)$

**Proposition A.3.** Let  $\varphi(\mathbf{x})$  be a convex and closed function defined for  $\mathbf{x} \in \mathbb{R}^m$ . Consider the affine operator

$$\begin{aligned} \mathcal{A} &: \mathbb{R}^n \rightarrow \mathbb{R}^m \\ \mathcal{A} &: \mathbf{x} \mapsto A(\mathbf{x}) + b \end{aligned}$$

where  $A \in \mathcal{M}_{n \times m}(\mathbb{R})$ ,  $b \in \mathbb{R}^m$ . Then  $f(\mathbf{x}) = \varphi(\mathcal{A}(\mathbf{x}))$  is convex and closed on the domain  $\text{dom}(f) = \{\mathbf{x} \in \mathbb{R}^n \mid \mathcal{A}(\mathbf{x}) \in \text{dom}(\varphi)\}$ .

**Definition A.5.** Let  $f$  be a convex function. The Fenchel conjugate of  $f$  is the function  $f^*$  defined by

$$f^*(\mathbf{y}) = \sup_{\mathbf{x} \in \mathbb{R}^n} \{\langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x})\}$$

$f^*$  has the following properties:

- $f^*$  is a closed convex function;
- if  $f$  is proper then  $f^*$  is proper;
- $f^{**} = \text{cl}(f)$ , where  $\text{cl}(f)$  is the closure of the function  $f$ ;
- if  $f$  is l.s.c and proper, then  $f^*$  is l.s.c. and  $f^{**} = f$ .

## A.2 $\varepsilon$ -Subgradients

**Definition A.6.** Let  $f$  be a convex function on  $\mathbb{R}^n$ . A vector  $\mathbf{p}$  is called a subgradient of  $f$  at the point  $\mathbf{x}_0 \in \text{dom}(f)$  if for any  $\mathbf{x} \in \text{dom}(f)$  we have

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) + \langle \mathbf{p}, \mathbf{x} - \mathbf{x}_0 \rangle$$

The set of all the subgradient of  $f$  at  $\mathbf{x}_0$  is called the subdifferential of  $f$  at the point  $\mathbf{x}_0$  and it is denoted as  $\partial f(\mathbf{x}_0)$ .

**Example A.1.** Let  $f(\mathbf{x}) = |\mathbf{x}|$ ,  $\mathbf{x} \in \mathbb{R}$ . Then  $\partial f(0) = [-1, 1]$ ; the subgradient of  $f$  in the origin is not unique.

**Proposition A.4.** Let  $f$  be a closed convex function and  $\mathbf{x} \in \text{int}(\text{dom}(f))$ . Then  $\partial f(\mathbf{x}_0)$  is a nonempty bounded set.

**Proposition A.5.** For any proper convex function  $f$  and any vector  $\mathbf{x}$ , the following assertions on a vector  $\mathbf{x}^*$  are equivalent to each other:

- ▷  $\mathbf{x}^* \in \partial f(\mathbf{x})$ ;
- ▷  $\langle \mathbf{z}, \mathbf{x}^* \rangle - f(\mathbf{z})$  achieves its supremum in  $\mathbf{z}$  at  $\mathbf{z} = \mathbf{x}$ ;
- ▷  $f(\mathbf{x}) + f^*(\mathbf{x}^*) \leq \langle \mathbf{x}, \mathbf{x}^* \rangle$
- ▷  $f(\mathbf{x}) + f^*(\mathbf{x}^*) = \langle \mathbf{x}, \mathbf{x}^* \rangle$

If  $\text{cl}f(\mathbf{x}) = f(\mathbf{x})$ , three conditions can be added:

- ▷  $\mathbf{x} \in \partial f^*(\mathbf{x}^*)$ ;
- ▷  $\langle \mathbf{x}, \mathbf{z} \rangle - f^*(\mathbf{z})$  achieves its supremum in  $\mathbf{z}$  at  $\mathbf{z} = \mathbf{x}^*$ ;
- ▷  $\mathbf{x}^* \in \partial \text{cl} f(\mathbf{x})$

**Proposition A.6.** If  $f$  is a closed proper convex function,  $\partial f^*$  is the inverse of  $\partial f$  in the sense of multivalued mappings, i.e  $\mathbf{x} \in \partial f^*(\mathbf{x}^*)$  if and only if  $\mathbf{x}^* \in \partial f(\mathbf{x})$ .

The following propositions regard some properties of subgradients, which are of main importance in optimization.

**Proposition A.7.** We have  $f(\mathbf{x}^*) = \min_{\mathbf{x} \in \text{dom}(f)} f(\mathbf{x})$  if and only if  $0 \in \partial f(\mathbf{x}^*)$ .

**Proposition A.8.** Let  $Q \subset \text{dom}(f)$  be a closed convex set,  $\mathbf{x}_0 \in Q$  and

$$\mathbf{x}^* = \arg \min \{f(\mathbf{x}) | \mathbf{x} \in Q\}$$

Then for any  $\mathbf{p} \in \partial f(\mathbf{x}_0)$  we have:  $\langle \mathbf{p}, \mathbf{x}_0 - \mathbf{x}^* \rangle \geq 0$ .

The theoretical results in the forthcoming allows us to effectively compute the subgradients of convex functions.

**Lemma A.3.** Let  $f$  be a closed convex function. Assume that it is differentiable on its domain. Then  $\partial f(\mathbf{x}) = \{f'(\mathbf{x})\}$  for any  $\mathbf{x} \in \text{int}(\text{dom}(f))$ .

**Proposition A.9.** Let function  $f$  be closed and convex with  $\text{dom}(f) \subset \mathbb{R}^m$ . Consider the affine operator

$$\begin{aligned} \mathcal{A} &: \mathbb{R}^n \rightarrow \mathbb{R}^m \\ \mathcal{A} &: \mathbf{x} \mapsto A(\mathbf{x}) + b \end{aligned}$$

where  $A \in \mathcal{M}_{n \times m}(\mathbb{R})$ ,  $b \in \mathbb{R}^m$ . Then the function  $\varphi(\mathbf{x}) = f(\mathcal{A}(\mathbf{x}))$  is closed and convex with domain  $\text{dom}(\varphi) = \{\mathbf{x} \in \mathbb{R}^n \mid \mathcal{A}(\mathbf{x}) \in \text{dom}(f)\}$  and for any  $x \in \text{int}(\text{dom}(\varphi))$  we have:

$$\partial\varphi(\mathbf{x}) = A^t \partial f(A\mathbf{x})$$

**Lemma A.4.** Let  $f_1(\mathbf{x})$  and  $f_2(\mathbf{x})$  are closed convex functions and  $\alpha_1, \alpha_2 \geq 0$ . Then the function

$$f(\mathbf{x}) = \alpha_1 f_1(\mathbf{x}) + \alpha_2 f_2(\mathbf{x})$$

is closed and convex and

$$\partial f(\mathbf{x}) = \alpha_1 \partial f_1(\mathbf{x}) + \alpha_2 \partial f_2(\mathbf{x})$$

for  $\mathbf{x} \in \text{int}(\text{dom} f) = \text{int}(\text{dom}(f_1)) \cap \text{int}(\text{dom}(f_2))$ .

**Definition A.7.** Let  $f$  be a convex function on  $\mathbb{R}^n$ . A vector  $\mathbf{p}$  is called a  $\varepsilon$ -subgradient of  $f$  at the point  $\mathbf{x}_0 \in \text{dom}(f)$  if for any  $\mathbf{x} \in \text{dom}(f)$  we have

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) + \langle \mathbf{p}, \mathbf{x} - \mathbf{x}_0 \rangle - \varepsilon$$

The set of all the subgradient of  $f$  at  $\mathbf{x}_0$  is called the  $\varepsilon$ -subdifferential of  $f$  at the point  $\mathbf{x}_0$  and it is denoted as  $\partial_\varepsilon f(\mathbf{x}_0)$ .

We observe that for  $\varepsilon = 0$ , we obtain the definition of subgradient and subdifferential.

**Proposition A.10.** Let  $f$  be a convex function on  $\mathbb{R}^n$  and  $\varepsilon_1 > \varepsilon_2 > 0$ . Then, for any  $\mathbf{x}_0 \in \text{dom}(f)$ , we have

$$\partial_{\varepsilon_1} f(\mathbf{x}_0) \supset \partial_{\varepsilon_2} f(\mathbf{x}_0) \supset \partial f(\mathbf{x}_0)$$

# Bibliography

- [1] Ya I Alber, A N Iusem, and M V Solodov. On the projected subgradient method for nonsmooth convex optimization in a Hilbert space. *Math. Prog.*, 81:23–35, 1998.
- [2] K J Arrow, L Hurwicz, and H Uzawa. *Studies in Linear and Non-Linear Programming*. Stanford University Press, Stanford, 1958.
- [3] J F Aujol, G Aubert, L Blanc-Feraud, and A Chambolle. Decomposition into a bounded variation component and an oscillating component. *Journal of Mathematical Imaging and Vision*, 22:71–88, 2005.
- [4] H Ayasso, T Rodet, and A Abergel. A variational Bayesian approach for unsupervised super-resolution using mixture models of point and smooth sources applied to astrophysical map-making. *Inverse Problems*, 28(12):125005, 2012.
- [5] P L Bartlett, M I Jordan, and J D McAuliffe. Convexity, classification and risk bounds. *Journal of the American Statistical Association*, 101(473):138–156, 2006.
- [6] J Barzilai and J M Borwein. Two point step size gradient methods. *IMA Journal of Numerical Analysis*, 8(1):141–148, 1988.
- [7] H H Bauschke and P L Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. CMS Books on Mathematics. Springer, 2011.
- [8] A Benfenati and V. Ruggiero. Inexact Bregman iteration for deconvolution of superimposed extended and point sources. *Communications in Nonlinear Science and Numerical Simulation*, 20:882–896.
- [9] A Benfenati and V. Ruggiero. Inexact Bregman iteration with an application to Poisson data reconstruction. *Inverse Problems*, 29(6):1–32, 2013.
- [10] M Bertero. Private Communications.
- [11] M Bertero and P Boccacci. *Introduction to Inverse Problems in Imaging*. Institute of Physics Publishing, Bristol and Philadelphia, 1998.
- [12] M Bertero, P Boccacci, G Desiderà, and G Vicidomini. Image deblurring with Poisson data: from cells to galaxies. *Inverse Problems*, 25:123006, 2009.
- [13] M Bertero, P Boccacci, G Talenti, R Zanella, and L Zanni. A discrepancy principle for Poisson data. *Inverse Problems*, 26:10500, 2010.

- 
- [14] M Bertero, H Lanteri, and L Zanni. Iterative image reconstruction: a point of view. *Proceedings of the Interdisciplinary Workshop on Mathematical Methods in Biomedical Imaging and Intensity-Modulated Radiation*, 2007.
- [15] D Bertsekas. *Convex Optimization Theory*, chapter 6 on Convex Optimization Algorithms, pages 251–489. Athena Scientific, Belmont, Massachusetts, 5 november 2012 edition, June 2009.
- [16] D Bertsekas and J Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Prentice-Hall, 1989.
- [17] S Bonettini, A Cornelio, and M Prato. A new semi-blind deconvolution approach for Fourier-based image restoration: an application in astronomy. *SIAM Journal on Imaging Sciences*, 6:1736–1757, 2013.
- [18] 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.
- [19] 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.
- [20] S Bonettini and V Ruggiero. An alternating extragradient method for Total Variation based image restoration from Poisson data. *Inverse Problems*, 27:095001, 2011.
- [21] 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.
- [22] S Bonettini, R Zanella, and L Zanni. A scaled gradient projection method for constrained image deblurring. *Inverse Problems*, 25:015002, 2009.
- [23] S Boyd, N Parikh, E Chu, B Peleato, and J Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, (3(1)):1–122, 2011.
- [24] U Brannlund, K C Kiwiel, and P O Lindberg. A descent proximal level bundle method for convex nondifferentiable optimization. *Oper. Res. Lett.*, (17):121–126, 1995.
- [25] L M Bregman. 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.
- [26] P Brucker. An  $\mathcal{O}(n)$  algorithm for quadratic knapsack problems. *Operations Research Letters*, 3(3):163–166, 1984.
- [27] A M Bruckstein, D L Donoho, and M Elad. From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Review*, 51(1):34–81, 2009.

- 
- [28] C Brune, A Sawatzky, and M Burger. *Scale Space and Variational Methods in Computer Vision*, volume 5567 of *LNCS*, chapter Bregman–TV–EM methods with application to optical nanoscopy, pages 235–246. Springer, 2009.
- [29] C Brune, A Sawatzky, and M Burger. Primal and dual Bregman methods with application to optical nanoscopy. *Int. J. Comput. Vis.*, 92(2):211–229, 2010.
- [30] C Brune, A Sawatzky, T Kösters, F Wübbeling, and M Burger. Forward-backward EM-TV methods for inverse problems with Poisson noise. 2011.
- [31] A La Camera, S Antonucci, M Bertero, P Boccacci, D Lorenzetti, and B Nisini. Image reconstruction for observations with an high dynamic range: LINC–NIRVANA simulations of a stellar jet. 8455:8455 3D, 2012.
- [32] A La Camera, S Antonucci, M Bertero, P Boccacci, D Lorenzetti, B Nisini, and C Arcidiacono. Reconstruction of high dynamic range images: Simulations of LBT observations of a stellar jet, a pathfinder study for future AO-assisted giant telescopes. *Publications Astron. Soc. Pacific*, 126:180–193, 2014.
- [33] M Carlván and L Blanc-Féraud. Two constrained formulations for deblurring Poisson noisy images. *Proc. IEEE International Conference on Image Processing (ICIP), Brussels, Belgium*, 2011.
- [34] M Carlván and L Blanc-Féraud. Sparse Poisson noisy image deblurring. *IEEE Transactions on Image Processing*, 21(4):1834–1846, 2012.
- [35] A Chambolle. An algorithm for Total Variation minimization and applications. *J. Math. Imag. Vis.*, 20:89–97, 2004.
- [36] A Chambolle. Total variation minimization and a class of binary mrf models. *EMMCVPR 05. Lecture Notes in Computer Sciences*, 3757:136–152, 2005.
- [37] 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.
- [38] P Charbonnier. *Reconstruction d’image: Regularization avec prise en compte des discontinuités*. PhD thesis, Univ. Nce-Sophia Antipolis, 1994.
- [39] P Charbonnier, L Blanc-Féraud, G Aubert, and A Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Trans. Image Processing*, 6:298–311, 1997.
- [40] R Ciak, B Shafei, and G Steidl. Homogeneous penalizers and constraints in convex image restoration. *J. Math. Imag. Vis.*, (47 (3)):210–230, 2012.
- [41] P L Combettes. Quasi-Féjerian analysis of some optimization algorithms. *Studies in Computational Mathematics*, 8(C):115–152, 2001.
- [42] P L Combettes and B C Vũ. Variable metric quasi-Féjer monotonicity. *Nonlinear Analysis: Theory, Methods, and Applications*, 78:17–31, feb 2013.

- [43] P L Combettes and B C Vũ. Variable metric forward-backward splitting with applications to monotone inclusions in duality. *Optimization*, 63(9):1289–1318, 2014.
- [44] A Cornelio, E Loli Piccolomini, and J Nagy. Constrained numerical optimization methods for blind deconvolution. *Numerical Algorithms*, 65:23–42, 2014.
- [45] Y H Dai and R Fletcher. New algorithms for singly linearly constrained quadratic programs subject to lower and upper bounds. *Math. Program., Ser. A*, 106:403–421, 2006.
- [46] J Eckstein. Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming. *Mathematics of Operations Research*, 18(1):202–226, 1993.
- [47] H W Engl, M Hanke, and A Neubauer. *Regularization of Inverse Problems*. Kluwer, Dordrecht, 1996.
- [48] Yu. M Ermoliev. Stochastic quasigradient methods and their application to system optimization. *Stochastics*, 9:1–36, 1983.
- [49] M A T Figueiredo and J M Bioucas-Dias. Deconvolution of Poissonian images using variable augmented Lagrangian optimization. *IEEE Workshop on Statistical Signal Processing, Cardiff*, 2009.
- [50] M A T Figueiredo and J M Bioucas-Dias. Restoration of Poissonian images using alternating direction optimization. *IEEE Transactions on Image Processing*, 19(12):3133–3145, 2010.
- [51] R Fletcher. A limited memory steepest descent method. *Math. Progr.*, 135(1–2):413–436, 2012.
- [52] G Frassoldati, L Zanni, and G Zanghirati. New adaptive stepsize selections in gradient methods. *Journal of Industrial and Management Optimization*, 4(2):299–312, 2008.
- [53] S Geman and D Geman. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.
- [54] S Geman, K Manbeck, and D E McClure. *Markov Random Fields: Theory and Applications*, chapter A comprehensive statistical model for single photon emission tomography, pages 93–130. Academic Press, 1993.
- [55] J F Giovanelli and A Coulais. Positive deconvolution for superimposed extended source and point sources. *Astr. Astrophys.*, 439:402–412, 2005.
- [56] J L Goffin and K C Kiwiel. Convergence of a simple subgradient level method. *Mathematical Programming*, (85):207–211, 1999.
- [57] J Hadamard. Sur les problemes aux derives partielles et leur signification physique. *Princeton University Bulletin*, pages 49–52, 1902.

- [58] M Hanke and P C Hansen. Regularization methods for large-scale problems. *Surveys on Mathematics for industry*, 3:189–199, 1993.
- [59] P C Hansen. Numerical tools for analysis and solution of fredholm integral equations of first kind. *Inverse Problems*, 8:956–972, 1992.
- [60] P C Hansen, J G Nagy, and D P O’Leary. *Deblurring images: matrices, spectra, and filtering*. SIAM, 2006.
- [61] P C Hansen and P D O’Leary. The use of l-curve in the regularization of discrete ill-posed problems. *SIAM Journal of scientific Computing*, 14:1487–1503, 1993.
- [62] T Hastie, R Tibshirani, and J Friedman. *The elements of statistical learning: Data mining, inference and prediction*. 2009.
- [63] B S He, Y Yang, and S L Wang. Alternating direction method with self-adaptive penalty parameters for monotone variational inequalities. *Journal of Optimization Theory and Applications*, (106–(2)):337–356, 2000.
- [64] J B Hiriart and C Lemarechal. *Convex Analysis and Minimization Algorithms*, volume vol. 1. 1993.
- [65] K C Kiwiel. Convergence of approximate and incremental subgradient methods for convex optimization. *SIAM J. Optim.*, 14(3):807–840, 2004.
- [66] K Koh, S J Kim, and S Boyd. An interior-point method for large-scale l-regularized logistic regression. *Journal of Machine Learning Research*, 1(8):1519–1555, 2007.
- [67] H Lanteri, M Roche, and C Aime. Penalized maximum likelihood image restoration with positivity constraints: multiplicative algorithms. *Inverse problems*, 18:1397–1419, 2002.
- [68] T Larsson, M Patriksson, and A-B Strömberg. On the convergence of conditional  $\epsilon$ -subgradient methods for convex programs and convex-concave saddle-point problems. *European Journal of Operational Research*, 151:461–473, 2003.
- [69] 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.
- [70] B Lucy. An iterative technique for the rectification of observed distributions. *Astronom. J.*, 79:745–754, 1974.
- [71] D G Luenberger and Y Ye. *Linear and Nonlinear Programming*. Springer, third edition edition, 2008.
- [72] N Maculan, C P Santiago, E M Macambira, and M H C Jardim. An  $\mathcal{O}(n)$  algorithm for projecting a vector on the intersection of a hyperplane and a box in  $\mathbb{R}^n$ . *Journal of Optimization Theory and Applications*, 117(3):553–574, 2003.
- [73] P Marcotte. Application of Khobotov’s algorithm to variational inequalities and network equilibrium problems. *INFOR.*, 29:258–270, 1991.

- [74] C De Mol and M Defrise. Inverse imaging with mixed penalties. In *Proceedings URSI EMTS 2004*, pages 798–800. Ed. PLUS Univ. Pisa, 2004.
- [75] A V Morozov. On the solution of functional equations by the method of regularization. *Soviet Math. Dokl.*, 7:414–417, 1996.
- [76] A Nedić. *Subgradient methods for convex minimization*. PhD thesis, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, 2002.
- [77] A Nedić and D P Bertsekas. Incremental subgradient methods for nondifferentiable optimization. *SIAM J. on Optim.*, 12:109–138, 2001.
- [78] Y Nesterov. *Introductory Lectures on Convex Optimization*. 87. Springer US, 1 edition, 2004.
- [79] S Osher, M Burger, D Goldfarb, J Xu, and W Yin. An iterative regularization method for total variation-based image restoration. *SIAM Journal on Multiscale Modeling and Simulation*, 4(2):460–489, 2005.
- [80] B Polyak. *Introduction to Optimization*. Optimization Software - Inc., Publication Division, N.Y., 1987.
- [81] F Porta, R Zanella, G Zanghirati, and L Zanni. Limited-memory scaled gradient projection methods for real-time image deconvolution in microscopy. *Communications in Nonlinear Science and Numerical Simulation*, (21):112–127, 2015.
- [82] B Reipurth, S Heathcote, J Morse, P Hartigan, and J Bally. Hubble space telescope images of the HH34 jet and bow shock: structure and proper motions. *Astron. J.*, 123:362–381, 2002.
- [83] W H Richardson. Bayesian-based iterative method of image restoration. *J. Opt. Soc. Amer. A*, 62(1):55–59, 1972.
- [84] R T Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, NJ, 1970.
- [85] R T Rockafellar. Monotone operators and the proximal point algorithms. *SIAM J. Control Optimization*, 14(5):877–898, 1976.
- [86] L Rudin, S Osher, and E Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
- [87] S Setzer, G Steidl, and T Teuber. Deblurring Poissonian images by split Bregman techniques. *J. Vis. Commun. Image R.*, 21:193–199, 2010.
- [88] L A Shepp and Y Vardi. Maximum likelihood reconstruction for emission tomography. *Trans. Med. Imaging*, MI-1:113–122, 1982.
- [89] D L Snyder, C W Helstrom, A D Lanterman, and R L White. Compensation for read-out noise in charge-coupled-device images. *J. of the Optical Society of America A*, pages 272–283, 1995.

- 
- [90] M V Solodov and S K Zaviev. Error stability properties of generalized gradient-type algorithms. *J. Opt. Theory and Appl.*, 98:663–680, 1998.
- [91] D M Strong and T F Chan. Exact solutions to Total Variation regularization problems. CAM Reports 96-41, UCLA Center for Applied Math., 1996.
- [92] T Teuber, G Steidl, and R H Chan. Minimization and parameter estimation for seminorm regularization models with I-divergence constraints. *Inverse Problems*, 29:035007, 2013.
- [93] A Tikhonov. On the stability of inverse problems. *Dokl. Akad. Nauk.*, 39:195–200, 1943.
- [94] A Tikhonov. Solution of incorrectly formulated problems and the regularization method. *Soviet Math. Dokl.*, 4:1035–1038, 1963.
- [95] A Tikhonov and V Y Arsenin. *Solution of ill-posed problems*. Winston, 1977.
- [96] V N Vapnik. *The Nature of Statistical Learning Theory*. Springer-Verlag, 2000.
- [97] C R Vogel. *Computational Methods for Inverse Problems*. SIAM, 2002.
- [98] S L Wang and L Z Liao. Decomposition method with a variable parameter for a class of monotone variational inequality problems. *Journal of Optimization Theory and Applications*, (109(2)):415–429, 2001.
- [99] R M Willett and R D Nowak. Platelets: A multiscale approach for recovering edges and surfaces in photon limited medical imaging. *IEEE Transactions on Medical Imaging*, 22:332–350, 2003.
- [100] W Yin, S Osher, D Goldfarb, and J Darbon. Bregman iterative algorithms for  $l_1$ -minimization with applications to compressed sensing. *Siam J. Imaging Sciences*, 1(1):143–168, 2008.
- [101] R Zanella, P Boccacci, L Zanni, and M Bertero. Efficient gradient projection methods for edge-preserving removal of Poisson noise. *Inverse Problems*, 25:045010, 2009.
- [102] L Zanni, A Benfenati, M Bertero, and V Ruggiero. Numerical methods for parameter estimation in poisson data inversion. *Journal of Mathematical Imaging and Vision*, 2014.
- [103] M Zhu and T F Chan. An efficient primal–dual hybrid gradient algorithm for Total Variation image restoration. CAM Report 08-34, UCLA, 2008.
- [104] M Zhu, S J Wright, and T F Chan. Duality-based algorithms for total-variation-regularized image restoration. *Computational Optimization and Applications*, 47:377–400, 2008.