Analisi di un metodo di minimizzazione alternata per problemi di deconvoluzione cieca in astronomia

Author: Simone Rebegoldi, Supervisor: Dr. Marco Prato

Università degli Studi di Modena e Reggio Emilia

1 Motivations

Blind deconvolution is the problem of image deblurring when both the original object and the blur are unknown. If we assume a space-invariant model, the naive problem formulation is to solve the equation

$$\boldsymbol{y} = \boldsymbol{K} * \boldsymbol{x}$$

where $\boldsymbol{y}$ is the detected image, $(\boldsymbol{x}, \boldsymbol{K})$ are respectively the unknown object and the unknown Point Spread Function (PSF) describing the blur of the imaging system and $*$ is the convolution product. It is obvious that this problem is extremely undetermined and that there is an infinite set of pairs solving the equation. Therefore the problem must be reformulated by introducing as far as possible all available *a priori* information on both the object and the PSF, in order to address the solving algorithm to the desidered solution.

In my degree thesis, I focused my attention on a particular astronomical imaging problem, in which images are acquired by using a Fizeau interferometer of the Large Binocular Telescope (LBT), known as LINC-NIRVANA (*Lbt INterferometric Camera and Near-InfraRed/Visible Adaptive iNterferometer for Astronomy*). In order to get an image with an isotropic resolution, LINC-NIRVANA acquires p images of the same astronomical object, which are convolved with p different PSFs. According to the maximum likelihood approach, the mathematical formulation of this imaging problem leads to a constrained minimization problem with $p+1$ blocks, namely the unknown object and the p point spread functions corresponding to the p acquisitions, which are subject to some constraints derived from the physical characteristics of the object and the technical features of the instrument. Due to the Poisson nature of the noise affecting the measured data, the objective function is the so-called Kullback-Leibler divergence of the unknown models from the acquired images. As it is known, this function is convex with respect to each block of variables for fixed values of the others, but is not convex with respect to the full set of variables. Therefore blind deconvolution is a difficult problem of non-convex optimization. However, the constraints have a nice separable structure, since they involve separately the blocks of variables, defining a feasible convex set for each of them. This allow to seek the solution of the problem by means of an inexact alternating minimization method, whose global convergence to stationary points of the objective function has been recently proved in a general setting [5]. The method is iterative and each iteration, called outer iteration, consists of alternating the updates of the object and the PSFs by means of fixed numbers of iterations, called inner iterations, of the scaled gradient projection (SGP) method.

In the case of single-dish telescopes, this alternating scheme has been introduced in [10] and has led to very promising results in the reconstruction of stellar fields, but it seems unable to provide satisfactory results in the case of diffuse objects. The final purpose of my degree thesis was to test the extension of this approach to the case of LINC-NIRVANA imaging by means of numerical experiments, in order to confirm its effectiveness in the case of stellar objects and to observe any improvement in the reconstruction of diffuse objects.

2 Method

First I considered the constrained optimization problem

$$\begin{aligned} &\min J(\boldsymbol{x}) \\ &\boldsymbol{x} \in \Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_m \subseteq \mathbb{R}^n \end{aligned} \tag{1}$$

where J is a differentiable objective function, Ω_i $(i = 1, \ldots, m)$ is a closed and convex subset of $\mathbb{R}^{n_i}$ with $n_1 + \ldots + n_m = n$ and any vector in the feasible set can be partitioned into vector components as $x = (x_1, x_2, \ldots, x_m)$, $x_i \in \Omega_i$.
The nonlinear Gauss-Seidel (GS) method, which is also known as the nonlinear block coordinate descent or the alternating optimization method, consists of solving (1) by successively minimizing the function f with respect to each block of variables over the corresponding constraints. Thus, given an initial point $\boldsymbol{x}^{(0)} \in \Omega$, for $k = 0, 1, 2, \ldots$, the iterate $\boldsymbol{x}^{(k+1)} = (\boldsymbol{x}_1^{(k+1)}, \ldots, \boldsymbol{x}_m^{(k+1)})$ is computed such that, for $i = 1, \ldots, m$, the block of variables $\boldsymbol{x}_i^{(k+1)}$ is a solution of the subproblem

$$\min_{\boldsymbol{y} \in \Omega_i} J(\boldsymbol{x}_1^{(k+1)}, \ldots, \boldsymbol{x}_{i-1}^{(k+1)}, \boldsymbol{y}, \boldsymbol{x}_{i+1}^{(k)}, \ldots, \boldsymbol{x}_m^{(k)}). \tag{2}$$

In [8], the authors proved that, when $m = 2$, every limit point of $\{\boldsymbol{x}^{(k)}\}$ is a critical point of (1). For $m \geq 3$, however, the convergence of the nonlinear GS method to a solution of (1) is not guaranteed, without additional convexity assumptions on the objective function J. Some convergence results are proved for example in [3, 8, 9] assuming the strict convexity of the function with respect to each block of variables.
All these convergence results, even in the case $m = 2$, are proved when each subproblem (2) is solved *exactly*, which is often impractical or too costly to compute. For that reason, in my degree thesis I considered a variation of the GS method, namely an *inexact* alternating scheme, where for $i = 1, \ldots, m$, the i-th subproblem (2) is solved *approximately* by means of a finite number of the scaled gradient projection (SGP) method, which applies to any problem of the form

$$\min_{\boldsymbol{x} \in \Omega} J(\boldsymbol{x}), \tag{3}$$

where Ω is a closed and convex set and $J(x) \in C^1(\Omega)$. Chosen an initial point $\boldsymbol{x}^{(0)} \in \Omega$, for $k = 0, 1, 2 \ldots$ the k-th iteration of the SGP method applied to problem (3) involves the computation of the feasible descent direction

$$\boldsymbol{d}^{(k)} = P_{\Omega,(D_k)^{-1}}(\boldsymbol{x}^{(k)} - \alpha_k D_k \nabla J(\boldsymbol{x}^{(k)})) - \boldsymbol{x}^{(k)}, \tag{4}$$

where D_k is a symmetric positive definite matrix, whose eigenvalues are bounded above and below by two constants that are independent of k, $P_{\Omega,(D_k)^{-1}}$ is the projection operator in the norm induced by the inverse of the matrix D_k and α_k is the positive steplength chosen in a bounded interval $[\alpha_{min}, \alpha_{max}]$ with $\alpha_{min} > 0$. The $(k+1)$-th iterate $\boldsymbol{x}^{(k+1)}$ is then computed as

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \lambda_k \boldsymbol{d}^{(k)} \tag{5}$$

where $\lambda_k = \theta^m$ with m being the smallest integer such that an Armijo-like successive stepsize reduction rule is satisfied.

This inexact alternating scheme, often referred to as the *Cyclic Block coordinate Gradient Projection* (CBGP) method, leads to stationary limit points, as seen in [5], for every number m of blocks and without any convexity assumption on the objective function.
The next step of my thesis was to adapt the CBGP algorithm to the LINC-NIRVANA imaging problem. LINC-NIRVANA is a Fizeau interferometer which combine the two beams from the primary mirrors of the large binocular telescope, acquiring higher resolution images of the astronomical target (see Figure 1). In order to get the maximum resolution in all directions, it is necessary to acquire p different images of the same object corresponding to p different rotations of the telescopes. Each image can be modeled according to the model proposed in [11], namely as a vector $\boldsymbol{y} = \boldsymbol{y}_{pe} + \boldsymbol{r}$ in $\mathbb{R}^m$, where $\boldsymbol{y}_{pe}$ is the number of photo-electrons due to object and background emission and is a realization of a Poisson random variable with expected value $\overline{\boldsymbol{y}} = \boldsymbol{K} * \boldsymbol{x} + \boldsymbol{b}$, where $\boldsymbol{x}$ is the original object, $\boldsymbol{K}$ is the PSF of the acquisition system and $\boldsymbol{b}$ is the background term (constant and known), while $\boldsymbol{r}$ represents the read-out noise (RON). As proposed in [12], RON can be approximated by a Poisson distribution with parameter σ^2 after addition of the variance σ^2 of the RON to the detected image and the corresponding background. Therefore all the detected images can be viewed as realizations of suitable Poisson random variables, namely

$$\boldsymbol{y}_j \sim Poisson(\boldsymbol{K}_j * \boldsymbol{x} + \boldsymbol{b}_j), \quad j = 1, \dots, p \tag{6}$$

where $\boldsymbol{K}_j$ and $\boldsymbol{b}_j$ are respectively the PSF and the background term corresponding to the jth detected image.
Since it is quite natural to assume that the p images are statistically independent, the likelihood of the problem is the product of the Poisson likelihoods of the different images. By taking its negative logarithm we obtain the following data-fidelity function which is the sum of p Kullback-Leibler (KL) generalized divergences, one for each image, i.e.

$$\begin{aligned} J(\boldsymbol{x}, \boldsymbol{K}_1, \dots, \boldsymbol{K}_p; \boldsymbol{y}, \boldsymbol{b}) &= \sum_{j=1}^{p} D_{KL}(\boldsymbol{y}_j; \boldsymbol{K}_j * \boldsymbol{x} + \boldsymbol{b}_j) \\ &= \sum_{j=1}^{p} \sum_{i=1}^{m} \left\{ \boldsymbol{y}_j(i) \ln \frac{\boldsymbol{y}_j(i)}{(\boldsymbol{K}_j * \boldsymbol{x})(i) + \boldsymbol{b}_j(i)} + (\boldsymbol{K}_j * \boldsymbol{x})(i) + \boldsymbol{b}_j(i) - \boldsymbol{y}_j(i) \right\}, \end{aligned} \tag{7}$$

where $(\boldsymbol{y}, \boldsymbol{b}) = \{(\boldsymbol{y}_j, \boldsymbol{b}_j)\}_{j=1}^{p}$.

Following the maximum likelihood approach, the problem of blind deconvolution consists in the minimization of this function with respect to $\boldsymbol{x}, \boldsymbol{K}_1, \dots, \boldsymbol{K}_p$ for given detected images $\boldsymbol{y}_j$ and background terms $\boldsymbol{b}_j$. The objective function is convex with respect to each block of variables for fixed values of the others, but is not globally convex. Moreover, this problem is highly ill-posed and allows uninteresting solutions. In order to avoid them, some suitable constraints on the unknown variables are introduced. Besides non-negativity of the object, we require that the object flux coincides with the average flux of the p detected images (after background subtraction) which is given by

$$c = \frac{1}{p} \sum_{j=1}^{p} \sum_{i=1}^{m} \{\boldsymbol{y}_j(i) - \boldsymbol{b}_j(i)\}. \tag{8}$$

Regarding the PSFs, as shown in [7] and [10], an important constraint is the upper bound derived from the knowledge of the so-called Strehl ratio (SR) of the instrument, characterizing

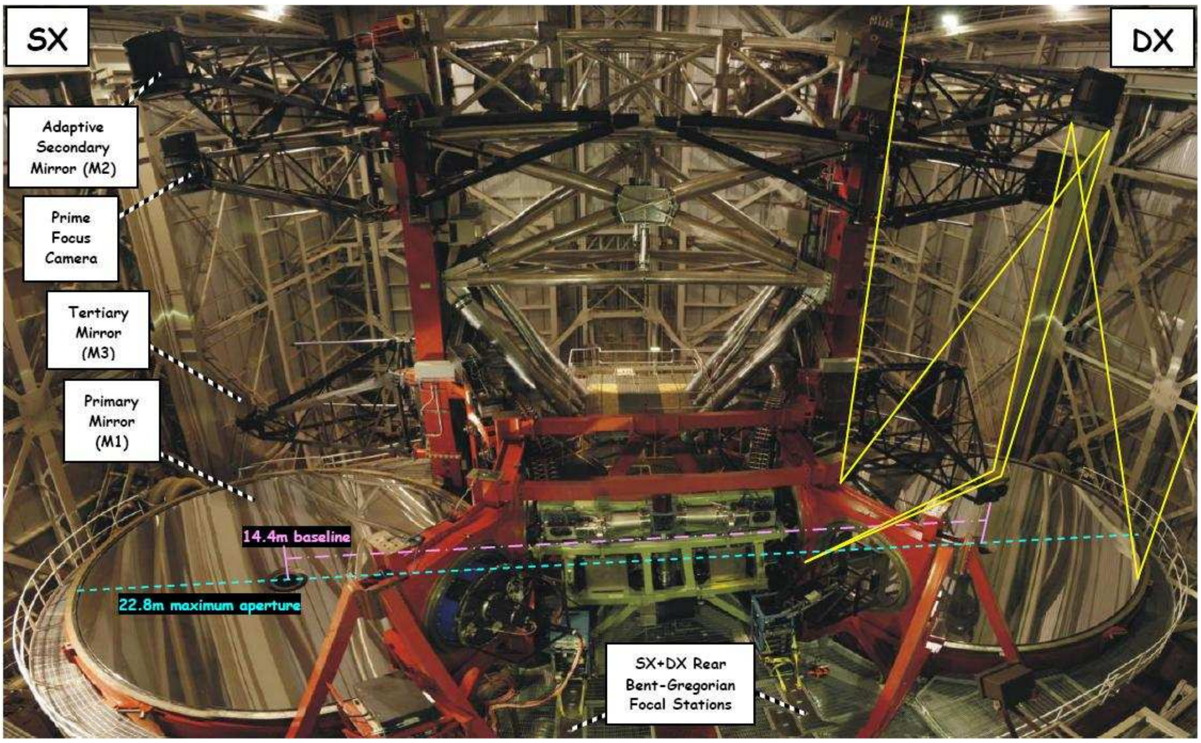

Figure 1: The Large Binocular Telescope at Mount Graham, Arizona. The two primary mirrors are clearly visible, placed at the basis of LBT, together with the secondary and tertiary mirrors, placed at different heights at sides of the structure. The path of the light is shown in yellow.

the correction of the atmospheric blur during the observation. Moreover, non-negativity and normalization to 1 provide additional constraints. In conclusion, the non-convex optimization problem considered can be formulated as follows

$$\begin{aligned} &\min J(\boldsymbol{x}, \boldsymbol{K}_1, \dots, \boldsymbol{K}_p; \boldsymbol{y}, \boldsymbol{b}) \\ &s.t.\ \boldsymbol{x} \geq 0,\ \sum_{\ell=1}^{n} x(\ell) = c;\ 0 \leq \boldsymbol{K}_j \leq s_j,\ \sum_{\ell=1}^{n} \boldsymbol{K}_j(\ell) = 1;\ j = 1, \dots, p, \end{aligned} \tag{9}$$

where s_j is the upper bound on the PSF $\boldsymbol{K}_j$ derived from the knowledge of the SR characterizing the acquisition of $\boldsymbol{y}_j$.
The problem, as formulated in (9), is of the form (1), since the constraints have a separable structure, involving separately the blocks of variables. Therefore, the CBGP method can be applied to problem (9). Concerning the SGP parameters choices for the numerical experiments, the scaling matrix D_k is taken diagonal with entries given by the components of the previous iterate $\boldsymbol{x}^{(k)}$, properly bound by two constants L_1 and L_2, the steplength α_k is selected by alternating the generalized Barzilai-Borwein rules [2] and finally the projection operator is performed with a secant-based method proposed in [6].

3 Numerical experiments and results

The CBGP method has been tested on three astronomical objects:

- a model of an open star cluster based on the image of the Pleiades, consisting of 9 stars with magnitudes ranging from about 13 (i.e. about 2.32×10^8 photon counts per pixel) to 16 (about 1.79×10^7 counts);

- two diffuse objects: the Crab Nebula NGC1952 and the planetary nebula NGC7027, both with magnitude 10.

For each object, three PSFs are generated by means of the software package LOST [1], corresponding to three different rotations of LINC-NIRVANA (respectively 0°, 60°and 120°) and each one with $SR = 0.73$. These PSFs are then convolved with the object, obtaining three images that are perturbed by adding a background term corresponding to about 13.5 $mag\ arcsec^{-2}$ in K-band and by corrupting the results with Poisson and additive Gaussian noise (RON) with variance $\sigma = 10\ e^{-}/px$. According to the approach proposed in [11], compensation for readout noise is obtained in the deconvolution algorithms by adding the constant $\sigma^2 = 100$ to the images and background. Finally the images corresponding to 60°and 120°are de-rotated in order to align them to the object.

Regarding the initial outer iterates, for the object a constant image with the average flux of the background-subtracted images is chosen, while for each PSF we pick the autocorrelation of the corresponding ideal PSF, which is band-limited and satisfies the SR constraint. Concerning the inner iterations, it seems quite natural to choose the same number of iterations for all the PSF's blocks. For the point-wise objects, we use 50 inner iterations on the object and 1 inner iteration for each PSF. This choice seems a sort of golden rule, since it always provides excellent reconstructions of both the object and the PSFs. However, we found no similar rule for the diffuse objects, for which the optimal numbers of inner iterations vary from case to case. Finally, the method is pushed to convergence in the case of point-wise objects, because of the sparsity property of the minimizers of the objective function, whereas it is earlier stopped for the diffuse objects.

From the numerical results (that we reported in Table 1), we observe that the CBGP method provides an excellent reconstruction of both the object and the PSFs profiles in the case of the point-wise object. Moreover, even if such excellent performances are not achieved with the diffuse objects, there are sensible improvements with respect to the single-dish telescope approach. In the particular case of the planetary nebula NGC7027, whereas single image blind deconvolution did not bring any improvement if compared with image deconvolution with given initial PSF, in the case of multiple images the reconstruction error decrease from 14.89% to 11.54% for the object and from 44.48% to 32.45% for the PSF.

Image	$RMSE^{obj}$	$RMSE_1^{obj}$	$RMSE_2^{obj}$	$RMSE_1^{psf}$	$RMSE_2^{psf}$
Pleiades	0.18%	33.19%	1.89%	44.48%	1.98%
NGC1952	12.43%	15.78%	15.17%	44.48%	30.33%
NGC7027	4.8%	14.89%	11.54%	44.48%	32.45%

Table 1: Root mean square errors (RMSE) for the three simulations. The best errors provided by SGP with the original PSFs and the autocorrelation of the ideal PSFs are reported in the first and second columns, respectively. The best error obtained with CBGP is shown in the third column. Finally, in the last two columns we report the mean errors between the original PSFs and the autocorrelation of the ideal PSFs, followed by the mean errors between the original PSFs and those obtained with CBGP.

References

[1] Arcidiacono C., Diolaiti E., Tordi M. et al. 2004, Layer-oriented Simulation Tool, *Appl. Opt.* **43**, 4288-4302.

[2] Barzilai J. e Borwein J.M. 1988, Two-point step size gradient methods, *IMA J. Numer. Anal.* **8**, 141-148.

[3] Bertsekas D.P. 1999, *Nonlinear Programming, 2nd edition.* Athena Scientific, Belmont.

[4] Bonettini S., Zanella R. e Zanni L. 2009, A scaled gradient projection method for constrained image deblurring, *Inverse Probl.* **25**, 015002.

[5] Bonettini S. 2011, Inexact block coordinate descent methods with application to non-negative matrix factorization, *IMA J. Numer. Anal.* **31**, 1431-1452.

[6] Dai Y.H. e Fletcher R. 2006, New algorithms for singly linearly constrained quadratic programs subject to lower and upper bounds, *Math. Program.* **106**, 403-421.

[7] Desiderà G. e Carbillet M. 2009, Strehl-constrained iterative blind deconvolution for post-adaptive-optics data, *Astron. Astrophys.* **507**, 1759-1762.

[8] Grippo L. e Sciandrone M. 2000, On the convergence of the block nonlinear Gauss-Seidel method under convex constraints, *Oper. Res. Lett.* **26**, 127-136.

[9] Luo Z.-Q. and Tseng P. 1992, On the convergence of the coordinate descent method for convex differentiable minimization, *J. Optim. Theory App.* **72**, 7-35.

[10] Prato M., La Camera A., Bonettini S. e Bertero M. 2013, A convergent blind deconvolution method for post-adaptive-optics astronomical imaging, *Inverse Probl.* **29**, 065017.

[11] Snyder D. L., Hammond A.M. e White R.L. 1993, Image Recovery from data acquired with a charge-coupled-device camera, *J. Opt. Soc. Am.* **A-10**, 1014-1023.

[12] Snyder D.L., Helstrom C.W., Lanterman A.D., Faisal M. e White R.L. 1995, Compensation for readout noise in ccd images, *J. Opt. Soc. Am.* **A-12**, 272-283.