A quasi-Newton projection method for nonnegatively constrained image deblurring

G. Landi* E. Loli Piccolomini†

Abstract

In this paper we present a quasi-Newton projection method for image deblurring. The mathematical problem is a constrained minimization problem, where the objective function is a regularization function and the constraint is the positivity of the solution. The regularization function is a sum of the Kullback-Leibler divergence, used to minimize the error in the presence of Poisson noise, and of a Tikhonov term. The Hessian of the regularization function is approximated in order to invert it using Fast Fourier Transforms. The numerical experiments on some astronomical images blurred by simulated Point Spread Functions show that the method gives very good results both in terms of relative error and computational efficiency.

Keywords: Nonnegatively constrained minimization, regularization, image deblurring, projected-Newton method, Poisson noise.

1 Introduction

In astronomical applications, the image formation model is:

$$Af + b = E\{z\} \tag{1}$$

where $f \in \mathbb{R}^n$ is the unknown image to be recovered, $z \in \mathbb{R}^n$ is the detected image that is the realization of a random process with expected value $E(z)$ and b is the expected value (usually constant) of the background. The matrix A contains the information on the Point Spread Function (PSF) determining the blur on the recorded image. The model is the discretization of a first kind Fredholm integral equation, whose kernel is the PSF. When the PSF is space invariant, as is our case, A is a block Toeplitz matrix with Toeplitz blocks (BTTB) [5]. Since A is usually ill-conditioned, the process for recovering the exact image f from (1) needs regularization. Moreover, we suppose that the components of the true solution f are known to be nonnegative, because they represent the intensities of

*Department of Mathematics, University of Bologna, e-mail: landig@dm.unibo.it
†Department of Mathematics, University of Bologna, e-mail: piccolom@dm.unibo.it

the imaged object at the various pixels. Hence, the image reconstruction problem can be formulated as the nonnegatively constrained minimization problem:

$$\begin{aligned} \min \quad & \mathcal{J}(f) = \mathcal{J}_0(f) + \frac{\lambda}{2}\|f\|_2^2 \\ \text{s.t.} \quad & f \geq 0, \end{aligned} \tag{2}$$

where $\mathcal{J}_0(f)$ is the Kullback-Leibler divergence representing consistency with the data in the presence of Poisson noise:

$$\mathcal{J}_0(f) = \sum_{j=1}^{n} \left\{ z_j \ln \frac{z_j}{(Af)_j + b_j} + (Af)_j + b_j - z_j \right\}.$$

The other term is the Tikhonov regularization term, that expresses prior knowledge on the desired solution. The positive parameter λ is the regularization parameter controlling the amount of regularization.

Problem (2) is a convex minimization problem since both the Kullback-Leibler divergence and the Tikhonov function are convex. In [6] a projected Newton-CG method has been presented for solving problem (2), using the Conjugate Gradient (CG) method for computing the search direction.

In this paper we present a quasi-Newton projection method where the search direction is computed by approximating the Hessian of $\mathcal{J}$ with a matrix that is easily invertible using Fast Fourier Transforms (FFTs) with low computational complexity.

The algorithm is tested on astronomical deblurring imaging problems and the results are compared with those obtained by two widely used algorithms: the Gradient Projection algorithm [7, 9] and the Projected Newton method of Bertsekas [3, 2, 9, 8].

The paper is organized as follows. In section 2 we present the proposed method and we outline the algorithm, in section 3 we report the numerical results obtained on some deblurring problems of astronomical images and finally section 4 contains the conclusions.

2 The method

The method proposed in this paper, that we call Quasi-Newton Projection (QNP) method, arises from the projected Newton-CG method presented in [6] for the solution of the nonnegatively constrained problem (2). In that method, the iteration has the general form of the projected Newton like methods proposed by Bertsekas in [3, 2]:

$$f^{k+1} = \left[f^k - \alpha^k p^k \right]^+$$

where p^k is the search direction, α^k is the steplength and $[\cdot]^+$ denotes the projection on the positive orthant.

For each iterate $f^k \geq 0$ we have defined the set of indices:

$$\mathcal{I}(f^k) = \{i \mid 0 \leq f_i^k \leq \varepsilon^k \text{ and } \frac{\partial \mathcal{J}(f^k)}{\partial f_i} > 0\}$$

where

$$\varepsilon^k = \min\{\varepsilon, w^k\}, \quad w^k = |f^k - [f^k - \nabla \mathcal{J}(f^k)]^+|$$

and ε is a small positive parameter. Throughout the paper, for notational simplicity, the gradient and the Hessian of $\mathcal{J}$ at f^k are denoted as $g^k \equiv \nabla \mathcal{J}(f^k)$ and $H^k \equiv \nabla^2 \mathcal{J}(f^k)$, respectively. Moreover, we denote $\mathcal{I}^k \equiv \mathcal{I}(f^k)$.

Given the set $\mathcal{I}^k$, we define the associated reduced gradient $g_{\mathcal{I}}^k$ as

$$\{g_{\mathcal{I}}^k\}_i = \begin{cases} g_i^k, & \text{if } i \notin \mathcal{I}^k; \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

At each iteration k, we consider the solution d_k of the linear system

$$H^k d^k = g_{\mathcal{I}}^k, \tag{4}$$

with

$$H^k = A^T D^k A + \lambda I, \quad D^k = \text{diag}(g^k./(Af^k + b).\hat{}2) \tag{5}$$

where ./ and $.\hat{}$ are the component-wise division and squaring and diag(v) indicates the diagonal matrix with diagonal vector v.

Then, the search direction p^k is given by

$$p_i^k = \begin{cases} d_i^k, & \text{if } i \notin \mathcal{I}^k; \\ g_i^k, & \text{otherwise;} \end{cases} \qquad i = 1, \ldots, n. \tag{6}$$

The step length α^k is computed with the variation of the Armijo rule discussed in [2]. In particular, α^k is the first number of the sequence $\{\beta^m\}_{m \in \mathbb{N}}$, $0 < \beta < 1$, such that

$$\mathcal{J}(f^k) - \mathcal{J}(f^k(\beta^m)) \geq \eta \left(\beta^m \sum_{i \notin \mathcal{I}^k} g_i^k p_i^k + \sum_{i \in \mathcal{I}^k} g_i^k \left(f_i^k - f_i^k(\beta^m) \right) \right) \tag{7}$$

where $f^k(\beta^m) = [f^k - \beta^m p^k]^+$ and $\eta \in (0, \frac{1}{2})$.

The solution of the linear system (4) is the most expensive computational kernel. In [6] the system has been solved inexactly, by using the CG method. In the QNP method presented here we solve the linear system (4) by inverting an approximation of the exact hessian H^k.

In order to reduce the computational cost of the algorithm for the system solution, we approximate the BTTB matrix A with a Block Circulant matrix with Circulant Blocks (BCCB matrix) C. The BCCB matrices arise in spatially invariant image deblurring when periodic boundary conditions are employed [5]. A BCCB matrix has some useful properties. It is normal, that is $C^T C = C C^T$, hence it has a unitary spectral decomposition:

$$C = F^* S F$$

where F is the two dimensional unitary Discrete Fourier Transform (DFT) matrix and S is a diagonal matrix. For this reason, the matrix-vector product involving C^{-1} can be performed without constructing C explicitly and using FFTs. See [5, 10, 9] for more details.

The diagonal matrix D^k of (5) is approximated with a diagonal matrix $\tilde{D}^k$ with all entries equal to the mean μ^k of the diagonal elements of D^k.

Finally the exact hessian H^k is substituted by:

$$\tilde{H}^k = C^T \tilde{D}^k C + \lambda I = \mu_k C^T C + \lambda I. \tag{8}$$

If

$$S = \text{FFT2}(C)$$

(where FFT2 is the Fast Fourier Transform of a matrix), then the system:

$$H^k d^k = g^k_{\mathcal{I}}$$

is approximated by:

$$\tilde{H}^k d^k = g^k_{\mathcal{I}}$$

which is solved in the Fourier space as:

$$d^k = \mu^k \text{IFFT2}\left(\text{FFT2}(g^k_{\mathcal{I}})./\left(S.\hat{}2 + \frac{\lambda}{\mu^k}\right)\right). \tag{9}$$

The QNP algorithm is formally outlined as follows.

Algorithm 2.1 (Quasi-Newton Projection algorithm)**.**
Choose $f^0 \geq 0$ and $\eta \in (0, \frac{1}{2})$.
Evaluate $\mathcal{I}^0$, g^0 and H^0.
Set $k = 0$.
Repeat until convergence

2. Computation of the search direction p^k
2.1 Compute the reduced gradient $g^k_{\mathcal{I}}$ by (3);
2.2 Compute d^k by (9);
2.3 Compute p^k by (6);

3. Computation of the steplength α^k
Find the smallest number $m \in \mathbb{N}$ satisfying (7);

4. Updates
Set $f^{k+1} = [f^k - \alpha^k p^k]^+$ and evaluate $\mathcal{I}^{k+1}$, g^{k+1}, H^{k+1};
Set $k = k + 1$;

end

Finally, let $g^k_{\mathcal{P}}$ be the projected gradient of $\mathcal{J}$ at f^k defined as

$$\{g^k_{\mathcal{P}}\}_i = \begin{cases} g^k_i, & \text{if } f^k_i > 0; \\ g^k_i, & \text{if } f^k_i = 0 \text{ and } g^k_i < 0; \\ 0, & \text{otherwise.} \end{cases}$$

The iterations of algorithm 2.1 are terminated according to the following stopping criteria:

i) when the value $\|g^k_{\mathcal{P}}\|_2$ is sufficiently close to zero; i.e. when:

$$\|g^k_{\mathcal{P}}\|_2 \leq \tau_1 \|g^0_{\mathcal{P}}\|_2 \tag{10}$$

where τ_1 is a given positive tolerance;

ii) when the relative distance between two successive iterates has become less than a given positive tolerance τ_2; i.e.

$$\frac{\|f^{k+1} - f^k\|_2}{\|f^{k+1}\|_2} \leq \tau_2; \tag{11}$$

iii) when a maximum number $K_{\max}$ of iterations has been performed.

The convergence of algorithm 2.1 can be proved as follows.

Definition 2.1. An $n \times n$ matrix P is *diagonal with respect to the subset of indices* $\mathcal{L} \subset \{1, \ldots, n\}$ if

$$P_{ij} = 0, \quad \forall i \in \mathcal{L},\ j = 1, 2, \ldots, N,\ i \neq j.$$

Proposition 2.1. *The search direction* p^k *in* (6) *is obtained as:*

$$p^k = P^k g^k$$

where P^k *is diagonal with respect to* $\mathcal{I}^k$ *and is defined as:*

$$P^k_{ij} = \begin{cases} \delta_{ij}, & \text{if either } i \in \mathcal{I}^k \text{ or } j \in \mathcal{I}^k; \\ \{(\tilde{H}^k)^{-1}\}_{ij}, & \text{otherwise;} \end{cases}$$

where $\tilde{H}^k$ *is the approximation* (8) *to the exact Hessian at iteration* k.

Proof. For a vector $v \in \mathbb{R}^n$ and the corresponding set $\mathcal{I}(v)$, let reduce be the operator such that

$$\{\mathsf{reduce}(v)\}_i = \begin{cases} v_i, & \text{if } i \notin \mathcal{I}(v); \\ 0, & \text{otherwise;} \end{cases} \qquad i = 1, \ldots, n.$$

Then, from (6) we have $p^k_i = g^k_i$ for $i \in \mathcal{I}^k$, while, for $i \notin \mathcal{I}^k$, we have

$$p^k_i = \sum_{j=1}^{n} \{(\tilde{H}^k)^{-1}\}_{ij} \{\mathsf{reduce}(g^k)\}_j = \sum_{j \notin \mathcal{I}^k} \{P^k\}_{ij} g^k_j.$$

Hence, $p^k = P^k g^k$. □

We can now prove the following result concerning some properties of the matrix P^k necessary to show the convergence of the QPN method.

Proposition 2.2. *If D^k is uniformly bounded for all k, then the matrices $\tilde{H}^k$ in* (8) *are positive definite and satisfy*

$$c_1\|y\|^2 \leq y^T\tilde{H}^k y \leq c_2\|y\|^2, \quad \forall y \in \mathbb{R}^n,\ k = 0, 1, 2, \ldots$$

for some positive scalars c_1 and c_2.

Proof. Let $0 < \nu_n \leq \ldots \leq \nu_1$ be the eigenvalues of C^TC. Then the eigenvalues of $\tilde{H}^k$ are:

$$\xi_j^k = \mu^k\nu_j + \lambda, j = 1, \ldots n, \quad \xi_n^k \leq \ldots \leq \xi_1^k$$

.

Since D^k is uniformly bounded, there exists $M > 0$ such that $0 < \mu^k \leq M$. Hence:

$$\lambda \leq \xi_j^k \leq M\nu_1 + \lambda, \quad j = 1, \ldots, n.$$

Then:

$$\lambda\|y\|_2^2 \leq \xi_n^k\|y\|_2^2 \leq y^T\tilde{H}^k y \leq \xi_1^k\|y\|_2^2 \leq (M\nu_1 + \lambda)\|y\|_2^2.$$

The thesis immediately follows by setting:

$$c_1 = \lambda \ \text{ and } \ c_2 = M\nu_1 + \lambda$$

□

From the proposition 2.2 it follows that $\tilde{c}_1$ and $\tilde{c}_2$ exist such that:

$$\tilde{c}_1\|y\|_2^2 \leq y^TP^k y \leq \tilde{c}_2\|y\|_2^2, \quad \forall y \in \mathbb{R}^n,\ k = 0, 1, 2, \ldots \tag{12}$$

Since the matrices P^k satisfy (12), then from the analysis in [1, Proposition 2, Proposition 4] it follows that the QNP method is convergent.

3 Numerical results

In this section the results of some numerical experiments are presented. They have been executed on a Pentium IV PC using Matlab 7.0 (Release 14).

The stopping criteria parameters of algorithm 2.1 have been fixed to the following values:

i) Projected gradient tolerance: $\tau_1 = 10^{-3}$;

ii) Successive iterates tolerance: $\tau_2 = 10^{-2}$;

iii) Maximum number of iterations: $K_{\max} = 30$.

The initial iterate f^0 has been chosen to be the constant image whose pixel values $f^0_{(i,j)}$ are equal to

$$f^0_{(i,j)} = \frac{\sum_{i=1}^{Nx}\sum_{j=1}^{Ny}(z_{i,j} - b)}{NxNy}$$

where $Nx \times Ny$ is the image size.

The regularization parameter λ has been heuristically chosen and the parameter β of (7) has been fixed equal to 0.25 in all the experiments.

In all the numerical experiments, an original 256×256 image has been convolved with an Adaptive Optic PSF (AOPSF) from the CAOS (Code for Adaptive Optics System) package [4], then a constant sky background term has been added and, finally, the resulting image has been corrupted with Poisson noise with different Signal to Noise Ratio (SNR) values.

In this work, we illustrate the results obtained for three test problems, called TP1, TP2 and TP3, respectively. The corresponding test images are the HST image of the galaxy NGC 1288 (figure 1(a)), the VLT image of the Crab Nebula (figure 2(a)) and a star cluster image (figure 3(a)). Some perturbed images are shown in figures 1(b), 2(b) and 3(b), for different SNR values. (In particular, SNR=38.20, SNR=44.23 and SNR=38.20, respectively).

Table 1 shows the results obtained with the previous images with varying SNR values. The table reports the relative Root Mean Square Error (RMSE) of the reconstructions defined as

$$\text{relative RMSE} = \frac{\|f_{\text{ex}} - f_{\text{rec}}\|_F}{\|f_{\text{ex}}\|_F}$$

where f_{ex} is the exact image, f_{rec} is the reconstructed one and $\|\cdot\|_F$ is the Frobenius norm of a matrix. Moreover, the table illustrates the total number of FFTs (fourth column) and the number of performed iterations (last column).

The numerical results in the table show that the QNP method is effective and efficient since it restores good quality images at a very low computational cost. Some restored images are depicted in figures 1(c), 2(c) and 3(c).

For some test problems, figures 1(d), 2(d) and 3(d) show the behavior of the relative reconstruction error as a function of the number of iterations; the circle indicates the error in the stopping iteration. These figures illustrate that the relative reconstruction error decreases very quickly during the first iterations and then it becomes stable. This behavior of the relative reconstruction error has been observed in all the performed experiments and indeed it represents an attractive characteristic of the QNP method.

The performance of the QNP method has been compared with that of the Gradient Projection (GP) method and the Projected Newton (PN) method. This last method requires, at each iteration, the solution of a linear system in order to compute the search direction. In our experiments, this linear system has been solved with the CG method stopped with a relative precision of 0.1 and a maximum number of 20 iterations allowed.

Figure 4 illustrates the behavior of the relative reconstruction error (left column) and the projected gradient norm (rigth column) as a function of the number of performed FFTs for the QNP method (dashed line), the GP method (solid line) and the PN method (dash-dotted line). From the graphs in the figure it is evident that the QNP method has the best performance in terms of relative reconstruction error reduction. On the other hand, all the methods have a slow decreasing of the projected gradient norm. We can conclude that the stopping

criteria ii), considering the relative distance between two successive iterates, is the most efficient one, in this application, to obtain a good reconstruction.

4 Conclusions

In this work, a quasi-Newton projection method has been presented for the nonnegatively constrained restoration of images degraded by blur and Poisson noise. In this method, determining the search direction only requires the computation of a pair of two-dimensional FFTs. The results of several numerical experiments show that the relative reconstruction error reaches the minimum value during the first iterations. This minimum value is smaller than that obtained by both the GP method and the PN method. Therefore, the QNP method appears to be effective and efficient since it provides good quality restored images at a low computational cost.

5 Acknowledgments

The authors are grateful to M. Bertero and P. Boccacci of the Department of Informatics of the University of Genova for kindly providing the data of the numerical experiments.

References

[1] D. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods.* Academic Press, 1982.

[2] D. Bertsekas. Projected Newton methods for optimization problem with simple constraints. *SIAM J. Control Optim.*, 20(2):221–245, 1982.

[3] D. Bertsekas. *Nonlinear Programming.* Athena Scientific, (2nd Edition), 1999.

[4] M. Carbillet et al. Determining surface temperatures from interior observations. *MNRAS*, 356:1263, 2005.

[5] P. C. Hansen, J.G. Nagy, and D.P. O'Leary. *Deblurring images: matrices, spectra and filtering.* SIAM, 2006.

[6] G. Landi and E. Loli Piccolomini. A projected Newton-CG method for nonnegative astronomical image deblurring. *Numer. Alg.*, 48:279–300, 2008.

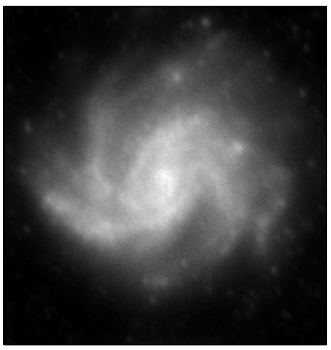
[7] J. Nocedal and S. J. Wright. *Numerical optimization.* Springer-Verlag New York, Inc., 1999.

[8] C. R. Vogel. Solution of linear systems arising in nonlinear image deblurring. In *in Workshop on Scientific Computing.* Springer, 1997.

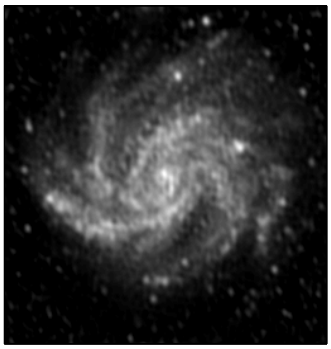
Test Pb.	SNR	rel. RMSE	FFTs	It.
TP1	44.23	0.235159	34	3
TP1	38.20	0.261398	26	3
TP1	34.01	0.287359	22	3
TP2	48.23	0.154689	6	1
TP2	44.23	0.162648	18	3
TP2	38.20	0.184557	18	3
TP3	44.23	0.404057	120	8
TP3	38.20	0.423696	66	5
TP3	34.01	0.558986	60	5

Table 1: Numerical results for the test problems.

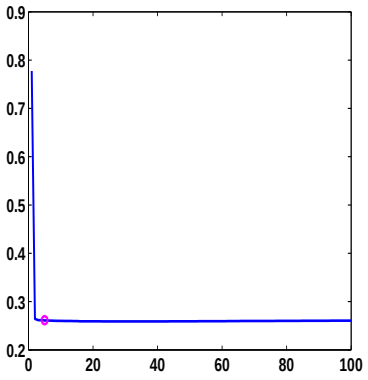
(a) Exact image.



(b) Blurred and noisy image (SNR=38.20).



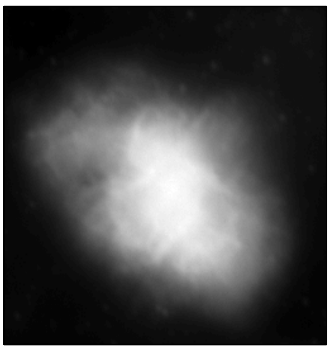
(c) Restored image.



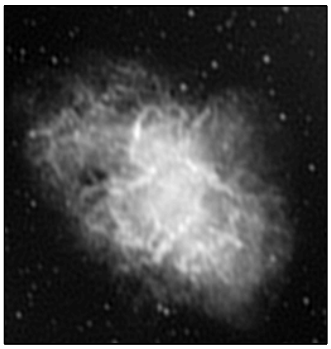
(d) Relative reconstruction error versus the iteration number. The circle indicates the stopping iteration.

Figure 1: TP1 test problem: the HST image of the galaxy NGC 1288.

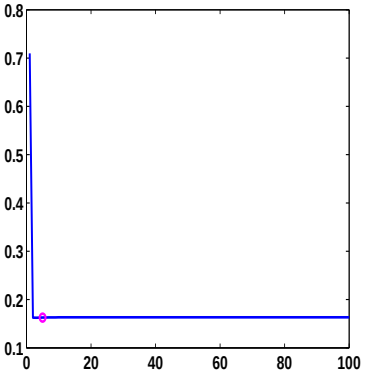
(a) Exact image.



(b) Blurred and noisy image (SNR=44.23).

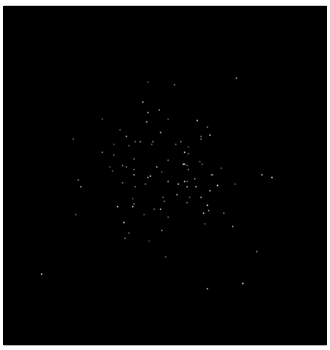


(c) Restored image.

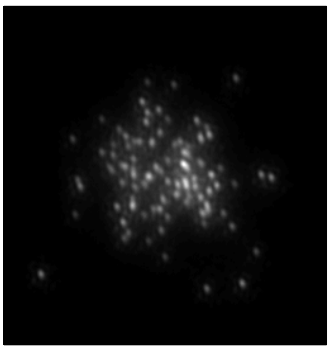


(d) Relative reconstruction error versus the iteration number. The circle indicates the stopping iteration.

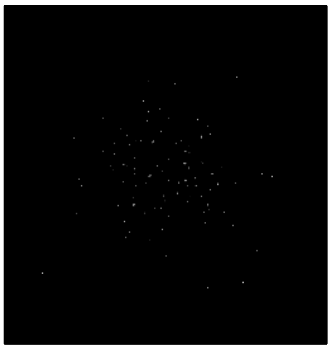
Figure 2: TP2 test problems: the VLT image of the Crab Nebula.



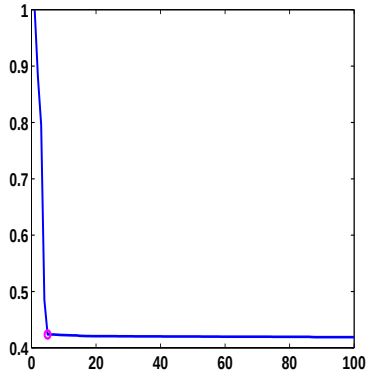
(a) Exact image.



(b) Blurred and noisy image (SNR=38.20).

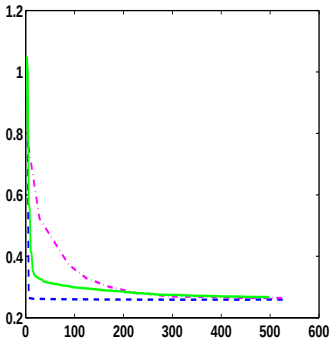


(c) Restored image.

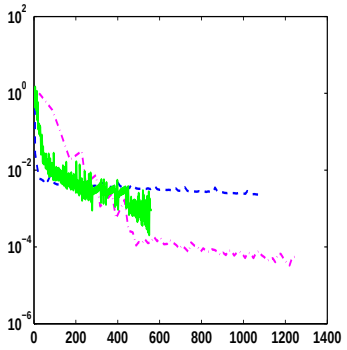


(d) Relative reconstruction error versus the iteration number. The circle indicates the stopping iteration.

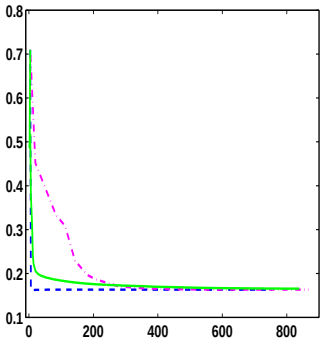
Figure 3: TP3 test problem: a star cluster image.



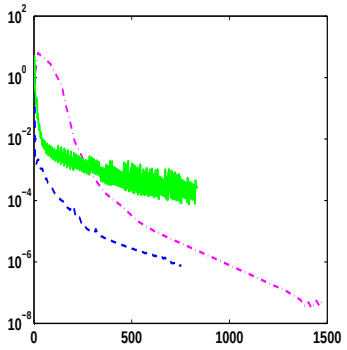
(a) TP1 test problem (SNR=38.20).



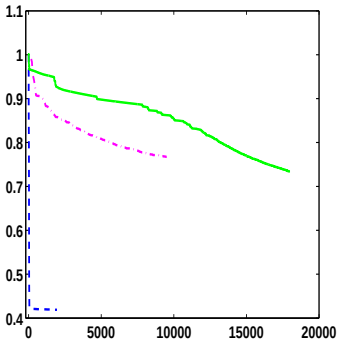
(b) TP1 test problem (SNR=38.20).



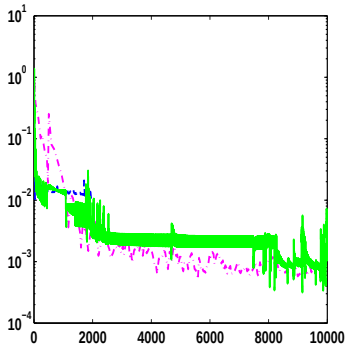
(c) TP2 test problems (SNR=44.23).



(d) TP2 test problems (SNR=44.23).



(e) TP3 test problem (SNR=38.20).



(f) TP3 test problem (SNR=38.20).

Figure 4: Comparisons between the QNP method (dashed line), the GP method (solid line) and the PN method (dashdotted line). The horizontal axis shows the number of FFTs. The vertical axis shows the relative reconstruction error (left column) and the projected gradient norm on a logarithmic scale (right column).

[9] C. R. Vogel. *Computational Methods for Inverse Problems*. SIAM, Philadelphia, PA, USA, 2002.

[10] C. R. Vogel and M. E. Oman. Fast, robust total variation–based reconstruction of noisy, blurred images. *IEEE Trans. on Image Proc.*, 7:813–824, 1998.