A discrepancy principle for Poisson data: uniqueness of the solution for 2D and 3D data

S Bonettini and V Ruggiero

Dipartimento di Matematica, Università di Ferrara, Polo Scientifico Tecnologico, Blocco B, Via Saragat 1, I-44100 Ferrara, Italy

E-mail: silvia.bonettini@unife.it, valeria.ruggiero@unife.it

Abstract. This paper is concerned with the uniqueness of the solution of a nonlinear equation, named *discrepancy equation*. For the restoration problem of data corrupted by Poisson noise, we have to minimize an objective function that combines a data-fidelity function, given by the generalized Kullback–Leibler divergence, and a regularization penalty function. Bertero et al. recently proposed to use the solution of the *discrepancy equation* as a convenient value for the regularization parameter. Furthermore they devised suitable conditions to assure the uniqueness of this solution for several regularization functions in 1D denoising and deblurring problems.

The aim of this paper is to generalize this uniqueness result to 2D and 3D problems for several penalty functions, such as an edge preserving functional, a simple case of the class of Markov Random Field (MRF) regularization functionals and the classical Tikhonov regularization.

1. Introduction

The maximum likelihood approach to image reconstruction from data corrupted by Poisson noise [9] leads to the minimization of a data–fidelity function, given by the generalized Kullback–Leibler (KL) divergence [1], both for denoising and deblurring problems. In the framework of the Bayesian paradigm [6], the *a–priori* information on the solution can be used to regularize this ill-conditioned problem by a suitable penalty function. The final problem can be formulated as the following constrained minimization problem

$$\begin{aligned} &\min \quad f_\beta(x) = f_0(x) + \beta f_1(x) \\ &\text{subject to } x \in \Omega \end{aligned} \tag{1}$$

where $f_0(x)$ is the KL data–fidelity function detailed in Section 2, $f_1(x)$ is the penalty function and the regularization parameter β is a positive scalar. Here $\Omega \subset \mathbb{R}^N$ is a closed and convex nonempty subset of the nonnegative orthant which depends on the domain of $f_\beta(x)$ and on the features of the application that we have to solve.

In [2] the authors propose a criterion to select a convenient value for β in both denoising and deblurring problems, that consists in determining the solution of the following nonlinear equation, named discrepancy equation:

$$D_y(\beta) = \frac{2}{M} f_0(x_\beta{}^*) = 1 \tag{2}$$

where $y \in \mathbb{R}^M$ is the vector of the input data of the restoration problem and $x_\beta{}^*$ is the solution of the problem (1). The uniqueness of the solution $x_\beta{}^*$ of (1) is crucial for the well–posedness of this *discrepancy principle*.

As discussed in Section 2, the generalized KL divergence $f_0(x)$ is a coercive function on the domain Ω (see [2]). Furthermore, when the data are positive ($y_i > 0$, $i = 1, ..., M$) and the null space of the blurring operator is trivial, $f_0(x)$ is strictly convex on Ω and, consequently, equation (2) has a unique solution. In this note we focus our interest on the general case, i.e. some of the y_i can be zero and, for deblurring problem, the blur matrix may have a nontrivial null space. In [2] suitable conditions are devised to assure that the proposed criterion provides a unique value of the regularization parameter for several regularization functions in 1D denoising and deblurring problems.

The aim of this note is to generalize this uniqueness result to 2D and 3D problems for several penalty functions. In particular, we consider:

- a regularization function from the class of edge preserving potentials [4] whose analytic form for 2D images is

$$f_1(x) = \sum_{i,j} \sqrt{(x_{i+1j} - x_{ij})^2 + (x_{ij+1} - x_{ij})^2 + \delta^2} \qquad \delta \neq 0$$

  which, for small values of δ, provides an approximation of the Total Variation (TV) functional [10];

- a simple case of the class of Markov Random Field (MRF) regularization functionals [7], given by

$$f_1(x) = \sum_{i,j} \sum_{k,l \in \mathcal{N}_{ij}} \sqrt{\left(\frac{x_{ij} - x_{kl}}{w_{kl}}\right)^2 + \delta^2} \qquad \delta \neq 0$$

  where $\mathcal{N}_{ij}$ denotes the set of the indices of the first 8 neighbors of the pixel ij and w_{kl} are positive weights;

- the classical Tikhonov regularization, which, for 2D images, is given by

$$f_1(x) = \frac{1}{2} \sum_{i,j} (x_{i+1j} - x_{ij})^2 + (x_{ij+1} - x_{ij})^2$$

For these penalty functions – nonnegative, differentiable and convex – we provide an explicit expression for the gradient vector and the Hessian matrix and we prove the following properties:

(P1) the null space of the Hessian of the penalty function $f_1(x)$ is given by $\{\alpha e_N : \alpha \in \mathbb{R}\}$, where e_N is the constant vector in R^N with all entries equal to 1;

(P2) the sum of the entries of the gradient of $f_1(x)$ is zero.

Property (P1) is crucial to prove the strict convexity of the function $f_\beta(x)$ over the domain Ω. Indeed, the generalized KL divergence $f_0(x)$ is a convex function; then, for any convex function $f_1(x)$, $f_\beta(x)$ is strictly convex if the intersection between the null spaces of the Hessians of $f_0(x)$ and $f_1(x)$ is trivial. If (P1) holds, this is equivalent to saying that $f_\beta(x)$ is strictly convex if the constant vector e_N does not belong to the null space of $f_0(x)$. As a consequence, the solution of the problem (1) is unique for fixed β and the hypotheses of the Lemma 2 in [2] hold.
Moreover, property (P2) is exploited in the proofs of the lemmas used in [2] to prove the theorems on the conditions that assure the existence of the solution of the discrepancy equation.
This paper concerns only discrete problems. Indeed, the main contribution of this note is to prove that the discrepancy principle can be applied to the discrete models for 2D and 3D image reconstruction problems.
The paper is organized as follows. In section 2 we introduce some notations and we recall some basic results on the generalized KL divergence and on the conditions for the strict convexity of the function $f_\beta(x)$. Furthermore, we state some preliminary results on the null space of the Hessian of special functions. In section 3 we show that (P1) and (P2) hold for an edge preserving regularization function in the 1D, 2D and 3D cases; analogous results are proved in section 4 and 5 for the MRF and Tikhonov regularization respectively. In section 6 we conclude that such results allow us to apply the analysis performed in [2] to the 2D and 3D cases, generalizing the conditions for the existence and uniqueness of the solution of the discrepancy equation (2).

2. Notation and preliminary results

We denote the ℓ_2 norm of a vector $x \in \mathbb{R}^N$ as $\|x\|$ and the null space of a matrix S by $\mathcal{N}[S]$. In the restoration problems, we indicate the entries of the unknown object $x \in \mathbb{R}^N$ as x_k, $k = 1, ..., N$ where, in the 2D and 3D cases, k corresponds to a multi-index. In particular for a 2D image $x \in \mathbb{R}^{n\times m}$ we have $N = nm$ and, assuming a column–wise ordering, the following correspondence holds

$$x_k \equiv x_{ij} \tag{3}$$

with $k = (j-1)n + i$; similarly, for a 3D image $x \in \mathbb{R}^{n\times m\times r}$ $(N = nmr)$ we have the equivalence

$$x_k \equiv x_{ijh} \tag{4}$$

with $k = (h-1)nm + (j-1)n + i$.
In the following we employ both notations, with a single index or a multi-index, taking into account the previous equivalences.
We consider two kind of boundary conditions:

P) periodic boundary conditions; for example, in the 2D case, we have: $x_{n+1j} = x_{1j}$, $x_{0j} = x_{nj}$, $j = 1, ..., m$ and $x_{im+1} = x_{i1}$, $x_{i0} = x_{im}$, $i = 1, ..., n$; we define also $x_{00} = x_{nm}$, $x_{n+1m+1} = x_{11}$, $x_{n+10} = x_{1m}$ and $x_{0m+1} = x_{n1}$;

R) reflexive (or Neumann) boundary conditions; this means, in the 2D case, that we assume $x_{n+1j} = x_{nj}$, $x_{0j} = x_{1j}$, $x_{in+1} = x_{in}$ and $x_{i0} = x_{i1}$, for $i = 1, ..., n$, $j = 1, ..., m$; we define also $x_{00} = x_{11}$, $x_{n+1m+1} = x_{nm}$, $x_{n+10} = x_{n1}$ and $x_{0m+1} = x_{1m}$.

We recall the definition of the generalized Kullback–Leibler divergence. Let $y \in \mathbb{R}^M$ be a vector of data. We assume $y \neq 0$, with nonnegative entries. Given the set of indices $\mathcal{I} = \{1, ..., M\}$, we define the following two disjoint subsets $\mathcal{I}_1$ and $\mathcal{I}_2$

$$\mathcal{I}_1 = \{k \in \mathcal{I}, y_k > 0\} \tag{5}$$

$$\mathcal{I}_2 = \{k \in \mathcal{I}, y_k = 0\} \tag{6}$$

Then the generalized KL divergence is defined as

$$f_0(x) = \sum_{k \in \mathcal{I}_1} \left\{ y_k \ln \frac{y_k}{(Hx+b)_k} + (Hx+b)_k - y_k \right\} + \sum_{k \in \mathcal{I}_2} (Hx+b)_k \tag{7}$$

where $b = \gamma e_M$ represents a positive constant background term and $H \in \mathbb{R}^{M \times N}$ is the blurring operator, with nonnegative entries which satisfies

$$\sum_{k=1}^{N} H_{lk} > 0 \qquad l = 1, ..., M \tag{8}$$

and normalized so that $H^T e_M = e_N$. A special case is image deconvolution, when we can take $M = N$ and the blurring operator is frequently approximated by a cyclic convolution of the object with a point spread function. In this case we also have $He_N = e_N$. When we consider a deblurring problem, the set Ω in (1) is defined as the nonnegative orthant of $\mathbb{R}^N$:

$$\Omega = \{x \in \mathbb{R}^N, x \geq 0\} \tag{9}$$

In the denoising case, the relationship between the data vector y and the object vector x is more simple. Indeed in (7) we have $N = M$, $H = I$, $b = 0$. Furthermore the domain Ω is given by

$$\Omega = \{x \in \mathbb{R}^N, x_k \geq \eta \quad k \in \mathcal{I}_1, x_k \geq 0 \quad k \in \mathcal{I}_2\} \tag{10}$$

where η is a positive small constant that must satisfy $\eta < \min\left\{\frac{1}{M}\sum_k y_k, \min_{k \in \mathcal{I}_1} y_k\right\}$. In both cases, it is well known that $f_0(x)$ is a nonnegative, differentiable, coercive and convex function [2].

The gradient and the Hessian of (7) are given by

$$\begin{aligned} \nabla f_0(x) &= e_N - H^T Z^{-1} y \\ \nabla^2 f_0(x) &= H^T Y Z^{-2} H \end{aligned}$$

where Z and Y are diagonal matrix of order M, such that $(Z)_{kk} = (Hx+b)_k$, $k \in \mathcal{I}_1$, and $(Z)_{kk} = 1$, $k \in \mathcal{I}_2$ and $(Y)_{kk} = y_k$, $k = 1, ..., M$. Thus, the null space of $\nabla^2 f_0(x)$ is

$$\mathcal{N}[\nabla^2 f_0(x)] = \{u \in \mathbb{R}^N : (Hu)_k = 0 \text{ if } k \in \mathcal{I}_1\}$$

and it contains the null space of H.

Remark. When y has only positive entries and, in the case of the deblurring problem, the matrix H has full column rank, $\mathcal{N}[\nabla^2 f_0(x)] = \{0\}$ and, consequently, $f_0(x)$ and $f_\beta(x)$ are strictly convex functions. In this note, we focus our interest in the nontrivial case.

To obtain the null space of $\nabla^2 f_\beta(x)$, we can invoke the following results.

Lemma 2.1 *Let A and B be two symmetric positive semidefinite matrices of order N. Then*

$$\mathcal{N}[A+B] = \mathcal{N}[A] \cap \mathcal{N}[B]$$

The proof of the previous Lemma can be easily obtained recalling that if a matrix S is positive semidefinite, we have $Sv = 0$ if and only if $v^T Sv = 0$ (see [8, p.400]).

We remark that for a denoising problem, since $H = I$, the null space of $\nabla^2 f_0(x)$ cannot contain the constant vector e_N; for a deblurring problem, the assumption (8) implies that $e_N \notin \mathcal{N}[\nabla^2 f_0(x)]$. As a consequence, if property (P1) holds for the penalty function $f_1(x)$, from the previous Lemma we have

$$\mathcal{N}[\nabla^2 f_\beta(x)] = \mathcal{N}[\nabla^2 f_0(x)] \cap \mathcal{N}[\nabla^2 f_1(x)] = \{0\}$$

We can summarize the previous remarks as follows.

Lemma 2.2 *When $f_1(x)$ is convex over Ω and (P1) holds, the Hessian of $f_\beta(x)$ in (1) is a positive definite matrix for all $x \in \Omega$ and, thus, $f_\beta(x)$ is strictly convex over Ω.*

In the following sections we will show that (P1) holds for several kinds of convex penalty functions in the 1D, 2D and 3D cases and, in addition, we will prove also property (P2).

3. Hypersurface (HS) regularization and Total variation (TV)

This section is focused on the edge preserving regularization. In particular, we consider a penalty function having the following form

$$f_1(x) = \sum_{k=1}^{N_1} \sqrt{D_k} \tag{11}$$

where in the 1-D case

$$D_k = (x_{k+1} - x_k)^2 + \delta^2, \tag{12}$$

while in 2D

$$D_k \equiv D_{ij} = (x_{i+1j} - x_{ij})^2 + (x_{ij+1} - x_{ij})^2 + \delta^2 \tag{13}$$

and in 3D

$$D_k \equiv D_{ijh} = (x_{i+1jh} - x_{ijh})^2 + (x_{ij+1h} - x_{ijh})^2 + (x_{ijh+1} - x_{ijh})^2 + \delta^2 \quad (14)$$

Here δ is a nonzero scalar parameter while $N_1 = N$ for periodic boundary conditions and $N_1 = n-1$, $N_1 = nm-1$ or $N_1 = (nm-1)r$ for reflexive boundary conditions in the 1D, 2D and 3D case respectively.
The model function in (11) is related to the class of edge preserving potentials considered in [12] and, for small values of δ, it can be considered as a discrete approximation of the TV functional [10]. Moreover, the expression (11) formally describes also the hypersurface potential proposed in [4].
Following [3], we observe that D_k can be written also as

$$D_k = \|A_k x\|^2 + \delta^2 \quad (15)$$

where A_k, $k = 1, ..., N_1$, is a matrix with N columns and $1, 2$ or 3 rows, depending on the dimension of the object x and on the assumed boundary conditions.
In particular, under periodic boundary conditions, A_k is a $d \times N$ matrix, where $d = 1, 2, 3$ is the dimension of the object x; for $d = 1$ ($N = n$), we define A_k as the $1 \times n$ vector with all zero components except the following two entries:

$$(A_k)_k = -1 \qquad (A_k)_{mod(k,n)+1} = 1 \qquad k = 1, ..., n \quad (16)$$

Similarly, for the 2D case ($N = nm$), we define $A_k \equiv A_{ij}$ for $i = 1, ..., n$, $j = 1, ..., m$ as the $2 \times N$ matrix with all zero entries except for the following four elements:

$$\begin{array}{lllll} (A_k)_{1\ (j-1)n+i} & = & -1 & (A_k)_{1\ (j-1)n+mod(i,n)+1} & = & 1 \\ (A_k)_{2\ (j-1)n+i} & = & -1 & (A_k)_{2\ mod(j,m)n+i} & = & 1 \end{array} \quad (17)$$

Finally, in the 3D case ($N = nmr$) for $i = 1, ..., n$, $j = 1, ..., m$ and $h = 1, ..., r$, we define $A_k \in \mathbb{R}^{3\times N}$ as

$$\begin{array}{llllll} (A_k)_{1\ (h-1)nm+(j-1)n+i} & = & -1 & (A_k)_{1\ (h-1)nm+(j-1)n+mod(i,n)+1} & = & 1 \\ (A_k)_{2\ (h-1)nm+(j-1)n+i} & = & -1 & (A_k)_{2\ (h-1)nm+mod(j,m)n+i} & = & 1 \\ (A_k)_{3\ (h-1)nm+(j-1)n+i} & = & -1 & (A_k)_{3\ mod(h,r)nm+(j-1)m+i} & = & 1 \end{array} \quad (18)$$

and zero otherwise.
When reflexive boundary conditions are assumed, we define the difference matrices as follows. For $d = 1$, we define the $n-1$ vectors A_k as in (16), for $k = 1, ..., n-1$. In the 2D case, for $i = 1, ..., n-1$, $j = 1, ..., m-1$ we define A_k as in (17), while, for the remaining cases, we have that $A_k \equiv A_{nj}$ and $A_k \equiv A_{im}$ are $1 \times N$ vectors with the following nonzero elements:

$$\begin{array}{llllll} (A_k)_{jn} & = & -1 & (A_k)_{jn+n} & = & 1 \quad i = n, \quad j = 1, ..., m-1 \\ (A_k)_{(m-1)n+i} & = & -1 & (A_k)_{(m-1)n+i+1} & = & 1 \quad i = 1, ..., n-1, \quad j = m \end{array}$$

In the 3D case ($N = nmr$) we define $A_k \in \mathbb{R}^{3\times N}$ as in (18) for $i = 1, ..., n-1$, $j = 1, ..., m-1$ and $h = 1, ..., r-1$. Furthermore, we define the nonzero entries of $A_k \equiv A_{imh} \in \mathbb{R}^{2\times N}$ as

$$\begin{array}{llllll} (A_k)_{1\ (h-1)nm+(m-1)n+i} & = & -1 & (A_k)_{1\ (h-1)nm+(m-1)n+i+1} & = & 1 \\ (A_k)_{2\ (h-1)nm+(m-1)n+i} & = & -1 & (A_k)_{2\ hnm+(m-1)n+i} & = & 1 \end{array}$$

for $i = 1, ..., n-1$ and $h = 1, ..., r-1$ and those of $A_k \equiv A_{njh} \in \mathbb{R}^{2\times N}$ as

$$\begin{array}{lll} (A_k)_{1\ (h-1)nm+(j-1)n+n} = -1 & & (A_k)_{1\ (h-1)nm+jn+n} = 1 \\ (A_k)_{2\ (h-1)nm+(j-1)n+n} = -1 & & (A_k)_{2\ hnm+(j-1)n+n} = 1 \end{array}$$

for $j = 1, ..., m-1$ and $h = 1, ..., r-1$, while $A_k \equiv A_{ijr}$ for $i = 1, ..., n-1$, $j = 1, ..., m-1$ have the following nonzero entries:

$$\begin{array}{lll} (A_k)_{1\ (r-1)nm+(j-1)n+i} = -1 & & (A_k)_{1\ (r-1)nm+(j-1)n+i+1} = 1 \\ (A_k)_{2\ (r-1)nm+(j-1)n+i} = -1 & & (A_k)_{2\ (r-1)nm+jn+i} = 1 \end{array}$$

Finally, we define the nonzero entries of $A_k \equiv A_{njr} \in \mathbb{R}^{1\times N}$

$$(A_k)_{(r-1)nm+(j-1)n+n} = -1 \qquad (A_k)_{(r-1)nm+(j-1)n+n+1} = 1$$

for $j = 1, ..., m-1$, those of $A_k \equiv A_{imr} \in \mathbb{R}^{1\times N}$

$$(A_k)_{(r-1)nm+(m-1)n+i} = -1 \qquad (A_k)_{(r-1)nm+(m-1)n+i+1} = 1$$

for $i = 1, ..., n-1$ and those of $A_k \equiv A_{nmh} \in \mathbb{R}^{1\times N}$

$$(A_k)_{(h-1)nm+mn} = -1 \qquad (A_k)_{hnm+mn} = 1$$

for $h = 1, ..., r-1$.

With these settings, formula (11) can be written also as

$$f_1(x) = \sum_{k=1}^{N_1} \sqrt{\|A_k x\|^2 + \delta^2} \tag{19}$$

which is the sum of the ℓ_2 norm of the vectors $v_k = \begin{pmatrix} A_k x \\ \delta \end{pmatrix}$, $k = 1, ..., N_1$. Since $\delta \neq 0$, $f_1(x)$ is a C^2 differentiable function.

The convexity of the norm operator implies that the penalty function (19) is convex in $\mathbb{R}^N$. Then the Hessian matrix is positive semidefinite for any $x \in \mathbb{R}^N$.

Now, exploiting the definition (19), we can derive a matrix expression for the gradient of the edge preserving function (11), which is formally independent of the dimension d and on the boundary conditions.

Indeed, each l–th element of the gradient of (19) is given by

$$\frac{\partial f_1(x)}{\partial x_l} = \sum_{k=1}^{N_1} \frac{{A_k}^T A_k x}{D_k^{1/2}} \qquad l = 1, ..., N \tag{20}$$

Moreover, we define the block matrix A stacking the difference matrices A_k by rows

$$A = \begin{pmatrix} A_1 \\ A_2 \\ ... \\ A_{N_1} \end{pmatrix} \tag{21}$$

The matrix A is a $q \times N$ matrix: for periodic boundary conditions $q = dN$, where d is the dimension of x, while for reflexive boundary conditions $q = n-1$, $q = 2(n-1)(m-1)+n+m-2$ or $q = 3(n-1)(m-1)(r-1)+2(n-1)(m-1)+2(r-1)(n-1)+2(r-1)(m-1)+n+m+r-3$ for the 1D, 2D and 3D case respectively. We define also a block diagonal matrix $E(x) \in \mathbb{R}^{q\times q}$ as follows

$$E(x) = \text{diag}(D_k^{1/2} I_d)_{k=1,\dots,N_1} \tag{22}$$

where the dimension d of the k–th diagonal block is the number of rows of the corresponding block A_k in the matrix A. Then, the gradient can be written as

$$\nabla f_1(x) = A^T E(x)^{-1} A x \tag{23}$$

(see also [5], [13] and references therein). We remark that A in (21) has constant entries, while the diagonal entries of $E(x)$ are dependent on the variable x. Furthermore the gradient of $f_1(x)$ has the special form

$$\nabla f_1(x) = L(x)x$$

with $L(x) = A^T E(x)^{-1} A$. Since the diagonal entries of $E(x)$ are positive for all $x \in \mathbb{R}^N$, $E(x)$ and $E(x)^{-1}$ are positive definite matrices. Then the rank of $L(x)$ and its null space are equal to those of constant matrix A.
Now, from (20)–(23) we derive the Hessian matrix of the function (11) (see also [3]):

$$\begin{aligned}\nabla^2 f_1(x) &= \sum_{k=1}^{N_1} \left(\frac{1}{D_k^{1/2}} {A_k}^T A_k - \frac{1}{D_k^{3/2}} {A_k}^T A_k x x^T {A_k}^T A_k \right) = \qquad (24)\\ &= \sum_{k=1}^{N_1} \left(\frac{1}{D_k^{1/2}} {A_k}^T (I_d - \frac{1}{D_k} A_k x x^T {A_k}^T) A_k \right)\end{aligned}$$

Then

$$\nabla^2 f_1(x) = A^T E(x)^{-1} F(x) A \tag{25}$$

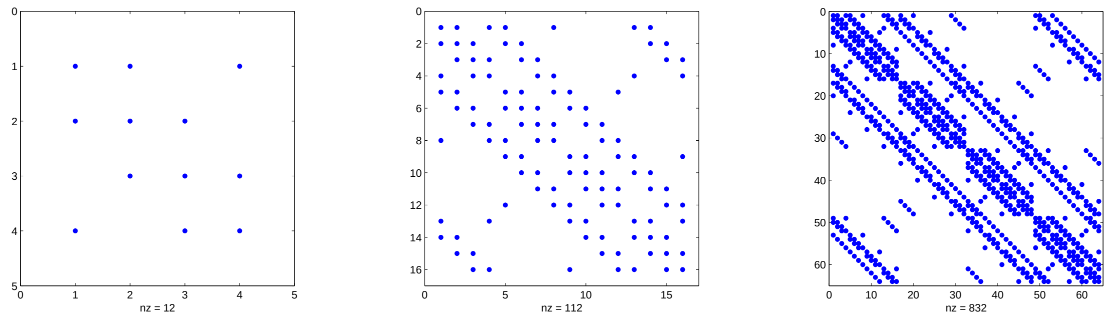
where A and $E(x)$ are given in (21) and $F(x)$ is the following square block diagonal matrix of order q

$$F(x) = \text{diag}\left(I_d - \frac{1}{D_k} A_k x x^T {A_k}^T \right)_{k=1,\dots,N_1}$$

where the dimension d of the k–th diagonal block is the number of rows of the matrix A_k.
In particular, we remark that each block of $F(x)$ is nonsingular and, consequently, $F(x)$ is nonsingular for all $x \in \mathbb{R}^N$. Indeed, the k–th diagonal block of $F(x)$ is the difference between the identity matrix and a dyadic product; from formula (15), we have that

$$\frac{1}{D_k} x^T {A_k}^T A_k x = \frac{\|A_k x\|^2}{D_k} < 1$$

Consequently, from the Sherman-Morrison theorem (see for example [8, p.19]), it follows that each diagonal block of $F(x)$ is nonsingular.



(a) (b) (c)

Figure 1. Structure of the Hessian matrix of the function (11) with periodic boundary conditions for $n = m = r = 4$: (a) $d = 1$; (b) $d = 2$; (c) $d = 3$.

Figure 1 shows the sparsity pattern of the Hessian matrix of $f_1(x)$ in (11) with periodic boundary conditions in the 1D, 2D and 3D cases. We observe that, for each row, we have three, seven and thirteen nonzero entries for $d = 1$, $d = 2$ and $d = 3$ respectively. The explicit expression of the nonzero entries of the 2D Hessian is given in Appendix A.

Theorem 3.1 *The null space of the Hessian matrix* $\nabla^2 f(x)$ *is given by the set of the minimum points of* $f_1(x)$*.*

Proof. We observe that, since $E(x)^{-1}F(x)$ is a symmetric positive definite matrix for all $x \in \mathbb{R}^N$, we have that

$$\mathcal{N}[\nabla^2 f_1(x)] = \mathcal{N}[A]$$

On the other hand, the minimum points of the convex function $f_1(x)$ satisfy the stationarity condition $\nabla f_1(x) = 0$, that is

$$A^T E(x)^{-1} A x = 0.$$

Then, the set of the minimum points of $f_1(x)$ is the subspace $\mathcal{N}[A]$. □

Since the minimum points of the functional (11) in both cases of periodic or reflexive boundary conditions are all the points $x \in \mathbb{R}^N$ such that $x_1 = x_2 = ... = x_N$, the previous theorem implies that the null space of the Hessian matrix is spanned by the vector $e_N \in \mathbb{R}^N$ with all entries equal to 1, that is

$$\mathcal{N}[\nabla^2 f_1(x)] = \{\alpha e_N : \alpha \in \mathbb{R}\} \quad \text{for all } x \in \mathbb{R}^N.$$

Thus, properties (P1) and (P2) hold for the edge preserving functional (11) with both periodic and reflexive boundary conditions, in the 1D, 2D and 3D cases.

4. Markov Random Field (MRF) regularization

For the 2D edge preserving image reconstruction, we can consider the following penalty function [7]

$$f_1(x) = \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{\ell \in \mathcal{N}_{ij}} \sqrt{\left(\frac{x_{ij} - x_{ij}^{(\ell)}}{w_{ij}^{(\ell)}} \right)^2 + \delta^2} \tag{26}$$

where $x_{ij}^{(\ell)}$, $\ell \in \mathcal{N}_{ij} \subseteq \{1, ..., 8\}$ denote the eight first neighbors of the pixel ij, while $w_{ij}^{(\ell)}$ are positive weights (typically, $w_{ij}^{(\ell)} = 1$ for the vertical and horizontal neighbors and $w_{ij}^{(\ell)} = \sqrt{2}$ for the diagonal neighbors). The function (26) is a simple example of the Markov random field regularization [6], where $f_1(x)$ represents a sum of potentials involving only the "nearest neighbor" interactions between the components of x.
When we assume periodic boundary conditions, $\mathcal{N}_{ij} = \{1, ..., 8\}$ for all $i = 1, ..., n$, $j = 1, ..., m$. Then for any $k = 1, ..., nm$ we can consider eight $1 \times nm$ vectors $A_k^{(1)}$,...,$A_k^{(8)}$, with all components equal to zero except for the following entries

$$\begin{array}{llllllll}
[A_k^{(1)}]_k & = 1 & [A_k^{(1)}]_{k-1} & = -1; & [A_k^{(2)}]_k & = 1 & [A_k^{(2)}]_{k+1} & = -1 \\
[A_k^{(3)}]_k & = 1 & [A_k^{(3)}]_{k-n} & = -1; & [A_k^{(4)}]_k & = 1 & [A_k^{(4)}]_{k+n} & = -1 \\
[A_k^{(5)}]_k & = \frac{1}{\sqrt{2}} & [A_k^{(5)}]_{k+n-1} & = -\frac{1}{\sqrt{2}}; & [A_k^{(6)}]_k & = \frac{1}{\sqrt{2}} & [A_k^{(6)}]_{k+n+1} & = -\frac{1}{\sqrt{2}} \\
[A_k^{(7)}]_k & = \frac{1}{\sqrt{2}} & [A_k^{(7)}]_{k-n-1} & = -\frac{1}{\sqrt{2}}; & [A_k^{(8)}]_k & = \frac{1}{\sqrt{2}} & [A_k^{(8)}]_{k-n+1} & = -\frac{1}{\sqrt{2}}
\end{array}$$

With these settings, the functional $f_1(x)$ can be written as

$$f_1(x) = \sum_{k=1}^{nm} \sum_{\ell=1}^{8} \sqrt{\|A_k^{(\ell)} x\|^2 + \delta^2} = \sum_{k=1}^{nm} \sum_{\ell=1}^{8} D_k^{(\ell)\frac{1}{2}}$$

where $D_k^{(\ell)} \equiv D_{ij}^{(\ell)} = ((x_{ij} - x_{ij}^{(\ell)})/w_{ij}^{(\ell)})^2 + \delta^2$. Also in this case, if $\delta \neq 0$, $f_1(x)$ is a C^2 differentiable convex function and its Hessian is a positive semidefinite matrix for any $x \in \mathbb{R}^{nm}$. From the previous expression we derive the gradient as

$$\nabla f_1(x) = \sum_{k=1}^{nm} \sum_{\ell=1}^{8} \frac{A_k^{(\ell)T} A_k^{(\ell)} x}{D_k^{(\ell)\frac{1}{2}}} \tag{27}$$

We now define the following square matrices of order nm

$$A^{(\ell)} = \begin{pmatrix} A_1^{(\ell)} \\ \vdots \\ A_{nm}^{(\ell)} \end{pmatrix} \qquad E(x)^{(\ell)} = \text{diag}\left(D_{ij}^{(\ell)\frac{1}{2}} \right) \qquad \ell = 1, ..., 8$$

When reflexive boundary conditions are assumed, we have that $\mathcal{N}_{ij} = \{1, ..., 8\}$ for $i = 2, ..., n-1$, $j = 2, ..., m-1$, while for the corners $((i, j) = (1, 1), (n, 1), (1, m), (n, m))$ $\mathcal{N}_{ij}$ contains only 3 indices, and for the remaining boundary pixels, $\mathcal{N}_{ij}$ has only 5 indices. Then, the number $N_1^{(\ell)}$ of rows of the matrices $A^{(\ell)}$ and $E(x)^{(\ell)}$ depends on the value of

ℓ. In any case, with both periodic and reflexive boundary conditions, after exchanging the two summations in (27), we have

$$\nabla f_1(x) = \left(\sum_{\ell=1}^{8} A^{(\ell)^T} E(x)^{(\ell)^{-1}} A^{(\ell)} \right) x \tag{28}$$

The gradient vector has the form $\nabla f_1(x) = L(x)x$, where $L(x) = \sum_{\ell=1}^{8} A^{(\ell)^T} E(x)^{(\ell)^{-1}} A^{(\ell)}$. Since $A^{(\ell)^T} E(x)^{(\ell)^{-1}} A^{(\ell)}$, $\ell = 1, ..., 8$ are symmetric positive semidefinite matrices, from the Lemma 2.1 we have $\mathcal{N}[L(x)] = \bigcap_{\ell=1}^{8} \mathcal{N}[A^{(\ell)}]$ for all $x \in \mathbb{R}^{nm}$. Since $A^{(\ell)} e_N = 0$, $\ell = 1, ..., 8$, we have that property (P2) holds

$$\nabla f_1(x)^T e_N = 0 \quad \text{ for all } x \in \mathbb{R}^{nm}$$

Moreover, the following expression for the Hessian can be derived:

$$\nabla^2 f_1(x) = \sum_{\ell=1}^{8} A^{(\ell)^T} E(x)^{(\ell)^{-1}} F(x)^{(\ell)} A^{(\ell)} \tag{29}$$

where $F(x)^{(\ell)} = \text{diag}(\delta^2 / D_k^{(\ell)})_{k=1,...,N_1}$ is a diagonal matrix with positive entries. Thus, the Hessian of the MRF regularization function can be written as the sum of eight positive semidefinite matrices $M^{(\ell)} = A^{(\ell)^T} E(x)^{(\ell)^{-1}} F(x)^{(\ell)} A^{(\ell)}$. From Lemma 2.1, we have $\mathcal{N}[\nabla^2 f_1(x)] = \bigcap_{\ell=1}^{8} \mathcal{N}[A^{(\ell)}]$ for all $x \in \mathbb{R}^{nm}$. Proceeding as in Theorem 3.1, we can conclude that the null space of the Hessian matrix is the set of the minimum points of (26), that is, $\{\alpha e_N : \alpha \in \mathbb{R}\} = \mathcal{N}[A^{(\ell)}] = \bigcap_{\ell=1}^{8} \mathcal{N}[A^{(\ell)}] = \mathcal{N}[\nabla^2 f_1(x)]$. In summary, we proved properties (P1) and (P2) for the MRF regularization function (26) in the 2D case; similar conclusions can be extended to 3D case.

5. Tikhonov regularization

We consider the standard Tikhonov regularization, based on the ℓ_2 norm of the discrete gradient of x, which can be expressed as

$$f_1(x) = \frac{1}{2} \sum_{k=1}^{N_1} D_k \tag{30}$$

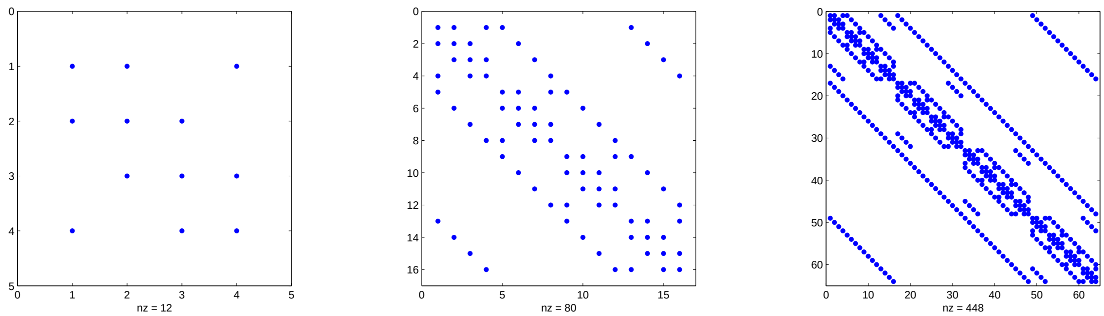
where D_k and N_1 are defined in (12)–(14) with $\delta = 0$. Using the same settings as in section 3, we have also

$$f_1(x) = \frac{1}{2} \sum_{k=1}^{N_1} \|A_k x\|^2 \tag{31}$$

The regularization function $f_1(x)$ is a convex quadratic function and its gradient is given by

$$\nabla f_1(x) = \sum_{k=1}^{N_1} A_k^T A_k x = A^T A x \tag{32}$$

where A is given in (21). The gradient vector of $f_1(x)$ has the special form $\nabla f_1(x) = L(x)x$, where $L(x) = A^T A$ and $L(x) \equiv L$ is a constant matrix.



(a) (b) (c)

Figure 2. Structure of the Hessian matrix of the function (30), with $n = m = r = 4$: (a) $d = 1$ (b) $d = 2$ (c) $d = 3$.

The Hessian of $f_1(x)$ is the matrix of order N given by

$$\nabla^2 f_1(x) = \sum_{k=1}^{N_1} A_k^T A_k x = A^T A \tag{33}$$

Under periodic or reflexive boundary condition, properties (P1) and (P2) hold for the Tikhonov regularization function (30) for $d = 1, 2, 3$. Indeed, we have that he sum of all entries of the gradient vector is equal to 0 ($e_N^T A^T A x = 0$) and $e_N \in \mathcal{N}[L] = \mathcal{N}[A] = \mathcal{N}[\nabla^2 f_1(x)]$. Following the proof of Theorem 3.1, we can prove that the rank of A is $N - 1$ for $d = 1, 2, 3$, and, consequently, its null space is spanned by the constant vector e_N.

In Figure 2 the sparsity pattern of the Hessian of Tikhonov function (30) for periodic boundary conditions is reported; for all $d = 1, 2, 3$ the off–diagonal elements are equal to -1, while the diagonal entries are equal to 2, 4 and 6 respectively. Furthermore, for each row, we have three, five and seven nonzero entries for $d = 1$, $d = 2$ and $d = 3$ respectively.

6. Conclusions

The results of the previous sections ensure that, when $f_0(x)$ is the generalized Kullback–Leibler divergence (7) and $f_1(x)$ is the TV-HS potential (11) or the MRF function (26) or the Tikhonov regularization term (30), properties (P1) and (P2) are satisfied. Lemma 2.1 and property (P1) enable us to conclude that the function $f_\beta(x) = f_0(x) + \beta f_1(x)$ is strictly convex over Ω for every $\beta > 0$.

This statement holds in 1D, 2D and 3D denoising and deblurring problems, only assuming that the nonnegative blurring matrix H is normalized such that $H^T e_M = e_N$ and $(He_N)_i > 0$ for any i.

As concerns the existence and uniqueness of the solution of the discrepancy equation (2), properties (P1) and (P2) allow us to apply the analysis proposed in [2] to all the

penalty functions considered in this paper.
In particular, in the settings described in Section 2, the following results can be stated (see Lemma 3–5, Theorem 2–3 in [2]):
Denoising: equation (2) has an unique solution if and only if the data satisfy the following condition

$$\frac{1}{N}\sum_{k\in I_1} y_k \ln y_k > \frac{1}{2} + \bar{y}\ln\bar{y} \tag{34}$$

where $\bar{y} = \frac{1}{N}\sum_{k=1}^{N} y_k$;
Deblurring: under the assumption that

$$\frac{1}{N}\sum_{j=1}^{N}(H^T y)_j > \gamma \tag{35}$$

equation (2) has an unique solution if and only if the following conditions are satisfied

$$f_0(x^*) < \frac{M}{2}, \tag{36}$$

$$f_0(\bar{c}e_N) > \frac{M}{2} \tag{37}$$

where x^* is a minimum point of $f_0(x)$ such that the minimizer of $f_\beta(x)$ converges to x^* as β converges to zero and $\bar{c}$ is the unique solution of the equation $\sum_{i\in I_1}\frac{(He_N)_i y_i}{\bar{c}(He_N)_i+b_i} = N$.
In the case of image deconvolution, since $M = N$ and $He_N = e_N$, the assumption (35) reduces to $\bar{y} > \gamma$, and condition (37) coincides with (34).

Appendix A

We give the explicit expression of the gradient and of the Hessian matrix for the 2D edge preserving function (11), assuming periodic boundary conditions. The k-th component of the gradient is given by

$$\frac{\partial f_1(x)}{\partial x_{ij}} = \frac{2x_{ij} - x_{i+1j} - x_{ij+1}}{D_{ij}^{1/2}} + \frac{(x_{ij} - x_{i-1j})}{D_{i-1j}^{1/2}} + \frac{(x_{ij} - x_{ij-1})}{D_{ij-1}^{1/2}}$$

where D_{ij} are defined in (13). Then, the nonzero entries of the Hessian are defined as follows:

$$\frac{\partial^2 f_1(x)}{\partial x_{ij}^2} = \frac{(x_{i+1j} - x_{ij+1})^2 + 2\delta^2}{D_{ij}^{3/2}} + \frac{(x_{i+1j-1} - x_{ij-1})^2 + \delta^2}{D_{ij-1}^{3/2}} + \frac{(x_{i-1j+1} - x_{i-1j})^2 + \delta^2}{D_{i-1j}^{3/2}}$$

$$\frac{\partial^2 f_1(x)}{\partial x_{ij}\partial x_{i+1j}} = \frac{(x_{ij+1} - x_{ij})(x_{i+1j} - x_{ij+1}) - \delta^2}{D_{ij}^{3/2}}$$

$$\frac{\partial^2 f_1(x)}{\partial x_{ij}\partial x_{ij+1}} = \frac{(x_{i+1j} - x_{ij})(x_{ij+1} - x_{i+1j}) - \delta^2}{D_{ij}^{3/2}}$$

$$\frac{\partial^2 f_1(x)}{\partial x_{ij}\partial x_{i-1j}} = \frac{(x_{ij} - x_{i-1j+1})(x_{i-1j+1} - x_{i-1j}) - \delta^2}{D_{i-1j}^{3/2}}$$

$$\frac{\partial^2 f_1(x)}{\partial x_{ij}\partial x_{ij-1}} = -\frac{(x_{ij}-x_{i+1j-1})(x_{i+1j-1}-x_{ij-1})-\delta^2}{D_{ij-1}^{3/2}}$$
$$\frac{\partial^2 f_1(x)}{\partial x_{ij}\partial x_{i-1j+1}} = -\frac{(x_{ij}-x_{i-1j})(x_{i-1j+1}-x_{i-1j})}{D_{i-1j}^{3/2}}$$
$$\frac{\partial^2 f_1(x)}{\partial x_{ij}\partial x_{i+1j-1}} = -\frac{(x_{ij}-x_{ij-1})(x_{i+1j-1}-x_{ij-1})}{D_{ij-1}^{3/2}}$$

References

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

[2] Bertero M, Boccacci P, Talenti G, Zanella R and Zanni L 2009 A discrepancy principle for Poisson data *submitted*

[3] Chan T F, Golub G H and Mulet P 1999 A nonlinear primal–dual method for total variation-based image restoration *SIAM J. Sci. Comput.* **20** 1964–1977

[4] Charbonnier P, Blanc–Féraud L, Aubert G and Barlaud A 1997 Deterministic edge–preserving regularization in computed imaging *IEEE Trans. Image Processing* **6** 298–311

[5] Chartrand R and Staneva V 2008 Total variation regularisation of images corrupted by non-Gaussian noise using a quasi-Newton method *Image Processing IET* **2** 295–303

[6] Geman S and Geman D 1984 Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images *IEEE Trans. Pattern Anal. Intell.* **6** 721–741

[7] Geman S, Manbeck K and McClure D E 1993 A comprehensive statistical model for single photon emission tomography *Markov Random Fields: theory and Applications* eds. Chellappa R and Jain A. (Boston: Academic Press) 93–130

[8] Horn R A and Johnson C R 1990 Matrix Analysis (Cambridge: Cambridge University Press)

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

[10] Vogel C R 2002 Computational Methods for Inverse Problems (Philadelphia: SIAM)

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

[12] Zanella R, Boccacci P, Zanni L and Bertero M 2009 Efficient gradient projection methods for edge–preserving removal of Poisson noise *Inverse Problems* **25** 045010

[13] Zhu M and Chan T 2008 An Efficient Primal–Dual Hybrid Gradient Algorithm For Total Variation Image Restoration *UCLA CAM Report* **08–34** 1–29