Annals of the University of Craiova, Mathematics and Computer Science Series
Volume 49(1), 2022, Pages 84–98, DOI:
ISSN: 1223-6934

Kronecker product approximation for the total variation regularization in image restoration

Abdeslem Hafid Bentbib, Abderrahman Bouhamidi, and Karim Kreit

Abstract. In this paper, we propose a new algorithm to restore blurred and noisy images based on the total variation regularization, where the discrete associated Euler-Lagrange problem is solved by exploiting the structure of the matrices and transforming the initial problem to a generalized Sylvester linear matrix equation by using a special Kronecker product approximation. Afterwards, global Krylov subspace methods are used to solve the linear matrix equation. Numerical experiments are given to illustrate the effectiveness of the proposed method.

2010 Mathematics Subject Classification. Primary : 94A08, 15A29 ; Secondary : 65F22.
Key words and phrases. Total variation, inverse problem, image restoration, image processing, regularization, global GMRES, L-curve and Kronecker product approximation.

1. Introduction

The problem of image restoration consists in obtaining a true image from an observed and degraded image by blur and additive noise which is assumed to be white and Gaussian. The noisy and blurred image problem is modeled by

$$\mathbf{z} = K\mathbf{u} + \mathbf{n}, \tag{1}$$

here, K is a known blur matrix, $\mathbf{n}$ is an additive Gaussian white noise, $\mathbf{u}$ and $\mathbf{z}$ are vectors representing the true image and the degraded image, respectively.

In general the blurring matrix K is ill-conditioned, then the problem of image restoration will be very sensitive to the noise vector. To overcome this problem we regularize the model (1). One of the most known regularization methods is the total variation (TV). The total variation based on the model of Rudin-Osher-Fatemi (ROF) [16] which is one of the well known problem in image restoration due to its effectiveness to preserve sharp features. The ROF model is given as follows

$$\min_{\mathbf{u}} \left(\lambda \int_{\Omega} \sqrt{|\nabla \mathbf{u}|} dxdy + \frac{1}{2} \|K\mathbf{u} - \mathbf{z}\|_{L^2}^2 \right), \tag{2}$$

where $\lambda > 0$ is the regularization parameter which controls the trade-off between the smoothness of $\mathbf{u}$ and the goodness of fit-to-the-data. $\nabla \mathbf{u}$ denotes the gradient operator of $\mathbf{u}$ and

$$\int_{\Omega} \sqrt{|\nabla \mathbf{u}|} dxdy, \tag{3}$$

Received June 07, 2021. Accepted January 9, 2022.

is the total variation semi norm [2]. Solving the corresponding partial equation of (2) is to hard owing to the degeneracy of the diffusion coefficients on edges. For this reason Chan, Golub and Mulet [6] replaced the norm $|\nabla \mathbf{u}|$ in (3) by the smoothed one $\sqrt{|\nabla \mathbf{u}|^2+\beta}$, where the real positive β is a smaller regularization parameter useful for avoiding numerical instabilities in the minimization flow and for the differentiabily of (3) when the $\nabla u = 0$. Therefore the model (2) is replaced by the new one

$$\min_{\mathbf{u}} \left(\lambda \int_{\Omega} \sqrt{|\nabla \mathbf{u}|^2+\beta} dxdy + \frac{1}{2}\|K\mathbf{u}-\mathbf{z}\|_{L^2}^2 \right). \tag{4}$$

The functional in (4) is strictly convex, then its global minimizer is unique. The well-posedness of problem (4) with $\beta \longrightarrow 0^+$ has been discussed in [1] and [14]. The corresponding Euler-Lagrange equation for (4) is given by:

$$-\lambda \nabla \cdot \left(\frac{\nabla \mathbf{u}}{\sqrt{|\nabla \mathbf{u}|^2+\beta}} \right) + K^*(K\mathbf{u}-\mathbf{z}) = 0, \tag{5}$$

where $\nabla \cdot F$ denotes the divergence operator of an element F and K^* is the adjoint operator of K with respect to the L_2 inner product. In this article, we will discuss the denoising and the deblurring problem. We suggest a special Kronecker product approximation to the corresponding discretization of the problem (5) to yield a Sylvester matrix equation that will be solved by using the global generalized minimum residual (Gl-GMRES) method [3, 12]. The outline of this paper is as follows. In Section 2, we apply the finite difference scheme and transform the problem (5) to a matrix problem. In Section 3, we review some notations, definitions and results relative to the Kronecker product approximation and apply this approach to our problem. We show how some linear matrix equations can be derived from TV regularization and combine our approach with the global-GMRES method to generalized Sylvester matrix equations. Section 4, is devoted to the choice of a suitable regularization parameters using the L-curve method. Finally, in Section 5, we give some experimental results and applications in image restoration to illustrate the effectiveness of our approach.

2. Finite difference scheme and discrete problem

Let us consider the Euler-Lagrange equation (5) with zero Neumann boundary conditions.

$$-\lambda \nabla \cdot \mathbf{P} + K^*(K\mathbf{u}-\mathbf{z}) = 0, \tag{6}$$

where $\mathbf{P}$ denotes the term $\dfrac{\nabla \mathbf{u}}{\sqrt{|\nabla \mathbf{u}|^2+\beta}}$ defined in (5).

To introduce the discrete problem we assume that the image is given by a matrix of size $n \times n$ i.e., an image $\mathbf{u}$ is an element of the matrix space $\mathbb{X} = \mathbb{R}^{n\times n}$. This space $\mathbb{X}$ is equipped with the inner product $\langle \cdot \rangle_{\mathbb{X}}$ and the associated norm $\|\cdot\|_{\mathbb{X}}$. For simplification, each pixel of $\mathbf{u}$ is denoted by $\mathbf{u}(x_\ell, y_k)$ or $u_{\ell,k}$. We consider the discrete Neumann boundary conditions:

$$u_{0,k} = u_{1,k},\ u_{n-1,k} = u_{n,k},\ u_{\ell,0} = u_{\ell,1} \text{ and } u_{\ell,n-1} = u_{\ell,n}.$$

The discrete gradient $\nabla \mathbf{u}$ is an element of $\mathbb{Y} = \mathbb{X} \times \mathbb{X}$ and is given by the forward differences [5], i.e.,

$$\nabla \mathbf{u}(x_\ell, y_k) = \begin{pmatrix} \partial_x \mathbf{u}(x_\ell, y_k) \\ \partial_y \mathbf{u}(x_\ell, y_k) \end{pmatrix}, \quad 1 \leqslant \ell, k \leqslant n,$$

where

$$\partial_x \mathbf{u}(x_\ell, y_k) \simeq \begin{cases} u_{\ell+1,k} - u_{\ell,k} & if \ \ \ell < n \\ 0 & if \ \ \ell = n \end{cases}, \tag{7}$$

and

$$\partial_y \mathbf{u}(x_\ell, y_k) \simeq \begin{cases} u_{\ell,k+1} - u_{\ell,k} & if \ \ k < n \\ 0 & if \ \ k = n \end{cases}. \tag{8}$$

Then by considering (7) and (8) for all $0 \leq \ell < n-1$ and $0 \leq k < n-1$, the elements of the term $\mathbf{P}$ is given by

$$\mathbf{P}_{\ell,k} \simeq \begin{pmatrix} \dfrac{u_{\ell+1,k} - u_{\ell,k}}{\sqrt{(u_{\ell+1,k} - u_{\ell,k})^2 + (u_{\ell,k+1} - u_{\ell,k})^2 + \beta}} \\ \dfrac{u_{\ell,k+1} - u_{\ell,k}}{\sqrt{(u_{\ell+1,k} - u_{\ell,k})^2 + (u_{\ell,k+1} - u_{\ell,k})^2 + \beta}} \end{pmatrix}. \tag{9}$$

In equation (6) we have to compute the divergence of $\mathbf{P}$. Then, taking into account the Neumann boundary, we introduce also a discrete divergence operator $\nabla \cdot : \mathbb{Y} \longrightarrow \mathbb{X}$ such that, for all $\mathbf{P} \in \mathbb{Y}$ and all $\mathbf{u} \in \mathbb{X}$, we have

$$\langle -\nabla \cdot \mathbf{P}, \mathbf{u} \rangle_{\mathbb{X}} = \langle \mathbf{P}, \nabla \mathbf{u} \rangle_{\mathbb{Y}},$$

where $\langle \cdot, \cdot \rangle_{\mathbb{Y}}$ denotes the inner product associated to the space $\mathbb{Y}$. Thus, it is easy to see that, for every $\mathbf{P} = \left(P^{(1)}, P^{(2)}\right) \in \mathbb{Y}$, the discrete divergence can be given by (see [5]):

$$\nabla \cdot \mathbf{P}_{\ell,k} = \begin{cases} P^{(1)}_{\ell,k} & \ell = 1 \\ -P^{(1)}_{\ell-1,k} & \ell = n \\ P^{(1)}_{\ell,k} - P^{(1)}_{\ell-1,k} & otherwise \end{cases} + \begin{cases} P^{(2)}_{\ell,k} & k = 1 \\ -P^{(2)}_{\ell,k-1} & k = n \\ P^{(2)}_{\ell,k} - P^{(2)}_{\ell,k-1} & otherwise \end{cases}$$

After discretization, the equation (6), can be written as

$$-\lambda \mathcal{A}(\mathbf{u})\, \mathbf{u} + K^T (K\mathbf{u} - \mathbf{z}) = 0, \tag{10}$$

here, $\mathbf{u} = (u_{1,1}, \ldots, u_{n,1}, \ldots, u_{1,n}, \ldots, u_{n,n})^T$ is the vector $\mathbf{u} = vec(U)$ obtained by stacking the columns of the matrix image $U = (u_{i,j})_{1 \leq i,j \leq n}$ from left to right. The matrix $\mathcal{A}(\mathbf{u})$ of size $n^2 \times n^2$ is a block tridiagonal matrix given by,

$$\mathcal{A}(\mathbf{u}) = \begin{pmatrix} A_1(\mathbf{u}) & B_1(\mathbf{u}) & O & \cdots & \cdots & O \\ B_1(\mathbf{u}) & A_2(\mathbf{u}) & B_2(\mathbf{u}) & \ddots & & \vdots \\ O & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & O \\ \vdots & & \ddots & B_{n-2}(\mathbf{u}) & A_{n-1}(\mathbf{u}) & B_{n-1}(\mathbf{u}) \\ O & \cdots & \cdots & O & B_{n-1}(\mathbf{u}) & A_n(\mathbf{u}) \end{pmatrix},$$

where the blocks $A_k(\mathbf{u})$ and $B_k(\mathbf{u})$ of size $n \times n$ are, respectively, the tridiagonal and diagonal matrices given by

$$\begin{aligned} A_k(\mathbf{u}) &= Tridiag(\mathbf{a}_k(\mathbf{u}), -\mathbf{d}_k(\mathbf{u}), \mathbf{a}_k(\mathbf{u})), \qquad \text{for } 1 \le k \le n \\ B_k(\mathbf{u}) &= Diag(\mathbf{b}_k(\mathbf{u})) \qquad\qquad\qquad\qquad\quad \text{for } 1 \le k \le n-1, \end{aligned}$$

here, the vectors $\mathbf{a}_k(\mathbf{u})$, $\mathbf{b}_k(\mathbf{u})$ and $\mathbf{d}_k(\mathbf{u})$ are given by

$$\begin{aligned} \mathbf{a}_k(\mathbf{u}) &= (a_{1,k}(\mathbf{u}), \cdots, a_{n-1,k}(\mathbf{u}))^T, \ 1 \le k \le n \\ \mathbf{b}_k(\mathbf{u}) &= (b_{1,k}(\mathbf{u}), \cdots, b_{n,k}(\mathbf{u}))^T, \quad 1 \le k \le n-1 \\ \mathbf{d}_k(\mathbf{u}) &= (d_{1,k}(\mathbf{u}), \cdots, d_{n,k}(\mathbf{u}))^T, \quad 1 \le k \le n. \end{aligned}$$

The coefficients $d_{\ell,k}(\mathbf{u})$ are given by

$$d_{\ell,k}(\mathbf{u}) = a_{\ell-1,k}(\mathbf{u}) + a_{\ell,k}(\mathbf{u}) + b_{\ell,k-1}(\mathbf{u}) + b_{\ell,k}(\mathbf{u}) \quad \text{for } 1 \le \ell, k \le n.$$

where the coefficients $a_{\ell,k}(\mathbf{u})$ and $b_{\ell,k}(\mathbf{u})$ are defined as follows

$$a_{\ell,k}(\mathbf{u}) = \begin{cases} 0 & \text{for } \ell = 0, 1, n \text{ and } 1 \le k \le n \\ \dfrac{1}{\sqrt{(u_{\ell+1,k}-u_{\ell,k})^2+\beta}} & \text{for } \begin{cases} 2 \le \ell \le n-1 \\ k = 1 \text{ or } k = n \end{cases} \\ \dfrac{1}{\sqrt{(u_{\ell+1,k}-u_{\ell,k})^2+(u_{\ell,k+1}-u_{\ell,k})^2+\beta}} & \text{for } 2 \le \ell, k \le n-1, \end{cases}$$

and

$$b_{\ell,k}(\mathbf{u}) = \begin{cases} 0 & \text{for } 1 \le \ell \le n \text{ and } k = 0, 1, n \\ \dfrac{1}{\sqrt{(u_{\ell,k+1}-u_{\ell,k})^2+\beta}} & \text{for } \begin{cases} \ell = 1 \text{ or } \ell = n \\ 2 \le k \le n-1 \end{cases} \\ \dfrac{1}{\sqrt{(u_{\ell+1,k}-u_{\ell,k})^2+(u_{\ell,k+1}-u_{\ell,k})^2+\beta}} & \text{for } 2 \le \ell, k \le n-1. \end{cases}$$

For the sake of simplicity, when there is no risque of confusion, we write $\mathcal{A}$ instead of $\mathcal{A}(\mathbf{u})$.

3. Kronecker product approximation and the associated Sylvester matrix equation

Let $A = (a_{ij})$ and $B = (b_{ij})$ be $n \times p$ and $s \times q$ matrices respectively. The Kronecker product of the matrices A and B is defined as the $(ns) \times (pq)$ matrix $A \otimes B = (a_{ij}B)$. The *vec* operator transforms the matrix A to a vector a of size $np \times 1$ by stacking the columns of A. Some properties of the Kronecker product are given below [13],

$$\begin{aligned} (A \otimes B)(C \otimes D) &= (AC) \otimes (BD) \\ (A \otimes B)^{-1} &= A^{-1} \otimes B^{-1}, \quad \text{if } A, B \text{ are nonsingular} \\ vec(AXB) &= (B^T \otimes A)vec(X) \\ (A \otimes B)^T &= A^T \otimes B^T \end{aligned}$$

$(A \otimes B)$ is orthogonal whenever A *and* B are orthogonal. (11)

Let A and B be two matrices in $\mathbb{R}^{n \times p}$, we define the following inner product $\langle A, B \rangle_F = tr(A^T B)$ where $tr(Z)$ denotes the trace of the square matrix Z. The corresponding well known Frobenius norm $\|\,.\,\|_F$ is given by $\|A\|_F = \sqrt{\langle A, A \rangle_F}$ and it is equivalent to the norm $\|\cdot\|_{\mathbb{X}}$. A system of matrices in $\mathbb{R}^{n \times p}$ is said to be F-orthogonal if it is orthogonal with respect to the scalar product $\langle\,.\,,\,.\,\rangle_F$.

The Kronecker product approximation (KPA) problem was first introduced by Pitsianis and Van Loan [15, 7]. For a matrix A of size $m_1m_2 \times n_1n_2$, the KPA problem consists in finding two matrices B and C of size $m_1 \times n_1$ and $m_2 \times n_2$, respectively, that is the Frobenius norm $||A - B \otimes C||_F$ is minimal. To this end, the authors define the matrix $B \otimes C$, which is a special rearrangement of A, relative to the blocking parameters m_1, m_2, n_1 and n_2. For our purpose, we use a vectorizing operation which turns a matrix into a vector by stacking the columns of the matrix. The KPA of $\mathcal{A}$ consists in finding a couple of matrices $(\widehat{\mathcal{A}}_1, \widehat{\mathcal{A}}_2)$ solution of the following problem

$$(\widehat{\mathcal{A}}_1, \widehat{\mathcal{A}}_2) = \underset{\mathcal{A}_1, \mathcal{A}_2}{\operatorname{argmin}} ||\mathcal{A} - \mathcal{A}_2 \otimes \mathcal{A}_1||_F. \tag{12}$$

Taking into account of the structure of the matrix $\mathcal{A}$ as a block tridiagonal matrix, the matrices $\mathcal{A}_1$ and $\mathcal{A}_2$ must be also block tridiagonal matrices. To simplify our problem, we can assume that the tridiagonal matrix $\mathcal{A}_1 = \widehat{A}$ is fix once for all. We can choose this matrix in the following form

$$\widehat{A} = Tridiag(\widehat{\mathbf{a}}, \widehat{\mathbf{d}}, \widehat{\mathbf{a}}), \tag{13}$$

where $\widehat{\mathbf{a}} = (a_1, \ldots, a_{n-1})^T$ and $\widehat{\mathbf{d}} = (d_1, \ldots, d_n)^T$. Then, we may choose the vectors $\widehat{\mathbf{a}}$ and $\widehat{\mathbf{d}}$ as the mean values of the vectors $\mathbf{a}_k(\mathbf{u})$ and $\mathbf{d}_k(\mathbf{u})$. Namely,

$$\widehat{\mathbf{a}} = \frac{1}{n}\sum_{k=1}^{n} \mathbf{a}_k(\mathbf{u}), \quad \widehat{\mathbf{d}} = \frac{1}{n}\sum_{k=1}^{n} \mathbf{d}_k(\mathbf{u}). \tag{14}$$

It is also possible to use the quadratic mean square values as following

$$a_\ell = \frac{1}{n}\sqrt{\sum_{k=1}^{n} a_{\ell,k}^2(\mathbf{u})}, \quad d_\ell = \frac{1}{n}\sqrt{\sum_{k=1}^{n} d_{\ell,k}^2(\mathbf{u})}, \text{ for } 1 \leq \ell \leq n. \tag{15}$$

Now, it remain to find a tridiagonal matrix $\widehat{\mathcal{A}}_2$ of the form

$$\widehat{\mathcal{A}}_2 = \widehat{E} = Tridiag(\widehat{\mathbf{b}}, \widehat{\mathbf{e}}, \widehat{\mathbf{b}}), \tag{16}$$

which is the solution of the KPA problem (12).

Theorem 3.1. *Assume that the first element $\widehat{\mathcal{A}}_1$ of the minimization (12) is given by the formulations (13) and ((14) or (15)), then the second element $\widehat{\mathcal{A}}_2$ is the tridiagonal matrix (16) where*

$$\widehat{e}_k = \frac{\langle A_k, \widehat{A}\rangle_F}{||\widehat{A}||_F^2}, \ 1 \leq k \leq n \quad \text{and} \quad \widehat{b}_k = \frac{\langle B_k, \widehat{A}\rangle_F}{||\widehat{A}||_F^2}, \ 1 \leq k \leq n-1. \tag{17}$$

Proof. We set $\widehat{\mathbf{e}} = (e_1, \ldots, e_n)^T$ and $\widehat{\mathbf{b}} = (b_1, \ldots, b_{n-1})^T$, using the Kronecker product properties (11), we can write

$$||\mathcal{A} - \widehat{E} \otimes \widehat{A}||_F^2 = \sum_{i=1}^{n} ||A_i - e_i\widehat{A}||_F^2 + 2\sum_{i=1}^{n-1} ||B_i - b_i\widehat{A}||_F^2. \tag{18}$$

Then, in this case, the solution of the problem (12) is obtained by solving each quadratic form in the right hand side of (18). The solution is reached for $\widehat{\mathbf{e}} = (\widehat{e}, \ldots, \widehat{e}_n)^T$

and $\widehat{\mathbf{b}} = (\widehat{b}_1, \ldots, \widehat{b}_{n-1})^T$ with

$$\widehat{e}_k = \frac{\langle A_k, \widehat{A}\rangle_F}{||\widehat{A}||_F^2},\ 1 \le k \le n \quad \text{and} \quad \widehat{b}_k = \frac{\langle B_k, \widehat{A}\rangle_F}{||\widehat{A}||_F^2},\ 1 \le k \le n-1.$$

It follows that the matrix $\mathcal{A}$ may be approximated in the mean-square sense by the Kronecker product

$$\mathcal{A} \approx \widehat{E} \otimes \widehat{A},$$

where $\widehat{E} = Tridiag(\widehat{\mathbf{b}}, \widehat{\mathbf{e}}, \widehat{\mathbf{b}})$. □

In the context of image restoration, when the point spread function (PSF) is separable, the blurring matrix K given in (4) is assumed to be decomposed as a Kronecker product $K = K_2 \otimes K_1$ of two blurring matrices of appropriate sizes. Finally, equation (10) may be approximated by the following equation

$$-\lambda\big(\widehat{E} \otimes \widehat{A}\big)\mathbf{u} + \big(K_2 \otimes K_1\big)^T\big(K_2 \otimes K_1\big)\mathbf{u} = \big(K_2 \otimes K_1\big)^T\mathbf{z}.$$

According to the properties of the Kronecker product and the fact that the matrix $\widehat{E}$ is symmetric, the last equation is equivalent to the following generalized Sylvester matrix equation

$$-\lambda\widehat{A}(U)U\widehat{E}(U) + (K_1^TK_1)U(K_2^TK_2) = K_1^TZK_2, \tag{19}$$

where U and Z are $n \times m$-matrices obtained from $\mathbf{u}$ and $\mathbf{z}$ respectively by using the *rechape* operator [7]. To solve the problem (19), we use the following iterative algorithm

Algorithm 1: Iterative solution of U.

`Choose` U_0;

`Solve for` U_{k+1} `the matrix equation`

$$-\lambda\widehat{A}(U_k)U_{k+1}\widehat{E}(U_k) + (K_1^TK_1)U_{k+1}(K_2^TK_2) = K_1^TZK_2;$$

`Return` U_{k+1};

At each iteration k, we have to solve a generalized Sylvester matrix equation:

$$AU_{k+1}D - C_kU_{k+1}B_k = E. \tag{20}$$

Where

$$A = K_1^TK_1,\ \ B_k = \widehat{E}(U_k),\ \ C_k = \lambda\widehat{A}(U_k),\ \ D = K_2^TK_2 \ \text{ and } \ E = K_1^TZK_2.$$

To solve the matrix equation (20) we use the global-generalized minimal residual (global-GMRES) method introduced by Jbilou, Mesaoudi and Sadok in [12], a good set of references may be found in [3], where A. Bouhamidi and K. Jbilou used the global-GMRES method which is an orthogonal projection method onto a matrix Krylov subspace. The following theorem (see [17, 3]) gives existence and uniqueness conditions of solution to the matrix equation (20).

Theorem 3.2. *Let $A-\lambda C$ and $B-\lambda D$ be two regular matrix pencils of order $n\times n$ and $p\times p$, respectively. Let $\mathcal{L}(A,C)$ and $\mathcal{L}(B,D)$ denote the sets of generalized eigenvalues for the pairs (A,C) and (B,D) respectively. Then, if*

$$\mathcal{L}(A,C) \cap \mathcal{L}(B,D) = \emptyset,$$

the linear matrix equation (20) *admits a unique solution.*

Let $\mathcal{M}$ be the linear operator defined from $\mathbb{R}^{n\times n}$ onto $\mathbb{R}^{n\times n}$ as

$$\mathcal{M}(U) = AUD - C_kUB_k.$$

Therefore, the problem (20) can be written as

$$\mathcal{M}(U) = E.$$

Let V be any $n \times p$ matrix and consider the matrix Krylov subspace $\mathcal{K}_k(\mathcal{M}, V) = span\{V, \mathcal{M}(V), \ldots, \mathcal{M}^{k-1}(V)\}$ associated to the pair $(\mathcal{M}, V)$. We note that $Z \in \mathcal{K}_k(\mathcal{M}, V)$ means that

$$Z = \sum_{i=0}^{k-1} \alpha_i \mathcal{M}^{i-1}V, \quad \alpha_i \in \mathbb{R}, \quad \text{for} \quad i = 0, \cdots, k-1,$$

where $\mathcal{M}^i(V)$ is defined recursively as $\mathcal{M}^i(V) = \mathcal{M}(\mathcal{M}^{i-1}(V))$. Remark also that the matrix Krylov subspace $\mathcal{K}_k(\mathcal{M}, V)$ is a subspace of $\mathbb{R}^{n\times p}$.

The modified global Arnoldi algorithm constructs an F-orthonormal basis $V_1, V_2, \ldots, V_k$ of the matrix Krylov subspace $\mathcal{K}_k(\mathcal{M}, V)$, i.e.,

$$tr(V_i^TV_j) = 0, \quad \text{for} \quad i \neq j, \;\; i, j = 1, \cdots, k, \quad \text{and} \quad tr(V_i^TV_i) = 1,$$

where $tr(V)$ denotes the trace of the matrix V. The algorithm is described as follows

Algorithm 2: Modified Global Arnoldi algorithm

```
Set V_1 = V/‖V‖_F;
for j = 1, ..., k do
    Ṽ = M(V_j);
    for i = 1, ..., j do
        h_{i,j} = ⟨V_i, Ṽ⟩_F;
        Ṽ = Ṽ − h_{i,j} V_i;
    end
    h_{j+1,j} = ‖Ṽ‖_F;
    V_{j+1} = Ṽ/h_{j+1,j};
end
```

Hereafter, we need some notations to define the global-GMRES algorithm given in [3]. For each k, let $\mathcal{V}_k$ denotes the $n \times kp$ matrix: $\mathcal{V}_k = [V_1, V_2, \ldots, V_k]$. $\tilde{H}_k$ is the $(k+1) \times k$ upper Hessenberg matrix whose nonzero entries $h_{i,j}$ are defined by Algorithm 2.

Starting from an initial guess $U_0 \in \mathbb{R}^{n\times p}$ and the corresponding residual $R_0 = E - \mathcal{M}(U_0)$, the approximate solution U_k is defined as follows

$$U_k = U_0 + F_k \;\text{ with }\; F_k \in \mathcal{K}_k(\mathcal{M}, R_0)$$

and

$$R_k = E - \mathcal{M}(U_k) \;\perp_F\; \mathcal{K}_k(\mathcal{M}, \mathcal{M}(R_0))$$

where the symbol $\perp_F$ denotes the orthogonality with respect to the scalar product $\langle ., . \rangle_F$. Observe that the residual $R_k = E - \mathcal{M}(U_k)$ is obtained by projecting orthogonaly R_0 onto the matrix Krylov subspace

$$\mathcal{K}_k(\mathcal{M}, \mathcal{M}(R_0)) = span\{\mathcal{M}(R_0), \mathcal{M}^2(R_0), \ldots, \mathcal{M}^k(R_0)\}.$$

This immediately shows that U_k can be obtained as the solution of the minimization problem

$$\min_{U - U_0 \in \mathcal{K}_k(\mathcal{M}, R_0)} \|E - \mathcal{M}(U)\|_F.$$

For the following theorem, the interest readers may consult [3] for a proof and further details.

Theorem 3.3. *At step k, the approximate solution U_k produced by the global-GMRES method is given by $U_k = U_0 + \mathcal{V}_k(y_k \otimes I_p)$, where y_k is the solution of the following small least-squares problem*

$$\min_{y \in \mathbb{R}^k} \left\| \|R_0\|_F \, e_1 - \tilde{H}_k y \right\|_2,$$

e_1 being the first unit vector of $\mathbb{R}^{k+1}$.

Consequently, The global-GMRES(m) algorithm for solving the generalized Sylvester matrix equation (20) is summarized as follows

Algorithm 3: Global-GMRES(m) for the generalized Sylvester matrix equation

Input: U_0, a tolerance ε and set $k = 0$, $kmax = 100$;
Compute: $R_0 = E - AU_0D + CU_0B$, $\beta = \|R_0\|_F$ and $V_1 = R_0/\beta$;
while $k \leqslant kmax$ **do**
  Construct the F-orthonormal basis $V_1, V_2, \ldots, V_m$;
    by applying algorithm 2 to the pair $(\mathcal{M}, V_1)$;
  Determine y_m as the solution of the least squares problem:;
    $$\min_{y \in \mathbb{R}^m} \left\| \|R_0\|_F \quad e_1 - \tilde{H}_m y \right\|_2;$$
  Compute: $U_m = U_0 + \mathcal{V}_m(y_m \otimes I_p)$ and $R_m = E - \mathcal{M}(U_m)$;
  if $||R_m||_F < \varepsilon$ **then**
    Stop;
  else
    $U_0 = U_m$, $R_0 = R_m$, $\beta = \|R_0\|_F$, $V_1 = R_0/\beta$, $k = k + 1$;
  end
end
Result: Return: U_0;

4. Parameter selection method for the total variation problem

The choice of the regularization parameter λ is important for getting good results in solving problem of Total Variation (4). Various techniques are available for determining such a value (see [9]). In what follows, we describe the L-curve criterion.

Since the size of the reduced Total Variation regularization problem (4) is small, the L-curve criterion [8, 11] is a popular method for determining a suitable value of λ. This method suggests to plot the curve

$$\lambda \longrightarrow \left(\|\big(K_2 \otimes K_1\big)\mathbf{u}_\lambda - \mathbf{z}\|_2; \|\big(\widehat{E} \otimes \widehat{A}\big)\mathbf{u}_\lambda\|_2 \right). \tag{21}$$

In this respect, we should solve the minimization problem

$$\min_{u \in \mathbf{R}^{n\times n}} \left\{ \|\big(K_2 \otimes K_1\big)\mathbf{u} - \mathbf{z}\|^2 + \lambda \|\big(\widehat{E} \otimes \widehat{A}\big)\mathbf{u}\|^2 \right\}. \tag{22}$$

In the process of doing so, we first transform the minimization problem (22) to the standard form

$$\min_{y \in \mathbf{R}^{n\times n}} \left\{ \|\big(\overline{A}_2 \otimes \overline{A}_1\big)\mathbf{y} - \mathbf{z}\|^2 + \lambda \|\mathbf{y}\|^2 \right\}. \tag{23}$$

Let $\widehat{A} = Q_1 R_1$ and $\widehat{E} = Q_2 R_2$ be the QR factorization of $\widehat{A}$ and $\widehat{E}$, respectively, where Q_1, Q_2 are orthogonal matrices and R_1, R_2 are upper triangular matrices. Using the KPA properties (11) and the fact that Q_1 and Q_2 are orthogonal matrices, we can write

$$\begin{aligned} \|\big(\widehat{E} \otimes \widehat{A}\big)\mathbf{u}\| &= \|(Q_2 R_2 \otimes Q_1 R_1)\mathbf{u}\|, \\ &= \|(Q_2 \otimes Q_1)(R_2 \otimes R_1)\mathbf{u}\|, \\ &= \|(R_2 \otimes R_1)\mathbf{u}\|. \end{aligned}$$

By setting

$$\mathbf{y} = (R_2 \otimes R_1)\mathbf{u},$$

we have

$$\mathbf{u} = (R_2 \otimes R_1)^{-1}\mathbf{y},$$

and therefore,

$$\begin{aligned} \|\big(K_2 \otimes K_1\big)\mathbf{u} - \mathbf{z}\| &= \|\big(K_2 \otimes K_1\big)(R_2^{-1} \otimes R_1^{-1})\mathbf{y} - \mathbf{z}\|, \\ &= \|\big(\big(K_2 R_2^{-1}\big) \otimes \big(K_1 R_1^{-1}\big)\big)\mathbf{y} - \mathbf{z}\|. \end{aligned}$$

And then, the required standard form (23)

$$\min_{y \in \mathbf{R}^{n\times n}} \left\{ \|\big(\overline{A}_2 \otimes \overline{A}_1\big)\mathbf{y} - \mathbf{z}\|^2 + \lambda \|\mathbf{y}\|^2 \right\},$$

where

$$\overline{A}_2 = K_2 R_2^{-1} \text{ and } \overline{A}_1 = K_1 R_1^{-1}.$$

The corresponding normal equation for this problem takes the form

$$\left(\big(\overline{A}_2 \otimes \overline{A}_1\big)^T \big(\overline{A}_2 \otimes \overline{A}_1\big) + \lambda I \right) \mathbf{y} = \big(\overline{A}_2 \otimes \overline{A}_1\big)^T \mathbf{z}.$$

Thus, the L-curve in terms of quantities of the minimization (23) becomes the map

$$\lambda \longrightarrow \left(\|\big(\overline{A}_2 \otimes \overline{A}_1\big)\mathbf{y}_\lambda - \mathbf{z}\|_2; \|\mathbf{y}_\lambda\|_2 \right).$$

Using the singular value decomposition (SVD) [7] for $\overline{A}_1$ and $\overline{A}_2$, we get

$$\overline{A}_1 = U_1 S_1 V_1^T \text{ and } \overline{A}_2 = U_2 S_2 V_2^T,$$

where $U_1, U_2 \in \mathbb{R}^{n\times n}$ and $V_1, V_2 \in \mathbb{R}^{n\times n}$ are orthogonal and unitary matrices. The matrices

$$S_1 = diag[s_{1_1}, s_{1_2}, ..., s_{1_n}] \text{ and } S_2 = diag[s_{2_1}, s_{2_2}, ..., s_{2_n}],$$

are rectangular diagonal matrices with non-negative real numbers on the diagonal containing singular values arranged from the largest to the smallest.

Let us set

$$S = S_2 \otimes S_1 \text{ and } \hat{Z} = U_1^T Z U_2,$$

pursuing [4] we get

$$\|(\overline{A}_2 \otimes \overline{A}_1)\mathbf{y}_\lambda - \mathbf{z}\|_2 = \sum_{i=1}^{k}\left(\frac{\lambda z_i}{s_i^2+\lambda}\right)^2 + \sum_{i=k+1}^{n} z_i^2, \tag{24}$$

and

$$\|\mathbf{y}_\lambda\|_2 = \sum_{i=1}^{k}\left(\frac{s_i z_i}{s_i^2+\lambda}\right)^2, \tag{25}$$

where $s = diag(S)$ contains the singular values, $z = vec(\hat{Z})$ and k is the final index of the nonzero singular values. The best regularization parameter λ should lie on the corner of the L-curve. The L-curve method chooses the regularization parameter corresponding to the point on the curve with maximum curvature. So, we use the formulation (24)-(25) to determine, for each iteration k, the points

$$p_{i,k} = \left\{\|(\overline{A}_2 \otimes \overline{A}_1)\mathbf{y}_{\lambda_i} - \mathbf{z}\|_2; \|\mathbf{y}_{\lambda_i}\|_2\right\}$$

on the curve. To compute the approximation of these point, a suitable algorithm has been defined in [18, Rodriguez and Theis].

We now give the Kronecker Product Approximation Total Variation Algorithm (KPA-TV) to restore the blurred and noisy image

Algorithm 4: KPA-TV algorithm.

```
Input: matrices of same size  U, U_0;
    and set k = 0 and itmax = 100;
while k ≤ itmax do
    Construct  Â(U_k) and Ê(U_k);
    Compute the optimal λ by applying  L-curve ;
    Compute:  X = GMRES(U_k) by applying algorithm 3;
    if ||X − U_k|| / ||X|| < ε then
        U_{k+1} = X,   k = k + 1
    end
end
```

Result: Return U_{k+1}.

5. Numerical results

In this section, we give some numerical tests to illustrate the effectiveness of the KPA-TV algorithm 4 to restore blurred and noisy images by using the total variation regularization and the Kronecker product approximation. All computation were carried out using MATLAB 16 on an Intel(R) core i3 CPU M370 with 4GB of RAM, the computation were carried out with approximately 16 decimal digits of accuracy. The exact gray scale image is denoted by U in all examples, and it is represented by an array of 256×256 pixels with value in $[0, 255]$. To determine the effectiveness of our approach, we evaluate the relative error,

$$e_k = \frac{\|U - U_k\|_F}{\|U\|_F},$$

of the computation approximation solution U_k and the peak-signal-to-noise ratio (PSNR),

$$PSNR(U, U_k) = 20 \log_{10} \left(\frac{255 * 256}{\|U - U_k\|_F} \right).$$

In the forthcoming examples, we looking for the optimal regularization parameter λ_{Opt} in the interval $[0, 1]$, and we give some maps of the L-curve and the optimal λ value.

5.1. Example 1. We let the exact (Blur - and noise - free) image be the " MUSICIAN" from Matlab, it's presented by an array 256×256 pixels and is shown on the left side of the figure 1. In this example, the blurring matrix K of size $256^2 \times 256^2$ is determined by the Point Spread Function (PSF) [10], which defines how each pixel is blurred, and the boundary condition, which specify our assumptions on the scene just outside our image. We assume that the horizontal and vertical components of the PSF can be separated. Then K can be expressed as a Kronecker product $K = K_2 \otimes K_1$; see [10]. The blurred and noise-free image $\hat{Z}$ is given by $K_1 U K_2^T$. The contaminated image by the blur K and a Gaussian white noise with 10^{-2} noise level is represented in the middle image of the figure 1.

Also when K can not be written as a tensor product of two matrices, it may be possible to approximate K quite accurately by such a product. The factors can be determined by solving the minimization problem

$$[\hat{K_1}, \hat{K_2}] = arg \min_{K_1, K_2} \|K - K_2 \otimes K_1\|_F$$

see [10].

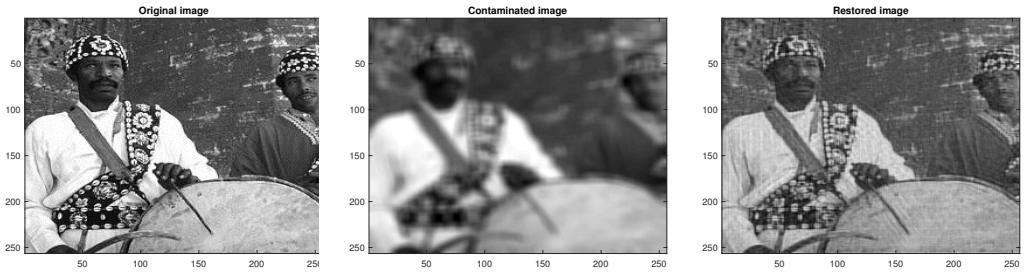

FIGURE 1. True image (left), the noisy image with $SNR = 12.51$ and $PSNR = 17.37$ (center), restored image with $SNR = 17.42$ and $PSNR = 22.27$ (right)

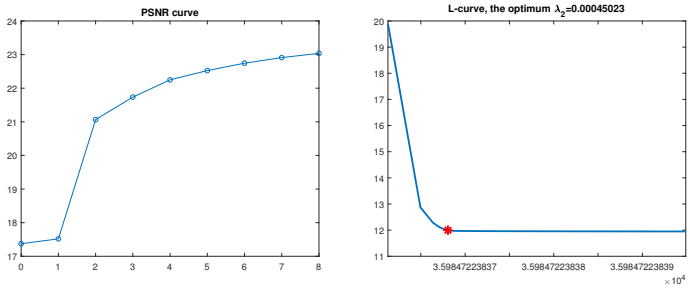

FIGURE 2. The PSNR curve (left), the L-curve with the optimal value of λ located in the red * (right)

Iteration	Relative Error	PSNR	SNR	$\lambda_{Opt.}$
0	6.3706e-01	17.37	12.51	1.0000e+00
1	4.7880e-01	17.6	12.72	1.0000e-02
2	1.5207e-01	21.1	16.23	4.5023e-04
3	9.0477e-02	21.89	17.02	1.0005e-05
4	8.5808e-02	21.9	17.03	1.0000e-02
5	5.6892e-02	22.27	17.42	1.6008e-04

TABLE 1. Relative errors, PSNR values and optimal value of λ

The figure 2 shows two curves, the left one represent the evolution of the PSNR trough the iterations. while the right curve stand for the determination of the regularization parameter λ by L-curve algorithm at iteration $k = 2$, and the vertex determined by the algorithm [18]. The vertex correspond to the regularization parameter $\lambda_2 = 4.5023e^{-4}$. The table 1 displays values of the relative errors and the PSNR values at each iteration k. The PSNR values are seen increase from 17.37 to 22.27, and the relatives errors e_k decrease, which are in accordance with the figures 1 and 2.

5.2. Example 2. In this example, we consider three benchmark images of size 256×256 which are "LENA", "MOON" and "CAMERAMAN" from Matlab, they are shown

in the left of the figure 3. We construct the blurring matrix K by the PSFDEFOCUS as in the first example. The blurred and noisy images shown in the middle of the figure 3 are constracted by the blurring matrix K and a noise vector with normally distributed random entries with zero mean and with variance chosen such that $\nu = 0.001$. The restored images are presented in the right of the figure 3 and they are obtained by applying the KPA-TV algorithm 4 to the problem (4). The table 2 gives for each image the number of iterations, the relatives errors and the PSNR of the contaminated and the restored images. We present also in the figure 4 the increasing curves of the PSNR values for each images.

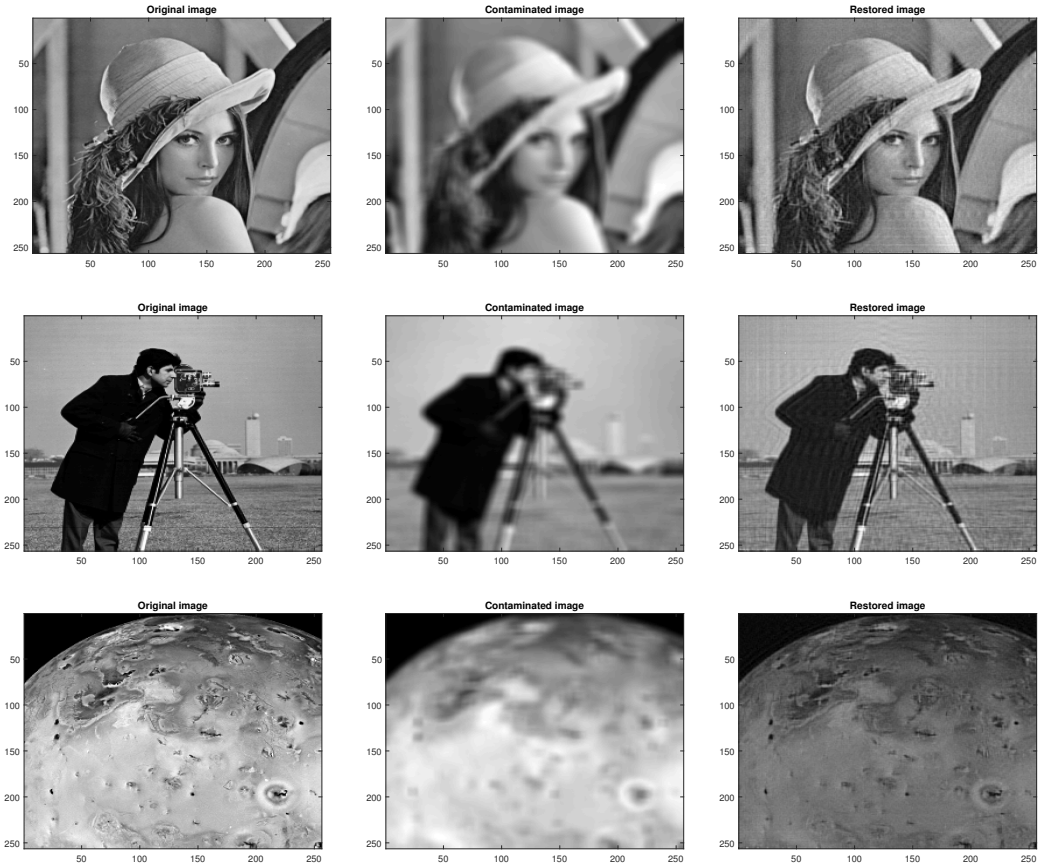

FIGURE 3. True images (left), the noisy images with $\nu = 0.001$ (center) and the restored images (right)

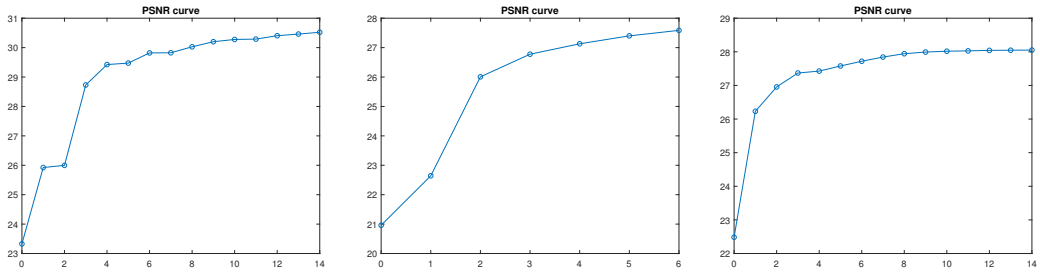

FIGURE 4. The PSNR curves for "LENA" image (left), "CAMERAMAN" image (center), "MOON" image (right)

Image	*Iter*	R_z	R_u	$PSNR(Z)$	$PSNR(U)$
Lena	12	6.3118×10^{-1}	7.3864×10^{-2}	23.33	30.52
Cameraman	6	5.7024×10^{-1}	6.2212×10^{-2}	20.96	27.59
Moon	13	4.1693×10^{-1}	5.8502×10^{-2}	22.48	28.05

TABLE 2. The computational results for some images with $\nu = 0.001$.

6. Conclusion

A new algorithm for image restoration using the total variation regularization is presented. The method is named Kronecker product approximation total variation (KPA-TV) and is based on the Kronecker product approximation by transforming the corresponding Euler equation of the initial problem to a generalized Sylvester linear matrix equation. Then, global Krylov subspace methods are used to solve the obtained linear matrix equation. Computed examples illustrate the effectiveness of the proposed approach.

References

[1] R. Acar and C.R. Vogel, Analysis of bounded variation penalty methods for ill-posed problems, *Inverse problems* **10** (1994), 1217–1229.

[2] G. Buttazzo, H. Attouch, and G. Michaille, *Variational analysis in Sobolev and BV spaces: applications to PDEs and optimization*, MOS-SIAM SERIES ON OPTIMIZATION 17, SIAM, Philadelphia, USA, 2014.

[3] A. Bouhamidi and K. Jbilou, Sylvester Tikhonov-regularization methods in image restoration, *Journal of Computational and Applied Mathematics* **206** (2007), 86–98. DOI: 10.1016/j.cam.2006.05.028

[4] A. Bouhamidi, K. Jbilou, L. Reichel, and H. Sadok, A generalized global Arnoldi method for ill-posed matrix equations, *Journal of Computational and Applied Mathematics* **236** (2012), 2078–2089. DOI: 10.1016/j.cam.2011.09.031

[5] A. Chambolle, An algorithm for total variation minimization and applications, *Journal of Mathematical imaging and vision* **20** (2004), 89–97. DOI: 10.1023/B:JMIV.0000011325.36760.1e

[6] Q. Chang and I.-L. Chern, Acceleration methods for total variation-based image denoising, *SIAM Journal on Scientific Computing* **25** (2003), 982–994. DOI: 10.1137/S106482750241534X

[7] G.H. Golub and C.F. Van Loan, *Matrix computations*, **3**, JHU Press, 2013.
[8] P.C. Hansen, Analysis of discrete ill-posed problems by means of the l-curve, *SIAM review* **34** (1992), 561–580.
[9] P. C. Hansen, *Rank-deficient and Discrete Ill-posed Problems: Numerical Aspects of Linear Inversion*, SIAM, Philadelphia, USA, 2005.
[10] P.C. Hansen, J.G. Nagy, and D.P. O'leary, *Deblurring images: matrices, spectra, and filtering*, SIAM, 2006.
[11] P.C. Hansen and D.P. O'Leary, The use of the l-curve in the regularization of discrete ill-posed problems, *SIAM Journal on Scientific Computing* **14** (1993), 1487–1503.
[12] K. Jbilou, A. Messaoudi, and H. Sadok, Global fom and gmres algorithms for matrix equations, *Applied Numerical Mathematics* **31** (1999), 49–63.
[13] P. Lancaster and L. Rodman, *Algebraic Riccati equations*, Clarendon press, 1995.
[14] L. Liu, Z.F. Pang, and Y. Duan, Retinex based on exponent-type total variation scheme, *Inverse Problems and Imaging* **12(5)** (2018), 1199–1217. DOI: 10.3934/ipi.2018050
[15] C.V. Loan and N. Pitsianis, Approximation with Kronecker products. In (M.S. Moonen, G.H. Golub, B.L.R. Moor, eds.) *Linear algebra for large scale and real-time applications*, Springer, Dordrecht, (1993), 293–314.
[16] L.I. Rudin, S. Osher, and E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D: Nonlinear Phenomena* **60** (1992), 259–268. DOI: 10.1016/0167-2789(92)90242-F
[17] G.W. Stewart, Error and perturbation bounds for subspaces associated with certain eigenvalue problems, *SIAM review* **15** (1973), 727–764.
[18] D. Theis and G. Rodriguez, An algorithm for estimating the optimal regularization parameter by the l-curve, *Rendiconti di matematica e delle sue applicazioni* **25** (2005), 69–84.

(A. H. Bentbib) Faculty of sciences and technics (LAMAI laboratory), Cadi Ayyad University, Marrakesh, Morocco. ORCID ID 0000-0002-4578-4377
E-mail address: a.bentbib@uca.ac.ma

(K. Kreit) Faculty of sciences and technics (LAMAI laboratory), Cadi Ayyad University, Marrakesh, Morocco. ORCID ID 0000-0002-8046-1610
E-mail address: k.kreit@uca.ma

(A. Bouhamidi) University of Littoral, L.M.P.A, 50 rue F Buisson BP699, F-62228 Calais-Cedex, France. ORCID ID 0000-0002-5675-7251
E-mail address: abderrahman.bouhamidi@lmpa.univ-littoral.fr