Tomographic Convex Time-Frequency Analysis

Rose F. Sfeir, Charbel H. Julien

Lebanese University Faculty of Sciences 2, Department of Computer science & Statistics, Fanar, Lebanon
E-mail: rosesfeir@gmail.com, charbeljulien@hotmail.com

Abstract—In this paper we aim to solve a problem of image reconstruction in tomography. In medical imaging, patients suffer from taking high dose of radioactive drug in order to get a well-qualified image. Our goal is to reduce this dose of radioactive drug given to the patients in PET scan and to get a well-qualified image. We use to modeling this problem using a convex function to minimize. In tomography, real problem requires a positive constraint and may get a blurred image due to poisson noise. Then, in order to get back a non blurred image of human body, we add to this function a wavelet regularization which is a non differentiable function. We introduce specific algorithms to get the minimum of the global function obtained. After presenting the classic algorithms with their conditions to solve the problem we find that Chambolle Pock's algorithm requires less properties than these algorithms and gives good results. Then, we propose its computation method with the proof.

Index Terms—Tomography deblurring, convex optimization, wavelet regularization, linear inverse problem.

I. INTRODUCTION

In this paper, we propose a method to compute Chambolle Pock's algorithm. This algorithm is efficient to solve a deblurring problem in tomography.

We aim to solve a problem of image restoration dealing with Positron emission tomography (PET). Our goal is to reduce the dose of radioactive drug given to the patient in PET scan and to get a non blurred image. With high dimensions, we want to find the representative components of the image. So, we take an image with some known, or estimated degradation and we restore it to its original appearance.

In order to reduce the noise, we have to minimize a functional in our case dealing with Poisson noise. Then, we have to solve a linear inverse problem.

Nesterov proposes the gradient method as cited in [1], Combettes and Wajs propose a proximal forward-backward splitting algorithm as cited in [2].

Beck and Teboulle, propose a fast iterative thresholding algorithm as cited in [3]. Daubechies et al propose an interative thresholding algorithm with a sparsity constraint as cited in [4]. And Chambolle Pock propose a first-order primal-dual algorithm as cited in [5]. We show that Chambolle Pock's algorithm is the most suitable for resolving our problem after comparing the algorithms cited above. Also, we give an explicit form of

its variables. We present four algorithms related to Chambolle Pock's algorithm:

Dupé et al propose a fast iterative forward backward splitting algorithm as cited in [6]. Moreau describes the duality in a Hilbert Space in [7]. Figueiredo and Nowak propose an expectation –maximization algorithm for Image Restoration based on a penalized likelihood formulated in the wavelet domain in [8]. Cohen et al propose numerical methods for the treatment of inverse problems based on adaptive wavelet Galerkin dicretization in [9].

We describe five algorithms proposed to get a well-qualified image, compared to the PSNR value: Nain et al propose the Peer Group Average algorithm as cited in [10]. Jiang et al propose the Stationary wavelet transform algorithm as cited in [11]. Mahmoud et al propose the Log Gabor filter as cited in [12]. Raj and Venkareswall propose the double density dual tree comp as cited in [13] and they propose the Split Bregman, total variation as cited in [14]. Zanella et al propose the scaled projected gradient method as cited in [15]. We explain the PET technique and formulate the problem.

In medical imaging, (PET) is a gamma imaging technique that uses radio tracers that emit positrons, the antimatter counterparts of electrons. In PET the gamma rays used for imaging are produced when a positron meets an electron inside the patient's body, an encounter that annihilates both electron and positron and produces two gamma rays traveling in opposite directions. By mapping gamma rays that arrive at the same time, the PET system is able to produce an image with high spatial resolution.

Our problem can be formulated as

$$Rf = p \qquad (1)$$

Where: $R \in \mathfrak{M}\left(\mathbb{R}^n, \mathbb{R}^m\right)$ is the projection matrix $R_{i,j}$ represents the weight of pixel j in the ray of projection i. $f \in \mathbb{R}^n$ the image to reconstruct. And $p \in \mathbb{R}^m$ the whole projections obtained in several directions or the sinogram, a radiograph to visualize any abnormal opening in the body, following the injection of contrast media into the opening.

The remainder of the paper is organized as follows. In section Ⅱ, a review of the necessary related work required to effectively implement our algorithm is presented for convex problems. The method and how we will compute our proposed algorithm is described in

Section Ⅲ. After that, application of the proposed algorithm and results is discussed in section Ⅳ, and we draw our conclusion in the last section.
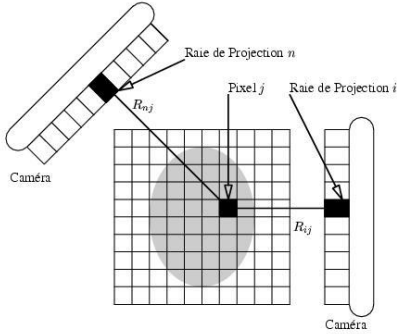

FIG. 2.6 – Construction de la matrice de projection R.
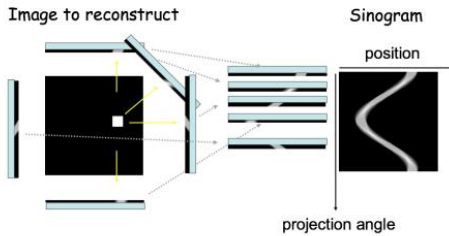
Fig.1. Construction of the projection matrix R.

Fig.2. Sinogram obtained after the projections over the image to reconstruct

II. RELATED WORK

A. Gradient Method

The gradient method is to minimize a convex differentiable function f. This method consists on backtracking lines search. And each line search minimizes $f\left(x - t\nabla f(x)\right)$ over t. Every iteration is inexpensive and it doesn't require second derivatives. Newton's method also minimize a convex differentiable function, uses the first few terms of the Taylor series of a function $f(x)$ in the vicinity of a suspected root.

The conjugate gradient method can be used to solve unconstrained optimization problems like the gradient method. It is a descent method.

The subgradient method is an algorithm for minimizing a non differentiable convex function. It is not a descent method; the function value can often increase.

$$\text{Solving } \min_{x \in X}\left(F(x) + G(x)\right) = \min_{x \in X} f(x) \quad (2)$$

where $F : X \to \mathbb{R}$ is a L-lipschitz differentiable function and $G : X \to \overline{\mathbb{R}}$ is a convex function non differentiable, requires the proximity operator as defined below. The operator

$$prox_G : \quad X \to X$$
$$x \to \arg\min_{y \in X}\left(G(y) + \frac{1}{2}|x - y|^2\right) \quad (3)$$

is called proximity operator associated to G. As cited in [2] and [3]. If G is differentiable, the subgradient projection method gives the solution. As cited in [1]. This method converges slowly, and the rate of convergence is $\frac{1}{\sqrt{k}}$

The projected subgradient method is similar to the subgradient method, we consider a convex set.

B. Extension of the Gradient Method

We will discuss two methods:ISTA, Iterative Shrinkage-Thresholding Algorithm and FISTA, Fast Iterative Shrinkage-Thresholding Algorithm, as cited in [3] and [4]. These methods are the extension of the gradient method.

ISTA consists on constructing on each iteration a regularization of linear differentiable functions. It means making some descent of gradient on F differentiable and computing the proximity operator on the non differentiable G.

FISTA is an accelerated version of ISTA. Chambolle Pock [5] show that solving the following problem $\min_{x \in X} F(Kx) + G(x)$ which is the sum of two convex functions, lower semi-continuous non differentiable with X, Y two finite spaces holding the scalar product. $K : X \to Y$ linear continuous operator with the inducted norm

$$\|K\| = \max_{\{x \in X \text{ with } \|x\| \leq 1\}} \|Kx\| \quad (4)$$

requires this convergent algorithm while $\tau\sigma\|K^2\| < 1$.

Initialization Choose $\tau\sigma > 0$, $\theta \in [0,1]$ $\left(x^0, y^0\right) \in X \times Y$ and $\overline{x}^0 = x^0$

$$\begin{cases} y^{n+1} = prox_{\sigma F^*}\left(y^n + \sigma K\overline{x}^n\right) \\ x^{n+1} = prox_{\tau G}\left(x^n - \tau K^* y^{n+1}\right) \\ \overline{x}^{n+1} = x^{n+1} + \theta\left(x^{n+1} - x^n\right) \end{cases} \quad (5)$$

While taking

$$\theta = 0 \quad (6)$$

in that algorithm we obtain the speed of convergence approximately of $(1/N)$. We can improve the rate of convergence to $(1/N^2)$. We summarize the algorithms their criteria and their rate of convergence on the table below.

Table1. Comparative table of iterative methods

Method	Criteria	Rate of convergence
Gradient method	f convex, L-Lipschitz continuously differentiable	$c(1-q)^k$
Newton method	f convex, L-Lipschitz continuously differentiable	$c(1-q)^{2k}$
Conjugate gradient method	f convex, L-Lipschitz differentiable	Not better than the gradient method
Subgradient method	f convex non differentiable	Slower than the Newton method
ISTA	f, g convex functions f L-lipschitz differentiable g possibly non smooth	$\frac{1}{k}$
MTWIST Monotone version of TWIST(Two step ISTA for extremely ill-conditioned problems)	f, g convex functions f L-lipschitz differentiable g possibly non smooth	$\frac{1}{k} \le r(k) \le \frac{1}{k^2}$
FISTA	f, g convex functions f L-lipschitz differentiable g possibly non smooth	$\frac{1}{k^2}$
Projected subgradient method	f convex non differentiable, $x \in C$ a convex set	Not better than $\frac{1}{\sqrt{k}}$
Chambolle Pock's algorithm	f, g lower semi-continuous conjugate convex functions	Can be accelerated to $\frac{1}{k^2}$ if f^* or g is uniformly convex.

In[6] they present a deconvolution algorithm, fast iterative thresholding algorithm, taking into consideration the Poisson noise. For denoising the images, they use the Anscombe transform. And finally, they select the regularizing parameter by using the generalized cross validation based model selection. They recommend using sparse-domain regularization in many deconvolution applications with Poisson noise.

In [7], Moreau propose dual functions and conjugate functions in the Hilbert space. He talks about proximity operator.

In [8], they propose a wavelet-based criterion for Image Deconvolution. This algorithm MPLE/ MAP criterion maximum penalized likelihood estimator, maximum a posteriori alternates between Fourier domain filtering and wavelet domain denoising.

In [9], they propose Galerkin methods for inverse problems based on wavelets. They combine a thresholding algorithm on the data with a Galerkin inversion on a fixed linear space and then they invert by constructing a smaller space adapted to the solution iteratively constructed.

In [10], many mixed noise removal techniques were compared such as Peer Group Averaging (PGA), Vector Median Filter (VMF), Vector Direction Filter (VDF), Fuzzy Peer Group Averaging (PGA), and Fuzzy Vector Median Filter (FVMF). (PGA) is a non-linear filter, whose aim is to average over a peer group rather than the whole window. (FVMF) performs a weighted averaging where the weight of each pixel is computed according to its similarity to the robust vector median. They prove that (PGA) is advisable when the image size is small and if the image size is large it is better to use (FVMF). Also, (PGA) gives best visual quality results while (FVMF) gives the worst visual quality results.

In [11], they develop a novel method to improve the Visual quality of X-ray CR images, consisting of a wavelet-based filter denoising based on SWT:Stationary Wavelet Transform and the wavelet thresholding then, they enhance the image contrast using Gamma Correction and after extract and classificate high frequency components into three feature images and then Image fusion obtained after adjusting the compression ratio of non-diagnosis feature component in the high signal range and calibrate the parameter.

In [12], they show that the Log Gabor Filter, a filter in the transform domain gives good images quality, PSNR values and CPU time, while Speckle Reducing Anisotropic Diffusion (SRAD) a spatial filtering is better than several commonly used filters including Gaussian, Gabor, Lee, Frost, Kuan, Weiner, Median, Visushrink, Sureshink and also the Homorphic Filter, but on the other hand, has a very high CPU compared to the other filters.

In [13], they show that in order to remove the noise of Poisson and Rician from medical images, the double density dual tree complex wavelet transform based on discrete wavelet transform is performing based on discrete wavelet transform is performing well than the other transforms.

In [14], they show that total variation based regularization split Bregman method from optimization method is well-suited to image restoration in medical images and it is better than the traditional spatial domain filtering methods.

In [15], they show that for Poisson noise removal the scaled gradient projection method based on a Bayesian approach is more efficient than the gradient projection methods.

III. METHOD AND THE PROPOSED ALGORITHM

The process of image construction is f_j : isotopes concentration on pixel j, $f_j \geq 0$. R_{ij} : the average number of photons detected on i per unit of isotopes concentrations on pixel j. We note

$$\tilde{p}_i = \sum_{j=1}^{n} R_{ij} f_j \qquad (7)$$

In fact, what we measure on each pixel i , is the realization of a random variable

$$p_i = \wp \left(\sum_{j \in \mathbb{R}^n} R_{ij} f_j \right) \qquad (8)$$

of Poisson parameter $\tilde{p}_i$. One way of finding f_j is to estimate the maximum of likelihood.

$$P\left(p/f \right) = \prod_i \frac{\left(\sum_j R_{ij} f_j \right)^{p_i}}{p_i!} e^{-\sum_j R_{ij} f_j} \qquad (9)$$

All p_i are independent. (And n are random variables). The maximum of likelihood estimate is the value that makes the observed data the most probable. Rather than maximizing this product, we often use the fact, that the logarithm is an increasing function so it will be equivalent to maximize the log likelihood. Maximizing likelihood on f means minimizing the criteria:

$$J(f) = \sum_i \left[\sum_j R_{ij} f_j - p_i log \left(\sum_j R_{ij} f_j \right) \right] \qquad (10)$$

We will minimize it under the constraint: $f \in \mathbb{R}_n^+$. We have to minimize the error. This functional is convex but not differentiable on $\mathbb{R}_n^+$, so if she takes a minimum, this minimum is global. J takes a global minimum, in this case, it is not unique because R is not invertible on prior. It is an ill-posed inverse problem.

A. Explanation of a basic linear inverse problem

A basic linear inverse problem leads us to study a discrete linear system of the form:

$$R x = p + w \qquad (11)$$

where $R \in \mathbb{R}^{m \times n}$ and $p \in \mathbb{R}^m$ are known, w is an unknown noise and x is the " true'' and unknown image to be estimated. In image blurring problems, $p \in \mathbb{R}^m$ represents the blurred image. In these applications, the matrix R describes the blur operator, which in the case of spatially invariant blurs represents a two-dimensional convolution operator. The problem of estimating x from the observed blurred and noisy image p is called an image deblurring problem.

The blur operator is unbounded, we are in ill-posed problems. The generalized inverse operator may be unbounded. So, it has to be replaced by bounded approximants, so that numerically stable solutions can be defined and used as meaningful approximations of the true solution corresponding to the exact data. This is the purpose of regularization.

In image deblurring applications, R is chosen as $R = AW$, where A is the blurring matrix and W contains a wavelet basis. R corresponds to performing inverse wavelet transform. We deal with the $l_{1?}$ norm regularization criterion because most images have a sparse representation in the wavelet domain. In our case, we are using medical data represented by photon images degraded by Poisson noise. Because of this, it is a difficult task to remove the noise. However, if we can remove Poisson noise effectively, we will acquire a fine image from the degraded image produced with a small amount of photons. This may lead to a reduction of the probability of not only medical exposure, but also medical errors. To accomplish the effective reduction of Poisson noise, the techniques based on wavelet transform have been proposed recently. The method gives thresholding in each wavelet domain, maintaining the features of an original signal. As seen in the related work, Chambolle Pock's algorithm requires less properties than the other algorithms.

We will choose to minimize $F(K x) + G(x)$

Where $F(K x) = J(x)$, $K = R$ and G is non differentiable. F and G are lower semi-continuous convex functions.

B. Choice Of The Functionals

$$J(x) = \sum_i \left[\sum_j R_{ij} x_j - p_i log \left(\sum_j R_{ij} x_j \right) \right] \qquad (12)$$

$\mathcal{X}_C$ is the function indicator of a convex set $C = \{ x : x \geq 0 \}$, $\mathcal{X}_C(x) = \begin{cases} 0 & if \quad x \in C \\ \infty & else. \end{cases}$ and $\|x\|_{1,W} = \sum_j \sum_k \left| < x, \psi_{j,k} > \right|$

$$G(x) = \mathcal{X}_C(x) + \lambda \|x\|_{1,W} \qquad (13)$$

We include on G the positive constraint and the wavelet $l_{1?}$ -regularization. G is the functional representing the denoising. The parameter λ is the regularization parameter, which is used to control the degree of regularization.

C. Application On Chambolle Pock's Algorithm

After choosing F and G convex functions as defined above, we compute: $_\sigma F^*$, $prox_{\sigma F^*}(x)$ and $prox_{\tau G}(x)$. The $prox$ have an explicit form. Knowing $p_i > 0$. After computing, we get:

$$prox_{\sigma F^*}(y) = \frac{1}{2}\left[y_i + 1 - \sqrt{(y_i - 1)^2 + 4\sigma p_i} \right] \quad (14)$$

$prox_{\tau G}(x)$ is the projection on a positive set of inverse wavelet transform of a threshold of parameter $\lambda\tau$ of the wavelet transform on x.

$$prox_{\tau G}(x) = \prod_C \left[W^{-1} S_{\lambda\tau} (Wx) \right] \quad (15)$$

C. Computing $prox_{\sigma F^*}(y)$

$$F(y) = \sum_i \left(y_i - p_i \log(y_i) + \mathcal{X}_{\mathbb{R}^*}(y_i) \right) \quad (16)$$

$$x = prox_{\sigma F^*(y)} \quad (17)$$

Then,

$$x_i = \frac{1}{2}\left[y_i + 1 - \sqrt{(y_i - 1)^2 + 4\sigma p_i} \right] \quad (18)$$

In our case, $p_i > 0$.
Proof:

$$f_i(y_i) = y_i - p_i \log(y_i) + \mathcal{X}_{\mathbb{R}^+}(y_i) \quad (19)$$

f_i is lower semi-continuous over $\mathbb{R}$ because $p_i > 0$ Then,

$$F(y) = \sum_i f_i(x_i) \quad (20)$$

is convex lower semi-continuous over $\mathbb{R}^n$.

For computing $prox_{\sigma F^*}(y)$, we will use Moreau's identity see [7].

$$prox_{\sigma F^*}(y) = y - \sigma\left(prox_{\frac{1}{\sigma}F}(Y) \right)\left(\frac{y}{\sigma} \right) \quad (21)$$

Then, we will have to compute

$$\tilde{y} = prox_{\tilde{\sigma}F}(Y) \text{ with } \tilde{\sigma} = \frac{1}{\sigma} \quad (22)$$

because we know the explicit form $F(y)^c$.

$$\begin{aligned}
\tilde{y} &= prox_{\tilde{\sigma}F}(Y) \\
&= \arg\min_y \frac{\|y - Y\|^2}{2\tilde{\sigma}} + F(y) \\
&= \arg\min_y \frac{\|y - Y\|^2}{2\tilde{\sigma}} + \langle y - p\log(x), 1 \rangle + \mathcal{X}_c(x) \\
&= \arg\min_y \sum_i \left[\frac{(y_i - Y_i)^2}{2\tilde{\sigma}} + f_i(y_i) \right] + \mathcal{X}_{\mathbb{R}^+}(y_i) \\
&= \arg\min_{y_i} \left(\frac{(y_i - Y_i)^2}{2\tilde{\sigma}} + f_i(y_i) + \mathcal{X}_{\mathbb{R}^+}(y_i) \right)
\end{aligned} \quad (23)$$

Differentiating $f_i(y)$ with respect to y_i, we get:

$$\frac{\partial f_i(y)}{\partial y_i} = 1 - \frac{p_i}{y_i} \text{ then, we have}$$

$$\frac{(y_i - Y_i)}{\tilde{\sigma}} + 1 - \frac{p_i}{y_i} = 0 \quad (24)$$

$$\frac{y_i^2 - y_i(Y_i - \tilde{\sigma}) - p_i\tilde{\sigma}}{y_i\tilde{\sigma}} = 0$$

By calculating the discriminant of this binomial

$$\Delta = (Y_i - \tilde{\sigma})^2 + 4p_i\tilde{\sigma} \geq 0 \quad (25)$$

delta being positive, admits two solutions:

$$y_{i1} = \frac{(Y_i - \tilde{\sigma}) - \sqrt{(Y_i - \tilde{\sigma})^2 + 4p_i\tilde{\sigma}}}{2} \quad (26)$$

$$y_{i2} = \frac{(Y_i - \tilde{\sigma}) + \sqrt{(Y_i - \tilde{\sigma})^2 + 4p_i\tilde{\sigma}}}{2} \quad (27)$$

We choose the second solution because for this value we get the minimum value of the $prox$. After Moreau's identity, we have:

$$\begin{aligned}
prox_{\sigma F^*}(y) &= y - \sigma\left(prox_{\frac{1}{\sigma}F}(Y) \right)\left(\frac{y}{\sigma} \right) \\
x_i &= y_i - \sigma\left[\frac{1}{2\sigma}\left(y_i - 1 + \sqrt{\frac{(y_i - 1)^2}{\sigma^2} + \frac{4\sigma p_i}{\sigma^2}} \right) \right] \\
x_i &= \frac{1}{2}\left[y_i + - \sqrt{(y_i -)^2 + \sigma p_i} \right]
\end{aligned} \quad (28)$$

And we get the result.

D. Computing $prox_{\tau G}(x)$

Before computing $prox_{\tau G}(x)$, we aim to minimize this following functional which will help us to compute $prox_{\tau G}(x)$.

$$J(\alpha) = \|y - A\alpha\|^2 + \lambda\|\alpha\|_1 \qquad (29)$$

Because,

$$K = \arg\min_{y \in X}\left(G(y) + \frac{1}{2\tau}\|x - y\|_X^2 \right)$$

$$= \arg\min_{y \in X}\left(\mathcal{X}_C(y) + \lambda\|Wy\|_1 + \frac{1}{2\tau}\|x - y\|_X^2 \right) \qquad (30)$$

We notice that this functional is convex being the sum of two convex functionals. In fact, let us compute the differential of $\|y - A\alpha\|^2$.

$$J_1(\alpha) = \|y - A\alpha\|^2 = \langle y - A\alpha, y - A\alpha \rangle \qquad (31)$$

$$\begin{aligned} J_1(\alpha + h) &- J_1(\alpha) \\ &= \langle y - A(\alpha + h), y - A(\alpha + h)\rangle \\ &\quad - \langle y - A\alpha, y - A\alpha \rangle \\ &= \langle y - A\alpha - Ah, y - A\alpha - Ah \rangle \\ &\quad - \langle y - A\alpha, y - A\alpha \rangle \\ &= \langle y - A\alpha, y - A\alpha \rangle - \langle y - A\alpha, Ah \rangle \\ &\quad - \langle Ah, y - A\alpha \rangle + \langle Ah, Ah \rangle - \langle y - A\alpha, y - A\alpha \rangle \\ &= -2\langle h, A^*(y - A\alpha) \rangle + \langle Ah, Ah \rangle \end{aligned} \qquad (32)$$

$$\nabla J_1(\alpha) = -2A^*(y - A\alpha) \qquad (33)$$

$$J_1''(\alpha) = 2AA^* \qquad (34)$$

Then, $J_1(\alpha)$ is convex. For $\|\alpha\|_1$ we have

$$\|\lambda\alpha + (1-\lambda)\alpha\|_1 \le \lambda\|\alpha\|_1 + (1-\lambda)\|\alpha\|_1 \qquad (35)$$

which shows that $\|\alpha_1\|$ is convex. Therefore, $J(\alpha)$ being the sum of two convex functions is convex. We know that J is convex, $0 \in \partial J(x)$ is equivalent to say that x is the minimizer of J. Then, the minimum is reached.

Another writing of:

$$\arg\min_{\alpha} \|y - A\alpha\|^2 + \lambda\|\alpha\|_1 \qquad (36)$$

We take: $y = \sum_k y_k \varphi_k$ and $A\alpha = \sum_k \alpha_k \varphi_k$ Then (36) becomes:

$$\arg\min_{\alpha} \left\| \sum_k y_k \varphi_k - \sum_k \alpha_k \varphi_k \right\|^2 + \lambda\|\alpha\|_1$$

$$= \arg\min_{\alpha} \left\| \sum_k |y_k - \alpha_k| \right\|^2 + \lambda\|\alpha\|_1 \qquad (37)$$

$$= \sum_k prox_{\frac{\lambda}{2}|\cdot|}(y_k)$$

E. Computing $prox_{\frac{\lambda}{2}|\cdot|}(y_k)$

$$J(\alpha_k) = |y_k - \alpha_k|^2 + \lambda|\alpha_k| \qquad (38)$$

We compute the subgradient of J on a point α_k. And we calculate at which condition $0 \in \partial J(\alpha_k)$. Let us consider these cases:

First case: J differentiable on $\alpha_k < 0$. We have

$$\partial J(\alpha_k) = \{\nabla J(\alpha_k)\} \qquad (39)$$

$$\nabla J(\alpha_k) = -2(y_k - \alpha_k) - \lambda$$
$$\nabla J(\alpha_k) = 0 \Leftrightarrow -2y_k - \lambda = -2\alpha_k \qquad (40)$$

Then, $\alpha_k = y_k + \dfrac{\lambda}{2}$

The critical point is:

$$\alpha_k = y_k + \frac{\lambda}{2} < 0 \qquad (41)$$

and this impose the following condition:

$$y_k + \frac{\lambda}{2} < 0$$

Second case: J is differentiable on $\alpha_k > 0$.
We have

$$\partial J(\alpha_k) = \{\nabla J(\alpha_k)\} \qquad (42)$$

we have

$$J(\alpha_k) = |y_k - \alpha_k|^2 + \lambda|\alpha_k|$$
$$\nabla J(\alpha_k) = -2(y_k - \alpha_k) + \lambda$$
$$\nabla J(\alpha_k) = 0 \Leftrightarrow -2y_k + \lambda = -2\alpha_k \qquad (43)$$

$$\alpha_k = y_k - \frac{\lambda}{2}$$

The critical point is

$$\alpha_k = y_k - \frac{\lambda}{2} \qquad (44)$$

and this impose $y_k > \frac{\lambda}{2}$ knowing that $\alpha_k > 0$. In both cases, $J(\alpha_k)$ is differentiable.

Third case: $\alpha_k = 0$

J is not differentiable. Let us calculate the subgradien

$$\text{ton } 0. \quad J(0) = |y_k|^2 \qquad (45)$$

If $p \in \partial J(0)$, we get:

$$
\begin{aligned}
&\forall \beta_k, \ J(\beta_k) - |y_k|^2 \geq \langle p, \beta_k \rangle \\
&|y_k - \beta_k|^2 - |y_k|^2 + \lambda |\beta_k| \geq \langle p, \beta_k \rangle \\
&\beta_k^2 - 2\beta_k y_k + \lambda |\beta_k| \geq \langle p, \beta_k \rangle \\
&\beta_k^2 - \beta_k (2y_k + p) + \lambda |\beta_k| \geq
\end{aligned}
\qquad (46)
$$

First case: $\beta_k > 0$

$$
\begin{aligned}
&\beta_k^2 - \beta_k (2y_k + p - \lambda) \geq 0 \\
&\beta_k [\beta_k - (2y_k + p - \lambda)] \geq 0, \ \forall \beta_k
\end{aligned}
$$

We must have: $2y_k + p - \lambda < 0$
Then,

$$p < \lambda - 2y_k \qquad (47)$$

Second case: $\beta_k < 0$

$$
\begin{aligned}
&\beta_k^2 - \beta_k (2y_k + p + \lambda) \geq 0 \\
&\beta_k [\beta_k - (2y_k + p + \lambda)] \geq 0
\end{aligned}
$$

We must have: $2y_k + p + \lambda > 0$
Then, $p > -\lambda - 2y_k$
Then, we must have:

$$-\lambda - 2y_k < p < \lambda - 2y_k \qquad (48)$$

Conclusion:

$$\partial J(0) = \{a : a = -2y + \lambda s, s \in [-1, 1]\} \qquad (49)$$

Then,

$$
prox_{\frac{\lambda}{2}\|\cdot\|}(y_k) =
\begin{cases}
y_k - \dfrac{\lambda}{2} & \text{if } y_k > \dfrac{\lambda}{2} \\[2mm]
0 & \text{if } |y_k| \leq \dfrac{\lambda}{2} \\[2mm]
y_k + \dfrac{\lambda}{2} & \text{if } y_k \leq -\dfrac{\lambda}{2}
\end{cases}
\qquad (50)
$$

Notations:

Let us note by: $\|y\|_{1,W}$ the norm of y on the wavelets basis. We will design by Wy the wavelet transform y. And we will note by

$$\tilde{G}(x) = \|Wx\|_1 \qquad (51)$$

The thresholding operator S_λ is defined by:

$$
S_\lambda(x) =
\begin{cases}
x - \lambda & \text{if } x \geq \lambda \\
0 & \text{if } |x| \leq \lambda \\
x + \lambda & \text{if } x \leq -\lambda
\end{cases}
\qquad (52)
$$

The previous functional is a thresholding of a parameter $\frac{\lambda}{2}$ on y_k. We aim to compute

$$K = prox_{\tau G}(x). \qquad (53)$$

We have:

$$
\begin{aligned}
G(y) &= \mathcal{X}_C(y) + \lambda \|y\|_{1,W} \\
&= \mathcal{X}_C(y) + \lambda \sum_{j,k} |\langle y, \psi_{j,k} \rangle| \\
&= \mathcal{X}_C(y) + \lambda \|Wy\|_1
\end{aligned}
\qquad (54)
$$

Then,

$$
\begin{aligned}
K &= \arg\min_{y \in X} \left(G(y) + \frac{1}{2\tau} \|x - y\|_x^2 \right) \\
&= \arg\min_{y \in X} \left(\mathcal{X}_C(y) + \lambda \|Wy\|_1 + \frac{1}{2\tau} \|x - y\|_x^2 \right) \\
&= \arg\min_{y \in X} \left(\mathcal{X}_C(y) + \lambda \|Wy\|_1 + \frac{1}{2\tau} \|Wx - Wy\|_x^2 \right) \\
&= \arg\min_{z \in X} \left(\mathcal{X}_C(W^{-1}z) + \lambda \|z\|_1 + \frac{1}{2\tau} \|Wx - z\|_x^2 \right) \\
&= \prod_C [W^{-1} S_{\lambda\tau}(Wx)]
\end{aligned}
\qquad (55)
$$

IV. APPLICATION ON PET AND RESULTS

To verify the performance of the proposed algorithm, we experiment with medical images. The first three images represent the simulated data $\left(f, p \text{ and } \tilde{p}\right)$. In the fourth and fifth images, we show results of Chambolle Pock's algorithm after 50 iterations with $\sigma = \tau = 0.05$, $\lambda = 0.5$ and $\theta = 0$.

The fourth image represents the constructed image x^{n+1} in Chambolle Pock's algorithm. And the fifth image represents the sinogram obtained y^{n+1} in Chambolle Pock's algorithm.

In the section method, by computing $prox_{\sigma F^*}\left(y\right)$ and $prox_{\tau G}\left(x\right)$, we removed the Poisson noise on the blurred projection operator and we get the unblurred image. Also, we get an explicit form of the parameters of Chambolle Pock's algorithm easy to implement in matlab.
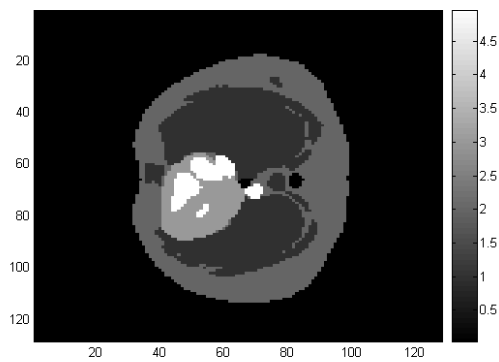
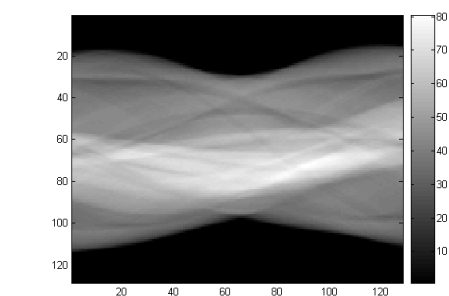Fig.3. Image f to reconstruct (unknown of the algorithm).

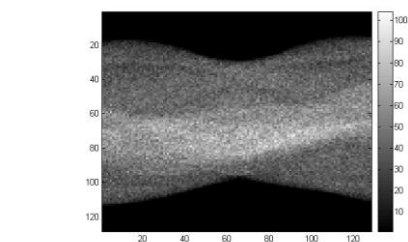Fig.4. Image of the non blurred sinogram $\tilde{p}$.
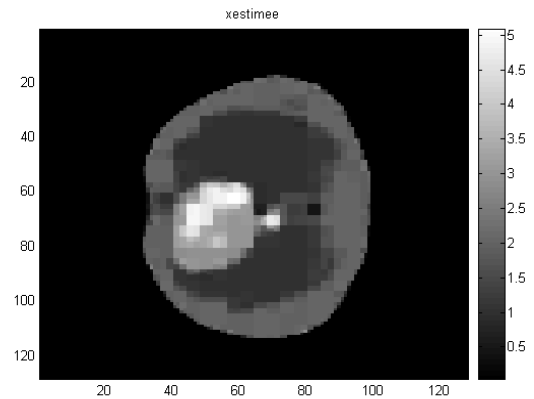

Fig.5. Image of the blurred sinogram p

Fig.6. Image x_{50} estimation of f after 50 iterations of the algorithm.
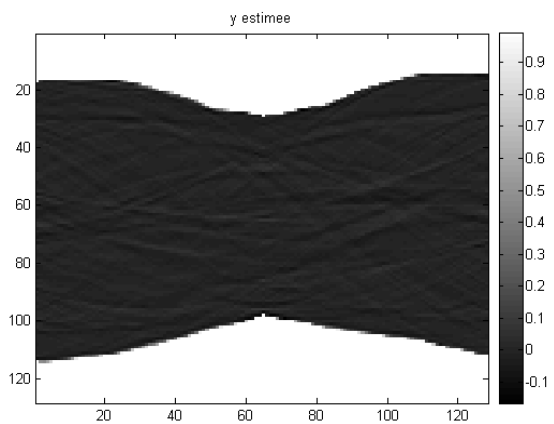

Fig.7. Image y_{50} after 50 iterations of the algorithm.

V. CONCLUSION

This study proposes a method to compute Chambolle Pock's algorithm, used to solve a deblurring problem in tomography, dealing with two convex, lower semi-continuous functions. After introducing the wavelet l_1-regularization on the functional representing the denoising, we compute the proximity operator associated to the functions in Chambolle Pock's algorithm and we get an explicit form of the parameters easy to implement on matlab.

The experiment results show that after computing 50 iterations of Chambolle Pock's algorithm with $\sigma = \tau = 0.05$ we get an unblurred image. This algorithm is efficient in removing Poisson noise from a tomographic image.

For the future work, this research will be extended further by involving other components of other related algorithms and we expect to get better results.

REFERENCES

[1] Y.Nesterov. *Introductory lectures on convex optimization a basic course,* Vol. 87, 2004.

[2] P. L. Combettes and V. R. Wajs, *Signal Recovery by proximal forward-backward splitting: Multiscale Modeling and Simulation*, 4 (4): 1168-1200, November 2005.

[3] A. Beck and M. Teboulle, "A fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse problems", 2009.

[4] I. Daubechies, M. Defrise and C. De Mol, " An iterative thresholding algorithm for linear inverse problems with a sparsity constraint", *Comm. Pure Appl. Math,* 57, pp. 1413-1541, 2008.

[5] A. Chambolle T. Pock, " A First-Order Primal –Dual Algorithm for Convex Problems with Applications to Imaging", december 2010.

[6] F.-X. Dupé, J. Fadili and J.-L. Starck, *A proximal iteration for deconvolving poisson noisy images using sparse representations*, IEEE TIP, Vol.18, 2009.

[7] J.J.Moreau, *Proximité et dualité dans un espace Hilbertien*, 1965.

[8] Mário Figueiredo and Robert D. Nowak, *An EM Algorithm for Wavelet-Based Image Restoration*, IEEE Transactions on Image Processing, Vol.12, no.8., August 2003.

[9] A.Cohen, M. Hoffmann and M. Reib "Adaptative Wavelet Galerkin Methods For Linear Inverse Problems". *SIAM*, 2003.

[10] A. K. Nain, S. Singhania, S. Gupta and Bharat Bhushan, "A comparative Study of mixed noise removal techniques", *International Journal of Signal Processing, Image Processing and Pattern Recognition*, Vol.7, no.1, pp.405-414, 2014.

[11] H. Jiang, Z. Wang, L. Ma, Y. Liu, P. Li. *A novel method to improve the visual quality of X-ray CR Images*. International Journal of Image, Graphics and Signal Processing, Vol.3, no.4, June 2011.

[12] A. A. Mahmoud, S. El Rabaie, T.E. Taha, O. Zahran, F. E. Abd El –Samie, W. Al-Nauimy, "Comparative study of different denoising filters for speckle noise reduction in ultrasonic B-mode". *International Journal of Image, Graphics and Signal Processing, Vol. 5, No.2*, February 2013.

[13] V. Naga Prudhvi Raj and T. Venkateswarlu, M.E. "Denoising of Poisson and Rician Noise from Medical Images using Variance Stability and Multiscale transforms". *International Journal of Computer Applications*, Vol 57/21, November 2012.

[14] V. N. Prudhvi Raj and Dr. T. Venkateswarlu, "Denoising of Medical Images using Total variational Method", Signal & Image Processing: An International Journal, Vol.3, No.2, April 2012.

[15] R. Zanella, P. Boccacci, L. Zanni and M Bertero.Ye. "Efficient gradient projection methods for edge-preserving removal of Poisson noise". *IOP Science, Inverse Problems*. Vol 25, No 4, February 2009.

Authors' Profiles

Rose Sfeir, female, was born in Ajaltoun, Lebanon on, 28/08/1988. She received the Masters degree in probability and statistics from The University of Provence, Aix-Marseille 1, France in 2011. Her research interests are in convex optimization and image processing.

Charbel Julien, male, was born in Jdeidet Marjayoun, South of Lebanon, on 31/08/1978. He received his Ph.D. degree in computer science in 2008 from Lumière Lyon 2's University, France. He is a Lecturer at the Lebanese University, Fanar, Computer Science and Statistics Department. His research interests are in machine learning and image processing.