International Journal of Tomography and Simulation [ISSN 2319-3336];
Year: 2016, Volume: 29, Issue Number: 2;
[Formerly known as “International Journal of Tomography & Statistics”
(ISSN 0972-9976; 0973-7294)];

Binary Tomography Reconstruction with Stochastic Diffusion Based on Level-set and Total Variation Regularization

B.Sixou [1], **L.Wang** [1], **S.Rit** [1], **F.Peyrin** [1,2]

[1]CREATIS, CNRS UMR 5220, Inserm U1044, INSA de Lyon, Universite de Lyon, F-69621, bruno.sixou@insa-lyon.fr

[2]ESRF,6 rue Jules Horowitz, F-38043, Grenoble Cedex France, peyrin@esrf.fr

ABSTRACT

In this work, we address the problem of the reconstruction of binary images from a small number of noisy tomographic projections. Recently, a new stochastic level-set approach was investigated to refine the reconstruction. The main limitation of this method is that it is only changing the boundaries of the reconstructed regions. In this work, we study a new stochastic approach based on Total Variation (TV) regularization with box constraints. The main advantage of this method is that random shape and boundaries variations can be included in a new way and that topology changes can be also added. The methods are tested on two complex bone micro-CT cross-section images for different noise levels and number of projections. While for the higher noise levels, the best reconstructions are obtained with a stochastic diffusion based on the Total Variation regularization, large decreases of the reconstruction errors are obtained when shape and topology noises are used simultaneously.

Keywords: X-ray imaging, discrete tomography, level-set regularization, total variation regularization, inverse problems.

2000 Mathematics Subject Classification: 65J20,65J22, 65K10, 65R10.

1 Introduction

The tomographic reconstruction problem from few projections is a highly ill-posed problem with many applications in medical imaging. It is crucial when the irradiation dose has to be reduced like for *in vivo* bone microstructure imaging. A number of attempts to solve this problem based on sparsity priors or Total Variation (TV) regularization have been proposed in the literature (Yu and Wang, 2010; Sidky and Pan, 2006; Sidky and Pan, 2008; Sidky and Pan, 2012; Defrise, 2011). Binary tomography methods may be proposed to simplify this challenging inverse problem (Herman and Kuba, 2007). These methods are associated to an under-determined linear system of equations with the linear Radon projection operator R and binary constraints:

$$Rf = p \quad f = (f_1, ...f_n) \in \{0,1\}^n \tag{1.1}$$

relating the pixel values $(f_i)_{1\leq i\leq n}$ of the image and the projections p. Only noisy projection values p^n are measured. Such discrete tomography methods are very useful for X-ray tomography of industrial objects (Schillinger, 2005) or bone imaging (VanGompel and Sijbers, 2010; Apostol, Boudousq, Basset, Odet, Yot, Tabary, Dinten, Boller, Kotzki and Peyrin, 2006; Bouxsein, Boyd, Christiansen, Guldberg, Jepsen and Müller, 2010).

Various methods have been proposed to obtain stable solutions for this reconstruction problem (Herman and Kuba, 2007; Batenburg and Sijbers, 2009; Cai and Ma, 2010) based on convex analysis optimization (Capricelli and Combettes, 2007; Schüle, Schnörr, Weber and Hornegger, 2005), Markov random fields (Liao and Herman, 2004) or Belief Propagation (Gouillart, Krzakala, Mezard and Zdeborová, 2013). Rather good reconstruction results are obtained with the TV regularization approach (Gouillart et al., 2013; Sixou, 2013). The TV regularization method was introduced by Rudin et al. (Rudin, Osher and Fatemi, 1992) and it is very useful for many image processing tasks. In a previous work, we have proposed to take into account the binary constraint with a level-set regularization approach (Sixou, 2013; Wang and Peyrin, 2014; Wang and Peyrin, 2015). The two regularization approaches have been compared for several noise levels and number of projections in (Sixou, 2013). The results obtained by the two reconstruction methods are very similar. Yet, for both methods, some errors are still present on the boundaries of the reconstructed regions and some regions are lost for a low number of projections. The reconstructed solution corresponds to a local minimum of the non-convex regularization functional.

It is thus interesting to escape the local minimum achieved with the level-set or TV regularization with global optimization methods. Simulated annealing methods are efficient but they are often very slow (Azencott, 1992; Catoni, 1992; Cot and Catoni, 1998). Algorithms based on stochastic differential equations have been proposed for the global optimization of non-convex functions (Gidas, 1995; Chiang and Sheu, 1987; Chow and Zhou, 1987; Parpas and Rustem, 2009). Stochastic partial differential equations methods have been used for image processing tasks like segmentation (Juan and Postelnicu, 2006). The convergence properties of the stochastic partial differential equation obtained with the sub-differential of the TV regularization semi-norm has also been studied in (Barbu and Rockner, 2009). Recently, stochastic level-set methods have been shown to improve the reconstructed images in binary tomogaphy (Wang and Peyrin, 2015). Stochastic approaches have not been investigated for the TV regularization which often outperforms the level-set regularization for binary tomography problems. Moreover, the level-set approach only modifies the shape of the 0 and 1 regions and their boundaries. With few projection angles, the reconstructed solution is a local minimum of the cost functional and some regions are lost. It is thus necessary to improve further the reconstruction.

The main contribution of this work is to present new stochastic approaches to improve the binary tomography reconstruction obtained with the deterministic TV minimization algorithm and to compare them with the stochastic level-set method investigated recently (Wang and Peyrin, 2014).

The principle of this new method is to use stochastic evolutions of a regularization Lagrangian based on the Alternate Direction of Minimization method (ADMM) approach to improve the re-

construction obtained with the TV regularization (Afonso, Bioucas-Dias and Figueiredo, 2011). The original part of the stochastic TV regularization method compared to the stochastic level-set method is twofold. First, the random change of the boundary is performed in a new way with the gradient of the image or with wavelets. Moreover, random topological changes are included to reveal new regions that can not be detected with shape changes. The numerical results are illustrated on thresholded bone micro-CT cross-sections for various noise levels and numbers of projections and compared with the ones obtained with the level-set stochastic approach (Wang and Peyrin, 2015).
This paper is organized as follows. After the introduction, the deterministic approaches used for the binary tomography problem are summarized. The TV regularization approach for binary tomography is detailed together with the ADMM algorithm, and the level-set regularization method is also briefly presented. The section 3 presents the new stochastic approaches considered and the various types of stochastic noises used for the stochastic partial differential equations in this work. The next section develop a linearization of the equations. The simulation details and the numerical results achieved on noisy bone CT cross-sections for different noise levels and numbers of projections are reported and discussed before the conclusion.

2 Level-set and Total Variation deterministic methods for binary tomography

2.1 The binary tomography problem

The tomographic reconstruction problem in modern CT scanners is 3D and the imaging geometry is often not parallel. Yet, we consider here a simple parallel geometry and the direct operator is the Radon projection operator (Natterer, 1986). The methods investigated in this paper can be generalized to other forward projection operators. Let $L(\phi, s)$ the straight line determined by the polar angle $\phi \in [0, \pi)$, and s the distance of the line from the origin. For an integrable function f, the Radon transform Rf is defined as the line integral:

$$Rf(\phi, s) = \int_{\vec{r} \in L(\phi,s)} f(\vec{r}) dl \tag{2.1}$$

where $\vec{r}$ is the spatial position. The discrete formulation of the binary tomography problem based on this continuous transform is associated to an under-determined linear system of equations with the linear Radon projection operator R and binary constraints (Eq.1.1). The noisy projections p^n and projections without noise p are assumed to be such that $\|p^n - p\| \approx n$, where n is the noise level. In the following, we will introduce successively the level-set regularization method and the Total Variation regularization method.

2.2 Level-set regularization for binary tomography

We assume that the function f to be reconstructed is the characteristic function of a regular set $D_1 \subset D$, $f = \chi_{D_1}$ (Sixou, 2013; Wang and Peyrin, 2015). It can be represented as $f = H(\theta)$, where θ is the level-set function belonging to the first-order Sobolev space $H_1(D)$.

The Heaviside distribution H is defined by:

$$H(\theta(x)) = \begin{cases} 1 & \text{if } \theta(x) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.2}$$

In a variational approach, a regularization functional $E(\theta)$ is minimized:

$$E(\theta) = \frac{\|RH(\theta) - p^n\|_2^2}{2} + F(\theta) \tag{2.3}$$

where F is a regularization term for the level-set function. In this work, we considered a Total Variation-H_1 regularization functional (Egger and Leitao, 2009; DeCezaro, Leitao and Tai, 2009), $F(\theta) = \beta_1 |H(\theta)|_{TV} + \beta_2 \|\theta\|_{H_1}^2$ where $|.|_{TV}$ is the Total Variation semi-norm, $\|.\|_{H_1}$ the Sobolev norm and β_1, β_2 the regularization parameters:

$$\|\theta\|_{H_1} = \|\theta\|_{L_2} + \|\nabla\theta\|_{L_2} \tag{2.4}$$

and

$$J_{TV}(H(\theta)) = \int_D |\nabla H(\theta)| d\mathbf{r} \tag{2.5}$$

The following smooth approximation of the Heaviside distribution has been used in this work: $H_\epsilon(x) = \frac{1+2\epsilon}{2}(erf(x/\epsilon) + 1) - \epsilon$ where ϵ is a real positive constant. The optimal level-set θ functions are found with a first-order optimality condition for the smoothed regularization functionals, $\frac{\delta E_\epsilon(\theta)}{\delta\theta} = 0$. From the current estimate θ_k, the update $\theta_{k+1} = \theta_k + \delta\theta$ is obtained iteratively with a Gauss-Newton method (Sixou, 2013).

For a small number of projections and the higher noise levels, the solution θ corresponds to a local minimum (Sixou, 2013). The reconstruction errors are localized at the boundaries of the binary image. In order to escape from this stationary point, a stochastic global optimization method has been proposed based on the level-set formulation (Wang and Peyrin, 2015).

2.3 The binary tomography problem and the Total Variation regularization

The TV regularization method is one of the most successful techniques for regularization in the field of image processing (Rudin et al., 1992; Acar and Vogel, 1994; Ng, Weiss and Yuan, 2010). In this work, the regularized solution is obtained with the minimization of a functional including a data fidelity term and the regularization term J_{TV}:

$$(P_0) \quad \text{minimize} \quad \frac{\alpha}{2}\|Rf - p^n\|_2^2 + J_{TV}(f) \tag{2.6}$$

where α is a parameter that determines the weight of the data fidelity term. The regularization term is given by $J_{TV}(f) = \int_D |\nabla f(\mathbf{r})| d\mathbf{r}$. The binary constraints lead to a non-convex inverse problem. Convexified models obtained by relaxation of the binary constraint have often been considered for segmentation tasks (Bresson and Osher, 2007; Brown and Bresson, 2011; Chan and Nikolova, 2006). We use here the same type of approach and the function f to be reconstructed is thus allowed to take values continuously from $[0, 1]$

Results of extensive numerical experiments show that algoritms based on the ADMM are among the state-of-the-art methods (Ng et al., 2010; Afonso et al., 2011; Chambolle and

Pock, 2011; Afonso, Bioucas-Dias and Figueiredo, 2010). In order to include convex constraints, $f \in C_v = [0,1]^n$, the following augmented Lagrangian $\mathcal{L}(f,(g_i),h,(\lambda_i),\lambda_C)$ is considered:

$$\mathcal{L} = \sum_i (\|g_i\|_2 - \lambda_i^t(g_i - D_i f) + \frac{\beta}{2}\|g_i - D_i f\|_2^2) + \frac{\alpha}{2}\|Rf - p^n\|_2^2 + I_{C_v}(h) + \frac{\beta}{2}\|h - f\|_2^2 - \lambda_C^t(h - f) \quad (2.7)$$

where α is a weighting parameter, β the Lagrangian parameter, I_{C_v} the indicator function of the convex set C_v, $I_{C_v}(x) = 0$ if $x \in C_v$, and ∞ otherwise. The Lagrange multipliers λ_i, λ_c are vectors in R^2 and R^{N^2}. For each pixel i, $D_i f \in R^2$ represents the first-order finite difference at pixel i in both horizontal and vertical directions, $(g_i)_{1\leq i\leq n}$ and h are the auxiliary unknowns corresponding to the gradient and to the convex constraint. Some sequences $(f^k,(g_i^k)_{1\leq i\leq n},h^k,(\lambda_i^k)_{1\leq i\leq n},\lambda_C^k)$ are constructed with successive minimizations to obtain the saddle point of the Lagrangian. The minimization scheme is detailed in (Sixou, 2013). For each pixel i:

$$g_i^{k+1} = max\{\|D_i f^k + \frac{1}{\beta}(\lambda_i^k)\| - \frac{1}{\beta}, 0\}\frac{D_i f^k + \frac{1}{\beta}(\lambda_i^k)}{\|D_i f^k + \frac{1}{\beta}(\lambda_i^k)\|} \quad (2.8)$$

The h^k update is:

$$h^{k+1} = \pi_C(f^k + \frac{\lambda_C^k}{\beta}) \quad (2.9)$$

where π_C is the projection on the convex set C. The new iterate f^{k+1} is obtained from the following linear system (Wang and Peyrin, 2014; Wang and Peyrin, 2016)::

$$(\sum_i D_i^t D_i + \frac{\mu}{\beta}R^t R + I)f^{k+1} = \sum_i D_i^t(g_i^{k+1} - \frac{1}{\beta}\lambda_i^k\} + \frac{\mu}{\beta}R^t p^\delta + h^{k+1} - \frac{\lambda_C^k}{\beta} \quad (2.10)$$

where I is the identity operator. The Lagrange multipliers $(\lambda_i),\lambda_C$ are updated with:

$$\lambda_i^{k+1} = \lambda_i^k - \beta(g_i^{k+1} - D_i f^{k+1}) \quad (2.11)$$

$$\lambda_C^{k+1} = \lambda_C^k - \beta(h^{k+1} - f^{k+1}) \quad (2.12)$$

The solution is trapped in a local minimum of the non convex binary problem (Sixou, 2013). For a small number of projections and the higher noise levels, the inverse problem becomes very ill-posed and some reconstruction errors are obtained on the boundaries of the binary image. Some regions are lost in the restored image. A new stochastic search approach to improve the reconstruction based on this deterministic algorithm is presented in Section 3.

3 Stochastic optimization based on level-set and TV regularization

3.1 Preliminaries

Our aim is to refine the local optima obtained with the former algorithms with stochastic optimization methods. Stochastic gradient methods are well-known in the field of convex optimization (Bertsekas.D. and Tsitsiklis.J., 2000). Let $(\Omega,\mathcal{F},P)$ be a probability space, in order to

obtain the global minimum of a function $g : \mathbb{R}^m \to \mathbb{R}$, a random trajectory $X(t)$ governed by the following diffusion process is often used (Gidas, 1995; Parpas and Rustem, 2009; Chow and Zhou, 2009; Chiang and Sheu, 1987; Geman and Hwang, 1986):

$$dX(t) = -\nabla g(X(t))dt + \mu(t)dW(t) \tag{3.1}$$

where W=$(W_1(t), ..., W_m(t))$ is the standard $m-$dimensional Brownian motion and $\mu(t)$ the stochastic noise strength. For an appropriate annealing schedule $\mu(t)$ and under appropriate conditions on the function g, the transition probability of $X(t)$ converges weakly to a probability measure which has its support on the set of global minimizers (Gidas, 1995; Parpas and Rustem, 2009; Chow and Zhou, 2009; Chiang and Sheu, 1987; Geman and Hwang, 1986). The main idea of this method is to combine a gradient flow and a stochastic perturbation to escape the traps of local minimizers. In order to escape the local minima generated by the level-set or TV regularizations, stochastic search algorithms based on the former regularization functionals will be used that generate random trajectories. At each step of the stochastic evolution, the binary discrepancy term $\|Rf_b - p^n\|$ is calculated, where f_b is the binary image.

3.2 Stochastic level-set evolution

To obtain a smooth evolution of the boundary curve between 0 and 1 regions, a new stochastic level-set method was studied recently (Wang and Peyrin, 2015). We summarize in this section the main aspects of the method. Level-set methods have been much studied for image processing tasks (Aubert and Kornprobst, 2006). A stochastic approach based on the level-set formalism has been proposed for segmentation purposes (Juan and Postelnicu, 2006). As demonstrated by Juan et al. (Juan and Postelnicu, 2006), the evolution of the boundary curve is independent of the level-set function used to represent it if the Stratanovich integral is used for the stochastic evolution. It was proposed to improve the reconstruction image with the following stochastic partial differential equation for the level-set function θ, for $x \in D$:

$$d\theta(x,t) = \delta\theta(x,t) + \mu(t)|\nabla\theta(x,t)| \circ dW(t) \tag{3.2}$$

where $\circ$ denotes the Stratanovitch convention (DaPrato and Zabczyk, 1992) and $\delta\theta$ is the gradient calculated as explained in Section II.B. Using the definition of the Stratonovich integral (DaPrato and Zabczyk, 1992), the equation can be transformed to get the following Itô stochastic differential equation (Juan and Postelnicu, 2006):

$$\begin{aligned} d\theta(x,t) &= \delta\theta(x,t) + \mu(t)|\nabla\theta(x,t)|dW(t) \\ &+ \frac{1}{2}\mu(t)(\triangle\theta(x,t) - |\nabla\theta(x,t)|div(\frac{\nabla\theta(x,t)}{|\nabla\theta(x,t)|})) \end{aligned} \tag{3.3}$$

As detailed in the simulation section, the stochastic search is performed with an intermittent diffusion method: level-set and stochastic level-set schemes are applied successively on random time intervals and with random diffusion strengths μ (Chow and Zhou, 1987; Chow and Zhou, 2009). This method was compared with the simulated annealing method in (Wang and Peyrin, 2015). A faster convergence is obtained with the stochastic level-set approach. Yet, with this method, only the boundaries between the 0 and 1 regions are modified. No new region is revealed by the algorithm and it not very efficient for the higher noise levels.

3.3 Stochastic TV based optimization

We propose to improve the solution obtained by TV regularization with stochastic search methods. Recently, some stochastic partial differential equations based on the TV regularization have been studied theoretically (Barbu and Rockner, 2009), with an evolution equation of the form:

$$dX(t) = \mathbf{div}(\mathbf{sgn}(\nabla X(t)))dt + \sigma(X(t))dW(t) \tag{3.4}$$

where $X \in H = L^2(\mathbb{R}^n)$, $W(t)$, $t \in [0, T[$, $T > 0$, is a Wiener process with covariance operator C (DaPrato and Zabczyk, 1992; Prevot and Rockner, 1992), the map $\sigma(., t)$ takes its values in the space of Hilbert-Schmidt operators and the multi-valued function $u \to \mathbf{sgn}(u)$ is the usual sign function. Some preliminary results with a noise term based on the gradient of the image have been presented in (Wang and Peyrin, 2014). In this section, we consider a coupling of ADMM with a stochastic diffusion for the augmented Lagrangian but we study several new noise terms. In the ADMM algorithm, the f iterate is obtained with a minimization of augmented Lagrangian $\mathcal{L}$, for the parameters $((g_i^{k+1}), h^{k+1}, (\lambda_i^k), \lambda_C^k)$:

$$f^{k+1} = \arg\min_f \mathcal{L}(f, (g_i^{k+1}), h^{k+1}, (\lambda_i^k), \lambda_C^k) \tag{3.5}$$

More precisely, the iterate f^{k+1} is obtained with the first-order optimality condition:

$$\nabla\mathcal{L}(f^{k+1}) = 0 \tag{3.6}$$

In order to improve the discontinuities of the reconstructed image and to reveal new regions, we propose to add several types of random perturbation to the gradient with respect to f of the Lagrangian regularization functional. We thus add some noise to the iterate f^{k+1}, and we consider the following stochastic partial differential equation for different types of noise σ :

$$df(t) = -\nabla\mathcal{L}(f, (g_i^{k+1}), h^{k+1}, (\lambda_i^k), \lambda_C^k)dt + \sigma(f(t), t)dW(t) \tag{3.7}$$

We introduce in the following three different noise terms corresponding to shape and topology changes.

1) A gradient dependent noise term written (Algorithm A_1):

$$\sigma(f(t), t)dW(t) = \mu_1(\frac{\partial f}{\partial x}dW_1(t) + \frac{\partial f}{\partial y}dW_2(t)) \tag{3.8}$$

where $(W_k(t))_{k=1,2}$ are independent Wiener random fields on H with a continuous covariance function C_k with a bounded integral kernel r_k, and μ_1 a positive constant that controls the strength of the noise. The gradient of the function f is used to detect the boundaries. This type of noise will be associated to a stochastic perturbation of the shape of the 0-1 regions with a change of their boundaries. The aim is to improve the efficiency of the method with random perturbations localized on the reconstruction errors, without changing the regions that are well restored.

2) An additive noise with an adapted covariance operator (Algorithm A_2):

Let us assume that the noise covariance operator $C : H \to H$ is a linear symmetric non-negative compact operator with eigenvalues μ_k and with a complete normalized eigenfunctions $(\phi_k)_{k\geq 1}$ system:

$$C\phi_k = \mu_k\phi_k \tag{3.9}$$

and that it is a trace class operator:

$$\|C\|_{L_1} = Tr(C) = \sum_{k=1}^{\infty} \mu_k < \infty \tag{3.10}$$

Under the former assumptions, the C-Wiener process $W(t)$ has the following series representation (DaPrato and Zabczyk, 1992; Prevot and Rockner, 1992):

$$W(t) = \sum_{k=1}^{\infty} \sqrt{\mu_k} \omega_t^k \phi_k \tag{3.11}$$

where $\{\omega_t^k\}$ is a sequence of independent, identically distributed standard Brownian motions in one dimension. In this work, the eigenfunctions ϕ_k will be the one obtained from a truncated wavelet decomposition of the boundary. The eigenvalues $\{\mu_k\}_{k\in J}$ corresponding to the high frequency wavelets used for the decomposition of the boundary are set to a constant value μ_2. The other eigenvalues of the covariance operator are set to zero. The noise term can thus be written:

$$W(t) = \mu_2 \sum_{k\in J} \omega_t^k \phi_k \tag{3.12}$$

where μ_2 is a positive constant. With this approach, the covariance of the noise is adapted to the boundary between the 0 and 1 regions since only the wavelets corresponding to the decomposition of the transition regions are taken into account. The aim is to apply the random perturbation on the discontinuities between the 0 and 1 regions. This type of noise term is expected to improve the performance since it is localized on boundaries reconstruction errors. 3) A nonlinear gradient dependent noise term is also considered given by:

$$\sigma(f(t))dW(t) = \mu_3(1 - f(t))R^*(Rf(t) - p^n)dW(t) \tag{3.13}$$

where μ_3 is a positive constant and $W(t)$ a C-Wiener random field with a bounded kernel. This noise term is proportional to the gradient of the data term of the objective functional. It is larger in the 0 regions of the image. The rationale behind this choice is to modify the topology of the 0 and 1 regions. With this type of random perturbation, new regions may be introduced depending on the value of the gradient of the data term. This type of noise can be added to the noise terms used in the algorithms (A_1) and (A_2) leading to the algorithms (A_3) or (A_4) respectively.

4 Linearisation of the evolution equations

Some stochastic algorithms based on the ADMM method have been investigated in (Ouyang and Gray.A.G., 2013). An additional regularization term is included in the Lagrangian to obtain a stable solution. It ensures the unicity of the solution but prevents the full exploration of the image solution space. Moreover, the algorithm does not use the reconstruction error distribution. The stochastic evolution equations proposed in our work are coupled and non-linear. Convergence results are difficult to demonstrate. In this section, we present an approximate linear system describing the dynamical evolution of the iterates. This type of linearization is similar to the one applied in the non-linear discrete Kalman filtering field (Einicke

and White.L.B., 1999; Einicke, 2012). We develop a linearized version of the system by starting with a reference trajectory which is calculated with the deterministic ADMM iterations without the stochastic noise. To obtain simple update formula, the convex constraints applied with Eq.2.9. and Eq.2.12. are not considered in the linearized equations. The iterates for the reference trajectory are denoted f_r^k, $g_{i,r}^k$, $\lambda_{i,r}^k$ and the stochastic iterates f^k, g_i^k, λ_i^k. The differences between the iterates obtained with or without noise f^k and f_r^k, $g_{i,r}^k$ and g_i^k, λ_i^k and $\lambda_{i,r}^k$ will be denoted as δf^k, δg_i^k, $\delta\lambda_i^k$ respectively. The Wiener noise is replaced by a n-dimensional Gaussian distributed random variable W_0 with mean zero and covariance Γ_0, $W_0 = \mathcal{N}(0, \Gamma_0)$. We assume that the Gaussian noise W_0 and the iterates f^k, g_i^k, λ_i^k are uncorrelated for all indexes k.

Proposition 4.1. *Under the former assumptions, the iterates* $\{\delta f^k\}$ *are Gaussian sequences and the steady state covariance is determined by the covariance equation* $H_\infty \Gamma_f^\infty H_\infty^t = G(f_r^\infty) \otimes \Gamma_0$ *with with* $H_\infty = (\frac{1}{\beta - 1}\sum_i D_i^t D_i + \frac{\mu}{\beta} R^t R + I)$ *and* G *is a function of the reference steady state* f_r^∞.

Proof. Eq. 2.10 is linear and the linear model that describes the dynamics of δf^k is:

$$(\sum_i D_i^t D_i + \frac{\mu}{\beta} R^t R + I)\delta f^{k+1} = \quad \sum_i D_i^t(\delta g_i^{k+1} - \tfrac{1}{\beta}\delta\lambda_i^k\} + G(f_r^k) \otimes W_0 \tag{4.1}$$

where $G(f_r^k)$ is a function of f_r^k describing the different types of noise and where $\otimes$ denotes the Kronecker product.

The equation 2.8 can be rewritten with the proximity operator of the Euclidean norm in $\mathbb{R}^n$, $\|.\|_2$, with $v_k = D_i f^k + \frac{1}{\beta}(\lambda_i^k)$.

$$prox_{\|.\|_2/\beta}(v_k) = (1 - \frac{1}{\beta\|v_k\|_2})_+ v_k = \begin{cases} (1 - \frac{1}{\beta\|v_k\|_2})v_k & \text{if} \quad v_k \geq 1/\beta \\ 0 & \text{otherwise} \end{cases} \tag{4.2}$$

The proximity operator may be linearized around $v_{k,r} = D_i f_r^k + \frac{1}{\beta}(\lambda_{i,r}^k)$:

$$prox_{\|.\|_2/\beta}(v_k) = (1 - \frac{1}{\beta\|v_{k,r}\|_2})v_{k,r} + (1 - \frac{1}{\beta\|v_{k,r}\|_2})h + \frac{1}{\beta}\frac{v_{k,r}}{\|v_{k,r}\|_2^3} < v_{k,r}, h > \tag{4.3}$$

with $h = D_i\delta f^k + \frac{1}{\beta}\delta\lambda_i^k$.

We thus find the linear approximation:

$$\delta g_i^{k+1} = (1 - \frac{1}{\beta\|v_{k,r}\|_2})(D_i\delta f^k + \frac{1}{\beta}\delta\lambda_i^k) + \frac{1}{\beta}\frac{v_{k,r}}{\|v_{k,r}\|_2^3} < v_{k,r}, D_i\delta f^k + \frac{1}{\beta}\delta\lambda_i^k > \tag{4.4}$$

The equation 2.11 is linear and thus we obtain the following update formula:

$$\delta\lambda_i^{k+1} = \delta\lambda_i^k - \beta(\delta g_i^{k+1} - D_i\delta f^{k+1}) \tag{4.5}$$

It is possible to show recursively that for all n the vector $((\delta f^k)_{0\leq k\leq n}, (\delta g_i^k)_{0\leq k\leq n}, (\delta\lambda_i^k)_{0\leq k\leq n})$ is Gaussian. Let us assume the first vector $(\delta f^0, (\delta g_i^0)_i, (\delta\lambda_i^0)_i)$ is Gaussian and that $((\delta f^k)_{0\leq k\leq n}, (\delta g_i^k)_{0\leq k\leq n}, (\delta\lambda_i^k)_{0\leq k\leq n})$ is Gaussian. From the linear relations Eq.4.4 and Eq.4.5, we see that $(\ (\delta g_i^k)_{0\leq k\leq n+1}, (\delta\lambda_i^k)_{0\leq k\leq n+1})$ is Gaussian. We have assumed that W_0 is uncorrelated with the vector $((\delta f^k)_{0\leq k\leq n}, (\delta g_i^k)_{0\leq k\leq n}, (\delta\lambda_i^k)_{0\leq k\leq n})$, thus it is also the case for $G(f_r^k) \otimes W_0$. Therefore the sequences $(\ (\delta g_i^k)_{0\leq k\leq n+1}, (\delta\lambda_i^k)_{0\leq k\leq n+1}, G(f_r^k) \otimes W_0)$ and $(\delta f^k)_{0\leq k\leq n+1}$ obtained with

Eq.4.1 are Gaussian.

We can thus assume that the iterates $\delta f^k, \delta g_i^k, \delta \lambda_i^k$ are Gaussian random variables with mean 0 and covariance matrices $\Gamma_f^k, \Gamma_{g,i}^k, \Gamma_{\lambda,i}^k$: $f_k = \mathcal{N}(0, \Gamma_f^k)$, $g_i^k = \mathcal{N}(0, \Gamma_{g,i}^k)$ and $\lambda_i^k = \mathcal{N}(0, \Gamma_{\lambda,i}^k)$.

From Eq.4.5 we obtain:

$$\Gamma_{\lambda,i}^{k+1} = \Gamma_{\lambda,i}^k - \beta(\Gamma_{g,i}^{k+1} - D_i \Gamma_f^{k+1} D_i^t) \tag{4.6}$$

Let us denote $H = (\sum_i D_i^t D_i + \frac{\mu}{\beta} R^t R + I)$, Equation 4.1 gives:

$$H\Gamma_f^{k+1} H^t = \sum_i D_i^t (\Gamma_{g,i}^{k+1} - \frac{1}{\beta}\Gamma_{\lambda,i}^k) D_i + G(f_r^k) \otimes \Gamma_0 \tag{4.7}$$

Ignoring the second term of Eq.4.3, to get a simple covariance equation, we obtain for large values of β:

$$\Gamma_{g,i}^{k+1} = (1 - \frac{1}{\beta \|v_{k,r}\|_2})(D_i \Gamma_f^k D_i^t + \frac{1}{\beta}\Gamma_{\lambda,i}^k) \tag{4.8}$$

The steady state equations obtained with $k \to \infty$ are the following:

$$\Gamma_{g,i}^{\infty} = D_i \Gamma_f^{\infty} D_i^t \tag{4.9}$$

$$H\Gamma_f^{\infty} H^t = \sum_i D_i^t (\Gamma_{g,i}^{\infty} - \frac{1}{\beta}\Gamma_{\lambda,i}^{\infty}) D_i + G(f_r^{\infty}) \otimes \Gamma_0 \tag{4.10}$$

$$\Gamma_{g,i}^{\infty} = (1 - \frac{1}{\beta \|v_{\infty,r}\|_2})(D_i \Gamma_f^{\infty} D_i^t + \frac{1}{\beta}\Gamma_{\lambda,i}^{\infty}) \tag{4.11}$$

From Eq.4.10, we deduce that:

$$\Gamma_{\lambda,i}^{\infty} = \frac{\beta}{\beta \|v_{\infty,r}\|_2 - 1} D_i \Gamma_f^{\infty} D_i^t \tag{4.12}$$

$$H\Gamma_f^{\infty} H^t = \frac{2 - \beta \|v_{\infty,r}\|_2}{1 - \beta \|v_{\infty,r}\|_2} \sum_i D_i^t D_i \Gamma_f^{\infty} D_i^t D_i + G(f_r^{\infty}) \otimes \Gamma_0 \tag{4.13}$$

and thus

$$H_{\infty} \Gamma_f^{\infty} H_{\infty}^t = G(f_r^{\infty}) \otimes \Gamma_0 \tag{4.14}$$

with $H_{\infty} = (\frac{1}{\beta - 1} \sum_i D_i^t D_i + \frac{\mu}{\beta} R^t R + I)$. □

We obtain a linear covariance equation with a forcing covariance Γ_0. If H_{∞} is inversible, the covariance matrix Γ_f^{∞} is a well-defined positive definite covariance matrix and the final iterate is a Gaussian vector fluctuating around f_r^{∞}, $\mathcal{N}(f_r^{\infty}, \Gamma_f^{\infty})$. With this linear model, it is possible to have a qualitative understanding of the properties of the steady state solution. With the gradient noise term or the wavelet decomposition of the boundaries, the forcing term $G(f_r^{\infty}) \otimes \Gamma_0$ is located on the boundaries bewteen the 0 and 1 regions. It the β parameter is large, then the inverse H_{∞}^{-1} can be approximated by $H_{\infty}^{-1} \sim I - \frac{\mu}{\beta} R^t R - \frac{1}{\beta - 1} \sum_i D_i^t D_i$. The discrete sum $\sum_i D_i^t D_i$ corresponds to the smoothing continuous operator Laplacian Δ. Let us assume the boundary region between the 0 and 1 regions may be decomposed with a shearlet basis. Such shearlet bases provides sparse representation for 2D functions which are smooth away from discontinuities along curves and they nearly diagonalizes the pseudo-differential operator $R^* R$. The boundary region is thus globally preserved by the operator H_{∞}^{-1}. In the framework of our linear model, the main part of the boundary noise is concentrated in the vicinity of the interface between the 0 and 1 regions.

5 Comparison of the algorithms: results and discussion

In the following, we have compared the convergence properties of the stochastic level-set method and of the TV based stochastic method. We present in this section some simulation details and the numerical results.

5.1 Simulations details

The TV and stochastic TV based methods, the level-set and stochastic level-set methods were applied to simulated projections of experimental bone cross-sections obtained with synchrotron micro-CT (voxel size: 15 μm) (Apostol et al., 2006). Fig.1 displays the bone cross-section images of size $N^2 = 256^2$ reconstructed from Filtered Back Projections (FBP) (Natterer, 1986) with 400 projections in a parallel beam geometry and 400 X-rays per projection and subsequently thresholded. These images correspond to high and low bone density cross-sections. These images are denoted f^* and are considered as ground truth images. The discrete approximation of the projection operator R is the Radon transform implemented on Matlab Image Processing Toolbox. For both stochastic methods, an intermittent diffusion is applied and the deterministic and stochastic schemes are applied successively and iteratively. At the end of each deterministic or stochastic run, the image is binarized.
The stochastic methods were tested for $M = 5$, $M = 10$, $M = 15$ or $M = 20$ equally spaced noisy projections, $N_r = 367$ rays per projection, with a Gaussian noise added to the projection. The noise distribution can be characterized by the standard deviation of the noise σ_p or peak to peak signal to noise ratio PPSNR:

$$PPSNR = 20log(\frac{f_{max}}{n_{max}}) \tag{5.1}$$

where f_{max} and n_{max} are the maximul signal and noise amplitude respectively. The noise standard deviations tested are $\sigma_p = 3$ (PPSNR=20/15 dB), $\sigma_p = 6.5$ (PPSNR=14/8.5 dB), $\sigma_p = 10$ (PPSNR=11/5 dB), $\sigma_p = 20$ (PPSNR=6/1.5 dB) or $\sigma_p = 30$ (PPSNR=3.7/1 dB) respectively for high and low density bone cross-sections. The noise level n can be estimated by $n = \sqrt{MN_r}\sigma_p$.

5.1.1 Level-set based methods

We summarize here the implementation of the method detailed in (Wang and Peyrin, 2014; Wang and Peyrin, 2015).

(a) (b)

Figure 1: Reconstructed images of the bone cross-section from 400 projections with the FBP algorithm. (a) High density bone image, (b) Low density bone image.

Stochastic level-set scheme

Step 1: Apply the deterministic level-set regularization scheme starting with initial level-set function $\theta_0 = 0$ to obtain an image f_0. The iterations are stopped when the iterates stagnate $\frac{\|f_{k+1}-f_k\|_2}{\|f_k\|_2} < 0.01$. At the end of this first optimization step, the Morozov discrepancy principle (Engl.H.W. and Neubauer.A, 1996) is not satisfied. The discrepancy term is much higher than the noise level, $\|p^n - Rf_b\| >> n$. After binarization, some reconstructions errors are still present at the boundaries between the 0 and 1 regions.

For k=1 to Maxiter do:

Step 2: Apply the stochastic algorithm. To discretize Eq.3.3, we have used an explicit scheme with finite differences, the WENO scheme (Jiang.G.S. and Peng.D, 2000) with spatial discretization step $\Delta x = 1$ and time step $\Delta t = 0.1$. The noise strength μ and the number of iterations N_{sto} are chosen randomly with a uniform distribution in $[0.01, 0.1]$ and $[1, 100]$.

Step 3: Apply a deterministic level-set step with 100 iterations.

Step 4: Binarize the image by thresholding and reinitialize the level-set function θ with the signed distance function.

To obtain a good accuracy, the ϵ parameter should be sufficiently small, $\epsilon = 0.03$. The regularization parameter β_1 was set to 0 because the H_1 term dominates the TV term (DeCezaro et al., 2009). And we tested many parameters (β_2) to obtain the best decrease of the level-set regularization functional Eq.2.3.

5.1.2 Total Variation based methods

Here are the details of the implementation of stochastic TV method.

Stochastic Total Variation scheme

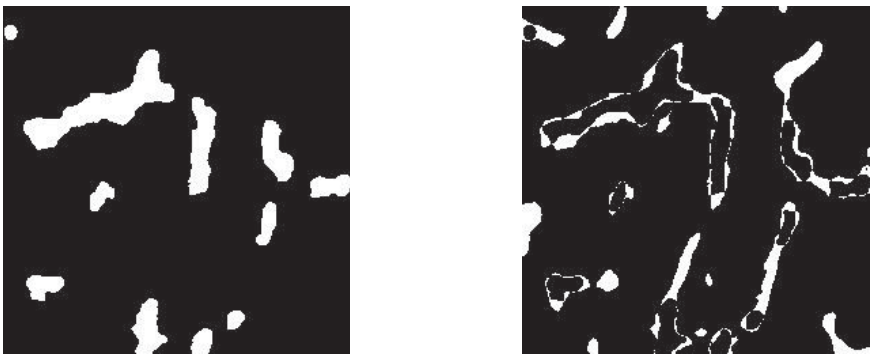
Step 1: Apply the deterministic TV regularization and the ADMM deterministic algorithm. The iterations are stopped when the iterates satisfies $\frac{\|f^{m+1}-f^m\|_2}{\|f^m\|_2} < 0.01$. The final image obtained at the end of the optimization process depends only on the parameter α. The α parameter is chosen such that the end of the minimization process $\|Rf(\alpha) - p^n\| \sim n$.

After binarization, the initial binary reconstructed image is denoted as f_0. The discrepancy term of f_0 is well-above the noise level and the Morozov principle is not fulfilled any longer. A local minimizer is again obtained.

For k=1 to Maxiter do:

Step 2: Apply the stochastic TV algorithm. From the iterate f^k_{STO}, the next iterate f^{k+1}_{STO} is calculated as the sum of the ADMM iterate f^{k+1} and one or several discretized noise terms on random time steps in the range $[0, T_{max}]$ with $T_{max} = 100$ and with stochastic noise strengths μ_i, (i=1,2 or 3). For each type of noise, the noise strength parameters μ_1, μ_2, μ_3 are chosen by trial and error to obtain the best decrease of the discrepancy term $\|Rf^k_b - p^n\|$, where f^k_b is the binarization of the grey-level image.

Step 3: Apply the deterministic TV minimization. The TV deterministic iterations are stopped when $\frac{\|f^{m+1}-f^m\|_2}{\|f^m\|_2} < 0.01$.



(a) (b)

Figure 2: (a) Reconstruction of the high density image f_0 obtained with the TV regularization for $\sigma_p = 20$ (PPSNR=6 dB) and M=10 projections, $MR = 8.12\%$; (b) Corresponding error map.

For comparison, some simulations have been performed in which the stochastic diffusions are replaced by a successive TV regularization minimizations separated by binarization steps (Algorithm A_0). The infinite-dimensional Wiener processes were approximated by Gaussian random field on the image grid. The discretization of the stochastic partial differential equation and of the Wiener processes were performed with classical finite difference methods and the Euler-Maruyama method (Kloeden.P.E and Platen.E, 1992). The C-Wiener processes were approximated with stationary Gaussian random fields with a correlation function C given by its Fourier transform $\tilde{C}(\mathbf{k}) = (|\mathbf{k}|^2 + 1)^{-2}$. These random fields were generated via FFT with independent normal distributed random numbers. For the algorithm (A_2), we assume that

(a) (b)

Figure 3: (a) Reconstruction of the low density image f_0 obtained with the TV regularization for $\sigma_p = 20$ (PPSNR=1.5 dB) and M=10 projections, $MR = 3.14\%$; (b) Corresponding error map.

the image f admits a sparse representation in an orthogonal wavelet basis $\{\phi_k, k \geq 0\}$. The index set describes the various levels of the resolution, the different positions and types of wavelet (Daubechies, 1992). The image f can thus be written $f = \mathcal{W}^*\mathbf{v}$, where $\mathbf{v} \in l_2$ is a wavelet coefficients vector, and $\mathcal{W}^*$ a synthesis operator. In this work, we have used the orthogonal Daubechies wavelet basis (Matlab implementation) and a 2-level wavelet decomposition of the images. Only the high frequency wavelet coefficients with the vertical, horizontal and diagonal details are taken into account in the noise term. These coefficients corresponds to the boundary between the 0 and 1 regions. For the algorithms A_3 and A_4, the homogeneous noise term is calculated at each iteration with the formula $\mu_3(1-f^k)Z_k\sqrt{\Delta t}R^*(Rf^k-p^n)$ where the $(Z_k)_{k\geq 0}$ are spatially correlated Gaussian random variables in R^{N^2}. The time step Δt is fixed to 0.1.

5.2 Numerical results

We present in this section the numerical results obtained with the different optimization methods. In order to have some quantitative results, the binarization of the grey-level image, f_b^k, is obtained at each iteration with a threshold of 0.5 and the data term $\frac{|\|Rf_b^k-p^n\|-n|}{n}$is calculated. The efficiency of the reconstruction process is evaluated with the misclassification rate MR:

$$MR = \frac{\sum |f_b(i) - f^*(i)|}{N^2} \times 100\% \tag{5.2}$$

where f_b is the binarized version of the reconstructed image and f^* is the ground truth. The uncertainty on the optimal misclassification rate, MR(%), estimated from several noise realizations and changes of 10% of the regularization parameters, is $\Delta MR = 0.05\%$. The negative rate $nMR(\%)$, positive rate $pMR(\%)$ are also evaluated.

$$\begin{cases} nMR = \frac{\sum_{i=1}^{N^2} |f_b(i)-f^*(i)|}{N^2} \times 100\% \quad if \;\; f_b(i) - f^*(i) < 0 \\ pMR = \frac{\sum_{i=1}^{N^2} |f_b(i)-f^*(i)|}{N^2} \times 100\% \quad if \;\; f_b(i) - f^*(i) > 0 \end{cases} \tag{5.3}$$

With the same input data, the FBP algorithm followed by thresholding leads to very bad reconstruction results with misclassification rates MR between 30% and 40%.

5.2.1 Level-set based method

The results are fully detailed in (Wang and Peyrin, 2015). They will be summarized here for a comparison with the stochastic Total Variation method. The initial value of the data term obtained after the first level-set scheme is well above the noise levels. The level-set algorithm can not escape this local minimum. With the iterations, a significant decrease of the data term is achieved towards these noise levels with the stochastic level-set method. A significant decrease of the misclassification rate as a function of the number of iterations is also obtained. The missclassification rates, negative rates, positive rates and data terms obtained at the end of the simulations are summarized in Table 1. As seen from the table, a large decrease of the misclassification rates was achieved with the stochastic level-set approach for the low noise levels and numbers of projections investigated (Wang and Peyrin, 2015). At the end of the simulations, the errors on the boundary of the images are much lower. The smooth evolution of the boundary proves to be more efficient than a Markov chain approach (Wang and Peyrin, 2015). The stochastic level-set method corresponds to a shape evolution and does not modify the topology of the 0 and 1 regions of the reconstructed images. The Table 1 also displays the misclassifications rates obtained with the TV methods. The TV deterministic optimization gives misclassifications rates that are a little higher than the ones achieved with the stochastic level-set method. Some improvement is obtained with the stochastic TV scheme (A_4) with a low noise strength. The best misclassification rates are similar for the two stochastic search methods for these low noise levels for M=15. The stochastic TV method outperforms the stochastic LS scheme for a lower number of projection angles M=10.
Yet, the improvement of the reconstruction with stochastic level-set method is not clear for low density images where many regions are missing in the first reoncstructed images and also for the higher noise levels. The stochastic TV method is much more effective in these cases, as detailed in the following subsection. It is thus interesting to study methods to add stochastic noise in the framework of the TV regularization and random perturbations of the topology of the images to reveal new regions.

5.2.2 Stochastic evolutions based on TV regularization

Two examples of the reconstructed images f_0 obtained with the deterministic TV algorithm and the corresponding error maps for the high density and low density bone images are displayed in Fig.2 and Fig.3 with the projections number $M = 10$ and the standard deviation of the noise on the projections $\sigma_p = 20$. Similarly to the reconstruction results obtained with the deterministic level-set method, some errors are still present on the boundary regions of the image. Especially, for the low density bone image, large regions have disappeared from the restored image.
For the dense bone image, the evolution of the discrepancy term $\|Rf_b^k - p^n\|$ for the binary image starting from the best image f_0 obtained with the TV regularization with box

Table 1: Misclassification rates MR ($\%$), negative rates nMR ($\%$), positive rates pMR ($\%$), $|\parallel Rf - p^n \parallel - n|/n$ obtained with the two level-set methods, and the two TV methods.

M, σ_p	LS	Stochastic LS	TV	Stochastic TV
	2.02	1.56	1.73	1.51
M=15	1.51	0.71	0.89	0.86
$\sigma_p = 3$	0.51	0.85	0.84	0.65
	0.36	0.08	0.16	0.07
	3.05	2.48	2.71	2.48
M=15	2.35	1.29	1.49	1.43
$\sigma_p = 6.5$	0.70	1.18	1.22	1.05
	0.18	0.03	0.02	0.01
	3.29	2.55	2.41	2.08
M=10	2.61	1.31	1.39	1.14
$\sigma_p = 3$	0.68	1.24	1.01	0.94
	0.86	0.23	0.34	0.09
	4.46	3.97	3.40	3.12
M=10	3.54	2.26	2.21	1.98
$\sigma_p = 6.5$	0.92	1.72	1.19	1.14
	0.37	0.09	0.09	0.03

constraints, with M=10 projection angles and the standard deviation of the noise on projection $\sigma_p = 20$ is displayed in Fig.4 as a function of the iteration number for the algorithms $(A_0), (A_1), (A_2), (A_3), (A_4)$. The time of every stochastic TV iteration is very similar for the algorithms $(A_1)....(A_4)$ and it is 20 $\%$ longer than the time of the TV iterations, and 30 times shorter than the time of one stochastic LS iteration.

With the iterations, some decrease of the discrepancy term related to the binary image towards the noise level $n = 1210$ ($\sigma_p = 20$) is observed for these stochastic algorithms. The results for $\sigma_p = 30$ and $M = 15$ are very similar. The change of the discrepancy term is below 5 $\%$ for different noise realizations.

The evolution of the misclassification rates are displayed on Fig.5 for the same noise level $\sigma_p = 20$, the number of projections $M = 10$ and the same algorithms. There is a clear correlation between the decrease of the discrepancy term and the decrease of the misclassification rate. Similar decreases are obtained for the higher noise levels and for a higher number of projections. The misclassification rates and values achieved for the minimum of the binary data term are summarized in Table 2 for the various number of projections and noise levels. The stochastic approaches are more efficient than that a TV regularization used repeatedly (A_0). A significant decrease of the reconstruction errors is obtained with the algorithms (A_1) and (A_2) where the noise is concentrated on the boundaries. The more effective boundary noise term is achieved with the wavelet implementation (algorithm A_2). Yet, better results are found with the schemes (A_3) and (A_4) when some additional noise proportional to the derivative of the gradient is also added to modify the image. These two types of noise may be understood as

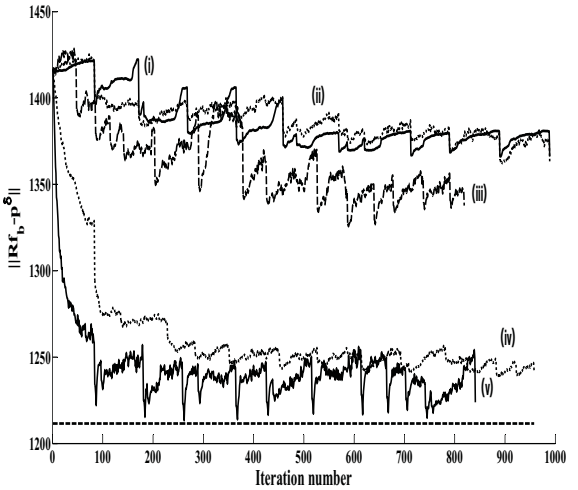


Figure 4: Evolution of the discrepancy term as the function of the iteration number for the high density bone image $M = 10, \sigma_p = 20$ and for the algorithms A_0(i), A_1(ii), A_2(iii), A_3(iv), and A_4(v). The noise level n is displayed for comparison with dotted lines.

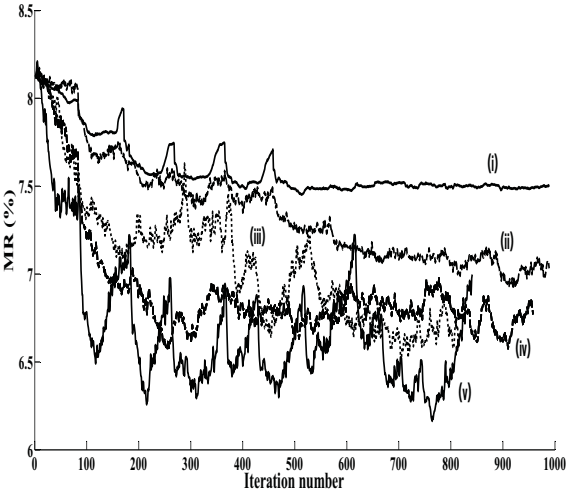


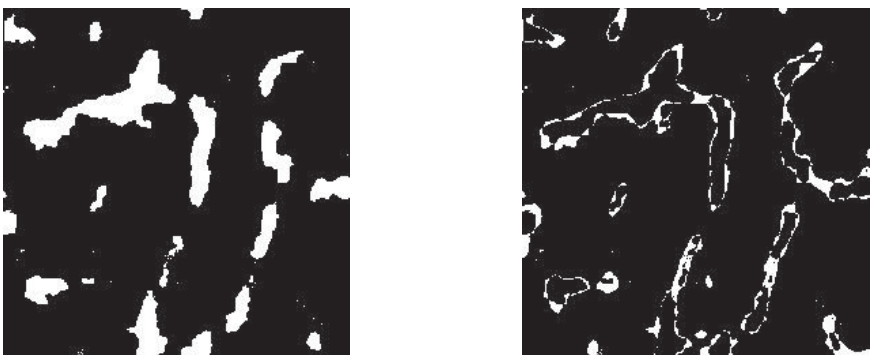
Figure 5: Evolution of the misclassification rate as the function of the iteration number for the high density bone image $\sigma_p = 20$ and for the algorithms A_0(i), A_1(ii), A_2(iii), A_3(iv), and A_4(v).

shape and topological stochastic modifications of the images. The image and the error map obtained for the minimum value of the discrepancy term for the algorithm (A_4) are displayed in Fig.6, for a number of projection $M = 10$ and the standard deviation of the noise on projection $\sigma_p = 20$. Very similar images are obtained with the algorithm (A_3). The missclassification rate calculated for these images is very close to the minimum rate achieved during the stochastic evolution. The boundary smoothness drops significantly but the errors on the boundaries have

Table 2: Misclassification rates MR ($\%$), negative rates nMR ($\%$), positive rates pMR ($\%$), $|\parallel Rf - p^n \parallel - n|/n$ obtained with the stochastic algorithm based on the Total Variation for high density bone image.

M, σ_p	TV	Stochastic TV				
		A_0	A_1	A_2	A_3	A_4
	7.18	6.89	6.82	6.60	5.03	5.03
M=15	1.68	1.68	1.44	1.27	1.86	1.84
$\sigma_p = 20$	5.50	5.22	5.38	5.33	3.17	3.19
	0.16	0.14	0.15	0.16	0	0
	9.54	9.36	7.29	9.17	6.58	6.30
M=15	1.33	1.21	1.29	0.92	2.11	1.73
$\sigma_p = 30$	8.21	8.15	7.99	8.25	4.47	4.57
	0.16	0.16	0.15	0.17	0.01	0.01
	8.12	7.45	6.92	6.51	6.57	6.16
M=10	1.87	1.69	1.32	1.28	2.42	2.02
$\sigma_p = 20$	6.26	5.76	5.61	5.23	4.15	4.15
	0.17	0.13	0.13	0.11	0.03	0.01
	10.18	9.60	9.30	8.34	7.70	7.67
M=10	1.19	1.30	1.34	1.74	2.13	2.24
$\sigma_p = 30$	9.0	8.30	7.96	6.60	5.56	5.43
	0.19	0.15	0.14	0.08	0.03	0.02

been much reduced. Some new regions have appeared, especially for the higher noise level. Similar simulations have been performed for M=15 projections. The best performance is again achieved with the combination of boundary noise and homogeneous noise.



(a) (b)

Figure 6: (a) Best reconstruction for the high density image obtained with the the nonlinear diffusion equation (A_4) for $\sigma_p = 20$, and $M = 10$ projections, $MR = 6.16\%$; (b) Corresponding error map.

For the low density bone image, the evolutions of discrepancy term $\parallel Rf_b^k - p^n \parallel$ and the

misclassification rate for the binary image starting from the best image f_0 obtained with TV regularization with box constraints for the algorithm (A_4) are displayed in Fig.7 and Fig.8 respectively. The misclassifications rates obtained are summarized in Table 3. The reconstructed images are displayed in Fig.9.

Table 3: Misclassification rates MR ($\%$), negative rates nMR ($\%$), positive rates pMR ($\%$), $|\parallel Rf - p^n \parallel - n|/n$ obtained with the stochastic algorithm based on the Total Variation (A_4) for sparse image.

M, σ_p	TV	Stochastic TV (A_4)
	3.05	2.91
M=5	0.17	0.2
$\sigma_p = 10$	2.88	2.71
	0.30	0.25
	2.39	2.08
M=10	0.47	0.51
$\sigma_p = 10$	1.91	1.56
	0.13	0.06
	1.96	1.67
M=20	0.23	0.30
$\sigma_p = 10$	1.72	1.37
	0.13	0.06
	3.14	2.62
M=20	0.07	0.34
$\sigma_p = 20$	3.06	2.29
	0.11	0.05

The correlation between the decrease of the discrepancy term and the decrease of the misclassification rate is clear for the four cases investigated. With the increase of the projection number and the decrease of the noise level, the inverse problem is less ill-posed and the stochastic method based on TV regularization is less useful to achieve a better reconstruction for this low density bone image. For example, for $M = 20, \sigma_p = 20$, the discrepancy term $\parallel Rf_0 - p^n \parallel$ of the first reconstructed image f_0 is very close to the noise level and it is not possible to improve much the reconstruction.

From these results, it is possible to compare the two stochastic methods presented. The stochastic level-set algorithm leads to a clear improvement of the reconstruction of the boundaries between the 0 and 1 regions. It is very useful for low noise levels but is does not yield improved reconstructions for the higher noise levels and for the low density image. For the higher noise levels, the stochastic TV based approach is more efficient because it leads to a modification of the boundaries but it also reveals new regions in the image when two types of noise are included. The random fluctuations lead to shape and topology changes of the 0 and 1 regions in the restored image.

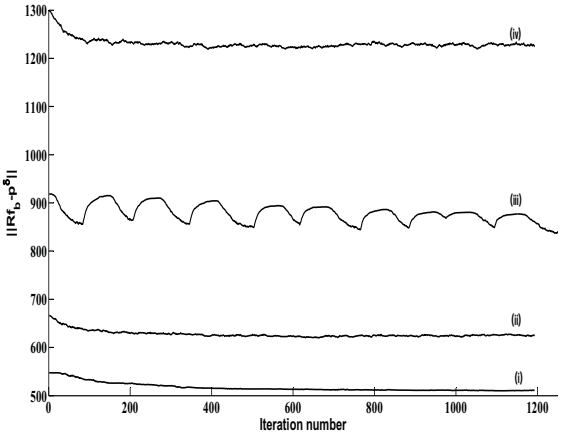


Figure 7: Evolution of the discrepancy term as the function of the iteration number for the low density bone image for the algorithms A_4, with $M = 5, \sigma_p = 10$ (i), $M = 10, \sigma_p = 10$ (ii), $M = 20, \sigma_p = 10$ (iii), $M = 10, \sigma_p = 20$ (iv).

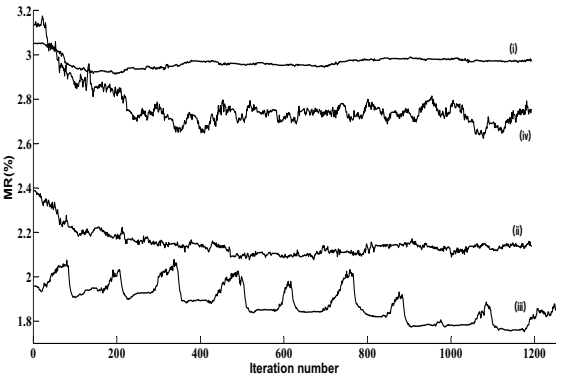


Figure 8: Evolution of the misclassification rate as the function of the iteration number for the low density bone image for the algorithms A_4, with $M = 5, \sigma_p = 10$ (i), $M = 10, \sigma_p = 10$ (ii), $M = 20, \sigma_p = 10$ (iii), $M = 10, \sigma_p = 20$ (iv).

6 Conclusion

This work compares new stochastic diffusion methods to reconstruct binary tomography cross-sections from few projection angles. A stochastic level-set method and a stochastic TV based method are used to improve the reconstructions obtained with deterministic algorithms on high and low density binary micro-CT bone sections. Stochastic level-set method leads to a decrease of the reconstruction errors localized on the boundaries for low noise levels and number of projections. And the stochastic search method based on the TV regularization is more efficient for the higher noise levels and for the low density image.

A first binary image is obtained with the deterministic level-set or TV regularization methods implemented with the Alternate Direction of Minimization method. The reconstructed image

(a) (b)

Figure 9: (a) Best reconstruction of the low density image obtained with the the nonlinear diffusion equation (A_4) for $\sigma_p = 20$, and $M = 10$ projections, $MR = 2.63\%$; (b) Corresponding error map.

is then refined with intermittent stochastic diffusion methods. In the level-set stochastic algorithm, the restoration is improved with a stochastic partial differential equation based on a Stratanovitch formulation. A linearization of the ADMM equations shows that the boundary noise remains located in the vicinity of the interface between the 0 and 1 regions with the iterations. In the TV based stochastic optimization method, the efficiency of different stochastic noise terms has been tested. The first type of noise is a boundary noise term. This noise is implemented with a finite difference estimation of the gradient or with a wavelet decomposition of the boundary. The second type of noise is an homogeneous noise proportional to the gradient of the data term. This stochastic algorithms leads also to a large decrease of the reconstruction errors localized on the boundaries. Some new regions are also revealed in the new reconstruction image. The best results are achieved when two types of stochastic noises are considered, corresponding to random change in the shape and in the topology of the images.

7 Appendix

In this section, we present some notions that are not commonl in the image processing and tomographic reconstruction literature. Further details can be found in (DaPrato and Zabczyk, 1992; Prevot and Rockner, 1992).
Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space. Let D be a bounded domain in $\mathbb{R}^2$ with a smooth boundary. A family of random variables $W(x,t) : \Omega \to \mathbb{R}^n$, $t \in [0,T]$, $x \in D$ is a finite-dimensional Wiener process with mean zero and covariance function $r(x,y)$ if

$$EW(x,t) = 0 \quad t \in [0,T] \quad x \in D \tag{7.1}$$

$$E(W(x,t)W(y,s)) = (t \wedge s)r(x,y) \quad t,s \in [0,T] \quad x,y \in D \tag{7.2}$$

Let $(\Omega, \mathcal{F}, \{\mathcal{F}\}_t, \mathcal{P})$ is a filtered probability spaces with filtration $\{\mathcal{F}\}_t, t \in [0,T]$. Let $W(t)$ be the

standard Brownian motion and $f(t)$ a continuous adapted process in $\mathbb{R}$ for $0 \leq t \leq T$. For any partition $\Delta_T = \{0 = t_0 < t_1 < ..... < t_n = T\}$, we define $|\Delta_T| = max_{1 \leq k \leq n}(t_k - t_{k-1})$ and

$$J_t^n = \Sigma_{k=1}^n \frac{1}{2}(f_{t_{k-1} \wedge t} + f_{t_k \wedge t})(W(t_k \wedge t) - W(t_{k-1} \wedge t)) \tag{7.3}$$

The sequence J_t^n converges uniformly in probability as $|\Delta_T| \to 0$ to J_t written as:

$$J_t = \int_0^t f(s) \circ dW(s) \tag{7.4}$$

which is known as the Stratonovitch integral of f with respect to W.
Let D be a bounded domain with a smooth boundary and $H = L_2(D)$ a Hilbert space, and $A : H \to H$ be a linear operator with a symmetric kernel $a(x, y) = a(y, x)$ defined by:

$$A\phi(x) = \int_D a(x, y)\phi(y)dy \quad \phi \in H \tag{7.5}$$

If a is square integrable so that:

$$\int_D \int_D |a(x, y)|^2 dxdy < \infty \tag{7.6}$$

then A is a self-adjoint Hilbert-Schmidt operator.

8 Acknowledgment

This work was performed within the framework of the LABEX PRIMES (ANR-11-LABX-0063) of University of Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR).

References

Acar, R. and Vogel, C. 1994. Analysis of total variation penalty methods, *Inverse Problems* **10**: 1217–1229.

Afonso, M., Bioucas-Dias, J. and Figueiredo, M. 2010. Fast image recovery using variable splitting and constrained optimization, *IEEE Transactions on Image Processing* **19**(9): 2345–2356.

Afonso, M., Bioucas-Dias, J. and Figueiredo, M. 2011. An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems, *SIAM Journal on Scientific computing* **20**(3): 681–695.

Apostol, L., Boudousq, V., Basset, O., Odet, C., Yot, S., Tabary, J., Dinten, J., Boller, E., Kotzki, P. and Peyrin, F. 2006. Relevance of 2d radiographic texture analysis for the assessment of 3d bone microarchitecture, *Medical Physics* **33**: 3546–3556.

Aubert, G. and Kornprobst, P. 2006. *Mathematical Problems in Image Processing:Partial Differential Equations and the Calculus of Variations*, Vol. 147, Applied Mathematical Sciences, Springer.

Azencott, R. 1992. Sequential simulated annealing:speed of convergence and acceleration techniques, *Simulated annealing:parallelization techniques, Wiley, New York* pp. 1–10.

Barbu, V. DaPrato, G. and Rockner, M. 2009. Stochastic nonlinear diffusion equations with singular diffusivity, *SIAM J.Math.Anal.* (41): 1106–1120.

Batenburg, K. J. and Sijbers, J. 2009. Generic iterative subset algorithm for discrete tomography, *Discrete Applied Mathematics* **157**: 438–451.

Bertsekas.D. and Tsitsiklis.J. 2000. Gradient convergence in gradient methods, *SIAM J.Optm.* **10**: 627–643.

Bouxsein, M. L., Boyd, S. K., Christiansen, B. A., Guldberg, R. E., Jepsen, K. J. and Müller, R. 2010. Guidelines for assessment of bone microstructure in rodents using micro-computed tomography, *Journal of Bone and Mineral Research* **25**: 1468–1486.

Bresson, X. Esedoglu, S. V. P. T. and Osher, S. 2007. Fast global minimization of the active countour/snake model, *Journal of Mathematical Imaging and Vision* **28**: 151–167.

Brown, E.S. Chan, T. and Bresson, X. 2011. Completely convex formulation of the chan-vese image segmentation model, *International Journal of Computer Vision* pp. 1–19.

Cai, W. and Ma, L. 2010. Comparaison of approaches based on optimization and algebraic iteration for binary tomography, *Computer Physics Communications* **181**: 1974–1981.

Capricelli, T. and Combettes, P. 2007. *Advances in discrete tomography and its applications: A convex programming algorithm for noisy discrete tomography*, Birkhauser, Boston, MA.

Catoni, O. 1992. Rough large deviation estimates for simulated annealing algorithms, *Annals of Probability* **20**: 1109–1146.

Chambolle, A. and Pock, T. 2011. A first-order primal-dual algorithm for convex problems with applications to imaging., *Journal fo Mathematical Imaging and Vision* **40**: 120–145.

Chan, T.F. Esedoglu, S. and Nikolova, M. 2006. Algorithms for finding global minimizers of image segmentation and denoising models, *SIAM Journal on Applied Mathematics* pp. 1632–1648.

Chiang, T.S. Hwang, C. and Sheu, S. 1987. Diffusion for global optimization in $\mathbb{R}^n$, *SIAM J.Control.Optim.* **25**: 737–753.

Chow, S. Yang, T. and Zhou, H. 1987. Diffusion for global optimization in $\mathbb{R}^n$, *SIAM J.Control.Optim.* **25**: 737–753.

Chow, S. Yang, T. and Zhou, H. 2009. Global optimization by intermittent diffusion, *Proceeding of the 8th Taiwan-Philippines Symposium on Analysis, Taiwan* .

Cot, C. and Catoni, . 1998. Piecewise constant triangular cooling schedule for generalized annealing algorithms, *Annals of Applied Probability* **8**: 375–396.

DaPrato, G. and Zabczyk, J. 1992. *Stochastic equations in infinite dimensions*, Cambridge University Press.

Daubechies, I. 1992. *Ten Lectures on Wavelets*, SIAM:Society for Industrial and Applied Mathematics, Philadelphia.

DeCezaro, A., Leitao, A. and Tai, X. C. 2009. On multiple level-set regularization methods for inverse problems, *Inverse Problems* **25**(035004).

Defrise, M, V. 2011. An algorithm for total variation regularization in high-dimensional linear problem, *Inverse Problems* .

Egger, A. and Leitao, L. 2009. Nonlinear regularization for ill-posed problems with piecewise constant or strongly varying solutions, *Inverse Problems* **25**(115014).

Einicke, G. 2012. *Smoothing, Filtering and Prediction: Estimating the Past, Present and Future*, Rijeka,Croatia:Intech.

Einicke, G. and White.L.B. 1999. Robust extended kalman filtering, *IEEE Trans.Signal Procssing* **47**: 2596–2599.

Engl.H.W., H. and Neubauer.A 1996. *Regularization of Inverse Problems*, Dordrecht, Kluwer Academic.

Geman, S. and Hwang, C. 1986. Diffusion for global optimization, *SIAM J.Control.Optim.* **24**: 1031–1043.

Gidas, B. 1995. Metropolis-type monte-carlo simulation algorithm and simulated annealing, *Topics in Contemporary Probability and Its Applications, Probablity Stochastics Series, CICS, Boca Raton* pp. 159–232.

Gouillart, E., Krzakala, F., Mezard, M. and Zdeborová, L. 2013. Belief propagation reconstruction for discrete tomography, *Inverse Problems* **29**(035003).

Herman, G. T. and Kuba, A. 2007. *Advances in discrete tomography and its applications*, Applied and Numerical Harmonic Analysis, Birkhauser, Boston, MA.

Jiang.G.S. and Peng.D 2000. Weighted eno schemes for hamiton-jacobi equations, *SIAM J.Sci.Comput.* **21**: 2126–2143.

Juan, 0. Keriven, R. and Postelnicu, G. 2006. Stochasti motion and the level-set method in computer vision: Stochastic active contour, *International Journal of Computer Vision* **69**: 7–25.

Kloeden.P.E and Platen.E 1992. *Numerical solution of stochastic differential equations*, Vol. 23, Application of Mathematics, Berlin:Springer.

Liao, H. Y. and Herman, G. 2004. Automated estimation of the parameters of the Gibbs priors to be uses in binary tomography, *Discrete Applied Mathematics* pp. 249–170.

Natterer, F. 1986. *The mathematics of computerized tomography*, Wiley-Teubner series in computer science, Vieweg+Teubner Verlag, UK.

Ng, M. K., Weiss, P. and Yuan, X. 2010. Solving constrained total-variation image restoration and recontruction problems via alternating direction methods, *SIAM Journal on Scientific computing* **32**: 2710–2736.

Ouyang, H. Hiao, N. T. and Gray.A.G. 2013. Stochastic alternating direction of method of multipliers, *Proceedings of the 30th International Conferenc on Machine Learning* pp. 80–88.

Parpas, P. and Rustem, B. 2009. Convergence analysis of a global optimization algorithm using stochastic differential equations, *J.Control. Optim.* **45**: 95–110.

Prevot, C. and Rockner, M. 1992. A concise course on stochastic partial differential equations, *Cambridge University Press* .

Rudin, L. I., Osher, S. and Fatemi, E. 1992. Nonlinear total variation based noise removal algorithms, *Phys.D.* **60**: 259–268.

Schillinger, B. 2005. Proposed combination of cad data and discrete tomography for the detection of coking and lubricants in turbine blades or engines, *Electonics Notes in Discrete Mathematics* **20**: 493–499.

Schüle, T., Schnörr, C., Weber, S. and Hornegger, J. 2005. Discrete tomography by convex-concave regularization and D.C programming, *Discrete Applied Mathematics* **151**: 229–243.

Sidky, E. Y. and Pan, X. 2006. Accurate image reconstrution feom few-views and limited-angle data in divergent-beam CT, *J.X-ray Sci.Techn.* **14**: 119–139.

Sidky, E. Y. and Pan, X. 2008. Image reconstruction in circular cone-beam computed tomography by constrained total variation minimization, *Phys.Med.Biol.* **53**: 4777–4807.

Sidky, E. Y. and Pan, X. 2012. Convex optimization problem prototyping for image reconstruction in computed tomography with the chambolle-pock algorithm., *Phys.Med.Biol.* **57**: 3065–3091.

Sixou, B. Wang, L. P. F. 2013. Bone microstructure reconstructions from few projections with level-set regularization, *IEEE International Symposium on Biomedical Imaging* pp. 1170–1173.

VanGompel, G. Batenburg, K. J. D. E. A. W. and Sijbers, J. 2010. A discrete tomography approach for superresolution micro-ct images:application to bone, *IEEE International Symposium on Biomedical Imaging* pp. 816–819.

Wang, L. Sixou, B. and Peyrin, F. 2014. Binary tomography reconstructions of bone microstructure from few projections with stochastic level-set methods, *IEEE International Conference Image Processing* .

Wang, L. Sixou, B. and Peyrin, F. 2015. Binary tomography reconstructions with stochastic level-set methods, *IEEE Signal Processing Letters* **22**: 922–924.

Wang, L. Sixou, B. R. S. and Peyrin, F. 2016. Binary tomography reconstructions from few projections with level-set regularization methods for bone microstructure study, *International Journal of Tomography and Simulation* **29**: 2345–2356.

Yu, H. and Wang, G. 2010. A soft-threshold filtering approach for reconstruction from a limited number of projections, *Phys.Med.Biol.* **55**: 3905–3916.