Efficient binary reconstruction for non destructive evaluation using gammagraphy

Marc Allain, Jérôme Idier

Efficient binary reconstruction for non destructive evaluation using gammagraphy

Marc Allain[1] **and Jérôme Idier**[2]

[1] Marc Allain was with the Institut de Recherche en Communications et en Cybernétique de Nantes (IRCCyN), 1 rue de la Noë, BP 92101, 44321 Nantes Cedex 03. He is now with Institut Fresnel, Domaine Universitaire de St-Jérôme, 13397 Marseille Cedex 20. E-mail: `marc.allain@fresnel.fr`.

[2] Jérôme Idier is with the Institut de Recherche en Communications et en Cybernétique de Nantes (IRCCyN), 1 rue de la Noë, BP 92101, 44321 Nantes Cedex 03. E-mail: `Jerome.Idier@irccyn.ec-nantes.fr`.

Abstract. The localization and the sizing of 3D flaws within a homogeneous metallic media is a major task for non destructive evaluation (NDE). This paper adresses the problem of the reconstruction of such flaws using an efficient binary algorithm. Basically, the method rests on the fact that a simple binary constraint suffices for an accurate and robust reconstructions in the context of NDE. A heuristic minimization, computationally attractive, is designed in order to provide fast reconstructions. The proposed algorithm is compared with standard binary (the *iterated conditional mode* algorithm) and non binary (penalized approach with convex potentials Gibbs random fields) reconstruction techniques.

PACS numbers: 81.70.-q

Submitted to: *Inverse Problems*

keywords : Non destructive evaluation, tomography, binary reconstruction.

1. Introduction

Reconstruction of binary volumes from noisy projections is an attractive tool for non destructive evaluation (NDE) of metal structures [2, 3]. It amounts to the simultaneous detection and sizing of small defects in an homogeneous material. However, such a tomographic reconstruction problem is challenging when only *limited-angle* projections are available, since it is then very hard to recover the geometry of the defects along the average direction of the rays. As far as *in situ* NDE is considered, this unfavorable observation context is often a direct consequence of the operational constraints imposed on the NDE setup. In [1], for instance, the radiographic inspection of in-service pipes in the nuclear industry is considered with a total angular excursion of 30°, see figure 1.

In limited-angle tomography, faithful reconstructions of the 3D attenuation map cannot be expected without adequate prior constraints on the solution [4, 5, 6]. One

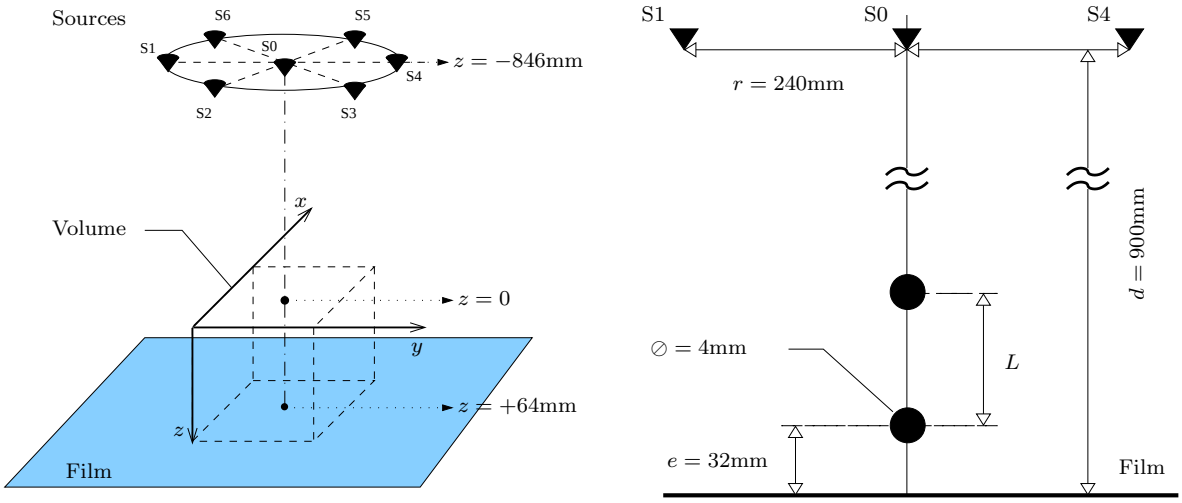

Figure 1. LEFT – 3D representation of the instrumental setup. RIGHT – Schematic representation of the measurement context for the synthetic problem considered in this paper. Note that all the dimensions are drawn from a real NDE setup in the nuclear industry, cf. [1, Fig. 1].

possible approach is to adopt a parametric model for the shape of the defect (for instance, that it is spherical), and to estimate the few corresponding parameters, say, in the least-square sense [7]. However, this approach relies on strong assumptions, which are not natural in the NDE context, where the number and/or the shape of the hypothetical defects are not known. In [8], a genetic-algorithm strategy is proposed to handle an unknown number of defects, but the constraint of a parameterized shape remains. A potentially more flexible extension is obtained using a description of objects in terms of *contours*. Again, parametric contour based methods [9, 10] hardly handle the case of an unknown number of contours, at completly unknown positions. Fully non-parametric modeling of the contour via a *level-set* formulation [11] yields a more suited approach. This leads to gradient-based iterations that aim to compute a local minimizer of the 3D partitioning problem [12].

A large part of the litterature rather assumes a voxel decomposition for the attenuation map. Then, the tomographic reconstruction $\widehat{\boldsymbol{x}}$ is usually defined as the minimizer of a penalized least-square criterion, such as

$$\begin{cases} \min\limits_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2 + \psi(\boldsymbol{x}; \boldsymbol{\theta}) & \text{(1a)} \\ \text{subject to} \quad \boldsymbol{x} \in \mathbb{X} \subseteq \mathbb{R}^N & \text{(1b)} \end{cases}$$

where $\boldsymbol{\theta}$ is a vector of hyperparameters. Within this framework, the solution $\widehat{\boldsymbol{x}}$ is clearly a trade-off: whereas the first term favors the solutions that fit the data, the second one discards the solutions that are not compatible with prior information.

A usual choice is to consider ψ as the energy of a Markov random field with positive interactions, in order to favor clusters of voxels in the reconstruction [13]. More specifically, the well-known *Ising model* is of special interest here, since the

spatial information to encode is binary: each voxel either belongs to an inclusion (of negligible attenuation), or to a sound region of the inspected object (of known attenuation). Only few works have been based on the Ising model to tackle NDE issues. Among the noticeable exceptions, let us mention [3], where it clearly appears that a major numerical obstacle must then be faced, of combinatorial nature, and that no simple, satisfying solution is available. Several authors have rather chosen to relax the binary constraint and to cope with NDE in the framework of continuous-valued Markovian reconstruction [2, 1]. The resulting problem is easier to deal from the numerical viewpoint, since gradient-based optimization algorithms are then available. Unfortunately, the regularizing character of the binary constraint is lost, so that the reconstructed defects appear as streched along some directions [1]. Let us remark here that the resort to Markov models strengthens such a tendancy to produce oversized defects. This motivates our choice of an independent, identically distributed (iid) binary model for the voxels: the binary constraint is then the only regularizing ingredient, while the iid assumption corresponds to the simplest possible spatial model. Although the resulting optimization problem is still intractable under an exact form, it lends itself to a simple and satisfactory solution. Its main features are as follows:

i) Firstly, a *region-of-interest* (ROI) is determined through straightforward computations, such that the maximum posterior mode of voxels outside the ROI is zero, *i.e.*, $\Pr(X_n = 0 \,|\, \boldsymbol{y}) > \Pr(X_n = 1 \,|\, \boldsymbol{y})$. This results in a dramatic reduction in the dimensionality of the problem.

ii) Secondly, we rely on a modified version of the *single most likely replacement* algorithm (SMLR) introduced in [14] to explore efficiently the binary configurations in the ROI. Whereas the original algorithm considers a "single site" update in the volume, the *block most likely replacement* (BMLR) proposed here considers "blockwise" updates within cubes of $2 \times 2 \times 2$ voxels. For limited-angle reconstructions, it is empirically observed that such a strategy allows to escape from local minima where the SMLR would be trapped.

Based on a realistic synthetic NDE problem, our numerical procedure is shown to be very efficient, both in terms of computational load and accuracy of reconstructed attenuation map. Moreover, the proposed method does not depend on any hyperparameter, in contrast with many high-resolution tomographic reconstruction methods.

The paper is organized as follows. In Section 2, the modeling of the tomographic measurement in NDE is introduced, and several Bayesian estimators are envisaged to solve the reconstruction problem This section ends with some remarks concerning the computation of these estimators that motivate the design of a computationally efficient heuristic. Section 3 deals with the construction of the ROI that allows a reduction of the voxels involved in the reconstruction process. In Section 4, the problem of binary object reconstruction within a ROI is stated. Standard binary reconstruction algorithms are introduced and the BMLR is derived. In Section 5, the performances of the BMLR are compared to the standard *iterated conditional mode* algorithm (ICM) and to a penalized

approach with convex potentials Gibbs random fields. Finally, a discussion of this work and some conclusions are presented in Section 6.

2. Binary constraint for NDE

Let $\mathcal{S} = \{1, \cdots, N\}$ be an arbitrary enumeration of the N voxels that constitute the unknown three-dimensional attenuation map $\boldsymbol{x}$. Under usual assumptions, the tomographic measurement along the m-th ray follows the Beer-Lambert relation [15, Section 4.1.1]: for all $m = 1, \cdots, M$,

$$y_m \stackrel{\text{def}}{=} -\log \frac{\tau_m}{\tau_s} = \boldsymbol{h}_{m\bullet}^{\mathrm{t}} \boldsymbol{x} + e_m, \tag{2}$$

where $\boldsymbol{h}_{m\bullet} \in \mathbb{R}_+^N$ stands for the projection process along the mth ray (see for instance [15, Section 7.1]), τ_s and τ_m are the photon emission rate of the source and the counting rate of detector m, respectively. e_m is a noise term that accounts for the statistical fluctuations in the measurement. Stacking all the measurements into a vector $\boldsymbol{y} \stackrel{\text{def}}{=} (y_m)$ yields the following observation model

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{e} \tag{3}$$

where $\boldsymbol{H}$ is a matrix of size $M \times N$, so that $\boldsymbol{h}_{m\bullet}^{\mathrm{t}}$ is the mth line of $\boldsymbol{H}$.

In this paper, it will be assumed that both the geometry and composition of the inspected object are known, up to the possible presence of defects, of course. In more precise terms, we restrict ourselves to the case where the object would be perfectly known if it was free of any defect. Let $\boldsymbol{r} \in \mathbb{R}_+^N$ denote the corresponding defect-free attenuation map, and let us introduce a binary map $\widetilde{\boldsymbol{x}} \in \{0,1\}^N$ to encode the geometry of the defects: voxels that belong to a defect are assigned the value '1', whereas the others are assigned the value '0'. Given that the defects are assumed to be composed of air, of negligible attenuation, the actual attenuation map $\boldsymbol{x}$ is made of voxels such that

$$x_m = \begin{cases} 0 & \text{if } \widetilde{x}_m = 1, \\ r_m & \text{otherwise,} \end{cases}$$

for all $m = 1, \cdots, M$. In compact notations, we have $\boldsymbol{x} = \boldsymbol{r} - \mathrm{diag}(\boldsymbol{r})\widetilde{\boldsymbol{x}}$, so that a new linear observation model can be deduced from (3):

$$\widetilde{\boldsymbol{y}} = \widetilde{\boldsymbol{H}}\widetilde{\boldsymbol{x}} + \widetilde{\boldsymbol{e}},$$

where $\widetilde{\boldsymbol{y}} = \boldsymbol{H}\boldsymbol{r} - \boldsymbol{y}$, $\widetilde{\boldsymbol{H}} = \boldsymbol{H}\mathrm{diag}(\boldsymbol{r})$, and $\widetilde{\boldsymbol{e}} = -\boldsymbol{e}$. Thus, the reconstruction problem can be reformulated as a linear inverse problem with binary unknowns. In the rest of the paper, the tilde signs will be omitted for sake of notational simplicity, so that the attenuation map itself will be considered as binary.

In the sequel, $\boldsymbol{e}$ is assumed to be an outcome of a normal random vector $\boldsymbol{E}$ with zero mean and iid components

$$\boldsymbol{E} \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}) \tag{4}$$

where $\boldsymbol{I}$ is the identity matrix and $\sigma > 0$ is the standard deviation of the noise. Provided that the counting rates at each detector are high and of same magnitude, the Gaussian additive observation model is of sufficient accuracy. These assumptions are realistic in NDE and the Gaussian model will be adopted in this paper. For low signal to noise ratio (SNR), however, a Poisson model is more appropriate to describe the random fluctuations in the measurements, see [16] for details.

Under assumptions (3) and (4), the likelihood of the observation vector reads

$$f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}\,|\,\boldsymbol{x}) \propto \exp\left\{-\frac{1}{2\sigma^2}\,\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|^2\right\} \tag{5}$$

where "$\propto$" stands for "proportional to". Considering the Bayesian framework, a simple iid model for the prior distribution is introduced:

$$\Pr(\boldsymbol{X}=\boldsymbol{x}) = \prod_{n=1}^{N} \Pr(X_n = x_n) = p_1^k\, p_0^{N-k}$$

where $\forall\, n \in \mathcal{S}$, $\Pr(X_n = 1) = p_1$, $\Pr(X_n = 0) = p_0 = 1 - p_1$, and k is the number of voxels that are equal to '1'. Then the posterior probability reads

$$\Pr(\boldsymbol{X}=\boldsymbol{x}\,|\,\boldsymbol{y}) = \frac{f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}\,|\,\boldsymbol{x})\,\Pr(\boldsymbol{X}=\boldsymbol{x})}{\sum_{\boldsymbol{x}'\in\mathbb{S}} f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}\,|\,\boldsymbol{x}')\,\Pr(\boldsymbol{X}=\boldsymbol{x}')} \tag{6}$$

where $\mathbb{S} = \{0,1\}^N$ is the set of the binary configurations. The *maximum a posteriori* (MAP) solution

$$\widehat{\boldsymbol{X}}^{\text{MAP}} \overset{\text{def}}{=} \arg\max_{\boldsymbol{x}\in\mathbb{S}} \Pr(\boldsymbol{X}=\boldsymbol{x}\,|\,\boldsymbol{y}) \tag{7}$$

amounts to minimizing the following quadratic criterion

$$\mathcal{J}(\boldsymbol{x}) \overset{\text{def}}{=} \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|^2 + 2\sigma^2\mu\mathbf{1}^t\boldsymbol{x} \tag{8}$$

under the binary constraint $\boldsymbol{x} \in \mathbb{S}$, where $\mathbf{1} \overset{\text{def}}{=} (1,\cdots,1)^t \in \mathbb{R}^N$ and $\mu \overset{\text{def}}{=} \log(p_0/p_1)$.

Beside the MAP solution, the Bayesian setting allows to define alternate estimators. In particular, the *posterior mean* of voxel n reads

$$E[X_n\,|\,\boldsymbol{Y}] = \sum_{x_n\in\{0,1\}} x_n \Pr(X_n = x_n\,|\,\boldsymbol{y}) = P_n$$

where

$$P_n \overset{\text{def}}{=} \Pr(X_n = 1\,|\,\boldsymbol{y}).$$

Following [17], let us also define the *marginal posterior mode* (MPM) estimator according to:

$$\widehat{X}_n^{\text{MPM}} \overset{\text{def}}{=} \arg\max_{x_n\in\{0,1\}} \Pr(X_n = x_n\,|\,\boldsymbol{y}), \tag{9}$$

which is easy to deduce from P_n:

$$\widehat{X}_n^{\mathrm{MPM}} = \begin{cases} 1 \text{ if } P_n > 1/2 \\ 0 \text{ otherwise.} \end{cases} \tag{10}$$

Unfortunately, none of the latter solutions are easily computable. In particular, (7) is a *binary quadratic programming problem*. In the terms and notations of statistical physics, it amounts to minimize the following energy:

$$H(\boldsymbol{s}) = -\sum_{(m,n)\in\mathcal{E}} J_{mn} s_m s_n + \sum_{n\in\mathcal{S}} F_n s_n, \tag{11}$$

where $s_n = 2x_n - 1 \in \{-1, +1\}$ are Ising spins, $J_{mn} = -\boldsymbol{h}_{\bullet m}^{\mathrm{t}} \boldsymbol{h}_{\bullet n}$,

$$\mathcal{E} = \left\{(m,n) \in \mathcal{S}^2, m < n, \boldsymbol{h}_{\bullet m}^{\mathrm{t}} \boldsymbol{h}_{\bullet n} \neq 0\right\}$$

is the set of interacting spins, and

$$F_n = -\boldsymbol{h}_{\bullet n}^{\mathrm{t}} \left(2\boldsymbol{y} - \sum_{m\in\mathcal{S}} \boldsymbol{h}_{\bullet m}\right) + 2\sigma^2\mu$$

plays the role of the nth component of an exterior magnetic field. In some particular cases, like the one encountered in binary image denoising [18], solving such a problem can be done in an acceptable number of operations, that is a polynomial function of N. On the contrary, our case falls within the class of NP-hard problems, for which it is very unlikely that a polynomial algorithm exists [19]. The first reason for NP-hardness is the fact that J_{mn} takes negative values (*i.e.*, the interactions are antiferromagnetic) in a nonzero magnetic field. Moreover, the non-planarity of the neighborhood graph $\mathcal{E}$ is another source of NP-hardness [20].

On the other hand, computing $\widehat{X}_n^{\mathrm{MPM}}$ requires the computation of partition functions involving summations over $\mathbb{S}$ or large subsets of $\mathbb{S}$. More specifically, the evaluation of the key quantity P_n can be envisaged as follows. Let $\mathbb{S}_n^i \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \mathbb{S} : x_n = i\}$, so that

$$P_n = \sum_{\boldsymbol{x}\in\mathbb{S}_n^1} \Pr(\boldsymbol{X} = \boldsymbol{x} \,|\, \boldsymbol{y}) = \frac{\sum_{\boldsymbol{x}\in\mathbb{S}_n^1} \Pr(\boldsymbol{X} = \boldsymbol{x} \,|\, \boldsymbol{y})}{\sum_{\boldsymbol{x}\in\mathbb{S}} \Pr(\boldsymbol{X} = \boldsymbol{x} \,|\, \boldsymbol{y})} = \frac{1}{1 + R_n(\boldsymbol{y})},$$

where

$$R_n(\boldsymbol{y}) \stackrel{\text{def}}{=} \frac{\sum_{\boldsymbol{x}\in\mathbb{S}_n^0} \Pr(\boldsymbol{X} = \boldsymbol{x} \,|\, \boldsymbol{y})}{\sum_{\boldsymbol{x}\in\mathbb{S}_n^1} \Pr(\boldsymbol{X} = \boldsymbol{x} \,|\, \boldsymbol{y})}. \tag{12}$$

Again, the computation of the sums entering the expression of R forms NP-hard problems [19, 20].

Despite the NP-hardness of the problem, some contributors resort to Markov Chain Monte-Carlo (MCMC) technics to solve such combinatory problems. In [17], a Gibbs sampler is used to compute the posterior mean estimator for binary denoising problems. In [21], the authors resort to simulated annealing in order to minimize a least-square

criterion under binary contraint. By nature, such MCMC technics converge in the asymptotic limit. When strong spatial correlations exist in the posterior probability, an efficient sampling is difficult and the resulting algorithms may ask for unrealistic computation time before a useful estimate is available. In practice, we found that this problem is a major obstacle in designing MCMC based techniques for binary reconstruction in limited-angle tomography. Hence, our main concern in this paper is to develop an efficient heuristic that computes a local maximizer of $\Pr(\boldsymbol{X} = \boldsymbol{x} \,|\, \boldsymbol{y})$ in a finite number of iterations. For this purpose, our strategy is twofold. Firstly, we aim at reducing the dimensionnality of the problem by an appropriate selection of the voxels that should be discared in the reconstruction process, see Section 3. Then, a local maximizer of $\Pr(\boldsymbol{X} = \boldsymbol{x} \,|\, \boldsymbol{y})$ is computed via a deterministic iterative procedure, see Section 4.

3. Region-of-interest (ROI) selection

As far as large 3D problems are considered, a first step toward a computationally attractive algorithm is to reduce the number of voxels involved in the reconstruction. As shown in the next section, it can be established that $\widehat{X}_n^{\mathrm{MPM}} = 0$ for a large amount of voxels, from costless preprocessing operations involving the data.

3.1. MPM based ROI in the Gaussian data case

Although P_n is not straightforward to calculate, it admits an easily computable upper bound $\overline{P}_n$ that will be fruitful to manipulate.

Let $\mathbf{0}$ denote the null vector of length N and $\boldsymbol{e}_n$ the nth canonical vector of $\mathbb{R}^N$. Indeed, the following proposition holds in the case of centered white Gaussian observation noise.

Proposition 1 *The posterior probability P_n admits the following upper bound:*

$$P_n \leq \overline{P}_n \stackrel{\text{def}}{=} \frac{1}{1 + \underline{R}_n(\boldsymbol{y})} \tag{13}$$

where

$$\underline{R}_n(\boldsymbol{y}) \stackrel{\text{def}}{=} \frac{\Pr(\boldsymbol{X} = \mathbf{0} \,|\, \boldsymbol{y})}{\Pr(\boldsymbol{X} = \boldsymbol{e}_n \,|\, \boldsymbol{y})} \tag{14}$$

$$= \exp\left(-\overline{\Delta}_n\right), \tag{15}$$

with

$$\overline{\Delta}_n \stackrel{\text{def}}{=} \frac{1}{2\sigma^2}\left(2\boldsymbol{h}_{\bullet n}^{\mathrm{t}}\boldsymbol{y} - \|\boldsymbol{h}_{\bullet n}\|^2\right) - \mu. \tag{16}$$

Proof: See Appendix A. ∎

Interestingly, it is clear from (10) that

$$\overline{P}_n \leq 1/2 \implies \widehat{X}_n^{\mathrm{MPM}} = 0. \tag{17}$$

The computation of $\overline{P}_n$ plays a pivotal role in our method since it allows a significant reduction of the effective dimension of the reconstruction problem. Let us introduce the set $\mathcal{I}$ of voxels for which implication (17) is inactive, *i.e.*,

$$\mathcal{I} = \{n \in \mathcal{S} : \overline{P}_n > 1/2\}.$$

The latter set defines a ROI in the sense that the reconstruction algorithm will be restricted to $\mathcal{I}$, while the other voxels will be assumed to match their MPM estimation, *i.e.*, $\widehat{X}_n^{\mathrm{MPM}} = 0$, $\forall n \in \mathcal{S} \setminus \mathcal{I}$.

Let us remark that

$$\begin{aligned} \overline{P}_n \leq 1/2 &\iff \overline{\Delta}_n \leq 0 \\ &\iff \boldsymbol{h}_{\bullet n}^{\mathrm{t}} \boldsymbol{y} \leq \frac{1}{2} \|\boldsymbol{h}_{\bullet n}\|^2 + \sigma^2 \mu \end{aligned} \tag{18}$$

so that a threshold test on the backprojection $\boldsymbol{H}^{\mathrm{t}}\boldsymbol{y}$ of the data allows to determine $\mathcal{I}$, where the threshold value is generally not spatially invariant. Moreover, the test does not depend on the noise variance provided that the binary states are considered equally probable, *i.e.*, $p_0 = p_1$.

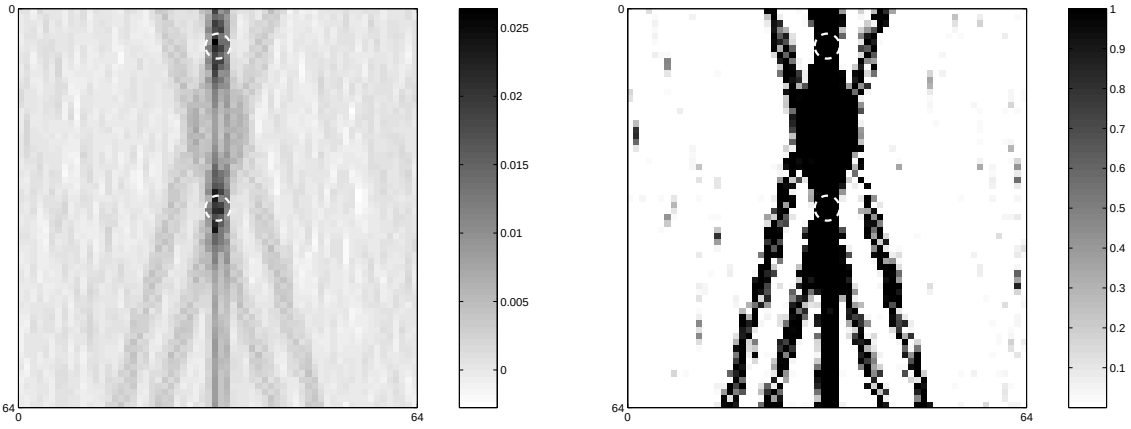

Figure 2. Central 2D section of the backprojection (left) and of the map of the upper bound $\overline{P}_n$ (right), repectively, obtained from noisy projections ($\sigma = 0.01$), for a pair of superimposed defects ($L = 26$mm).

A direct extension of Proposition 1 can be established in the general Gaussian case of a noise with mean $\boldsymbol{m}$ and covariance $\boldsymbol{R}$. A less trivial but richer extension is provided in Subsection 3.3 to encompass some cases of log-concave likelihoods.

Finally, let us note that Proposition 1 admits a formal counterpart, which yields a sufficient condition to ensure $\widehat{X}_n^{\mathrm{MPM}} = 1$. The latter condition is hardly met in our situation, because too few voxels are equal to one, but it would be interesting to implement Condition (18) in more balanced situations. The proof is omitted, since it is almost a paraphrase of that of Proposition 1.

Proposition 2 *The posterior probability P_n admits the following lower bound:*

$$P_n \geq \underline{P}_n \stackrel{\text{def}}{=} \frac{1}{1+\overline{R}_n(\boldsymbol{y})}$$

where

$$\overline{R}_n(\boldsymbol{y}) \stackrel{\text{def}}{=} \frac{\Pr(\boldsymbol{X}=\mathbf{1}-\boldsymbol{e}_n \,|\, \boldsymbol{y})}{\Pr(\boldsymbol{X}=\mathbf{1} \,|\, \boldsymbol{y})} = \exp\left(-\underline{\Delta}_n\right),$$

with

$$\underline{\Delta}_n \stackrel{\text{def}}{=} \frac{1}{2\sigma^2}\left(2\boldsymbol{h}_{\bullet n}^{\mathrm{t}}\left(\boldsymbol{y}-\boldsymbol{H}\mathbf{1}\right)+\|\boldsymbol{h}_{\bullet n}\|^2\right)+\mu.$$

3.2. Synthetic example

For illustrative purpose, let us compute the threshhold value in a simple, yet realistic context. The inspected object is made of an homogeneous material, with two homogeneous defects of spherical shape. Its discrete representation contains $64\times64\times64$ voxels, while each of the two defects spreads over 32 voxels. Some schematic illustrations of the measurement context are provided in Fig. 1. The data $\boldsymbol{y}$ consists in seven 2D projections (each one being a 128 $\times$ 128 image) that were corrupted by zero mean iid Gaussian perturbations with standard deviation $\sigma = 0.01$, which is close to the maximum amplitude of the projection of a single voxel. Both the ROI $\mathcal{I}$ and the backprojection were deduced from this data set. The former was computed with the additional assumption that $p_1 = p_0$ (*i.e.*, $\mu = 0$).

Fig. 2(a) and (b) depict the "central" 2D cross-section of the backprojection $\boldsymbol{H}^{\mathrm{t}}\boldsymbol{y}$ and of the upper bound $\overline{P}_n$, respectively.

In this case, 17.3% of the 64×64 voxels of the central cross-section happen to belong to the ROI, and only 1.5% of the $64\times64\times64$ voxels of the object belong to it, while the ROI incorporates all the 64 voxels of the defects. This example shows that ROI selection can lead to an impressive reduction of the problem dimension. However, it is worth to stress that such a preliminary operation does not raise the main ambiguity concerning the position of the defects, since the voxels that belong to the ROI are organized along the vertical axis, which is the average direction of the rays. It is only aimed at lowering the dimensionality of the reconstruction problem, not at solving it. However, as far as combinatory issues are involved, the size of the vector of unkowns is known as a crucial parameter.

From (18), it should be clear that the achievable reduction depends on the number of flaws in the volume, the noise in the data, and the number and the angular excursion of the projections. Therefore, it is not easy to assess quantitative results concerning the size of the ROI. However, our experience indicates that increasing the number of projections and/or their angular excursion decreases the size of the ROI. Likewise, a smaller SNR yields a larger ROI.

3.3. Generalization to log-concave likelihoods

Proposition 1 can actually be stated in more general terms, provided that the observations $\boldsymbol{y}$ be independent given the object $\boldsymbol{x}$, and that the likelihood of y_m be a log-concave function‡ of $\boldsymbol{h}_{m\bullet}^{\mathrm{t}}\boldsymbol{x}$.

Proposition 3 *Let the likelihood function take the following form*

$$f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}\,|\,\boldsymbol{x}) = \prod_{m=1}^{M} \phi_m(\boldsymbol{h}_{m\bullet}^{\mathrm{t}}\boldsymbol{x}), \tag{19}$$

where ϕ_m are log-concave functions. Then (13)-(14) *still hold.*

Proof: See Appendix B. ∎

The main non Gaussian model we have in mind is the Poisson data model, which is the acknowledged model to describe that the measurements are related to a counting process. In the present case of transmission tomography, the corresponding data likelihood reads

$$\Pr(\boldsymbol{Y} = \boldsymbol{y}\,|\,\boldsymbol{X} = \boldsymbol{x}) = \prod_{m=1}^{M} \frac{(z_m)^{y_m}}{y_m!} \exp(z_m)\,. \tag{20}$$

In the latter expression, $z_m = \tau_s \exp(-\xi_m)$, where τ_s is the emission rate of the source and ξ_m is related to the attenuation along the ith ray according to

$$\xi_m = \boldsymbol{h}_{m\bullet}^{\mathrm{t}}\boldsymbol{x} + g_m,$$

where $\boldsymbol{g} = (g_m)$ stands for a background component that accounts for the Compton scattering effect. Taking the scattering into account ($g_m \neq 0$) leads to non-logconcave functions. However, when the scattering can be neglected the Poisson likelihood (20) is a product of log-concave functions of $\boldsymbol{x}$ and Proposition 3 applies with the corresponding expression for $\underline{R}_n(\boldsymbol{y})$

$$\underline{R}_n(\boldsymbol{y}) = \prod_{m=1}^{M} \exp\left(\tau_s\left(\exp(-h_{mn}) - 1\right) + y_m h_{mn}\right). \tag{21}$$

Here we shall not investigate the Poisson data model further since our main concern is to consider gamma or X ray tomographic measurements with high counting rates, which are well approximated by the simple Gaussian described in Section 2.

4. Reconstruction using a BMLR approach

In the case of limited-angle tomography, designing a computationaly attractive algorithm that leads to high quality binary reconstructions is a challenge. Here, the proposed approach is twofold. Firstly, the reconstructed volume is reduced to the ROI as defined

‡ By definition, ϕ is a log-concave function if $\log\phi$ is a concave function.

in the previous section. Our aim is then to minimize the penalized least-square criterion (8) with respect to the binary voxels that pertain to the ROI, while the others are forced to zero. For this purpose, we propose a deterministic procedure called block most likely replacement (BMLR). It is an heuristic procedure comparable to ICM. However, it provides significantly better results than ICM, while being still of a moderate computational cost for the considered NDE problems.

4.1. ROI constrained formulation

Let us introduce the set $\mathbb{I} = \{\boldsymbol{x} \in \mathbb{S} : x_n = 0 \text{ if } n \notin \mathcal{I}\}$. Hereafter, the assumption $\boldsymbol{x} \in \mathbb{I}$ is taken for granted. It is then natural to consider the problem of maximizing (8) under the constraint $\boldsymbol{x} \in \mathbb{I}$. Let us remark here that such a constrained formulation is not necessarily equivalent to the unconstrained one. Actually, it is easy to check that the two problems admit the same solution if $\widehat{\boldsymbol{X}}^{\text{MAP}} \in \mathbb{I}$ only. Unfortunately, it is far from easy to check the latter condition since $\widehat{\boldsymbol{X}}^{\text{MAP}}$ is unknown. Nonetheless, maximizing (8) under the constraint $\boldsymbol{x} \in \mathbb{I}$ remains coherent and likely to produce a meaningful solution. Moreover, it can be checked by practice that the proposed maximization procedure with or without the ROI provides close reconstruction results — see Section 5.

4.2. ICM and SMLR heuristics

>From the sake of low computational complexity, the ICM procedure introduced in [13] is of particular interest since it amounts to simple successive scalar relaxation of voxels. Let $\mathcal{N} \stackrel{\text{def}}{=} \{n_1, \cdots, n_{|\mathcal{I}|}\}$ be an arbitrary permutation of the set $\mathcal{I}$. From any initial solution $\boldsymbol{x}^{(1)} \in \mathbb{I}$, the kth iteration of ICM can be formally defined by

$$\begin{aligned} \boldsymbol{x}^{(k,1)} &= \boldsymbol{x}^{(k)} \\ \ell = 1 \cdots |\mathcal{I}|, \quad \boldsymbol{x}^{(k,\ell+1)} &= \begin{cases} \boldsymbol{\zeta}^{(k,\ell)} & \text{if } \mathcal{J}(\boldsymbol{\zeta}^{(k,\ell)}) < \mathcal{J}(\boldsymbol{x}^{(k,\ell)}), \\ \boldsymbol{x}^{(k,\ell)} & \text{otherwise,} \end{cases} \\ \boldsymbol{x}^{(k+1)} &= \boldsymbol{x}^{(k,|\mathcal{I}|+1)}, \end{aligned}$$

where $\boldsymbol{\zeta}^{(k,\ell)}$ is a *proposition vector* defined by

$$\boldsymbol{\zeta}^{(k,\ell)} \stackrel{\text{def}}{=} \boldsymbol{x}^{(k,\ell)} \oplus \boldsymbol{e}_{n_\ell}, \tag{22}$$

and $\oplus$ stands for the componentwise *logical exclusive or*.

The ICM iterations can be implemented efficiently in a recursive fashion, as follows. Let $\boldsymbol{x}$ and $\boldsymbol{x}'$ be two binary N-length vectors that only differ at index n, *i.e.*, $x'_n = x_n \pm 1$. Then, it is easy to establish that

$$\mathcal{J}(\boldsymbol{x}') = \mathcal{J}(\boldsymbol{x}) + \boldsymbol{h}_{\bullet n}^{\mathrm{t}} (\boldsymbol{h}_{\bullet n} \mp 2\boldsymbol{\varepsilon}) \pm 2\sigma^2\mu, \tag{23}$$

where $\boldsymbol{\varepsilon} = \boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}$. Moreover, we have also

$$\boldsymbol{\varepsilon}' = \boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}' = \boldsymbol{\varepsilon} \mp \boldsymbol{h}_{\bullet n} \tag{24}$$

In the implementation of ICM, Equations (23) and (24) can be used to compare $\mathcal{J}(\boldsymbol{\zeta}^{(k,\ell)})$ to $\mathcal{J}(\boldsymbol{x}^{(k,\ell)})$ and to recursively compute $\boldsymbol{\varepsilon}^{(k,\ell)} = \boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}^{(k,\ell)}$, respectively.

By construction, the ICM ensures that the criterion $\mathcal{J}$ never increases. Moreover, convergence occurs after a finite number of iterations, as soon as $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)}$. However, an obvious drawback of this heuristic is that the solution is not invariant with respect to the scanning order in $\mathcal{I}$.

The latter limitation of ICM can be made for by a *single most likely replacement* (SMLR) procedure, as introduced by Kormylo and Mendel [14] in the field of sparse spike train restoration. According to SMLR, each scan yields the swap of only one voxel, the one that produces the largest decrease of the criterion:

$$\boldsymbol{x}^{(k+1)} = \underset{\boldsymbol{x}\in\mathbb{V}(\boldsymbol{x}^{(k)})}{\arg\max}\, \mathcal{J}(\boldsymbol{x}), \tag{25}$$

where $\mathbb{V}(\boldsymbol{x}^{(k)}) = \{\boldsymbol{\zeta}^{(k,\ell)} : 1 \leq \ell \leq |\mathcal{I}|\}$ is the set of all proposition vectors generated by elementary modifications of $\boldsymbol{x}^{(k)}$. SMLR shares several features of ICM:

- An efficient implementation is obtained thanks to the recursive expressions (23) and (24).
- The criterion $\mathcal{J}$ never increases, and convergence occurs after a finite number of iterations, as soon as $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)}$.

Although very attractive from the computational side, none of these heuristics provide faithfull reconstructions in our context. In particular, we have found that the set $\mathbb{V}(\boldsymbol{x}^{(k)})$ is not rich enough to prevent ICM and SMLR to get stuck into bad solutions. To gain efficiency, richer sets of propositions should prioritarily allow neighboring voxels to swap their values, so that the position of the hypothetical defects can adjust more easily along the iterations. A possible solution would be to adapt Chi and Mendel's single-spike-shift detector [22]. We found more appropriate to develop an original variation of it, which is developed in the next section.

4.3. The BMLR heuristic

Whereas the SMLR only explores the two possible values of a single voxel at each step, the new heuristic explores the $2^8 = 256$ configurations of a basic $2 \times 2 \times 2$ cube of voxels in the ROI. The resulting procedure is called *block most likely replacement* (BMLR).

4.3.1. Formal expression of the heuristic Let Ω stands for the set of nonempty intersections between the ROI and the $2 \times 2 \times 2$ cubes. Hereafter, such intersections will be called *blocks*. Most of the blocks contain eight voxels, but some of those located at the boundaries of the ROI may contain a fewer number of voxels.

In what follows, the compact notation $\boldsymbol{x}_{\boldsymbol{\omega}}$ is employed for the subvector $\{x_n, n \in \boldsymbol{\omega}\}$, where $\boldsymbol{\omega}$ is a set of sites, *i.e.*, a subset of $\{1, \ldots, N\}$. Let us also introduce the notation

$\boldsymbol{\zeta}^{(k,\boldsymbol{\omega},\boldsymbol{b})}$ for the proposition vector defined by

$$\begin{aligned}\boldsymbol{\zeta}_{\boldsymbol{\omega}}^{(k,\boldsymbol{\omega},\boldsymbol{b})} &= \boldsymbol{x}_{\boldsymbol{\omega}}^{(k)} \oplus \boldsymbol{b},\\ \boldsymbol{\zeta}_{\mathcal{I}\setminus\boldsymbol{\omega}}^{(k,\boldsymbol{\omega},\boldsymbol{b})} &= \boldsymbol{x}_{\mathcal{I}\setminus\boldsymbol{\omega}}^{(k)},\end{aligned}$$

where $\boldsymbol{\omega} \in \Omega$ and $\boldsymbol{b} \in \{0,1\}^{|\boldsymbol{\omega}|}$. Then, the updating equation (25) still holds for the BMLR heuristic, with an extended set of proposition vectors:

$$\mathbb{V}(\boldsymbol{x}^{(k)}) = \left\{\boldsymbol{\zeta}^{(k,\boldsymbol{\omega},\boldsymbol{b})} : \boldsymbol{\omega} \in \Omega,\, \boldsymbol{b} \in \{0,1\}^{|\boldsymbol{\omega}|}\right\}.$$

Again, the criterion $\mathcal{J}$ never increases, and convergence occurs after a finite number of iterations, as soon as $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)}$. However, the examination of each block $\boldsymbol{\omega}$ requires $2^{|\boldsymbol{\omega}|}$ evaluations of the criterion, where the bloc-size $|\boldsymbol{\omega}|$ varies between 1 and 8. The resulting computational cost can be prohibitive if the implementation issue is not carefully dealt. The next subsection proposes an efficient solution based on Gray codes.

4.3.2. Efficient implementation using Gray codes A Gray code (or *reflected* code) is a binary numeral system where two successive values differ in only one digit [23, Section 20.2]:

$$(0, 1, 11, 10, 110, 111, 101, 100, 11000, 11001, \ldots)$$

Let us remark that the first 2^n elements are obtained by some permutation of the first 2^n natural numbers.

For each block $\boldsymbol{\omega}$, we propose to explore the $2^{|\boldsymbol{\omega}|}$ binary configurations in the Gray code order rather than in the usual one. Then, two successive trials $\boldsymbol{b}$ and $\boldsymbol{b}'$ respectively correspond to two vectors $\boldsymbol{\zeta}^{(k,\boldsymbol{\omega},\boldsymbol{b})}$ and $\boldsymbol{\zeta}^{(k,\boldsymbol{\omega},\boldsymbol{b}')}$ that only differ at a single index. As a consequence, the recursive expressions (23) and (24) can be applied to explore the $2^{|\boldsymbol{\omega}|}$ values of $\mathcal{J}(\boldsymbol{\zeta}^{(k,\boldsymbol{\omega},\boldsymbol{b})})$ at an affordable numerical cost, for all $\boldsymbol{\omega} \in \Omega$.

An even more efficient version consists in introducing $f_n = \boldsymbol{h}_{\bullet n}^{\mathrm{t}}\boldsymbol{\varepsilon}$ and $g_{mn} = \boldsymbol{h}_{\bullet m}^{\mathrm{t}}\boldsymbol{h}_{\bullet n}$ for all $m, n \in \boldsymbol{\omega}$, so that (23) can be replaced by

$$\mathcal{J}(\boldsymbol{x}') = \mathcal{J}(\boldsymbol{x}) + g_{nn} \mp 2f_n \pm 2\sigma^2\mu,$$

while the $|\boldsymbol{\omega}|$ scalars f_n can be recursively computed using

$$f'_m = \boldsymbol{h}_{\bullet n}^{\mathrm{t}}\boldsymbol{\varepsilon}' = f_m \mp g_{mn}.$$

While each of the $2^{|\boldsymbol{\omega}|}$ applications of (23) involves a scalar product, none is required anymore in the new recursion. Only $|\boldsymbol{\omega}|\,(|\boldsymbol{\omega}|+1)/2$ scalar products are required to compute g_{mn} for all $m, n \in \boldsymbol{\omega}$.

Table 1 depicts the pseudo-code for the central part of the BMLR scheme. For the sake of brevity, the set of blocks Ω is assumed available from previous calculation, as well as the projection matrix $\boldsymbol{H}$. For large size problems, $\boldsymbol{H}$ cannot be stored and its entries must rather be recalculated when required.

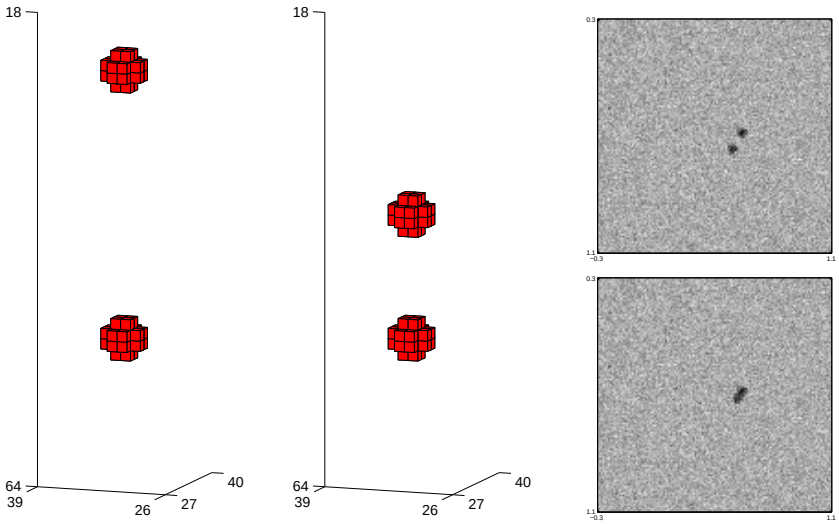

Figure 3. LEFT – Pairs of voxelized spheres considered in the data generation process, in configuration `separated` ($L = 26$mm) and `close` ($L = 12$mm), respectively. RIGHT – Corresponding noisy projection from source S6 with noise standard deviation $\sigma = 0.01$.

In Table 1, the initial configuration is the null object $\boldsymbol{x}^0 = \mathbf{0}$. However, a different initialization point could be easily considered. In particular, a faster but less optimal method could provide an initial solution, *e.g.,* ICM, that would be refined using BMLR. For the sake of computational time, such a two-step approach would be of particular interest to reconstruct objects of larger extent than localized defects.

5. Test on synthetic data sets

The 3D synthetic NDE problem presented in Section 3.2 is now considered in order to test the proposed reconstruction algorithm. The aim is to illustrate the detection, sizing and separation capabilities of the method on the ground of simulations.

5.1. The simulation context

The considered reconstruction problem involves two superimposed defects of spherical shape, in an homogeneous material. The volume of interest contains $64 \times 64 \times 64$ voxels, while each of the two defects spreads over 32 voxels. Two different values have been considered for the distance between the two defects: $L \in \{26\text{mm}, 12\text{mm}\}$ (see Figure 3(a)). The data set $\boldsymbol{y}$ consists in seven projections that were corrupted by an uncorrelated Gaussian noise with standard deviation $\sigma \in \{0, 0.005, 0.01\}$ (see Figure 3(b)). The geometric dimensions of the problem are given by Figure 1. They correspond to a realistic situation of in-service pipe inspection in the nuclear industry [1, Fig. 1].

Table 1. Pseudo-code of BMLR.

$K \leftarrow 8$; $G_1 \leftarrow 1$; $\lambda \leftarrow 2\sigma^2\mu$
for $i = 1 : K-1$ **do**
  $k \leftarrow 2^i$; $G_k \leftarrow i+1$ % $G \in \{0,7\}$: position of
  for $j = 1 : k-1$ **do** % changing bit between
    $G_{k+j} \leftarrow G_{k-j}$ % successive elements in
  end for % Gray code of length 2^8
end for
$\boldsymbol{x}^0 \leftarrow \mathbf{0}; \boldsymbol{\varepsilon}^0 \leftarrow \boldsymbol{y}; \mathcal{J}^0 \leftarrow \|\boldsymbol{y}\|^2$ % Initialization
repeat % Main loop
  $\mathcal{J}_{\max} \leftarrow \mathcal{J}^0$;
  for all $\boldsymbol{\omega} \in \Omega$ **do** % Loop on all blocks
    for all $n \in \boldsymbol{\omega}$ **do**
      $f_n \leftarrow \boldsymbol{h}_{\bullet n}^{\mathrm{t}} \boldsymbol{\varepsilon}$
      $g_{nn} \leftarrow \boldsymbol{h}_{\bullet n}^{\mathrm{t}} \boldsymbol{h}_{\bullet n}$
      for all $m \in \boldsymbol{\omega}, m < n$ **do**
        $g_{mn} \leftarrow g_{nm} \leftarrow \boldsymbol{h}_{\bullet m}^{\mathrm{t}} \boldsymbol{h}_{\bullet n}$
      end for
    end for
    $\mathcal{J} \leftarrow \mathcal{J}^0$; $\boldsymbol{x} \leftarrow \boldsymbol{x}^0$
    for $k = 0 : 2^{|\omega|} - 1$ **do** % Loop on all block states
      $n \leftarrow \boldsymbol{\omega}_{G_k}$
      $s \leftarrow 2x_n - 1$ % $s = 1$ if $x_n = 1 \rightarrow 0$,
      $x_n \leftarrow 1 - x_n$ % $s = -1$ if $x_n = 0 \rightarrow 1$
      $\mathcal{J} \leftarrow \mathcal{J} + g_{nn} + s(2f_n - \lambda)$
      for all $m \in \boldsymbol{\omega}$ **do**
        $f_m \leftarrow f_m + s g_{mn}$
      end for
      if $\mathcal{J} < \mathcal{J}_{\max}$ **then** % New best configuration
        $\mathcal{J}_{\max} \leftarrow \mathcal{J}$; $n_{\max} \leftarrow n$; $s_{\max} \leftarrow s$
      end if
    end for % End of loop on block states
  end for % End of loop on blocks
  if $\mathcal{J}_{\max} < \mathcal{J}^0$ **then** % New configuration
    $\mathcal{J}^0 \leftarrow \mathcal{J}_{\max}$; $x^0_{n_{\max}} \leftarrow 1 - x^0_{n_{\max}}$; $\boldsymbol{\varepsilon}^0 \leftarrow \boldsymbol{\varepsilon}^0 + s_{\max} \boldsymbol{h}_{\bullet n}$
  end if
until $\mathcal{J}_{\max} = \mathcal{J}^0$ % End of main loop. The best
% found configuration is $\boldsymbol{x}^0$.

5.2. Non-binary reconstruction results

Here, a non-binary penalized approach tested in [1] is used as a comparative tool. It consists of considering an optimization problem of the form (1) with $\mathbb{X} = \mathbb{R}_+^N$ and

$$\psi(\boldsymbol{x};\boldsymbol{\theta}) \;=\; \lambda \sum_{i\sim j} \phi(x_i - x_j;\delta) + \alpha \sum_{n=1}^{N} |x_n| \,, \tag{26}$$

where $\lambda, \alpha \geq 0$, $i \sim j$ stands for all pairs of neighboring sites in the first-order neighborood system (hence, each voxel has six neighbors) and $\phi(t;\delta)$ is a Huber function [24]. The role of such a penalty function is to favor piecewise homogeneous solutions. In practice, the penalized approach requires the tuning of the three hyperparameters $\boldsymbol{\theta} = (\lambda, \delta, \alpha)^{\mathrm{t}}$. Here, they have been empirically adjusted according to qualitative appreciation. The corresponding reconstruction results are depicted by Fig. 4. As expected, the estimated attenuations are underestimated, while the extension of the defects is overestimated along the vertical axis. In particular, the result does not allow to clearly discriminate the two spheres in the case $L = 12$mm.

The penalized least-square criterion being convex, the global minimizer can be computed from an initial guess $\boldsymbol{x}^{(0)}$ by an appropriate iterative descent algorithm. Here, $\boldsymbol{x}^{(0)}$ has been chosen as the backprojection of the data and the minimization step was performed using a projected-gradient method, *i.e.*, a straightforward adaptation of *steepest-descent* that handles the positivity constraint in a natural way [25, p.203]. However, this algorithm inherits the slow asymptotic convergence of a gradient method. Table 2 shows that several hours are required to reach the adopted convergence criterion $\psi(\boldsymbol{x}^{(k-1)}) - \psi(\boldsymbol{x}^{(k)}) < 10^{-6}$, where $k \in \mathbb{N}$ the iteration number. Let us remark here that the same constrained minimization problem could probably be adressed using more efficient, but more complex algorithms (see for instance [26]).

5.3. Binary reconstruction results

In this section, the proposed BMLR algorithm is compared to the ICM. In all cases, the initial state is chosen as a defect-free object, *i.e.*, $\boldsymbol{x}^{(0)} = \mathbf{0}$. Both heuristics require the tuning of the hyperparameter $\mu = \log(p_0/p_1)$. Since the defects are expected to constitute a very small part of the total volume, a strict Bayesian viewpoint would lead to choose prior probabilities p_0 and $p_1 = 1 - p_0$ such that $p_0 \gg p_1$. For two distinct reasons, we rather set $p_0 = p_1$.

- On the one hand, false positives are preferable to false negatives, which justifies to overvalue p_1/p_0. This could be reformulated more rigorously according to the Bayesian cost theory, by replacing the MAP estimator by another Bayesian estimator that accounts for distinct costs for false negatives and false positives, respectively c_0 and c_1. Then our choice corresponds to the assumption that $p_0/c_0 = p_1/c_1$.
- On the other hand, choosing $p_0 = p_1$ leads to a fully unsupervised version of BMLR.

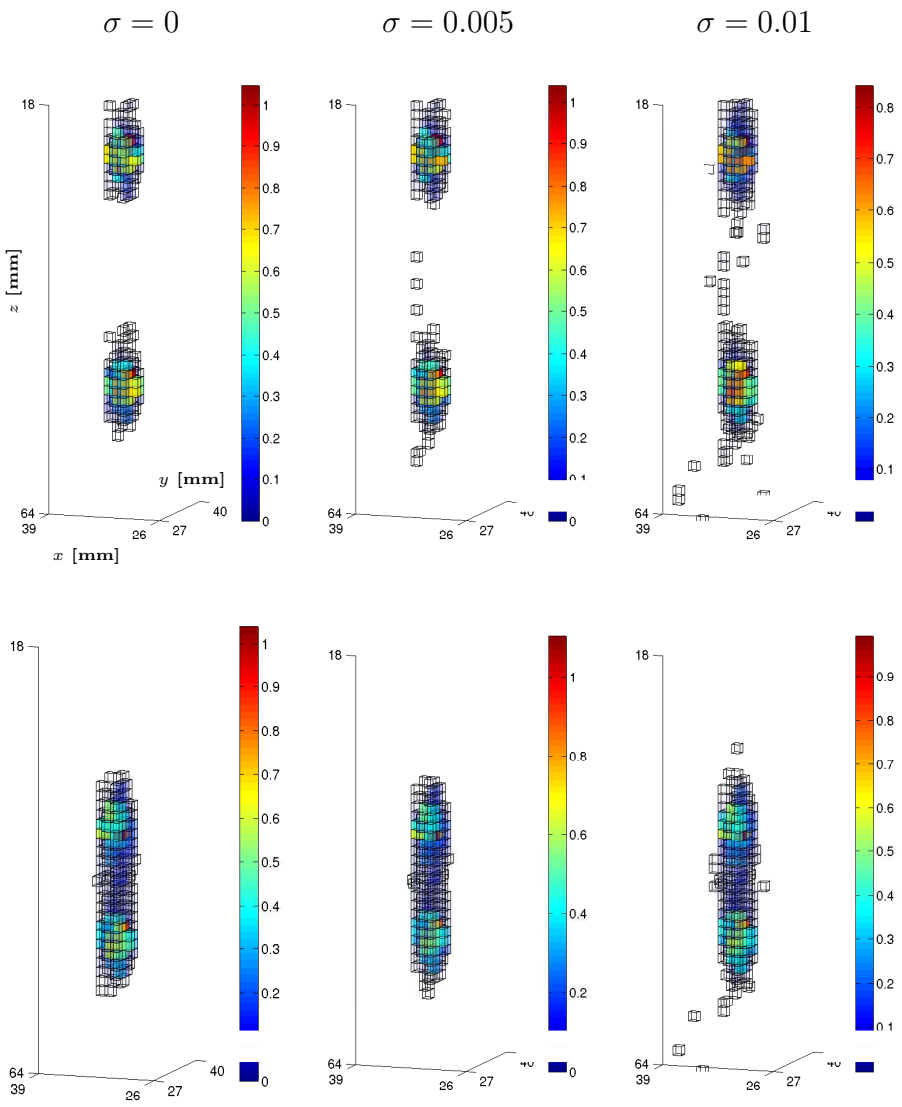

Figure 4. Reconstructions from noisy projections of a pair of voxelized spheres using a penalized approach.

The reconstruction results are shown in Figs. 5 and 6 for the ICM and the BMLR algorithms, respectively. Even in the noiseless case, the ICM fails to provide a faithful reconstruction in the case of two close spheres. In comparison, the BMLR algorithm achieves a perfect reconstruction in all cases of high or moderate SNR.

On the other hand, the number of false positive voxels is subject to a threshold effect: when the noise standard deviation reaches the magnitude of the projection of a single voxel (whose maximum value is $\sigma_{\text{th}} = 1/64 \approx 0.015$ in the present simulation), many spurious, isolated voxels appear in the reconstruction. This is clearly visible on Fig. 6 in the case $\sigma = 0.01$. In this case, two distinct strategies may be adopted to reduce the number of false positive voxels.

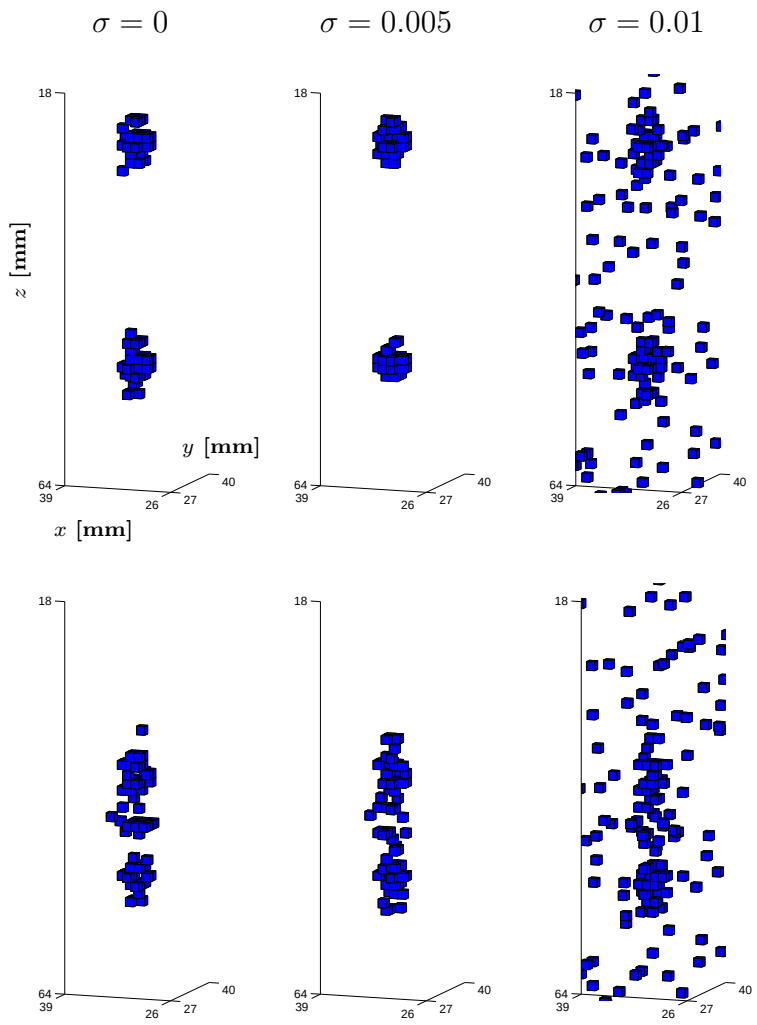

Figure 5. Reconstructions from noisy projections of a pair of voxelized spheres using an ICM approach.

The first one rests upon the fact that the BMLR reconstruction remains acceptable, let alone those spurious voxels. As far as real defects can be assumed to extend over several neighboring voxels, the idea is then to remove isolated voxels by simple post-processing. For sake of computational burden, it is interesting to also discard isolated voxels from the ROI. In the considered cases where $\sigma = 0.01$, a large amount of iterations are spent to activate voxels that are isolated in the ROI, and can thus be saved. The resulting reconstruction is depicted in the first column of Fig. 8.

An alternate strategy is to introduce a penalization of active voxels by assigning a positive value to the regularization parameter $\lambda \overset{\text{def}}{=} 2\sigma^2\mu$ in criterion (8). The tuning of λ has a great impact on the final quality of the 3D attenuation map. As illustrated in Fig. 9, a correct tuning of λ significantly reduces the number of spurious voxels and produces a satisfactory reconstruction of the flaws. However, over estimating λ leads to false negative voxels, so that the tuning of λ must be cautiously handled.

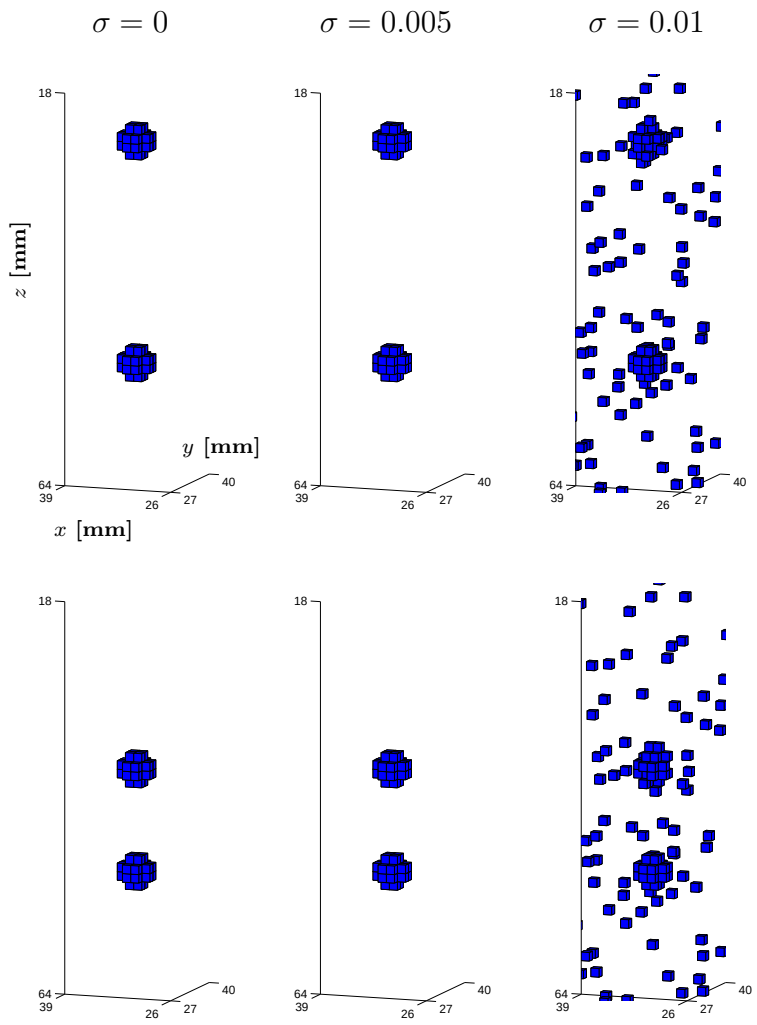

Figure 6. BMLR reconstructions from noisy projections of a pair of discretized defects.

The number of iterations and the CPU time required for convergence for the ICM and BMLR heuristics are shown in Table 2. Not surprinsingly, ICM is a very fast method. However, BMLR remains very attractive in terms of computing time. The step of building the ROI is not expensive since it results from a simple thresholding of the backprojection map, see (18). Depending on the SNR and on the positions of the spheres, the ROI varies from 1.1% to 1.6% of the total number of voxels in the considered examples. The total number of iterations is roughly proportional to the number of activated voxels in the solution. Obviously, the latter depends on the volume of the defects, which is expected to be small in NDE. However, it is also dependent on the SNR: the BMLR heuristic converges in a dozen of iterations if $\sigma < \sigma_{\text{th}}$ whereas a few hundreds are required if $\sigma > \sigma_{\text{th}}$. In the former case, the convergence is fast and the solution is reached within two seconds. In the latter case, most of the iterations activate only one isolated site of the ROI and the time needed for convergence increases significantly. As shown in Table 2, removing isolated sites from the ROI allows to speed-

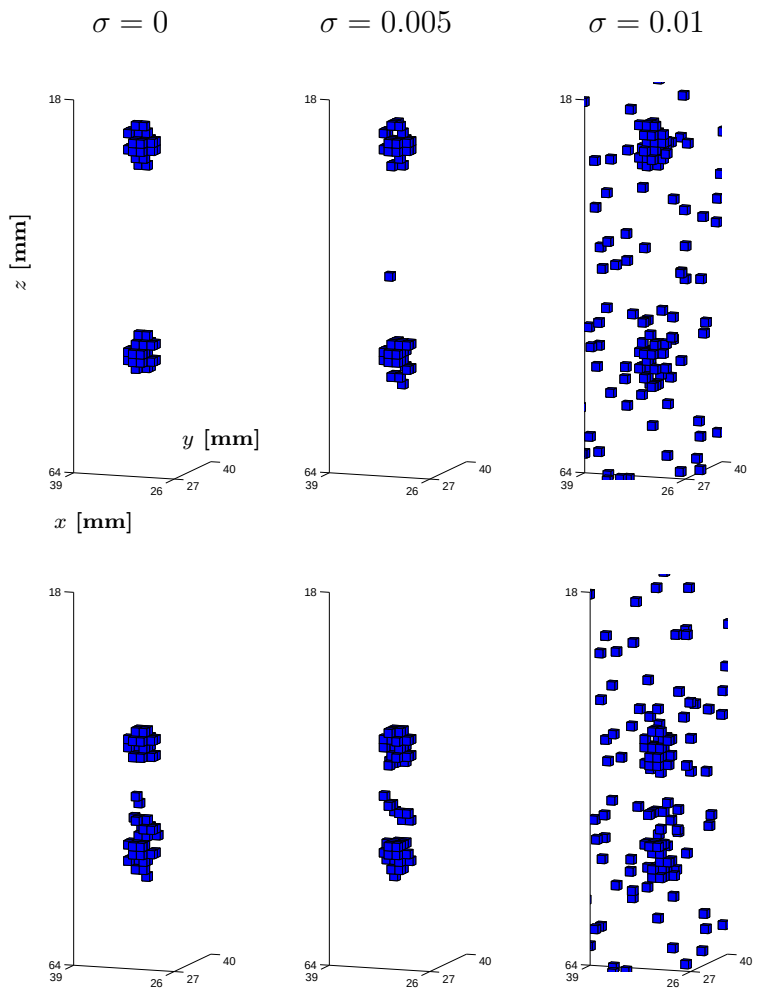

Figure 7. BMLR reconstructions from noisy projections of a pair of continuous defects.

up the convergence of the BMLR in this case. As stated earlier in this section, this pre-processing of the ROI is natural provided that flaws extend over more than a single voxel in the volume.

In principle, the BMLR heuristic can provide different solutions whether the ROI is considered or not. In order to test the variability of the solution, we also ran the BMLR heuristic with $\mathbb{I} = \mathbb{S}$, *i.e.*, we ignored the ROI. For high to moderate SNR values, the results were strictly identical to the previous ones, while only marginal differences appeared for the low SNR case. Finally, let us remark that the BMLR heuristic asks for an unrealistic computation time of several hours when the ROI is ignored, or if all the voxels belong to the ROI. In such a defavorable situation, the BMLR could still be efficiently implemented using a parallel structure, since the blocks can be explored independently one from each other.

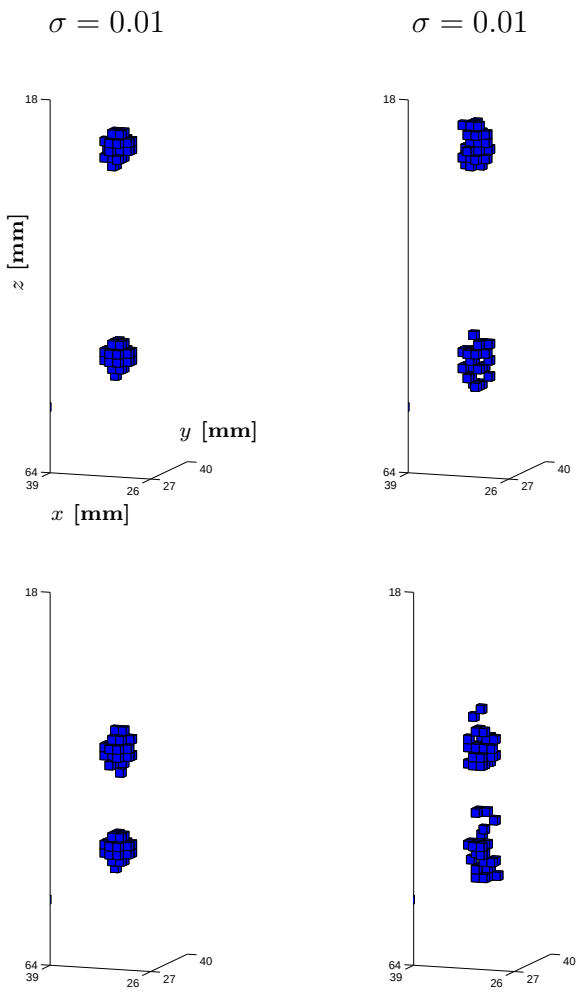

Figure 8. For the lowest SNR values, removing isolated voxels in the final BMLR reconstruction provides a better solution to both discretized (left) and continuous (right) defect problems.

5.4. BMLR: robustness with respect to modeling errors

In the previous subsection, the data are noisy projections of discretized objects. Obviously, real objects are continuous and (3) only constitutes an approximate model. Therefore, we found appropriate to further investigate the robustness of the method by considering projections from perfectly spherical objects, given that closed-form expressions are available for the projection of ellipsoids. The reconstructions provided by the BMLR are shown in Figure 7. In the case of distant objects, a good reconstruction is obtained whatever the SNR — up to the fact that spurious, isolated voxels must be removed for $\sigma = 0.01$, as shown on the second column of Figure 8. In the case of closer objects, the quality of the reconstruction deteriorates slightly, but the BMLR still provides meaningful results, since the two objects are still clearly distinguishable one from each other.

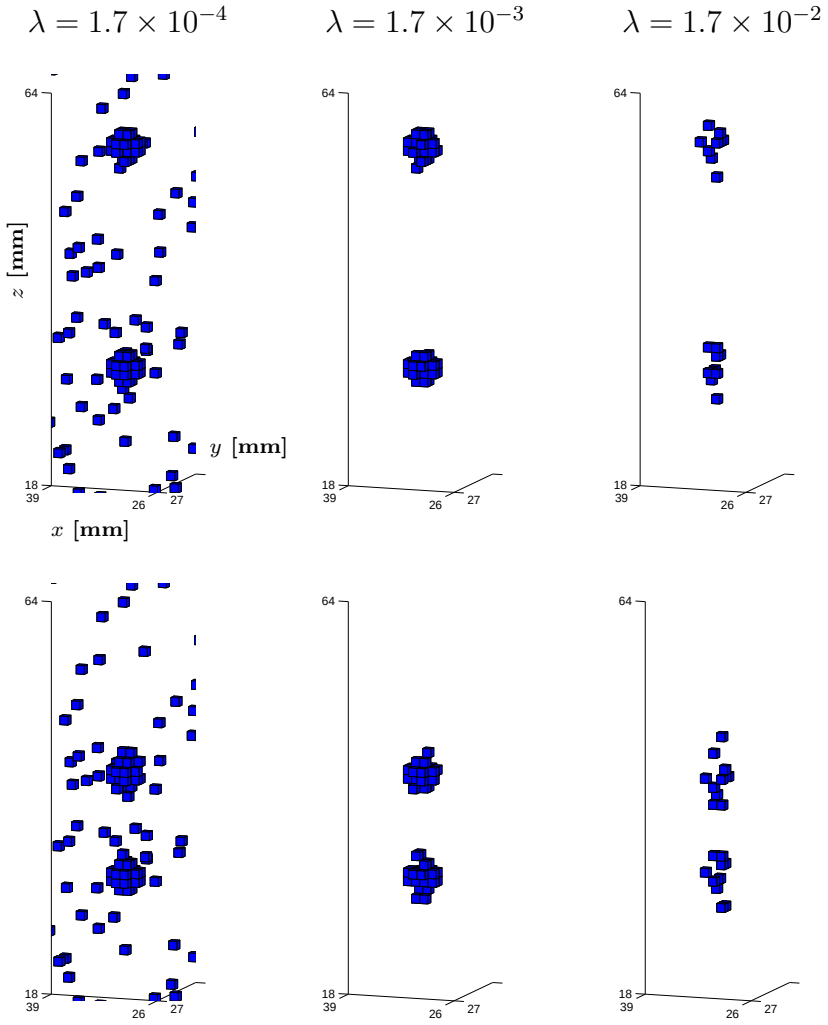

Figure 9. Impact of the regularization parameter $\lambda = 2\sigma^2 \log(p_0/p_1)$ on BMLR reconstruction at the lowest SNR (*i.e.*, $\sigma = 0.01$). Whereas under regularization (left) does not efficiently eliminate spurious voxels and over regularization (right) cancels most of the voxels in the flaws, a correct tuning of λ (center) produces a satisfactory reconstruction.

6. Conclusion

Spatially interacting models such as Markov random fields are usually employed within a Bayesian approach to reconstruction. Here, we rather resorted to a simpler iid model, and we proposed an efficient binary reconstruction algorithm, both in terms of robustness and computational burden. For high to moderate SNR, the binary constraint is sufficient to regularize the inversion of the tomographic problem. For low SNR, however, the number of false positive voxels increases dramatically. A simple strategy to deal with this problem is to cancel isolated voxels. Another strategy is to perform a penalization of active voxels by assigning a positive value to the regularization parameter $\lambda = 2\sigma^2\mu$ in criterion (8). However, the supervised tuning of λ must then be handled.

Table 2. Number of iterations and CPU time for the penalized approach and the ICM and BMLR heuristics on a Pentium III 2.4 Ghz (586 series) with 1Go RAM.

Algorithm	L [mm]	σ	# it.	CPU time
Penalized approach	26	0	333	3h 5min
		5.10^{-3}	97	1h 25min
		10^{-2}	198	2h 55min
	12	0	325	3h 42min
		5.10^{-3}	356	5h 11min
		10^{-2}	410	5h 58min
ICM	26	0	9	0.408s
		5.10^{-3}	11	0.452s
		10^{-2}	11	0.501s
	12	0	7	0.306s
		5.10^{-3}	8	0.331s
		10^{-2}	8	0.333s
BMLR	26	0	30	1.57s
		5.10^{-3}	27	1.55s
		10^{-2}	346	16.94s
	12	0	19	1.18s
		5.10^{-3}	19	1.23s
		10^{-2}	339	17.08s
BMLR (with isolated voxels discarded from the ROI)	26	0	30	1.57s
		5.10^{-3}	27	1.53s
		10^{-2}	173	8.37s
	12	0	19	1.17s
		5.10^{-3}	19	1.29s
		10^{-2}	170	8.43s

With a view to reconstruction from real data, several comments can be made about the physical model considered here. Firstly, the additive Gaussian noise assumption may not be appropriate for low counting rates τ_m. In such cases, the Poisson model (20) provides a more accurate description of the data. The ROI should then be computed using (21), and straightforward modifications should be brought to the BMLR algorithm.

Secondly, it should be noted that the Beer-Lambert relation (2) is only valid for a monochromatic emission of photons under a non-diffractive propagation assumption. In particular, Compton scattering is then negligible, which is not necessarily true for high energy gamma-ray sources (beyond 100 keV). In the latter case, further investigations would be required to define an appropriate ROI, for instance on the basis of the scattering model developped in [27].

Acknowledgments

This work was partially supported by Electricité de France (EDF) under contract number E961/P10AY0/E3.

Appendix A. Proof of Proposition 1

First, it is immediate to check that

$$\overline{\Delta}_n = -\log \underline{R}_n(\boldsymbol{y}) = \frac{1}{2\sigma^2}\left(\|\boldsymbol{y}\|^2 - \|\boldsymbol{y} - \boldsymbol{h}_{\bullet n}\|^2\right) - \mu,$$

which readily gives expression (16).

Now let us show that

$$\underline{R}_n(\boldsymbol{y}) \leq \frac{f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}\,|\,\boldsymbol{x})}{f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}\,|\,\boldsymbol{x}')} \tag{A.1}$$

for all couples $(\boldsymbol{x}, \boldsymbol{x}') \in \mathbb{S}_n^0 \times \mathbb{S}_n^1$ such that $\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{e}_n$: given (5), the latter inequality holds, since

$$\frac{1}{\sigma^2}\|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}'\|^2 - \frac{1}{\sigma^2}\|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2 + \overline{\Delta}_n = \frac{2}{\sigma^2}\boldsymbol{h}_{\bullet n}^{\mathrm{t}}\boldsymbol{H}\boldsymbol{x} \geq 0.$$

Given (12), inequalities (A.1) jointly imply that (13) holds.

Appendix B. Proof of Proposition 3

In the more general setting of Proposition 3, let us show that (A.1) is still valid for all couples $(\boldsymbol{x}, \boldsymbol{x}') \in \mathbb{S}_n^0 \times \mathbb{S}_n^1$ such that $\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{e}_n$. Given (19), this amounts to prove that

$$\prod_{m=1}^{M} \frac{\phi_m(a)\,\phi_m(d)}{\phi_m(b)\,\phi_m(c)} \leq 1,$$

with $a = 0$, $b = \mathcal{H}_{mn}$, $c = \boldsymbol{h}_{m\bullet}^{\mathrm{t}}\boldsymbol{x}$, and $d = b + c$, where $\mathcal{H}_{mn}$ denotes the (m, n)-entry of matrix $\boldsymbol{H}$. It is actually true that each term of the above product belongs to $[0, 1]$. This is obvious for those terms for which $\mathcal{H}_{mn} = \boldsymbol{h}_{m\bullet}^{\mathrm{t}}\boldsymbol{x} = 0$, since we have then $a = b = c = d$. Otherwise, let us remark that $a \leq b \leq d$ and $a \leq c \leq d$, so that

$$\log \phi_m(b) \geq \theta \log \phi_m(a) + (1 - \theta) \log \phi_m(d) \tag{B.1}$$

$$\log \phi_m(c) \geq \theta' \log \phi_m(a) + (1 - \theta') \log \phi_m(d) \tag{B.2}$$

where

$$\theta = \frac{d-b}{d-a} = \frac{c}{b+c}, \quad \theta' = \frac{d-c}{d-a} = \frac{b}{b+c} = 1-\theta.$$

Finally, term-to-term summation of inequalities (B.1) and (B.2) allows to conclude that $\phi_m(a)\,\phi_m(d) \leq \phi_m(b)\,\phi_m(c)$.

References

[1] L. Chatellier, L. Fournier, B. Charbonnier, L. Robillard, and B. Chassignole. Characterization improvement through signal processing application to radiography inspection: 3D recconstruction. In *Pressure Vessels and Piping Division Conference*, Denver, CO, 2005. ASME PVP.

[2] B. Chalmond, F. Coldefy, and B. Lavayssière. Tomographic reconstruction from non-calibrated noisy projections in non-destructive evaluation. *Inverse Problems*, 15:399–411, 1999.

[3] F. Retraint, F. Peyrin, and J.-M. Dinten. Three-dimensional regularized binary image reconstruction from three two-dimensional projections using a randomized ICM algorithm. *Int. J. Imag. Syst. Tech.*, 9:139–146, 1998.

[4] K. M. Hanson and G. W. Wechsung. Bayesian approach to limited-angle reconstruction in computed tomography. *J. Opt. Soc. Amer.*, 73:1501–1509, Nov. 1983.

[5] R. Rangayyan, A. P. Dhawan, and R. Gordon. Algorithms for limited-view computed tomography: An annotated bibliography and a challenge. *Applied Optics*, 24(23):4000–4012, Dec. 1985.

[6] E. T. Quinto. Tomographic reconstructions from incomplete data — numerical inversion of the exterior Radon transform. *Inverse Problems*, 4:867–876, 1988.

[7] D. J. Rossi and A. S. Willsky. Reconstruction from projection based on detection and estimation of objects — Part I and II: Performance analysis and robustness analysis. *IEEE Trans. Acoust. Speech, Signal Processing*, 32:887–906, 1984.

[8] M. Benedetti, M. Donelli, and A. Massa. Multicrack detection in two-dimensionnal structures by means of GA-based strategies. *IEEE Trans. Ant. Propag.*, 55:205–215, 2007.

[9] K. M. Hanson, G. S. Cunningham, G. R. J. Jennings, and D. R. Wolf. Tomographic reconstruction based on flexible geometric models. In *Proc. IEEE ICIP*, volume 2, pages 145–147, Austin, TX, Nov. 1994.

[10] C. Soussen and A. Mohammad-Djafari. Polygonal and polyhedral contour reconstruction in computed tomography. *IEEE Trans. Image Processing*, 13(11):1507–1523, Nov. 2004.

[11] J. A. Sethian. *Level set methods and fast marching methods*. Cambridge Monograph on Applied and Computational Mathematics. Cambridge University Press, 1999.

[12] J.-P. Bruandet, J.-M. Dinten, F. Peyrin, and M. Barlaud. 3D tomographic reconstruction of binary images from cone beam projections: a fast level set approach. In *Proc. IEEE ISBI*, Washington, USA, July 2002.

[13] J. E. Besag. On the statistical analysis of dirty pictures (with discussion). *J. R. Statist. Soc. B*, 48(3):259–302, 1986.

[14] J. J. Kormylo and J. M. Mendel. Maximum-likelihood detection and estimation of Bernoulli-Gaussian processes. *IEEE Trans. Inf. Theory*, 28:482–488, 1982.

[15] A. C. Kak and M. Slaney. *Principles of Computerized Tomographic Imaging*. IEEE Press, New York, NY, 1988.

[16] K. D. Sauer and C. A. Bouman. A local update strategy for iterative reconstruction from projections. *IEEE Trans. Signal Processing*, 41(2):534–548, Feb. 1993.

[17] J. L. Marroquin, S. K. Mitter, and T. A. Poggio. Probabilistic solution of ill-posed problems in computational vision. *J. Amer. Statist. Assoc.*, 82:76–89, 1987.

[18] D. M. Greig, B. T. Porteous, and A. H. Seheult. Exact maximum a posteriori estimation for binary images. *J. R. Statist. Soc. B*, 51(2):271–279, 1989.

[19] F. Barahona. On the computational complexity of Ising spin glass models. *Journal of Physics A: Mathematical and General*, 15(10):3241–3253, 1982.

[20] S. Istrail. Statistical mechanics, three-dimensionality and NP-completeness: I. Universality of intractability of the partition functions of the Ising model across non-planar lattices. In *Proc. 32nd ACM Symp. on Theory of Comp. (STOC00)*, pages 87–96, Portland, OR, May 2000.

[21] N. Robert, F. Peyrin, and M. J. Yaffe. Binary vascular reconstruction from limited number of cone beam projections. *Med. Phys.*, 12:1839–1851, 1994.

[22] C. Y. Chi and J. M. Mendel. Improved maximum-likelihood detection and estimation of Bernoulli-Gaussian processes. *IEEE Trans. Inf. Theory*, 30:429–435, 1984.

[23] H. Press, William, B. P. Flannery, S. A. Teukolsky, and W. T. Vettering. *Numerical Recipes: The Art of Scientific Computation*. Cambridge Univ. Press, New York, 1989.

[24] P. J. Huber. *Robust Statistics*. John Wiley, New York, NY, 1981.

[25] D. P. Bertsekas. *Nonlinear programming*. Athena Scientific, Belmont, MA, 1995.

[26] C. A. Johnson and A. Sofer. A primal-dual method for large-scale image reconstruction in emission tomography. *SIAM J. Optimization*, 11(3):691–715, 2000.

[27] K. D. Sauer, J. J. Sachs, and C. Klifa. Bayesian estimation of 3-D objects from few radiographs. *IEEE Trans. Nuclear Sciences*, 41(5):1780–1790, Oct. 1994.