UNSUPERVISED BAYESIAN 3D RECONSTRUCTION FOR NON-DESTRUCTIVE EVALUATION USING GAMMAGRAPHY

Adrien Trillon[a,b], Jérôme Idier[b], and Pierre Peureux[a]

[a]Electricite de France, Research and Development Division,
6 Quai Watier, BP 49, 78401 Chatou, France
phone: + (33)130878506, fax: + (33)130877589,
email: adrien-externe.trillon@edf.fr,
pierre.peureux@edf.fr

[b]IRCCyN (CNRS UMR 6597)
1 rue de la Noë, BP 92101, 44321 Nantes Cedex 3, France
phone: + (33)240376909, fax: + (33)240376930,
email: Jerome.Idier@irccyn.ec-nantes.fr
web: http://www.irccyn.ec-nantes.fr/~idier

ABSTRACT

In this paper, a fully non-supervised extension of a recent reconstruction method is proposed. Previous contributions to Bayesian estimation of Gibbs hyperparameters are mostly based on the tabulation of the partition function. However, such a usual approach becomes impracticable when the dimensions of the object are large and variable across reconstructions. Our solution is to extrapolate the previously tabulated partition function to the specific dimensions of the region of interest. Validation tests show that this idea reveals efficient and realistic.

1. INTRODUCTION

Non-destructive evaluation (NDE) is used for the in-service inspection of metal structures [1, 2]. In the context of nuclear power station steel pipe inspection, the use of narrow-beam gamma-rays is described in [2]. Due to operational constraints, the inspection is performed with a total angular excursion of 30°.

With a view to a quantitative characterization of the potential defects, digitized tomography of the three-dimensional (3D) region of interest (ROI) is preferable to the simple interpretation of the radiographs. However, such a tomographic reconstruction problem is challenging since only *limited-angle* projections are available. Faithful reconstructions of the 3D attenuation map cannot be expected without adequate prior constraints on the solution [3, 4]. More specifically, recent contributions rely on

- the binary character of the attenuation map [5];
- the positivity of the attenuation map and on the sparsity of the defects, jointly with an edge-preserving Markov random field (MRF) prior model [6];
- an underlying 3D Markov-Potts prior model for region labels [7].

In the latter contribution, Mohammad-Djafari and Robillard propose an interesting reconstruction method that is nearly unsupervised, *i.e.,* all statistical hyperparameters but one are estimated rather than empirically adjusted. The remaining parameter is the strength of interactions in the Ising model. Unfortunately, this parameter plays a paramount role in the quality of reconstruction, so there is much at stake in a fully unsupervised extension of Mohammad-Djafari and Robillard's method (hereafter called "MDR method"). This is the motivation of the present work, special attention being paid to the practicability in the NDE context.

2. PROBLEM STATEMENT

In order to process the radiograms numerically, the films are usually digitized, and the inspected object is assumed to be decomposable into voxels. Up to pre-processing steps for calibration, Compton diffusion correction and compensation of misalignments between films, the observation equation reads $\boldsymbol{g} = \mathbf{H}\boldsymbol{f} + \varepsilon$, where $\boldsymbol{g} \in \mathbb{R}_+^M$ is the data vector gathering all projections, $\boldsymbol{f} = \{f(r), r \in \mathscr{R}\} \in \mathbb{R}_+^N$ contains the attenuation values corresponding to the N voxels of the ROI $\mathscr{R}$ scanned in an arbitrary order, and $\mathbf{H}$ is a projection matrix whose structure is characteristic of tomography problems. ε is a vector of noise, standing for all kinds of errors. Following [7], it will be assumed centered, white and Gaussian with a variance σ_ε^2:

$$p(\boldsymbol{g}\,|\,\boldsymbol{f}, \sigma_\varepsilon) = (2\pi\sigma_\varepsilon^2)^{-M/2} \exp\Big(-\frac{1}{2\sigma_\varepsilon^2}\,\|\boldsymbol{g} - \mathbf{H}\boldsymbol{f}\|^2\Big). \quad (1)$$

As a prior model, [7] makes use of a Gaussian mixture model, defined as a doubly Markov field:

H1) a spatial field of labels $\boldsymbol{z}$ is introduced to describe an object composed of K distinct materials, and $\boldsymbol{z}$ is assumed distributed according to a Potts model. Here, $K = 2$, so the latter identifies with an Ising model:

$$p(\boldsymbol{z}\,|\,\alpha) = p_\alpha(\boldsymbol{z}) = \exp\left(-\alpha J(\boldsymbol{z})\right)/k(\alpha) \quad (2)$$

where $\boldsymbol{z} = \{z(r), r \in \mathscr{R}\} \in \{0,1\}^N$,

$$J(\boldsymbol{z}) = \sum_{(r,s)\in\mathscr{C}} \delta(z(r) - z(s)),$$
$$k(\alpha) = \sum_{\boldsymbol{z}} \exp\left(-\alpha J(\boldsymbol{z})\right) \quad (3)$$

is the partition function (*i.e.,* the normalization constant), and $\mathscr{C}(r)$ is the set of cliques, *i.e.,* of neighbor pairs.

H2) the distribution of $\boldsymbol{f}$ conditional on $\boldsymbol{z}$ is defined as a first-order Gaussian random field. A decorrelation assumption is made for voxels belonging to distinct materials (*i.e.,* $z(r) \neq z(r')$), and the field is finally described by $2K+1$ parameters: the mean and variance parameters (m_k, v_k) of each material, and a correlation parameter β.

One of the interesting features of the resulting MDR method is to provide Bayesian estimates of $\boldsymbol{f}$ and $\boldsymbol{z}$ that do not depend on hyperparameters $\theta = (\sigma_\varepsilon, (m_k, v_k)_{1,\ldots,K})$, the latter being either estimated in a joint maximum *a posteriori* (JMAP) approach, or marginalized in a Markov chain Monte Carlo (MCMC) version. However, the proposed solutions still depend on hyperparameters α and β: typically,

the MCMC version of the MDR method allows to compute posterior expectations such as $\widehat{\boldsymbol{z}}(\alpha,\beta) = \mathrm{E}[\boldsymbol{z}\,|\,\boldsymbol{g},\alpha,\beta]$ and $\widehat{\boldsymbol{f}}(\alpha,\beta) = \mathrm{E}[\boldsymbol{f}\,|\,\boldsymbol{g},\alpha,\beta]$ or $\widehat{\boldsymbol{f}}(\alpha,\beta,\widehat{\boldsymbol{z}}) = \mathrm{E}[\boldsymbol{f}\,|\,\boldsymbol{g},\widehat{\boldsymbol{z}},\alpha,\beta]$. Whereas the influence of β can be assumed negligible, α plays the crucial role of a regularization parameter:

- For arbitrary large values of α, $p(\boldsymbol{z}\,|\,\alpha)$ tends to zero for all label vectors $\boldsymbol{z}$, but for constant ones. Therefore, the estimated object will be composed of only one material. In other words, the possible presence of defects will be missed. Such a solution is typically *over-regularized*.
- For arbitrary small values of α, no spatial regularization is imposed on the estimated object. Given the strongly ill-posed character of the problem, the corresponding solution will be typically *under-regularized*.

Therefore, the present paper focuses on a convenient way to deal with α in an unsupervised framework.

To begin with, a simplification is introduced with respect to Assumption H2: here, the distribution of $\boldsymbol{f}$ conditional on $\boldsymbol{z}$ is assumed white Gaussian, which corresponds to the particular case where $\beta = 0$. In the NDE context, it can be checked that the value of β has very little influence on the estimated solution. Thus, we have:

$$p(\boldsymbol{f}\,|\,\boldsymbol{z},\theta) = \prod_r \left(2\pi\sigma^2_{z(r)}\right)^{-1/2} \exp\left(-\frac{(f(r)-m_{z(r)})^2}{2\sigma^2_{z(r)}}\right). \quad (4)$$

On the other hand, we concentrate on the MCMC approach, which is likely to produce more reliable estimates than the JMAP method [8]. In principle, α can be incorporated in the sampling steps of a MCMC scheme, so that $\boldsymbol{z}$ and $\boldsymbol{f}$ be estimated according to $\widehat{\boldsymbol{z}} = \mathrm{E}[\boldsymbol{z}\,|\,\boldsymbol{g}]$ and $\widehat{\boldsymbol{f}} = \mathrm{E}[\boldsymbol{f}\,|\,\boldsymbol{g}]$ or $\widehat{\boldsymbol{f}}(\widehat{\boldsymbol{z}}) = \mathrm{E}[\boldsymbol{f}\,|\,\boldsymbol{g},\widehat{\boldsymbol{z}}]$. The joint posterior law of $(\boldsymbol{f},\boldsymbol{z},\theta,\alpha)$ decomposes as

$$p(\boldsymbol{f},\boldsymbol{z},\theta,\alpha\,|\,\boldsymbol{g}) \propto p(\boldsymbol{g}\,|\,\boldsymbol{f},\sigma_\varepsilon)p(\boldsymbol{f}\,|\,\boldsymbol{z},\theta)p(\boldsymbol{z}\,|\,\alpha)p(\theta)p(\alpha),$$

where the rhs terms are defined according to (1), (4), (2) and to hyperprior assignments, respectively. Therefore, the additional sampling step should be performed according to

$$p(\alpha\,|\,\boldsymbol{g},\boldsymbol{f},\boldsymbol{z},\theta) \propto p(\boldsymbol{z}\,|\,\alpha)p(\alpha).$$

However, there is a notorious difficulty here: $p(\boldsymbol{z}\,|\,\alpha)$ incorporates the partition function $k(\alpha)$ defined by (3), whose dependence on α is neither explicit nor computable in a direct way, except for a domain $\mathscr{R}$ of very small size, since there are 2^N terms in (3). Moreover, the exact calculation of $k(\alpha)$ is known to be a NP-hard problem for non-planar graphs [9].

In the next section, a Monte Carlo method based on importance sampling is proposed, which is directly inspired from [10]. It allows to compute $k(\alpha)$ on a given grid of reasonable values of α. The numerical cost of such an approach is acceptable, provided that the computation of $k(\alpha)$ be performed once for all. Unfortunately, this is not possible in our NDE context because $k(\alpha)$ depends on the size of the ROI, which is specific to each acquisition geometry, and partly user-defined according to the data.

Therefore, Section 4 examines the possibility of extrapolating $k(\alpha)$ from values tabulated for ROI of given sizes, to the value corresponding to an ROI of another size.

3. DIRECT EVALUTION OF THE PARTITION FUNCTION

The partition function cannot be estimated as a simple average of samples using a Monte Carlo method, since we have

$$k(\alpha) = \sum_{\boldsymbol{z}} \exp\left(-\alpha J(\boldsymbol{z})\right) = \sum_{\boldsymbol{z}} k(\alpha) p_\alpha(\boldsymbol{z}) = \mathrm{E}_\alpha\left[k(\alpha)\right],$$

where E_α denotes the expectation relative to distribution p_α. However, the importance sampling principle allows to get around this limitation. Indeed, the following identity is easy to establish for all α', α:

$$\frac{k(\alpha)}{k(\alpha')} = \mathrm{E}_{\alpha'}\Big[\frac{\exp\left(-\alpha J(\boldsymbol{z})\right)}{\exp\left(-\alpha' J(\boldsymbol{z})\right)}\Big]. \quad (5)$$

Moreover, the partition function is explicit for $\alpha = 0$ according to $k(0) = \sum_z \exp(0) = 2^N$, which allows to determine the remaining multiplicative constant. In practice, however, the identity (5) provides numerically accurate sample average only if α and α' take close values. We are thus driven to the following method to tabulate $k(\alpha)$ over a fine grid $(\alpha_0 = 0, \alpha_1, \ldots, \alpha_I)$: for $i = 1, \ldots, I$,

1. generate samples $(\boldsymbol{z}_\ell)_{\ell=1,\ldots,L}$ of $p_{\alpha_{i-1}}$ using an appropriate sampler;
2. compute $\log k_{\mathrm{MCMC}}(\alpha_i)$

$$= \log k_{\mathrm{MCMC}}(\alpha_{i-1}) + \log\Big(\frac{1}{L}\sum_\ell \exp\big((\alpha_i - \alpha_{i-1})J(\boldsymbol{z}_\ell)\big)\Big)$$

where the logarithm is taken to prevent numerical overflow problems.

In our context, resorting to the Swendsen-Wang sampler [11] is far more efficient than the usual Gibbs sampler to perform step 1, especially for large values of α, *i.e.,* when there are strong interactions between voxels.

Finally, let us remark that the described method is sufficiently accurate for our purpose. However, in more complex situations, there exist more powerful extensions of the same principle [10].

4. EXTRAPOLATION METHOD FOR ROI OF VARIABLE SIZE

4.1 Principle based on the number of cliques

To our best knowledge, there is no exact way of extrapolating $k(\alpha)$ from an array of given size, to another of different size. In this paper, we consider that, as a function of the array dimensions (α being considered as a constant here),

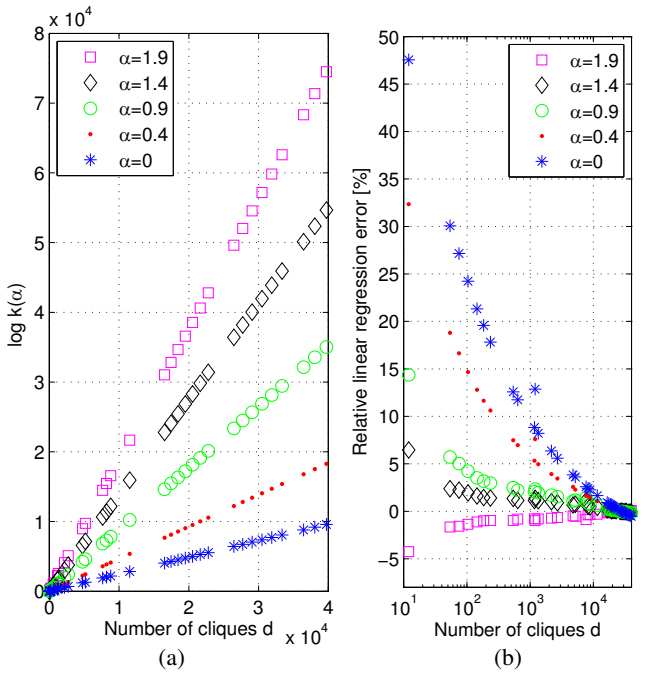
$$k(\alpha) \textit{ mainly depends on the number of cliques,} \quad (6)$$

i.e., of $d = (a-1)\times b\times c + a\times(b-1)\times c + a\times b\times(c-1)$ for an array of dimensions $a\times b\times c$.

Actually, $k(\alpha)$ is not only a function of d, since, for instance, $k(0) = 2^{a\times b\times c}$. However, Fig. 1(a) indicates that (6) is an operational rule of thumb, the dependency of $\log k(\alpha)$ with respect to d appearing as approximately linear. In this figure, $\log k_{\mathrm{MCMC}}(\alpha)$ is displayed for 5 distinct values of α and 36 parallepipedic arrays.

Fig. 1(b) displays the relative linear regression error:

$$\frac{\log k_{\mathrm{MCMC}}(\alpha) - S(\alpha)\alpha}{\log k_{\mathrm{MCMC}}(\alpha)},$$



(a) (b)

Figure 1: (a) Normalization constant $\log_{\text{MCMC}} k(\alpha, d)$ as a function of the number of cliques d for 5 values of α and 36 parallepipedic arrays[1]; (b) corresponding relative linear regression error values.

where $S(\alpha)$ corresponds to the best L_2 estimation.

Let us remark that a linear regression approach rather based on the number of sites N turns out to produce larger error values.

In what follows, the interpolation/extrapolation formula $\log k(\alpha) \approx S(\alpha)\alpha$ is adopted and tested with a view to estimate α on grids of arbitrary dimensions, in the practical range of ROI dimensions.

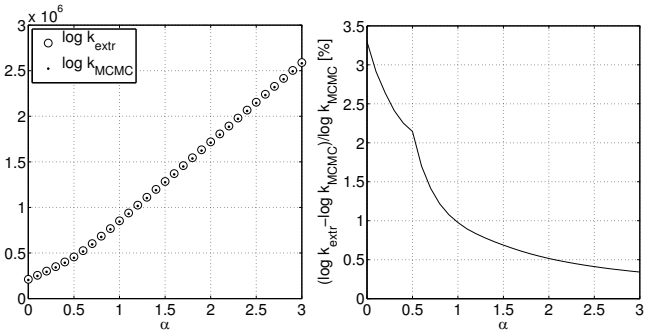
4.2 Practical validation tests

4.2.1 Extrapolation capacity

To test the extrapolation capacity of the method, an array of size $76 \times 51 \times 76$ has been considered, which amounts to more than 3×10^5 sites and nearly 9×10^5 cliques.

On the one hand, the direct evaluation method of Section 3 has been used to compute $\log k_{\text{MCMC}}$ for $\alpha = \alpha_0, , \ldots, \alpha_I$. On the other hand, the extrapolation formula has been applied to produce indirectly estimated values $\log k_{\text{extr}}$ on the same grid. The two results and the relative difference are displayed on Fig. 2. The difference between the two versions is at worst slightly higher than 3%, which seems perfectly acceptable in our context.

Let us remark that the position of the critical point $\alpha_c \approx 0.4336$ with respect to the phase transition phenomenon [12] is visible on Fig. 2(b).



(a) (b)

Figure 2: Comparison between MCMC-based and extrapolation-based estimation of the partition function.

4.2.2 Estimation of α given the label field z

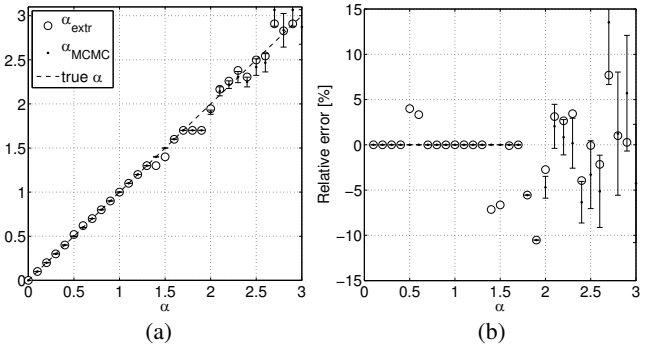
Let us consider the problem of estimating α in a simulated complete data case, *i.e.,* from a realization of an Ising field. For each α in the grid $\alpha_0, , \ldots, \alpha_I$, a $76 \times 51 \times 76$ realization z has been generated using the Swendsen-Wang sampler. Once $k(\alpha)$ has been computed for the same values of α, by either method, it is easy to obtain approximate estimated values of α. For instance,

$$\mathrm{E}[\alpha \,|\, z] \approx \frac{\sum_i p(z \,|\, \alpha_i) p(\alpha_i) \alpha_i}{\sum_i p(z \,|\, \alpha_i) p(\alpha_i)}$$

corresponds to the posterior mean for a given prior law $p(\alpha)$. On Fig. 3(a), black dots and white circles indicate posterior mean values respectively computed using k_{MCMC} and k_{extr}, both for a uniform prior. We have also reported standard deviation values, deduced from a similar sample average for the variance.

The performance of the two methods are comparable, except around the critical point α_c, where the extrapolation based method produces an additional error probably due to the higher variability of $k(\alpha)$ near α_c.

On the other hand, both perform less accurately for large values of α. This does not reveal growing numerical errors, but rather a loss of information: when the true value of α



(a) (b)

Figure 3: Estimation of α either based on k_{MCMC} or k_{extr}. α_{MCMC} is plotted along with error bars of one standard deviation.

[1] dimensions (a, b, c) of the 36 parallepipedic arrays, sorted by increasing number of cliques (from 12 to 39744):

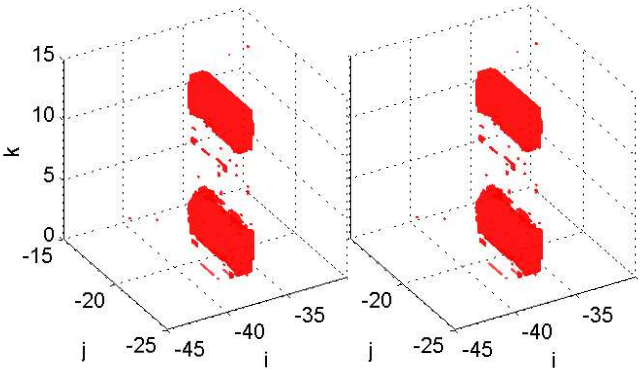
a	2	3	4	4	4	4	4	6	6	7	20	8	9	10	12	12	14	14	14	16	18	18	18	19	19	19	20	21	21	21	22	22	22	23	23	24
b	2	3	3	4	4	4	5	6	6	8	8	8	9	10	12	12	14	14	15	16	18	18	19	19	19	20	20	21	21	22	22	22	23	23	24	24
c	2	3	3	3	4	5	5	6	7	8	3	8	10	10	12	13	14	15	15	16	18	19	19	19	20	20	20	21	22	22	22	23	23	24	24	24

is large, accurate estimation of α becomes impossible from a realization z of constant size, since the latter tends to be uniform. Such a loss of information is clearly indicated by error bars of length growing with α in Fig. 3(b).

Finally, let us remark that the computation of k_{MCMC} is already demanding for a $76 \times 51 \times 76$ array: several hours of computing time are necessary on an up-to-date desktop PC (actually, much more than the total time to compute k_{MCMC} on the 36 arrays used to generate Fig. 1, while k_{extr} is almost immediate.

4.2.3 Estimation of α in the tomography context

The last test involves the simulated reconstruction of a 3D object. The object corresponds to two defects, one on top of the other, in an ROI of $76 \times 51 \times 76$ voxels (more details can be found in [2]). As seen on Fig. 4, the use of k_{extr} in place of k_{MCMC} produces only slight differences in the reconstructed object. This is not surprising since the estimated values of α are also close: $\alpha_{\text{MCMC}} = 1.4$ and $\alpha_{\text{extr}} = 1.3$, the difference being equal to one grid step. The result found is also very close from the one presented in [2, Fig. 4], obtained after empirical tuning of α (and β).



(a) Result based on k_{MCMC} (b) Result based on k_{extr}

Figure 4: Reconstructed object using a MCMC method based on either k_{MCMC} or k_{extr}.

5. CONCLUSION

In this paper, we have proposed a non-supervised version of Mohammad-Djafari and Robillard's reconstruction method [7]. Following Higdon [10], we have resorted to a tabulation of the partition function. However, Higdon's approach becomes impracticable when the size of the object is large and variable across reconstructions. Our solution is to extrapolate the tabulated partition function to the specific dimensions of the region of interest. Validation tests show that this idea reveals efficient and realistic.

The extrapolation technique could probably be improved by accounting for both the number of cliques and the number of sites. The numerical asymptotic results available for the cubic Ising model could also be taken into account [12].

However, the main limitation of the reconstruction method is to rely on a prior assigning equiprobabilities to the two types of label: by symmetry, it is clear that (2) implies $p_\alpha(z_s = 0) = p_\alpha(z_s = 1) = 1/2,\ \forall s, \alpha$. Obviously, the potential defects obviously extend over a very small part of the reconstructed volume. Therefore, a more suited prior for z could incorporate an "external field", *i.e.,* $p(z \mid \alpha, \gamma) \propto \exp(-\alpha J(z) + \gamma \sum_s z_s)$. Nonetheless, the price to pay would be the handling of a bivariate partition function, together with additional difficulties regarding efficient sampling [13].

REFERENCES

[1] 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.*, vol. 9, pp. 139–146, 1998.

[2] L. Fournier, L. Chatellier, P. Peureux, A. Mohammad-Djafari, and J. Idier, "3-D reconstruction from narrow-angle radiographs", in *QNDE*, Golden, CO, July 2007.

[3] K. M. Hanson and G. W. Wechsung, "Bayesian approach to limited-angle reconstruction in computed tomography", *J. Opt. Soc. Am.*, vol. 73, pp. 1501–1509, Nov. 1983.

[4] R. Rangayyan, A. P. Dhawan, and R. Gordon, "Algorithms for limited-view computed tomography: An annotated bibliography and a challenge", *Applied Optics*, vol. 24, no. 23, pp. 4000–4012, Dec. 1985.

[5] M. Allain and J. Idier, "Efficient binary reconstruction for non destructive evaluation using gammagraphy", *Inverse Problems*, vol. 23, no. 4, pp. 1371–1393, 2007.

[6] C. Soussen and J. Idier, "Reconstruction of 3D localized objects from limited angle X-ray projections: an approach based on sparsity and multigrid image representation", to appear, *J. Electr. Imag.*, 2008.

[7] A. Mohammad-Djafari and L. Robillard, "Hierarchical markovian models for 3d computed tomography in non destructive testing applications", in *Proc. EUSIPCO 2006*, Florence, Italy, Sep. 2006, pp. 111–115.

[8] R. J. A. Little and D. B. Rubin, "On jointly estimating parameters and missing data by maximizing the complete-data likelihood", *Amer. Statist.*, vol. 37, pp. 218–220, Aug. 1983.

[9] 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.*, Portland, OR, May 2000, pp. 87–96.

[10] D. M. Higdon, J. E. Bowsher, V. E. Johnson, T. G. Turkington, D. R. Gilland, and R. J. Jaszczak, "Fully Bayesian estimation of Gibbs hyperparameters for emission computed tomography data", *IEEE Trans. Medical Imaging*, vol. 16, no. 5, pp. 516–526, Oct. 1997.

[11] D. M. Higdon, "Auxiliary variable methods for markov chain monte carlo with applications", *J. Amer. Statist. Assoc.*, vol. 93, no. 442, pp. 585–595, 1998.

[12] F. Y. Wu, "The Potts model", *Rev. Mod. Phys.*, vol. 54, no. 1, pp. 235–268, Jan. 1982.

[13] A. Barbu and S. Zhu, "Generalizing Swendsen-Wang for image analysis", *J. Comput. Graph. Statist.*, vol. 16, no. 4, pp. 877–900, 2007.