ISING FIELD PARAMETER ESTIMATION FROM INCOMPLETE AND NOISY DATA

J.-F. Giovannelli

Groupe Signal – Image
Laboratoire de l'Intégration du Matériau au Système (Université de Bordeaux – CNRS)
351 Cours de la Libération, 33405 Talence Cedex, France
E-mail: Giova@IMS-Bordeaux.fr

ABSTRACT

The present paper deals with the estimation problem of the Ising field parameter and extends a previous one [1]. It proposes an estimate from indirect observation (incomplete and noisy), whereas the previous paper proposed an estimate from direct observation (complete and noiseless). Both of them are based on an explicit expression for the partition function, known for a long time [2] but, to the best of our knowledge, never used for parameter estimation (except in our previous paper [1]). Both of them are developed in a Bayesian framework. In our previous study (direct observation), the posterior law is explicit but in the present case (indirect observation) the posterior law is not explicit due to the hidden structure. The proposed approach relies on a full Bayesian strategy and a stochastic sampling algorithm (Gibbs sampler including a Metropolis-Hastings step) for posterior exploration. The paper proposes a numerical evaluation of the proposed method.

Index Terms — Ising field, parameter estimation, incomplete data, hidden variable, Bayesian, partition function.

1. ISING FIELD AND UNSUPERVISED INVERSION

Over the past decades, Bayesian methods for ill-posed image reconstruction problems have become increasingly popular because they are able to coherently account for both measurements and image properties. Within this framework, we are interested in advanced problems of joint reconstruction-segmentation, i.e. including a segmentation operation jointly performed with the reconstruction. To this end, methods resort to the Ising / Potts model for region labels (introduced in [3, 4] for segmentation). Such reconstruction-segmentation problems have recently been adressed and solutions have been proposed based on this model [5, 6, 7, 8].

Anyway, the complete solution requires the adjustment of the Ising / Potts field parameter. The generally investigated solutions for Markov field parameter estimation resort to statistics [9, Part.VI], [10, Ch.8] whether the field is directly or indirectly observed. They are potentially powerful but they come up against two major difficulties related to the explicitation of the observation law.

1. *The partition function of the field.* For general random fields, the partition function is in unknown relation with the field parameter. Nevertheless, there are three specific cases where this is not: (1) the well-known Gaussian field, (2) a class of non-Gaussian fields based on continuous mixture of Gaussians, recently proposed [11] and (3) the Ising field, of interest here. For the latter, the partition is explicitly known from [2], but surprisingly, a certain part of the image processing and computer science literature seems unaware of its existence. Moreover, to the best of our knowledge, it has never been used in parameter estimation methods (except in our previous paper [1]).

2. *The marginalization over the unobserved field.* When the field is directly observed, the parameter likelihood (in a standard sense or in a non-Bayesian sense) is available and [1] is devoted to this case. When the field is indirectly observed, the observation law is based on the marginalization of the unobserved field. Except for rare cases (e.g. Gaussian), it is a difficult task, for general field structure and observation scheme. In the case of the Ising field, the marginalization is impossible. Moreover, general tools, such as the class of Expectation-Maximization schemes are untractable in this case. We resort to a full Bayesian strategy and a stochastic sampler for posterior law exploration.

The paper is organized as follows. Section 2 introduces the notations, the Ising model including its partition function and the observation scheme. Section 3 proposes the parameter estimation method and the developed algorithm. They are numerically illustrated and evaluated in Section 3.3. Conclusions and perspectives are delivered in Section 4.

2. ISING FIELD AND OBSERVATION SCHEME

2.1. The Ising field and partition function

The homogeneous Ising field is a binary Markov random field driven by a unique parameter $\beta \in \mathbb{R}_+$. Here, it is considered

978-1-4577-1302-6/11/$26.00 ©2011 IEEE

on a finite $N \times N$ regular lattice, with $P = N^2$ sites, and denoted by $\boldsymbol{X} = \left[X_s\,,\, s \in \mathbb{N}_N^2\right]$, where $\mathbb{N}_N = \{1, 2, \ldots, N\}$. The probability for a configuration $\boldsymbol{x}_0$ writes:

$$\Pr\left[\,\boldsymbol{X} = \boldsymbol{x}_0 \,|\, \beta\,\right] = Z_P(\beta)^{-1}\, \exp\left[\,2P\beta\,\rho\left(\boldsymbol{x}_0\right)\right]\,. \tag{1}$$

From a statistical standpoint, the law is clearly in the exponential family: β is the natural parameter and $\rho\left(\boldsymbol{x}\right)$ is a sufficient statistic. It describes the neighborhood structure and the pixels interactions:

$$\rho\left(\boldsymbol{x}\right) = \frac{1}{2P}\sum_{r\sim s}\delta\left(x_r, x_s\right)\,,$$

where $\delta\left(x, y\right)$ is 1 if $x = y$ and 0 otherwise and the $\sim$ symbol stands for neighbor relationship in a 4-connexity system.

The keystone for the inference on β is the partition function $Z_P(\beta)$ involved in Eq. (1). It is given by a formidable summation over the 2^P possible configurations of the field:

$$Z_P(\beta) = \sum_{\boldsymbol{x}} \exp\left[\,2P\beta\,\rho\left(\boldsymbol{x}\right)\right]\,. \tag{2}$$

Its theoretical calculation seems a desperate task, however, a salient contribution due to Onsager [2] provides the result for both infinite and finite lattices. Starting from Onsager's paper and other literature [12], it has been a task to drive out the partition function. It is given and studied in our previous paper [1]: the huge summation of Eq. (2), becomes a one dimensional finite short summation, which is considerably simpler and enables the subsequent developements.

2.2. Observation scheme

We are interested in a standard scheme of hidden Markov fields: the field $\boldsymbol{X}$ is indirectly and incomplitely observed. The observed field is denoted by $\boldsymbol{Y}$ and it is defined over a set of sites denoted by $\mathcal{S}$. The observation scheme is defined by the set of probabilities $\Pr\left[\,\boldsymbol{Y} \,|\, \boldsymbol{X}\,\right]$ for each configuration of the hidden field $\boldsymbol{X}$ and the observed field $\boldsymbol{Y}$.

The case of interest is a separable perturbation, that is to say: $\Pr\left[\,\boldsymbol{Y} \,|\, \boldsymbol{X}\,\right]$ is the product of the $\Pr\left[\,Y_s \,|\, X_s\,\right]$. Moreover, it is defined by an error probability π, such that:

$$\Pr\left[\,Y_s = x \,|\, X_s = x\,\right] \;=\; 1 - \pi \tag{3}$$

$$\Pr\left[\,Y_s \neq x \,|\, X_s = x\,\right] \;=\; \pi \tag{4}$$

for all observed site $s \in \mathcal{S}$. In other words: each variable Y_s is the variable X_s swapped with probability π. As a whole, the probability for the entire observed field, given the hidden field can be written

$$\Pr\left[\,\boldsymbol{Y} = \boldsymbol{y} \,|\, \boldsymbol{X} = \boldsymbol{x}\,, \pi\,\right] = \pi^Q\, \exp\left[\,\bar{\pi}\,\nu(\boldsymbol{x}, \boldsymbol{y})\,\right] \tag{5}$$

where $\bar{\pi} = -\mathrm{logit}\left[\pi\right]$, Q is the number of observed site (cardinality of $\mathcal{S}$) and $\nu(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\mathcal{S}} \delta\left(x_s, y_s\right)$ counts the number of non swapped pixels (over the observed sites).

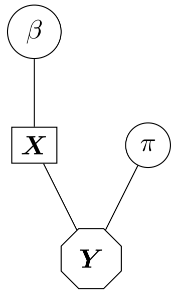


Fig. 1. Hierarchical structure, graphical model. β and π are the parameters of interest, $\boldsymbol{X}$ is the hidden (Ising) field and $\boldsymbol{Y}$ is the observed field.

2.3. Numerical example

This Section proposes a numerical example, for a size $N = 32$, a parameter value $\beta^\star = 0.8$ and a probability $\pi^\star = 0.2$. An Ising configuration $\boldsymbol{x}_0$ is produced according to (1) by a common parallel Gibbs sampler (in a chessboard-like manner) and an observed configuration $\boldsymbol{y}_0$ is simulated according to (3)-(4). The configurations are shown in Fig. 2.

(a) Hidden field (b) Observed field

Fig. 2. Hidden configuration $\boldsymbol{x}_0$ (left) and observed configuration $\boldsymbol{y}_0$ (right). The gray pixels mark the unobserved sites (about 10 %). The field parameter is $\beta^\star = 0.8$ and the error probability is $\pi^\star = 0.2$. Field size is $N = 32$.

3. PARAMETER ESTIMATION

3.1. Posterior density

This section tackles the estimation problem for (β, π) within a Bayesian framework. It relies on the posterior density based on the observation law and a prior density:

$$f\left(\beta, \pi \,|\, \boldsymbol{Y} = \boldsymbol{y}_0\right) = \frac{\Pr\left[\,\boldsymbol{Y} = \boldsymbol{y}_0 \,|\, \beta, \pi\,\right] f\left(\beta, \pi\right)}{\Pr\left[\,\boldsymbol{Y} = \boldsymbol{y}_0\,\right]}\,. \tag{6}$$

The used prior $f\left(\beta\right)$ is uniform over finite interval $[0, B]$. Practically, we set $B = 2$: larger values of β are of no interest since the field is quasi-surely in a uniform configuration. Regarding π, the used prior $f(\pi)$ is uniform over $[0, 1]$. Moreover, β and π are a priori independent.

Remark 1 — *It is possible to show an ambiguity between π and $1 - \pi$ that is not explored in this paper.*

The point is the observation law $\Pr[\boldsymbol{Y} = \boldsymbol{y}_0 \,|\, \beta, \pi]$. It is built from the law for the hidden field (1) and the law for the observation (5):

$$\begin{aligned}\Pr[\boldsymbol{Y} = \boldsymbol{y}_0 \,|\, \beta, \pi] &= \sum_{\boldsymbol{x}} \Pr[\boldsymbol{Y} = \boldsymbol{y}_0, \boldsymbol{X} \,|\, \beta, \pi] \\ &= \sum_{\boldsymbol{x}} \Pr[\boldsymbol{Y} = \boldsymbol{y}_0, \,|\, \boldsymbol{X}, \pi] \Pr[\boldsymbol{X} \,|\, \beta]\end{aligned}$$

As mentionned in the introduction, there are usually two difficulties. The first one is related to the partition function of the law for the hidden component $\Pr[\boldsymbol{X} \,|\, \beta]$. In the present case, the partition is known (see section 2.1). The second difficulty is related to the marginalization, i.e. the summation over all the possible configurations of the hidden field. It is not tractable neither theoretically nor numerically and we resort to a sampling strategy.

3.2. Sampling the posterior density

This section describes a sampling method in order to explore the posterior law of interest $f(\beta, \pi \,|\, \boldsymbol{Y} = \boldsymbol{y}_0)$. It is a usual strategy [13, 14] based on the full posterior law for all the unobserved quantities given the observed one:

$$\Pr[\beta, \pi, \boldsymbol{X} \,|\, \boldsymbol{Y} = \boldsymbol{y}_0] = \frac{\Pr[\beta, \pi, \boldsymbol{X}, \boldsymbol{Y} = \boldsymbol{y}_0]}{\Pr[\boldsymbol{Y} = \boldsymbol{y}_0]}$$

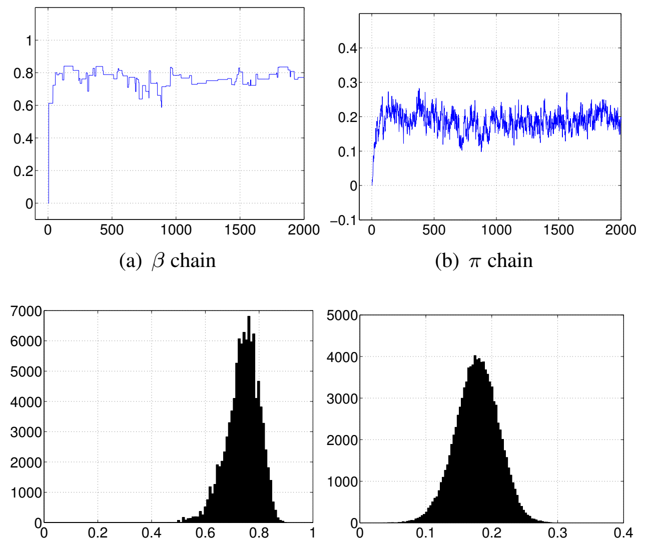
and the joint law writes:

$$\Pr[\beta, \pi, \boldsymbol{X}, \boldsymbol{Y}] = \Pr[\boldsymbol{Y} \,|\, \boldsymbol{X}, \pi] \Pr[\boldsymbol{X} \,|\, \beta] f(\beta) f(\pi)$$

and must be read in relation with the hierarchical structure shown in Fig. 1. It is exactly known from the law for the hidden field (1), the law for the observation (5) and the prior laws for β and π.

Direct sampling seems unfeasible due to the intricate dependence of the different variables. The proposed strategy resorts to Monte-Carlo Markov Chain algorithms based on Gibbs sampler. It iteratively samples $\boldsymbol{X}$, π and β under their respective conditional probabilities:

- $\boldsymbol{X} \sim \boldsymbol{X} \,|\, \boldsymbol{Y}, \beta, \pi$. It is an inhomogeneous Ising field that is itself sampled by means of a parallel Gibbs sampler (in a chessboard-like manner).
- $\pi \sim \pi \,|\, \boldsymbol{Y}, \boldsymbol{X}, \beta \equiv \pi \,|\, \boldsymbol{Y}, \boldsymbol{X}$. It is a Beta density $\mathcal{B}(Q - \nu(\boldsymbol{x}, \boldsymbol{y}) + 1, \nu(\boldsymbol{x}, \boldsymbol{y}) + 1)$.
- $\beta \sim \beta \,|\, \boldsymbol{Y}, \boldsymbol{X}, \pi \equiv \beta \,|\, \boldsymbol{X}$. It is not a standard density so direct sampling is impossible. We use a Metropolis-Hastings sampling step with the prior as the proposal.



(a) β chain (b) π chain

(c) β histogram (d) π histogram

Fig. 3. Chains (top) and histograms (bottom) of β (left) and π (right).

3.3. Simulation results

In order to monitor convergence, the empirical averages of the generated samples (for β and π) are recursively computed. The algorithm is stopped when their variations become smaller than a given threshold T. In the presented example $T = 10^{-5}$ and the algorithm produced 11086 samples. The first 2000 samples are shown in Fig. 3: after about 500 iterations (burn-in time) the parameters are stabilized and seem to be under the stationary law of the chain, i.e. the posterior.

The iterates are also shown in Fig. 3 as histograms: they are representative of the marginal posterior densities $f(\beta \,|\, \boldsymbol{Y} = \boldsymbol{y}_0)$ and $f(\pi \,|\, \boldsymbol{Y} = \boldsymbol{y}_0)$. They are also plotted in Fig. 4 as 2-D histogram: it is representative of the joint

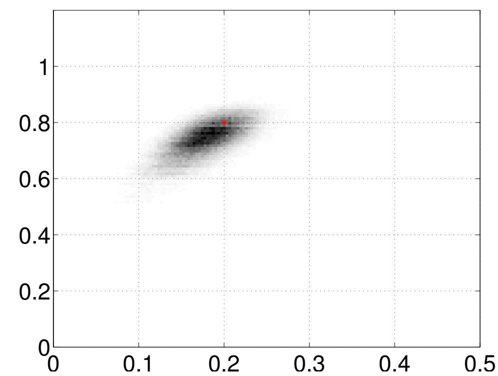


Fig. 4. 2-D histogram representative of the joint posterior $f(\beta, \pi \,|\, \boldsymbol{Y} = \boldsymbol{y}_0)$. Vertical axis β and horizontal axis π.

	True	Mean	Std Deviation	Error
β	0.8	0.7394	0.0606	7.5%
π	0.2	0.1776	0.0325	11.2%

Table 1. Numerical results.

posterior density $f(\beta, \pi \,|\, \boldsymbol{Y} = \boldsymbol{y}_0)$.

The considered point estimate is the Posterior Mean and the Posterior Standard Deviation is also computed. Results are given in Tab. 1: the observed error is relatively small (7.5% and 11.2% for β and π respectively). Anyway, an important point is that the method provides an idea of the uncertainties via the posterior standard deviation: it is shown, on this example, that each true value is inside the one standard deviation interval around the estimated value.

Remark 2 — *It requires about one minute to compute*[1] *the whole set of samples, i.e. to explore the posterior density and deduce the point estimates and standard deviations.*

4. CONCLUSION

This paper provides an advance in the area of Ising field manipulation and possible use for unsupervised image processing, reconstruction and segmentation. It focuses on a parametric estimation problem from indirect observation (incomplete and noisy). The problem presents two major difficulties related to the explicitation of the observation law: the first one is related to the partition function and the second one is related to the marginalization.

The estimate is developed in a full Bayesian framework for the Ising field parameter and the noise parameter. The proposed strategy is based: (1) on the exact law for the field, thanks to the exact partition function and (2) on marginalization through stochastic sampling (Metropolis-Hastings within Gibbs). The posterior mean and standard deviation are computed for both parameters.

The explicit expression for the partition function has been known for a long time [2] but to the best of our knowledge it has never been used for parameter estimation (except in our previous paper [1]). The novelty of the present paper, with respect to the previous one, is a methodological and algorithmic solution for indirect observation.

The proposed work is limited to the Ising field itself and it does not apply to the Potts field. A possible development is to approximate the partition function of the Potts field and to extend the present contribution. Another possible developement deals with the unsupervised aspect in Bayesian reconstruction-segmentation method [5, 7, 8, 15], taking advantage of the present contribution.

[1]The algorithm is implemented within the environment Matlab on a PC with a 3 GHz CPU and 3 GB of RAM. Code is about 100 lines long.

5. REFERENCES

[1] J.-F. Giovannelli, "Estimation of the Ising field parameter thanks to the exact partition function," in *Proc. IEEE ICIP*, Hong-Kong, September 2010, pp. 1441–1444.

[2] L. Onsager, "A two-dimensional model with an order-disorder transition," *Phys. Rev.*, vol. 65, no. 3 & 4, pp. 117–149, February 1944.

[3] J. E. Besag, "On the statistical analysis of dirty pictures (with discussion)," *J. R. Statist. Soc. B*, vol. 48, no. 3, pp. 259–302, 1986.

[4] S. Geman and C. Graffigne, "Markov random field image models and their applications to computer vision." in *Proceedings of the International Congress of Mathematicians*, Berkeley, CaliforniaUSA, 1986, pp. 1496–1517.

[5] O. Féron, B. Duchêne, and A. Mohammad-Djafari, "Microwave imaging of inhomogeneous objects made of a finite number of dielectric and conductive materials from experimental data," *Inverse Problems*, vol. 21, no. 6, pp. 95–115, december 2005.

[6] T. Vincent, L. Risser, P. Ciuciu, and J. Idier, "Spatially unsupervised analysis of within-subject fMRI data using multiple extrapolations of 3D Ising field partition functions," in *Machine Learning for Signal Processing*, Grenoble, France, September 2009.

[7] H. Ayasso and A. Mohammad-Djafari, "Variational Bayes with Gauss-Markov-Potts prior models for joint image restoration and segmentation," in *Int. Conf. on Computer Vision and Applications*, ser. Visapp Proceedings, Funchal, Madaira, Portugal, 2008.

[8] A. Mohammad-Djafari and H. Ayasso, "Variational Bayes and mean field approximations for Markov field unsupervised estimation," in *Machine Learning for Signal Processing*, Grenoble, France, September 2009.

[9] G. Winkler, *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods*. Springer Verlag, BerlinGermany, 2003.

[10] J. Idier, Ed., *Bayesian Approach to Inverse Problems*. London: ISTE Ltd and John Wiley & Sons Inc., 2008.

[11] J.-F. Giovannelli, "Unsupervised Bayesian convex deconvolution based on a field with an explicit partition function," *IEEE Trans. Image Processing*, vol. 17, no. 1, pp. 16–26, January 2008.

[12] K. Huang, *Statistical mechanics*. New-York, USA: Wiley, 1987.

[13] J.-M. Marin and C. Robert, *Bayesian Core. A Practical Approach to Computational Bayesian Statistics*, ser. Texts in statistics. Paris, France: Springer, 2007.

[14] A. Gelman, J. C. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, 2nd ed. Boca Raton, USA: Chapman & Hall/CRC, 2004.

[15] S. K. Kazuyuki Tanaka and M. Yasuda, "Statistical performance analysis by loopy belief propagation in Bayesian image modeling," in *International Workshop on Statistical-Mechanical Informatics*, ser. Journal of Physics: Conference Series, March 2010.