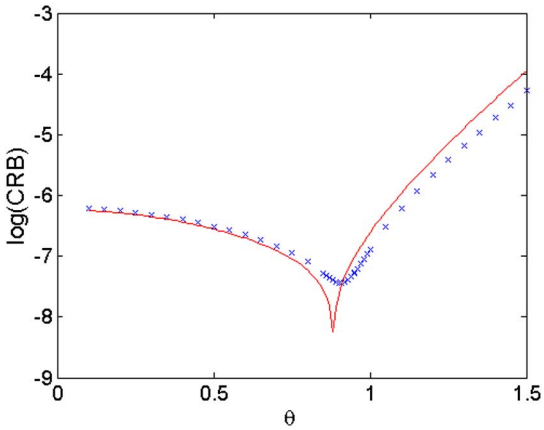

Fig. 1. Cramer–Rao bound for an Ising (MRF) defined on a toroidal graph. True CRB (solid red), estimates obtained by MC integration (blue crosses).

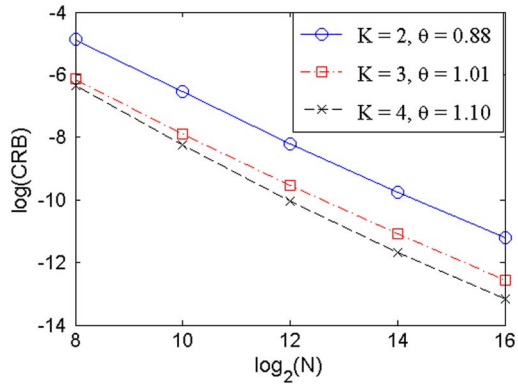

Fig. 2. Cramer–Rao bound for an Ising ($K = 2$) and two Potts MRF ($K = 3$ and $K = 4$) close to phase-transition and for different field sizes N. Results are displayed in *log-log* scales.

granularity coefficient or *inverse temperature* parameter. The Gibbs distribution (10) corresponds to the Ising MRF when $K = 2$, and to the Potts MRF for $K \geq 3$. In our experiments $\mathcal{V}(n)$ will be considered to be a bidimensional first-order (i.e., 4-pixel) neighborhood structure. However, the proposed method is valid for any correct neighborhood structure (see [12] for more details). Finally, note that despite their simplicity these models are extensively used in modern image segmentation and/or classification applications (see [20], [21] and references therein) and that the estimation of the granularity parameter θ is still an active research topic [9].

B. Validation With Ground Truth

To validate the proposed MC method under controlled conditions (i.e., for a known CRB), the proposed methodology has been first applied to an Ising model defined on a *toroidal* graph (i.e., with cyclic boundary conditions) of size $N = 32 \times 32$. Unlike most MRF models, this particular MRF has a known normalizing constant and FIM [8].

Fig. 1 compares the estimates obtained for different values of θ with the true CRB [8]. These estimates have been computed from Markov chains of 1 000 000 samples generated with a Gibbs sampler. More details about this experiment are provided in a separate technical report [19].

We observe in Fig. 1 that the estimates obtained with the proposed method are in good agreement with the true values of the CRB. We also observe that the error introduced by using a Monte Carlo approximation varies slightly with the value of θ and is larger at approximately $\theta = 0.9$, coinciding with the phase-transition temperature of the Ising MRF ($\theta_c = \log(1 + \sqrt{2}) \approx 0.88$). These variations with θ are due to the fact that the mixing properties of the Gibbs sampler deteriorate at temperatures close to θ_c due to long range dependencies between the elements of $\boldsymbol{z}$ [13, p. 339]. As a result the Markov chains associated with different values of θ have different effective samples sizes [13, p. 499] (i.e., different numbers of equivalent independent samples) and produce estimates with different accuracies. Indeed, the effective sample size, measured from the chain's autocorrelation function, is 870 000 samples for $\theta = 0.1$, it decreases progressively to 20 000 samples for $\theta = 0.9$ and then increases to 235 000 samples for $\theta = 1.5$.

C. Asymptotic Study of the CRB

The second set of experiments shows the evolution of the CRB with respect to the size of observation vector $\boldsymbol{z}$ (i.e., the number of field components N). The CRB has been computed for the following 5 field sizes $N = (2^8, 2^{10}, 2^{12}, 2^{14}, 2^{16})$, corresponding to bidimensional MRFs of size 16×16, 32×32, 64×64, 128×128 and 256×256. Experiments have been performed using an Ising MRF, a 3-state and a 4-state Potts MRF (i.e., $K = 2$, $K = 3$ and $K = 4$ respectively) defined on a regular lattice (not a toroid). CRB estimates have been computed from Markov chains of 2500 000 samples. Finally, for each model, the parameter θ was set close to the critical phase-transition value, i.e., $\theta_c = \log(1 + \sqrt{K})$ to introduce a strong dependency between the components of the MRF. Fig. 2 shows the resulting CRBs versus the size of the MRF N in logarithmic scales.

We observe that for all models the logarithm of the CRB decreases almost linearly with the logarithm of the number of field components. This result shows that the strong dependency between the field components does not modify significantly the linear behavior that is generally observed for models defined by statistically independent components. We also observe that the CRB decreases with the number of states K, indicating that an accurate estimation of θ for the Ising model is more difficult than for a Potts MRF.

D. Evaluation of State-of-the Art Estimators of θ

The third set of experiments compares the CRB to the empirical variance of three state-of-the art estimation methods, the *auxiliary variable* [1], *exchange* [2] and *ABC* [9] algorithms. As explained previously, the CRB is often used as a means to measure the performance of unbiased estimators in terms of mean square error. In this letter, the three algorithms studied in [1], [2], [9] have been used to compute an approximate maximum-likelihood (ML) estimation of θ for the 3-state Potts MRF (additional results with an Ising MRF are provided in a separate technical report [19]).

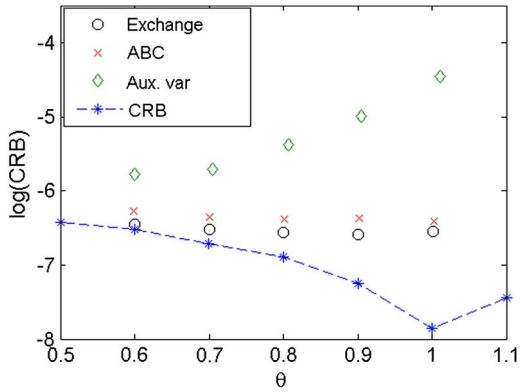

Fig. 3. Cramer–Rao bounds for a 3-state Potts models of size 32×32. Results are displayed in logarithmic scale.

Experiments were conducted as follows. First $N_{ML} = 2500$ synthetic observation vectors $\boldsymbol{z}^{(i)} \sim f_\theta(\boldsymbol{z})$, $i = 1, \ldots, N_{ML}$ were generated using an appropriate Gibbs sampler. Then, for each observation $\boldsymbol{z}^{(i)}$, three ML estimates $\hat{\theta}^{(i)}_{EXCH}$, $\hat{\theta}^{(i)}_{ABC}$ and $\hat{\theta}^{(i)}_{AUX}$ were computed using the three estimation methods mentioned above. Finally, the variance of each estimator was approximated by computing the sample variance, e.g., $\mathrm{Var}(\hat{\theta}_{ABC}) = (N_{ML}-1)^{-1} \sum_{i=1}^{N_{ML}} (\hat{\theta}^{(i)}_{ABC} - \bar{\theta}_{ABC})^2$ with $\bar{\theta}_{ABC} = N_{ML}^{-1} \sum_{i=1}^{N_{ML}} \hat{\theta}^{(i)}_{ABC}$. More details about this experiment are provided in a separate technical report [19].

Fig. 3 compares the CRB estimated with our method for the 3-state Potts MRF with the variance of the ML estimates obtained with the state-of-the art algorithms. These values have been computed for $\theta < \theta_c$ which is the range of interest for this model (for $\theta > \theta_c$ all the field components have almost surely the same color). We observe the good performance of the ML estimators based on the exchange [2] and the ABC [9] for small values of θ (i.e., $\theta < 0.8$). However, their performance decreases progressively for $\theta > 0.8$, a behavior that is explained by the fact that the estimators use a Gibbs sampler to approximate the intractable likelihood. As explained previously, the mixing properties of this sampler deteriorate as θ increases towards the critical value θ_c. This results in a degradation of the approximation of the likelihood and in a larger ML variance. Moreover, one can also see that the estimator based on the auxiliary variable method [1] has a larger variance than the others for all values of θ. This result is in accordance with the experiments reported in [2], [9].

IV. Conclusion

This letter studied the problem of computing the CRB for the parameters of Markov random fields. For these distributions the CRB depends on the derivatives of the normalizing constant or partition function $C(\boldsymbol{\theta})$, which is generally intractable. This difficulty was addressed by exploiting an interesting property of the exponential family that relates the derivatives of the normalizing constant $C(\boldsymbol{\theta})$ to expectations of the MRF potential. Based on this, it was proposed to estimate the Fisher information matrix of the MRF (and therefore the CRB) using a Monte Carlo method. The proposed approach was successfully applied to the Ising and the Potts models, which are frequently used in signal processing applications. The resulting bounds have been used, in turn, to assess the statistical efficiency of three state-of-the art estimation methods that are interesting for image processing applications. An extension of the proposed method to hidden MRFs is currently under investigation. Perspectives for future work include the derivation of Bayesian bounds for intractable models whose unknown parameters are assigned prior distributions.

References

[1] J. Moller, A. N. Pettitt, R. Reeves, and K. K. Berthelsen, "An efficient Markov chain Monte Carlo method for distributions with intractable normalising constants," *Biometrika*, vol. 93, no. 2, pp. 451–458, Jun. 2006.

[2] I. Murray, Z. Ghahramani, and D. MacKay, "MCMC for doubly-intractable distributions," in *Proc. (UAI 06) 22nd Annu. Conf. Uncertainty in Artificial Intelligence*, Cambridge, MA, USA, Jul. 2006, pp. 359–366.

[3] C. Andrieu and G. O. Roberts, "The pseudo-marginal approach for efficient Monte Carlo computations," *Ann. Statist.*, vol. 37, no. 2, pp. 697–725, 2009.

[4] C. Andrieu, A. Doucet, and R. Holenstein, "Particle Markov chain Monte Carlo methods," *J. Roy. Statist. Soc. B*, vol. 72, no. 3, May 2010.

[5] X. Descombes, R. Morris, J. Zerubia, and M. Berthod, "Estimation of Markov random field prior parameters using Markov chain Monte Carlo maximum likelihood," *IEEE Trans. Image Process.*, vol. 8, no. 7, pp. 945–963, Jun. 1999.

[6] F. Forbes and G. Fort, "Combining Monte Carlo and mean-field-like methods for inference in hidden Markov random fields," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 824–837, Mar. 2007.

[7] C. McGrory, D. Titterington, R. Reeves, and A. Pettitt, "Variational bayes for estimating the parameters of a hidden potts model," *Statist. Comput.*, vol. 19, no. 3, pp. 329–340, Sep. 2009.

[8] J.-F. Giovannelli, "Ising field parameter estimation from incomplete and noisy data," in *Proc. IEEE Int. Conf. Image Proc. (ICIP)*, Sep. 2011, pp. 1853–1856.

[9] M. Pereyra, N. Dobigeon, H. Batatia, and J.-Y. Tourneret, "Estimating the granularity parameter of a Potts-Markov random field within an MCMC algorithm," *IEEE Trans. Image Process.*, vol. 22, no. 6, pp. 2385–2397, Jun. 2013.

[10] R. G. Everitt, "Bayesian parameter estimation for latent Markov random fields and social networks," *J. Comput. Graph. Statist.*, vol. 21, no. 4, pp. –960, 2012, to appear.

[11] H. L. V. Trees, *Detection, Estimation, and Modulation Theory: Part I.* New York, NY, USA: Wiley, 1968.

[12] S. Z. Li, *Markov Random Field Modeling in Image Analysis*. New York, NY, USA: Springer-Verlag, 2001.

[13] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*. New York, NY, USA: Springer-Verlag, 1999.

[14] C. P. Robert, *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation (2nd ed.)*. New York, NY, USA: Springer-Verlag, 2001.

[15] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer-Verlag, 2006.

[16] D. M. Kay, *Information Theory, Inference and Learning Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[17] C. J. Geyer, "Practical Markov chain Monte Carlo," *Statist. Sci.*, vol. 7, no. 4, pp. 473–483, Nov. 1992.

[18] R. Swendsen and J. Wang, "Nonuniversal critical dynamics in Monte Carlo simulations," *Phys. Rev. Lett.*, vol. 58, no. 2, pp. 86–88, Jan. 1987.

[19] M. Pereyra, N. Dobigeon, H. Batatia, and J.-Y. Tourneret, Computing the Cramer-Rao bound of Markov random field parameters: Application to the Ising and the Potts models Univ. Toulouse, IRIT/INP-ENSEEIHT, Toulouse, France, Tech. Rep., Jun. 2013 [Online]. Available: http://arxiv.org/abs/1206.3985/=0pt

[20] T. Vincent, L. Risser, and P. Ciuciu, "Spatially adaptive mixture modeling for analysis of fMRI time series," *IEEE Trans. Med. Imag.*, vol. 29, no. 4, pp. 1059–1074, Apr. 2010.

[21] M. Pereyra, N. Dobigeon, H. Batatia, and J.-Y. Tourneret, "Segmentation of skin lesions in 2D and 3D ultrasound images using a spatially coherent generalized Rayleigh mixture model," *IEEE Trans. Med. Imag.*, vol. 31, no. 8, pp. 1509–1520, Aug. 2012.