Comment on MCMC methodology

Chuanhai Liu Jun Liu
Department of Statistics
Harvard University
Cambridge, MA 02138

The authors are to be congratulated for providing an enjoyable survey of the very important area of the MCMC methodology. Our comment will focus on the assessment of the convergence of iterative sampling using multiple runs, as advocated by Gelman and Rubin (1992); our proposal has benefited from insightful conversations with Professor D.B. Rubin.

When multiple parallel runs are used, it may be revealing to utilize information across parallel chains. In particular, by combining the transition ratios and density ratios across m parallel chains, an interesting monitoring statistic can be constructed as follows. At the t-th iteration, $m(m-1)$ values of the control variable U are computed, each using two different parallel chains i and j,

$$U^{(i,j,t)} = \frac{\pi(X^{(j,t)})}{\pi(X^{(i,t)})} \cdot \frac{K_\pi(X^{(j,t-1)},\ X^{(i,t)})}{K_\pi(X^{(j,t-1)},\ X^{(j,t)})} \quad i \neq j,\ i,j = 1,\ldots,m, \tag{1}$$

where $\pi(\cdot)$ is the target distribution, $K_\pi(\cdot,\cdot)$ is the transition function as in Section 3.2 of Smith and Roberts (1993); assumed time homogeneous as with the Gibbs sampler; and the ratio $\pi(X^{(j,t)})/\pi(X^{(i,t)})$ is computed, for example using (2.2) of Besag (1974) for the Gibbs sampler. If $p_t(\cdot)$ denotes the current density,

$$p_t(Y) = \int K_\pi(X,Y)p_{t-1}(X)dX,$$

then one can prove the following equation demonstrating the relationship between U and the convergence of p_t to π:

$$E_0(U^{(i,j,t)}) = \mathrm{var}_\pi(\frac{p_t(X)}{\pi(X)}) + 1. \tag{2}$$

The expectation on the left hand side of (2) is taken under the initial distribution and the variance on the right hand side is taken under the target distribution. Therefore, each $U^{(i,j,t)}$

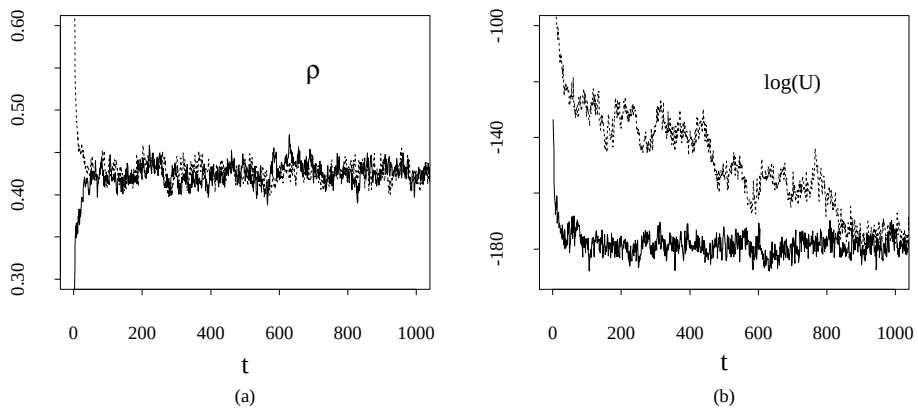

Figure 1: The Gibbs sampler for the Ising Model; dotted lines are the means for the 20 series starting from $\rho = 1$, and solid lines are the means for the 20 series starting at $\rho = 0$. (a) Correlation coefficient; (b) $\log(U)$.

is an unbiased estimate of a distance between distribution p_t and the target distribution π, and is potentially useful in monitoring the convergence of an iterative simulation.

To illustrate this method, we applied it to a Gibbs sampling scheme on a 3-dimensional Ising model of size $10 \times 10 \times 10$, where 40 parallel runs are simulated, half of them starting from the extreme situation corresponding to $\rho = 1$, and the other half from the other extreme of $\rho = 0$. Our result, displayed in Figure 1, suggests that monitoring the convergence of the entire underlying random variable by the value of the nearest-neighbor correlation, ρ, may not be sufficient. In particular, Figure 1(a) shows that the distribution of mean ρ seems to have converged after 100 iterations, whereas Figure 1(b) shows that the entire underlying random variable converges much slower.

References

[1] Besag, J. (1974), Spatial Interaction and the Statistical Analysis of Lattice systems (with discussion). *J. R. Statist. Soc.* **B, 36,** 192–236.

[2] Gelman A. and Rubin D. B. (1992), Honest Inferences from Iterative Simulation. To appear on Statistical Science.