trivariate distribution which is $N(\mu_1, \Sigma_1)$ with probability 0.5 and $N(\mu_2, \Sigma_1)$ with probability 0.5 where $\mu_2^T = (-1.5, -7, 7)$. In both cases the Metropolis-within-Gibbs method appears to be superior. Combined with its relatively simple set-up it becomes attractive for many problems.

All runs were performed in a commonly available computing environment—an IBM-PC 80286 with mathematics coprocessor. The Gibbs sampler was run with 300 parallel strings in case 1 and 500 in case 2. Convergence of the sampler is assessed relative to the known mean and covariance matrix of the joint distribution. Since the number of iterations to convergence is itself a random variable we simulated this as well, reporting the median number of iterations and associated performance. No claims are made for our simulation schedule or convergence criterion other than a common specification for all procedures in a given case.

Andrew Gelman (University of California, Berkeley) **and Donald B. Rubin** (Harvard University, Cambridge): We congratulate the authors and the Royal Statistical Society for gathering these interesting papers on the increasingly important topic of iterative simulation. Since we have our own forum for advocating the use of multiple (but not short) series to draw inferences from iterative simulation (Gelman and Rubin, 1992), we confine ourselves here to a simple but potentially important point. In the spirit of much of the discussion in these papers (e.g. Section 7 of Smith and Roberts), we would like to mention yet another variant of the Metropolis algorithm that may be useful in Bayesian simulation: an approximate Gibbs sampler.

As noted by various researchers (e.g. Tierney (1991)), one iteration of the Gibbs sampler can be viewed as d steps of the Metropolis–Hastings algorithm, where each step corresponds to one of the d conditional distributions $\pi(x_i|x_{-i})$, which define the Gibbs sampler and the joint distribution $\pi(x)$. For some problems, sampling from some, or all, of the correct conditional distributions is impossible, although approximations, $g_i(x_i|x_{-i})$, are available. Performing the Gibbs sampler with the approximate conditional distributions instead of the correct distributions will not work, in that the iterations will not converge to the desired joint distribution π. The Metropolis–Hastings algorithm, however, can be used for each of the d steps of an approximate Gibbs sampler and will converge to the correct joint distribution. The transition probability function at the ith Metropolis step at iteration t is then, in the notation of Smith and Roberts,

$$q_{it}(x, x') = \begin{cases} g_i(x_i'|x_{-i}) & \text{if } x'_{-i} = x_{-i}, \\ 0 & \text{otherwise,} \end{cases}$$

and the ratio of importance ratios is

$$\begin{aligned} r &= \frac{\pi(x')\, q_{it}(x', x)}{\pi(x)\, q_{it}(x, x')} \\ &= \frac{\pi(x_i'|x_{-i})\, g_i(x_i|x_{-i})}{\pi(x_i|x_{-i})\, g_i(x_i'|x_{-i})}, \end{aligned}$$

which is identically equal to 1 only if $g_i(x_i|x_{-i}) \equiv \pi(x_i|x_{-i})$. When g_i is an approximation, the Metropolis step will have a positive probability of not jumping.

This Metropolis–approximate Gibbs sampler should be useful in at least two situations: to correct for an analytical approximation and when using the values of the distribution computed at a discrete set of points. An example of an analytical approximation before the Gibbs sampler appears in Dempster *et al.* (1983), who approximate a binomial likelihood by a normal distribution on the logit scale to be conjugate with a normal prior distribution on the parameters. Discrete approximations commonly arise in the Gibbs sampler when a conditional density can be computed at several values (perhaps with difficulty) but not directly sampled from. In this case, if a method such as adaptive rejection sampling (Gilks and Wild, 1992) is unavailable, an approximate conditional density can be created by interpolating the density calculated at a few points, and then the Metropolis–approximate Gibbs sampler step can be applied.

Donald Geman (University of Massachusetts, Amherst): Recently, we hear that the Markov chain Monte Carlo (MCMC) method is 'revolutionizing' Bayesian statistics, and these papers suggest that. Applications outside spatial statistics (let alone image analysis) are proliferating, and the MCMC method has become an area of study in its own right.