YANAN FAN, DAVID S. LESLIE and MATTHEW P. WAND
(*University of New South Wales, Australia and University of Bristol, UK*)

We would like to congratulate the authors on their efforts in presenting a unified approached to the use of sequential Monte Carlo (SMC) samplers in Bayesian computation. In this, and the companion publication (Del Moral et al. 2006), the authors illustrate the use of SMC as an alternative to Markov chain Monte Carlo (MCMC).

These methods have several advantages over traditional MCMC methods. Firstly, unlike MCMC, SMC methods do not face the sometimes contentious issue of diagnosing convergence of a Markov chain. Secondly, in problems where mixing is chronically slow, this method appear to offer a more efficient alternative, see Sisson *et al* (2006) for example. Finally, as the authors point out, adaptive proposals for transition kernels can easily be applied since the validity of SMC does not rely on ergodic properties of any Markov chain. This last property may give the SMC approach more scope for improving algorithm efficiency than MCMC.

In reference to the binary probit regression model presented in Section 4.2, the authors chose to use a multivariate normal distribution as the initial importance distribution, with parameter value given by simulated estimates from an MCMC sampler. An alternative, more efficient strategy may be to estimate the parameters of the multivariate normal distribution by fitting the frequentist binary probit regression model. One can obtain maximum likelihood and the associated variance-covariance estimates for this, and many other, models using standard statistical software packages. We are currently designing a SMC method to fit a general design generalized linear mixed model (GLMM) (Zhao et al. 2006) and find this approach to work well.

The authors adopt an MCMC kernel, and update the coefficients of the covariates β from its full conditional distribution. If one cannot sample directly from the full conditional distributions, a Metropolis–Hastings kernel may be used. The choice of scaling parameters in such kernels can greatly influence the performance of the sampler, and it is not clear if optimal scaling techniques developed in the MCMC literature are immediately applicable here. In our current work to use these techniques for GLMMs we have found that using the variance–covariance estimate from the frequentist model as a guide for scaling the MH proposal variance works well. In general, can the authors offer any guidance on the properties of optimal MCMC kernels for use with SMC samplers?