The Annals of Statistics
2006, Vol. 34, No. 4, 1646–1652
DOI: 10.1214/009053606000000524
Main article DOI: 10.1214/009053606000000515

REJOINDER

BY S. C. KOU, QING ZHOU AND WING H. WONG

Harvard University, Harvard University and Stanford University

We thank the discussants for their thoughtful comments and the time they have devoted to this project. As a variety of issues have been raised, we shall present our discussion in several topics, and then address specific questions asked by particular discussants.

1. Sampling algorithms. The widely used state-of-the-art sampling algorithms in scientific computing include temperature-domain methods, such as parallel tempering and simulated tempering, energy-domain methods, such as multicanonical sampling and the EE sampler, and methods involving expanding the sampling/parameter space. The last group includes the Swendsen–Wang type algorithms for lattice models, as Wu and Zhu pointed out, and the group Monte Carlo method [1]. If designed properly, these sampling-space-expansion methods could be very efficient, as Wu and Zhu's example in computer vision illustrated. However, since they tend to be problem-specific, we did not compare the EE sampler with them. The comparison in the paper is mainly between the EE sampler and parallel tempering. Atchadé and Liu's comparison between the EE sampler and the multicanonical sampling thus complements our result. It has been more than 15 years since multicanonical sampling was first introduced. However, we feel that there are still some conceptual questions that remain unanswered. In particular, the key idea of multicanonical sampling is to produce a flat distribution in the energy domain. But we still do not have a simple intuitive explanation of (i) why focusing on the energy works, (ii) why a distribution flat in the energy is sought, and (iii) how such a distribution helps the sampling in the original sample space. The EE sampler, on the other hand, offers clear intuition and a visual picture: the idea is simply to "walk" on the equi-energy sets, and hence focusing on the energy directly helps avoid local trapping. In fact, the numerical results in Atchadé and Liu's comment clearly demonstrate the advantage of EE over multicanonical sampling in the 20 normal mixture example. Specifically, their Table 1 shows that in terms of estimating the probabilities of visiting each mode, the EE sampler is about two to three times more efficient. We think that estimating the probability of visiting individual modes provides a more sensitive measure of the performance, the reason being that even if a sampler misses two or three modes in each run, the sample average of the first and second moments could still be quite

Received December 2005.