BAYESIAN STATISTICS 9,
J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid,
D. Heckerman, A. F. M. Smith and M. West (Eds.)

Parameter inference for stochastic kinetic models of bacterial gene regulation: a Bayesian approach to systems biology: a discussion

Nicolas Chopin & Christian P. Robert
CREST (ENSAE), Paris and Université Dauphine
nicolas.chopin@ensae.fr, xian@ceremade.dauphine.fr

Summary

In this discussion, we reflect on the links between the likelihood-free method of the author and of recent developments by Møller et al. (2006) and Friel and Pettitt (2008), as well as the ABC literature (Beaumont et al., 2002).

While very much impressed by the scope of the chimical reaction models handled by Professor Wilkinson, we will (presumably predictably!) focus on the simulation aspects of his paper.

First, the solution proposed by the author to overcome the difficulties of handling the complex likelihood $\pi(x|\theta)$ reminds us of the auxiliary completion of Møller et al. (2006), who created (as well) an auxiliary duplicate of the data x and a pseudo-posterior on the duplicate to overcome computing the normalising constant in $\pi(x|\theta)$. As pointed out in Cucala et al. (2009), the choice of the completion distribution in Møller et al. (2006) may be detrimental to the convergence of the algorithm and we wonder if the same happens to the likelihood-free algorithm of the author.

Second, the dismissal of ABC (Approximate Bayesian computation, see, e.g., Grelaud et al., 2009) as being difficult to calibrate and to automatise is slightly unfair in that the summary statistics used in ABC are generaly suggested by the practitioners. Sequential ABC has been studied in Beaumont et al. (2009) as well, bringing a correction to Sisson et al. (2007) and building up a population Monte Carlo scheme for the approximation of $\pi(\theta, \mathbf{x}|\mathcal{D})$.

Third, when considering the sequential solution of Professor Wilkinson, we wonder about the approximation effects due to (a) the use of a kernel at each time t and (b) the lack of correction of the paths up to time t when given the new data

N. Chopin and C. Robert are supported by the 2007–2010 grant ANR-07-BLAN-0237-01 "SP Bayes".

d_{t+1}, because this particle approach is bound to diverge quite rapidly from the true distribution.

Specifically, consider the same algorithm, but (a) with a fixed parameter θ (no estimation), and (b) no "slight perturbation" (Step 2a). Add a re-weighting step, where the current weight of each sampled trajectory is multiplied by the partial observation likelihood $P(\mathcal{D}_{t+1} \mid \mathcal{D}_t, x_{t+1})$ (which should depend in most cases on the part of the trajectory sampled between time t and $t+1$). Then one obtains (at a negligible extra cost), a valid sequential Monte Carlo (SMC) algorithm for a continuous-time hidden Markov model, as in e.g. Chopin and Varini (2007). If θ is included, then the algorithm remains valid (in the sense that the Monte Carlo error goes to zero as the number of trajectories goes to infinity), but it is likely to diverge over time (in the sense that the asymptotic variance typically grows quickly over time, see also our dicussions on the paper by Lopes et al. in this volume). The PMCMC approach of Andrieu et al. (2010), while expensive, may be a more reasonable approach in this case.

REFERENCES

Andrieu, C., A. Doucet, and R. Holenstein. 2010. Particle Markov chain Monte Carlo methods. *J. Royal Statist. Society Series B* 72(3): 269–342.

Beaumont, M., J.-M. Cornuet, J.-M. Marin, and C. Robert. 2009. Adaptive approximate Bayesian Computation. *Biometrika* 96(4): 983–990.

Beaumont, M., W. Zhang, and D. Balding. 2002. Approximate Bayesian Computation in Population Genetics. *Genetics* 162: 2025–2035.

Chopin, N. and E. Varini. 2007. Particle filtering for continuous-time hidden Markov models. *ESAIM: Proceedings* 19: 12–17.

Cucala, L., J.-M. Marin, C. Robert, and D. Titterington. 2009. Bayesian inference in k-nearest-neighbour classification models. *J. American Statist. Assoc.* 104 (485): 263–273.

Friel, N. and A. Pettitt. 2008. Marginal likelihood estimation via power posteriors. *J. Royal Statist. Society Series B* 70(3): 589–607.

Grelaud, A., J.-M. Marin, C. Robert, F. Rodolphe, and F. Tally. 2009. Likelihood-free methods for model choice in Gibbs random fields. *Bayesian Analysis* 3(2): 427–442.

Møller, J., A. Pettitt, R. Reeves, and K. Berthelsen. 2006. An efficient Markov chain Monte Carlo method for distributions with intractable normalising constants. *Biometrika* 93: 451–458.

Sisson, S. A., Y. Fan, and M. Tanaka. 2007. Sequential Monte Carlo without likelihoods. *Proc. Natl. Acad. Sci. USA* 104: 1760–1765.