



Parallel Tempering

dr Błażej Miasojedow and mgr Mateusz Łacki

Wydział Matematyki, Informatyki i Mechaniki

University of Warsaw, Poland

B.Miasojedow@mimuw.edu.pl

mateusz.lacki@biol.uw.edu.pl

Bayesian Modelling for Bioinformatics

Applications

- Hierarchical modelling for identification of co-expression patterns in microarray data by cluster analysis (Medvedovic *et al.*, 2004; Stingo and M., 2010)
- Assessing the importance of explanatory variables (Stingo and M., 2010)
- Model Selection

MCMC

- MCMC algorithms are used to simulate samples out of analytically untractable posterior distributions
 - Most popular algorithm: **Green-Metropolis-Hastings** (Geyer, 2012)
 - Generates a sequence of points that are thought of as being an instantiation of a Markov Chain

$$X \equiv \{X^{[k]}\}_{k=0}^\infty$$

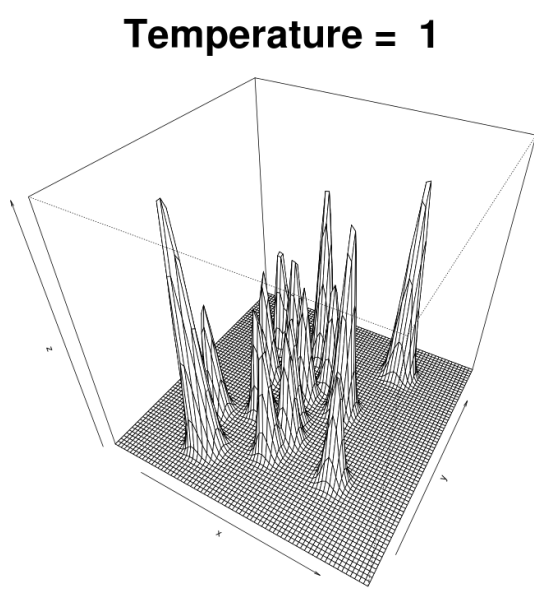
→ Approximates, thanks to Ergodic Theory, integrals

$$\mathcal{E}g(X) = \int_{\Omega} g(x) d\pi \approx \frac{1}{N} \sum_{i=1}^N g(X^{[i]}),$$

where π is the requested a posteriori distribution. In particular: we can approximate probabilities of any measurable set, $\mathcal{P}(A)$

- **GMH** estimates may suffer from poor mixing
 - Chain X restricted to user-provided number of iterations N could get stuck in a probability cluster
 - Multimodal priors result in multimodal posteriors
 - Multimodal priors are selected when we suspect that the phenomenon under study is not concentrated around a particular point

Example



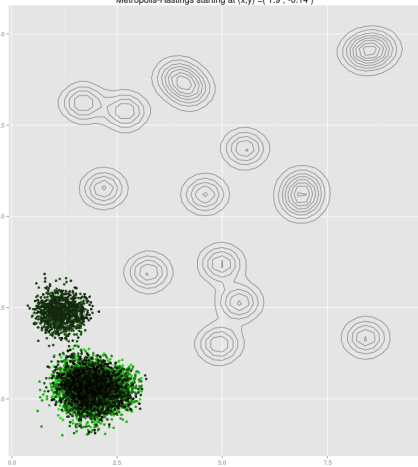
- Consider π given by density f that is a mixture of normal distributions

$$f(x) = \sum_{i=1}^{20} \frac{\omega_i}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(x - \mu_i)^2}{2\sigma_i^2}\right)$$

where σ_i are standard deviations, ω_i are weights, and μ_i are means.

- Some of the peaks mingle together to form bigger ones

- **GMH** draws sample points from only two modes
- Estimates of probabilities and moments are totally fallacious



¿Question?

How can we enhance mixing so that the STATE SPACE is better searched for probability clusters?

Parallel Tempering

Hello.

Hug Multimodality

Parallel Tempering rocks

Bye

bkas Bye World Medvedovic *et al.* (2004) Bye World Baragatti *et al.* (2013) Bye World Miasojedow *et al.* (2013)

References

BARAGATTI, M., GRIMALDI, A. and POMMERET, D. (2013). Parallel tempering with equi-energy moves. *Statistics and Computing*, 23 (3), 323–339.

GEYER, C. J. (2012). *Markov Chain Monte Carlo Lecture Notes*. Unpublished.

MEDVEDOVIC, M., YEUNG, K. and BUMGARDNER, R. (2004). Bayesian mixture model based clustering of replicated microarray data. *Bioinformatics*, 20 (8), 1222–1232.

MIASOJEDOW, B., MOULINES, E. and VIHOLA, M. (2013). An adaptive parallel tempering algorithm. *Journal of Computational and Graphical Statistics*, 1 (January).

STINGO, F. and M., V. (2010). *Bayesian Statistics 9*, Oxford University Press, chap. Bayesian Models for Variable Selection that Incorporate Biological Information.