Bayesian Model Comparison via Sequential Monte Carlo and Path Sampling

Yan Zhou[1], Adam M. Johansen and John A. D. Aston

Methodology

Bayesian Model Comparison

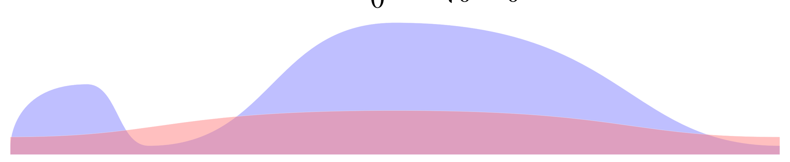
$$\pi(M_k|\mathbf{x}) = \frac{p(\mathbf{x}|M_k)\pi(M_k)}{p(\mathbf{x})} \qquad \text{(Model probability)}$$

$$p(\mathbf{x}|M_k) = \int_{\theta_k} p(\mathbf{x}|\theta_k, M_k)\pi(\theta_k|M_k)\,\mathrm{d}\theta_k \qquad \text{(Evidence, Marginal likelihood)}$$

Sequential Monte Carlo Sampler

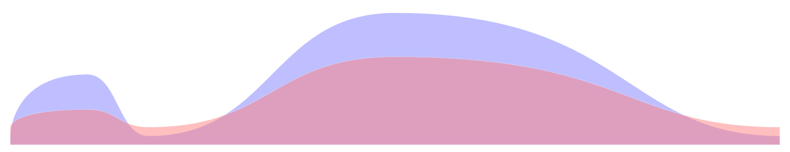
$$\gamma_n(\theta_k) = \frac{\pi_n(\theta_k)}{Z_n} \propto \pi(\theta_k|M_k)[p(\mathbf{x}|\theta_k, M_k)]^{\alpha(t_n/T)}$$

Draw $X_0^{(i)} \sim \eta_0(x_0)$



$$W_0^{(i)} \propto \gamma(X_0^{(i)})/\eta_0(X_0^{(i)})$$

Draw $X_n^{(i)} \sim K_n(X_{n-1}^{(i)}, x2)$



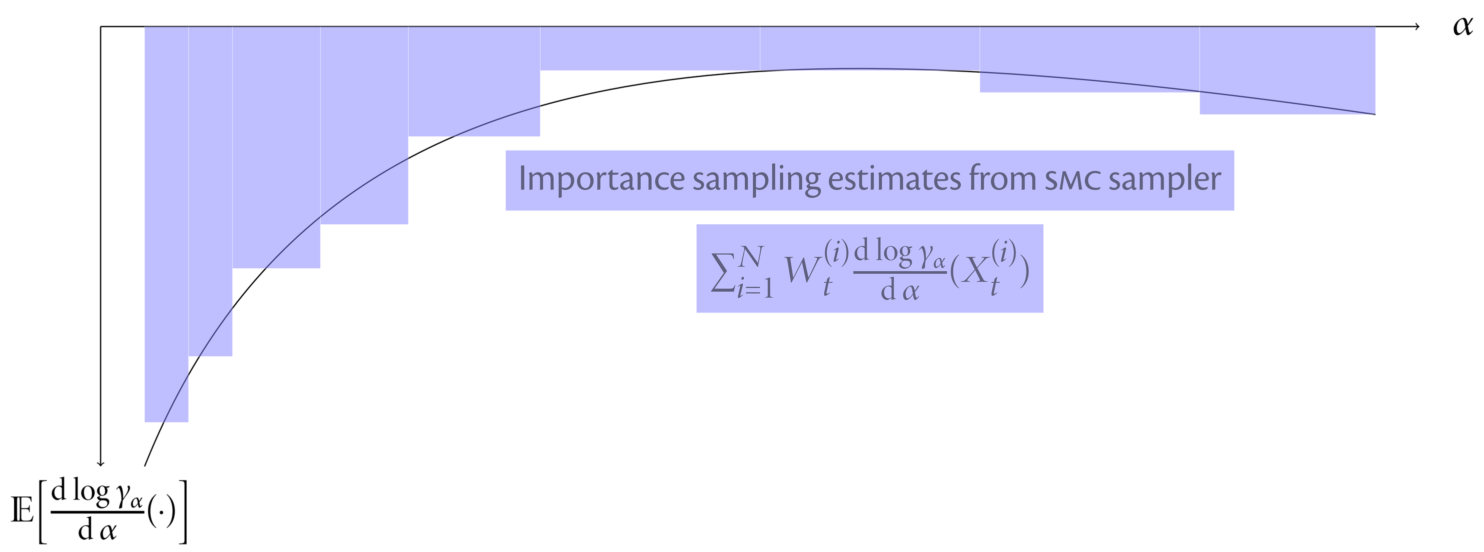
$$W_n^{(i)} \propto \tilde{w}_n^{(i)} W_{n-1}^{(i)}$$

$$\tilde{w}_n(x_{n-1}, x_n) = \frac{\gamma_n(x_n)L_{n-1}(x_n, x_{n-1})}{\gamma_{n-1}(x_{n-1})K_n(x_{n-1}, x_n)}$$

Estimate the Evidence – Normalizing Constants

$$\frac{\hat{Z}_n}{Z_{n-1}} = \sum_{i=1}^{N} W_{n-1}^{(i)}\tilde{w}_n(X_{n-1:n}^{(i)}) \qquad \text{(Direct estimator)}$$

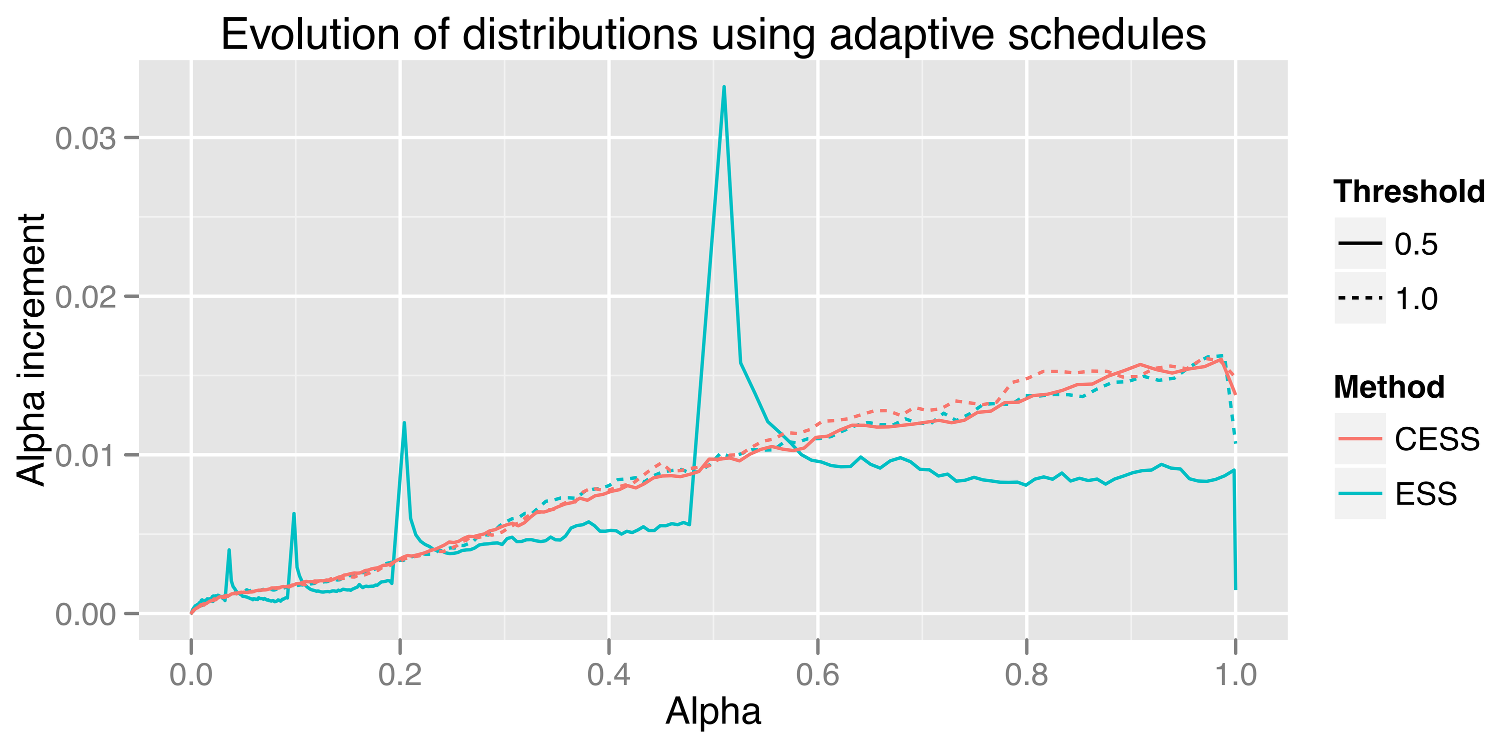
$$\log\Big(\frac{Z_1}{Z_0}\Big) = \int_0^1 \mathbb{E}_\alpha\Big[\frac{\mathrm{d}\log\gamma_\alpha}{\mathrm{d}\alpha}(\cdot)\Big]\,\mathrm{d}\alpha \qquad \text{(Path sampling estimator)}$$



Adaptive Algorithms

Adaptive Specifications of Distributions

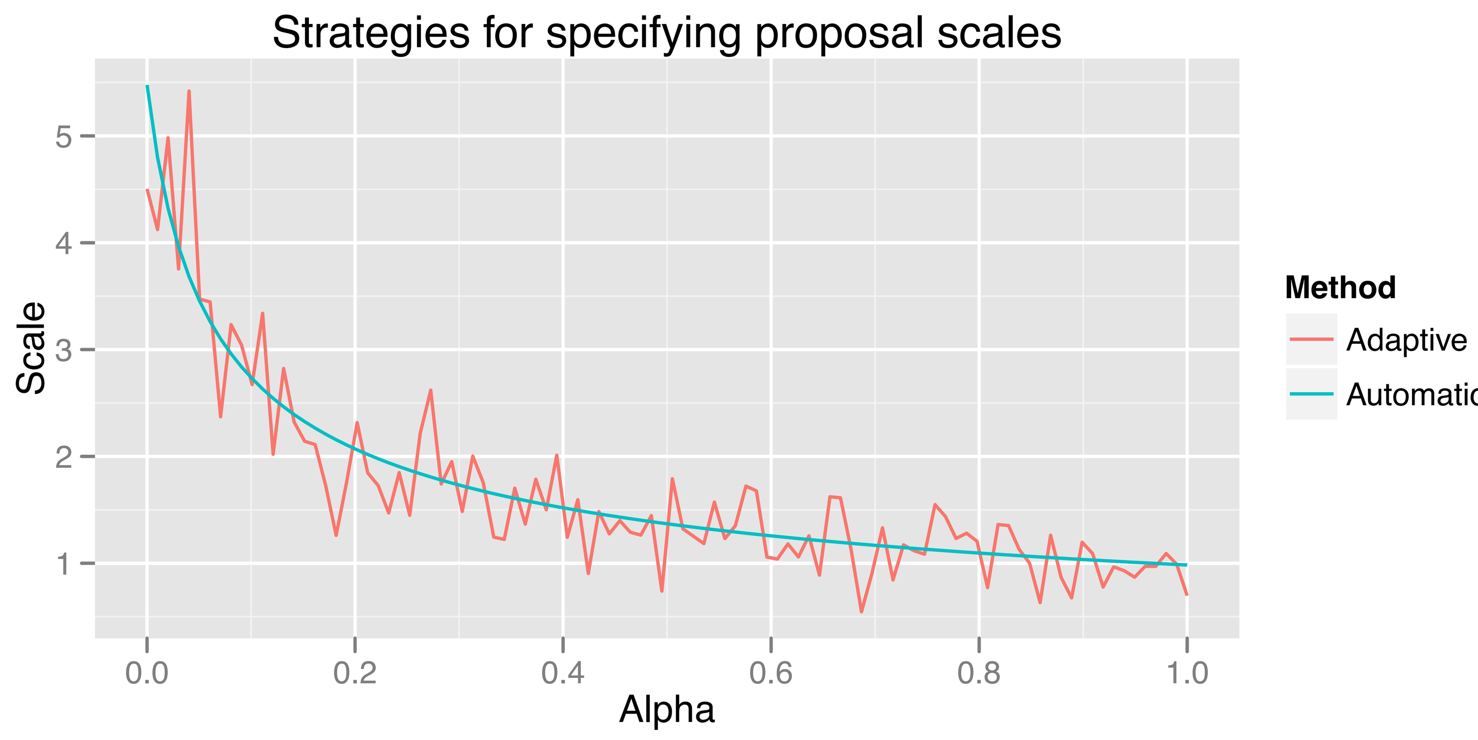
$$\mathrm{ESS}_n = \frac{(\sum_{j=1}^N W_{n-1}^{(j)}\tilde{w}_n^{(j)})^2}{\sum_{j=1}^N W_{n-1}^{(j)} W_{n-1}^{(j)}(\tilde{w}_n^{(j)})^2} \qquad \mathrm{CESS}_n = \frac{(\sum_{j=1}^N W_{n-1}^{(j)}\tilde{w}_n^{(j)})^2}{\sum_{j=1}^N \frac{1}{N} W_{n-1}^{(j)}(\tilde{w}_n^{(j)})^2} \qquad (1)$$



A threshold of ESS$/N = 1$ is equivalent to resampling at every time step.

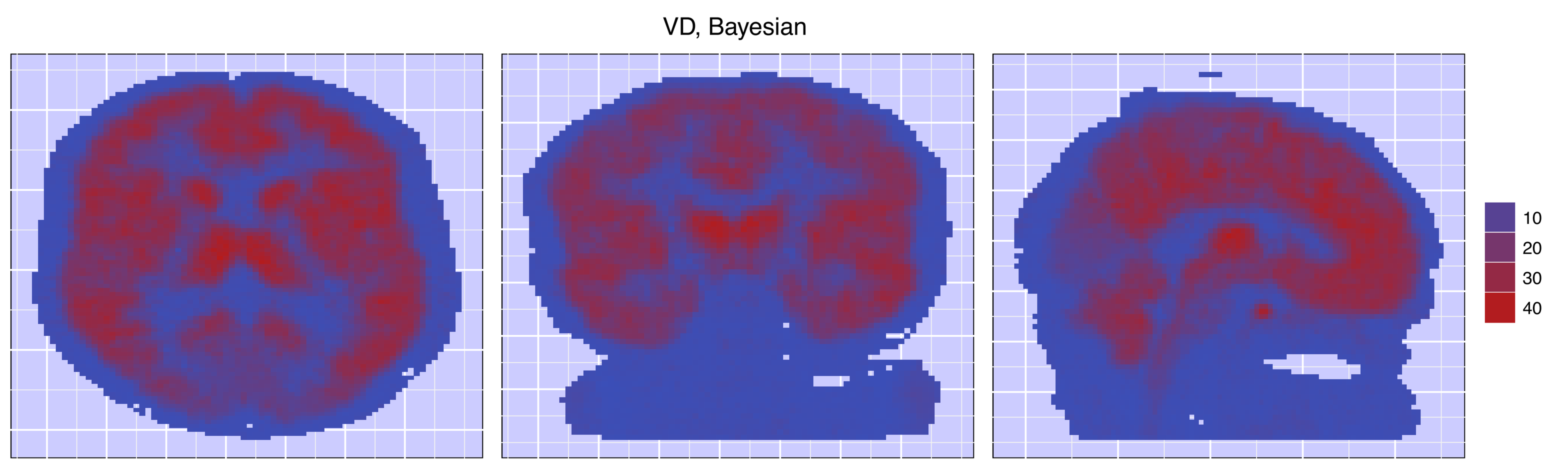
Automatic and Adaptive Specifications of Proposal Scales

Method		Pros	Cons
Automatic	$\sigma_t = \sigma^*\Big[\frac{b}{a + b\alpha(t/T)}\Big]^p$	Reliable if proper constants are found	Finding constants need a little efforts
Adaptive	σ_t = importance sampling estimates	Conceptually appealing	Can be unreliable

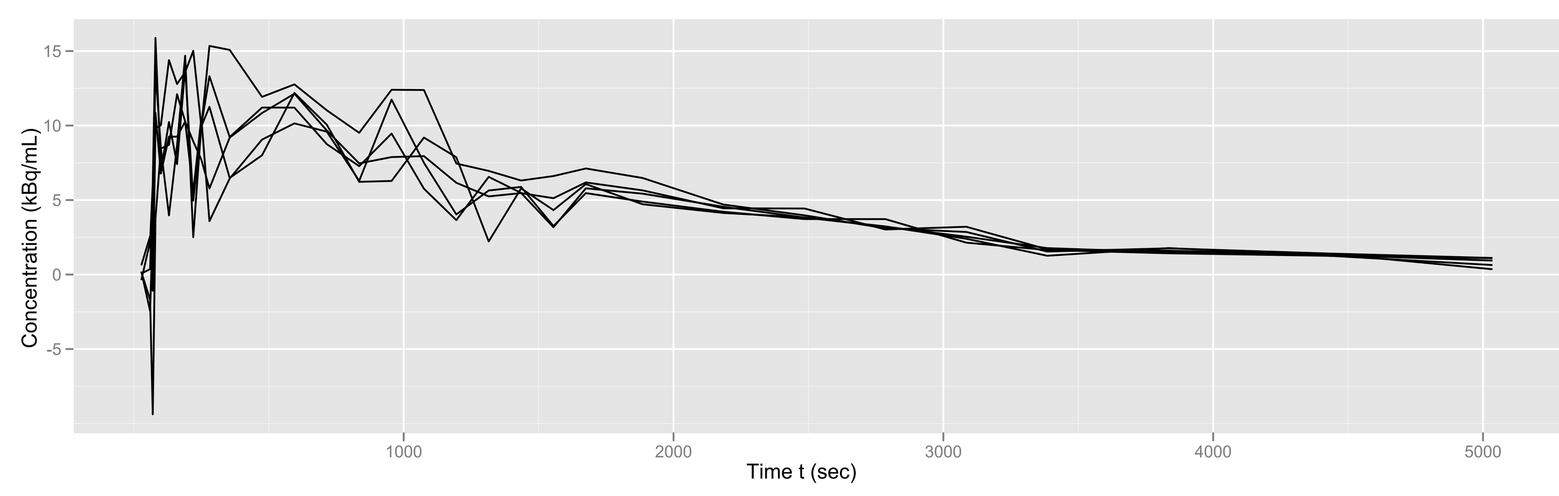


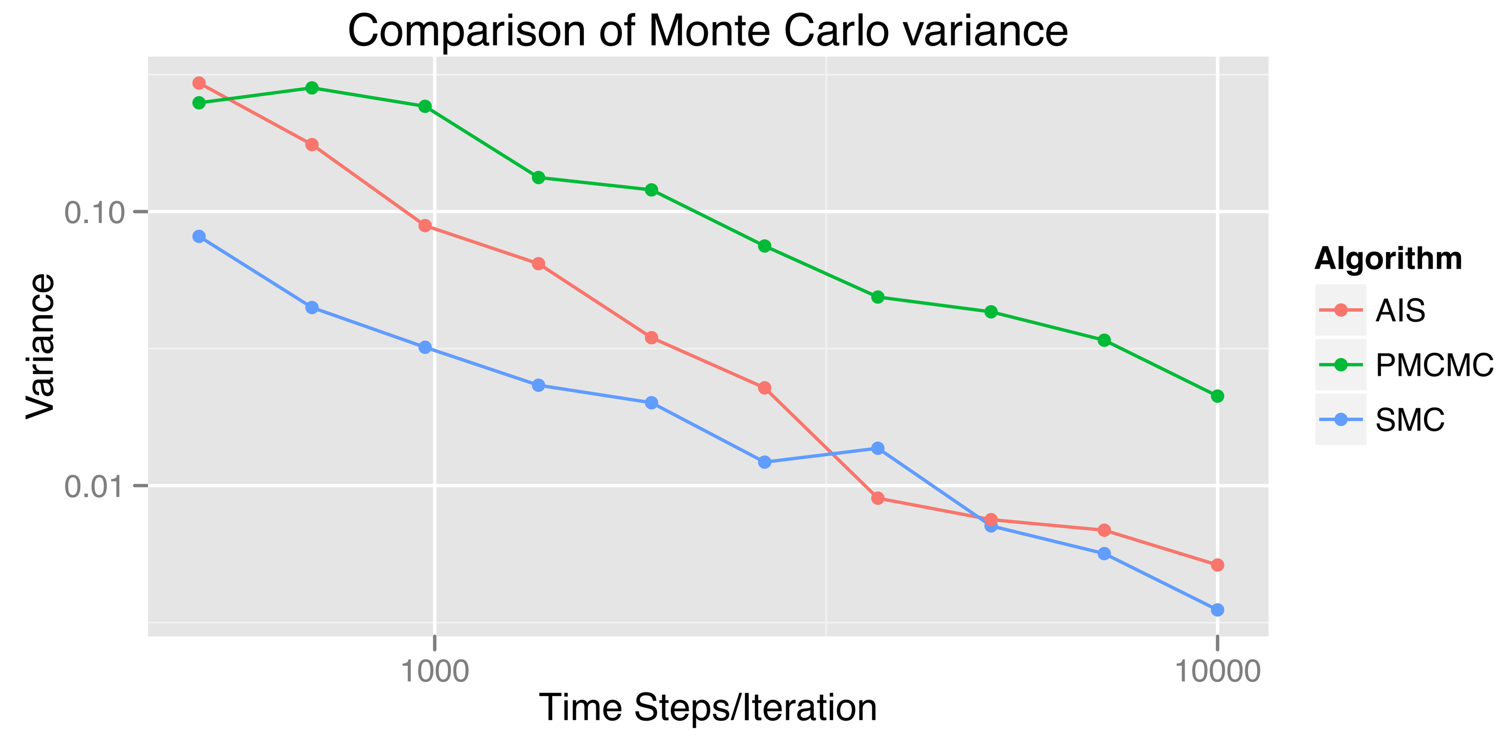
Results

Positron Emission Tomography Model



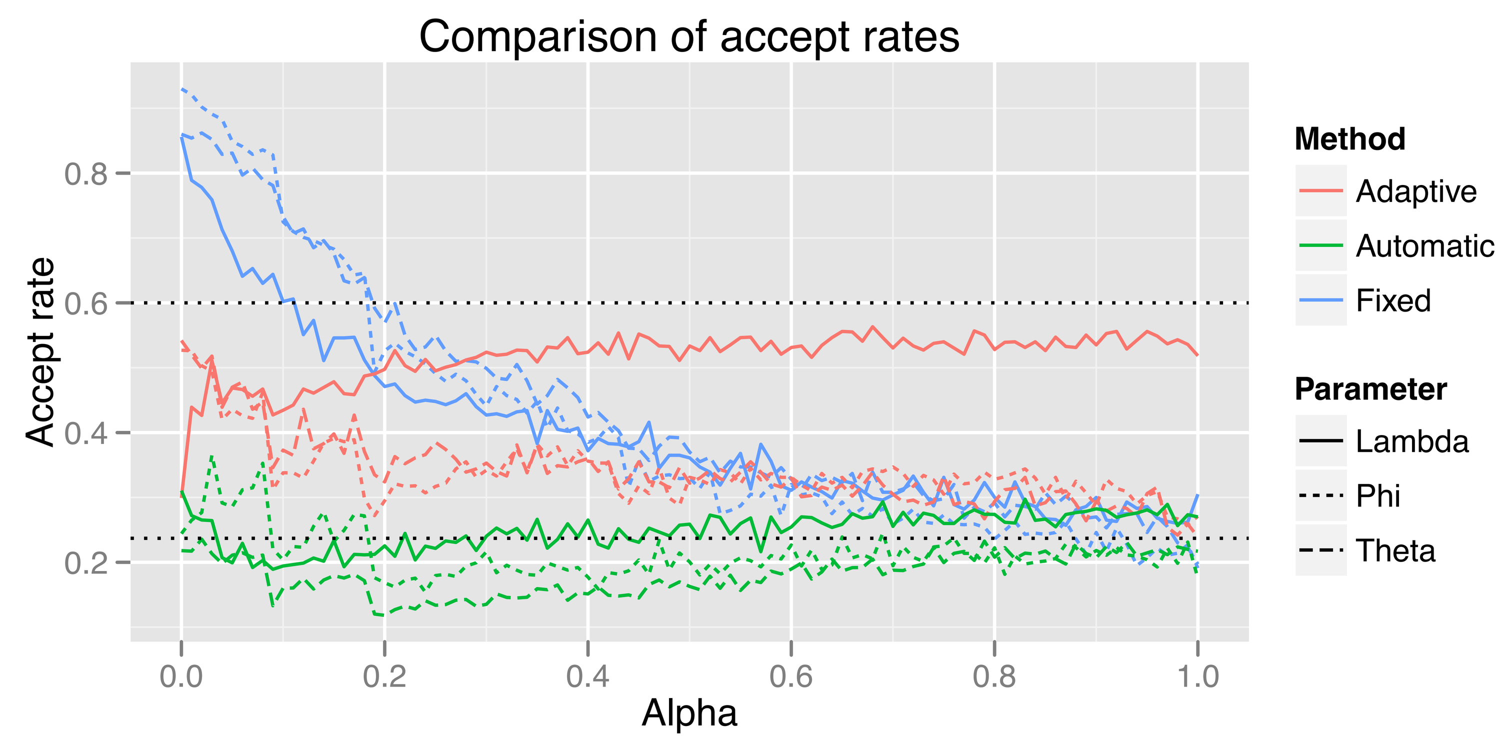
Typical PET data

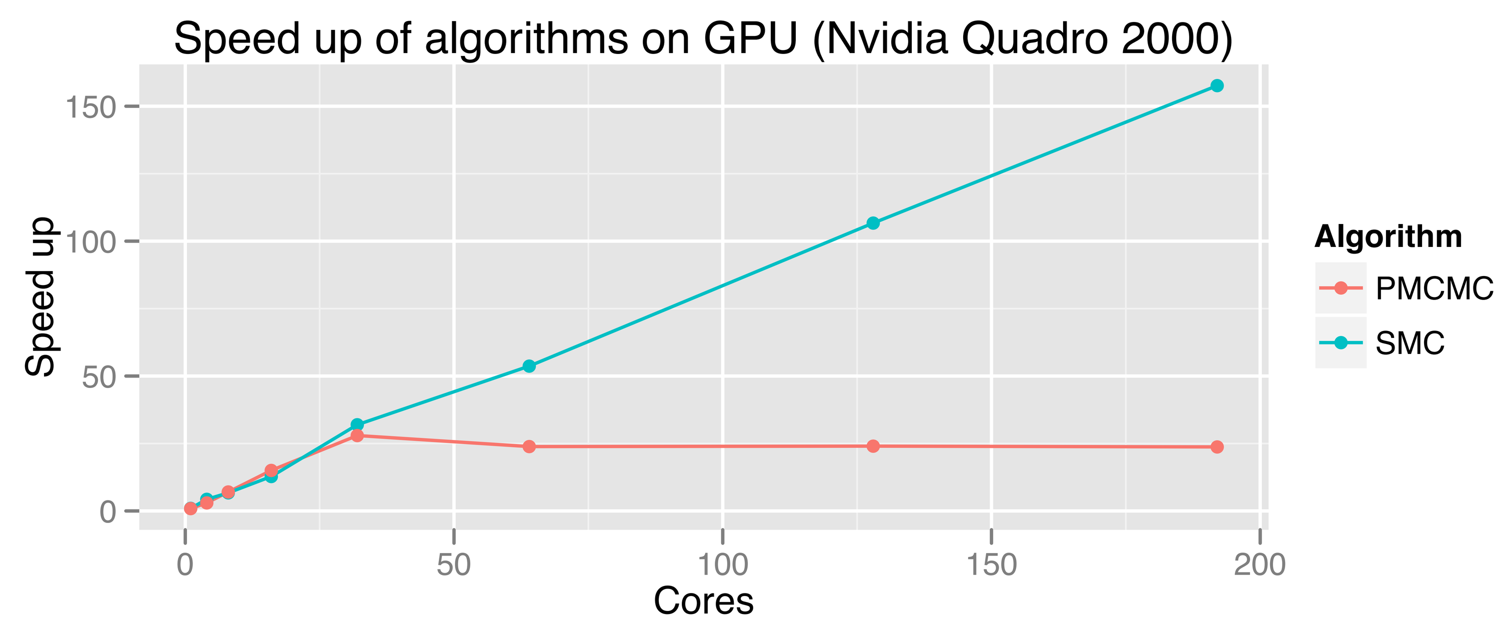




Comparison of marginal likelihood estimator using different schedules

	Scheduler			
Time steps T	Linear	Prior	Adaptive	$\overline{T}_{\text{Adaptive}}$
500	-49.9 ± 4.47	-39.1 ± 0.27	-39.0 ± 0.28	498
1000	-44.3 ± 1.87	-39.1 ± 0.20	-39.1 ± 0.21	987
2000	-41.5 ± 1.01	-39.1 ± 0.14	-39.1 ± 0.14	1986
5000	-40.0 ± 0.46	-39.1 ± 0.07	-39.1 ± 0.05	4875





[1] Email address: Yan.Zhou@warwick.ac.uk