Bank Sampling

by

Ben Allanach (University of Cambridge)

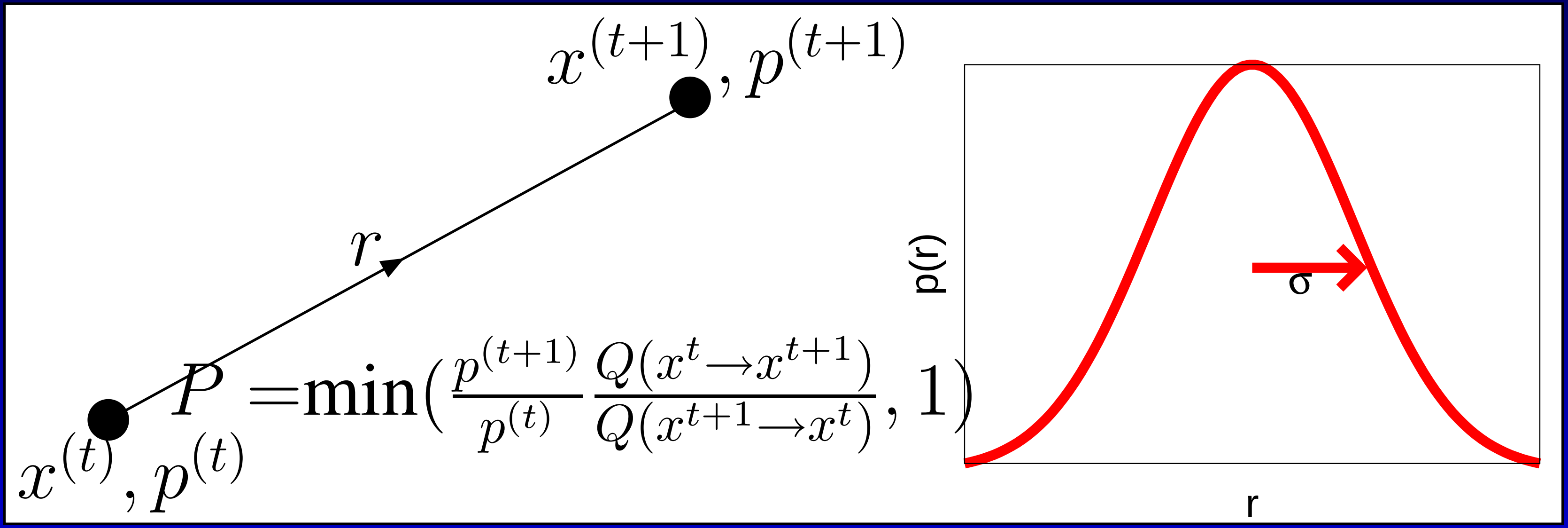
Talk outline [a]

- Isolated maxima
- A ‘bank’ of clues
- Multi-modal ellipsoidal sampling

[a] Based on work with C G Lester arXiv:0705.0486, to appear in Comp. Phys. Comm.

Metropolis-Hastings Sampling

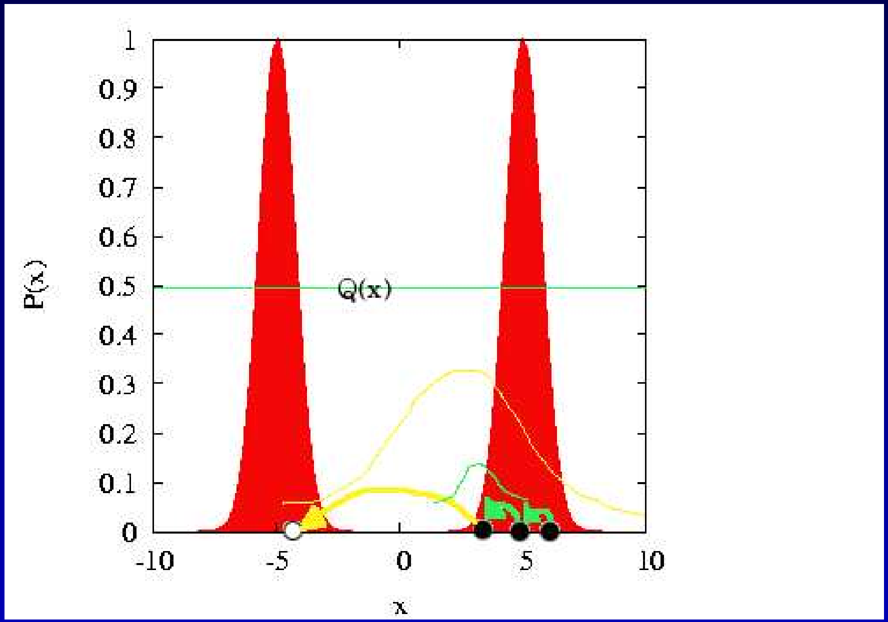
Metropolis-Hastings Markov chain consists of list of parameter points $x^{(t)}$ and associated posterior probabilities $p^{(t)}$.



Final density of x points $\propto p$. Required number of points goes *linearly* with number of dimensions.

Potential Problem

Often, people use a flat $Q(x)$. The trouble with this *"blind drunk"* sampling is the following situation:



Either large or small proposal widths σ lead to low efficiencies of sampling. Our proposal is to determine a $Q(x)$ closer to $P(x)$ *semi-automatically*.

Bank Sampling

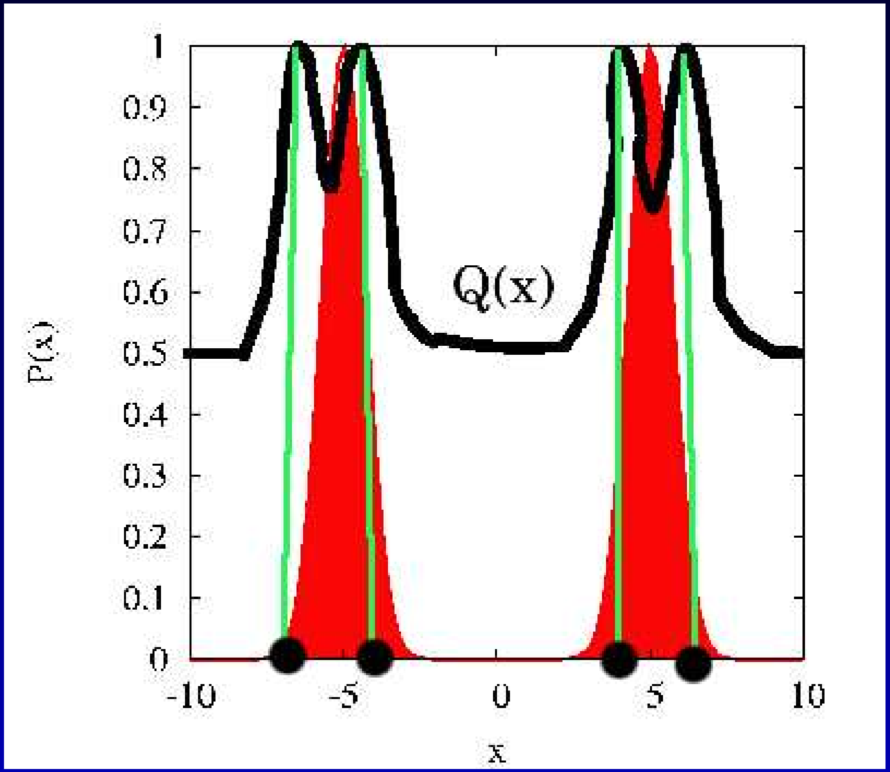


Figure 1: Bank points determined from previous runs: want to have at least one point in each maximum. *Knowledgeable drunk*

Proposal Distribution

$$Q_{bank}(\mathbf{x};\mathbf{x}^{(t)}) = (1-\lambda)K(\mathbf{x};\mathbf{x}^{(t)}) + \lambda \sum_{i=1}^{N} w_i K(\mathbf{x};\mathbf{y}^{(i)})$$

w_i are a set of N weights: $\sum_{i=1}^{N} w_i = 1$, $0 < \lambda < 1$, while K is the proposal distribution.

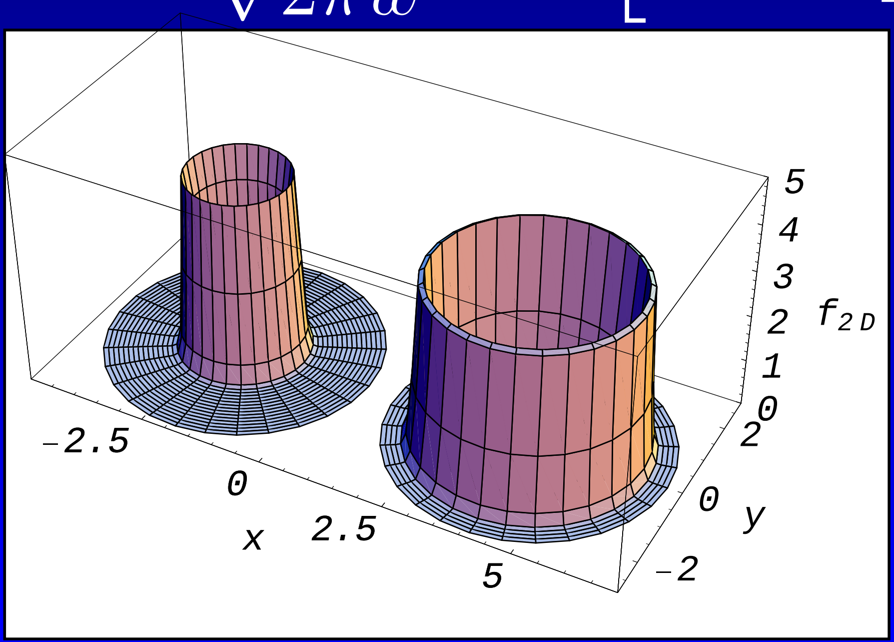
With probability $(1-\lambda)$ propose a local Metropolis step of the usual kind, i.e. "close" to the last point in the chain. With probability λ, teleport to the vicinity of one of the number of "banked" points, chosen with weight w_i from within the bank.

Example Distribution

$$f_{2D}(\mathbf{x}) = \text{circ}(\mathbf{x}; c_1, r_1, w_1) + \text{circ}(\mathbf{x}; c_2, r_2, w_2)$$

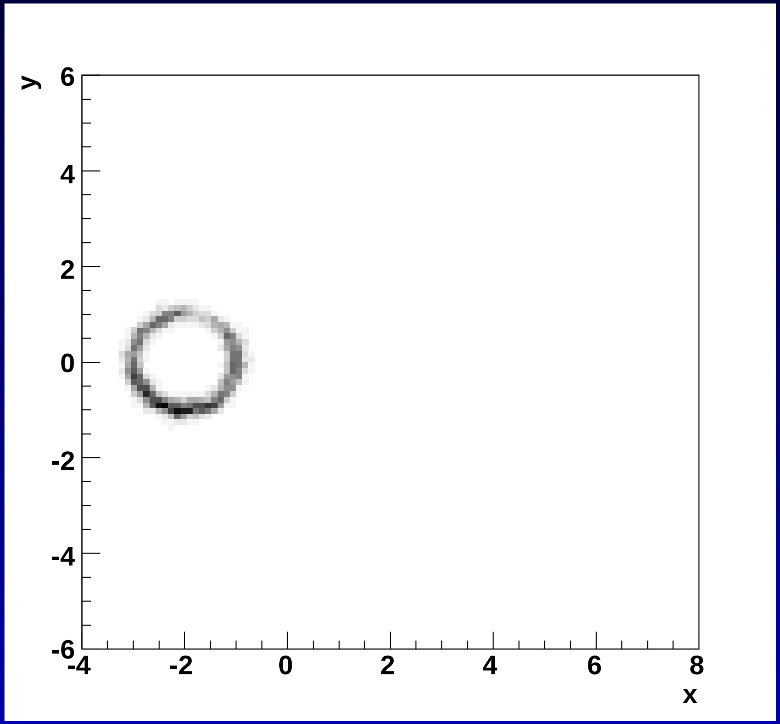
where $c_1 = (-2, 0)$, $r_1 = 1$, $w_1 = 0.1$, $c_2(+4, 0)$, $r_2 = 2$, $w_2 = 0.1$ and

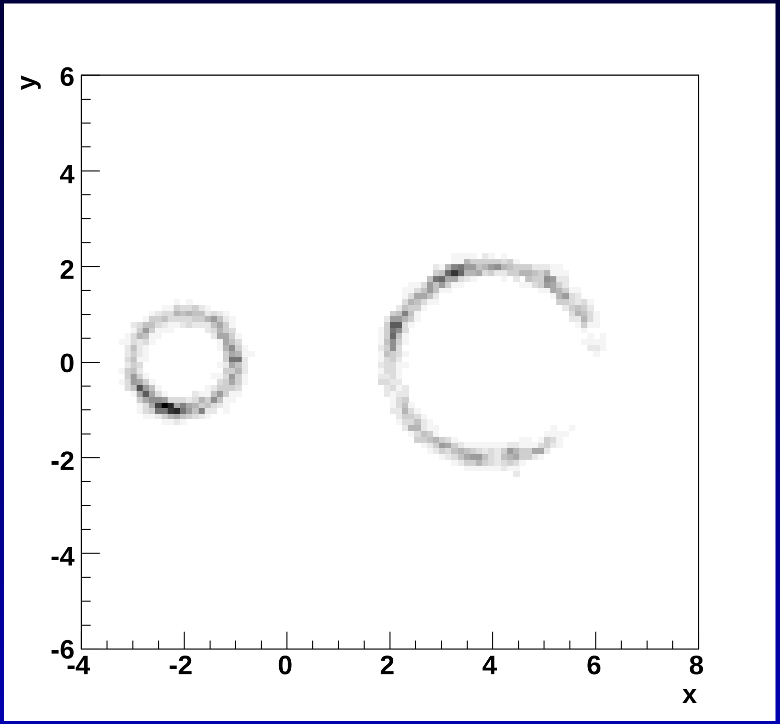
$$\text{circ}(\mathbf{x}; \mathbf{c}, r, w) = \frac{1}{\sqrt{2\pi w^2}} \exp\left[-\frac{(|\mathbf{x} - \mathbf{c}| - r)^2}{2w^2}\right].$$



Bank vs Metropolis

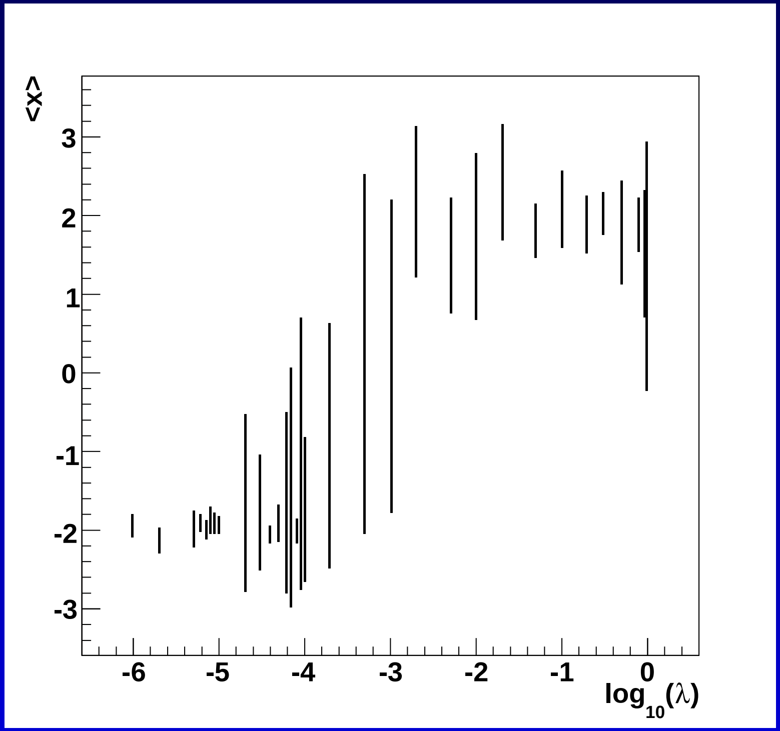
10 000 samples for MCMC and bank sampling:

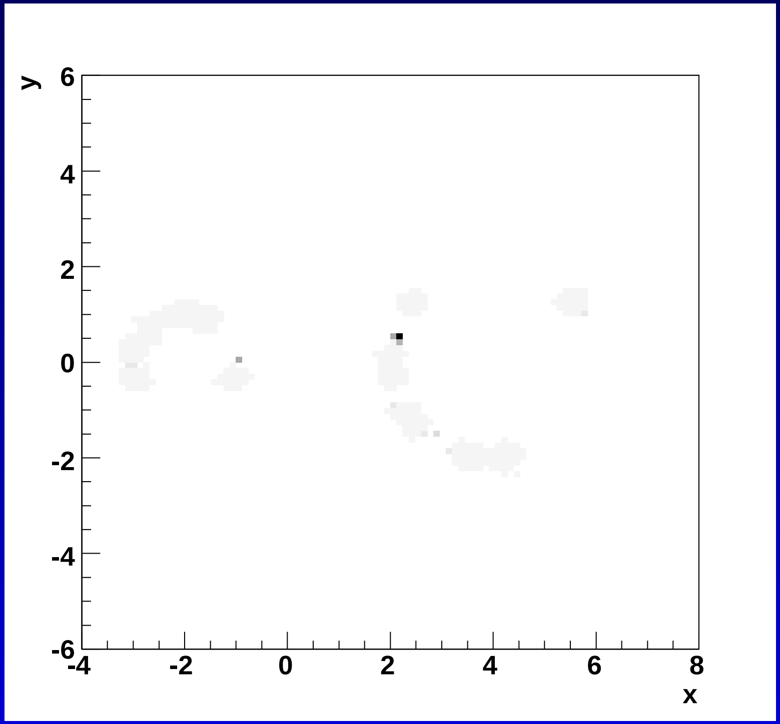




Safety with respect to λ

10 bank samplers, with 10 bank points generated in each circle: 10 000 samples. All started from $x = -2$. Correct $\langle x \rangle = 2$. $\lambda \approx 1$ is importance sampling limit.





Q: What values of λ are "safe"?

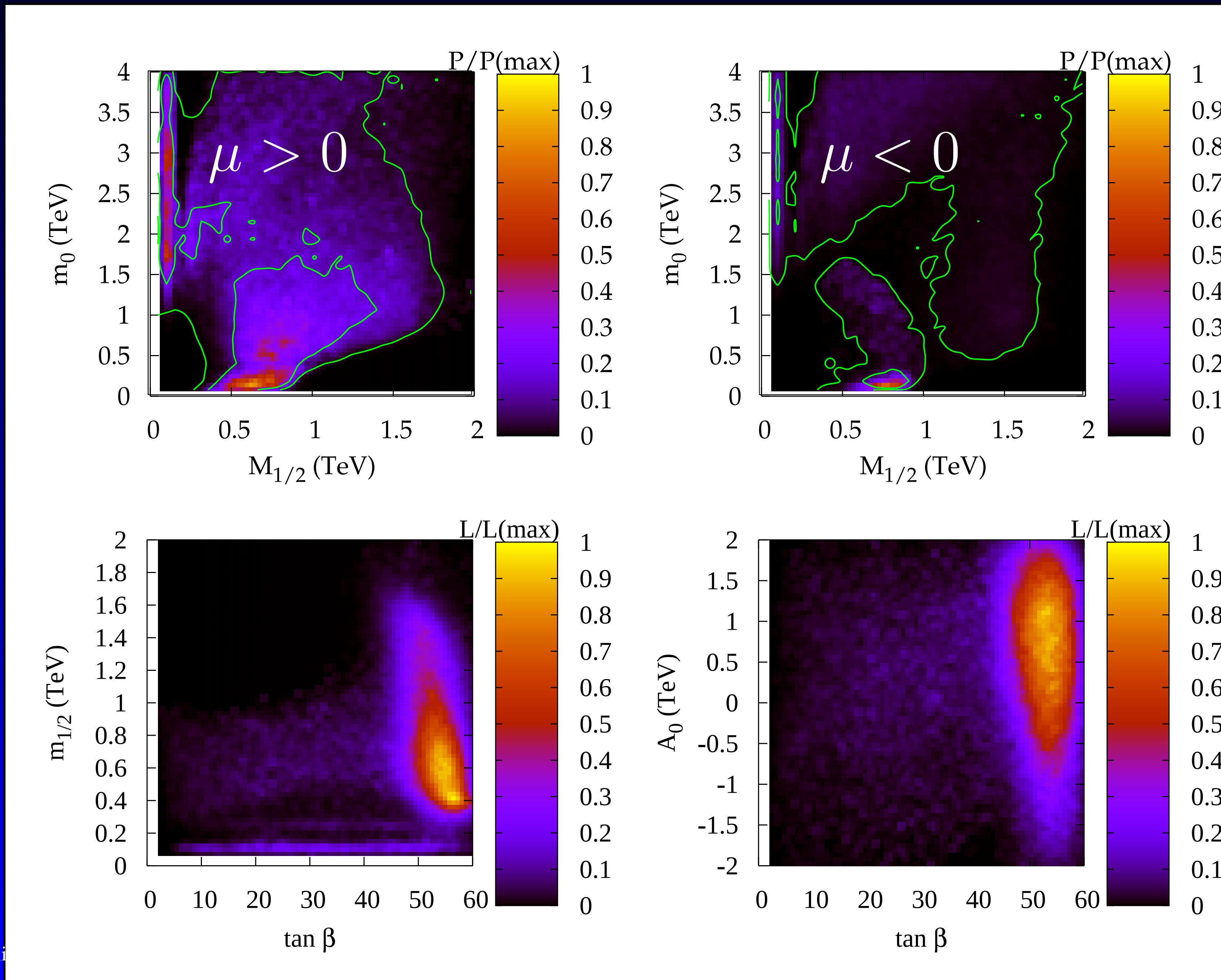
A: [0.001, 0.9]

Example Schedule

The following schedule was used by us in a particle physics context:

- 10 ordinary Metropolis runs with 100 000 steps and random starting points.
- Obtain 5000 random samples from the above to form the bank. *Make sure local maxima are represented*
- Weights were set proportional to the number of repetitions of the point in the bank sample (could also use the posterior probability).
- 10 bank sampling runs of 500 000 steps with the above bank.

Particle Physics fits



Nested Sampling

Skilling 2004 Evidence Z, parameters $\Theta_{d=1,...D}$, prior $\pi(\Theta)$ and likelihood L. Define the prior volume above a certain likelihood λ:

$$X(\lambda) = \int_{L(\Theta)>\lambda} \pi(\Theta) d^D\Theta$$

$$Z \equiv \int L(\Theta)\pi(\Theta) d^D\Theta = \int_0^1 L(X) dX$$

If one could sample from X, one could calculate Z.

We shall have M *live points*, $X_{i=0,...,M}$, $L_{i=0,...,M}$.

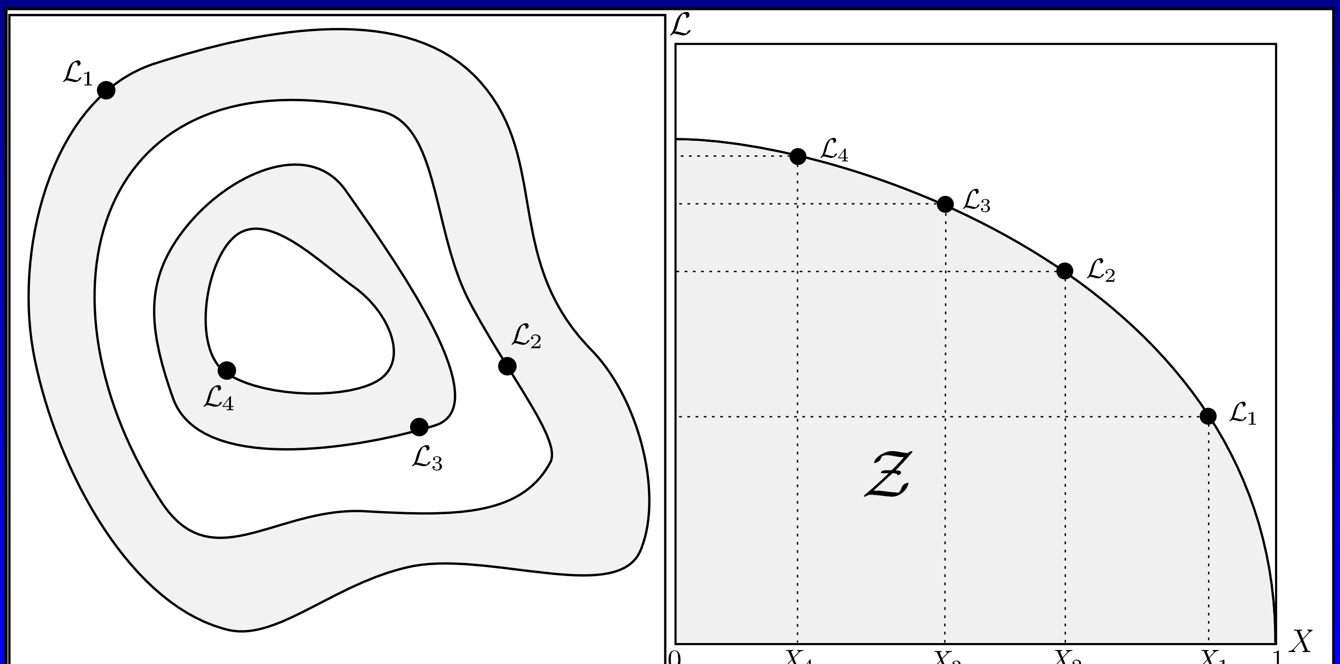
Science & Technology Facilities Council

Supersymmetry Cambridge Working group

Evidence Evaluation

$$0 < X_M < \cdots < X_1 < X_0 = 1,\ Z \approx \sum_i^M \frac{L_i}{2}(X_{i-1}-X$$

Sample from within each contour: $X_i = t_i X_{i-1}$ where $p(t_i) = M t_i^{M-1}$.



Ellipsoids/Metropolis

- Approximate iso-likelihood contour by N live points sampled [a] from prior with D-dimensional elllipsoid determined by covariance matrix of live points enlarged by a factor f: select new points from within this contour.
- Use a *clustering* algorithm (X means) to identify distinct well-separated ellipsoids. If a proposed point is in n overlapping ellispoids, accept with probability $1/n$.

[a] F Feroz and M Hobson arXiv:0704.3704

Two Chimneys Model

D	$N_{like}/1000$	efficiency/%	$N_{like}/1000$	eff/%
2	28	16	77	6
5	69	10	106	6
10	579	2	179	6
20	43 000	0.05	391	5
100			3007	4

Can use Metropolis with 20 steps instead of ellipsoids for very high D (on the right).

Summary

Bank sampling and multi-modal ellipsoidal sampling both deal efficiently with isolated maxima and allow determination of the ratios of the evidence. Compared to MME, bank sampling is:

- Unable to calculate the evidence
- Able to calculate evidence ratio *less* accurately
- Easier to implement (specific form of MH)
- Able to efficiently deal with cases where prior is numerically determined for each point