New sampler for cosmology

Anže Slosar, Brookhaven National Laboratory

BNL group meeting, 12/16

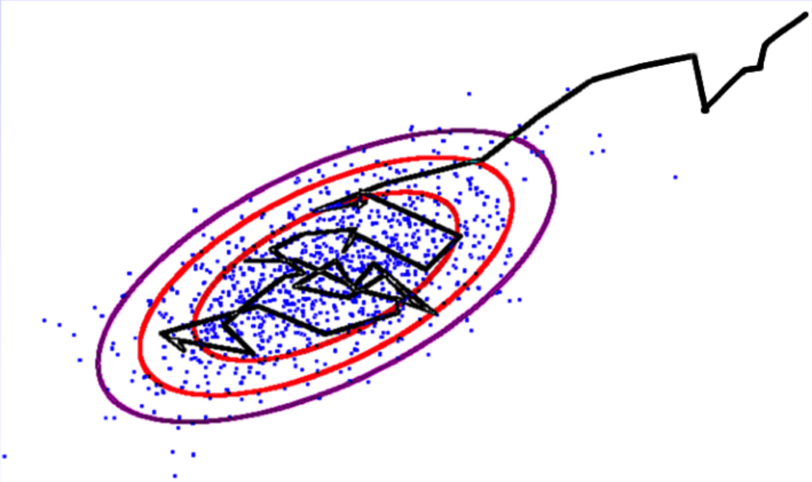
Introduction

- This is a half-baked project that I've been sitting on for over a year
- Now it is handed to Jose, who is fumbling around trying to incorporate it into SimpleMC
- It is a method to sample likelihood surface, whose main advantage is *scalability*: you don't win by having fewer likelihood evaluations but instead by having a algorithm that scales better to thousands of cores.
- Allows one to get answers in minutes on a couple of thousand of cores

The problem

- Question: How to get marginalised constraints for cosmological parameters when you have likelihood has $N > 10$ dimension and each evaluation is computationally expensive ($>$1s)
- Answer: use `CosmoMC` which runs Markov Chain Monte Carlo (MCMC)
- MCMC is an algorithm that "walks" around the likelihood and produces *samples*
- Integrals over likelihood can be converted to sums over samples

MCMC



Problems with MCMC

Markov Chain Monte Carlo does not scale very well:

- Scales perfectly for small number of chains, but not on modern architectures with 1000s of cores
- To run a `CosmoMC` chain you still run on 64 cores and wait for two days, instead of running on 10000 cores and wait 20 mins.
- But can't you run 1000 chains?
- Yes, but burn-in is a constant time process: one always needs to throw away some ~ 1000 steps, because either:
  - You start chains randomly – they need to burn in
  - You start chains at high-likelihood region - they need to decorrelate
- Both are inefficient: You take 1000 samples to burn-in, but then 100 samples on each chain is enough to get 100,00 samples – quite inefficient

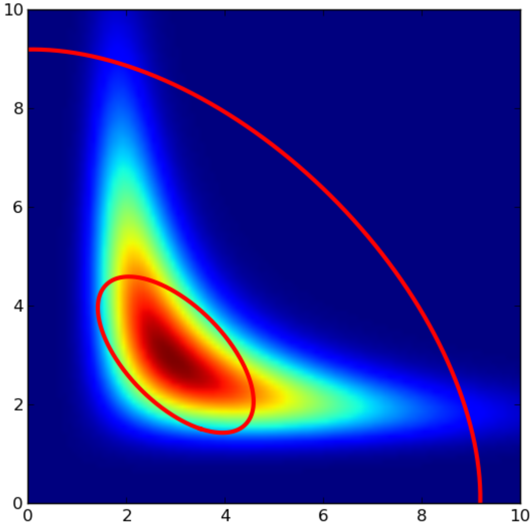
Importance Sampling

Assuming that one can sample from a known distribution, then one can weight samples to recover the effective sampling from a target distribution (whose properties one would like to study)

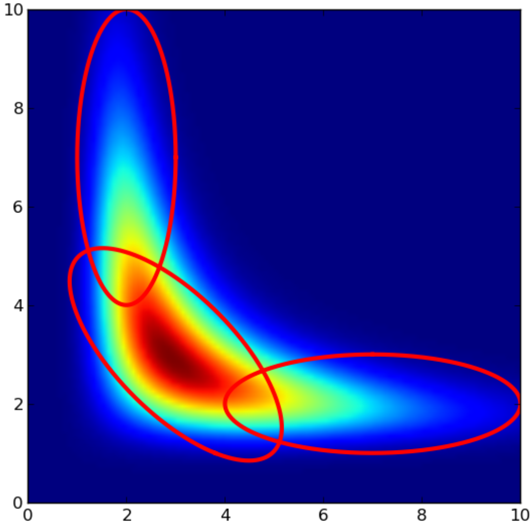
$$w_i = A\frac{L_\mathrm{t}(\mathbf{x}_i)}{L_\mathrm{s}(\mathbf{x}_i)}, \quad (1)$$

- Used to add a dataset to existing chains
- People tried to use it to sample cosmological likelihood directly using a Gaussian, but it fails miserably with bananas:
  - Either your Gaussian does not encompass the banana: weights blow up at the edges
  - Your Gaussian covers the banana, but also empty volume around it: most weights zero.

Why naive importance sampling doesn't work



But if you could do something like this?



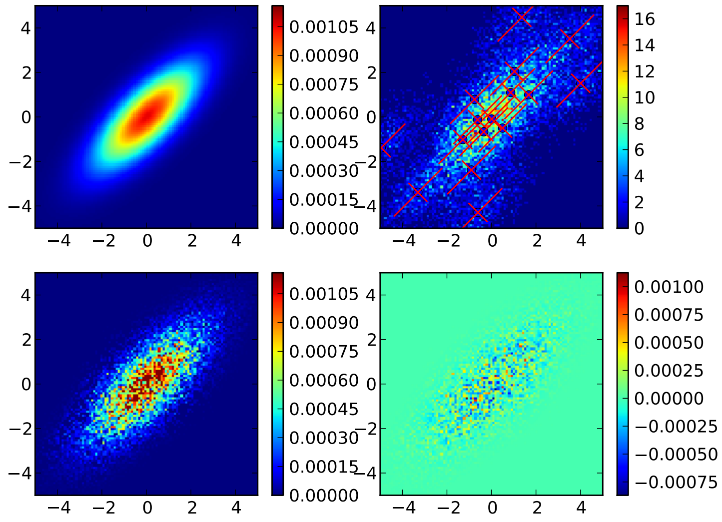
Guassian embedding sampler

1. Populate a list of Gaussians with a single Gaussian centered at a chosen starting point with suitable covariance.
2. Take N samples from the most recently added Gaussian in the list.
3. Calculate importance sample weights,

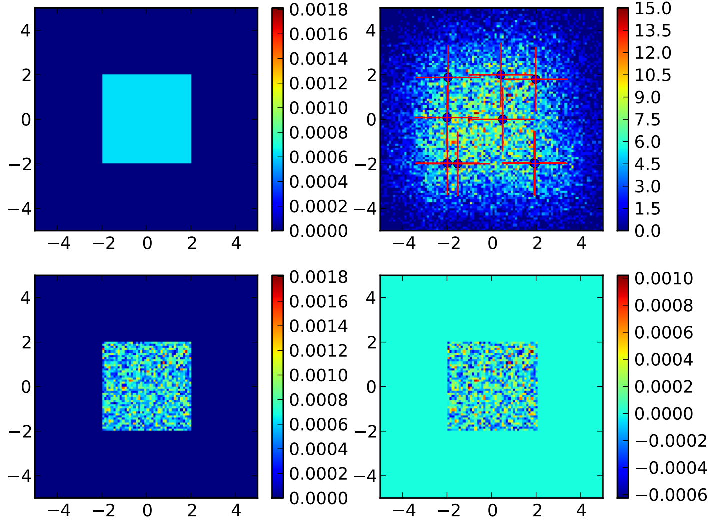
$$w_i = A \frac{L_{\mathrm{t}}(\mathbf{x}_i)}{\sum_{j=1\ldots M} G_j(\mathbf{x}_i - \mu_j, \mathrm{C}_j)}, \tag{2}$$

4. Add new Gaussian at the position of the largest importance weight
5. GOTO step 2

Test 1: Gaussian

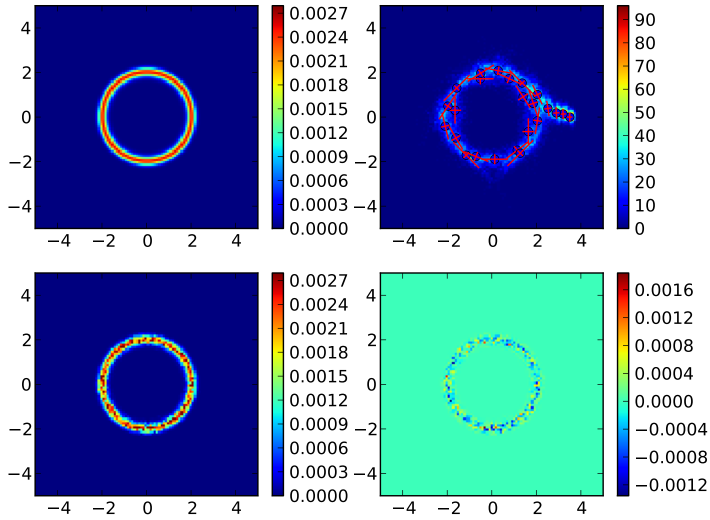


Test 2: Box



0.0018
0.0016
0.0014
0.0012
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000
15.0
13.5
12.0
10.5
9.0
7.5
6.0
4.5
3.0
1.5
0.0
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000
−0.0002
−0.0004
−0.0006

Test 3: Doughnut shaped banana

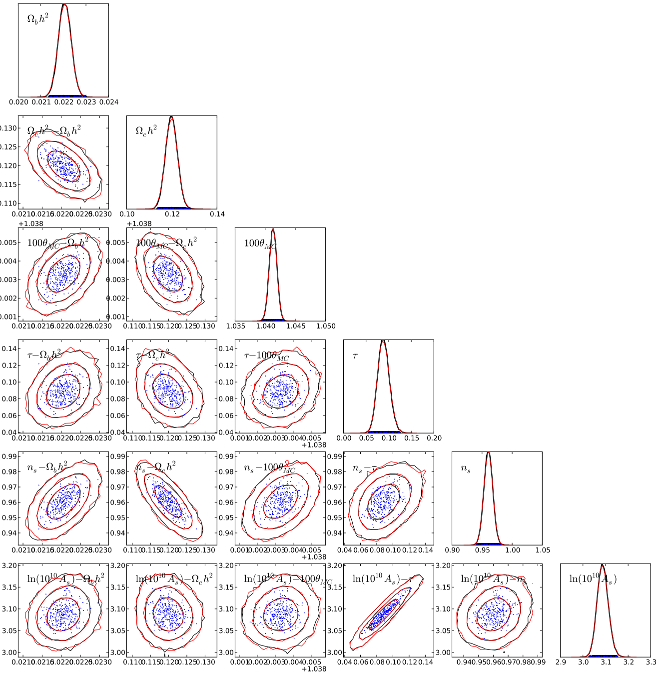


Convergence

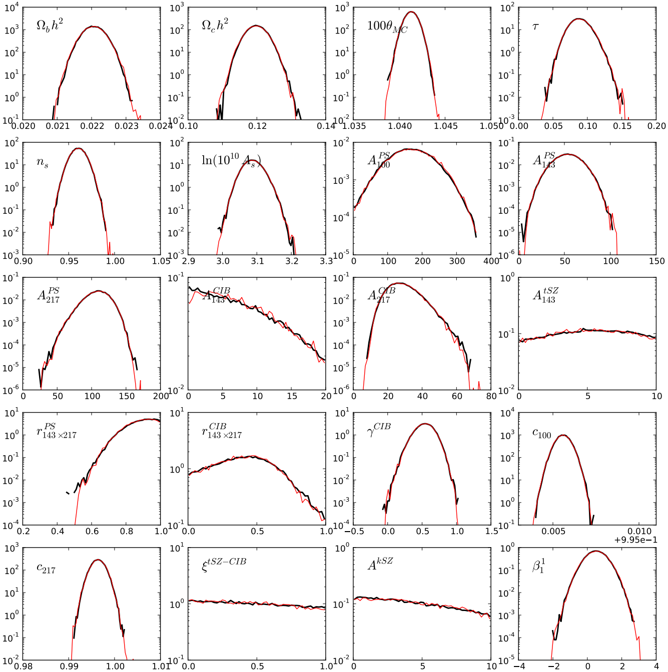
Tried several convergence tests. Jose is now working on my latest idea, namely that

$$c = \frac{\Delta\theta \mathrm{C}^{-1} \Delta\theta}{N_p}, \tag{3}$$

where $\Delta\theta$ is the shift in mean parameter values when the last Gaussian is added, is smaller than a certain number.



$\Omega_b h^2$
$\Omega_c h^2$
$\Omega_c h^2 - \Omega_b h^2$
$100\theta_{MC}$
$100\theta_{MC} - \Omega_b h^2$
$100\theta_{MC} - \Omega_c h^2$
τ
$\tau - \Omega_b h^2$
$\tau - \Omega_c h^2$
$\tau - 100\theta_{MC}$
n_s
$n_s - \Omega_b h^2$
$n_s - \Omega_c h^2$
$n_s - 100\theta_{MC}$
$n_s - \tau$
$\ln(10^{10} A_s)$
$\ln(10^{10} A_s) - \Omega_b h^2$
$\ln(10^{10} A_s) - \Omega_c h^2$
$\ln(10^{10} A_s) - 100\theta_{MC}$
$\ln(10^{10} A_s) - \tau$
$\ln(10^{10} A_s) - n_s$
+1.038



$\Omega_b h^2$
$\Omega_c h^2$
$100\theta_{MC}$
τ
n_s
$\ln(10^{10} A_s)$
A_{100}^{PS}
A_{143}^{PS}
A_{217}^{PS}
A_{143}^{CIB}
A_{217}^{CIB}
A_{143}^{tSZ}
$r_{143\times217}^{PS}$
$r_{143\times217}^{CIB}$
γ^{CIB}
c_{100}
+9.95e-1
c_{217}
$\xi^{tSZ-CIB}$
A^{kSZ}
β_1^1