Hybrid (Hamiltonian) Monte Carlo

- We would like to increase the acceptance rate to improve efficiency, and explore the target distribution efficiently ('good mixing')
- We have a hard problem in many dimensions. Solution:
- Make things harder: add in M auxiliary variables, one for each parameter in the model.
- Imagine each of the parameters in the problem as a coordinate.
- Target distribution ➔ effective potential
- For each coordinate HMC generates a generalised momentum.
- It then samples from the extended target distribution in 2M dimensions.

HMC

- HMC explores this 2M-dimensional space by treating the problem as a dynamical system, and evolving the phase space coordinates by solving the dynamical equations.
- Finally, it ignores the momenta (marginalising, as in MCMC), and this gives a sample of the original target distribution.
- May help with decorrelating the points in the chain.
- Invented by particle physicists (Duan et al 1987)

Theory

- Potential $U(\theta) = -\ln p(\theta)$
- For each θ_α, generate a momentum u_α.
- K.E. $K = u^T u/2$
- Define a Hamiltonian

$$H(\boldsymbol{\theta}, \mathbf{u}) \equiv U(\boldsymbol{\theta}) + K(\mathbf{u})$$

- and define an extended target density

$$p(\boldsymbol{\theta}, \mathbf{u}) = \exp\left[-H(\boldsymbol{\theta}, \mathbf{u})\right]$$

Magic of HMC

William Rowan Hamilton

- Evolve as a dynamical system

$$\dot{\theta}_\alpha = u_\alpha$$
$$\dot{u}_\alpha = -\frac{\partial H}{\partial \theta_\alpha}$$

- H remains constant, so extended target density is uniform – all points get accepted!
- Also, you can make big jumps – good mixing, if you generate a new u each time

Complications

- Evolving the system takes time. Take big steps.
- We don't know the complete U = - ln p (it's what we are looking for)
- If we can quickly evaluate the derivative of U, fine.
- We might *approximate* U (from a short MCMC)

$$U = \frac{1}{2}(\theta - \theta_0)_\alpha C^{-1}_{\alpha\beta}(\theta - \theta_0)_\beta$$

- H is therefore not constant
- Use Metropolis-Hastings. Accept new point with probability

$$min\left\{1, \exp\left[-H(\boldsymbol{\theta}^*, \mathbf{u}^*) + H(\boldsymbol{\theta}, \mathbf{u})\right]\right\}$$

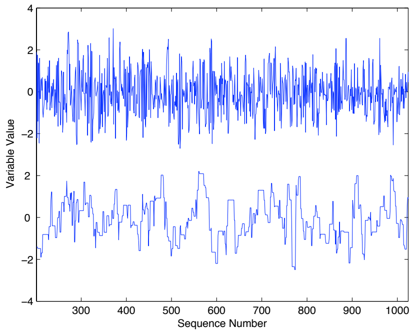
Algorithm

Hamiltonian Monte Carlo

```
initialize x_(0)
for i = 1 to N_samples
        u ~ N(0,1)
        (x*_(0), u*_(0)) = (x_(i-1), u)
        for j = 1 to N
                make a leapfrog move: (x*_(j-1), u*_(j-1)) → (x*_(j), u*_(j))
        end for
        (x*, u*) = (x_(N), u_(N))
        draw α ~ (0,1)
        if α < min{1, e^{-(H(x*,u*)-H(x,u))}}
                x_(i) = x*
        else
                x_(i) = x_(i-1)
end for
```

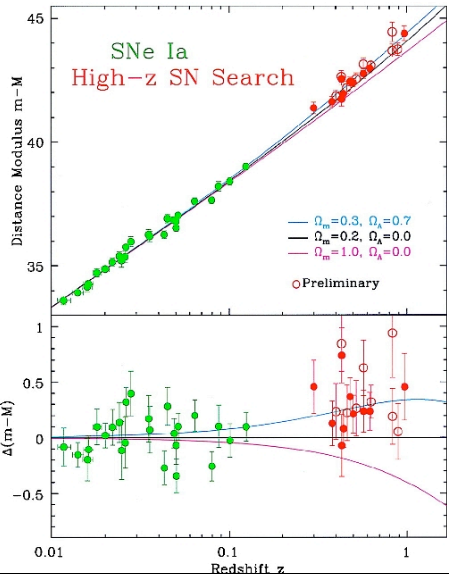
From Hajian 2006

HMC vs MCMC



Typical speed-ups: factor 4.

Supernova distances

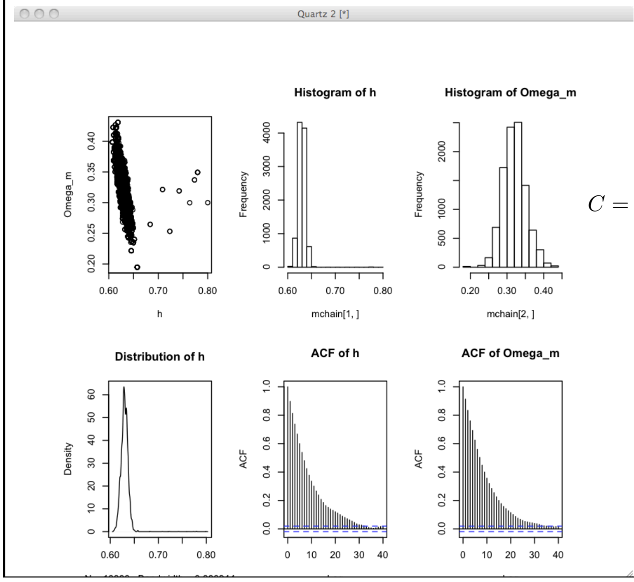


- Luminosity Distance depends on Ω_m and H_0 (for flat Universe)

$$f = \frac{L}{4\pi D_L^2}$$

$$D_L = \frac{c(1+z)}{H_0}\int_0^z \frac{dz'}{\sqrt{\Omega_m(1+z')^3 + 1 - \Omega_m}}$$

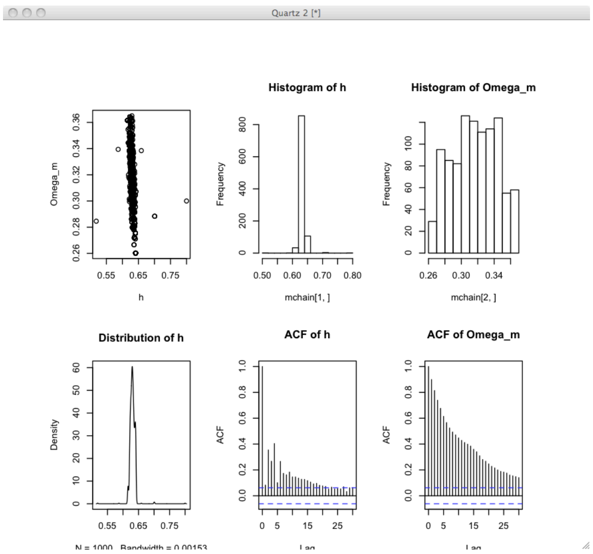
MCMC run



$$C = \begin{pmatrix} 7.1 \times 10^{-5} & -1.9 \times 10^{-4} \\ -1.9 \times 10^{-4} & 1.0 \times 10^{-3} \end{pmatrix}$$

Acceptance rate 0.15

HMC



Acceptance rate 0.4

Still work to do…