Introduction to Hamiltonian Monte Carlo

Patrick Ding, James Dole, Naveed Merchant

Physics Recap

- Kinetic energy - Energy an object has because it is in motion
  - Example: A drop of rain falling
  - Example: A wheel spinning

- Potential Energy - Energy an object has stored as a result of its position.
  - Example: A person holding a coin above the ground. When the coin is dropped, the potential energy is converted to kinetic energy and the coin falls.
  - Example: The voltage measured across the terminals of a battery.

Physics Recap - Hamiltonian

$$H(q, p) = K + U$$

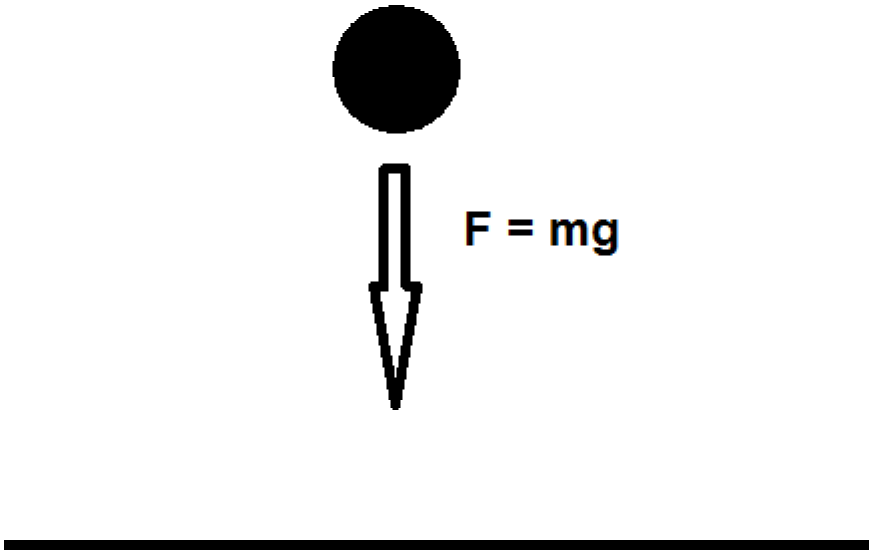
- q - position
- p - momentum

$$H(q, p) = K + U$$

$$\frac{dp}{dt} = -\frac{\partial H}{\partial q} \qquad \frac{dq}{dt} = \frac{\partial H}{\partial p}$$

Physics Recap - Hamiltonian Example

Example: object in free fall



Physics Recap - Hamiltonian Example

$$\begin{aligned} H &= K + U \\ &= \frac{1}{2}mv^2 + mgh \\ &= \frac{1}{2}m\left(\frac{p}{m}\right)^2 + mgq \\ &= \frac{p^2}{2m} + mgq \end{aligned}$$

Physics Recap - Hamiltonian Example

$$H = \frac{p^2}{2m} + mgq$$

$$\frac{dp}{dt} = -\frac{\partial H}{\partial q} \qquad \frac{dq}{dt} = \frac{\partial H}{\partial p}$$

$$\frac{dp}{dt} = -mg \qquad \frac{dq}{dt} = \frac{p}{m}$$

$$\frac{d(mv)}{dt} = -mg \qquad v = \frac{p}{m}$$

$$ma = -mg \qquad v = \frac{mv}{m}$$

$$a = -g \qquad v = v$$

Leapfrog Algorithm

Now we have $p(q, t)$ and $q(p, t)$.

Need to approximate with discrete time steps

Naive approach:

$$p(t+\epsilon) = p(t) + \frac{dp}{dt}\epsilon$$
$$q(t+\epsilon) = q(t) + \frac{dq}{dt}\epsilon$$

Issues with convergence. p and q depend on each other

Leapfrog Algorithm

Leapfrog Algorithm:

$$p(t+0.5\epsilon) = p(t) + \frac{dp}{dt}0.5\epsilon$$
$$q(t+\epsilon) = q(t) + \frac{dq}{dt}\epsilon$$
$$p(t+\epsilon) = p(t+0.5\epsilon) + \frac{dp}{dt}0.5\epsilon$$

Better convergence! Only one extra step is needed.

Introduction to Hamiltonian Monte Carlo

Suppose we wish to sample from D dimensions $(q_1, q_2, \ldots, q_D)$

We can cleverly construct D addition dimensions $(p_1, p_2, \ldots, p_D)$

$$
\begin{aligned}
\pi(q, p) &= \exp(-H(q, p)) \\
H(q, p) &= -\log(\pi(q, p)) \\
H(q, p) &= -\log(\pi(p|q)\pi(q)) \\
H(q, p) &= -\log(\pi(p|q)) - log(\pi(q)) \\
H(q, p) &= K(p, q) + U(q)
\end{aligned}
$$

Introduction to Hamiltonian Monte Carlo

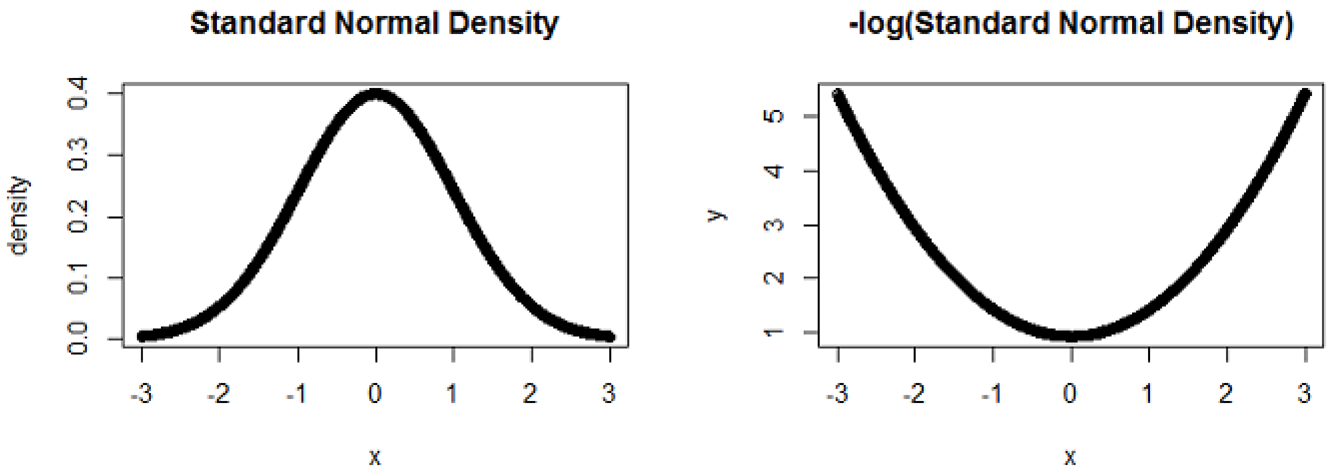
$$
\begin{aligned}
H(q,p) &= K(p,q) + U(q) \\
H(q,p) &= K(p) + U(q) \\
\pi(q,p) &= e^{-K(p)-U(q)} \\
&= e^{-K(p)}e^{-U(q)}
\end{aligned}
$$

To find marginal distribution of q, drop p

Hamiltonian Monte Carlo Algorithm

1. Transform density into potential energy

$$U = -\log(f)$$



Hamiltonian Monte Carlo Algorithm

1 Transform density into potential energy

2 Solve Hamilton's equations. Let $K = \frac{1}{2}mv^2$. Calculate $\frac{\partial U}{\partial q}$

3 Initialize q_o

4 Sample p (e.g. MVN)

5 Calculate proposal p and q. Use leapfrog algorithm

6 Accept-reject according to $\min\left(1, e^{(H_{new}-H_{old})}\right)$

HMC connection to MH

HMC is like an "intelligent" MH algorithm

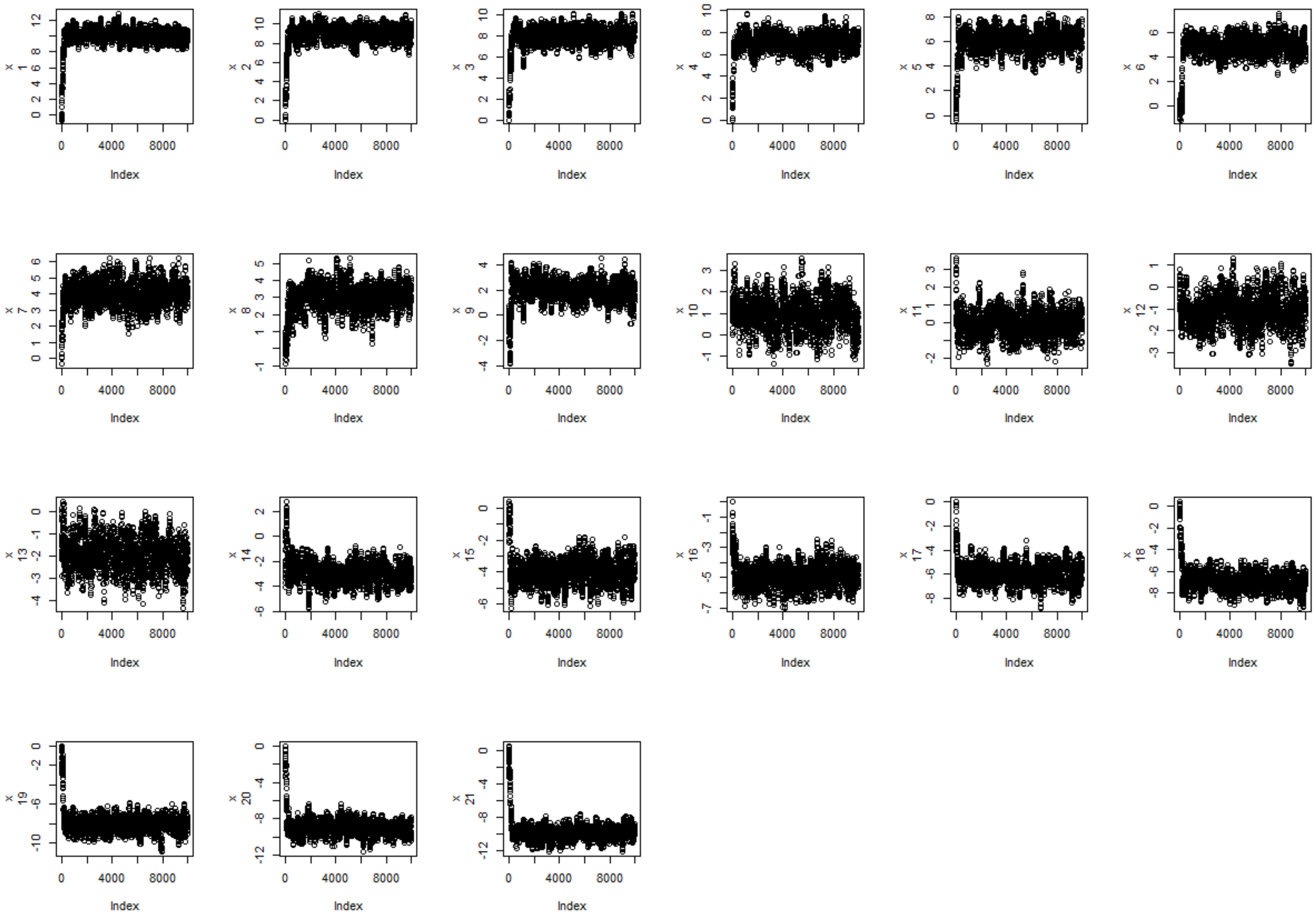
Proposal is symmetric and reversible if we negate p_{new}, which does not affect Hamiltonian

$$\min\left(1, \frac{\pi(p_{new}, q_{new})}{\pi(p_{old}, q_{old})}\right) = \min\left(1, \frac{e^{-H_{new}}}{e^{-H_{old}}}\right)$$
$$= \min\left(1, e^{(H_{new} - H_{old})}\right)$$

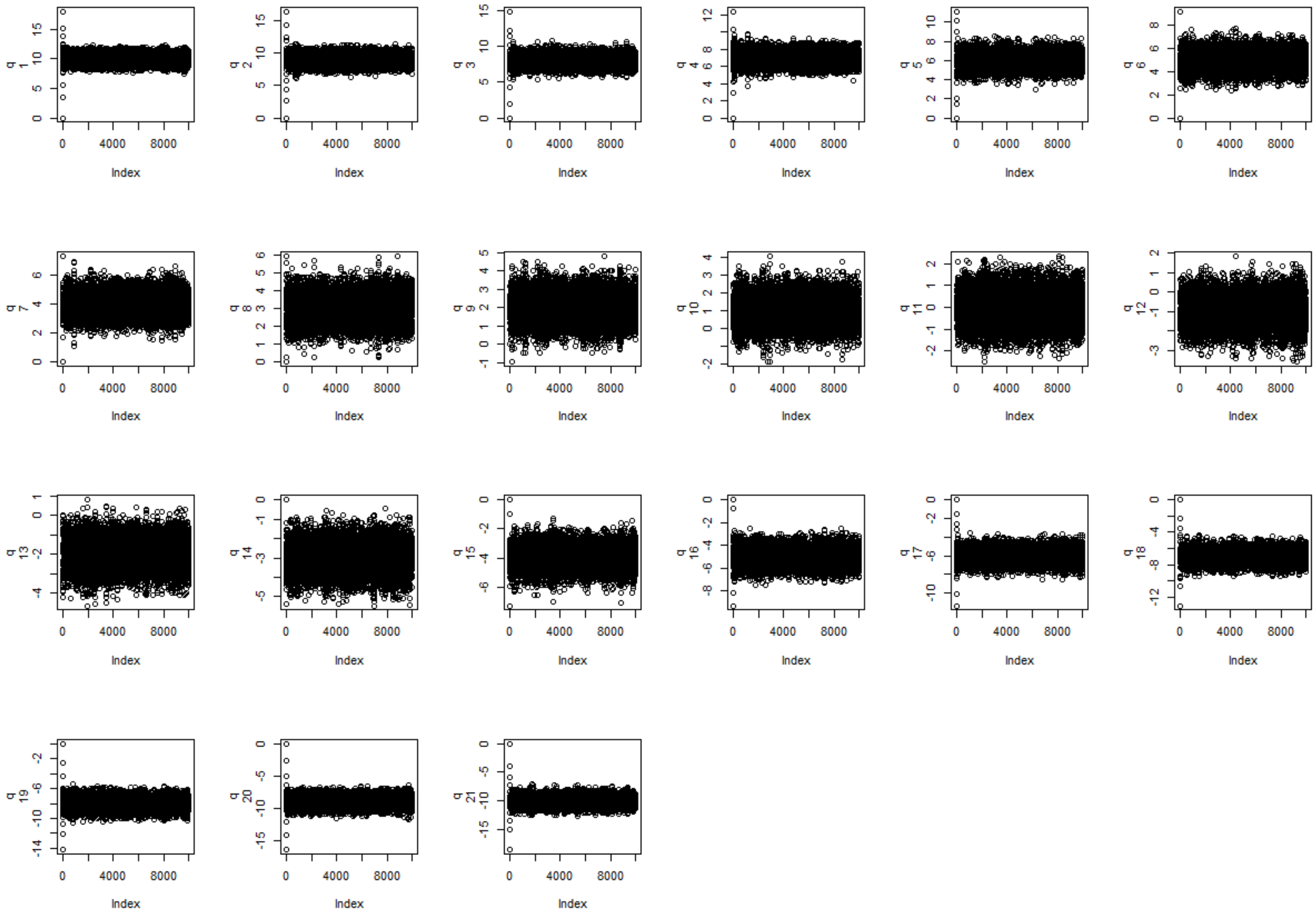
HMC Example - Multivariate Normal

$$X \sim N(\mu, \Sigma)$$
$$\mu = (10, 9, \dots, 0, \dots, -9, -10)^T$$
$$\Sigma = \frac{1}{2} I_{20}$$

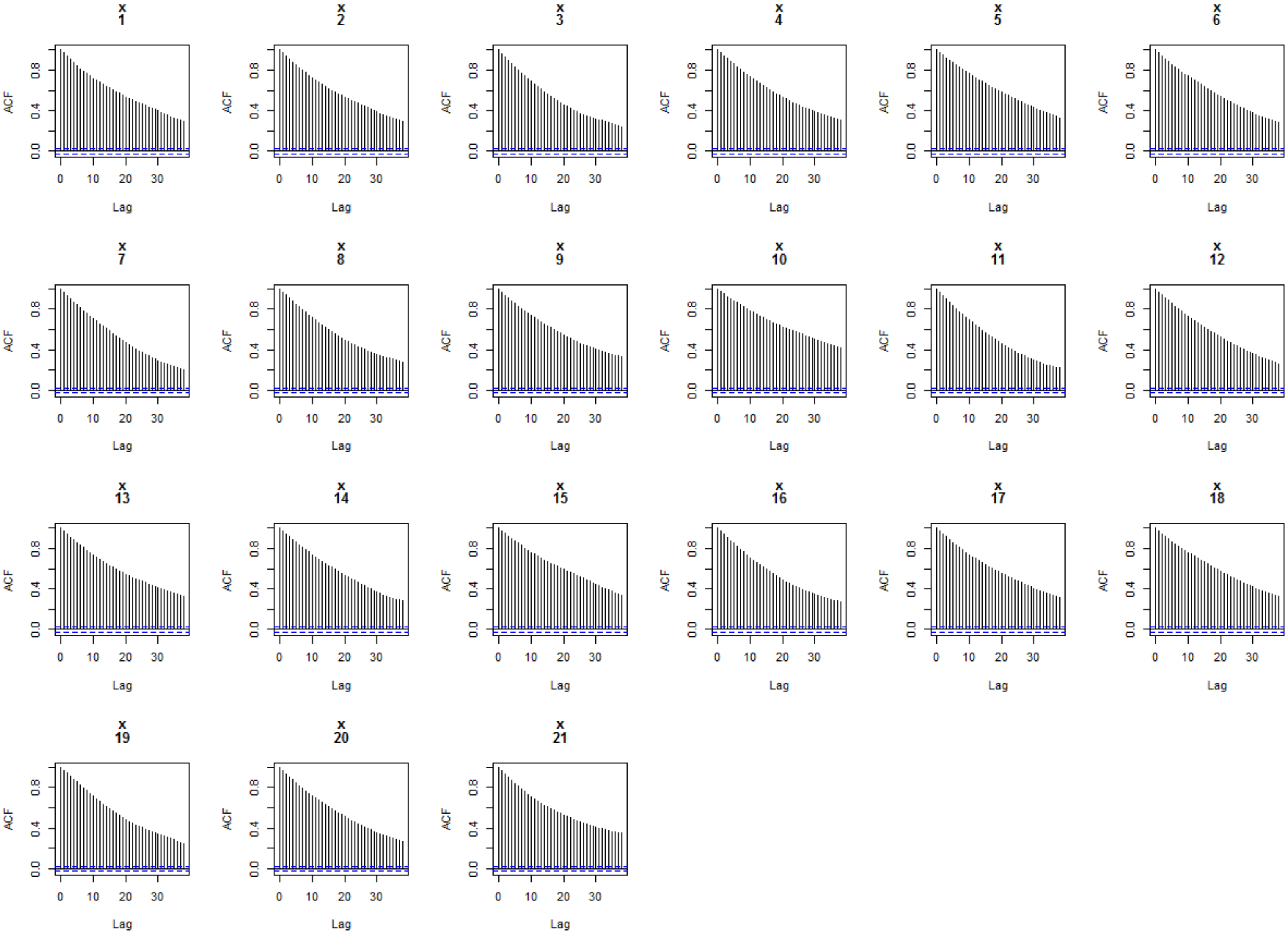
HMC MVN: MH Trace Plots



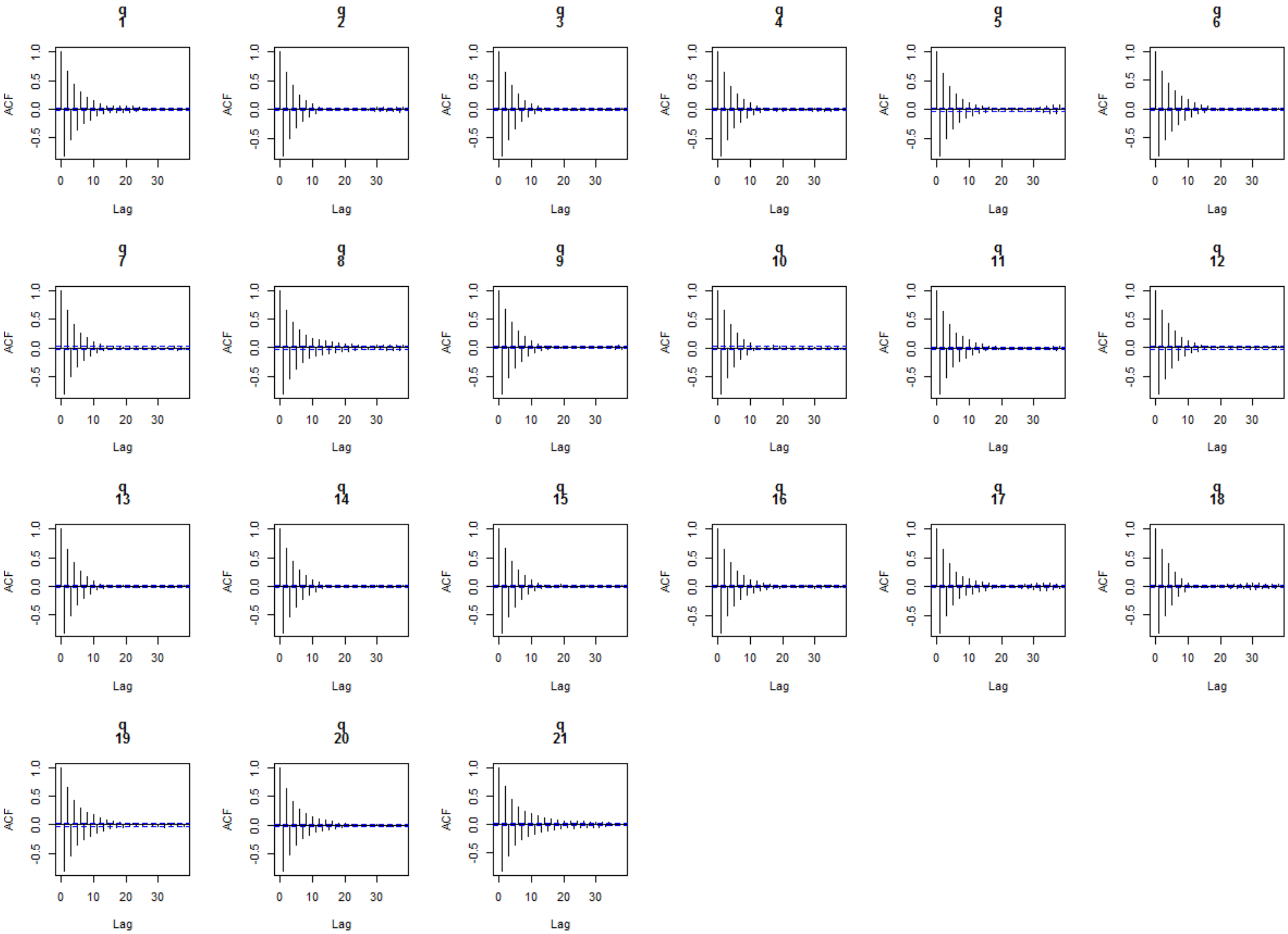
HMC MVN: HMC Trace Plots



HMC MVN: MH Autocorrelation



HMC MVN: HMC Autocorrelation



HMC MVN: Effective Sample Size

- 7500 samples
- 2500 burn in
- Negative autocorrelation

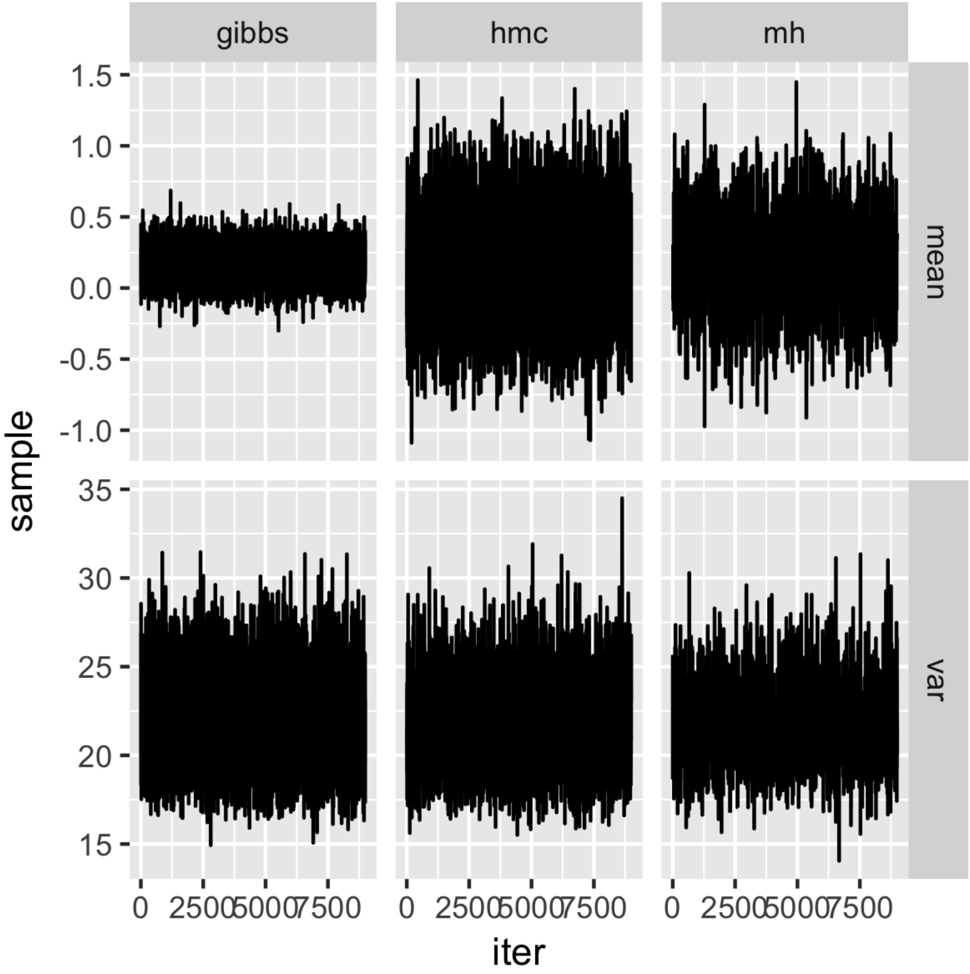
Method	Effective Sample Size[1]
MH	119
HMC	68820

1. ESS averaged over for the dimensions

HMC Example - Normal Inverse Gamma

$$
\begin{aligned}
x_1 \ldots x_n | \mu &\sim N(\mu, \sigma^2) \\
\pi(\mu, \sigma^2) &\propto \frac{1}{\sigma^2} \\
\mu, \sigma^2 | x_1 \ldots x_n &\sim N\text{-}\Gamma^{-1}(\tau = \bar{x}, \lambda = n, \alpha = (n+4)/2, \beta = \frac{1}{2}(\sum x_i^2 - n\bar{x}^2)) \\
\sigma^2 | x_1 \ldots x_n &\sim \Gamma^{-1}(\alpha, \beta) \\
\mu | x_1 \ldots x_n &\sim t_{2\alpha}\big(\tau, \beta/(\alpha\lambda)\big)
\end{aligned}
$$

HMC Example - Normal Inverse Gamma



Method	Effective Sample Size
MH	1380.9
Gibbs	8999.0
HMC	4698.4