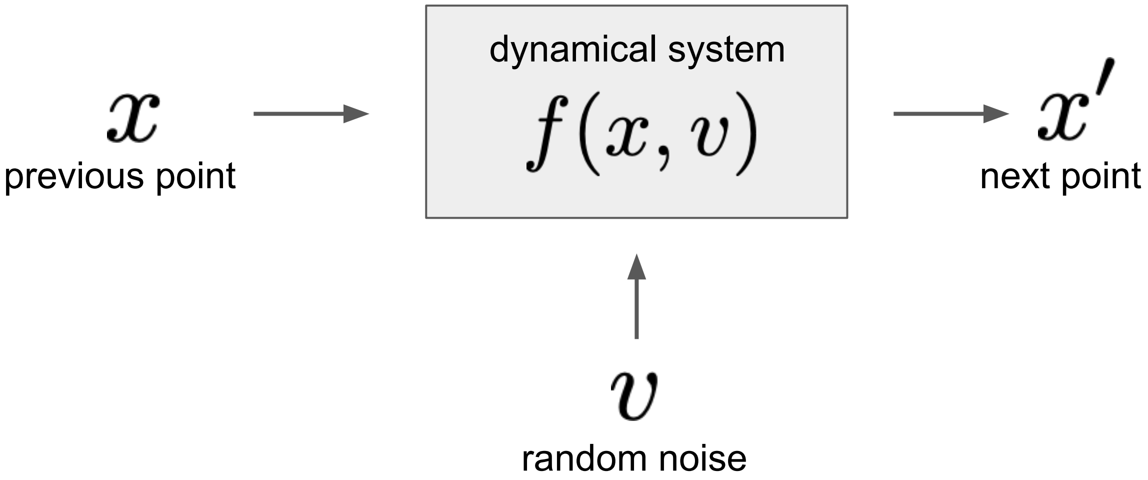
Orbital MCMC

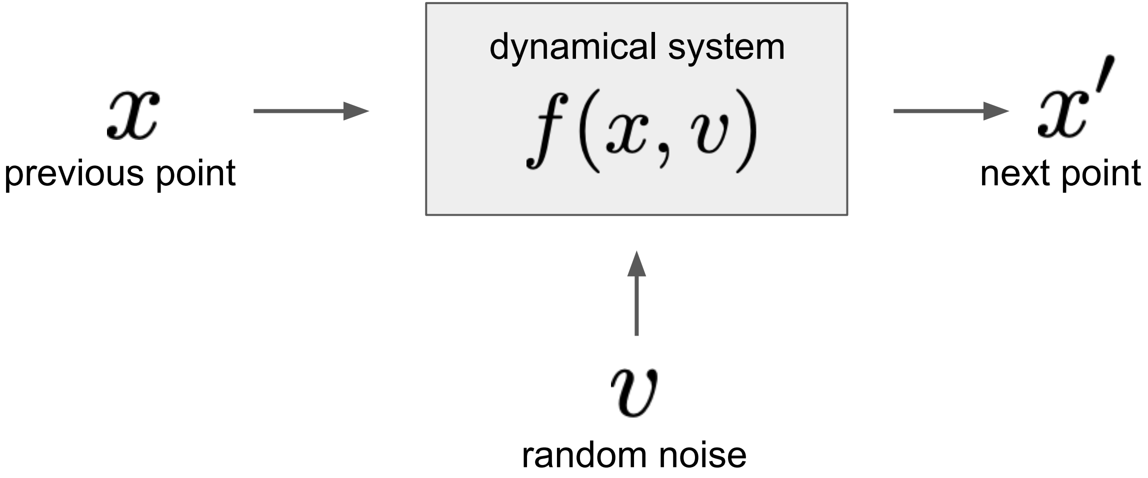
Kirill Neklyudov and Max Welling

Markov Chains Monte Carlo (MCMC)



~~Involutive MCMC*~~

Orbital MCMC



~~Necessary condition: $f(f(x, v)) = [x, v]$~~

*Kirill Neklyudov, Max Welling, Evgenii Egorov, and Dmitry Vetrov, "Involutive mcmc: a unifying framework." ICML, 2020.

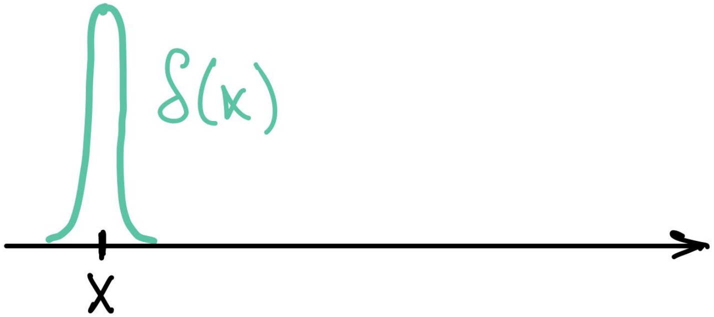
Outline

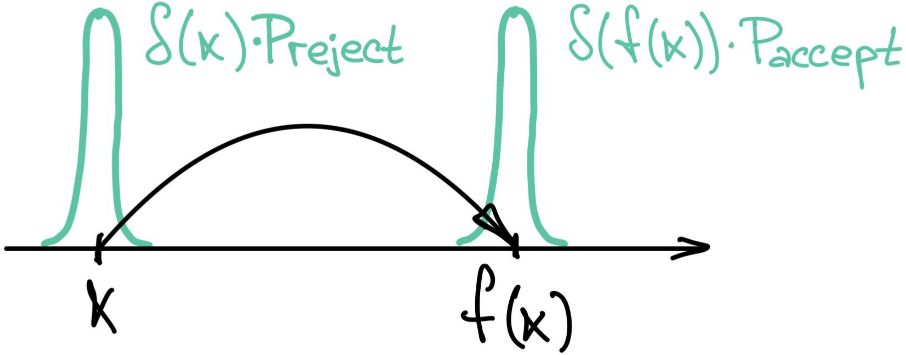
- Background on Involutive MCMC
  (General recipe for many MCMC algorithms)
- Orbital MCMC
  (Generalization of the Metropolis-Hastings-Green test)
- Illustrative Example
  (Application of Orbital MCMC for the Hamiltonian dynamics)

Outline

- Background on Involutive MCMC
  (General recipe for many MCMC algorithms)
- Orbital MCMC
  (Generalization of the Metropolis-Hastings-Green test)
- Illustrative Example
  (Application of Orbital MCMC for the Hamiltonian dynamics)

Involutive MCMC





Stochastic kernel:

$$t(x'|x) = \delta(x' - f(x)) \overbrace{\min\left\{1, \frac{p(f(x))}{p(x)}\left|\frac{\partial f}{\partial x}\right|\right\}}^{P_{\text{accept}}} +$$
$$+\delta(x' - x) \underbrace{\left(1 - \min\left\{1, \frac{p(f(x))}{p(x)}\left|\frac{\partial f}{\partial x}\right|\right\}\right)}_{P_{\text{reject}}}$$

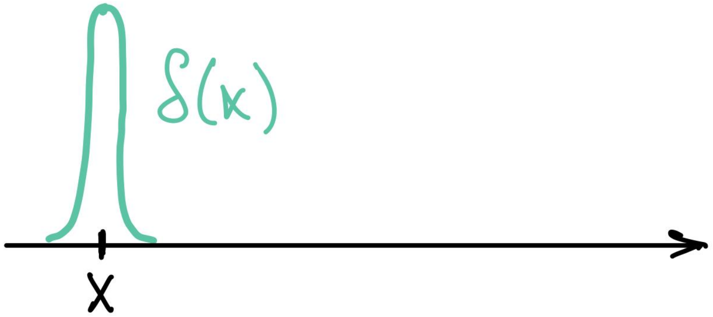
We want to keep the target invariant:

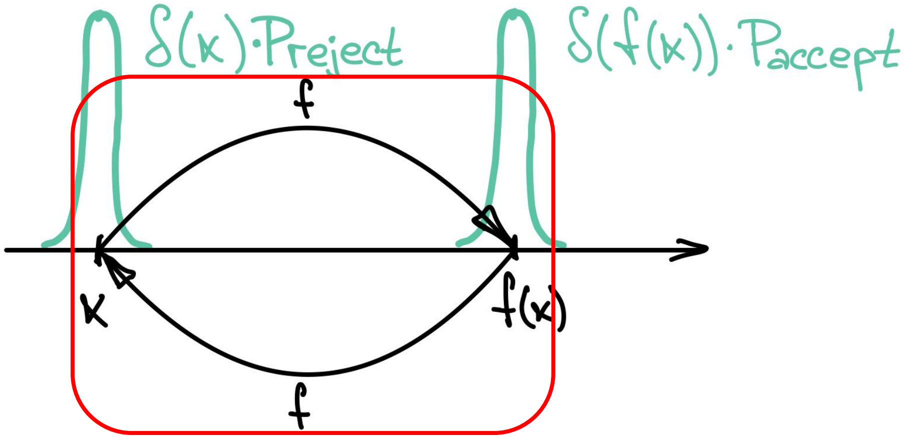
$$\int dx\; t(x'|x)p(x) = p(x')$$

Measure-preserving condition:

$$p(x) = p(f(f(x)))\left|\frac{\partial f(f(x))}{\partial x}\right|.$$

Involutive function





Measure-preserving condition:

$$p(x) = p(f(f(x)))\left|\frac{\partial f(f(x))}{\partial x}\right|.$$

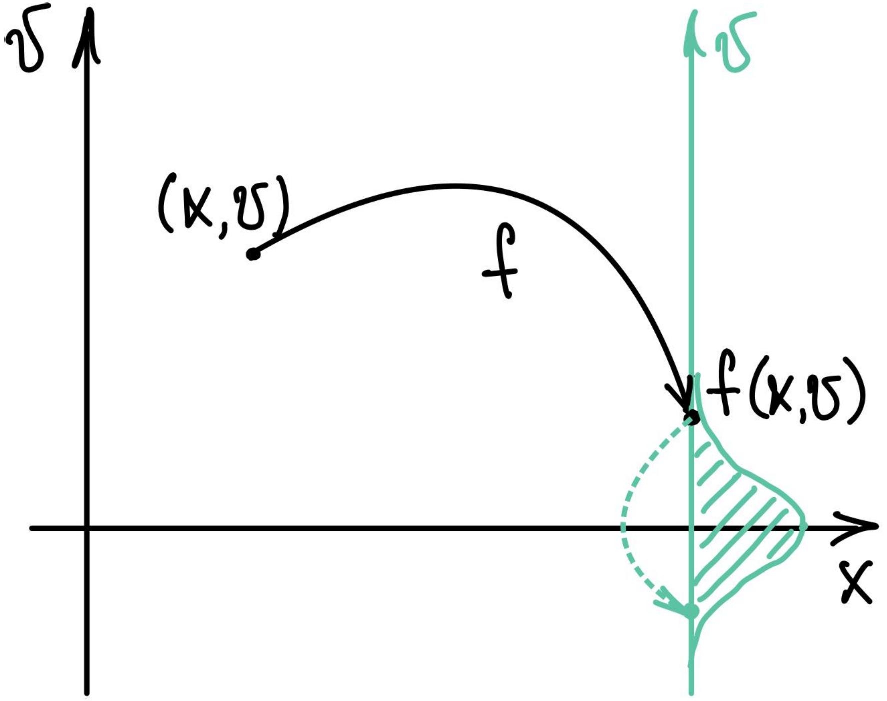
Trivial solution:

$$f(f(x)) = x$$

In other words:

$$f(x) = f^{-1}(x)$$

Auxiliary random variable



Extended target distribution:

$$p(x, v) = p(x)p(v|x)$$

Choose the auxiliary distribution:

- easy to sample
- easy to evaluate the density

For instance:

$$p(x, v) = p(x)\mathcal{N}(v|0, 1)$$

(like in HMC)

Involutive MCMC

Algorithm

0. initial point is x

1. sample $v \sim p(v|x)$

2. propose $[x', v'] = f(x, v)$

3. evaluate $P = \min\left\{1, \frac{p(f(x,v))}{p(x,v)} \left|\frac{\partial f(x,v)}{[x,v]}\right|\right\}$

4. accept $\begin{cases} x', & \text{with prob. } P \\ x, & \text{with prob. } (1-P) \end{cases}$

Name & Citation
Metropolis-Hastings (Hastings, 1970)
Mixture Proposal (Habib & Barber, 2018)
Multiple-Try Metropolis (Liu et al., 2000)
Sample-Adaptive MCMC (Zhu, 2019)
Reversible-Jump MCMC (Green, 1995)
Hybrid Monte Carlo (Duane et al., 1987)
RMHMC (Girolami & Calderhead, 2011)
NeuTra (Hoffman et al., 2019)
A-NICE-MC (Song et al., 2017)
L2HMC (Levy et al., 2017)
Persistent HMC (Horowitz, 1991)
Gibbs (Geman & Geman, 1984)
Look Ahead (Sohl-Dickstein et al., 2014)
NRJ (Gagnon & Doucet, 2019)
Lifted MH (Turitsyn et al., 2011)

Outline

- Background on Involutive MCMC
  (General recipe for many MCMC algorithms)
- Orbital MCMC
  (Generalization of the Metropolis-Hastings-Green test)
- Illustrative Example
  (Application of Orbital MCMC for the Hamiltonian dynamics)

Involutive MCMC –> Orbital MCMC

Acceptance test: Dynamics:

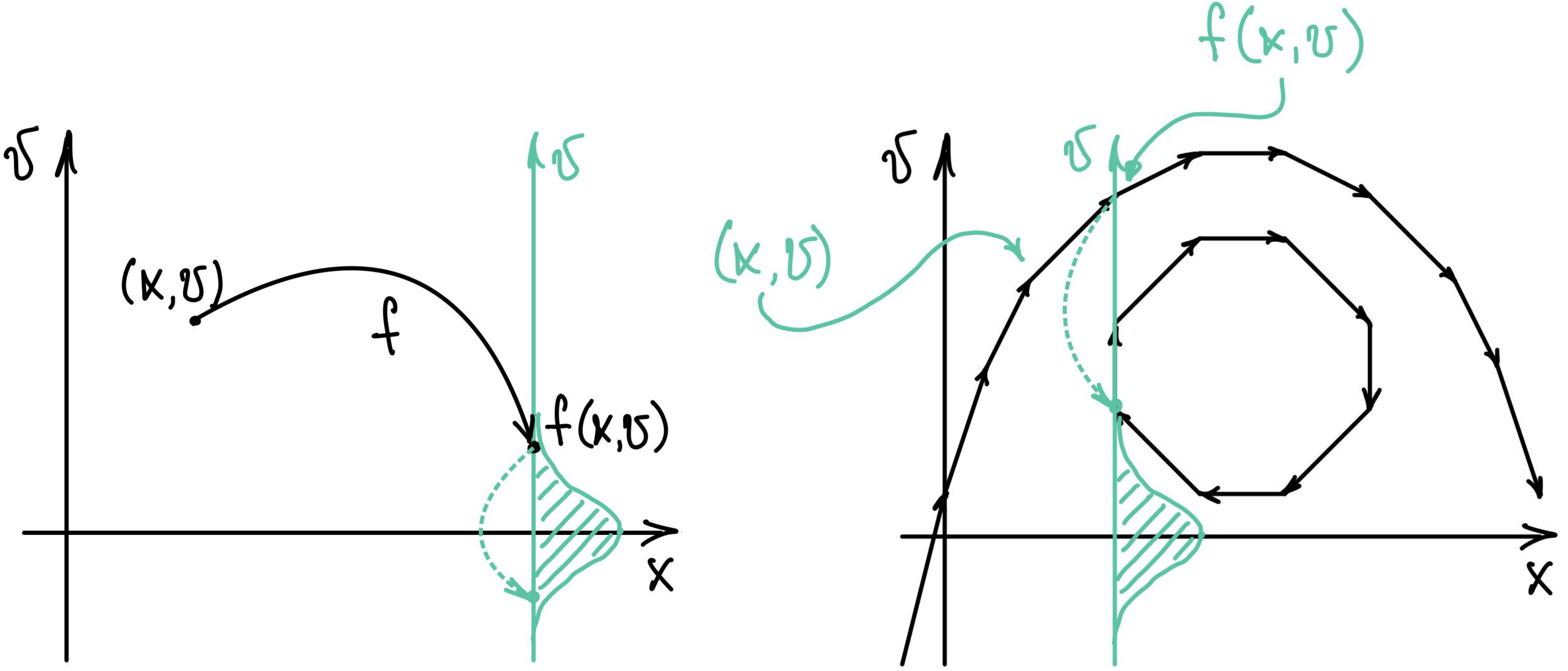
$$P_{\text{accept}} = \min\left\{1, \frac{p(f(x))}{p(x)}\left|\frac{\partial f}{\partial x}\right|\right\} \longrightarrow f(f(x)) = x$$

Involutive MCMC

$$P_{\text{accept}} = \inf_{k\in\mathbb{Z}}\left\{\frac{p(f^k(x))}{p(x)}\left|\frac{\partial f^k}{\partial x}\right|\right\} \longleftarrow \text{any } f$$

Orbital MCMC

Involutive MCMC vs Orbital MCMC



~~Involutive MCMC~~

Orbital MCMC

Algorithm

0. initial point is x

1. sample $v \sim p(v|x)$

2. propose $[x', v'] = f(x, v)$

3. evaluate $P = \inf_{k \in \mathbb{Z}} \left\{ \frac{p(f^k(x,v))}{p(x,v)} \left| \frac{\partial f^k(x,v)}{\partial(x,v)} \right| \right\}$

4. accept $\begin{cases} x', & \text{with prob. } P \\ x, & \text{with prob. } (1-P) \end{cases}$

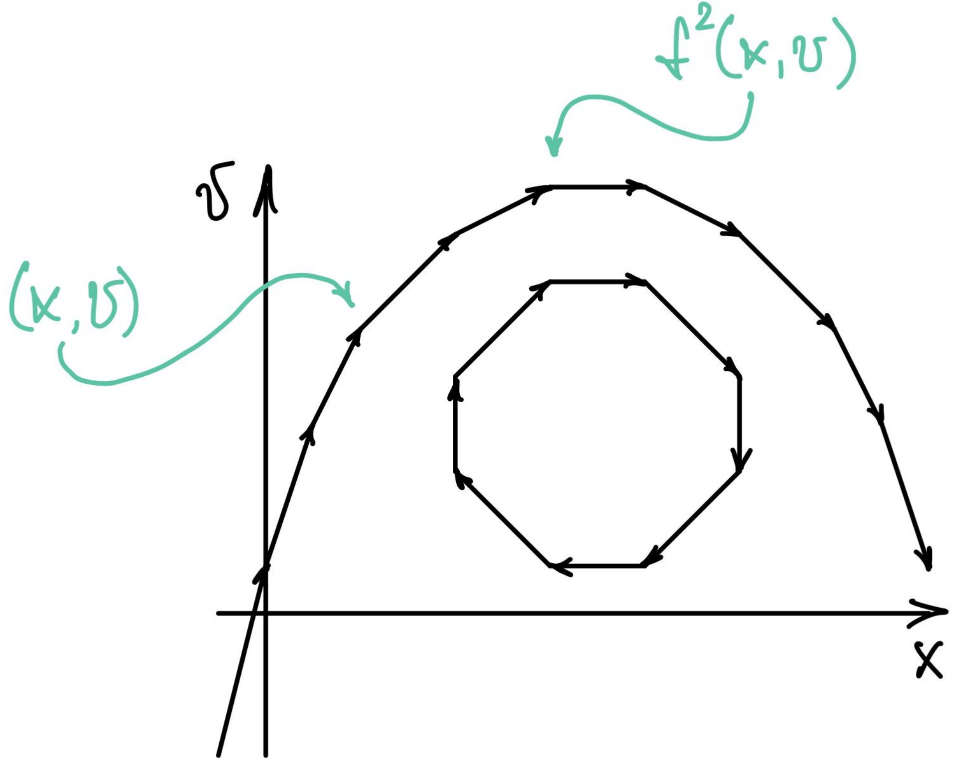
For Involutive MCMC

$$t(x'|x)p(x) = t(x|x')p(x')$$

For Orbital MCMC

$$t(x'|x)p(x) \neq t(x|x')p(x')$$

Problems with the test



In general:

pt

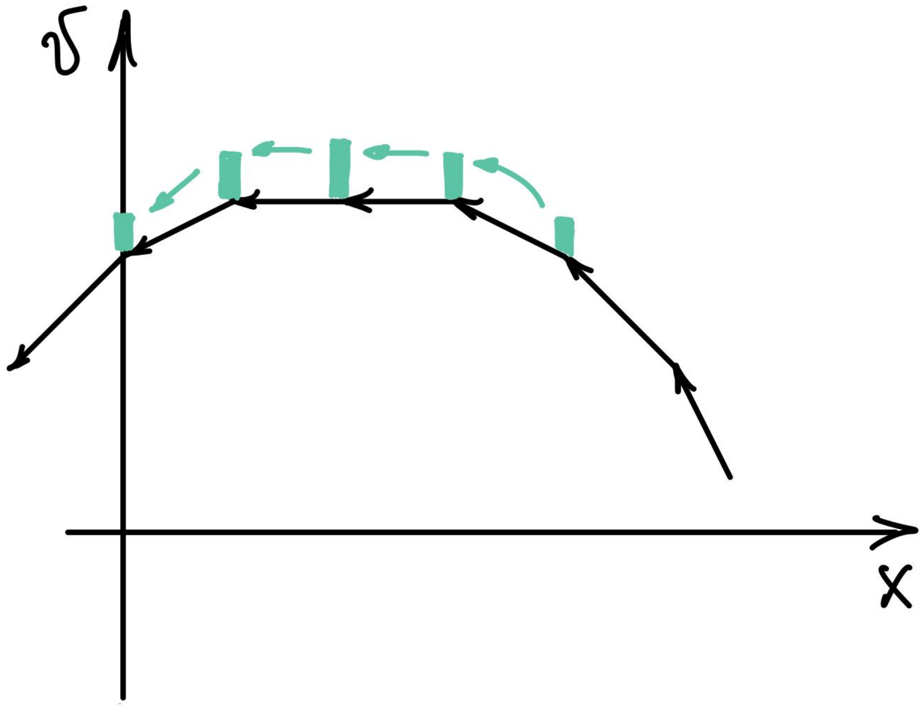
$$P_{\text{accept}} = \inf_{k \in \mathbb{Z}} \left\{ \frac{p(f^k(x))}{p(x)} \left| \frac{\partial f^k}{\partial x} \right| \right\}$$

For periodic orbits:

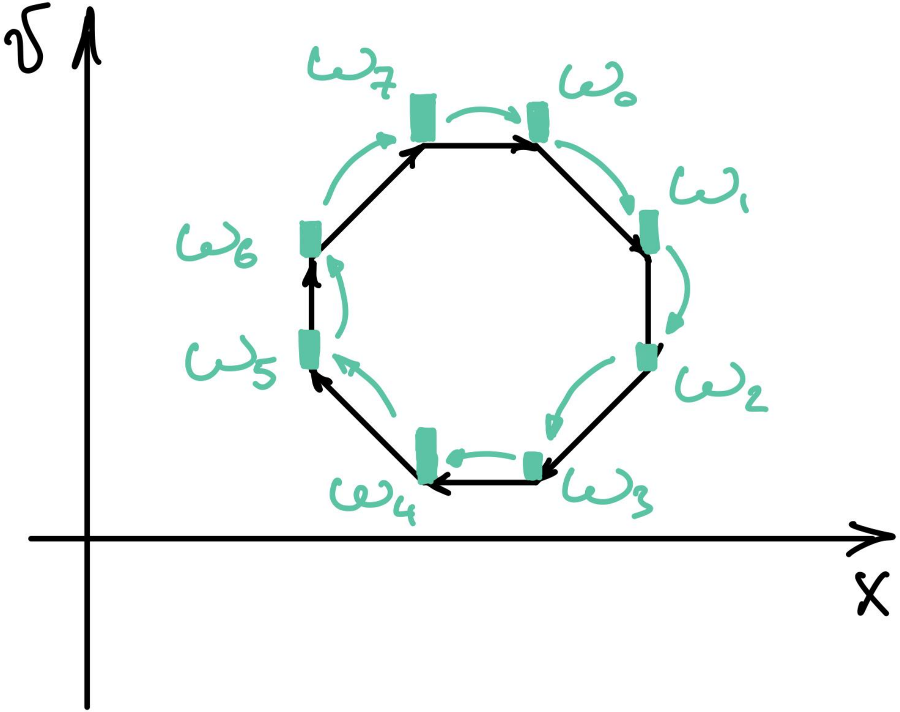
$$P_{\text{accept}} = \min_{k \in [0 \ldots T-1]} \left\{ \frac{p(f^k(x))}{p(x)} \left| \frac{\partial f^k}{\partial x} \right| \right\}$$

Still T-1 evaluations per 1 sample

Iterating on the same orbit



Accepting the whole orbit



Application of a kernel:

$$[Kp](x') = \int dx\ t(x'|x)p(x)$$

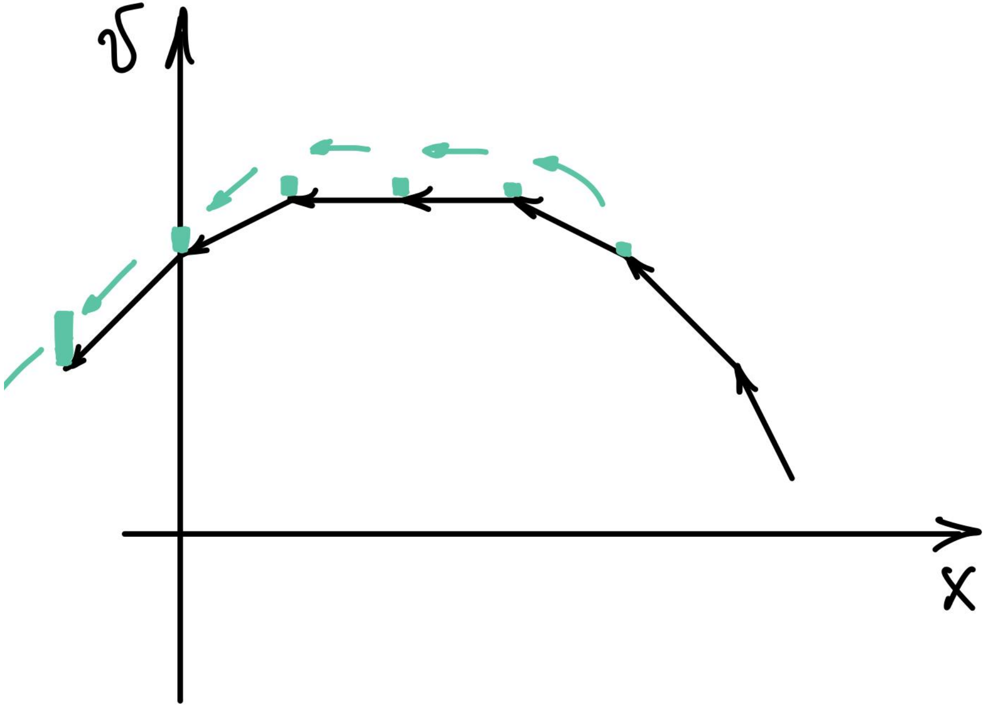
In the limit for a single orbit:

$$\lim_{t\to\infty} \frac{1}{t} \sum_{i=0}^{t-1} K^i p_0 = \sum_{i=0}^{T-1} \omega_i \delta(x - f^i(x_0))$$

where:

$$\omega_i = \frac{p(f^i(x_0))\left|\frac{\partial f^i}{\partial x}\right|_{x=x_0}}{\sum_{j=0}^{T-1} p(f^j(x_0))\left|\frac{\partial f^j}{\partial x}\right|_{x=x_0}}$$

Infinite orbits



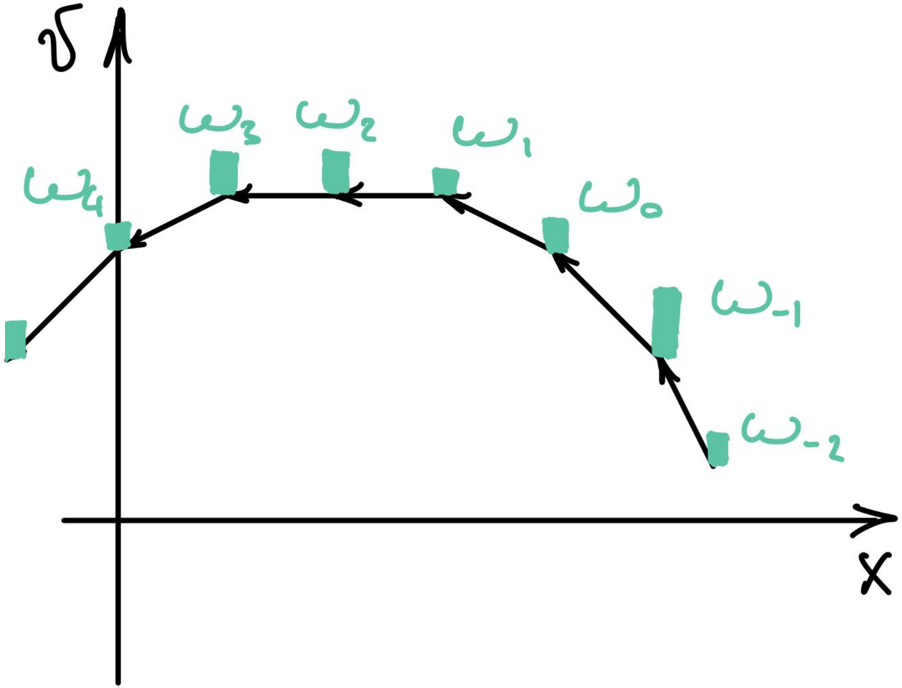
Application of a kernel:

$$[K^t p_0](x) = \sum_{i=0}^{t} \omega_i^t \delta(x - f^i(x_0))$$

In the limit for a single orbit:

$$\lim_{t\to\infty} \frac{1}{t} \sum_{t'=0}^{t} \omega_i^{t'} = 0$$

Infinite orbits



The time-average on a single orbit:

$$\lim_{t\to\infty} \frac{1}{t} \sum_{i=0}^{t-1} K^i p_0 = \sum_{i=-\infty}^{+\infty} \omega_i \delta(x - f^i(x_0))$$

Formula for weights:

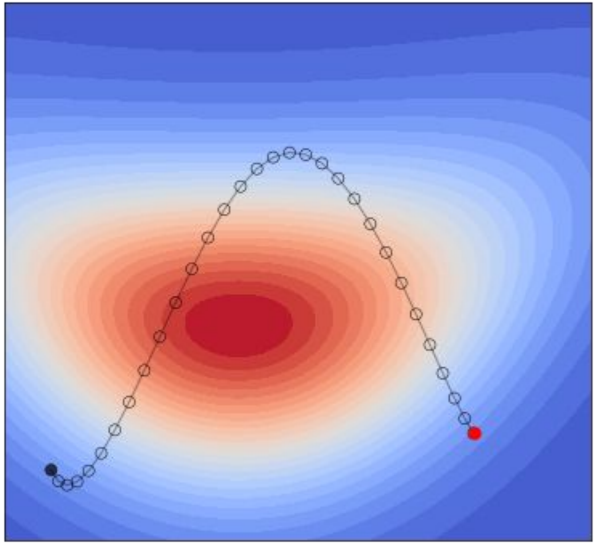
$$\omega_i = \frac{p(f^i(x_0)) \left| \frac{\partial f^i}{\partial x} \right|_{x=x_0}}{\sum_{j=-\infty}^{+\infty} p(f^j(x_0)) \left| \frac{\partial f^j}{\partial x} \right|_{x=x_0}}$$

Outline

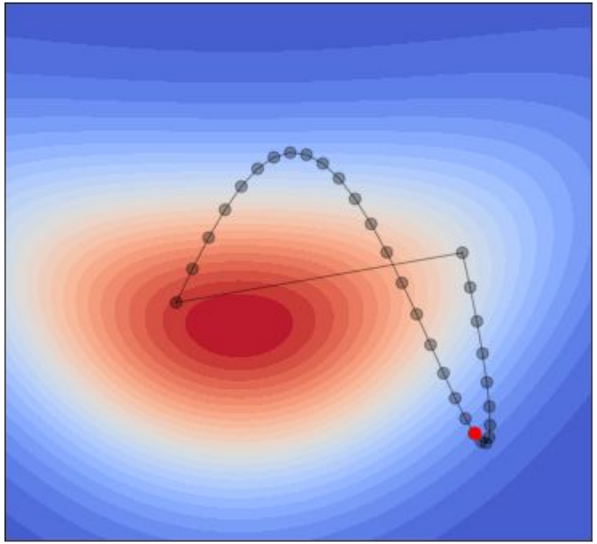
- Background on Involutive MCMC
  (General recipe for many MCMC algorithms)
- Orbital MCMC
  (Generalization of the Metropolis-Hastings-Green test)
- Illustrative Example
  (Application of Orbital MCMC for the Hamiltonian dynamics)

Example (HMC)

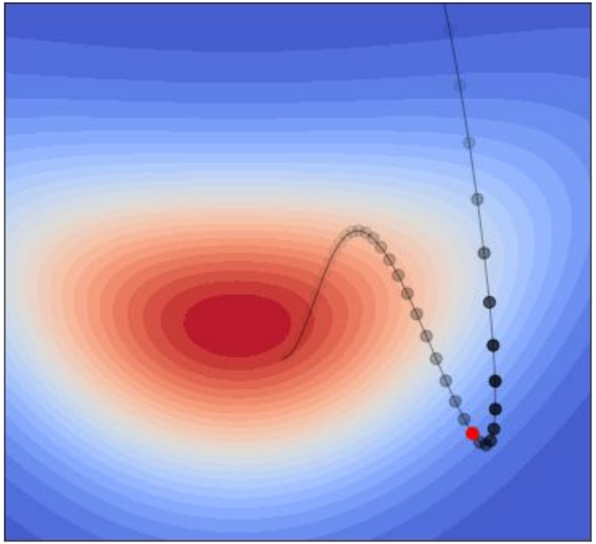
HMC (standard)



Orbital MC (periodic orbit)



Orbital MC (infinite orbit)

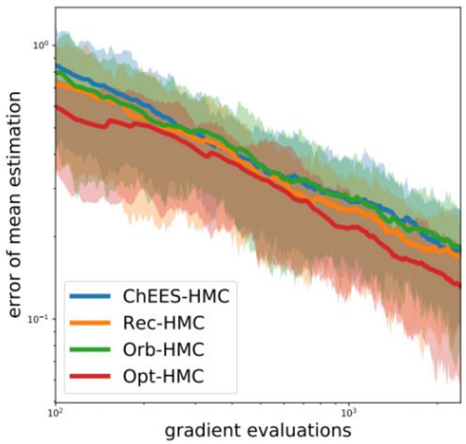


$$t(x'|x) = \delta(x' - f(x)) \overbrace{\min\left\{1, \frac{p(f(x))}{p(x)}\left|\frac{\partial f}{\partial x}\right|\right\}}^{P_{\text{accept}}} + \delta(x' - x)\underbrace{\left(1 - \min\left\{1, \frac{p(f(x))}{p(x)}\left|\frac{\partial f}{\partial x}\right|\right\}\right)}_{P_{\text{reject}}}$$

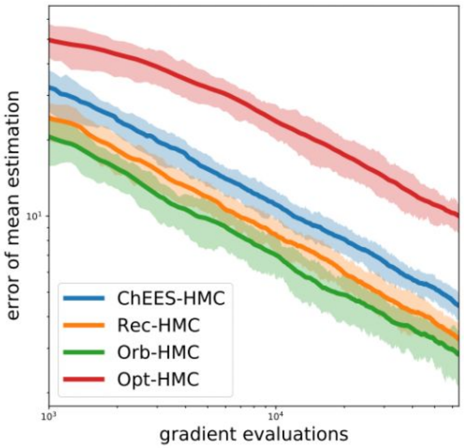
$$\omega_i = \frac{p(f^i(x_0))\left|\frac{\partial f^i}{\partial x}\right|_{x=x_0}}{\sum_{j=0}^{T-1} p(f^j(x_0))\left|\frac{\partial f^j}{\partial x}\right|_{x=x_0}}$$

$$\omega_i = \frac{p(f^i(x_0))\left|\frac{\partial f^i}{\partial x}\right|_{x=x_0}}{\sum_{j=-\infty}^{+\infty} p(f^j(x_0))\left|\frac{\partial f^j}{\partial x}\right|_{x=x_0}}$$

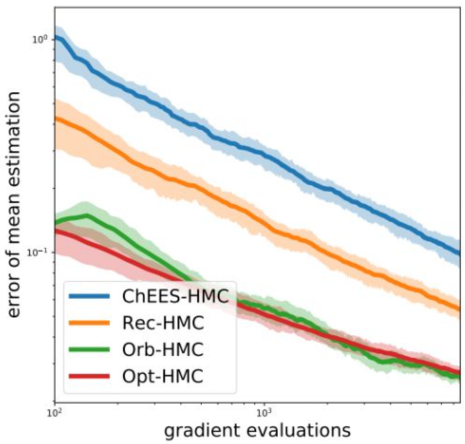
Empirical Results for Hamiltonian Dynamics



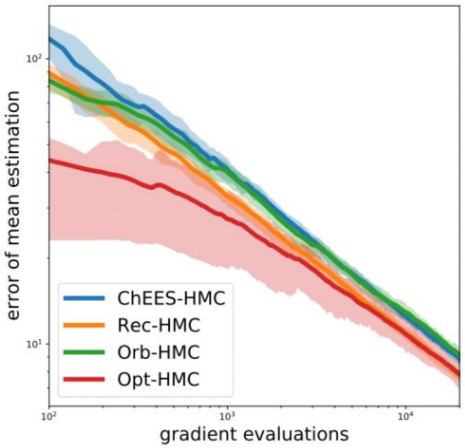
Synthetic (2-D)



Synthetic (50-D)



Bayesian logreg (25-D)



Item-response (501-D)

***ChEES-HMC** = Hoffman, M., Radul, A., Sountsov, P.
"An adaptive-MCMC scheme for setting trajectory lengths in HMC." AISTATS, 2021.

****Rec-HMC** = Nishimura, A., Dunson, D.
"Recycling intermediate steps to improve Hamiltonian Monte Carlo". Bayesian Analysis, 2020.

Orbital MCMC (summary)

- Generalizes Involutive MCMC, which describes many MCMC methods
- Generalizes Metropolis-Hastings-Green test
- Allows for usage of any dynamical system for sampling

Orbital MCMC (future directions)

- Optimal number of steps for a single orbit
- Examples with more complicated (chaotic/learnable) dynamical systems