Evaluation:

 $(r_0, r_0) \leftarrow (M^{1/2} q^{(l)}, M^{-1/2} p^{(l)})$ $(hh)r_i + h^{-1} \sin(hh) \dot{r}_i - 0.5h^2 \mathbb{E}(hh) F(w(hh)r_i)$ $\cos(hh)r_i$ $-0.5h(\mathbb{E}_{0}(h\negthinspace h)F(w(h\negthinspace h) r_{j})+\mathbb{E}_{1}(h\negthinspace h)F(w(h\negthinspace h) r_{j+1}))$ $(q^*, p^*) \leftarrow (M^{-1/2}r_L, M^{1/2}r_L)$ x is supported by ARO Award # W911NF-12-1-0241,

1000141210066, Alfred P. Sloan Fellowship, and the

ck Center for Visual Computing and Communication.

1/2 $q^{(i)}$, $M^{-1/2} p^{(i)}$

1/2 $q^{(i)}$, $M^{-1/2} p^{(i)}$

1/2 $($ $+1$
 1 2 $\begin{aligned} &\frac{\mathsf{MNC}\left(\mathsf{proposed}\right)}{q^{(i)}, M^{-1/2}p^{(i)}}\ =&1\ &\left.\left.\right)r_{j}+\mathsf{h}^{-1}\sin(h\mathsf{h})\dot{r}_{j}-0.5h^{2}\mathsf{E}\left(h\mathsf{h}\right)F\left(\mathsf{w}(h\mathsf{h})\right)\dot{r}\right.\\ &\left.\left.\mathsf{h}\mathsf{h}\right)r_{j}+\cos(h\mathsf{h})\dot{r}_{j}\ =&\mathsf{E}_{\mathsf{o}}(h\mathsf{h})F\left(\mathsf{w}(h\mathsf{h})r_{j}\right)+\mathsf{E}_{\mathsf{1}}$ $\begin{split} &\frac{\mathbf{proposed)}}{\mathbf{p}^{(l)}}\Big)^{1}\mathrm{sin}(h\mathsf{h})\dot{\mathsf{r}}_{j}-0.5h^{2}\mathsf{E}\left(h\mathsf{h}\right)\mathbf{F}\left(\mathsf{w}(h\mathsf{h})\mathbf{r}\right)\ \mathrm{cos}(h\mathsf{h})\dot{\mathsf{r}}_{j}\ \mathbf{F}\left(\mathsf{w}(h\mathsf{h})\mathbf{r}_{j}\right)+&\mathsf{E}_{1}(h\mathsf{h})\mathbf{F}\left(\mathsf{w}(h\mathsf{h})\mathbf{r}_{j+1}\right)\ \frac{1}{2}\dot{\mathsf{r}}_{l}}$ This work is supported by ARO Award # W911NF-12-1-0241,

ONR # N000141210066, Alfred P. Sloan Fellowship, and the

Max Planck Center for Visual Computing and Communication.
 ponential HMC (proposed)
 \vec{r}_0) \leftarrow $(M^{$ This work is supported by ARO Award # W911NF-12-1-0241,

ONR # N000141210066, Alfred P. Sloan Fellowship, and the

Max Planck Center for Visual Computing and Communication.
 Exponential HMC (proposed)
 $(r_o, r_o) \leftarrow (M^{1/2}$ This work is supported by ARO Award # W911NF-12-1-0241,

ONR # N000141210066, Alfred P. Sloan Fellowship, and the

Max Planck Center for Visual Computing and Communication.
 Exponential HMC (proposed)
 $(r_o, r_o) \leftarrow (M^{1/2}$ This work is supported by ARO Award # W911NF-12-1-0241,

ONR # N000141210066, Alfred P. Sloan Fellowship, and the

Max Planck Center for Visual Computing and Communication.

 ponential HMC (proposed)
 \vec{r}_0) $\leftarrow (M^{1$ upported by ARO Award # W911NF-12-1-0241,
41210066, Alfred P. Sloan Fellowship, and the
enter for Visual Computing and Communication
 $\overline{MC \textbf{(proposed)}}$
 $\overline{MN} = \sum_{i=1}^{N} p^{(i)}$ ONR # N000141210066, Alfred P. Sloan Fellowship, and the

Max Planck Center for Visual Computing and Communication.
 j i j $\leftarrow (M^{1/2}q^{(i)}, M^{-1/2}p^{(i)})$
 $\leftarrow 0, 1, ..., L-1$
 $\leftarrow 1 = \cos(hh)r_j + h^{-1}\sin(hh)r_j - 0.5h^2(E(hh))F(W(hh)r_j)$
 \leftarrow *j* iviax Planck Center for Visual Computing and Communication.
 j \vec{r}_0) \leftarrow $\left(M^{1/2}q^{(i)}, M^{-1/2}p^{(i)}\right)$
 \vec{r}_0) \leftarrow $\left(M^{1/2}q^{(i)}, M^{-1/2}p^{(i)}\right)$
 \leftarrow 0,1,...,L-1
 \leftarrow j_{+1} = $\cos(hh)r_j + h^{-1}\sin(hh)r_j - 0.$ *j*
 j j + C₁(hh) $F(W(hh)r_j)$
 j) + C₁(hh) $F(W(hh)r_{j+1})$
 j lue) are pre-computable 1C (proposed)
 L $M^{-1/2}p^{(i)}$
 $+h^{-1}\sin(hh) \dot{r}_j - 0.5h^2 \mathbb{E}(hh)F(w(hh) r_j)$
 $\dot{r}_j + \cos(hh) \dot{r}_j$
 $(hh)F(w(hh) r_j) + \mathbb{E}_1(hh)F(w(hh) r_{j+1})$
 \vdots , $M^{1/2} \dot{r}_i$
 \vdots functions in blue) are pre-computable This work is supported by ARO Award # W911NF-12-1-0241,

ONR # N000141210066, Alfred P. Sloan Fellowship, and the

Max Planck Center for Visual Computing and Communication.
 j \vec{r}_0) \leftarrow $(M^{1/2}q^{(i)}, M^{-1/2}p^{(i)}\right)$ *h* h h \mathbf{r}_j + cos(*h*) \mathbf{h} = 0.5*h* \mathbf{f} (*w*) + \mathbf{f} (*w*) + \mathbf{h} (*h*) \mathbf{r}_j + $\mathbf{$ supported by ARO Award # W911NF-12-1-0241,

1141210066, Alfred P. Sloan Fellowship, and the

Center for Visual Computing and Communication.
 HMC (proposed)
 $q^{(i)}$, $M^{-1/2}p^{(i)}$
 \rightarrow
 \rightarrow
 \rightarrow
 \rightarrow
 \rightarrow
 \rightarrow
 $\$ *h* k is supported by ARO Award # W911NF-12-1-0241,

h h h h r i + 1-1 sin (h h) r i + 0.5 km Fellowship, and the

did HMC (proposed)
 $h^{1/2}q^{(i)}$, $M^{-1/2}p^{(i)}$
 $h^{(1)}r_j + h^{-1}\sin(hh)r_j - 0.5h^2E(hh)F(w(hh)r_j)$

in (h h) r i + $^{-1}$ sin (hh) \dot{r} , $-$ 0.5h²($E(hh)$ F $(h$ ← $(M^{1/2}q^{(i)}, M^{-1/2}p^{(i)})$

0,1,...,L-1

= cos(hh)r_j + h⁻¹sin(hh)r_j - 0.5h²E (hh)F (w(hh)r_j)

= -hsin(hh)r_j + cos(hh)r_j

-0.5h(E_o(hh)F (w(hh)r_j)+E₁(hh)F (w(hh)r_{j+1})) This work is supported by ARO Award # W911NF-12-1-0241,

ONR # N000141210066, Alfred P. Sloan Fellowship, and the

Max Planck Center for Visual Computing and Communication.
 onential HMC (proposed)
 $\begin{array}{l}\n\downarrow_0 \rightarrow (\mathbf{$ is work is supported by ARO Award # W911NF-12-1-0241,

NR # N000141210066, Alfred P. Sloan Fellowship, and the

ax Planck Center for Visual Computing and Communication.
 nential HMC (proposed)
 $\left(\frac{M^{1/2}q^{(i)}, M^{-1/2}p^{$ is work is supported by ARO Award # W911NF-12-1-0241,
NR # N000141210066, Alfred P. Sloan Fellowship, and the
ax Planck Center for Visual Computing and Communication.
nential HMC (proposed)
 $\left(\frac{M^{1/2}q^{(i)}, M^{-1/2}p^{(i)}}{$ work is supported by ARO Award # W911NF-12-1-0241,

R# N000141210066, Alfred P. Sloan Fellowship, and the

R Planck Center for Visual Computing and Communication.

That HMC (proposed)
 $-(M^{1/2}q^{(i)}, M^{-1/2}p^{(i)})$
 $(1, ..., L-1)$ $\leftarrow (M^{-1/2}r_{\!\scriptscriptstyle 1}, M^{1/2}\dot{r}_{\!\scriptscriptstyle 1})$ This work is supported by ARO Award # W911NF-12-1-0241,
ONR # N00141210066, Alfred P. Sloan Fellowship, and the
Max Planck Center for Visual Computing and Communication.

 ponential HMC (proposed)
 \vec{r}_0) \leftarrow $(M^{1$ \dot{r}_L **Finally and the UP and Set of the UP and** This work is supported by ARO Award # W911NF-12-1-0241,

ONR # N000141210066, Alfred P. Sloan Fellowship, and the

Max Planck Center for Visual Computing and Communication.
 ponential HMC (proposed)
 \vec{r}_0) $\leftarrow (M^{1/$ This work is supported by ARO Award # W911NF-12-1

ONR # N000141210066, Alfred P. Sloan Fellowship, an

Max Planck Center for Visual Computing and Commun

ponential HMC (proposed)
 \vec{r}_0) $\leftarrow (M^{1/2}q^{(i)}, M^{-1/2}p^{(i)})$
 Fractional Burnock and the Walletter of Nisaan Fellowship, and the for Visual Computing and Communication.
 P r *F* $\left(\frac{p}{r}\right)^{2} p^{(i)}$ and the first final Fellowship, and the formulation.
 P r $\left(\frac{p}{r}\right)^{2}$ $\left(\$ This work is supported by ARO Award # W9111

ONR # N000141210066, Alfred P. Sloan Fellow.

Max Planck Center for Visual Computing and C
 Exponential HMC (proposed)
 $\mathbf{r}_0, \dot{\mathbf{r}}_0 \} \leftarrow \left(M^{1/2} q^{(i)}, M^{-1/2} p^{(i)} \right)$
 x
 x
 w W911NF-12-1-0241,
 x Ellowship, and the
 x and Communication.
 x
 x (*hh*) $F(w(hh)r_j)$
 x (*hh*) $F(w(hh)r_j)$ **EXERCUTE:**
 Supported by ARO Award # W911NF-12-1-0241,

0141210066, Alfred P. Sloan Fellowship, and the

center for Visual Computing and Communication.
 HMC (proposed)
 $P_q^{(i)}$, $M^{-1/2} p^{(i)}$
 $P_q^{(i)}$, $M^{-1/2} p^{(i)}$ Work is supported by ARO Award # W911NF-12-1-0241,

Work is supported by ARO Award # W911NF-12-1-0241,

N000141210066, Alfred P. Sloan Fellowship, and the

Planck Center for Visual Computing and Communication.
 (M^{1/2} For the Barry of the NARO Award # W911NF-12-1-0241,

210066, Alfred P. Sloan Fellowship, and the

er for Visual Computing and Communication.
 C (proposed)
 $I^{-1/2}p^{(i)}$
 $h^{-1}\sin(hh)\dot{r}_j - 0.5h^2E(hh)F(w(hh)r_j)$
 $r_j + \cos(hh)\dot{r$

Exponential Integration for Hamiltonian Monte Carlo Wei-Lun Chao¹, Justin Solomon^{2,3}, Dominik L. Michels³, and Fei Sha¹ and the Mork ¹University of Southern California, ²Princeton University, ³Stanford University **ion for Hamiltonian Monte Carlo**

ia, ²Princeton University, ³Stanford University

ia, ²Princeton University, ³Stanford University
 al

state, large h w to tuning M

state, large h w to tuning M

state, large h

Highlights

 \blacksquare Improve the technique of integration **Fast step, large** *h* **w/o tuning** *M*

- HMC belongs to Metropolis-Hastings MCMC: proposal + acceptance test Proposal based on classical mechanics **Final positions of motion (after time** *t*) as proposals, governed by ODEs ■ In practice: divide *t* into *L* steps *h* Acceptance: $H(q^*, p^*) \approx H(q^{(i)}, p^{(i)})$ | | Gaussian approximately Challenges: **numerical integrators** Leapfrog: fast step, limited *h* Preconditioning: suitable *M* $\sim p(q) \Longrightarrow (q, p) \sim p(q, p) = p(q)p(p)$ | | 1) Linear: high $H(q, p)$ = $-\log p(q) - \log p(p)$ $U(q)$ + $\frac{1}{2}p^{T}M^{-1}p$ + const. $\begin{bmatrix} \blacksquare \end{bmatrix}$ lntegrate the using variation (q) + $\frac{1}{2}p^{\prime}M^{-1}p$ + const. **Explicit tor, which is more robust

DDEs than the standard

CDEs than the standard

gintegrator

te effective samples with

teeptance in short time
** $\frac{duction}{d}$ **
** $\frac{dfunction}{dx}$ **
** $\frac{dfunction}{dx}$ **
** $\frac{dfunction}{dx}$ **
** $\frac{dfunction}{dx}$ **
** $\frac{dfunction}{dx}$ **
 Example 19 and 19 Exploring the set of a exponential**
 Propose the use of a exponential
 Propose the use of a exponential
 **o stiff ODEs than the standard

eaptrog integrator**
 approach
 approach
 approach
 approach
 approac Propose the use of a exponential
 ntegrator, which is more robust
 **to stiff ODEs than the standard

eapfrog integrator**
 **Senerate effective samples with

nigh acceptance in short time

ntroduction**
 MNC belongs **Expansion Apple 11**
 Expansion $p(pq) \Rightarrow (q, p) \Rightarrow p(q) \Rightarrow (pq)$
 $= \frac{U(q)}{p(q)} + \frac{1}{2}p^TM^{-1}p + \text{const.}$
 Expansion of motion
 Expansion 6
 Expansion *p***: momentum

Expansion** *p***: momentum**
 Expansion *p***: momentum**
 **Expan than the standard

igrator

iective samples with

ince in short time

tion**

to Metropolis-Hastings

sal + acceptance test

d on classical mechanics
 p) ~ $p(q, p) = p(q)p(p)$

momentum
 $p(q) - log p(p)$
 q) + $\frac{1}{2}p^TM^{-1}p$ + , log log $1 \n\begin{array}{ccc}\nT & T & T^{-1} & \n\end{array}$ $= U(q) + \frac{1}{2}p^{\prime}M^{-1}p + \text{const.}$ $2¹$ and $1¹$ us potential kinetic $U(q)$ + $-p^{T}M^{-1}p$ + const. \parallel **F** If **Propose the use of a exponential regrator**
 **Constant to be the simples with

Experiment Constant Approach**
 Experiment Constant Approach
 Experiment Constant Approach

HMC belongs to Metropolis-Hastings
 Proposal *h* $\left(\rho, p\right) = -V_q H = H^{-1} p$
 $\left(\gamma + H = M^{-1} p\right)$
 $\gamma + \gamma = 0$ acceptance (stable)
 $\left(\gamma + H = M^{-1} p\right)$
 $\left(\gamma + H = M$ (q) (a) and q (a) and q (b) and q (c) and q $\dot{q} = \nabla_p H = M^{-1} p$ **Introduction**

HMC belongs to Metropolis-Hastings

MCMC: proposal + acceptance test

Proposal based on classical mechanics
 $q \sim p(q) \Rightarrow (q, p) \sim p(q, p) = p(q)p(p)$
 $H(q, p) = -\log p(q) - \log p(p)$
 $= \frac{U(q)}{\log q} + \frac{1}{2} p^r M^{-1} p + \text{const.}$

There **Introduction**

HMC belongs to Metropolis-Hastings

MCMC: proposal + acceptance test

Proposal based on classical mechanics
 $q \rightarrow p(q) \Rightarrow (q, p) \rightarrow p(q, p) = p(q)p(p)$
 $q \rightarrow p(q) \rightarrow p(q, p) = p(q)p(p)$
 $= \underbrace{U(q)}_{\text{position}} + \underbrace{1}_{\text{momentum}} \underbrace{1}_{\text{linear: high frequency (un$ **Introduction**

HMC belongs to Metropolis-Hastings

MCMC: proposal + acceptance test

Proposal based on classical mechanics
 $q \sim p(q) \Rightarrow (q, p) \sim p(q, p) = p(q)p(p)$
 $H(q, p) = -\log p(q) - \log p(p)$
 $= U(q) + \frac{1}{2}p^TM^{-1}p + \text{const.}$

Final position $\ddot{\boldsymbol{q}} + \boldsymbol{M}^{-1} \nabla_{\boldsymbol{q}} U(\boldsymbol{q}) = 0$ $\dot{\boldsymbol{p}} = - \nabla_{_{\boldsymbol{a}}} \boldsymbol{\mathcal{H}} = - \nabla_{_{\boldsymbol{a}}} \boldsymbol{U}\big(\boldsymbol{q}\big)$ p^{\prime} ρ^{\prime} ρ q^{\bigcup} (*Y*) \bigcup $y = -\log p(q) - \log p(p)$
 $= U(q) + \frac{1}{2} p^T M^{-1} p$

potential

sitions of motion (after oposals, governed by C
 $f = M^{-1} p$
 $qH = -\nabla_q U(q)$

tice: divide t into L step

prince: $H(q^*, n^*) \approx H(q)$ **ntroduction**

HMC belongs to Metropolis-Hastings

MCMC: proposal + acceptance test

Proposal based on classical mechanics
 $q \sim p(q) \Rightarrow (q, p) \sim p(q, p) = p(q)p(p)$
 $q:$ position p: momentum
 $H(q, p) = -\log p(q) - \log p(p)$
 $= U(q) + \frac{1}{2} p^{T} M$ **Approach**

reptance test
 $(p, p) = p(q)p(p)$
 $p^T M^{-1} p + \text{const.}$
 $p^T M^{-1} p$ **HMC** belongs to Metropolis-Hastings
 p *p* (*q*) \Rightarrow (*q*, *p*) \Rightarrow *p*(*q*, *p*) = *p*(*q*)*p*(*p*)
 q: position *p*: momentum
 p q: position *p*: momentum
 p q: position *p*: momentum
 p q p = -log $H = M^{-1} p$ **Exponential Integrator:**

Provided and Cassical mechanics
 $\begin{aligned}\n&\text{Exponential Integrate the ODE in}\\
&\text{Use the ODE$ *i* on classical mechanics
 $\frac{\partial}{\partial p}(q, p) = p(q)p(p)$
 $\frac{1}{2} \frac{1}{p} p^{i} M^{-1} p + \text{const.}$
 $\frac{1}{2} \frac{p^{i} M^{-1} p + \text{const.}}{k \text{inetic}}$
 iii $\frac{1}{2} \frac{1}{p} p^{i} M^{-1} p + \text{const.}$
 iii $\frac{1}{2} \frac{1}{p} p^{i} M^{-1} p + \text{const.}$
 iii $\frac{1}{2} \frac{1$
	- **≻ RMHMC: adaptive** *M***, slow step**
- **Investigate numerical integration of ODEs for Hamiltonian Monte Carlo (HMC)**
- **Propose the use of a** *exponential integrator*, **which is more robust leapfrog integrator**
- **high acceptance in short time**

Algorithm

Introduction

 \blacksquare Explicitly consider different frequencies Decompose the ODE into two terms

 $:$ position $p:$ momentum $|$ $|$ 2) Nonlinear: low frequency (stable) Integrate the linear term analytically \blacksquare Integrate the nonlinear term numerically using *variation of constants*

Goal

$L_{\mu}L_{\mu}$
 $\frac{1}{2}$ **Posterior of Bayesian logistic regression (BLR) ≻** Prior: $\mathcal{N}(0, T)$ with $T \in \{0.01, 1, 100\}$ *L* = 100, *h* selected to ensure suitable AR 2000 1500 200 $500 \vert$ \square 0.8 0.9 0.9 0.7 Acceptance rate Acceptance rate BLR: scalability & preconditioning & splitting

 \triangleright Posterior of Independent Component Analysis

 $\sqrt[1]{v_q U(q)} = 0$ \int $\sqrt[1]{\frac{1}{\sqrt{q}}}$ Pr time

DEs
 $\begin{bmatrix} \mathbf{F} \mathbf{x} \mathbf{p} \mathbf{p} \mathbf{p} \end{bmatrix}$
 $\begin{bmatrix} \mathbf{F} \mathbf{x} \mathbf{p} \mathbf{p} \mathbf{p} \end{bmatrix}$
 $\begin{bmatrix} \nabla_q U(q) = \dot{y}^{-1} q + f(q) = \frac{\dot{y}}{T} - \nabla_q \log \frac{\dot{y}}{T} \end{bmatrix}$
 $\begin{bmatrix} \nabla_q U(q) = \dot{y}^{-1} q + f(q) = \frac{\dot{y}}{T} - \nabla_q \log \frac$ **Decomposition** Gaussian approximation: 1) Laplace approximation (expHMC) $\begin{array}{c|c} \hline \end{array}$ 2) Empirical statistics Integration: ensure convergence of HMC Explicit integrators [Hairer & Lubich, 2000] Filter function [Garcia-Archilla et al., 1998] $\nabla_q U(q) = \dot{Y}^{-1} q + f(q) = \dot{Y}^{-1} q + (\nabla_q U(q) - \dot{Y}^{-1} q)$ $(q) = \mathcal{N}(q; -1, y) \times \frac{f'(1, y)}{f'(1, y)}$ *q* **pinential Integrator:**

plicitly consider different frequenc

ecompose the ODE into two terms

Linear: high frequency (unstable)

Nonlinear: low frequency (stable)

regrate the linear term analytically

tegrate the nonli *q*; ~, \dot{y}) $\left(\frac{1}{2}\right) \times \frac{P(1)}{P(1)}$ \ddot{y} $\ddot{\text{S}}$ $\mathcal{N}(\boldsymbol{q};-, \dot{y}) \times \frac{\rho(\mathbf{q})}{\mathcal{N}(\boldsymbol{q};-, \dot{y})}$; \sim , \dot{y}) $\times \frac{r(1)}{r}$; \sim , \ddot{y}) *p* $p(q) = \mathcal{N}(q; -1, y) \times \frac{P(1, y)}{P(1, y)}$ $-\nabla_q \log \frac{1}{r}$ (expHMC)

 $\ddot{q} + M^{-1}(\dot{y}^{-1}q + f(q)) = 0 \Longrightarrow \ddot{r} + \ln^2 r + F(r) = 0$

Approach

Exponential Integrator:

1) Linear: high frequency (unstable)

[1] R. Neal. MCMC using Hamiltonian dynamics. *Handbook of Markov Chain Monte Carlo*, 2011 [2] M. Girolami and B. Calderhead. Riemann Manifold Langevin and Hamiltonian Monte Carlo methods. *Journal of the Royal Statistical Society*, 2011

Experiments

Exponential HMC:

This work is supported by ARO Award # W911NF-12-1-0241, ONR # N000141210066, Alfred P. Sloan Fellowship, and the Max Planck Center for Visual Computing and Communication.