Generalized Dynamics for HMC

K. Harrison[1,2], X.-Y. Jin[1] and J.C. Osborn[1]

[1] Argonne National Laboratory

[2] North Carolina State University

July 28, 2026

Lattice 2026, UMD

Metropolis-Hastings (M-H) Algorithm

N. Metropolis, et al. (1953),
W.K. Hastings (1970)

- Want to sample from distribution: $W(z)$
- Choose some update probability function: $P(y|x)$
- Choose starting configuration $z^{(0)}$
  For i = 1..N:
  - Get configuration y from $P(y|z^{(i-1)})$
  - Calculate Acceptance Probability $A\left(y, z^{(i-1)}\right) = \min\left(1, \frac{P\left(z^{(i-1)}\middle|y\right) W(y)}{P\left(y\middle|z^{(i-1)}\right) W(z^{(i-1)})}\right)$
  - With probability $A\left(y, z^{(i-1)}\right)$ accept: $z^{(i)} = y$
    else reject: $z^{(i)} = z^{(i-1)}$

Hybrid Monte Carlo Algorithm

S. Duane, et al. (1987)

- Want to sample x from $e^{-V(x)}$
- Introduce auxiliary variables p
  - $z = (x, p)$
  - $W(z) = e^{-H(z)}$
  - $H(z) = T(p) + V(x)$
- Update probability function $P(u|z)$ from Hamilton dynamics
  - Choose new p from distribution $e^{-T(p)}$
  - Approximately, but reversibly, integrate equations of motion to time τ, obtain $u = (y, q)$
    - $\dot{x}_i = \frac{dH}{dp_i} \qquad \dot{p}_i = -\frac{dH}{dx_i}$
  - Accept/reject using $A(u, z) = \frac{P(Z|\mathcal{U})W(u)}{P(\mathcal{U}|Z)W(z)} = \frac{W(u)}{W(z)} \left|\frac{du}{dz}\right|$

Generalized Dynamics HMC (GDHMC)

- Consider general dynamics equations

$$\dot{z}_i = F_i(z)$$

- Can work out constraint for perfect M-H acceptance

$$\left(\frac{dH}{dz_i} - \frac{d}{dz_i}\right) F_i(z) = 0$$

- Note that for a symplectic update

$$\frac{d}{dz_i} F_i(z) = 0$$

which gives energy conservation

$$0 = \frac{dH}{dz_i} F_i(z) = \frac{dH}{dz_i} \dot{z}_i = \dot{H}$$

Generalized Dynamics HMC (GDHMC)

- Dynamics constraint

$$\left(\frac{dH}{dz_i} - \frac{d}{dz_i}\right) F_i(z) = 0$$

- A general solution

$$F_i(z) = C_i(z) e^{H_i} + \left(\frac{dH}{dz_j} - \frac{d}{dz_j}\right) G_{ij}(z)$$

$H_i(z)$ is all parts of $H(z)$ that depend on z_i

$C_i(z)$ does not depend on z_i

$G_{ij}(z) = -G_{ji}(z)$

Also need to ensure reversibility

Generalized Dynamics HMC (GDHMC)

- Expanding in original (x) and auxiliary (q) variables (dropping C_i terms)

$$F_i^x(x,q) = \left(\frac{dH}{dx_j} - \frac{d}{dx_j}\right) G_{ij}^{xx}(x,q) + \left(\frac{dH}{dq_j} - \frac{d}{dq_j}\right) G_{ij}^{xq}(x,q)$$
$$F_i^q(x,q) = \left(\frac{dH}{dx_j} - \frac{d}{dx_j}\right) G_{ij}^{qx}(x,q) + \left(\frac{dH}{dq_j} - \frac{d}{dq_j}\right) G_{ij}^{qq}(x,q)$$

$$G_{ij}^{xx} = -G_{ji}^{xx} \quad G_{ij}^{qq} = -G_{ji}^{qq} \quad G_{ij}^{qx} = -G_{ji}^{xq}$$

- Note that the number of q variables is arbitrary (including 0, 1, …)
- Standard HMC (q=p): $G_{ij}^{xx} = G_{ij}^{qq} = 0 \quad G_{ij}^{xq} = -G_{ji}^{qx} = \delta_{ij}$

Relation to other HMC methods

- Riemannian Manifold HMC (RMHMC)

S. Duane, et al. (1986),
M. Girolami, et al. (2011)

$$H_R = \frac{1}{2} p^T M(x)^{-1} p + \frac{1}{2} \ln|M(x)| + V(x)$$

$$\dot{x}_i = \frac{dH_R}{dp_i} \qquad \dot{p}_i = -\frac{dH_R}{dx_i}$$

- Dynamics equivalent to GDHMC with

$$M^{-1} = G\, G^T \qquad q = G^T p \qquad H_G = \frac{1}{2} q^T q + V(x)$$

$$G^{xx} = 0 \quad G^{xq} = G \quad G^{qq} = \cdots$$

Relation to other HMC methods

M. Lüscher (2010)

- Field Transformation HMC (FTHMC)

$$H_F = \frac{1}{2} q^T q + V(f(y)) - \ln\left|\frac{df}{dy}\right|$$

$$\dot{y}_i = \frac{dH_F}{dq_i} \qquad \dot{q}_i = -\frac{dH_F}{dy_i}$$

- Dynamics equivalent to GDHMC with

$$x_i = f_i(y) \qquad H_G = \frac{1}{2} q^T q + V(x)$$

$$G^{xx} = 0 \qquad G^{xq} = \left(\frac{df^{-1}}{dx}\right)^{-1} \qquad G^{qq} = 0$$

Relation to other HMC methods

- Nambu dynamics

Y. Nambu (1973)

  - Two sets of auxiliary variables: p, r
  - Additional constraint function G_N, require $\dot{G}_N = 0$
  - Equations of motion

$$\dot{x}_i = \frac{dH}{dp_i}\frac{dG_N}{dr_i} - \frac{dH}{dr_i}\frac{dG_N}{dp_i} \qquad \dot{p}_i = \frac{dH}{dr_i}\frac{dG_N}{dx_i} - \frac{dH}{dx_i}\frac{dG_N}{dr_i} \qquad \dot{r}_i = \frac{dH}{dx_i}\frac{dG_N}{dp_i} - \frac{dH}{dp_i}\frac{dG_N}{dx_i}$$

- Dynamics equivalent to GDHMC with

$$G_{ij}^{xp} = -G_{ji}^{px} = \frac{dG_N}{dr_i}\delta_{ij} \quad G_{ij}^{pr} = -G_{ji}^{rp} = \frac{dG_N}{dx_i}\delta_{ij} \quad G_{ij}^{rx} = -G_{ji}^{xr} = \frac{dG_N}{dp_i}\delta_{ij}$$

Relation to other sampling methods

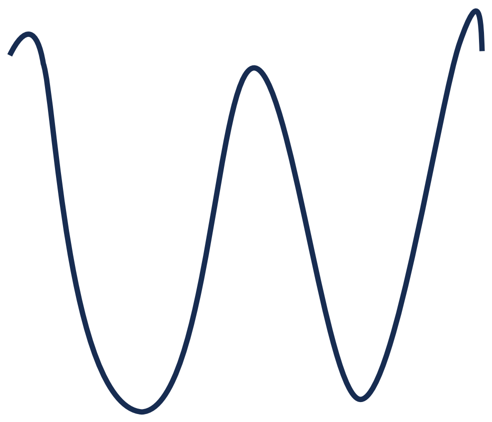
- Consider one variation of GDHMC

$$G^{xx}_{ij} = G^{pp}_{ij} = 0 \qquad G^{xp}_{ij} = -G^{px}_{ji} = e^{V_i(x)}\delta_{ij}$$

$$\dot{x}_i = e^{V_i(x)} p_i$$

$$\dot{p}_i = 0$$

  - No force term, doesn't avoid potential barriers
  - Instead, velocity becomes large, chance of stopping low



$V_i(x_i)$

- Similar to heatbath if integration time were random

GDHMC experiments: double well potential

- Double well potential

$$V(x) = x^4 - \alpha x^2$$

- Can't use full potential in G^{xp} since x would diverge
- Instead use just middle barrier

$$G^{xp} = e^{-\alpha x^2}$$

- Cancels force from barrier, keeps confining potential
- Performed simulations to test feasibility of method
- Leapfrog integrator with adaptive ODE solver for x update: $\dot{x} = e^{-\alpha x^2} p$
- Found GDHMC performed well when HMC failed

GDHMC experiments: 1d XY + periodic double well

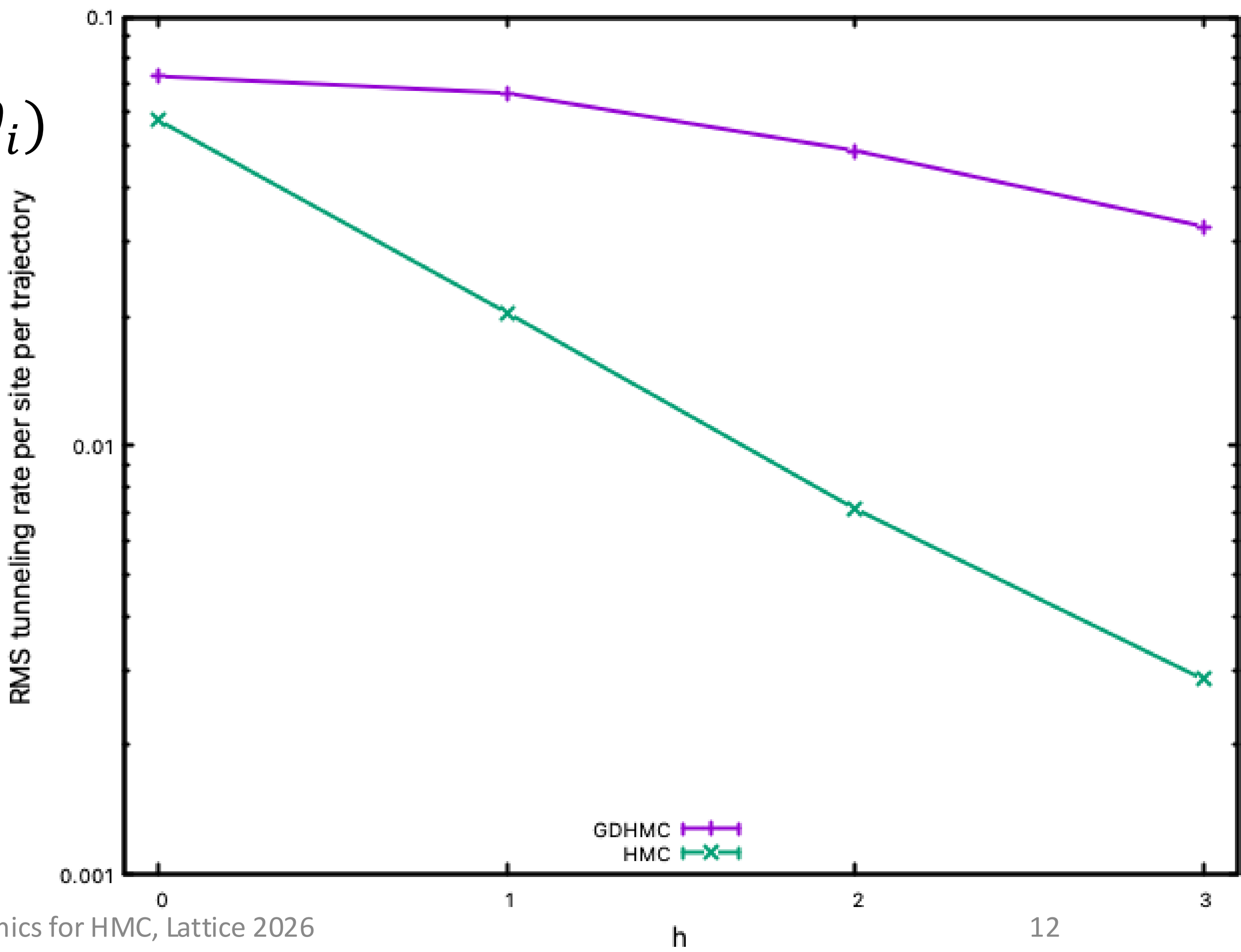
- 1d XY model with extra local potential

$$V = \sum_{<ij>} 1 - \cos(\theta_i - \theta_j) + h \sum_i \cos(2\theta_i)$$

- Compact variables, can use full potential

$$G_{ij}^{\theta p} = e^{V_i \delta_{ij}}$$

- Adaptive ODE solver for θ_i update, one site at a time (alternating even/odd)
- Plot of tunneling rate per site for 128 sites
- GDHMC has higher tunneling rate
- Also tried $G_i^{\theta p} = e^{h \cos(2\theta_i)}$ found full potential worked best
- GDHMC with adaptive solve still slower needs improvement



Tests in 2d U(1) pure gauge theory

- 2d U(1) pure gauge theory

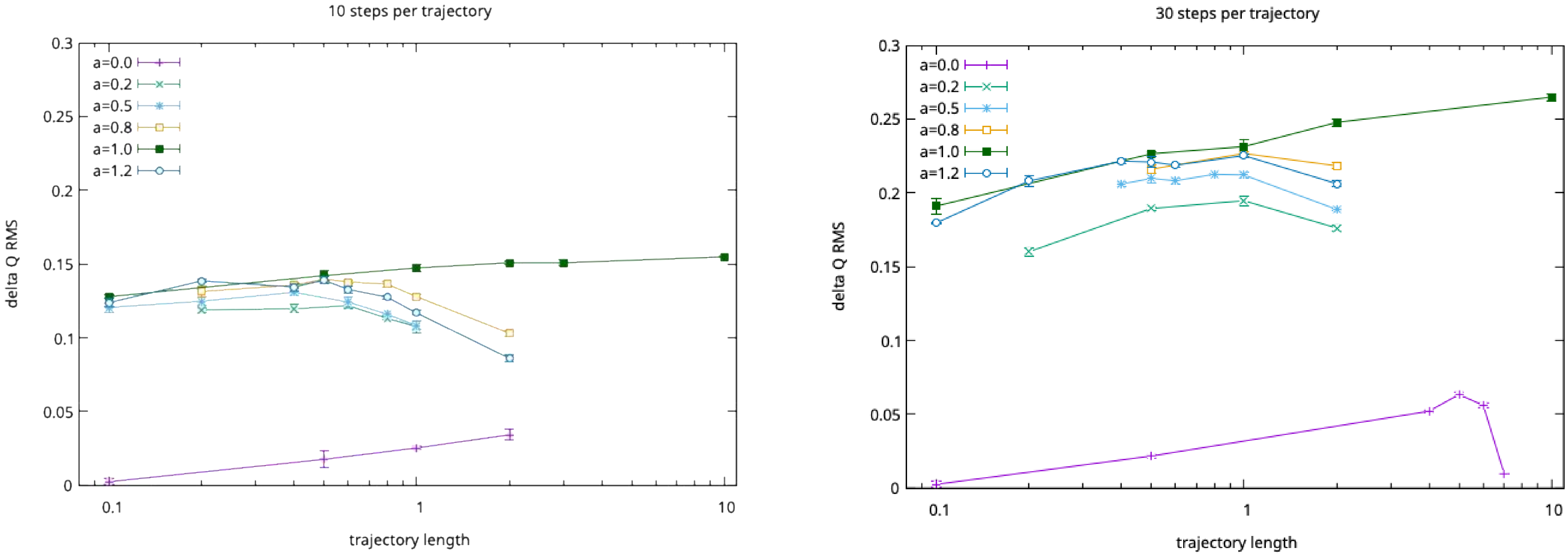
$$G_{ii}^{\theta p} = (1-a) + ae^{V_i}$$

$$\dot{p}_i = -(1-a)\frac{dV}{d\theta_i}$$

- Parameters
  - L = 64 , β = 8
  - Adaptive ODE for θ_i update, alternating even/odd sites and direction
  - 10,000 trajectories per run
  - 4 runs with different seeds per set of parameters
  - All runs start from the same initial thermalized configuration

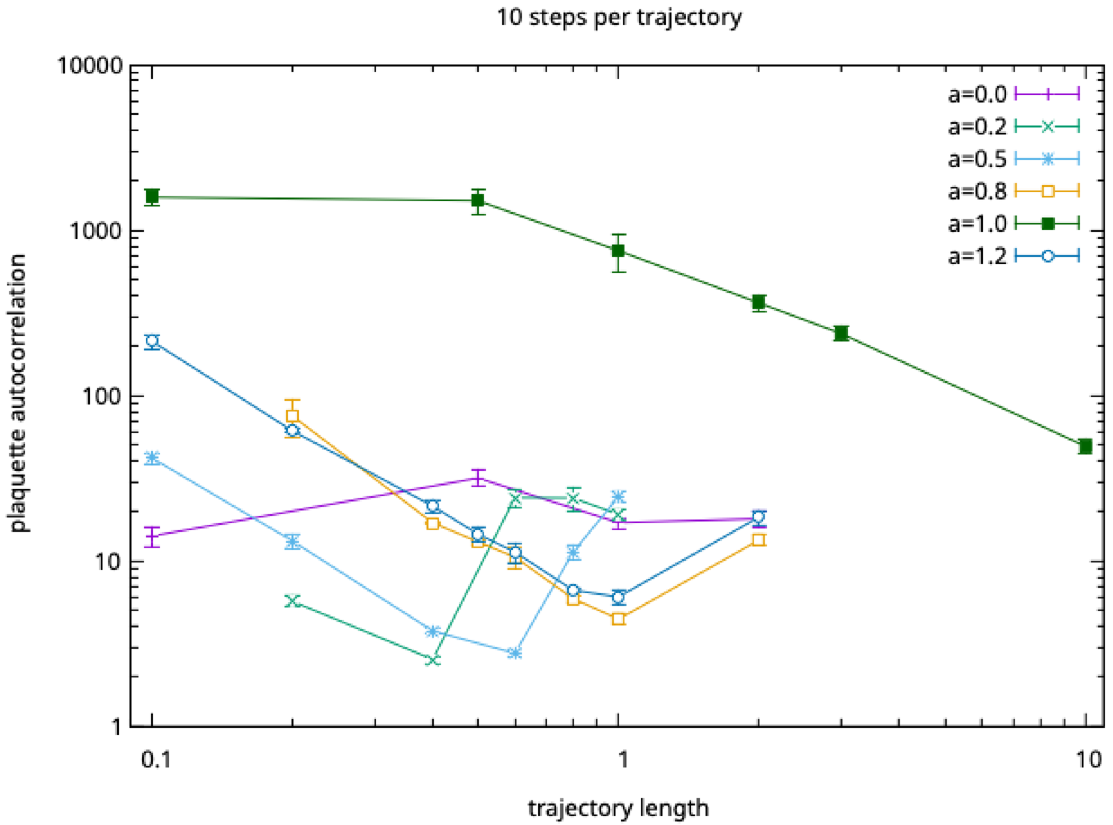
Topology tunneling rate

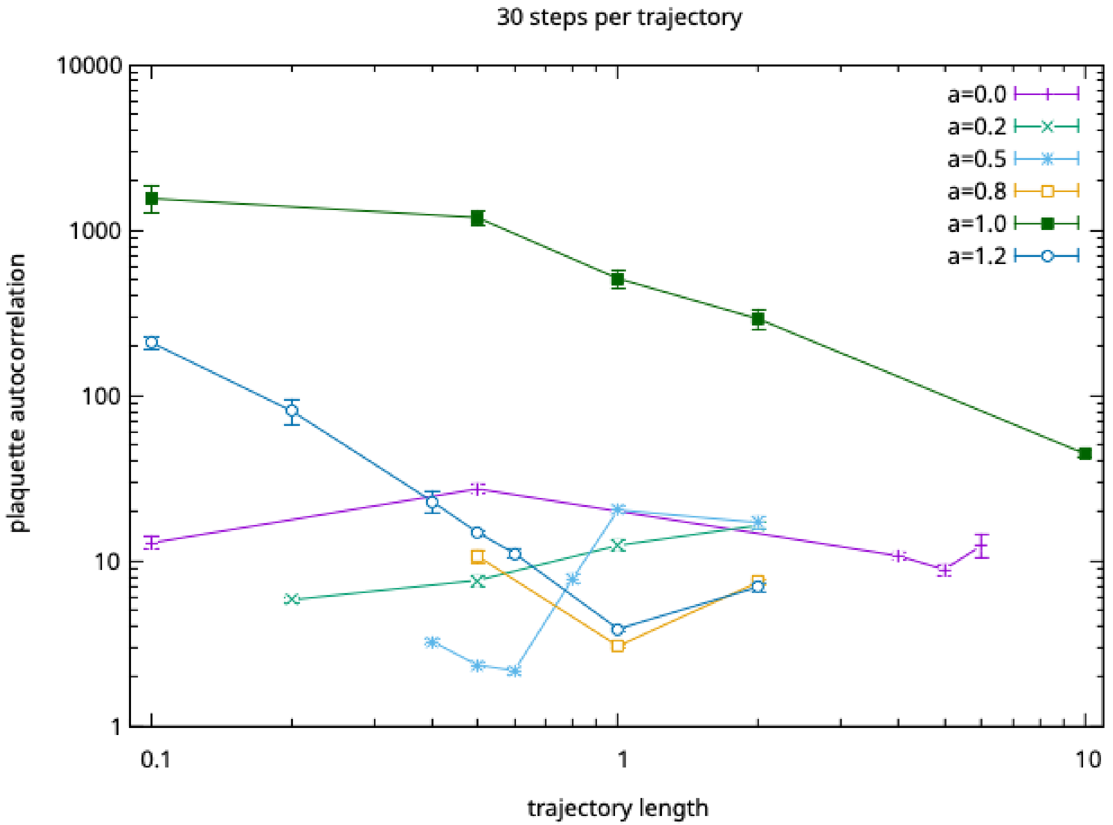
- Tunneling rate
  delta Q RMS = $\sqrt{\sum(Q_{i+1} - Q_i)^2/(N-1)}$
- GDHMC has 4-5x tunneling rate of HMC



Plaquette autocorrelation

- Plaquette integrated autocorrelation time
- GDHMC a≠1 can reduce plaquette autocorrelation over HMC
- GDHMC a=1 has much larger autocorrelation
- Plaquette force helps evolve plaquette
- Maybe adding a topology force would help evolve topology?



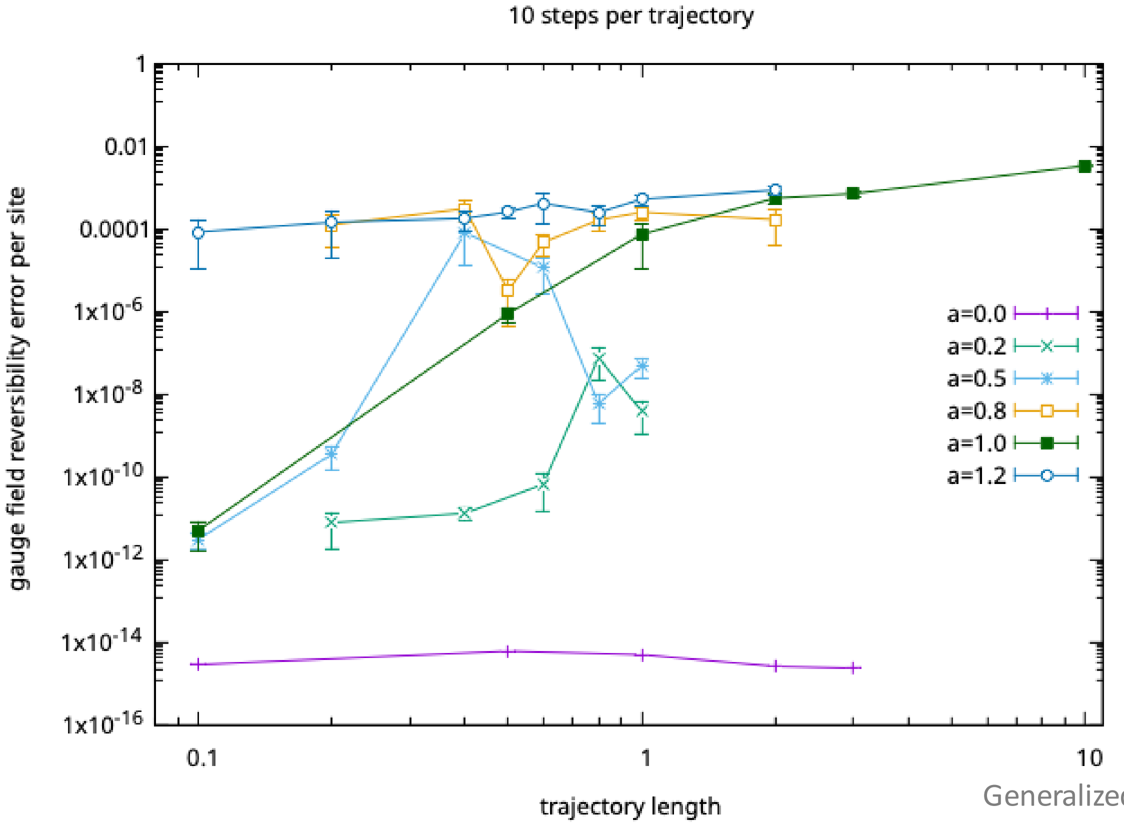


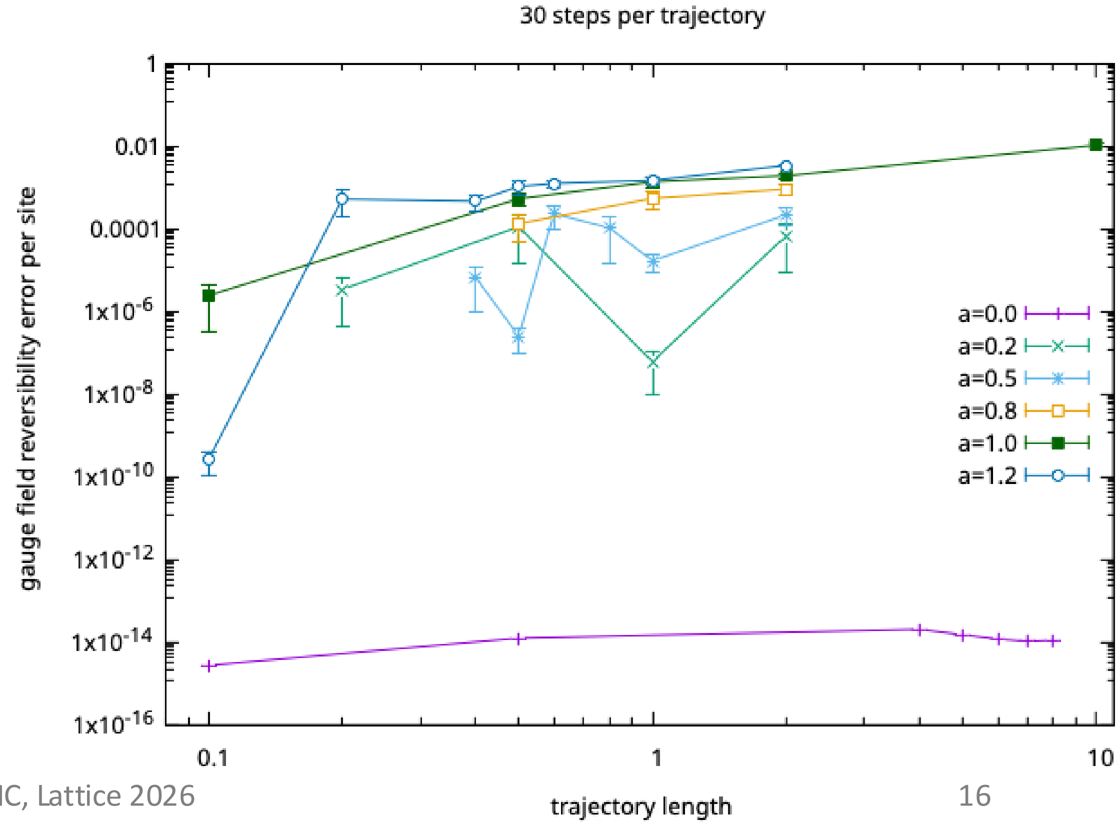
Reversibility error

- RMS error in gauge field reversibility per site $\left\langle \sqrt{\sum \left(U_i^0 - U_i^R\right)^2 / V} \right\rangle_i$

  U_i^0: field at beginning of trajectory,

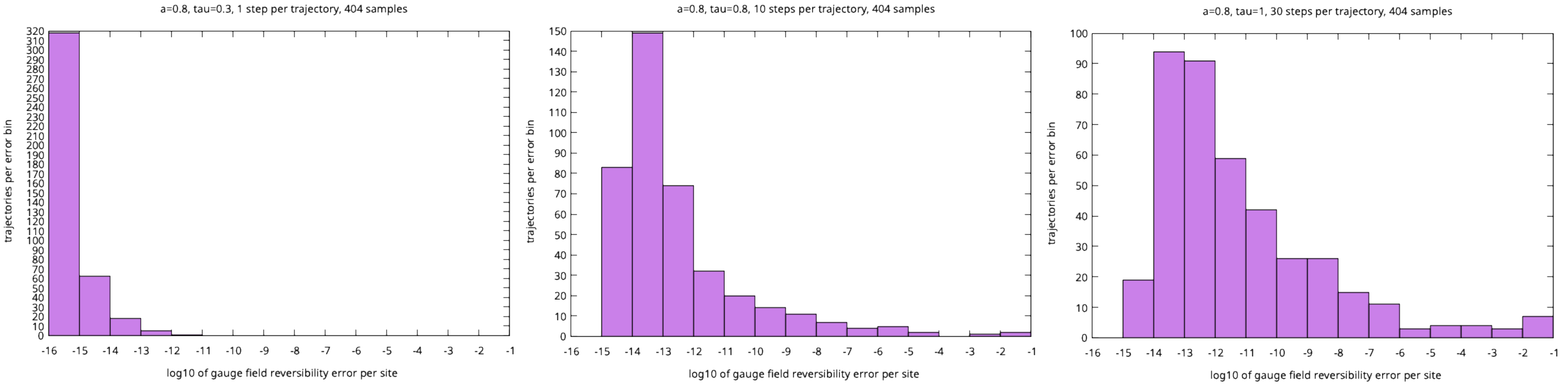
  U_i^R: field after running in reverse from end of trajectory
- HMC reversibility error very low
- GDHMC reversibility error grows quickly with number of steps





Reversibility error

- Histogram over trajectories of RMS error in gauge field reversibility per site
- Including results from 1 step (left) for reference, with 10 steps (center) and 30 steps (right)
- All plots for a = 0.8
- Error for single step small, but errors accumulate quickly for multiple steps



Summary

- Formulated equations for generalized dynamics within HMC constrained by M-H acceptance criterion
- Presented general solution to constraint equation parameterized by general functions, allowing easy construction of variations
- Showed how several other HMC variants map to the general solution
- Tested one variant of GDHMC designed to reduce or eliminate force
  - Found that solving for coordinate update is manageable in 2d U(1) theory
  - Need to understand reversibility error better and control it (fewer steps)
- Explore large space of variations
  - Try range of forms for $G_{ij}(z)$, including topology, …
  - More auxiliary variables
  - Can include variables shared across sites (i.e. coarse grained/collective variables)
  - lots of other possibilities

ODE methods

- Currently solving single-site ODE
  - Adaptive ODE (Runge-Kutta)
    - Solve to high precision to make reversible
  - Implicit integrator
    - Need to keep step size small to ensure reversibility
  - Series expansion
    - Expand ODE in series and integrate
    - Have for a=1 case, faster than adaptive
- Could try evolving all sites together
  - Use implicit or adaptive methods
  - Likely requires smaller average step size,
    but evolving all sites together may be worthwhile