Johann Ostmeyer



Helmholtz-Institut für Strahlen- und Kernphysik, University of Bonn, Germany

May 17, 2024









0. Goal: sample from distribution $x \sim e^{-S(x)}$

0. Goal: sample from distribution $x \sim e^{-S(x)}$

1. Define artificial Hamiltonian $\mathcal{H} = \frac{1}{2}p^2 + S(x)$

- 0. Goal: sample from distribution $x \sim \mathrm{e}^{-S(x)}$
- 1. Define artificial Hamiltonian $\mathcal{H} = \frac{1}{2}p^2 + S(x)$
- 2. Sample momenta $p \sim e^{-\frac{p^2}{2}}$

- 0. Goal: sample from distribution $x \sim e^{-S(x)}$
- 1. Define artificial Hamiltonian $\mathcal{H} = \frac{1}{2}p^2 + S(x)$
- 2. Sample momenta $p \sim e^{-\frac{p^2}{2}}$
- 3. Solve equations of motion (EOM) numerically

$$\dot{x} = \frac{\partial \mathcal{H}}{\partial p}$$
$$\dot{p} = -\frac{\partial \mathcal{H}}{\partial x}$$

- 0. Goal: sample from distribution $x \sim e^{-S(x)}$
- 1. Define artificial Hamiltonian $\mathcal{H} = \frac{1}{2}p^2 + S(x)$
- 2. Sample momenta $p \sim e^{-\frac{p^2}{2}}$
- 3. Solve equations of motion (EOM) numerically

$$\dot{x} = -\frac{\partial \mathcal{H}}{\partial p}$$
$$\dot{p} = -\frac{\partial \mathcal{H}}{\partial x}$$

4. Accept new x with probability $p_{\sf acc} = \min\left(1, {
m e}^{-\varDelta {\cal H}}
ight)$

0. Presume action $S(x) = \frac{1}{2}x^{\mathsf{T}}Mx$

0. Presume action
$$S(x) = \frac{1}{2}x^{\mathsf{T}}Mx$$

1. Diagonalise
$$M = \Omega \cdot \operatorname{diag}(\omega^2) \cdot \Omega^{\dagger}$$

0. Presume action
$$S(x) = \frac{1}{2}x^{\mathsf{T}}Mx$$

1. Diagonalise
$$M = \Omega \cdot \operatorname{diag}(\omega^2) \cdot \Omega^{\dagger}$$

2. Sample $r \sim \mathcal{N}(0,1)^{\dim(M)}$

0. Presume action
$$S(x) = \frac{1}{2}x^{\mathsf{T}}Mx$$

1. Diagonalise
$$M = \Omega \cdot \operatorname{diag}(\omega^2) \cdot \Omega^{\dagger}$$

2. Sample
$$r \sim \mathcal{N}(0,1)^{\dim(M)}$$

3. Then
$$x = \Omega \cdot \operatorname{diag}(\omega) \cdot \Omega^{\dagger} \cdot r$$
 follows $x \sim \mathrm{e}^{-S(x)}$

Sampling from an *almost* normal distribution [JO et al. *cond-mat.mtrl-sci*/2312.14914; Xing et al. *PRL* 126 (2021)]

Now
$$S(x) = \frac{1}{2}x^{\mathsf{T}}Mx + \varepsilon V(x)$$

[JO et al. cond-mat.mtrl-sci/2312.14914; Xing et al. PRL 126 (2021)]

Now
$$S(x) = \frac{1}{2}x^{\mathsf{T}}Mx + \varepsilon V(x)$$

No direct sampling!

[JO et al. cond-mat.mtrl-sci/2312.14914; Xing et al. PRL 126 (2021)]

Now
$$S(x) = \frac{1}{2}x^{\mathsf{T}}Mx + \varepsilon V(x)$$

- ► No direct sampling!
- ► Use HMC?

[JO et al. cond-mat.mtrl-sci/2312.14914; Xing et al. PRL 126 (2021)]

Now
$$S(x) = \frac{1}{2}x^{\mathsf{T}}Mx + \varepsilon V(x)$$

- ► No direct sampling!
- ► Use HMC?

[JO et al. cond-mat.mtrl-sci/2312.14914; Xing et al. PRL 126 (2021)]



4/12

$$\blacktriangleright \mathcal{H} = \frac{1}{2} \sum_{i} \left[\omega_i^2 x_i^2 + \frac{p^2}{m_i^2} \right]$$

•
$$\mathcal{H} = \frac{1}{2} \sum_{i} \left[\omega_i^2 x_i^2 + \frac{p^2}{m_i^2} \right]$$

• Equations of motion

$$\dot{x}_i = \frac{p_i}{m_i^2}$$
$$\dot{p}_i = -\omega_i^2 x_i$$

►
$$\mathcal{H} = \frac{1}{2} \sum_{i} \left[\omega_i^2 x_i^2 + \frac{p^2}{m_i^2} \right]$$

► Equations of motion

$$\dot{x}_i = \frac{p_i}{m_i^2}$$
$$\dot{p}_i = -\omega_i^2 x_i$$



$$x_i(t) = x_i^0 \cos\left(\frac{\omega_i}{m_i}t\right) + \frac{1}{m_i\omega_i} p_i^0 \sin\left(\frac{\omega_i}{m_i}t\right)$$
$$p_i(t) = p_i^0 \cos\left(\frac{\omega_i}{m_i}t\right) - m_i\omega_i x_i^0 \sin\left(\frac{\omega_i}{m_i}t\right)$$

►
$$\mathcal{H} = \frac{1}{2} \sum_{i} \left[\omega_i^2 x_i^2 + \frac{p^2}{m_i^2} \right]$$

► Equations of motion

$$\dot{x}_i = \frac{p_i}{m_i^2}$$
$$\dot{p}_i = -\omega_i^2 x_i$$



$$x_i(t) = x_i^0 \cos\left(\frac{\omega_i}{m_i}t\right) + \frac{1}{m_i\omega_i} p_i^0 \sin\left(\frac{\omega_i}{m_i}t\right)$$
$$p_i(t) = p_i^0 \cos\left(\frac{\omega_i}{m_i}t\right) - m_i\omega_i x_i^0 \sin\left(\frac{\omega_i}{m_i}t\right)$$

• Choose $m_i = \omega_i$ and trajectory length $T = \frac{\pi}{2}$

"Fourier acceleration"

[Batrouni et al. PRD 32 (1985); Cohen-Stead et al. PRE 105 (2022); JO & Buividovich hep-lat/2404.09723]

Theorem (Optimal HMC trajectory length and kinetic term) *Given the harmonic action*

$$S(x) = \frac{1}{2} x^{\mathsf{T}} M x \,,$$

then the Hamiltonian

$$\mathcal{H} = \frac{1}{2}p^{\mathsf{T}}M^{-1}p + S(x)$$

together with the HMC trajectory length $T = \frac{\pi}{2}$ make HMC equivalent to uncorrelated direct sampling.

[JO & Buividovich hep-lat/2404.09723; Urbach et al. CPC 174 (2006)]

Algorithm: Leap-Frog step for
$$\mathcal{H} = \frac{1}{2}p^{\mathsf{T}}M^{-1}p + \frac{1}{2}x^{\mathsf{T}}Mx + V(x)$$

[JO & Buividovich hep-lat/2404.09723; Urbach et al. CPC 174 (2006)]

Algorithm: Leap-Frog step for
$$\mathcal{H} = \frac{1}{2}p^{\mathsf{T}}M^{-1}p + \frac{1}{2}x^{\mathsf{T}}Mx + V(x)$$

 $(x,p) \leftarrow \mathsf{EFA}\left(x^{0},p^{0},h/2\right);$
 $p \leftarrow p - h \cdot \nabla V(x);$
 $(x(h),p(h)) \leftarrow \mathsf{EFA}\left(x,p,h/2\right);$

[JO & Buividovich hep-lat/2404.09723; Urbach et al. CPC 174 (2006)]

Algorithm: Leap-Frog step for
$$\mathcal{H} = \frac{1}{2}p^{\mathsf{T}}M^{-1}p + \frac{1}{2}x^{\mathsf{T}}Mx + V(x)$$

 $(x,p) \leftarrow \mathsf{EFA}\left(x^{0},p^{0},h/2\right);$
 $p \leftarrow p - h \cdot \nabla V(x);$
 $(x(h),p(h)) \leftarrow \mathsf{EFA}\left(x,p,h/2\right);$

Algorithm: EFA with $M = \Omega \cdot \operatorname{diag}(\omega^2) \cdot \Omega^{\dagger}$

[JO & Buividovich hep-lat/2404.09723; Urbach et al. CPC 174 (2006)]

Algorithm: Leap-Frog step for
$$\mathcal{H} = \frac{1}{2}p^{\mathsf{T}}M^{-1}p + \frac{1}{2}x^{\mathsf{T}}Mx + V(x)$$

 $(x,p) \leftarrow \mathsf{EFA}\left(x^{0},p^{0},h/2\right);$
 $p \leftarrow p - h \cdot \nabla V(x);$
 $(x(h),p(h)) \leftarrow \mathsf{EFA}\left(x,p,h/2\right);$

Algorithm: EFA with $M = \Omega \cdot \operatorname{diag}(\omega^2) \cdot \Omega^{\dagger}$

$$y^0 \leftarrow \Omega^\dagger \cdot x^0; q^0 \leftarrow \Omega^\dagger \cdot p^0;$$

[JO & Buividovich hep-lat/2404.09723; Urbach et al. CPC 174 (2006)]

Algorithm: Leap-Frog step for
$$\mathcal{H} = \frac{1}{2}p^{\mathsf{T}}M^{-1}p + \frac{1}{2}x^{\mathsf{T}}Mx + V(x)$$

 $(x,p) \leftarrow \mathsf{EFA}\left(x^{0},p^{0},h/2\right);$
 $p \leftarrow p - h \cdot \nabla V(x);$
 $(x(h),p(h)) \leftarrow \mathsf{EFA}\left(x,p,h/2\right);$

Algorithm: EFA with $M = \Omega \cdot \operatorname{diag}(\omega^2) \cdot \Omega^{\dagger}$

 $y^{0} \leftarrow \Omega^{\dagger} \cdot x^{0}; q^{0} \leftarrow \Omega^{\dagger} \cdot p^{0};$ $y_{i}(h) \leftarrow \cos(h) y_{i}^{0} + \frac{1}{\omega_{i}^{2}} \sin(h) q_{i}^{0};$ $q_{i}(h) \leftarrow \cos(h) q_{i}^{0} - \omega_{i}^{2} \sin(h) y_{i}^{0};$

[JO & Buividovich hep-lat/2404.09723; Urbach et al. CPC 174 (2006)]

Algorithm: Leap-Frog step for
$$\mathcal{H} = \frac{1}{2}p^{\mathsf{T}}M^{-1}p + \frac{1}{2}x^{\mathsf{T}}Mx + V(x)$$

 $(x,p) \leftarrow \mathsf{EFA}\left(x^{0},p^{0},h/2\right);$
 $p \leftarrow p - h \cdot \nabla V(x);$
 $(x(h),p(h)) \leftarrow \mathsf{EFA}\left(x,p,h/2\right);$

Algorithm: EFA with $M = \Omega \cdot \operatorname{diag}(\omega^2) \cdot \Omega^{\dagger}$

$$y^{0} \leftarrow \Omega^{\dagger} \cdot x^{0}; q^{0} \leftarrow \Omega^{\dagger} \cdot p^{0};$$

$$y_{i}(h) \leftarrow \cos(h) y_{i}^{0} + \frac{1}{\omega_{i}^{2}} \sin(h) q_{i}^{0};$$

$$q_{i}(h) \leftarrow \cos(h) q_{i}^{0} - \omega_{i}^{2} \sin(h) y_{i}^{0};$$

$$x(h) \leftarrow \Omega \cdot y(h); p(h) \leftarrow \Omega \cdot q(h);$$

Benchmarking the SSH model [JO & Buividovich hep-lat/2404.09723]

$$S_{\text{SSH}} = \sum_{t} \left[\omega_0^2 x_t^2 + N_t^2 \left(x_{t+1} - x_t \right)^2 \right] + \text{weak electron interactions}$$



Critical slowing down in the classical Ising model

[JO & Buividovich hep-lat/2404.09723]



Critical slowing down in the classical Ising model [JO & Buividovich *hep-lat*/**2404**.09723]



$$\tau_{\rm int}(T) = \frac{\tau_{\rm int}(T = \pi/2)}{1 - \cos(T)^{\pi/2T}}$$

Gauge Theories with Fourier acceleration

[Borsanyi et al. Science 347 (2015); Duane & Pendleton Phys. Lett. B 206 (1988)]

 $2D, U(1), \beta = 10$



Gauge Theories with Fourier acceleration

[Borsanyi et al. Science 347 (2015); Duane & Pendleton Phys. Lett. B 206 (1988)]

 $2D, U(1), \beta = 10$

$$4D, SU(3), L = 10$$

10

в

30



[Apers et al. stat.ML/2209.12771; Meyer et al. CPC 176 (2007)]

[Apers et al. stat.ML/2209.12771; Meyer et al. CPC 176 (2007)]

 \blacktriangleright Randomise trajectory length T

[Apers et al. stat.ML/2209.12771; Meyer et al. CPC 176 (2007)]

 \blacktriangleright Randomise trajectory length T

• Scale
$$T \propto \sqrt{c_M}$$
, $c_M = \frac{\lambda_{\max}}{\lambda_{\min}}$ is condition number of M

[Apers et al. stat.ML/2209.12771; Meyer et al. CPC 176 (2007)]

 \blacktriangleright Randomise trajectory length T

• Scale
$$T \propto \sqrt{c_M}$$
, $c_M = rac{\lambda_{\max}}{\lambda_{\min}}$ is condition number of M

▶ T too long ⇒ unnecessary work $T_{\text{CPU}} \propto T \propto \sqrt{c_M}$

[Apers et al. stat.ML/2209.12771; Meyer et al. CPC 176 (2007)]

 \blacktriangleright Randomise trajectory length T

• Scale
$$T \propto \sqrt{c_M}$$
, $c_M = rac{\lambda_{\max}}{\lambda_{\min}}$ is condition number of M

- T too long \Rightarrow unnecessary work $T_{\rm CPU} \propto T \propto \sqrt{c_M}$
- ► T too short \Rightarrow random walk $T_{\text{CPU}} \propto \tau_{\text{int}} \propto c_M$

[Apers et al. stat.ML/2209.12771; Meyer et al. CPC 176 (2007)]

 \blacktriangleright Randomise trajectory length T

▶ Scale
$$T \propto \sqrt{c_M}$$
, $c_M = rac{\lambda_{\max}}{\lambda_{\min}}$ is condition number of M

- ▶ T too long ⇒ unnecessary work $T_{\rm CPU} \propto T \propto \sqrt{c_M}$
- ▶ T too short \Rightarrow random walk $T_{CPU} \propto \tau_{int} \propto c_M$
- \Rightarrow better too long than too short

For action

$$S(x) = \frac{1}{2}x^{\mathsf{T}}Mx + \varepsilon V(x)$$
choose Hamiltonian

$$\mathcal{H} = \frac{1}{2}p^{\mathsf{T}}M^{-1}p + S(x)$$

For action

$$S(x) = \frac{1}{2}x^{\mathsf{T}}Mx + \varepsilon V(x)$$
choose Hamiltonian

$$\mathcal{H} = \frac{1}{2}p^{\mathsf{T}}M^{-1}p + S(x)$$

▶ Trajectory length $T = \frac{\pi}{2}$

For action

$$S(x) = \frac{1}{2}x^{\mathsf{T}}Mx + \varepsilon V(x)$$
choose Hamiltonian

$$\mathcal{H} = \frac{1}{2}p^{\mathsf{T}}M^{-1}p + S(x)$$

- ▶ Trajectory length $T = \frac{\pi}{2}$
- Integrate harmonic EOM exactly

For action

$$S(x) = \frac{1}{2}x^{\mathsf{T}}Mx + \varepsilon V(x)$$
choose Hamiltonian

$$\mathcal{H} = \frac{1}{2}p^{\mathsf{T}}M^{-1}p + S(x)$$

• Trajectory length
$$T = \frac{\pi}{2}$$



$$S_{\text{SSH}} = \sum_{t} \left[\omega_0^2 x_t^2 + N_t^2 \left(x_{t+1} - x_t \right)^2 \right] + \dots$$

The Su-Schrieffer-Heeger (SSH) model [Su et al. PRL 42 (1979)] |

Without Fourier acceleration $au_{ ext{int}} \propto 1 + \left(rac{2N_t}{eta\omega_0}
ight)^2$.

$$\begin{split} H_{\text{SSH}} &= \omega_0 \sum_{i,\alpha} \left(a_{i,\alpha}^{\dagger} a_{i,\alpha} + \frac{1}{2} \right) \\ &- \sum_{i,\alpha} J_{\alpha} \left(1 - \lambda_{\alpha} x_{i,\alpha} \right) \left(c_i^{\dagger} c_{i+\alpha} + c_{i+\alpha}^{\dagger} c_i \right) - \mu \sum_i c_i^{\dagger} c_i \end{split}$$

The Su-Schrieffer-Heeger (SSH) model [Su et al. PRL 42 (1979)] II

$$\begin{split} S_{\text{SSH}} &= \frac{1}{2} x^{\text{T}} M_{\text{SSH}} x + \text{electron interactions} \\ &= \frac{\beta}{2N_t} \sum_t \left[\omega_0^2 x_t^2 + \frac{N_t^2}{\beta^2} \left(x_{t+1} - x_t \right)^2 \right] + \text{electron interactions} \\ &= \frac{1}{2\beta N_t} \sum_{\xi} \left[\left(\beta \omega_0 \right)^2 + 4N_t^2 \sin^2 \left(\frac{\pi}{N_t} \xi \right) \right] y_{\xi}^2 + \text{electron interactions} \,, \\ &y_{\xi} &= \frac{1}{\sqrt{N_t}} \sum_t e^{-i\frac{2\pi}{N_t}\xi t} x_t \,, \quad \xi = 0, \dots, N_t - 1 \,. \end{split}$$

Autocorrelation [Wolff CPC 156 (2004)]

Integrated autocorrelation time

$$\tau_{\text{int}} \equiv \frac{1}{2} + \sum_{t=1}^{\infty} \rho_{\mathcal{A}}(t)$$
$$\rho_{\mathcal{A}}(t) = \frac{1}{\sigma_{\mathcal{A}}^2 N} \sum_{t_0=1}^{N} \left(\mathcal{A}(t_0 + t) - \langle \mathcal{A} \rangle \right) \left(\mathcal{A}(t_0) - \langle \mathcal{A} \rangle \right)$$

No Fourier acceleration [JO & Buividovich hep-lat/2404.09723]

Corollary

If no FA is used, i.e. $\mathcal{H} = \frac{1}{2}p^2 + \frac{1}{2}x^TMx$, with fixed trajectory length, then

$$au_{
m int} \propto \left(rac{\omega_{
m max}}{\omega_{
m min}}
ight)^2 + \mathcal{O}\left(1
ight) \,,$$

where ω_{\min}^2 , ω_{\max}^2 are the smallest/largest eigenvalue of M.

Too short trajectory [JO & Buividovich hep-lat/2404.09723]

Corollary

Presuming FA is used together with short trajectory length $T < \frac{\pi}{2}$ and observable is measured in intervals of fixed HMC time, then

$$au_{
m int}(T) = rac{ au_{
m int}(T = \pi/2)}{1 - \cos(T)^{\pi/2T}} \,.$$

Bibliography I

- ¹ S. Apers, S. Gribling, D. Szilágyi, "Hamiltonian Monte Carlo for efficient Gaussian sampling: long and random steps", 10.48550/arXiv.2209.12771 (*stat.ML*/2209.12771).
- ² G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky, K. G. Wilson, "Langevin simulations of lattice field theories", Phys. Rev. D 32, 2736–2747 (*PRD* 32 (1985)).
- ³ S. Borsanyi, S. Durr, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, L. Lellouch, T. Lippert, A. Portelli, K. K. Szabo, B. C. Toth, "Ab initio calculation of the neutron-proton mass difference", Science 347, 1452–1455 (*Science* 347 (2015)).
- ⁴ B. Cohen-Stead, O. Bradley, C. Miles, G. Batrouni, R. Scalettar, K. Barros, "Fast and scalable quantum Monte Carlo simulations of electron-phonon models", Phys. Rev. E. 105, 065302 (*PRE* 105 (2022)).
- ⁵ S. Duane, A. D. Kennedy, B. J. Pendleton, D. Roweth, "Hybrid Monte Carlo", Phys. Lett. B 195, 216 –222 (*PhysLetB* 195 (1987)).
- ⁶ S. Duane, B. J. Pendleton, "Gauge invariant fourier acceleration", Physics Letters B 206, 101–106 (*Phys. Lett. B* 206 (1988)).

Bibliography II

- ⁷ JO, P. Buividovich, "Minimal Autocorrelation in Hybrid Monte Carlo simulations using Exact Fourier Acceleration", 10.48550/arXiv.2404.09723 (*hep-lat*/2404.09723).
- ⁸ JO, T. Nematiaram, A. Troisi, P. Buividovich, "First-principle quantum Monte-Carlo study of charge carrier mobility in organic molecular semiconductors", 10.48550/arXiv.2312.14914 (cond-mat.mtrl-sci/2312.14914).
- ⁹ H. Meyer, H. Simma, R. Sommer, M. Della Morte, O. Witzel, U. Wolff, "Exploring the HMC trajectory-length dependence of autocorrelation times in lattice QCD", Computer Physics Communications 176, 91–97 (CPC 176 (2007)).
- ¹⁰W. P. Su, J. R. Schrieffer, A. J. Heeger, "Solitons in Polyacetylene", Phys. Rev. Lett. 42, 1698–1701 (*PRL* 42 (1979)).
- ¹¹C. Urbach, K. Jansen, A. Shindler, U. Wenger, "HMC algorithm with multiple time scale integration and mass preconditioning", Computer Physics Communications 174, 87–98 (CPC 174 (2006)).

Bibliography III

¹²U. Wolff, "Monte Carlo errors with less errors", Computer Physics Communications 156, 143–153 (*CPC* 156 (2004)).

¹³B. Xing, W.-T. Chiu, D. Poletti, R. T. Scalettar, G. Batrouni, "Quantum Monte Carlo Simulations of the 2D Su-Schrieffer-Heeger Model", Phys. Rev. Lett. **126**, 017601 (*PRL* **126** (2021)).