# Hybrid Monte Carlo methods for sampling on submanifolds

**Gabriel STOLTZ**

(CERMICS, Ecole des Ponts & MATHERIALS team, INRIA Paris)

*In collaboration with T. Lelièvre and M. Rousset*

*Work also supported by ANR Funding ANR-14-CE23-0012 ("COSMOS")*

# Outline

Generalize[1] Zappa/Holmes-Cerfon/Goodman (2017): large timesteps

- **Motivation**
  - Computational statistical physics
  - Where constraints appear
  - Metropolis & standard Generalized Hybrid Monte Carlo
- **RATTLE dynamics with reverse projection check** (truly reversible)
  - Standard RATTLE scheme
  - "Abstract" reversible RATTLE scheme
  - Local and theoretical realization through the implicit function theorem
  - A more practical scheme based on Newton's method
- **Generalized Hybrid Monte Carlo algorithms** (Reversibility is key!)
- **Some numerical results**

[1] T. Lelièvre, M. Rousset, G. Stoltz, *arXiv preprint* **1807.02356**

# Motivation

# Computational statistical physics

- *Predict macroscopic properties of matter from its microscopic description*

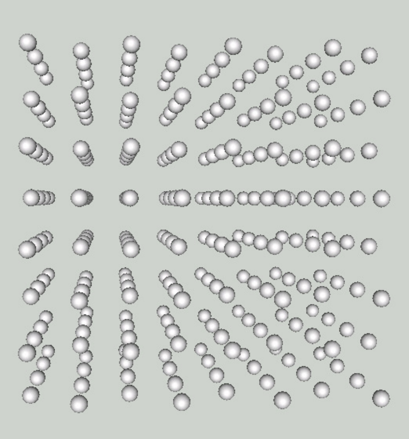
- **Microstate**
  - positions $q = (q_1, \dots, q_N)$ and momenta $p = (p_1, \dots, p_N)$
  - energy $V(q) + \sum_{i=1}^{N} \frac{p_i^2}{2m_i}$

- **Macrostate**
  - described by a probability measure $\mu$
  - constraints fixed exactly or in average (number of particles, volume, energy)

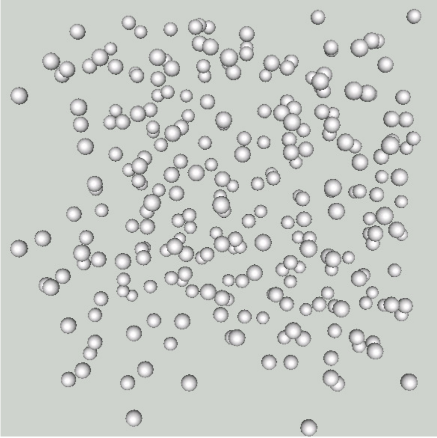
- **Properties :**
  - static $\langle A \rangle = \int_{\mathcal{E}} A(q,p)\,\mu(dq\,dp)$ (equation of state, heat capacity,...)
  - dynamic (transport coefficient, transition pathway, etc)

# Examples of molecular systems (1)

What is the melting temperature of Argon?



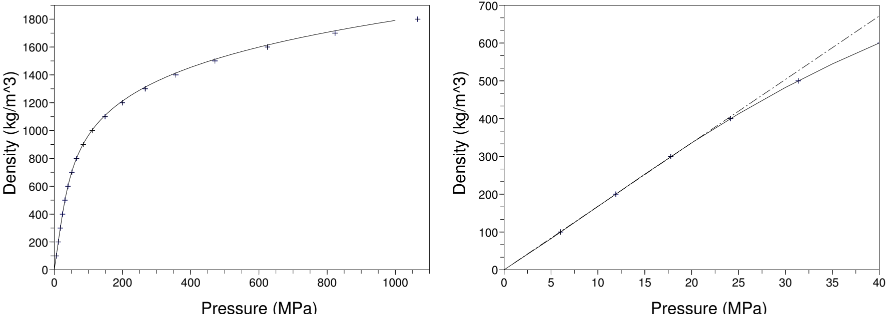
(a) Solid Argon (low temperature)



(b) Liquid Argon (high temperature)

# Examples of molecular systems (2)

Equation of state of Argon: density as a function of pressure at fixed temperature $T = 300$ K



# Sampling measures with constraints

- **Typical probability measures in stat. physics/Bayesian statistics**
  - unknowns = parameters in statistics, atomic coordinates for stat phys
  - position space measure $Z^{-1}\mathrm{e}^{-\beta V(q)}\,dq$ with $\beta^{-1} = k_{\mathrm{B}}T$
  - phase-space measure

$$\mu(dq\,dp) = Z^{-1}\mathrm{e}^{-\beta H(q,p)}\,dq\,dp, \quad H(q,p) = V(q) + \sum_{i=1}^{N} \frac{p_i^2}{2m_i}$$

- Equality constraints arise from
  - molecular constraints (fixed bond lengths, angles, etc)
  - fixed values of reaction coordinates $\xi(q)$ [free energy]

- Inequality constraints could be considered as well

# Metropolis-Hastings algorithm (1)

- Markov chain method[2,3] to sample $\nu(dq) = Z^{-1}\mathrm{e}^{-\beta V(q)}\, dq$

  - Given $q^n$, propose $\tilde{q}^{n+1}$ according to transition probability $T(q^n, \tilde{q})$
  - Accept with probability

$$\min\left(1, \frac{T(\tilde{q}^{n+1}, q^n)\, \nu(\tilde{q}^{n+1})}{T(q^n, \tilde{q}^{n+1})\, \nu(q^n)}\right),$$

  and set in this case $q^{n+1} = \tilde{q}^{n+1}$; otherwise, set $q^{n+1} = q^n$.

- Example of proposals
  - Gaussian displacement $\tilde{q}^{n+1} = q^n + \sigma\, G^n$ with $G^n \sim \mathcal{N}(0, \mathrm{Id})$
  - Biased random walk[4,5] $\tilde{q}^{n+1} = q^n - \alpha \nabla V(q^n) + \sqrt{2\alpha\beta^{-1}}\, G^n$

---

[2] Metropolis, Rosenbluth (×2), Teller (×2), *J. Chem. Phys.* (1953)
[3] W. K. Hastings, *Biometrika* (1970)
[4] G. Roberts and R.L. Tweedie, *Bernoulli* (1996)
[5] P.J. Rossky, J.D. Doll and H.L. Friedman, *J. Chem. Phys.* (1978)

# Metropolis-Hastings algorithm (2)

- Transition kernel

$$P(q, dq') = \min\left(1, r(q, q')\right) T(q, q')\, dq' + \left(1 - \alpha(q)\right)\delta_q(dq'),$$

where $\alpha(q) \in [0, 1]$ is the probability to accept a move starting from $q$:

$$\alpha(q) = \int_{\mathcal{D}} \min\left(1, r(q, q')\right) T(q, q')\, dq'.$$

- The canonical measure is reversible with respect to $\nu$, hence invariant:

$$P(q, dq')\nu(dq) = P(q', dq)\nu(dq')$$

- Pathwise ergodicity[6] when the chain is irreducible

$$\lim_{N \to +\infty} \frac{1}{N} \sum_{n=1}^{N} A(q^n) = \int_{\mathcal{D}} A(q)\, \nu(dq)$$

Allows for unbiased sampling and stabilization of numerical schemes

[6] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability* (1993)

# How GHMC works (1)

• **Aim:** sample the phase-space measure through Hamiltonian dynamics + momentum resampling

$$\begin{cases} \dot{q}(t) = M^{-1} p(t), \\ \dot{p}(t) = -\nabla V(q(t)) \end{cases}$$

Reversibility: $\phi_t \circ S = S \circ \phi_{-t}$ where $S(q,p) = (q,-p)$ and $\phi_t$ flow

• In practice, discretization using a reversible scheme, e.g. Verlet

$$\begin{cases} p^{n+1/2} = p^n - \dfrac{\Delta t}{2} \nabla V(q^n) \\ q^{n+1} = q^n + \Delta t \, M^{-1} p^{n+1/2} \\ p^{n+1} = p^{n+1/2} - \dfrac{\Delta t}{2} \nabla V(q^{n+1}) \end{cases}$$

• Two importants properties of the scheme: reversible and preserves the Lebesgue measure

# How GHMC works (2)

- Transition kernel $T(x, x')$ with $x = (q, p)$
- Assume that $r(x, x') = \dfrac{T(S(x'), S(dx))\,\pi(dx')}{T(x, dx')\,\pi(dx)}$ is defined and positive[7]

**Generalized Hybrid Monte Carlo (Horowitz, 1991)**

- given $x^n$, propose a new state $\tilde{x}^{n+1}$ from $x^n$ according to $T(x^n, \cdot)$;
- accept the move with probability $\min\left(1, r(x^n, \tilde{x}^{n+1})\right)$, and set in this case $x^{n+1} = \tilde{x}^{n+1}$; otherwise, set $x^{n+1} = S(x^n)$.

- Reversibility up to $S$, *i.e.* $P(x, dx')\,\mu(dx) = P(S(x'), S(dx))\,\mu(dx')$

- Standard HMC: $T(q, dq') = \delta_{\Phi_\tau(q)}(dq')$, momentum reversal upon rejection (not important since momenta are resampled, but is important when momenta are partially resampled)

[7] T. Lelièvre, M. Rousset, and G. Stoltz, *Free Energy Computations: A Mathematical Perspective*

# How GHMC works (3)

**Complete algorithm:** starting from $(q^0, p^0)$,

(i) update the momentum as $\widetilde{p}^{n+1} = \alpha p^n + \sqrt{\frac{m(1-\alpha^2)}{\beta}}\, G^n$

(ii) propose $(\widetilde{q}^{n+1}, p^{n+1}) = \Phi_{\Delta t}(q^n, \widetilde{p}^{n+1})$

(iii) accept with probability $\min\left(1, \mathrm{e}^{-\beta[H(\widetilde{q}^{n+1},p^{n+1})-H(q^n,p^n)]}\right)$ and set $(q^{n+1}, p^{n+1}) = (\widetilde{q}^{n+1}, p^{n+1})$ in this case; otherwise set $(q^{n+1}, p^{n+1}) = (q^n, -\widetilde{p}^{n+1})$

• **Limiting case $\alpha = 0$:** one-step HMC = MALA = Euler-Maruyama discretization of the overdamped Langevin dynamics + Metropolis

$$\widetilde{q}^{n+1} = q^n - h\nabla V(q^n) + \sqrt{\frac{2h}{\beta}}\, G^n, \qquad h = \frac{\Delta t^2}{2}$$

• Possible application: sampling eigenvalues of random matrices[8]

[8] D. Chafaï and G. Ferre, *arXiv preprint* **1806.05985**

# (Truly) Reversible RATTLE dynamics

# Constrained Gibbs measures (1)

- Submanifold: level set of smooth function $\xi : \mathbb{R}^d \to \mathbb{R}^m$ with $m < d$:

$$\mathcal{M} = \left\{ q \in \mathbb{R}^d, \, \xi(q) = 0 \right\}$$

- $M \in \mathbb{R}^{d \times d}$ fixed symmetric positive definite matrix

**Assumption**

The matrix $G_M(q) = [\nabla\xi(q)]^T \, M^{-1} \nabla\xi(q) \in \mathbb{R}^{m \times m}$ is invertible in a neighborhood of $\mathcal{M}$ in $\mathbb{R}^d$

- Associated cotangent space

$$T_q^*\mathcal{M} = \left\{ p \in \mathbb{R}^d, \, [\nabla\xi(q)]^T \, M^{-1} p = 0 \right\} \subset \mathbb{R}^d$$

and cotangent bundle

$$T^*\mathcal{M} = \left\{ (q,p) \in \mathbb{R}^d \times \mathbb{R}^d, \, \xi(q) = 0 \text{ and } [\nabla\xi(q)]^T \, M^{-1} p = 0 \right\} \subset \mathbb{R}^d \times \mathbb{R}^d$$

# The RATTLE integrator

- Second order discretization of the constrained Hamiltonian dynamics

$$\begin{cases} \quad dq_t = M^{-1} p_t \, dt \,, \\ \quad dp_t = -\nabla V(q_t) \, dt + \nabla \xi(q_t) \, d\lambda_t \,, \\ \xi(q_t) = 0 \end{cases}$$

## RATTLE scheme (Andersen, 1983)

$$\begin{cases} p^{n+1/2} = p^n - \dfrac{\Delta t}{2} \nabla V(q^n) + \nabla \xi(q^n) \, \lambda^{n+1/2}, \\ q^{n+1} = q^n + \Delta t \, M^{-1} \, p^{n+1/2}, \\ \xi(q^{n+1}) = 0, & (C_q) \\ p^{n+1} = p^{n+1/2} - \dfrac{\Delta t}{2} \nabla V(q^{n+1}) + \nabla \xi(q^{n+1}) \, \lambda^{n+1}, \\ \left[ \nabla \xi(q^{n+1}) \right]^T M^{-1} p^{n+1} = 0, & (C_p) \end{cases}$$

- Momentum constraint always satisfied, but not the position constraint

# Formal reversibility of RATTLE

- Start from $(q^{n+1}, -p^{n+1})$ and go to $(q^n, -p^n)$
- Initially $\left[\nabla\xi(q^{n+1})\right]^T M^{-1} p^{n+1} = 0$ and $\xi(q^{n+1}) = 0$
- Call $\widetilde{\lambda}^{n+1}$ and $\widetilde{\lambda}^{n+1/2}$ the Lagrange multipliers

$$
\begin{cases}
\left[\nabla\xi(q^{n+1})\right]^T M^{-1} p^{n+1} = 0, \\
-p^{n+1/2} = -p^{n+1} - \dfrac{\Delta t}{2} \nabla V(q^{n+1}) + \nabla\xi(q^{n+1})\, \widetilde{\lambda}^{n+1/2}, \\
\xi(q^{n+1}) = 0, \\
q^n = q^{n+1} - \Delta t\, M^{-1}\, p^{n+1/2}, \\
\xi(q^n) = 0, & (C_q) \\
-p^n = -p^{n+1/2} - \dfrac{\Delta t}{2} \nabla V(q^n) + \nabla\xi(q^n)\, \widetilde{\lambda}^{n+1}, \\
\left[\nabla\xi(q^n)\right]^T M^{-1} p^n = 0, & (C_p)
\end{cases}
$$

- Suggests $\widetilde{\lambda}^{n+1} = \lambda^{n+1/2}$ and $\widetilde{\lambda}^{n+1/2} = \lambda^{n+1}$

# Admissible Lagrange multiplier functions

- Note that $q^{n+1} = \tilde{q}^n + \Delta t\, M^{-1} \nabla \xi(q^n) \lambda^{n+1/2}$ (unconstrained move $\tilde{q}^n$)

- Lagrange multipliers $\Delta t \lambda^{n+1/2} = \Lambda(q^n, \tilde{q}^n)$ function of current position (direction of projection) and unconstrained move $\tilde{q}^n$ (can be far off $q^n$)

### Admissible Lagrange multiplier function $\Lambda$

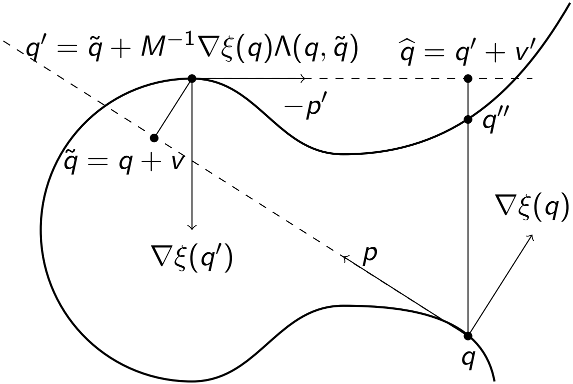
$C^1$ function defined on an open set $\mathcal{D}$ of $\mathcal{M} \times \mathbb{R}^d$ with values in $\mathbb{R}^m$ with

- projection property: $\forall (q, \tilde{q}) \in \mathcal{D}, \quad \tilde{q} + M^{-1} \nabla \xi(q) \Lambda(q, \tilde{q}) \in \mathcal{M}$
- non-tangential projection property: for all $(q, \tilde{q}) \in \mathcal{D}$,

$$\left[ \nabla \xi \left( \tilde{q} + M^{-1} \nabla \xi(q) \Lambda(q, \tilde{q}) \right) \right]^T M^{-1} \nabla \xi(q) \in \mathbb{R}^{m \times m} \text{ is invertible.}$$

- $\mathcal{D}$ contain elements $(q, \tilde{q}) \in \mathcal{M} \times \mathcal{M}$ for which $\left[ \nabla \xi \left( \tilde{q} \right) \right]^T M^{-1} \nabla \xi(q)$ is invertible (in this case, $\Lambda(q, \tilde{q}) = 0$)

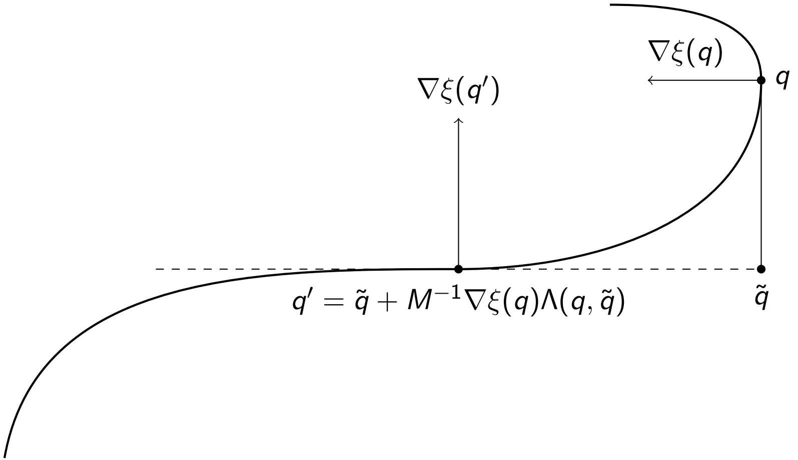
# What can go wrong with the projection?




The projection may not exist, or may not be unique

RATTLE may not be reversible for large timesteps due to the choice of projection

# About the non-tangential property




There may be infinitely many possible projections (not isolated points)

# Towards a reversible RATTLE scheme

• Composing RATTLE with momentum reversal (involution = good for Metropolis!)

• Admissible set (open) of moves which can be projected back onto $\mathcal{M}$

$$A = \left\{ (q,p) \in T^*\mathcal{M}, \ \left( q, q + \Delta t \, M^{-1} \left[ p - \frac{\Delta t}{2} \nabla V(q) \right] \right) \in \mathcal{D} \right\}$$

Can be proved to be non-empty!

• Define $\Psi_{\Delta t}(q,p) = (q^1, -p^1)$ for $(q,p) \in A$ where $(q^1, p^1)$ is obtained from $(q,p)$ by one step of the RATTLE scheme

### Properties of $\Psi_{\Delta t}$

The application $\Psi_{\Delta t} : A \to T^*\mathcal{M}$ is a $C^1$ local diffeomorphism, locally preserving the phase-space measure[a] $\sigma_{T^*\mathcal{M}}(dq\, dp)$

[a] to be defined later on...

# The reversible RATTLE scheme

- Difficulty: analysis at fixed $\Delta t$, for all configurations $(q, p)$

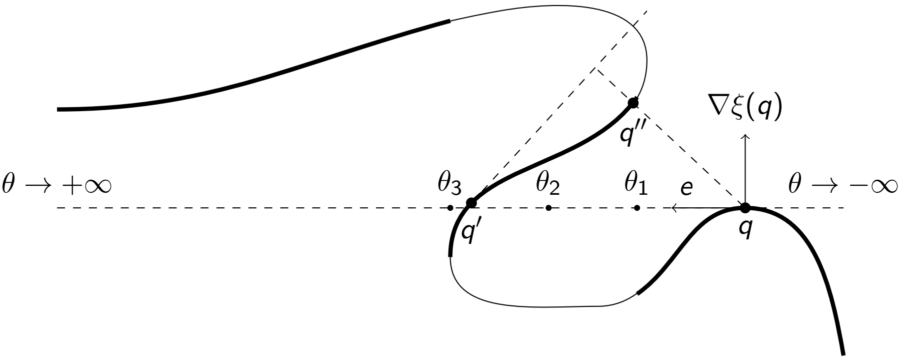
## Reversible RATTLE scheme

Define $\Psi_{\Delta t}^{\mathrm{rev}}(q, p) = \Psi_{\Delta t}(q, p) 1_{\{(q,p) \in B\}} + (q, p) 1_{\{(q,p) \notin B\}}$ where

$$B = \left\{ (q, p) \in A, \ \Psi_{\Delta t}(q, p) \in A \text{ and } (\Psi_{\Delta t} \circ \Psi_{\Delta t})(q, p) = (q, p) \right\}$$

- Explicitly, for any $(q, p) \in T^*\mathcal{M}$,
  - (i) check if $(q, p)$ is in $A$; if not return $(q, p)$;
  - (ii) when $(q, p) \in A$, compute the configuration $(q^1, p^1)$ obtained by one step of the RATTLE scheme;
  - (iii) check if $(q^1, -p^1)$ is in $A$; if not, return $(q, p)$;
  - (iv) compute the configuration $(q^2, -p^2)$ obtained by one step of the RATTLE scheme starting from $(q^1, -p^1)$;
  - (v) if $(q^2, p^2) = (q, p)$, return $(q^1, -p^1)$; otherwise return $(q, p)$.

# Illustrating the reverse projection check




Reverse projection check...

- not successful for increments corresponding to $\theta \in (\theta_2, \theta_3)$
- successful for small increments (corresponding to $\theta < \theta_2$) or for sufficiently large ones (corresponding to $\theta > \theta_3$)

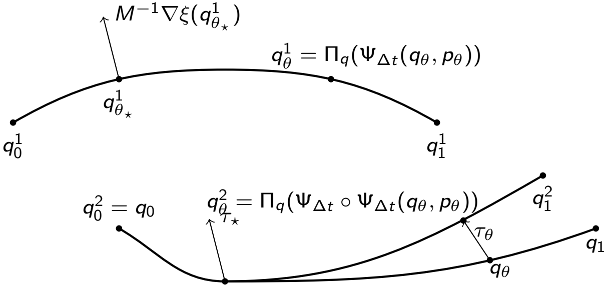
# Properties of the reversible RATTLE integrator (1)

## On the structure of the set $B$

Let $C$ be a path connected component of $A \cap \Psi_{\Delta t}^{-1}(A)$. If there is $(q,p) \in C$ such that $(\Psi_{\Delta t} \circ \Psi_{\Delta t})(q,p) = (q,p)$, then

$$\forall (q,p) \in C, \qquad (\Psi_{\Delta t} \circ \Psi_{\Delta t})(q,p) = (q,p).$$

As a corollary, the set $B$ is the union of path connected components of the open set $A \cap \Psi_{\Delta t}^{-1}(A)$. In particular, it is an open set of $T^*\mathcal{M}$.



# Properties of the reversible RATTLE integrator (2)

## Reversibility and measure preservation

The map $\Psi_{\Delta t}^{\mathrm{rev}} : T^*\mathcal{M} \to T^*\mathcal{M}$ is globally well defined, and satisfies

$$\Psi_{\Delta t}^{\mathrm{rev}} \circ \Psi_{\Delta t}^{\mathrm{rev}} = \mathrm{Id}.$$

Moreover, both $\Psi_{\Delta t}^{\mathrm{rev}} : B \to B$ and $\Psi_{\Delta t}^{\mathrm{rev}} : B^c \to B^c$ are $C^1$-diffeomorphisms which preserve the measure $\sigma_{T^*\mathcal{M}}(dq\,dp)$. As a consequence, $\Psi_{\Delta t}^{\mathrm{rev}} : T^*\mathcal{M} \to T^*\mathcal{M}$ globally preserves the measure $\sigma_{T^*\mathcal{M}}(dq\,dp)$.

# Practical reversible RATTLE dynamics

# Theorerical realization: implicit function theorem (1)

- Assume for simplicity $\{q \in \mathbb{R}^d, \|\xi(q)\| \leqslant \alpha\}$ compact for some $\alpha > 0$

There exists an open subset $\mathcal{D}_{\rm imp}$ of $\mathcal{M} \times \mathbb{R}^d$ and an admissible Lagrange multiplier function $\Lambda : \mathcal{D}_{\rm imp} \to \mathbb{R}^m$ such that

- $\mathcal{G}_M(q,\tilde{q}) = [\nabla\xi(q)]^T M^{-1} \nabla\xi(\tilde{q}) \in \mathbb{R}^{m\times m}$ is invertible on $\mathcal{D}_{\rm imp}$;
- $\mathcal{E} = \Big\{(q,\tilde{q}) \in \mathcal{M}^2,\ \mathcal{G}_M(q,\tilde{q}) \text{ is invertible}\Big\} \subset \mathcal{D}_{\rm imp}$ and $\Lambda = 0$ on $\mathcal{E}$;
- For any $(q_0,\tilde{q}_0) \in \mathcal{D}_{\rm imp}$, there is a neighborhood $\mathcal{V}_0$ of $(q_0,\tilde{q}_0)$ in $\mathcal{D}_{\rm imp}$ and $\alpha_0 > 0$ such that

$$\forall (q,\tilde{q}) \in \mathcal{V}_0,\ \|\Lambda(q,\tilde{q})\| < \alpha_0 \text{ and } \forall \lambda \in \mathbb{R}^m \setminus \{\Lambda(q,\tilde{q})\},$$
$$\xi(\tilde{q} + M^{-1}\nabla\xi(q)\lambda) = 0 \implies \|\lambda\| \geq \alpha_0$$

- A few comments...
  - points $q$ and $\tilde{q}$ in $\mathcal{D}_{\rm imp}$ are not required to be close (but $\tilde{q}$ should still be close to $\mathcal{M}$)
  - the Lagrange multiplier is the smallest solution in norm

# Theorerical realization: implicit function theorem (2)

- Introduce the sets

$$A_{\mathrm{imp}} = \left\{ (q,p) \in T^*\mathcal{M}, \ \left( q, q + \Delta t\, M^{-1} \left[ p - \frac{\Delta t}{2} \nabla V(q) \right] \right) \in \mathcal{D}_{\mathrm{imp}} \right\}$$

$$B_{\mathrm{imp}} = \left\{ (q,p) \in A_{\mathrm{imp}} \cap \Psi_{\Delta t}^{-1}(A_{\mathrm{imp}}), \ (\Psi_{\Delta t} \circ \Psi_{\Delta t})(q,p) = (q,p) \right\}$$

- Non empty for $\Delta t$ sufficiently small ($\Psi_{\Delta t}^{\mathrm{rev}}(q,p) = (q,p)$ for some $p$)

## Local reversibility result

There exists $\beta > 0$ (independent of $\Delta t > 0$) such that, if (for $(q^1, -p^1) = \Psi_{\Delta t}(q,p)$)

$$\left\| \Delta t M^{-1} \left( p - \frac{\Delta t}{2} \nabla V(q) \right) \right\|, \left\| \Delta t M^{-1} \left( -p^1 - \frac{\Delta t}{2} \nabla V(q^1) \right) \right\| < \beta,$$

then $(q,p)$ and $(q^1, -p^1)$ belong to $A_{\mathrm{imp}}$ and $(\Psi_{\Delta t} \circ \Psi_{\Delta t})(q,p) = (q,p)$

# A more practical realization

• Use Newton iterations to bring the unconstrained move sufficiently close to the submanifold, and rely then on the "theoretical" projection provided by the implicit function theorem

Iterate on $n = 0, \dots, N_{\rm newt}$,

(1) If $[\nabla\xi(\tilde{q} + M^{-1}\nabla\xi(q)\theta^n)]^T M^{-1}\nabla\xi(q)$ is not invertible then set $(q, \tilde{q}) \notin \mathcal{D}_{\rm newt}$ and exit the loop;

(2) Otherwise, $\theta^{n+1} = \theta^n - \left([\nabla\xi(\tilde{q} + M^{-1}\nabla\xi(q)\theta^n)]^T M^{-1}\nabla\xi(q)\right)^{-1} \xi\left(\tilde{q} + M^{-1}\nabla\xi(q)\theta^n\right)$

If these iterations are successful, set $\widehat{q} = \tilde{q} + M^{-1}\nabla\xi(q)\theta^{N_{\rm newt}}$:

(3) If $(q, \hat{q}) \notin \mathcal{D}_{\rm imp}$ then set $(q, \tilde{q}) \notin \mathcal{D}_{\rm newt}$ and exit.

(4) Otherwise, set $(q, \tilde{q}) \in \mathcal{D}_{\rm newt}$ and $\Lambda_{\rm newt}(q, \tilde{q}) = \Lambda(q, \widehat{q})$

The function $\Lambda_{\rm newt}$ defines an admissible Lagrange multiplier function on the open set $\mathcal{D}_{\rm newt}$

# Generalized HMC schemes

# Constrained Gibbs measures (2)

- Phase space Liouville measure $\sigma_{T^*\mathcal{M}}(dq\,dp) = \sigma_{\mathcal{M}}^{M}(dq)\,\sigma_{T_q^*\mathcal{M}}^{M^{-1}}(dp)$ with $\sigma_{\mathcal{M}}^{M}(dq)$ Riemannian measure on $\mathcal{M}$ induced by scalar product $\langle \cdot, \cdot \rangle_M$ on $\mathbb{R}^d$ (similarly for $\sigma_{T_q^*\mathcal{M}}^{M^{-1}}(dp)$)

**Target measure to sample (independent of $M$)**

$$\mu(dq\,dp) = Z_\mu^{-1} \mathrm{e}^{-H(q,p)}\,\sigma_{T^*\mathcal{M}}(dq\,dp) = \nu(dq)\,\kappa_q(dp),$$

with $\kappa_q$ Gaussian and $\nu(dq) = Z_\nu^{-1}\,\mathrm{e}^{-V(q)}\,\sigma_{\mathcal{M}}^{M}(dq)$

- Coarea (conditioning): $\delta_{\xi(q)}(dq) = (\det M)^{-1/2}\,|\det G_M(q)|^{-1/2}\,\sigma_{\mathcal{M}}^{M}(dq)$

# Sampling the constrained Gibbs measure

- **Algorithm:** Starting from $(q^n, p^n) \in T^*\mathcal{M}$,

  (i) Evolve the momenta according to the mid-point Euler scheme

$$\begin{cases} p^{n+1/4} = p^n - \dfrac{\Delta t}{2}\,\gamma\, M^{-1}\left(p^n + p^{n+1/4}\right) + \sqrt{2\gamma\Delta t}\, G^n + \nabla\xi(q^n)\,\lambda^{n+1/4}\,, \\ \left[\nabla\xi(q^n)\right]^T M^{-1} p^{n+1/4} = 0\,. \end{cases}$$

  (ii) Evolve with reversible RATTLE: $(\tilde{q}^{n+1}, \tilde{p}^{n+3/4}) = \Psi^{\mathrm{rev}}_{\Delta t}(q^n, p^{n+1/4})$

  (iii) Draw a random variable $U^n$ with uniform law on $(0,1)$:
    - if $U^n \leqslant \exp(-H(\tilde{q}^{n+1}, \tilde{p}^{n+3/4}) + H(q^n, p^{n+1/4}))$, accept the proposal
    - else reject the proposal: $(q^{n+1}, p^{n+3/4}) = (q^n, p^{n+1/4})$.

  (iv) Reverse momenta $p^{n+1} = -p^{n+3/4}$.

- Preserves $\mu$ by construction

# A simple numerical example

# Three-dimensional system, one-dimensional constraint

- $q = (x, y, z) \in \mathbb{R}^3$ and $\xi(q) = \left(R - \sqrt{x^2 + y^2}\right)^2 + z^2 - r^2$

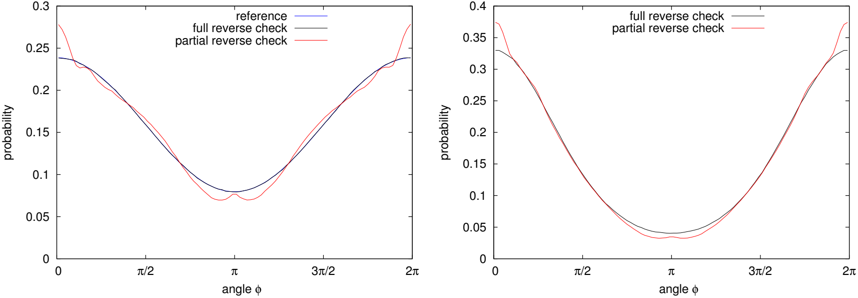
- GHMC, analytical integration of momenta (with $\alpha = \mathrm{e}^{-\gamma \Delta t}$)

$$p^{n+1/4} = P(q^n) \left[\alpha p^n + \sqrt{1 - \alpha^2}\, G^n\right], \quad P(q) = \mathrm{Id} - \frac{\nabla \xi(q) \otimes \nabla \xi(q)}{|\nabla \xi(q)|^2}$$

- Potential $V(q) = k|q|^2/2$

- Partial reverse check: $\Psi_{\Delta t} \circ \Psi_{\Delta \mathrm{t}}$ is well defined but do not check whether $\Psi_{\Delta t} \circ \Psi_{\Delta \mathrm{t}}(q, p) = (q, p)$

# The need for reversibility checks

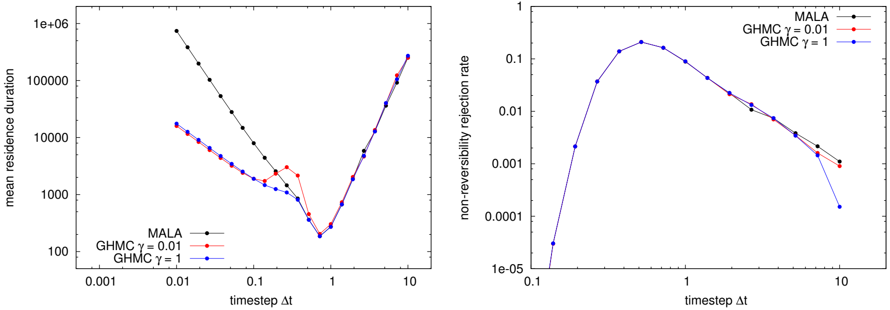



Histograms of the sampled angles $\phi$ with the GHMC scheme, with full or partial reverse projection check, for $\Delta t = 1$. Left: $k = 0$. Right: $k = 1$.

# Analysis of the rejection rate

| Method | Total | Newton | Newton rev. | non-rev. | Metropolis |
|---|---|---|---|---|---|
| MRW $\Delta t = 1$ | 0.675 | 0.562 | $3.02 \cdot 10^{-4}$ | 0.0742 | 0.0385 |
| MALA $\Delta t = 1$ | 0.675 | 0.509 | $5.83 \cdot 10^{-4}$ | 0.149 | 0.0167 |
| GHMC $\Delta t = 1$, $\alpha = 0.1$ | 0.675 | 0.509 | $5.83 \cdot 10^{-4}$ | 0.149 | 0.0167 |
| GHMC $\Delta t = 1$, $\alpha = 0.5$ | 0.675 | 0.509 | $5.83 \cdot 10^{-4}$ | 0.149 | 0.0167 |
| GHMC $\Delta t = 1$, $\alpha = 0.9$ | 0.675 | 0.509 | $5.83 \cdot 10^{-4}$ | 0.149 | 0.0167 |
| MRW $\Delta t = 0.3$ | 0.158 | 0.0803 | $1.06 \cdot 10^{-4}$ | 0.0127 | 0.0652 |
| MALA $\Delta t = 0.3$ | 0.107 | 0.0763 | $1.22 \cdot 10^{-4}$ | 0.0138 | 0.0168 |
| GHMC $\Delta t = 0.3$, $\alpha = 0.1$ | 0.107 | 0.0763 | $1.22 \cdot 10^{-4}$ | 0.0138 | 0.0168 |
| GHMC $\Delta t = 0.3$, $\alpha = 0.5$ | 0.107 | 0.0763 | $1.22 \cdot 10^{-4}$ | 0.0138 | 0.0168 |
| GHMC $\Delta t = 0.3$, $\alpha = 0.9$ | 0.107 | 0.0763 | $1.22 \cdot 10^{-4}$ | 0.0138 | 0.0168 |
| MRW $\Delta t = 0.1$ | 0.0259 | $5 \cdot 10^{-7}$ | 0 | $7 \cdot 10^{-8}$ | 0.0259 |
| MALA $\Delta t = 0.1$ | $6.73 \cdot 10^{-4}$ | $5 \cdot 10^{-7}$ | $10^{-9}$ | $5 \cdot 10^{-8}$ | $6.73 \cdot 10^{-4}$ |
| GHMC $\Delta t = 0.1$, $\alpha = 0.1$ | $6.72 \cdot 10^{-4}$ | $5 \cdot 10^{-7}$ | $10^{-9}$ | $6 \cdot 10^{-8}$ | $6.72 \cdot 10^{-4}$ |
| GHMC $\Delta t = 0.1$, $\alpha = 0.5$ | $6.73 \cdot 10^{-4}$ | $5 \cdot 10^{-7}$ | $2 \cdot 10^{-9}$ | $8 \cdot 10^{-8}$ | $6.72 \cdot 10^{-4}$ |
| GHMC $\Delta t = 0.1$, $\alpha = 0.9$ | $6.74 \cdot 10^{-4}$ | $5 \cdot 10^{-7}$ | 0 | $7 \cdot 10^{-8}$ | $6.73 \cdot 10^{-4}$ |

# Metastability analysis for a double-well potential




Left: mean residence duration as a function of the timestep. Right: non-reversibility rejection rate

Maximal (and non negligible) non reversibility rejection rate at the optimal timestep!