Path integral molecular dynamics (with surface hopping)

Jianfeng Lu 鲁剑锋

Duke University
jianfeng@math.duke.edu

Casa Matemática Oaxaca, November 2018

Based on joint works with
Zhennan Zhou (BICMR, Peking University)

Duke
UNIVERSITY

Quantities of interests for nuclear quantum effects

- Time propagation:
$$\langle u | e^{iHt} A e^{-iHt} | u \rangle \, ;$$
- Thermal equilibrium average:
$$\frac{\mathrm{tr}\left(e^{-\beta H} A\right)}{\mathrm{tr}\, e^{-\beta H}} \, ;$$
- Dynamical correlation function:
$$\frac{\mathrm{tr}\left(e^{-\beta H} A e^{-iHt} B e^{iHt}\right)}{\mathrm{tr}\, e^{-\beta H}} \, ,$$

where $H = -\frac{1}{2}\Delta + V$ is a Hamiltonian, A and B are observables, u is some wave function, β is inverse temperature and t is time.

Difficulty: Curse of dimensionality.

R. Feynman: Turn this into a sampling problem of a classical system.

Classic example of cross-fertilization → Feynman-Kac rep.

Lie-Trotter / Strang splitting for $H = -\frac{1}{2}\Delta + V$:

$$\operatorname{tr} e^{-\beta H} \approx \operatorname{tr}\left(\exp\left(\frac{\beta}{2N}\Delta\right)\exp\left(-\frac{\beta}{N}V\right)\cdots\exp\left(\frac{\beta}{2N}\Delta\right)\exp\left(-\frac{\beta}{N}V\right)\right)$$

R. Feynman: Turn this into a sampling problem of a classical system.

Classic example of cross-fertilization → Feynman-Kac rep.

Lie-Trotter / Strang splitting for $H = -\frac{1}{2}\Delta + V$:

$$\begin{aligned}
\operatorname{tr} e^{-\beta H} &\approx \operatorname{tr}\left(\exp\left(\frac{\beta}{2N}\Delta\right)\exp\left(-\frac{\beta}{N}V\right)\cdots\exp\left(\frac{\beta}{2N}\Delta\right)\exp\left(-\frac{\beta}{N}V\right)\right) \\
&= \int\cdots\int \langle q_1|e^{\frac{\beta}{2N}\Delta}|p_N\rangle\langle p_N|e^{-\frac{\beta}{N}V}|q_N\rangle\times\cdots \\
&\qquad\qquad \times\langle q_2|e^{\frac{\beta}{2N}\Delta}|p_1\rangle\langle p_1|e^{-\frac{\beta}{N}V}|q_1\rangle\,\mathrm{d}p\,\mathrm{d}q \\
&= \int\cdots\int \langle q_1|p_N\rangle e^{-\frac{\beta}{2N}|p_N|^2}\langle p_N|q_N\rangle e^{-\frac{\beta}{N}V(q_N)}\times\cdots \\
&\qquad\qquad \times\langle q_2|p_1\rangle e^{-\frac{\beta}{2N}|p_1|^2}\langle p_1|q_1\rangle e^{-\frac{\beta}{N}V(q_1)}\,\mathrm{d}p\,\mathrm{d}q
\end{aligned}$$

R. Feynman: Turn this into a sampling problem of a classical system.

Classic example of cross-fertilization → Feynman-Kac rep.

Lie-Trotter / Strang splitting for $H = -\frac{1}{2}\Delta + V$:

$$\begin{aligned}
\operatorname{tr} e^{-\beta H} &\approx \operatorname{tr}\left(\exp\left(\frac{\beta}{2N}\Delta\right)\exp\left(-\frac{\beta}{N}V\right)\cdots\exp\left(\frac{\beta}{2N}\Delta\right)\exp\left(-\frac{\beta}{N}V\right)\right) \\
&= \int\cdots\int \langle q_1|e^{\frac{\beta}{2N}\Delta}|p_N\rangle\langle p_N|e^{-\frac{\beta}{N}V}|q_N\rangle \times \cdots \\
&\qquad\qquad \times \langle q_2|e^{\frac{\beta}{2N}\Delta}|p_1\rangle\langle p_1|e^{-\frac{\beta}{N}V}|q_1\rangle \,\mathrm{d}p\,\mathrm{d}q \\
&= \int\cdots\int \langle q_1|p_N\rangle e^{-\frac{\beta}{2N}|p_N|^2}\langle p_N|q_N\rangle e^{-\frac{\beta}{N}V(q_N)} \times \cdots \\
&\qquad\qquad \times \langle q_2|p_1\rangle e^{-\frac{\beta}{2N}|p_1|^2}\langle p_1|q_1\rangle e^{-\frac{\beta}{N}V(q_1)} \,\mathrm{d}p\,\mathrm{d}q \\
&\propto \int\cdots\int \exp\left(-\sum_{i=1}^{N}\frac{|q_i - q_{i+1}|^2}{2\beta_N} - \beta_N\sum_{i=1}^{N}V(q_i)\right)\mathrm{d}q,
\end{aligned}$$

with $\beta_N = \frac{\beta}{N}$ and the convention $q_{N+1} = q_1$.

Ring-polymer representation:

$$\operatorname{tr} e^{-\beta H} \propto \int \cdots \int \exp\left(- \sum_{i=1}^{N} \frac{|q_i - q_{i+1}|^2}{2\beta_N} - \beta_N \sum_{i=1}^{N} V(q_i) \right) \mathrm{d}q.$$

Sampling problem on a high-dimensional energy landscape.

- Path-integral Monte Carlo
- Path-integral Molecular Dynamics
- Extension to dynamical quantities (ring polymer MD, centroid MD, Matsubara dynamics, Liouville dynamics, etc.)

Path-integral molecular dynamics (PIMD): Re-introducing momentum

$$\operatorname{tr} e^{-\beta H} \propto \int \cdots \int \exp\left(-\beta_N \sum_{i=1}^{N} \frac{|p_i|^2}{2} - \sum_{i=1}^{N} \frac{|q_i - q_{i+1}|^2}{2\beta_N} - \beta_N \sum_{i=1}^{N} V(q_i)\right) \mathrm{d}p\, \mathrm{d}q$$

$$=: \int \cdots \int e^{-\beta_N H_N(p,q)}\, \mathrm{d}p\, \mathrm{d}q$$

with

$$H_N(p,q) = \sum_{i=1}^{N} \frac{|p_i|^2}{2} + \sum_{i=1}^{N} \frac{|q_i - q_{i+1}|^2}{2\beta_N^2} + \sum_{i=1}^{N} V(q_i).$$

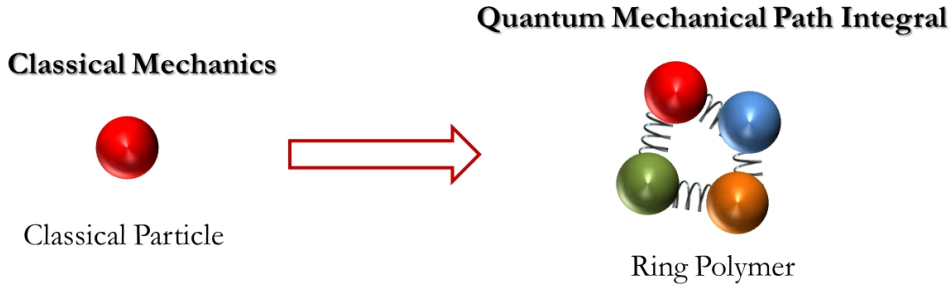


Figure: Image credit: Tom Markland @Standford

Taking $N \to \infty$, we obtain the path-integral representation as a continuum limit of the ring polymers

$$\begin{aligned} \operatorname{tr} e^{-\beta H} &\propto \int \mathrm{D}[\mathfrak{q}] \exp\left(- \int_0^\beta \frac{|\dot{\mathfrak{q}}(\tau)|^2}{2} + V(\mathfrak{q}(\tau)) \, \mathrm{d}\tau \right) \\ &=: \int \mathrm{D}[\mathfrak{q}] \exp\left(-\left(\tfrac{1}{2} \langle \mathfrak{q}, L\mathfrak{q} \rangle + V(\mathfrak{q}) \right) \right) \end{aligned}$$

where the path $\mathfrak{q} : [0, \beta] \to \mathbb{R}^d$ takes periodic boundary conditions.

Cross-Fertilization: It is quite similar to diffusion bridge sampling except for the boundary condition.

Taking $N \to \infty$, we obtain the path-integral representation as a continuum limit of the ring polymers

$$\operatorname{tr} e^{-\beta H} \propto \int \mathrm{D}[\mathfrak{q}] \exp\Big(- \int_0^\beta \frac{|\dot{\mathfrak{q}}(\tau)|^2}{2} + V(\mathfrak{q}(\tau)) \, \mathrm{d}\tau \Big)$$
$$=: \int \mathrm{D}[\mathfrak{q}] \exp\Big(-\big(\tfrac{1}{2} \langle \mathfrak{q}, L\mathfrak{q} \rangle + V(\mathfrak{q}) \big) \Big)$$

where the path $\mathfrak{q} : [0, \beta] \to \mathbb{R}^d$ takes periodic boundary conditions.

For the auxiliary momentum, using kinetic energy with "diagonal mass" does not make sense in the continuum limit. Instead, it is better to use (similar to [Beskos, Pinski, Sanz-Serna, Stuart 2011] for diffusion bridges)

$$\operatorname{tr} e^{-\beta H} = \int \mathrm{D}[(\mathfrak{q}, \mathfrak{p})] \exp\Big(-\big(\tfrac{1}{2} \langle \mathfrak{q}, L\mathfrak{q} \rangle + \tfrac{1}{2} \langle \mathfrak{p}, (L^\alpha)^{-1} \mathfrak{p} \rangle + V(\mathfrak{q}) \big) \Big)$$

where $L^\alpha = L + \alpha \mathrm{Id}$ to avoid the singularity due to the periodic Laplacian.

Regularizing the position part as well, we get

$$\begin{aligned}\operatorname{tr} e^{-\beta H} &= \int \mathrm{D}[(\mathfrak{q},\mathfrak{p})] \exp\Big(-\big(\tfrac{1}{2}\langle \mathfrak{q}, L\mathfrak{q}\rangle + \tfrac{1}{2}\langle \mathfrak{p}, (L^\alpha)^{-1}\mathfrak{p}\rangle + V(\mathfrak{q})\big)\Big)\\ &= \int \mathrm{D}[(\mathfrak{q},\mathfrak{p})] \exp\Big(-\big(\tfrac{1}{2}\langle \mathfrak{q}, L^\alpha\mathfrak{q}\rangle + \tfrac{1}{2}\langle \mathfrak{p}, (L^\alpha)^{-1}\mathfrak{p}\rangle + U^\alpha(\mathfrak{q})\big)\Big)\end{aligned}$$

with $U^\alpha(q) = V(q) - \alpha|q|^2$ (assumed to be growing at infinity).

The associate (inf. dim.) (underdamped) Langevin dynamics reads

$$\begin{aligned}\mathrm{d}\mathfrak{q} &= (L^\alpha)^{-1}\mathfrak{p}\,\mathrm{d}t;\\ \mathrm{d}\mathfrak{p} &= -L^\alpha\mathfrak{q}\,\mathrm{d}t - \nabla U^\alpha\,\mathrm{d}t - \gamma\mathfrak{p}\,\mathrm{d}t + \sqrt{2\gamma L^\alpha}\,\mathrm{d}\omega.\end{aligned}$$

Langevin dynamics

$$\mathrm{d}\mathfrak{q} = (L^\alpha)^{-1}\mathfrak{p}\,\mathrm{d}t;$$
$$\mathrm{d}\mathfrak{p} = -L^\alpha \mathfrak{q}\,\mathrm{d}t - \nabla U^\alpha\,\mathrm{d}t - \gamma\mathfrak{p}\,\mathrm{d}t + \sqrt{2\gamma L^\alpha}\,\mathrm{d}\omega.$$

The dynamics is rather stiff due to L^α, causing trouble for sampling. Introducing $\mathfrak{v} = (L^\alpha)^{-1}\mathfrak{p}$ and preconditioning, we arrive at

$$\mathrm{d}\mathfrak{q} = \mathfrak{v}\,\mathrm{d}t;$$
$$\mathrm{d}\mathfrak{v} = -\mathfrak{q}\,\mathrm{d}t - (L^\alpha)^{-1}\nabla U^\alpha\,\mathrm{d}t - \gamma\mathfrak{v}\,\mathrm{d}t + \sqrt{2\gamma(L^\alpha)^{-1}}\,\mathrm{d}\omega.$$

This leads to a much better sampling scheme when the number of bead is large (in preparation with Zhennan Zhou).

Question: Convergence rate?
(ongoing project with Yulong Lu and Jonathan Mattingly)

In the above, we have assumed that $H = T + V$, where V is just a single potential surface for the nuclei degree of freedom. This is known as the adiabatic approximation or Born-Oppenheimer approximation.

The adiabatic approximation is justified if $E_0(q)$ (ground state) as a potential energy surface is well separated from $E_1(q)$ (excited state), and hence *the transition of electrons to excited states can be neglected*.

This is however often *not* the case, for applications like photoexcited dynamics, electron transfer and surface chemistry, where the Born-Oppenheimer approximation falls apart.

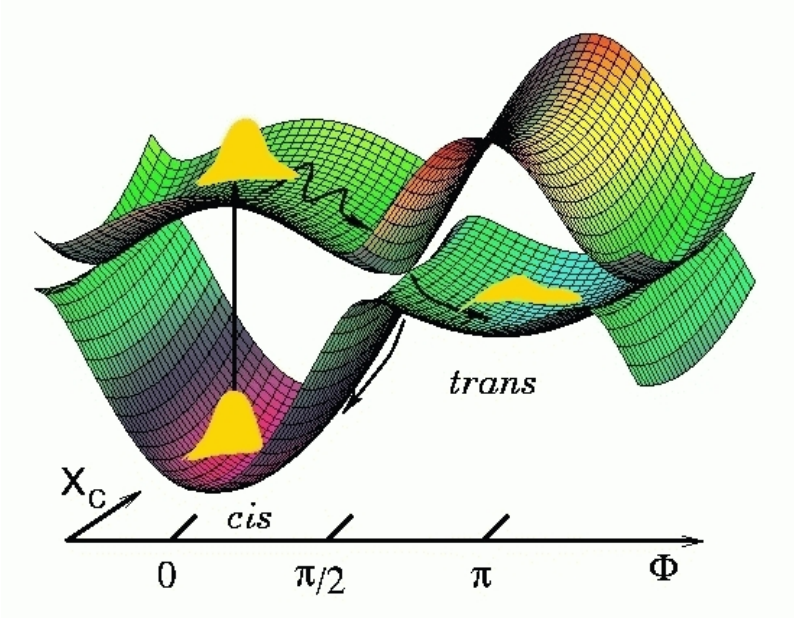


Figure: Schematic of ultrafast non-adiabatic photoreaction.

This thus leads us to the non-adiabatic scenario, for which, under the diabatic representation, the Hamiltonian takes

$$H = -\frac{1}{2}\begin{pmatrix} \Delta & \\ & \Delta \end{pmatrix} + \begin{pmatrix} V_{00} & V_{01} \\ V_{10} & V_{11} \end{pmatrix}$$

where a two-level system is assumed for simplicity, generalizations to multi-level is possible.

Thermal equilibrium sampling for multi-level systems

$$\langle A \rangle = \frac{\mathrm{tr}_{\mathrm{ne}}(e^{-\beta H} A)}{\mathrm{tr}_{\mathrm{ne}}(e^{-\beta H})},$$

where

$$\mathrm{tr}_{\mathrm{ne}} = \mathrm{tr}_{L^2(\mathbb{R}^d)}\, \mathrm{tr}_{\mathbb{C}^2}\,.$$

Question: How to extend the PIMD sampling approach to multi-level systems?

Previous works on non-adiabatic ring-polymer representation:

- Meyer-Miller-Stock-Thoss mapping variable approach, by mapping the discrete electronic degree freedom to continuous variables using uncoupled harmonic oscillators. Then just apply what we did before.

  In particular, recent works by the groups of Ananth and T. Miller on PIMD based on mapping variable approaches.

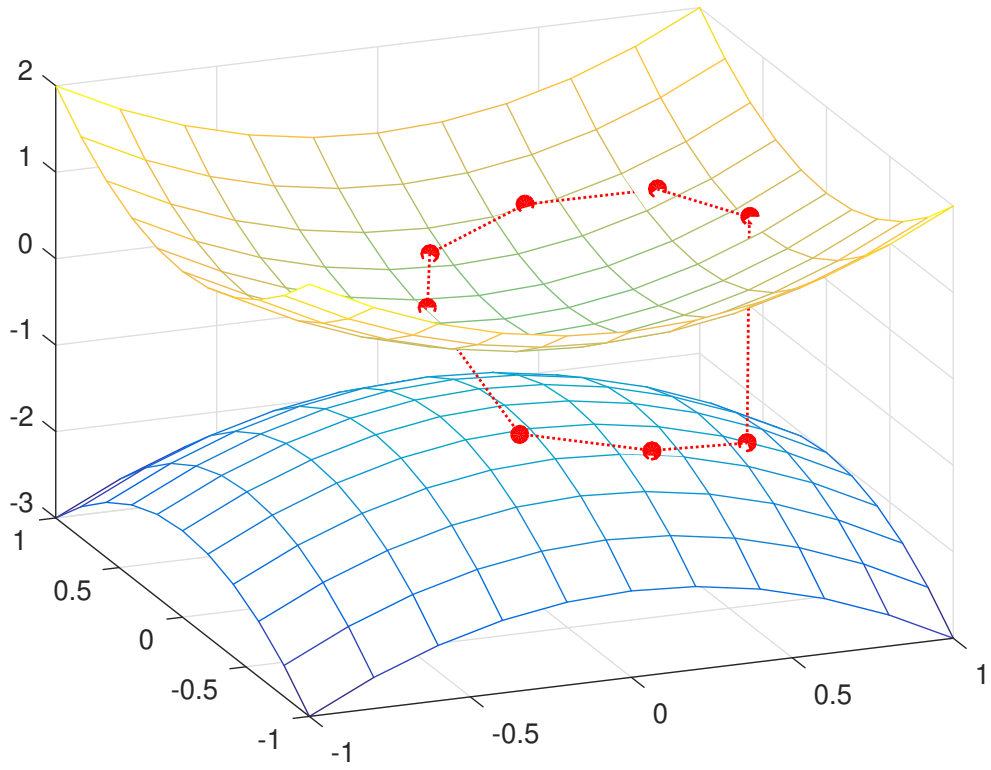
- Design a ring-polymer representation that keeps the discreteness of the electronic levels.

  A much less traveled road. Only previous work we are aware of is [Schmidt, Tully 2007], which used Monte Carlo to sample the non-adiabatic ring polymer configuration.

  We will take this route, as it is more advantageous when combining with the dynamics.

Idea: Further extend the phase space to encode the discrete level.

$\tilde{z} = (p, q, \ell) \in \mathbb{R}^{dN} \times \mathbb{R}^{dN} \times \{0, 1\}^N$, so now each bead (p_i, q_i, ℓ_i) lives on two copies of the classical phase space, with $\ell_i \in \{0, 1\}$ a level index.



Thus, phase space $L^2(\mathbb{R}^d) \otimes \mathbb{C}^2 \to \mathbb{R}^{dN} \times \mathbb{R}^{dN} \times \{0, 1\}^N$.

From adiabatic to non-adiabatic ring polymer represenation.

Strang / Trotter splitting can be still used:

$$\operatorname{tr} e^{-\beta H} \approx \operatorname{tr}\left(\exp\left(-\frac{\beta}{N}T\right)\exp\left(-\frac{\beta}{N}V\right)\cdots\exp\left(-\frac{\beta}{N}T\right)\exp\left(-\frac{\beta}{N}V\right)\right)$$

where T and V are matrices of operators:

$$T = -\frac{1}{2}\begin{pmatrix}\Delta & \\ & \Delta\end{pmatrix} \qquad \text{and} \qquad V = \begin{pmatrix}V_{00} & V_{01}\\ V_{10} & V_{11}\end{pmatrix}$$

and continue the process of inserting resolution of identities, but at this time, we will also need to inserting ℓ to track the matrix components.

From adiabatic to non-adiabatic ring polymer represenation.

Partition function (in a simpler case; ratios of time average estimator needed in general):

$$\operatorname{tr} e^{-\beta H} \approx \iint e^{-\beta_N H_N(p,q)} \,\mathrm{d}p \,\mathrm{d}q$$

$$\operatorname{tr}_{\mathrm{ne}} e^{-\beta H} \approx \iint \sum_{\ell \in \{0,1\}^N} e^{-\beta_N H_N(p,q,\ell)} \,\mathrm{d}p \,\mathrm{d}q$$

$$\operatorname{tr}(A e^{-\beta H}) \approx \iint e^{-\beta_N H_N(p,q)} W_N[A](p,q) \,\mathrm{d}p \,\mathrm{d}q$$

$$\operatorname{tr}_{\mathrm{ne}}(A e^{-\beta H}) \approx \iint \sum_{\ell \in \{0,1\}^N} e^{-\beta_N H_N(p,q,\ell)} W_N[A](p,q,\ell) \,\mathrm{d}p \,\mathrm{d}q$$

The expression of H_N and $W_N[A]$ is complicated; but follows from straightforward calculations.
We will skip the details as our main focus today is sampling.

Goal: Sample from $\varrho(p,q,\ell) \propto e^{-\beta_N H_N(p,q,\ell)}$ on $\mathbb{R}^{dN} \times \mathbb{R}^{dN} \times \{0,1\}^N$.

The evolution of (p,q) follows the Hamiltonian dynamics with a Langevin thermostat

$$\begin{cases} \mathrm{d}q = \nabla_p H_N(p,q,\ell)\,\mathrm{d}t \\ \mathrm{d}p = -\nabla_q H_N(p,q,\ell)\,\mathrm{d}t - \gamma p\,\mathrm{d}t + \sqrt{2\gamma\beta_N^{-1}}\,\mathrm{d}W \end{cases}$$

The evolution of ℓ follows a surface hopping type dynamics as a Markov jump process with $(z = (p,q))$

$$\mathbb{P}\big(\ell(t+\delta t) = \ell' \mid \ell(t) = \ell, z(t) = z\big) = \delta_{\ell'\ell} + \eta\lambda_{\ell'\ell}(z)\delta t + o(\delta t),$$

where $z = (p,q)$ and η is an overall hopping intensity scaling parameter.

Surface hopping dynamics

$$\mathbb{P}\Big(\ell(t+\delta t)=\ell' \mid \ell(t)=\ell, z(t)=z\Big)=\delta_{\ell'\ell}+\eta\lambda_{\ell'\ell}(z)\delta t+o(\delta t).$$

To guarantee that the dynamics samples the desired distribution, we may choose $\lambda_{\ell'\ell}$ such that the detailed balance condition is satisfied:

$$\lambda_{\ell'\ell}(z)e^{-\beta_N H_N(z,\ell)}=\lambda_{\ell\ell'}(z)e^{-\beta_N H_N(z,\ell')}.$$

For instance,

$$p_{\ell'\ell}(z)=A_{\ell'\ell}\exp\Big(\frac{\beta_N}{2}\big(H_N(z,\ell)-H_N(z,\ell')\big)\Big)$$

with $A_{\ell'\ell}$ a symmetric "accessibility matrix" (e.g., only allow one level change of the beads, etc.).

PIMD with surface hopping (PIMDSH):

$$\begin{cases} \mathrm{d}q = \nabla_p H_N(p,q,\ell)\,\mathrm{d}t \\ \mathrm{d}p = -\nabla_q H_N(p,q,\ell)\,\mathrm{d}t - \gamma p\,\mathrm{d}t + \sqrt{2\gamma\beta_N^{-1}}\,\mathrm{d}W \\ \mathbb{P}\big(\ell(t+\delta t) = \ell' \mid \ell(t) = \ell, z(t)\big) = \delta_{\ell'\ell} + \eta\lambda_{\ell'\ell}(z)\delta t + o(\delta t). \end{cases}$$

Cross-Fertilization: Anything similar in Bayesian statistics?

PIMD with surface hopping (PIMDSH):

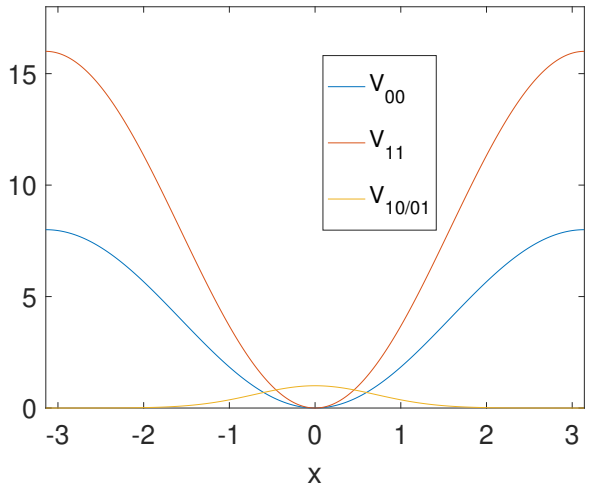
$$\begin{cases} \mathrm{d}q = \nabla_p H_N(p,q,\ell)\,\mathrm{d}t \\ \mathrm{d}p = -\nabla_q H_N(p,q,\ell)\,\mathrm{d}t - \gamma p\,\mathrm{d}t + \sqrt{2\gamma\beta_N^{-1}}\,\mathrm{d}W \\ \mathbb{P}\big(\ell(t+\delta t) = \ell' \mid \ell(t) = \ell, z(t)\big) = \delta_{\ell'\ell} + \eta\lambda_{\ell'\ell}(z)\delta t + o(\delta t). \end{cases}$$

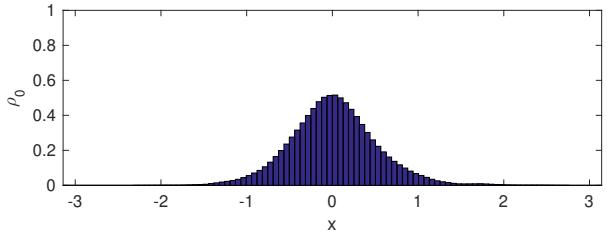
Numerical discretization (2nd weak order):

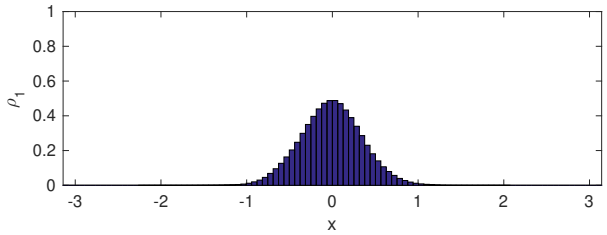
- Operator splitting scheme between (q,p) and ℓ;
- BAOAB splitting for Langevin [Leimkuhler, Matthews 2013] (see also [Liu, Li, Liu 2016] for BAOAB type scheme for PIMD);

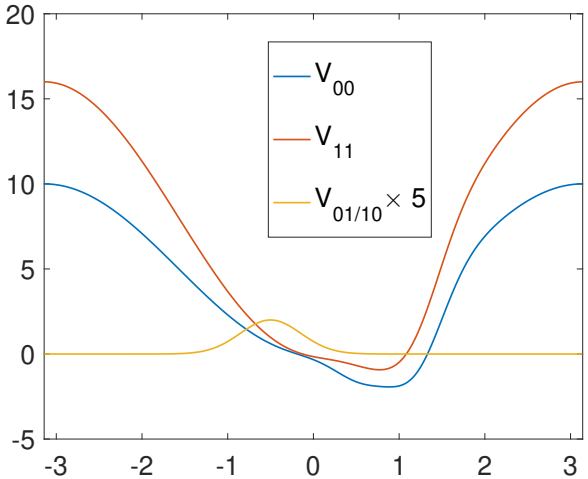
$$(\mathrm{B})\begin{cases} \mathrm{d}q = 0 \\ \mathrm{d}p = -\nabla_q H_N\,\mathrm{d}t \end{cases} (\mathrm{A})\begin{cases} \mathrm{d}q = \nabla_p H_N\,\mathrm{d}t \\ \mathrm{d}p = 0 \end{cases} (\mathrm{O})\begin{cases} \mathrm{d}q = 0 \\ \mathrm{d}p = -\gamma p\,\mathrm{d}t + \sqrt{2\gamma\beta_N^{-1}}\,\mathrm{d}W \end{cases}$$

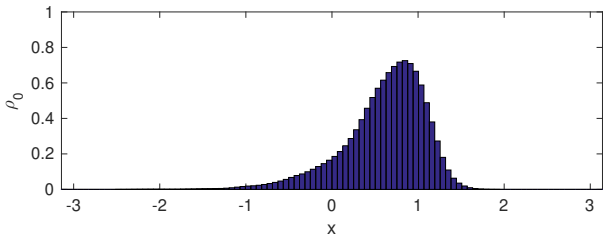
Numerical tests: two cases in the diabatic representation

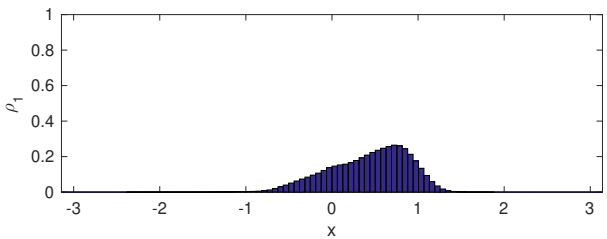








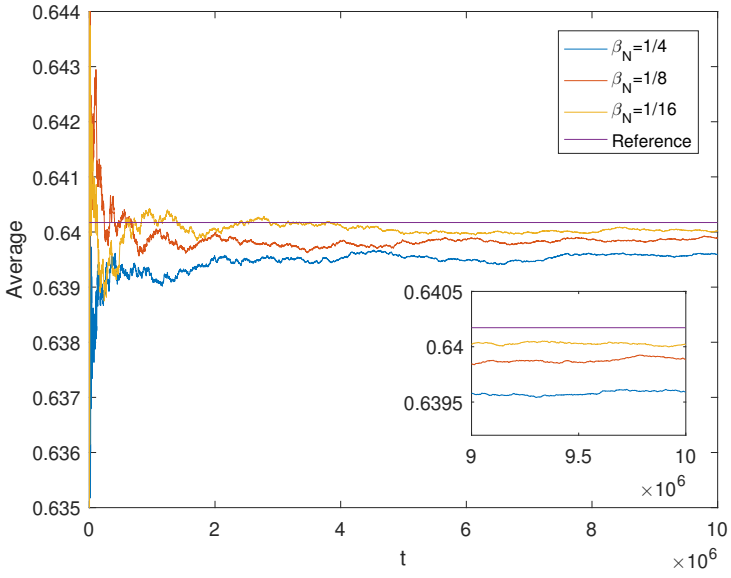




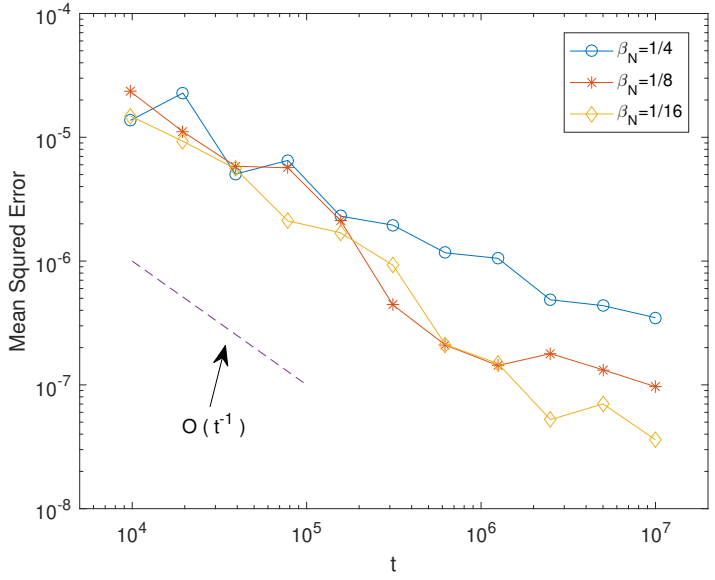
Also consider two observables

$$A_{\text{diag}} = \begin{pmatrix} e^{-q^2} & 0 \\ 0 & e^{-q^2} \end{pmatrix} \text{(diagonal)}, \quad A_{\text{offd}} = \begin{pmatrix} 0 & e^{-q^2} \\ e^{-q^2} & 0 \end{pmatrix} \text{(off-diagonal)}$$

Convergence with number of beads for diagonal observable:
For $\beta = 1$ and $\beta_N = \frac{1}{4}$, $\frac{1}{8}$ and $\frac{1}{16}$.

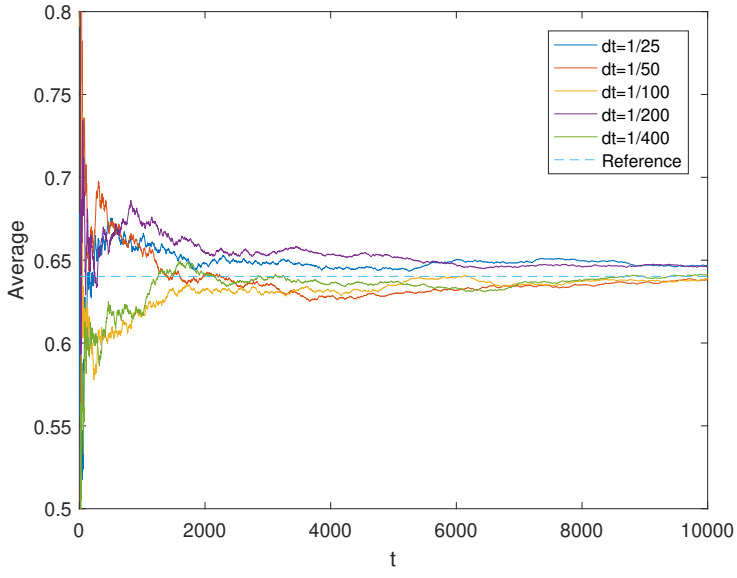


(A) Running average

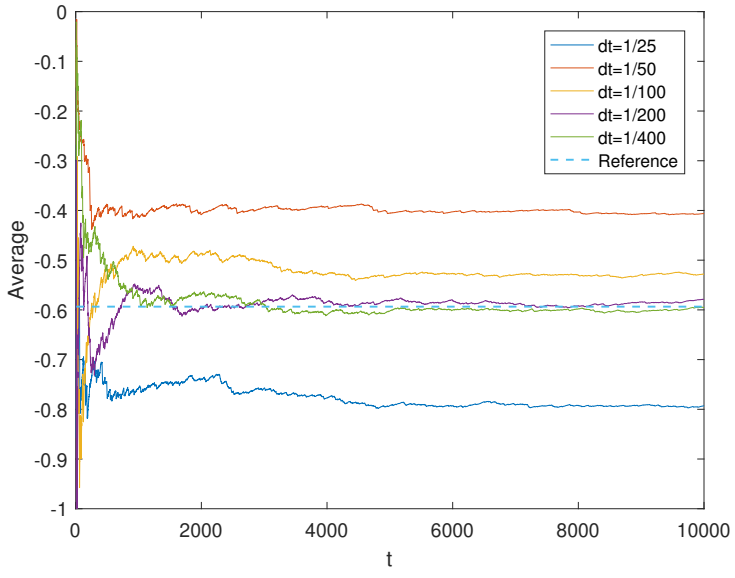


(B) Convergence rate

Convergence with different Δt:



(A) diagonal observable



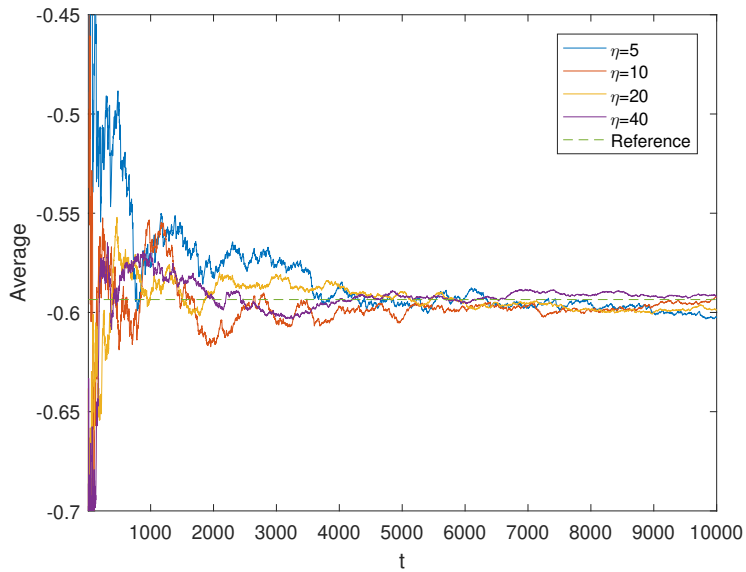
(B) off-diagonal observable

Why much worse performance in the off-diagonal case?

Let us increase the hopping intensity parameter η. Recall

$$\mathbb{P}\big(\ell(t+\delta t)=\ell' \mid \ell(t)=\ell, z(t)\big)=\delta_{\ell'\ell}+\eta\lambda_{\ell'\ell}(z)\delta t+o(\delta t).$$

	$\eta=5$	$\eta=10$	$\eta=20$	$\eta=40$
Error	8.27e-3	1.38e-3	4.97e-3	1.44e-3
95% C.I.	1.09e-2	7.67e-3	3.59e-3	1.87e-3
M.S.E.	9.93e-5	1.32e-5	2.80e-5	2.98e-6



For the off-diagonal observable $\left[\begin{smallmatrix} 0 & * \\ * & 0 \end{smallmatrix}\right]$, only kinks (consecutive beads on different levels) matter. Those can be only sampled efficiently if $\eta \gg 1$.

This suggests to increase η in the PIMDSH dynamics

$$\begin{cases} \mathrm{d}q = \nabla_p H_N(p,q,\ell)\,\mathrm{d}t \\ \mathrm{d}p = -\nabla_q H_N(p,q,\ell)\,\mathrm{d}t - \gamma p\,\mathrm{d}t + \sqrt{2\gamma\beta_N^{-1}}\,\mathrm{d}W \\ \mathbb{P}\big(\ell(t+\delta t) = \ell' \mid \ell(t) = \ell, z(t)\big) = \delta_{\ell'\ell} + \eta\lambda_{\ell'\ell}(z)\delta t + o(\delta t). \end{cases}$$

The price to pay though is that large η means stiffer dynamics and hence smaller time steps.

How about directly sending the hopping frequency $\eta \to \infty$?

How about directly sending the hopping frequency $\eta \to \infty$?

The situation is in fact very similar to the case of enhanced tempering scheme in recent years, known as **infinite swapping**:

- Parallel tempering is more efficient with infinite swapping by empirical distribution large deviation theory [Dupuis et al 2012] (earlier numerical evidence by [Sindhikara et al 2008])
- The limiting dynamics can be reformulated as a mixture potential, so enables simplified implementation [Lu, Vanden-Eijnden 2013]
- More efficient implementation based on multiscale integrator [Yu, Lu, Abrams, Vanden-Eijnden 2016] [Lu, Vanden-Eijnden 2017+]
- Similar observation and analysis in the context of irreversible "Langevin" sampler [Rey-Bellet, Spiliopoulous 2015] [Lu, Spiliopoulous 2018+]

- Extensions to simulated tempering
  (see the talk by Anton Martinsson, Thursday afternoon)

Infinite switching limit $\eta \to \infty$

$$\begin{cases} \mathrm{d}q = \nabla_p H_N(p,q,\ell)\,\mathrm{d}t \\ \mathrm{d}p = -\nabla_q H_N(p,q,\ell)\,\mathrm{d}t - \gamma p\,\mathrm{d}t + \sqrt{2\gamma\beta_N^{-1}}\,\mathrm{d}W \\ \mathbb{P}\big(\ell(t+\delta t) = \ell' \mid \ell(t) = \ell, z(t)\big) = \delta_{\ell'\ell} + \eta\lambda_{\ell'\ell}(z)\delta t + o(\delta t). \end{cases}$$

The position and momentum $z(t)$ evolves much more slowly than the surface index $\ell(t)$. Thus, an averaging lemma gives

$$\begin{cases} \mathrm{d}q = p\,\mathrm{d}t; \\ \mathrm{d}p = -\nabla_q \bar{H}_N(q,p)\,\mathrm{d}t - \gamma p\,\mathrm{d}t + \sqrt{2\gamma\beta_N^{-1}}\,\mathrm{d}W. \end{cases}$$

The forcing terms is replaced by its averaged counterpart

$$-\nabla_q \bar{H}_N(q,p) = -\sum_{\ell\in\{0,1\}^N} \nabla_q H_N(q,p,\ell)\pi(\ell \mid q,p).$$

Proposition: The averaged dynamics converges faster to equilibrium and corresponding estimator has lower variance.

The remaining bottleneck is the averaged force

$$-\nabla_q \bar{H}_N(q,p) = -\sum_{\ell \in \{0,1\}^N} \nabla_q H_N(q,p,\ell)\pi(\ell \mid q,p).$$

since explicit summation wrt $\pi(\ell \mid q,p) \propto \exp(-\beta_N H_N(q,p,\ell))$ has exponential cost 2^N.

Similar to the case of parallel tempering in our earlier work, we can use a multiscale integrator *a la* heterogeneous multiscale method (HMM), which consists of a **microsolver**, a **macrosolver** and an **estimator** passing parameters between the two.

- The macrosolver is for the slow averaged Langevin dynamics where the averaged version is replaced by an approximation from the estimator.
- The microsolver is for the fast dynamics of ℓ (surface hopping);
- The microsolver takes R steps before one step of macrosolver; R is a control parameter.

The overall algorithm goes as following for each macro time step k:

1. Microsolver. Evolve ℓ_k via a stochastic simulation algorithm from t_k to $t_{k+1} := t_k + \Delta t$. A path in ℓ: $\ell_{k,1}, \cdots \ell_{k,J}$ associated with $\tau_1, ..., \tau_J$.
2. Estimator. Given the trajectory of ℓ, we estimate the averaged force term by

$$-\nabla_q \bar{H}(z) \approx -\frac{1}{\Delta t} \sum_{j=1}^{J} \nabla_q H_N(z, \ell_{k,j}) \tau_j.$$

   The weighted average is approximated by

$$\widetilde{W}[A](z_k) \approx \frac{1}{\Delta t} \sum_{j=1}^{J} W_N[A](z_k, \ell_{k,j}) \tau_j.$$

3. Macrosolver. Evolve z_k to z_{k+1} using one time-step of size Δt using some integrator for the Langevin equations (BAOAB) with the force term replaced by that obtained from the estimator.

Tests for diagonal and off-diagonal observables with different Δt

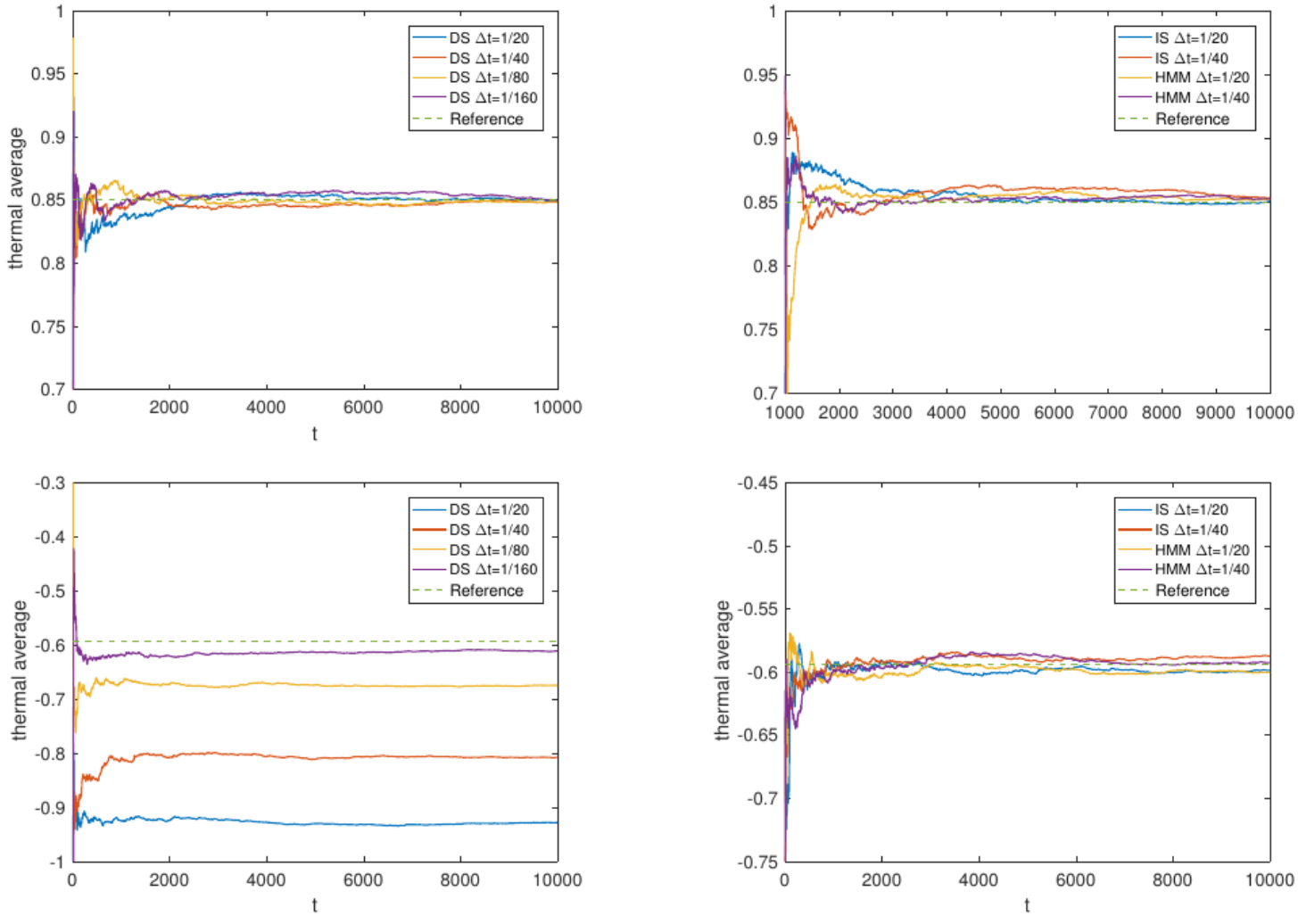
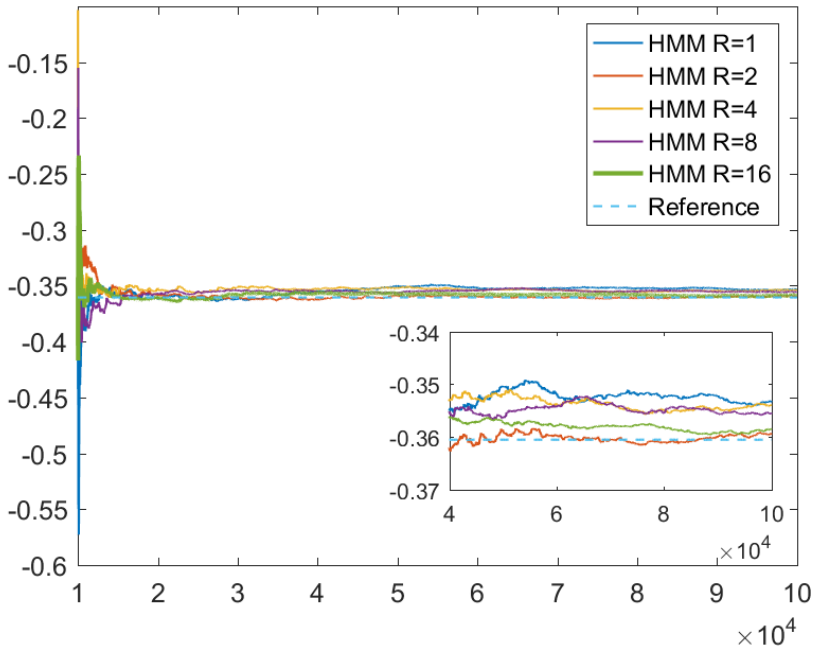


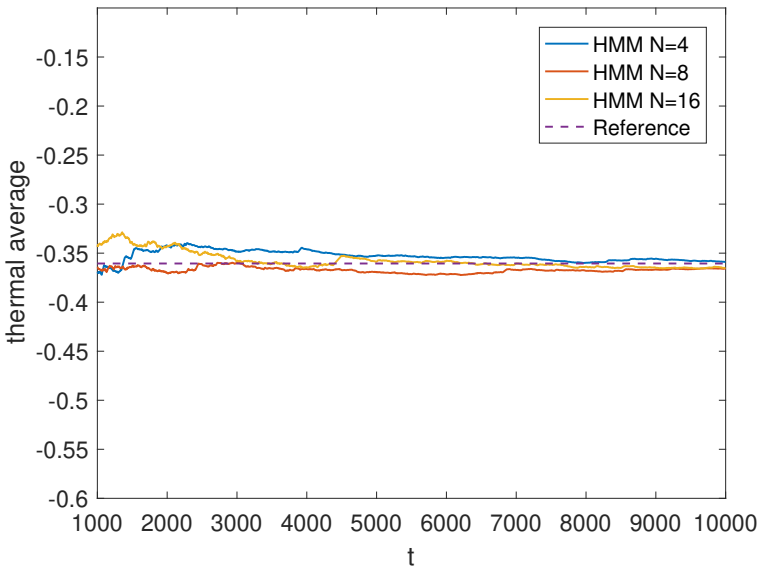
Figure: (left) direct simulation; (right) infinite switching

Tests for off-diagonal observables with different R

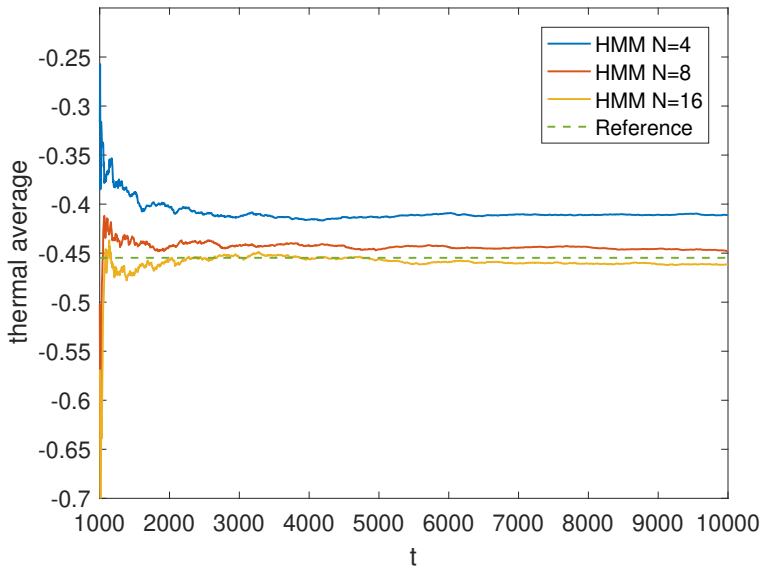


DS	$\Delta t = \frac{1}{20}$	$\Delta t = \frac{1}{40}$	$\Delta t = \frac{1}{80}$	$\Delta t = \frac{1}{160}$
M.S.E.	6.5666e-2	2.6749e-3	9.7038e-4	8.4174e-4
HMM	$\frac{\Delta t}{R} = \frac{1}{20}$	$\frac{\Delta t}{R} = \frac{1}{40}$	$\frac{\Delta t}{R} = \frac{1}{80}$	$\frac{\Delta t}{R} = \frac{1}{160}$
M.S.E.	1.2851e-4	3.5028e-5	5.8980e-5	3.7058e-5

Tests for off-diagonal observables with different N (or β_N)



(A) potential 1



(B) potential 2 (more challenging)

The overall error is more sensitive to the number of beads N (which also affects the asymptotic error) than to the time step ratio R.

Summary:

- We propose a ring-polymer representation for non-adiabatic quantum systems (both adiabatic and diabatic representations); each bead has a level index besides the classical phase space configuration.
- The ring polymer configuration space can be sampled by path-integral molecular dynamics with surface hopping.
- The sampling efficiency can be enhanced by exploiting the limit of infinite switching and numerical implementation based on HMM multiscale integrator.

Future directions:

- Quantitative results on convergence
- Continuum limit of the PIMD with surface hopping
- Dynamical correlation function; how to design sampling schemes with systematic controllable error?
- Open quantum systems, non-equilibrium sampling for quantum dynamics, etc.

Thanks for your attention!

Email: jianfeng@math.duke.edu

URL: http://www.math.duke.edu/~jianfeng/

- L. and Zhou, Accelerated sampling by infinite swapping of path integral molecular dynamics with surface hopping, J. Chem. Phys. 2018
- L. and Zhou, Path integral molecular dynamics with surface hopping for thermal equilibrium sampling of nonadiabatic systems, J. Chem. Phys. 2017