

Frozen Gaussian Approximation with Surface Hopping: a **stochastic semiclassical algorithm** for **high dimensional Schrödinger equations**

Jianfeng Lu 鲁剑锋

Duke University
jianfeng@math.duke.edu

IPAM, UCLA, February 2017

Based on joint works with
Zhenning Cai (National Univ. of Singapore)
Zhe Wang (Duke University)
Zhennan Zhou (Duke → BICMR, Peking University)



The Schrödinger equation for nuclei and electrons (in atomic unit)

$$i\partial_t u = -\frac{1}{2M}\Delta_x u - \frac{1}{2}\Delta_r u + V(x, r)u,$$

where $x \in \mathbb{R}^m$ the position of nuclei and $r \in \mathbb{R}^n$ the position of electrons.
After rescaling

$$i\varepsilon\partial_t u = -\frac{\varepsilon^2}{2}\Delta_x u - \frac{1}{2}\Delta_r u + V(x, r)u,$$

where $\varepsilon = \sqrt{1/M} \ll 1$ (about 1/40 for hydrogen nucleus).

P.A.M. Dirac, 1929

The fundamental laws necessary to the mathematical treatment of large parts of physics and the whole of chemistry are thus fully known, and the difficult lies **only** in the fact that application of these laws leads to equations that are **too complex to be solved**.

The dimension is too high even for small molecules.

Let us try to separate the nucleus and electron degrees of freedom

$$i\varepsilon\partial_t u = -\frac{\varepsilon^2}{2}\Delta_x u - \frac{1}{2}\Delta_r u + V(x,r)u =: -\frac{\varepsilon^2}{2}\Delta_x u + H_e(x)u.$$

Denote $\Psi_k(r; x)$ as the eigenstates of $H_e(x)$

$$H_e(x)\Psi_k(r; x) = E_k(x)\Psi_k(r; x),$$

where x enters as a parameter. These are called **adiabatic states**.

It is a whole different story how to approximate the above eigenvalue problem, *aka* the electronic structure theory, which we will not go into here ... see the talks of Claudia Draxl, Gregory Beylkin, and Francois Gygi.

The dimension is too high even for small molecules.

Let us try to separate the nucleus and electron degrees of freedom

$$i\varepsilon\partial_t u = -\frac{\varepsilon^2}{2}\Delta_x u - \frac{1}{2}\Delta_r u + V(x,r)u =: -\frac{\varepsilon^2}{2}\Delta_x u + H_e(x)u.$$

Denote $\Psi_k(r; x)$ as the eigenstates of $H_e(x)$

$$H_e(x)\Psi_k(r; x) = E_k(x)\Psi_k(r; x),$$

where x enters as a parameter. These are called [adiabatic states](#).

It is a whole different story how to approximate the above eigenvalue problem, *aka* the electronic structure theory, ~~which we will not go into here ...~~ see the talks of Claudia Draxl, Gregory Beylkin, and Francois Gygi.

Let us make a digression anyway, since there are many interested in HPC, electronic structure calculations, and scalable libraries here ...



Let us try to separate the nucleus and electron degrees of freedom

$$i\varepsilon\partial_t u = -\frac{\varepsilon^2}{2}\Delta_x u - \frac{1}{2}\Delta_r u + V(x,r)u =: -\frac{\varepsilon^2}{2}\Delta_x u + H_e(x)u.$$

Denote $\Psi_k(r; x)$ as the eigenstates of $H_e(x)$

$$H_e(x)\Psi_k(r; x) = E_k(x)\Psi_k(r; x),$$

where x enters as a parameter. These are called [adiabatic states](#).

(time-dependent) Born-Oppenheimer approximation:

$$u(t, x, r) \approx u_0(t, x)\Psi_0(r; x),$$

with u_0 solving the nuclei Schrödinger equation (dropping terms on $\mathcal{O}(\varepsilon)$)

$$i\varepsilon\partial_t u_0(t, x) = -\frac{\varepsilon^2}{2}\Delta_x u_0 + E_0(x)u_0.$$

Nuclei Schrödinger equation on the energy surface $E_0(x)$:

$$i\varepsilon\partial_t u_0(t, x) = -\frac{\varepsilon^2}{2}\Delta_x u_0 + E_0(x)u_0.$$

It is still impractical to solve directly, as this is a PDE on \mathbb{R}^m .

To make it practical, recall that $\varepsilon \ll 1$. The semiclassical limit $\varepsilon \rightarrow 0$ gives the Hamiltonian dynamics for the nuclei with Hamiltonian $\frac{1}{2}|p|^2 + E_0(q)$.

$$\begin{aligned}\frac{dp}{dt} &= -\nabla_q E_0(q); \\ \frac{dq}{dt} &= p.\end{aligned}$$

This is known as the **ab initio molecular dynamics**, since the interaction potential E_0 is given by a first principle calculations.

Thus, from **Schrödinger equations** to **molecular dynamics**, we have made two approximations:

- Adiabatic approximation: $u(t, x, r) = u_0(t, x)\Psi_0(r; x)$;
- Semiclassical approximation: $\varepsilon \downarrow 0$.

The adiabatic approximation is justified if $E_0(x)$ (ground state) as a **potential energy surface** is well separated from $E_1(x)$ (excited state), and hence *the electron will not go to the excited state $E_1(x)$* .

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

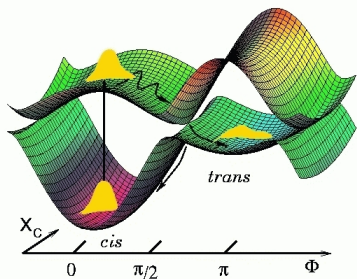


Figure: Schematic of ultrafast non-adiabatic photoreaction.

© Gerhard Stock

Thus, from **Schrödinger equations** to **molecular dynamics**, we have made two approximations:

- Adiabatic approximation: $u(t, x, r) = u_0(t, x)\Psi_0(r; x)$;
- Semiclassical approximation: $\varepsilon \downarrow 0$.

The adiabatic approximation is justified if $E_0(x)$ (ground state) as a **potential energy surface** is well separated from $E_1(x)$ (excited state), and hence *the electron will not go to the excited state $E_1(x)$* .

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

Instead, we would need to consider components of excited states:

$$u(t, x, r) = \sum_i u_i(t, x)\Psi_i(r; x),$$

and to make the notation simple, we will just consider two states (generalization is straightforward):

$$u(t, x, r) = u_0(t, x)\Psi_0(r; x) + u_1(t, x)\Psi_1(r; x).$$

Two state wave function:

$$u(t, x, r) = u_0(t, x)\Psi_0(r; x) + u_1(t, x)\Psi_1(r; x).$$

Substitute into the Schrödinger equation, we get

$$i\varepsilon\partial_t \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = -\frac{\varepsilon^2}{2}\Delta_x \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} + \begin{pmatrix} E_0 & \\ & E_1 \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} - \varepsilon^2 \sum_{j=1}^m \begin{pmatrix} d_{00} & d_{01} \\ d_{10} & d_{11} \end{pmatrix}_j \partial_{x_j} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} - \frac{\varepsilon^2}{2} \begin{pmatrix} D_{00} & D_{01} \\ D_{10} & D_{11} \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix},$$

where for $k, l = 0, 1$

$$D_{kl}(x) = \langle \Psi_l(r; x), \Delta_x \Psi_k(r; x) \rangle_r,$$
$$(d_{kl}(x))_j = \langle \Psi_l(r; x), \partial_{x_j} \Psi_k(r; x) \rangle_r.$$

As before, it is almost hopeless trying to directly simulate the matrix Schrödinger equation in high dimensions, and thus we still hope to resort to the semiclassical approximation.

As $\varepsilon \rightarrow 0$, u is highly oscillatory, and thus

$$\partial_t u = \mathcal{O}(\varepsilon^{-1}), \quad \partial_x u = \mathcal{O}(\varepsilon^{-1}), \quad \Delta_x u = \mathcal{O}(\varepsilon^{-2}), \quad \text{while } u = \mathcal{O}(1).$$

Therefore, to the order of $\mathcal{O}(\varepsilon)$, we may neglect the D term (with rigorous control of error):

$$i\varepsilon \partial_t \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = -\frac{\varepsilon^2}{2} \Delta_x \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} + \begin{pmatrix} E_0 & \\ & E_1 \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} - \varepsilon^2 \sum_{j=1}^m \begin{pmatrix} d_{00} & d_{01} \\ d_{10} & d_{11} \end{pmatrix}_j \partial_{x_j} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} - \frac{\varepsilon^2}{2} \begin{pmatrix} D_{00} & D_{01} \\ D_{10} & D_{11} \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}.$$

We should have also introduced another small parameter δ indicating the gap between E_0 and E_1 , as the $\mathcal{O}(\varepsilon)$ contribution only matters when the gap δ is small; the dynamics will be otherwise adiabatic.

For simplicity, let us suppress the δ dependence in the notation.

Our results hold uniformly for arbitrarily small spectral gap, as long as d and D stay $\mathcal{O}(1)$.

As $\varepsilon \rightarrow 0$, u is highly oscillatory, and thus

$$\partial_t u = \mathcal{O}(\varepsilon^{-1}), \quad \partial_x u = \mathcal{O}(\varepsilon^{-1}), \quad \Delta_x u = \mathcal{O}(\varepsilon^{-2}), \quad \text{while } u = \mathcal{O}(1).$$

Therefore, to the order of $\mathcal{O}(\varepsilon)$, we may neglect the D term (with rigorous control of error):

$$i\varepsilon \partial_t \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = -\frac{\varepsilon^2}{2} \Delta_x \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} + \begin{pmatrix} E_0 & \\ & E_1 \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} - \varepsilon^2 \sum_{j=1}^m \begin{pmatrix} d_{00} & d_{01} \\ d_{10} & d_{11} \end{pmatrix}_j \partial_{x_j} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}$$

where $d_{kl}(x) = \langle \Psi_l(r; x), \nabla_x \Psi_k(r; x) \rangle_r$.

Instead of solving the matrix Schrödinger equation, we still need to employ semiclassical approximation and get something like the molecular dynamics, but now, we would need to take into account

the coupling terms between the two energy surfaces.

This question is, how?

What theoretical chemists proposed ...

Tully's fewest switch surface hopping [JChemPhys 1990]

Semiclassical approximation to the matrix Schrödinger equation

$$i\varepsilon\partial_t \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = -\frac{\varepsilon^2}{2}\Delta_x \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} + \begin{pmatrix} E_0 & \\ & E_1 \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} - \varepsilon^2 \sum_{j=1}^m \begin{pmatrix} d_{00} & d_{01} \\ d_{10} & d_{11} \end{pmatrix}_j \partial_{x_j} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}$$

Intuition: Without the coupling d term, it is really just propagating on one of the energy surfaces, and hence very similar to the ab initio molecular dynamics, however, we need to **modify** the dynamics so to account the coupling terms.

Denote a trajectory as $(q(t), p(t))$, the FSSH algorithm proposes to solve along the trajectory a ODE system of $(k = 1, 2)$

$$i\varepsilon\dot{c}_k(t) = E_k(q(t))c_k(t) - i\varepsilon \sum_j c_j(t)p(t) \cdot d_{kj}(q(t)),$$

and $(|c_0(t)|^2, |c_1(t)|^2)$ is understood as “population” on each energy surface along the trajectory.

Tully's fewest switch surface hopping [JChemPhys 1990]

Population evolution (let us accept this equation):

$$i\varepsilon\dot{c}_k(t) = E_k(q(t))c_k(t) - i\varepsilon \sum_j c_j(t)p(t) \cdot d_{kj}(q(t)).$$

Except at hopping, the trajectory evolves on one of the energy surface l_t as a Hamiltonian flow

$$\dot{q}(t) = p(t), \quad \dot{p}(t) = -\nabla E_{l_t}(q(t)).$$

If $l_t = 0$ and $(|c_0(t)|^2)' < 0$, the trajectory “hops” to the other surface (so $l_{t+\delta t} = 1$) with probability $\delta t (|c_0(t)|^2)' / |c_0(t)|^2$; if $(|c_0(t)|^2)' \geq 0$, no hopping occurs. Similarly for $l_t = 1$.

When hopping occurs, a momentum shift is used to conserve the classical energy (detailed strategy not too important for our discussion).

Motivation of our work

- While the surface hopping algorithm is widely used, the derivation of the algorithm from the Schrödinger equation remains mysterious. Even chemists, including Tully himself, think so ...

The situation was so confusing that it is not even clear how to interpret the computational results obtained by the algorithm.

Motivation of our work

- While the surface hopping algorithm is widely used, the derivation of the algorithm from the Schrödinger equation remains mysterious. Even chemists, including Tully himself, think so ...

The situation was so confusing that it is not even clear how to interpret the computational results obtained by the algorithm.

- After the birth of the FSSH, it is found that it has issues in certain situations. As a result, there have been many attempts try to “correct” it, for example, by introducing “decoherence”, etc.

Without a systematic understanding, it is very hard to see which direction one should follow to improve the algorithms.

Motivation of our work

- While the surface hopping algorithm is widely used, the derivation of the algorithm from the Schrödinger equation remains mysterious. Even chemists, including Tully himself, think so ...

The situation was so confusing that it is not even clear how to interpret the computational results obtained by the algorithm.

- After the birth of the FSSH, it is found that it has issues in certain situations. As a result, there have been many attempts try to “correct” it, for example, by introducing “decoherence”, etc.

Without a systematic understanding, it is very hard to see which direction one should follow to improve the algorithms.

- Besides the surface hopping, there is a whole family of non-adiabatic methods in the literature with many active developments.

Lots of opportunities for applied mathematicians.

high dimensional numerics / stochastic algorithms / applied harmonic analysis / modern numerical linear algebra

Previous works

- In the chemistry / physics literature, often the starting point to “understand” the FSSH algorithm is the quantum classical Liouville equation [Kapral-Ciccotti 1999], which is an extension of the Moyal’s evolution for Wigner distribution to the matrix Schrödinger equations. See e.g., the works of R. Kapral, Ch. Schütte, J. Subotnik, J. Tully and others.

Here we will take a different route, which is based on wavepackets (coherent states).

On the other hand, the idea behind our algorithm also leads to rigorous surface hopping algorithms for QCLE and in fact more general high dimensional hyperbolic systems (w/ **Zhenning Cai**).

- We actually have not yet fully understood Tully’s original FSSH algorithm; one difficulty is to understand the c equation, which is in fact Ehrenfest type (mean field approximation)??

We will offer an algorithm in the same spirit with rigorous error control.

Previous mathematical works

- Somehow rather confusingly, “surface hopping” is also used for a different “older” algorithm [Tully-Preston 1971], which relies on Landau-Zener asymptotics. This “original” surface hopping algorithm deals with a special situation of hopping at conical intersections (i.e., Landau-Zener transition).

The FSSH type algorithms are for more general situations, and is quite different in many ways.

The Landau-Zener asymptotics has been rigorously justified by the work of G. Hagedorn, A. Joye, S. Teufel and others

Numerical analysis done by F. Bornemann, C. Lasser, Shi Jin, Ch. Schütte and others.

Our contribution:

An algorithm in the spirit of FSSH rigorously derived from Schrödinger equation ...

Surface hopping as a path integral

The surface hopping is an asymptotically convergent **path integral** representation for the matrix Schrödinger equation!

Similar to using Feynman-Kac formula to stochastically solve reaction-diffusion equations (for physicists: diffusion Monte Carlo).

$$i\varepsilon\partial_t \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = -\frac{\varepsilon^2}{2}\Delta_x \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} + \begin{pmatrix} E_0 & \\ & E_1 \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} - \varepsilon^2 \sum_{j=1}^m \begin{pmatrix} d_{00} & d_{01} \\ d_{10} & d_{11} \end{pmatrix}_j \partial_{x_j} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}$$

We have the path integral representation up to semiclassical error

$$u(t, x, r) = \mathbb{E} F(x, r; \{\tilde{z}_s\}_{0 \leq s \leq t}) + \mathcal{O}_{L^2}(\varepsilon),$$

where

- \mathbb{E} stands for expectation over the path space of \tilde{z}_t , and
- F is a functional depends on the trajectory $\{\tilde{z}_s\}_{0 \leq s \leq t}$ (exact expression, unfortunately quite long and requires more notations, will be given later).

Surface hopping as a path integral

The surface hopping is an asymptotically convergent **path integral** representation for the matrix Schrödinger equation!

Similar to using Feynman-Kac formula to stochastically solve reaction-diffusion equations (for physicists: diffusion Monte Carlo).

$$i\varepsilon\partial_t \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = -\frac{\varepsilon^2}{2}\Delta_x \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} + \begin{pmatrix} E_0 & \\ & E_1 \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} - \varepsilon^2 \sum_{j=1}^m \begin{pmatrix} d_{00} & d_{01} \\ d_{10} & d_{11} \end{pmatrix}_j \partial_{x_j} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}$$

We have the path integral representation up to semiclassical error

$$u(t, x, r) = \mathbb{E} F(x, r; \{\tilde{z}_s\}_{0 \leq s \leq t}) + \mathcal{O}_{L^2}(\varepsilon),$$

Theorem (L.-Zhou, arXiv:1602.06459)

With some regularity assumptions on E, d and D , we have (uniformly wrt spectral gap)

$$\|u - \mathbb{E} F(x, r; \{\tilde{z}_s\}_{0 \leq s \leq t})\|_{L^2} = \mathcal{O}(\varepsilon).$$

Surface hopping trajectory ensemble

Path space average

$$u(t, x, r) = \mathbb{E} F(x, r; \{\tilde{z}_s\}_{0 \leq s \leq t}) + \mathcal{O}_{L^2}(\varepsilon).$$

The trajectory follows a Markov switching process

$$dz_t = d(q_t, p_t) = (p_t, -\nabla_q E_{l_t}(p_t, q_t)) dt,$$

$$\mathbb{P}(l_{t+\delta t} = l \mid l_t = k, z_t = z) = \delta_{kl} + \lambda_{kl}(z)\delta t + o(\delta t)$$

with rate matrix

$$\lambda(z) = \begin{pmatrix} \lambda_{00}(z) & \lambda_{01}(z) \\ \lambda_{10}(z) & \lambda_{11}(z) \end{pmatrix} = \begin{pmatrix} -|p \cdot d_{01}(q)| & |p \cdot d_{01}(q)| \\ |p \cdot d_{10}(q)| & -|p \cdot d_{10}(q)| \end{pmatrix}.$$

- l_t follows a (nonhomogeneous) Poisson process with state dependent jumping intensity $\lambda(z)$; it contains a.s. finite number of jumps.
- z_t follows the Hamiltonian flow on energy surface l_t ;
 z_t is **continuous** on the phase space, in **sharp contrast** with existing surface hopping algorithms, which try to conserve “classical energy” of each trajectory at hopping (Why?? only the ensemble average matters after all?).

Surface hopping trajectory ensemble

Path space average

$$u(t, x, r) = \mathbb{E} F(x, r; \{\tilde{z}_s\}_{0 \leq s \leq t}) + \mathcal{O}_{L^2}(\varepsilon).$$

The trajectory follows a Markov switching process

$$dz_t = d(q_t, p_t) = (p_t, -\nabla_q E_{l_t}(p_t, q_t)) dt,$$

$$\mathbb{P}(l_{t+\delta t} = l \mid l_t = k, z_t = z) = \delta_{kl} + \lambda_{kl}(z)\delta t + o(\delta t)$$

with rate matrix

$$\lambda(z) = \begin{pmatrix} \lambda_{00}(z) & \lambda_{01}(z) \\ \lambda_{10}(z) & \lambda_{11}(z) \end{pmatrix} = \begin{pmatrix} -|p \cdot d_{01}(q)| & |p \cdot d_{01}(q)| \\ |p \cdot d_{10}(q)| & -|p \cdot d_{10}(q)| \end{pmatrix}.$$

Equivalently, in terms of Kolmogorov forward equation (ϱ being the distribution)

$$\frac{\partial}{\partial t} \varrho_t(z, l) + \{h^{(l)}, \varrho_t(z, l)\} = \sum_{k=0}^1 \lambda_{kl}(z) \varrho_t(z, k),$$

where $\{\cdot, \cdot\}$ stands for the Poisson bracket (of the Hamiltonian flow).

A rigorous surface hopping algorithm

The path integral representation readily suggests an algorithm for approximating u :

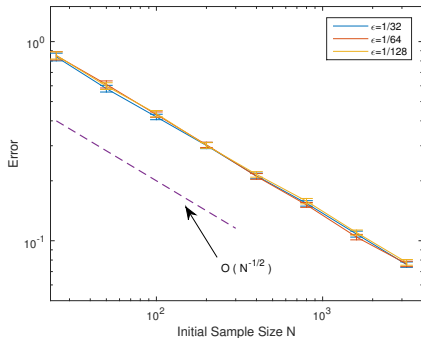
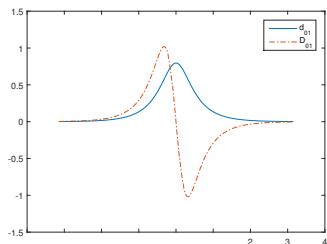
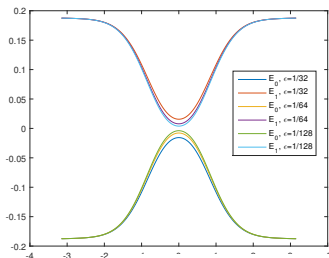
$$\begin{aligned}u(t, x, r) &= \mathbb{E} F(x, r; \{\tilde{z}_s\}_{0 \leq s \leq t}) + \mathcal{O}_{L^2}(\varepsilon) \\ &= \frac{1}{M} \sum_{i=1}^M F(x, r; \{\tilde{z}_s^i\}_{0 \leq s \leq t}) + \mathcal{O}_{L^2}\left(\frac{1}{\sqrt{M}}\right) + \mathcal{O}_{L^2}(\varepsilon),\end{aligned}$$

where \tilde{z}^i is i.i.d. realization of the trajectory.

With our mathematical understanding of the surface hopping algorithm (in particular, what it really tries to compute), it is natural to design more efficient sampling algorithms based on e.g., interacting walkers.

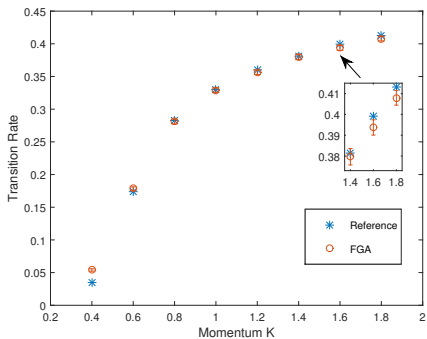
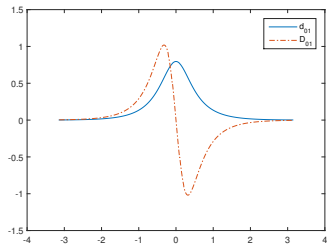
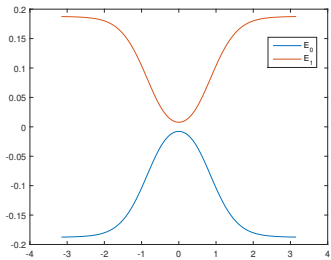
[L.-Zhou, JChemPhys 2016]

Our algorithm actually offers an approximation to the wave function (and hence nuclear quantum effect, physical observables, transition rate, etc.)

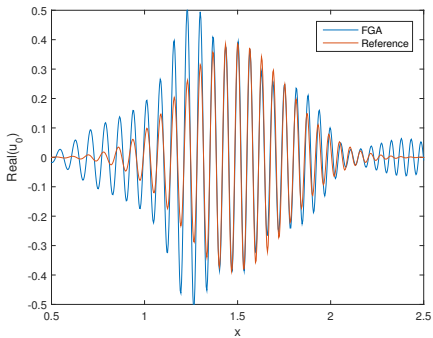
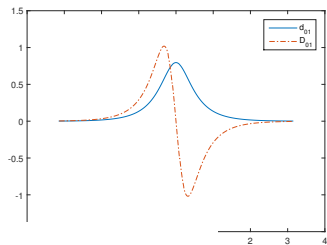
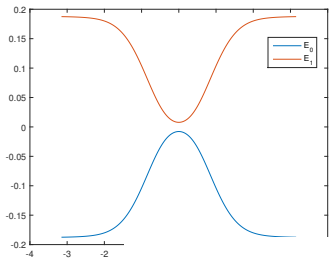


L^2 error as function of N and ϵ (sampling error dominates)

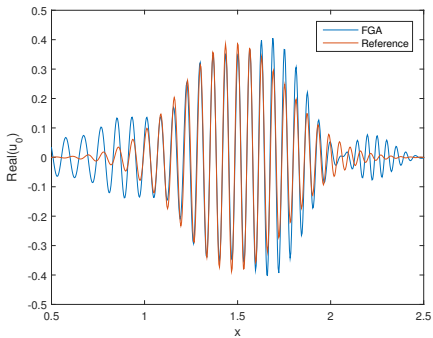
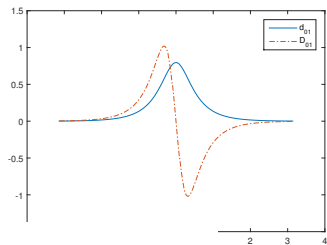
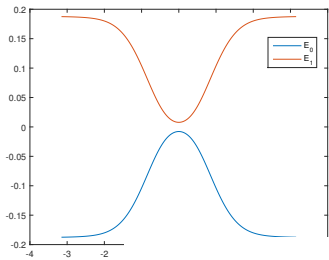
Example from [Tully, JChemPhys 1990]



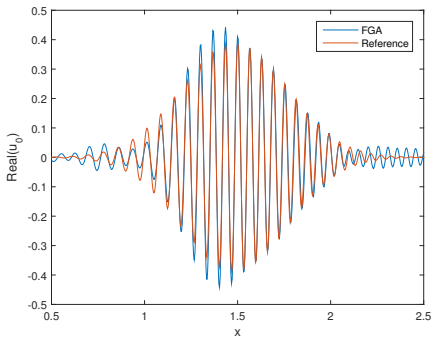
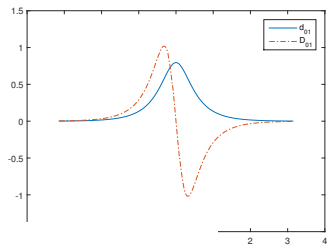
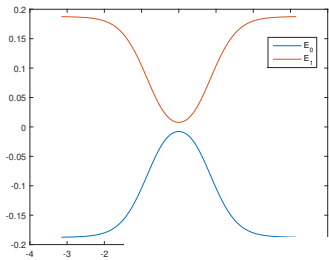
$N = 1600$



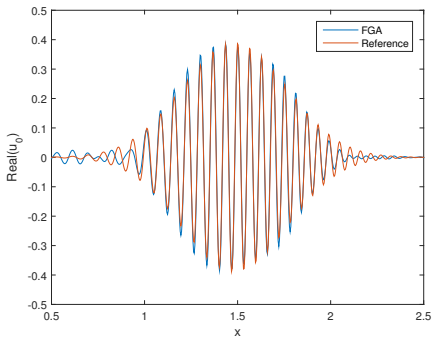
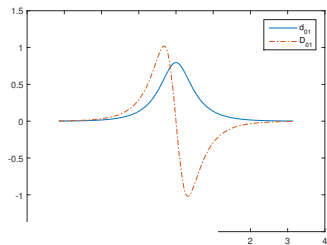
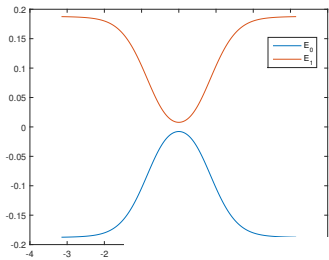
200 initial trajectories



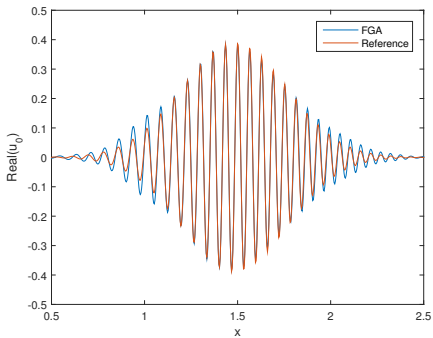
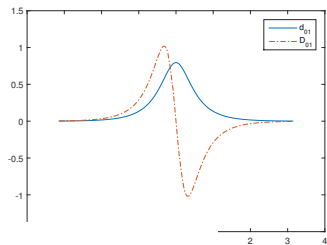
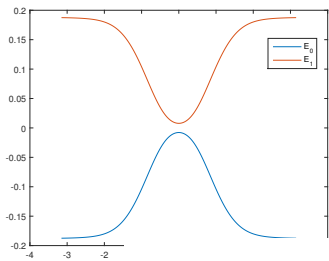
400 initial trajectories



800 initial trajectories



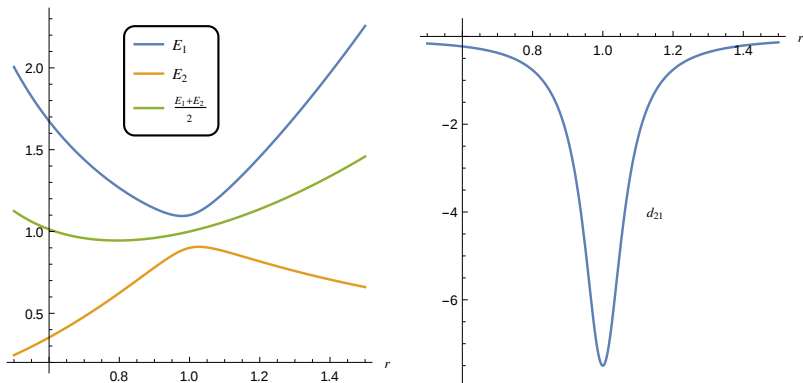
1600 initial trajectories



3200 initial trajectories

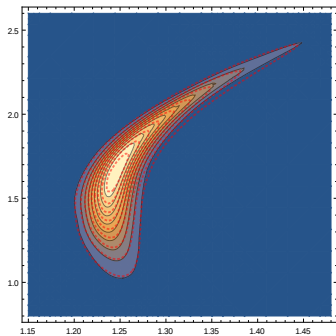
We can extend the algorithm for a wide class of high dimensional linear hyperbolic systems (w/ Zhenning Cai).

Application to QCLE with $V(r) = \begin{pmatrix} x^2 & 1/10 \\ 1/10 & 1/x \end{pmatrix}$ and $\varepsilon = 0.01$.

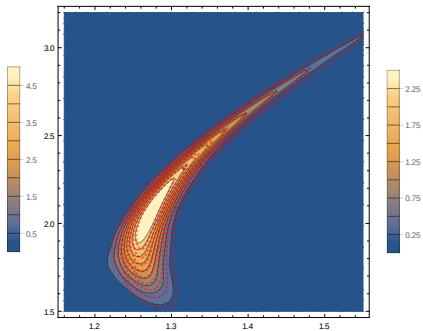


Example from [Horenko, Salzmann, Schmidt, Schütte, JChemPhys 2002]

Application to QCLE with $V(r) = \begin{pmatrix} x^2 & 1/10 \\ 1/10 & 1/x \end{pmatrix}$ and $\varepsilon = 0.01$.



Left: W_{11} ;



Right: W_{22} .

Note that this is a 2D simulation (since the QCLE is on the phase space).

We also have numerical results for much higher dimension systems (spin Boson systems up to 400 dimension).

Surface hopping as a path integral

$$i\varepsilon\partial_t \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = -\frac{\varepsilon^2}{2}\Delta_x \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} + \begin{pmatrix} E_0 & \\ & E_1 \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} - \varepsilon^2 \sum_{j=1}^m \begin{pmatrix} d_{00} & d_{01} \\ d_{10} & d_{11} \end{pmatrix}_j \partial_{x_j} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}$$

We have

$$u(t, x, r) = \mathbb{E} F(x, r; \{\tilde{z}_s\}_{0 \leq s \leq t}) + \mathcal{O}_{L^2}(\varepsilon),$$

Key idea:

- Semiclassical methods for single surface Schrödinger equation:

$$i\varepsilon\partial_t u = -\frac{\varepsilon^2}{2}\Delta_x u + E_0 u - \varepsilon^2 d_{00} \cdot \nabla u.$$

e.g., Herman-Kluk propagator (frozen Gaussian approximation), Gaussian beam methods, or other approaches ...

- To deal with the off-diagonal term, let us rewrite the system as

$$i\varepsilon\partial_t \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = H \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} + \varepsilon \begin{pmatrix} 0 & H_{01} \\ H_{10} & 0 \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}$$

$$i\varepsilon\partial_t \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = H \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} + \varepsilon \begin{pmatrix} 0 & H_{01} \\ H_{10} & 0 \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}$$

By Duhamel's principle

$$\begin{pmatrix} u_0 \\ u_1 \end{pmatrix}(t) = e^{-\frac{i}{\varepsilon}tH} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}(0) - i \int_0^t e^{-\frac{i}{\varepsilon}(t-t_1)H} \begin{pmatrix} 0 & H_{01} \\ H_{10} & 0 \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}(t_1) dt_1$$

Thus we get the Dyson series:

$$\begin{aligned} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}(t) &= e^{-i\frac{t}{\varepsilon}H} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}(0) \\ &- i \int_0^t e^{-\frac{i}{\varepsilon}(t-t_1)H} \begin{pmatrix} 0 & H_{01} \\ H_{10} & 0 \end{pmatrix} e^{-\frac{i}{\varepsilon}t_1H} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}(0) dt_1 \\ &- \int_0^t \int_0^{t_2} e^{-\frac{i}{\varepsilon}(t-t_2)H} \begin{pmatrix} 0 & H_{01} \\ H_{10} & 0 \end{pmatrix} e^{-\frac{i}{\varepsilon}(t_2-t_1)H} \begin{pmatrix} 0 & H_{01} \\ H_{10} & 0 \end{pmatrix} \times \\ &\quad \times e^{-\frac{i}{\varepsilon}t_1H} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}(0) dt_1 dt_2 \\ &+ \dots \end{aligned}$$

Thus we get the Dyson series:

$$\begin{aligned} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} (t) &= e^{-i\frac{t}{\hbar}H} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} (0) \\ &- i \int_0^t e^{-i\frac{i}{\hbar}(t-t_1)H} \begin{pmatrix} 0 & H_{01} \\ H_{10} & 0 \end{pmatrix} e^{-i\frac{i}{\hbar}t_1H} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} (0) dt_1 \\ &- \int_0^t \int_0^{t_2} e^{-i\frac{i}{\hbar}(t-t_2)H} \begin{pmatrix} 0 & H_{01} \\ H_{10} & 0 \end{pmatrix} e^{-i\frac{i}{\hbar}(t_2-t_1)H} \begin{pmatrix} 0 & H_{01} \\ H_{10} & 0 \end{pmatrix} \times \\ &\quad \times e^{-i\frac{i}{\hbar}t_1H} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} (0) dt_1 dt_2 \\ &+ \dots \end{aligned}$$

Observe that one can view each integrand as a particular hopping time sequence ($0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq t$). The series then corresponds to a summation / integration over all possible hopping time sequences.

This is a very high dimensional integral; thus calls for Monte Carlo algorithms. If we “poissonize” the hopping times, we arrive at the semiclassical path integral representation after some calculations.

$$\begin{aligned}
u_{\text{FGA}}(t, x, r) = & \mathcal{Z} \mathbb{E}_{\tilde{z}_t} \left[\Psi_{n \bmod 2}(r; x) \left(\prod_{k=1}^n \frac{\tau^{(k)}(T_{k:1}, z_0)}{|\tau^{(k)}(T_{k:1}, z_0)|} \right) \right. \\
& \times \frac{C^{(n)}(t, T_{n:1}, z_0)}{|C^{(0)}(0, z_0)|} \exp\left(\frac{i}{\varepsilon} \Theta^{(n)}(t, T_{n:1}, z_0, x)\right) \\
& \left. \times \exp\left(\int_{t_n}^t |\tau^{(n+1)}(s, T_{n:1}, z_0)| ds\right) \prod_{k=1}^n \exp\left(\int_{t_{k-1}}^{t_k} |\tau^{(k)}(s, T_{k-1:1}, z_0)| ds\right) \right].
\end{aligned}$$

While this expression looks complicated, it is actually easy to understand

- The first line is because $\tau^{(k)}(T_{k:1}, z_0)$ (such as $d_{01} \cdot p$) is in general complex, and therefore we have to use its modulus as intensity in the Poisson process, and hence the phase corrections;
- The second line is the integrand of the deterministic ansatz, except that we reweight by the initial sampling;
- The third line is correction due to our non-homogeneous and state dependent Poisson process (for likelihood of hopping sequences).

Looking back at the fewest switches surface hopping algorithms:

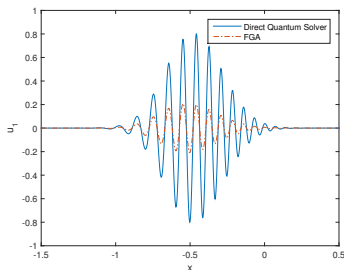
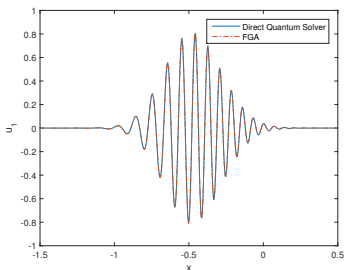
- The computational cost of each trajectory in our algorithm is similar to Tully's FSSH.

It is an interesting future direction to better understand the difference and connections, in particular, to design the "best" corrections to Tully's approach (it is possible from the Monte Carlo point of view).

- While similar in spirit, Tully's FSSH is perhaps more like a geometric optics / WKB approach, which only tries to capture the "semiclassical trajectory", but does not provide information on wavefunctions (since e.g., the amplitude factor C is missing in the ansatz).
- The momentum jumps for conserving the "classical energy" along trajectory require some further thoughts. Note that for our ansatz, each packet is not a solution to the Schrödinger equation itself (this is the same in the Herman-Kluk propagator).

The momentum jump idea is perhaps putting too much emphasis on a single trajectory, rather than the ensemble average.

- The non-homogeneous and state dependent Poisson process (which is involved in essentially all version of surface hopping algorithms) requires the crucial reweighting terms in averaging.



- As for ongoing and future works, the understanding of the path integral perspective of the surface hopping algorithms opens doors to many possible extensions for non-adiabatic systems: open systems, thermal equilibrium average, dynamical correlation, etc.

It also connects to other flavors of quantum Monte Carlo for ground states and time propagation.

Thanks for your attention!

Email: jianfeng@math.duke.edu

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

- with Zhennan Zhou, Frozen Gaussian approximation with surface hopping for mixed quantum-classical dynamics: A mathematical justification of fewest switches surface hopping algorithms, arXiv:1602.06459
- with Zhennan Zhou, Improved sampling and validation of frozen Gaussian approximation with surface hopping algorithm for nonadiabatic dynamics, J. Chem. Phys. 2016