A DIABATIC SURFACE HOPPING ALGORITHM BASED ON TIME DEPENDENT PERTURBATION THEORY AND SEMICLASSICAL ANALYSIS

DI FANG AND JIANFENG LU

ABSTRACT. Surface hopping algorithms are popular tools to study dynamics of the quantum-classical mixed systems. In this paper, we propose a surface hopping algorithm in diabatic representations, based on time dependent perturbation theory and semiclassical analysis. The algorithm can be viewed as a Monte Carlo sampling algorithm on the semiclassical path space for piecewise deterministic path with stochastic jumps between the energy surfaces. The algorithm is validated numerically and it shows good performance in both weak coupling and avoided crossing regimes.

1. INTRODUCTION

The quantum-classical dynamics is known as an important tool to investigate non-adiabatic systems. Among them, surface hopping algorithms, in particular fewest switches surface hopping (FSSH) first proposed by John Tully [28, 25, 26], have been applied to many different practical situations. While its success and improvements under the adiabatic representation have been well documented in both chemical and mathematical literatures, its performance in the diabatic picture seems to be less satisfactory, in which case a full theory or partial explanation remains as an interesting open problem. In fact, it was once claimed by Tully himself [27] that the surface hopping algorithm "does not appear justified with a diabatic basis". Many efforts have been devoted to the improvements of surface hopping algorithm under the diabatic representation (see, e.g., [30, 31, 21, 19, 4]). The focus of this paper is to propose a diabatic surface hopping method using asymptotic derivation from the Schrödinger equation in a diabatic basis.

A nature question to ask is why to consider the diabatic representation. There are various reasons with the three main motivations listed as follows. First, as mentioned in [23, 29], a diabatic representation is quite attractive or convenient in terms of simulations, because the diabatic energies and couplings are smooth functions of the nuclear coordinates involving multiplicative instead of derivative operators. Hence, it would make more sense to consider the diabtic representation near the transition zone where the adiabatic potential energy surface is ill-defined or the adiabatic coupling becomes very large. Moreover, in some cases a diabatic basis may arise as a nature choice. For example, consider the spin transitions, and one could identify the potential energy surfaces with pure spin character and transitions induced by the spin-orbit interaction, instead of the mixed spin character in the adiabatic basis on the contrary. Another example is the photoreaction, where the molecular system is prepared in a diabatic electronic state, instead of an adiabatic one, by the laser field [19, 3, 9]. Third, as is well-known there are

at least two mainstreams of surface hopping algorithms – the fewest-switch type (in the spirits of FSSH) and Landau-Zener type (based on the famous Landau-Zener asymptotics [10, 32]). Fruitful applications as there are of both types, a clear connection between the two remains rather mysterious. Due to the diabatic nature of the original form of the Landau-Zener formula, a rigorously justified diabatic surface hopping algorithm may bridge the gap for future study, which also serves as one of our motivations.

In this paper, we propose a new diabatic surface hopping algorithm in the same spirit of the FSSH algorithm. Our algorithm is derived based on semiclassical analysis of the Schrödinger equation, adapted from the adiabatic picture in the recent work of one of us [16, 2] to the diabatic setting here. This algorithm not only agrees with a clear intuition but also is justified by an asymptotic derivation. Inheriting the merits of the Schrödinger equation based derivation, our algorithm is time-reversible and also the commonly mentioned drawbacks of surface hopping, such as frustrated hops and decoherence, do not arise as issues for the algorithm. The dynamics is described by a system of ODEs with the hoppings represented by a stochastic jump process. In simulations, we use a direct Monte Carlo method, and the solution of the Schrödinger equation can be simply recovered by a trajectory average, *aka* a path integral stochastic representation.

The rest of the paper is organized as follows. Starting with an introduction of the Schrödinger equation in the diabatic representation that is studied here, we describe an intuitive understanding of the diabatic surface hopping based on time pertubation theory, and propose a diabatic surface hopping algorithm that matches the expansion. In Section 3, the algorithm is tested on various numerical examples, and the sources of errors are discussed. Finally, in the last two sections, we derive the algorithm asymptotically and stochastically, where an asymptotic derivation of the algorithm is given in Section 3, based on the frozen Gaussian approximation, *aka* the Herman-Kluk propagator, and a stochastic interpretation is presented in Section 4, where the solution of the Schrödinger equation is represented by a stochastic path integral.

2. THE DIABATIC SURFACE HOPPING METHOD

2.1. **Two level Schrödinger equation in a diabatic representation.** For simplicity of notations, let us focus in this work on a two-level Schrödinger equation in a diabatic representation:

$$(2.1) \qquad i\varepsilon\partial_t \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = -\frac{\varepsilon^2}{2}\Delta_x \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} + \begin{pmatrix} V_{00} & \delta V_{01} \\ \delta V_{10} & V_{11} \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix},$$

with initial condition

$$(2.2) \qquad u(0,x) = u_{\text{in}}(x) = \begin{pmatrix} u_0(0,x) \\ 0 \end{pmatrix},$$

where $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^m$, $u(t,x) = [u_0(t,x),\ u_1(t,x)]^\top$ is the (two component) wave function, and we have assumed without loss of generality that the initial condition concentrates on the first energy surface. The Hamiltonian is given by $H = -\frac{\varepsilon^2}{2}\Delta_x + V(x)$, where

$$V(x) = \begin{pmatrix} V_{00}(x) & \delta V_{01}(x) \\ \delta V_{10}(x) & V_{11}(x) \end{pmatrix}$$

is a matrix potential. V is Hermitian and hence $V_{01} = V_{10}^*$ for every x.

In this setup, we have two parameters: ε is the semiclassical parameter and δ is a scaling parameter for the amplitude of the off-diagonal terms in the matrix potential. A small δ means weak coupling between the two diabatic surfaces. For a given problem in practice, ε and δ are small but fixed parameters. Here for our algorithm, we would consider the regimes that ε and δ are small, but we will not delve much into the asymptotic limit that they go to 0. The asymptotic limits of these parameters going to zero are studied for instance for the scaling that $\delta = \sqrt{\varepsilon} \to 0$ [5], where the Landau-Zener asymptotic limit was studied; see also [13, 12, 6]. Another well known case is the scaling limit of $\delta \to 0$ with fixed ε, which is related to the celebrated Marcus theory [18]. We will leave the rigorous study of the asymptotic limit of our algorithm to future works; while we will test the algorithm on these cases in our numerical examples.

2.2. **Surface hopping from the view of time dependent perturbation theory.** As usual for diabatic surface hopping algorithms, we view the diagonal elements of the potential matrix $V(x)$ as two potential energy surfaces and the off-diagonal elements as coupling terms. One key observation behind our diabatic surface hopping is a time dependent perturbation theory expansion of the solution to the Schrödinger equation. To start with, let us rewrite (2.1) to separate the diagonal and off-diagonal terms

$$(2.3) \qquad i\varepsilon \partial_t \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = \begin{pmatrix} -\frac{\varepsilon^2}{2}\Delta_x + V_{00} & 0 \\ 0 & -\frac{\varepsilon^2}{2}\Delta_x + V_{11} \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} + \delta \begin{pmatrix} 0 & V_{01} \\ V_{10} & 0 \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}.$$

Then the wave function $u(t,x) = [u_0(t,x), u_1(t,x)]^\top$ can be given by the Duhamel's principle by viewing the last term above as a forcing term

$$(2.4) \qquad u(t,x) = \mathcal{U}(t,0) u(0,x) - \frac{i\delta}{\varepsilon} \int_0^t \mathcal{U}(t,t_1) \begin{pmatrix} 0 & V_{01} \\ V_{10} & 0 \end{pmatrix} u(t_1,x) \, dt_1,$$

where

$$\mathcal{U}(t,s) = \begin{pmatrix} \mathcal{U}_0(t,s) & \\ & \mathcal{U}_1(t,s) \end{pmatrix} = \begin{pmatrix} e^{-\frac{i}{\varepsilon}(t-s)H_0} & \\ & e^{-\frac{i}{\varepsilon}(t-s)H_1} \end{pmatrix}$$

is the propagator of the diagonal part of the equation

$$i\varepsilon \partial_t \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = \begin{pmatrix} -\frac{\varepsilon^2}{2}\Delta_x + V_{00} & \\ & -\frac{\varepsilon^2}{2}\Delta_x + V_{11} \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} =: \begin{pmatrix} H_0 & \\ & H_1 \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}.$$

The last equality gives the definition of H_k, the Hamiltonian of the $k-$th diabatic energy surface. To simplify the notations, we will use below the shorthand for the off-diagonal matrix potential

$$M := \begin{pmatrix} 0 & V_{01} \\ V_{10} & 0 \end{pmatrix}.$$

Substitute (2.4) recursively into the right hand side, we get

$$u(t,x) = \mathcal{U}(t,0) u(0,x) - \frac{i\delta}{\varepsilon} \int_0^t dt_1 \, \mathcal{U}(t,t_1) M \mathcal{U}(t_1,0) u(0,x)$$

$$(2.5) \qquad + \left(\frac{-i\delta}{\varepsilon}\right)^2 \int_0^t dt_2 \int_0^{t_2} dt_1 \, \mathcal{U}(t,t_2) M \mathcal{U}(t_2,t_1) M \mathcal{U}(t_1,0) u(0,x)$$

$$+ \left(\frac{-i\delta}{\varepsilon}\right)^3 \int_0^t dt_3 \int_0^{t_3} dt_2 \int_0^{t_2} dt_1 \, \mathcal{U}(t,t_3) M \mathcal{U}(t_3,t_2) M \mathcal{U}(t_2,t_1) M \mathcal{U}(t_1,0) u(0,x)$$

$$+ \cdots.$$

Recall that we have assumed that the initial datum (2.2) is non-zero only in the first component, the diagonal propagator $\mathscr{U}$ will keep the non-zero structure, while the action of M will "flip" the non-zero entries in the vector. Following this, a simple calculation shows that all terms on the right hand side of (2.5) containing even number (including 0) of M terms would contribute at time t to the first entry of $u(t, x)$, while the rest contributes to the second entry, *i.e.*,

$$
\begin{aligned}
u_0(t, x) = {} & \mathscr{U}_0(t, 0) \, u_0(0, x) \\
& + \left(\frac{-i\delta}{\varepsilon}\right)^2 \int_0^t dt_2 \int_0^{t_2} dt_1 \mathscr{U}_0(t, t_2) \, V_{01} \mathscr{U}_1(t_2, t_1) \, V_{10} \mathscr{U}_0(t_1, 0) \, u_0(0, x) \\
& + \left(\frac{-i\delta}{\varepsilon}\right)^4 \int_0^t dt_4 \int_0^{t_4} dt_3 \int_0^{t_3} dt_2 \int_0^{t_2} dt_1 \, \mathscr{U}_0(t, t_4) \, V_{01} \mathscr{U}_1(t_4, t_3) \, V_{10} \mathscr{U}_0(t_3, t_2) \, V_{01} \times \\
& \hspace{6cm} \times \mathscr{U}_1(t_2, t_1) \, V_{10} \mathscr{U}_0(t_1, 0) \, u_0(0, x) \\
& + \cdots,
\end{aligned}
$$
(2.6)

and,

$$
\begin{aligned}
u_1(t, x) = {} & \left(\frac{-i\delta}{\varepsilon}\right) \int_0^t dt_1 \mathscr{U}_1(t, t_1) \, V_{10} \mathscr{U}_0(t_1, 0) \, u_0(0, x) \\
& + \left(\frac{-i\delta}{\varepsilon}\right)^3 \int_0^t dt_3 \int_0^{t_3} dt_2 \int_0^{t_2} dt_1 \mathscr{U}_1(t, t_3) \, V_{10} \mathscr{U}_0(t_3, t_2) \, V_{01} \mathscr{U}_1(t_2, t_1) \, V_{10} \times \\
& \hspace{6cm} \times \mathscr{U}_0(t_1, 0) \, u_0(0, x) \\
& + \cdots.
\end{aligned}
$$
(2.7)

Note that the sequence $\{t_j\}_{j=1}^n$ corresponds to the time that the non-zero entries in the wave function are swapped; in between the consecutive t_j's, the wave function propagates on one of the diabatic energy surfaces, and thus the switching can be viewed as a hop from one energy surface to the other. The series (2.6) and (2.7) thus take into account the contributions from even number of hops (including no hop) and odd number of hops, respectively.

To make the above observation more transparent, let us consider the first several terms of both series. The first term of (2.6) $\mathscr{U}_0(t, 0) \, u_0(0, x)$ represents the contribution of 0 hops, since it corresponds to wave packets starting on diabatic surface 0, following the dynamic on surface 0 till the final time t. For the first term of (2.7), for a fixed t_1, the integrand corresponds to wave packets, initially starting on surface 0, propagating along the same potential energy surface from time 0 to t_1, "hopping" to surface 1 at time t_1 according to some hopping coefficent $\frac{i\delta}{\varepsilon} V_{01}$ (to be made precise in the next section), and then propagating on surface 1 till the final time t. Similarly, the second term of (2.6) contributes to hopping twice, with hopping times t_1 and t_2, and etc.

Therefore, the total wave function is a summation of the contributions of all number of hops:

$$
u(t, x) = \begin{pmatrix} u_0^{(0)}(t, x) + u_0^{(2)}(t, x) + u_0^{(4)}(t, x) + \cdots \\ u_0^{(1)}(t, x) + u_0^{(3)}(t, x) + u_0^{(5)}(t, x) + \cdots \end{pmatrix},
$$
(2.8)

where $u_0^{(n)}(t, x)$ denotes the contribution initiated at surface 0, with n hoppings by time t, which contains an integral with respect to all possible ordered hopping times $T_{n:1} := (t_n, \cdots, t_1)$. It is a good place to point out that this leads to a high dimensional integral to calculate $u_0^{(n)}$ and thus u, not mention the high dimensional phase space integral already involved in the semiclassical approximation. To deal with

such high dimensional integrals numerically, it is natural to design a Monte Carlo scheme, combined with semiclassical approximation for the integrand. This is exactly the idea behind the surface hopping algorithm in this work.

The series expansion (2.6) converges absolutely for any fixed ε and δ provided $V_{01} \in L_x^\infty(\mathbb{R}^m)$, since we only integrate over ordered time sequence and thus a combinatorial factor in the denominator as the number of hops (off-diagonal M) increases. To be specific, one could easily do a loose estimate on each term in this series:

$$
(2.9) \qquad \left\| u_0^{(n)} \right\|_{L_x^2(\mathbb{R}^m)} \le \left(\frac{\delta}{\varepsilon} \right)^n \int_0^t dt_n \int_0^{t_n} dt_{n-1} \cdots \int_0^{t_2} dt_1 \, \| V_{01} \|_{L_x^\infty(\mathbb{R}^m)}^n \, \| u_0(0,x) \|_{L_x^2(\mathbb{R}^m)}
$$

$$
= \left(\frac{\delta}{\varepsilon} \right)^n \frac{t^n}{n!} \, \| V_{01} \|_{L_x^\infty(\mathbb{R}^m)}^n \, \| u_0(0,x) \|_{L_x^2(\mathbb{R}^m)},
$$

which implies the absolute convergence of (2.8) for any fixed ε and δ, due to the dominant convergence theorem. On the other hand, for larger δ or smaller ε, we need to sum over more terms in the series expansion to make sure that the remainder is below certain threshold. While we do not sum over the series deterministically term by term in our algorithm, this is still a good indication that the numerical algorithm would face challenge when δ/ε is large, for example, if we consider fix δ and send ε to 0. More precisely, in our algorithm, a large δ/ε would lead to a large variance of the estimator and thus requires increasing number of surface hopping trajectories. In fact, in the literature, it is well known that the diabatic surface hopping encounters difficulties when the off-diagonal term is large compared with the semiclassical parameter, see e.g., [20]. Thus this observation is perhaps not limited to our own version of the surface hopping algorithm, but can be viewed as a generic quantitative difficulty index for the diabatic surface hopping.

2.3. **A diabatic surface hopping algorithm based on frozen Gaussian approximation.** The diabatic surface hopping algorithm we propose is based on a stochastic summation strategy of the series expansion in the last subsection. The asymptotic justification of the algorithm is postponed to Section 4. An overview of the algorithm goes as follows: We initiate an ensemble of trajectories according to the initial datum $u(0,x)$; for each trajectory, it evolves on one of the diabatic energy surfaces, and stochastically hops to the other diabatic surface from time to time; finally, when all the trajectories are evolved to the final desired time, the wave function (or more practically the desired observable) is estimated by taking into account the information from all the trajectories. We will describe below the two key components of the algorithm:

(1) stochastic dynamics of the trajectories; and
(2) estimation of the wave function at the final simulation time.

The initial sampling of the trajectories is analogous to the adiabatic case, so we refer the readers to [17, 16] for details, which we omit here.

Each trajectory evolves in the extended phase space and is thus characterized by position Q_t, momentum P_t, and surface index l_t to keep track of the current potential energy surface. The equation of motion of (Q,P) is given by

$$
(2.10) \qquad \frac{d}{dt} Q = P,
$$

$$
\frac{d}{dt} P = -\nabla V_{l_t l_t}(Q),
$$

where the force depends on the current surface l_t. The surface index l_t follows a Markov jump process on the state space $\{0, 1\}$, where infinitesimal transition probability is given by

$$(2.11) \qquad P\left(l_{t+\delta t} = k \mid l_t = j, Q_t = q\right) = \delta_{jk} + \lambda_{jk}\left(q\right)\delta t + o\left(\delta t\right),$$

with transition rate

$$(2.12) \qquad \lambda\left(q\right) = \begin{pmatrix} \lambda_{00}\left(q\right) & \lambda_{01}\left(q\right) \\ \lambda_{10}\left(q\right) & \lambda_{11}\left(q\right) \end{pmatrix} = \begin{pmatrix} -\frac{\delta}{\varepsilon}\left|V_{01}\left(q\right)\right| & \frac{\delta}{\varepsilon}\left|V_{01}\left(q\right)\right| \\ \frac{\delta}{\varepsilon}\left|V_{10}\left(q\right)\right| & -\frac{\delta}{\varepsilon}\left|V_{10}\left(q\right)\right| \end{pmatrix}.$$

Here the modulus of V_{01} is taken since we interpret that as a non-negative rate. Therefore, (Q_t, P_t, l_t) can be viewed as a Markov switching process whose probability distribution $F_t\left(q, p, l\right)$ follows the following Kolmogorov forward equation

$$(2.13) \qquad \partial_t F_t\left(q, p, l\right) + p \cdot \nabla_q F_t\left(q, p, l\right) - \nabla_p V_{ll}(q) \cdot \nabla_p F_t\left(q, p, l\right) = \sum_{j=0}^{l} \lambda_{jl}\left(q\right) F_t\left(q, p, j\right),$$

where the last two terms of the left hand side is induced by the Hamiltonian flow of (Q_t, P_t), and the right hand side stands for the contribution of the Markov jump process. Note that l_t is a piecewise constant and almost surely each trajectory has finite number of jumps. Consider a realization of a trajectory with n jumps, we will denote the discontinuity set of l_t as $\{t_j\}_{j=1}^n$, which is the ordered set of time of jumps of the trajectory.

Along each trajectory, we also solve for the classical action S_t and weighting factor A_t given by

$$(2.14) \qquad \begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}S &= \frac{1}{2}|P|^2 - V_{l_t l_t}\left(Q\right), \\ \frac{\mathrm{d}}{\mathrm{d}t}A &= \frac{1}{2}A\operatorname{tr}\left(Z^{-1}\left(\partial_z P - i\partial_z Q\nabla_Q^2 V_{l_t l_t}\left(Q\right)\right)\right), \end{aligned}$$

where we have used the short hand

$$\partial_z := \partial_q - i\partial_p, \quad \text{and} \quad Z := \partial_z(Q + iP).$$

For the trajectory with initial motion $(Q_0, P_0) = \left(q_0, p_0\right)$, the initial action and amplitude are assigned as

$$(2.15) \qquad S_0 = 0 \quad \text{and} \quad A_0 = 2^{m/2}\int_{\mathbb{R}^m} u_0\left(0, y\right) e^{\frac{i}{\varepsilon}\left(-p_0\left(y-q_0\right)+\frac{i}{2}|y-q_0|^2\right)}\,dy,$$

where the integral for A_0 is either analytically given or numerically estimated.

Finally, the wave function is reconstructed via a trajectory average, that is, an expectation over all of the trajectories,

$$(2.16) \qquad \tilde{u}\left(t, x\right) = C_{\mathcal{N}}\mathbb{E}\left[(-i)^n\left(\prod_{j=1}^n \frac{V_{l_{t_j} l_{t_{j-1}}}\left(Q_{t_j}\right)}{\left|V_{l_{t_j} l_{t_{j-1}}}\left(Q_{t_j}\right)\right|}\right)\frac{A_t}{|A_0|}\exp\left(\frac{i}{\varepsilon}\Theta_t\right)\exp\left(\int_0^t \frac{\delta}{\varepsilon}|V_{01}\left(Q_s\right)|ds\right)\begin{pmatrix} \mathbb{1}_{n \text{ is even}} \\ \mathbb{1}_{n \text{ is odd}} \end{pmatrix}\right],$$

where the normalization constant reads

$$(2.17) \qquad C_{\mathcal{N}} = \frac{1}{(2\pi\varepsilon)^{3m/2}}\int_{\mathbb{R}^{2m}}|A_0\left(q_0, p_0\right)|dq_0 dp_0,$$

and we have used $\mathbb{1}$ to denote the indicator function and

$$(2.18) \qquad \Theta_t = S_t + P_t\left(x - Q_t\right) + \frac{i}{2}|x - Q_t|^2.$$

While the expression (2.16) appears complicated, all the terms involved in the expectation are obtained along the trajectories. In the algorithm, the expectation over path ensemble will be estimated by Monte Carlo algorithm sampling the path space, as will be explained in more details below.

3. NUMERICAL IMPLEMENTATION AND EXAMPLES

3.1. **Numerical implementation.** Numerical description of our scheme is as follows: Given initial datum $u_{\mathrm{in}}(x) = [u_0(0,x)\ 0]^T$, one first calculates the initial amplitude according to (2.15) by numerical quadratures (or even analytically, for special initial data). One samples the initial position and momentum according to $|A_0(q_0, p_0)|$, viewed as a joint density function of (q_0, p_0). Each pair of (q_0, p_0) corresponds to a trajectory, with the initial position, momentum and action given as

$$(3.1) \qquad\qquad Q_0 = q_0, \ \ P_0 = p_0, \ \ S_0 = 0.$$

Then one marches the ODE system (2.10) and (2.14) by a ODE integrator (high-order ones are preferred and we choose here the fourth order Runge-Kutta scheme). For each time step $t = t^k$, the hopping probability of this time step is calculated as $\Delta t \delta |V_{01}(Q_{t^k})|$, where Δt is the time step size chosen sufficiently small. We thus draw a random number uniformly distributed on the interval $[0, 1]$; a hop happens if the probability is larger than the drawn random number, upon which we switch the potential energy surface index ℓ_t, with all of the trajectory coefficients (Q_t, P_t, S_t, A_t) being continuous. Moreover, at each of the hopping time t_n, we record the hopping coefficients $V_{01}(Q_{t_n})$. Repeat this procedure until the desired final time, the wave function (or expectation of an observable) is estimated according to the trajectory average (2.16).

Before we turn to numerical examples, let us discuss the possible source of errors of the numerical solution compared to the exact solution to the Schrödinger equation:

  (i) Initial error, coming from the numerical estimation of A_0, if analytic expression is not available. Note that if the phase space is high dimensional, numerical quadrature to obtain A_0 is not feasible, and hence we would need to restrict to specific initial data such that analytic expression is available.

 (ii) Asymptotic error, introduced by the semiclassical approximation of the evolution on single diabatic surfaces, where we neglect terms with orders higher than ε in the frozen Gaussian approximation. This error is of order $\exp\left(\frac{\delta T}{\varepsilon}\right)\varepsilon$, where T is the final time (see Section 4.2 for further discussion).

(iii) Numerical integration error of the ODEs. This error can be controlled by choosing high order ODE solvers. In our case the fourth order Runge-Kutta method seems to provide good numerical accuracy.

(iv) Sampling error, incurred by the application of the Monte Carlo method. This error decays as $N^{-1/2}$, where N denotes the number of trajectories. A numerical validation of the sampling error is given in Fig. 1. Typically the sampling error dominates the total error of the numerical algorithm.

3.2. **Numerical examples.** In this section, the numerical examples are chosen in the spirit of the three standard test cases first considered in Tully's paper [25]. The reference solutions are all computed by a
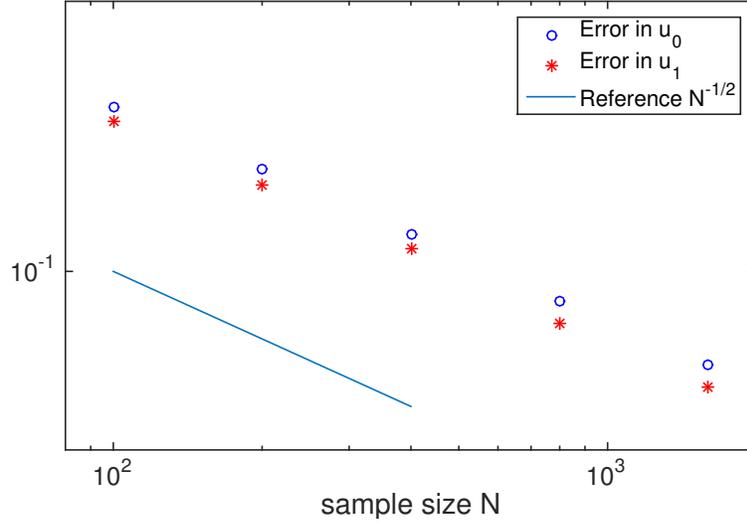
FIGURE 1. Sampling error incurred by the Monte Carlo method. Here ε is fixed as 0.04, and test on various numbers of trajectories $N = 100, 200, 400, 800, 1600$. The plot confirms the convergence rate of $N^{-1/2}$ and that the sampling error dominates the total numerical error.

time-splitting pseudo-spectral scheme [1] (this is possible as these test cases are in low spatial dimension), with the meshing strategy $\Delta t = o(\varepsilon)$, $\Delta x = o(\varepsilon)$ and refined enough to guarantee that the reference solution is close to the true solution.

Example 3.1 (Simple Crossing). Consider the matrix potential

$$(3.2) \qquad \begin{pmatrix} \tanh x & \delta \\ \delta & -\tanh x \end{pmatrix},$$

whose diabatic potential energy surfaces is shown in Fig. 2a. In our test, we choose $\varepsilon = \delta = 0.04$, the final time $T = 1.2$, and the initial datum given by

$$u_0(0, x) = e^{-12.5(x+1.5)^2} e^{\frac{2i}{\varepsilon}(x+1.5)}.$$

The number of trajectories N is 5000. It can be seen in Fig. 2b that the numerical solutions match nicely with the reference solutions. This validates that our algorithm can correctly reconstruct the wave function.

Example 3.2 (Dual Crossing). In the second example, we consider the case of dual crossing; it is a more challenging example as it exhibits two avoided crossing. So to get the correct solution, the algorithm has to capture the phase of the wave function during the evolution. The matrix potential reads

$$\begin{pmatrix} 0 & 0.015e^{-0.06x^2} \\ 0.015e^{-0.06x^2} & -0.1e^{-0.28x^2} + 0.05 \end{pmatrix}.$$

We take in this example $\varepsilon = \frac{1}{\sqrt{2000}}$, and initial datum

$$u_0(0, x) = e^{-\sqrt{500}(x+2.5)^2} e^{\frac{2i}{\varepsilon}(x+2.5)}.$$
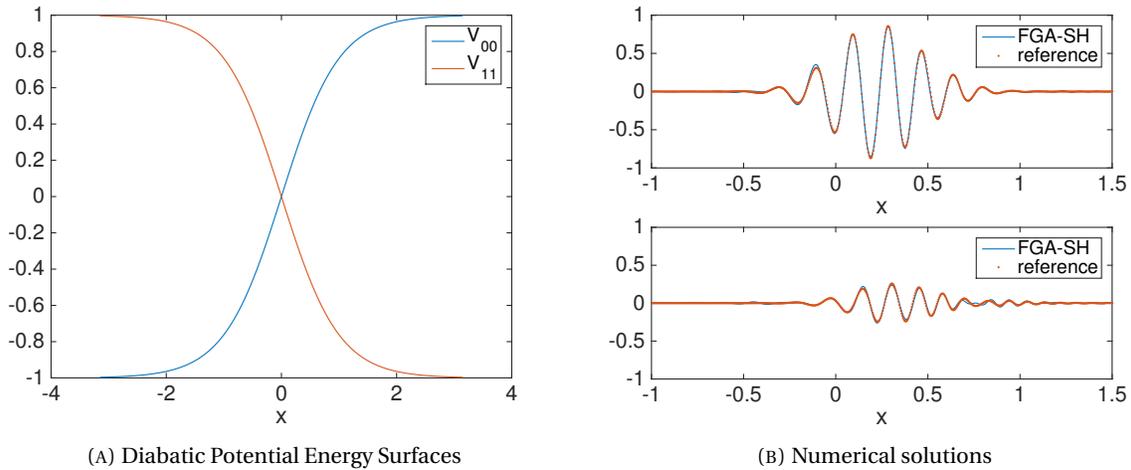
(A) Diabatic Potential Energy Surfaces

(B) Numerical solutions

FIGURE 2. Simple Crossing. (Left panel): the diabatic potential energy surfaces. (Right panel): the numerical solutions (blue solid) of u_0 (top) and u_1 (bottom) agree well with the reference solutions (red dashed). The reference solutions is computed by the time-splitting pseudo-spectral scheme with $\Delta t = \frac{\varepsilon}{32}$ and $\Delta x = \frac{(b-a)\varepsilon}{32}$.

We run the algorithm until the wave packets pass both crossing areas around $T = 2.2$. The number of the trajectories is 10000. For this example, we observe again that the numerical and reference solutions are in good agreement.
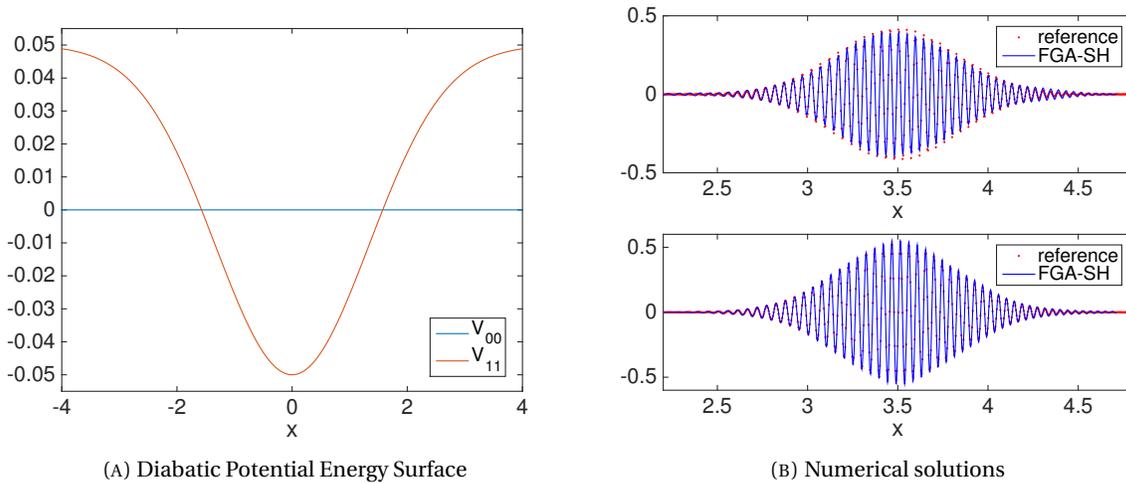

(A) Diabatic Potential Energy Surface

(B) Numerical solutions

FIGURE 3. Dual Crossing. (Left panel): the diabatic potential energy surfaces. (Right panel): the numerical solutions (blue solid) of u_0 (top) and u_1 (bottom) agree well with the reference solutions (red dashed). The reference solutions is computed by the time-splitting pseudo-spectral scheme with $\Delta t = \frac{\varepsilon}{32}$ and $\Delta x = \frac{(b-a)\varepsilon}{32}$.

Example 3.3 (Extended Coupling with Reflection)**.** In this example, we consider the matrix potential

$$\begin{pmatrix} \arctan(10x) + \frac{\pi}{2} & \delta \\ \delta & -\arctan(10x) - \frac{\pi}{2} \end{pmatrix}.$$

This is often considered as a challenging example, since it contains an extended region of strong nona-diabatic coupling. An interesting feature of this example is that the wave packet evolves on the upper surface would be partially reflected (bounced back), as shown in Fig. 4b (note the two wave packets on the upper diabatic energy surface). Consider $\varepsilon = \delta = 0.04$, the final time $T = 1.4$, and the initial data same as in Example 3.1. The number of the trajectories is 30000. It can be seen in Fig. 4 that our algorithm still keeps great agreement with the reference solutions. Both the transmitted and reflected wave packets on the second energy surface are accurately captured.
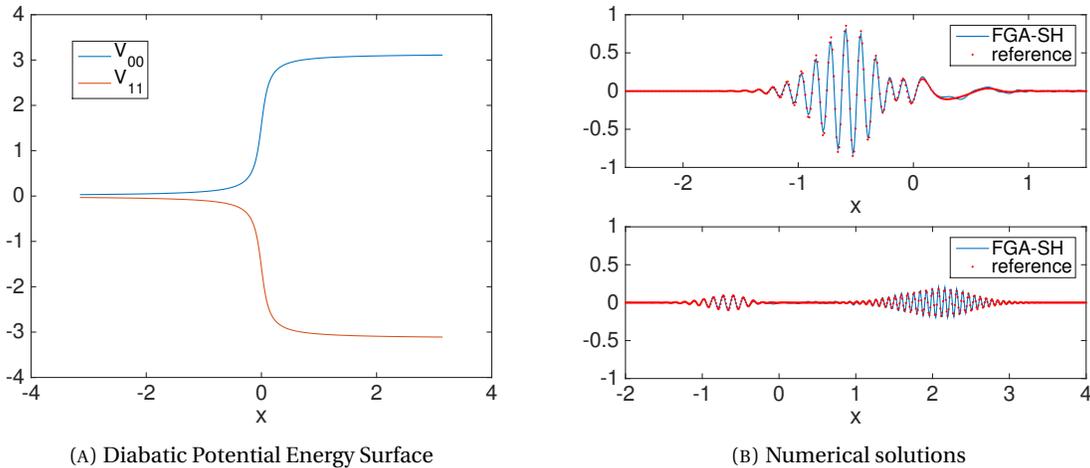

(A) Diabatic Potential Energy Surface   (B) Numerical solutions

FIGURE 4. Extended Coupling with Reflection. (Right panel): the numerical solutions (blue solid) of u_0 (top) and u_1 (bottom) show good agreements with the reference solutions (red dashed).

Example 3.4 (Weak Coupling Regime)**.** In this example, we focus on the weak coupling regime when $\delta \to 0$ and ε fixed. This regime is related to the Marcus theory [18] in the theoretical chemistry literature, it is particularly interesting since as shown in [11], the original Tully's FSSH does not capture the right asymptotic limit. Let us define the transition rate as

$$(3.3) \qquad\qquad R = \frac{\int_{\mathbb{R}^m} |u_1(T,x)|^2 \, dx}{\int_{\mathbb{R}^m} |u_0(0,x)|^2 \, dx},$$

where T is some time when the wave packets bypass the transition zone. The Marcus theory states that the transition rate is proportional to the square of the off diagonal term $\delta|V_{01}|$. Unfortunately, as is shown in [11], the standard FSSH gives the rate proportional to $\delta|V_{01}|$ instead of $\delta^2|V_{01}|^2$. Our algorithm is able to capture the correct asymptotic limit, as shown in Fig. 5; the numerically predicted transition rate is quadratic in $\delta|V_{01}|$. In this test, we take the matrix potential (3.2) with various δ and initial datum

$$u_0(0,x) = e^{-12.5(x+1)^2} e^{\frac{2i}{\varepsilon}(x+1)}$$
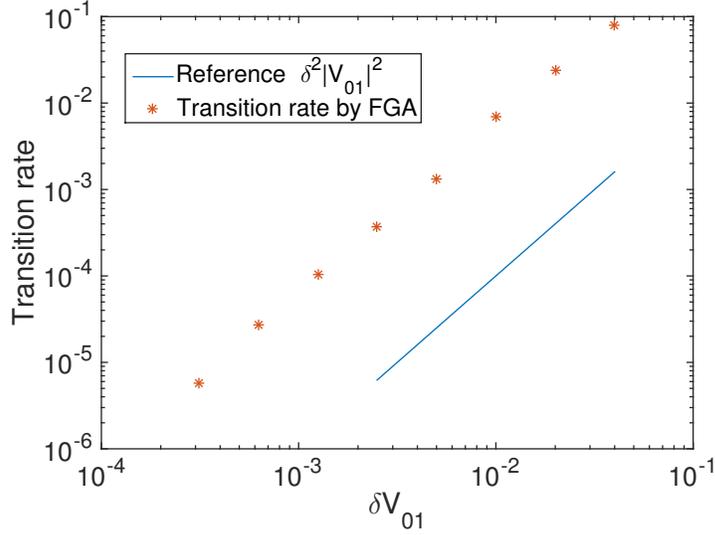
for the final time $T = 1$.

FIGURE 5. Weak Coupling Regime. The transition rate computed by our numerical scheme (red $*$) grows proportional to $\delta^2 |V_{01}|^2$, which agrees with the Marcus theory.

Example 3.5 (Landau-Zener Regime)**.** We shall now test the performance of our algorithm in the Landau-Zener regime as in [5], i.e., when the off-diagonal term of the potential matrix V is $O(\varepsilon^{\frac{1}{2}})$. Consider the matrix potential

(3.4)
$$\begin{pmatrix} \tanh x & \varepsilon^{\frac{1}{2}} \\ \varepsilon^{\frac{1}{2}} & -\tanh x \end{pmatrix},$$

with the semiclassical parameter $\varepsilon = 0.04$, final time $T = 0.5$, and initial datum given by

$$u_0(0, x) = e^{-12.5(x+1)^2} e^{\frac{3i}{\varepsilon}(x+1)}.$$

We take 20000 trajectories in the numerical test. As shown in Fig. 6, our scheme still produces good agreement with the reference solution. On the other hand, compared to Example 3.1, where the off-diagonal terms of the matrix potential has magnitude $O(\varepsilon)$, we do need significantly more trajectories to achieve similar accuracy. We thus further study in the next two examples the relation of the number of trajectories versus the ratio δ/ε.

Example 3.6 (Number of Trajectories versus $\frac{\delta}{\varepsilon}$)**.** As have seen in the previous example, more trajectories are needed as $\frac{\delta}{\varepsilon}$ becomes larger. In this example, we consider fixing $\varepsilon = 0.02$ and increasing the values of δ. The potential used in this example is (3.2) (the previous example corresponds to $\delta = \varepsilon^{1/2}$) and initial datum

$$u_0(0, x) = e^{-25(x+.4)^2} e^{\frac{2i}{\varepsilon}(x+.4)}$$

for the time $T = 0.5$. We record the number of trajectories needed for the L^2 error of the wave function to achieve certain fixed error threshold. One can see that the number of trajectories grows with respect to $\frac{\delta}{\varepsilon}$ exponentially. As predicted before, it becomes more challenging as $\frac{\delta}{\varepsilon}$ become larger, and one would need more trajectories to get accurate wave function.
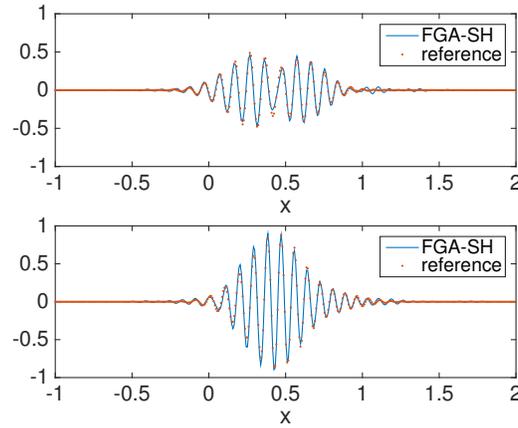
FIGURE 6. Landau-Zener Regime. The numerical solutions (blue solid) show good agreements with the reference solutions (red dashed).
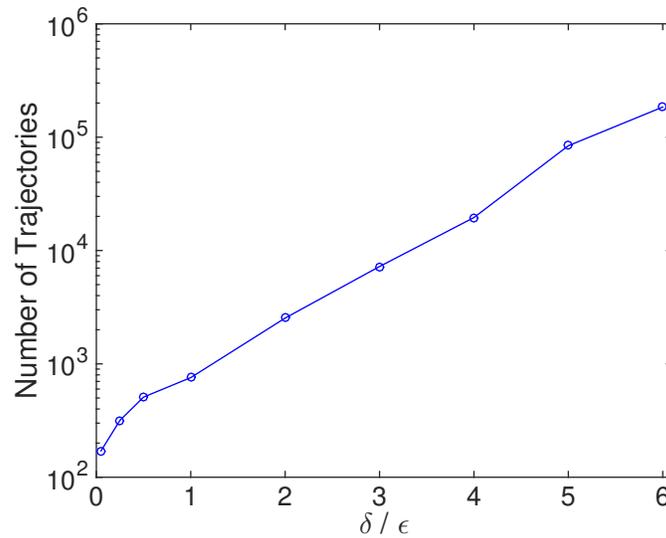

FIGURE 7. Number of trajectories to achieve certain threshold with various $\frac{\delta}{\varepsilon}$ in semi-log scale. Here we fix $\varepsilon = 0.02$, and choose $\delta = 0.001, 0.005, 0.01, 0.02, 0.04, 0.06, 0.08, 0.1, 0.12$. The number of trajectories grows exponentially with respect to δ to achieve the same L^2 relative error threshold 0.08.

Example 3.7 (Avoided Crossing Regime with varying ε)**.** To continue, we test our algorithm in the set-ting of avoided crossing as studied in [5], where δ is chosen as the order of $\sqrt{\varepsilon}$. This particular choice of δ generates a family of problems, where the transition rate gives an order one contribution due to the Landau-Zener formula. As mentioned before, the diabatic representation is most useful inside the transition zone where the off-diagonal terms are relatively small, while adiabatic picture can be applied outside the transition zone. In this example, the propagation of the solution near the transition zone is considered. Since the gap is of order $\sqrt{\varepsilon}$, the time for the wave packets to pass the transition zone is of

the same order. In our numerical tests, for $\varepsilon = 0.04, 0.01, 0.0025, 0.00125$, one considers $\delta = \sqrt{\varepsilon}$, initial datum

$$u_0(0, x) = e^{-\frac{1}{2\varepsilon}(x-\mu_q)^2} e^{\frac{2i}{\varepsilon}(x-\mu_q)}, \tag{3.5}$$

where $\mu_q = -2\sqrt{\varepsilon}$, and run till the final time $T = 3\sqrt{\varepsilon}$, where the wave packet passes the transition zone. As can be observed from Fig. 8, the number of trajectories needed does not increase noticeably as ε decreases, which shows the validity of the algorithm for small ε.
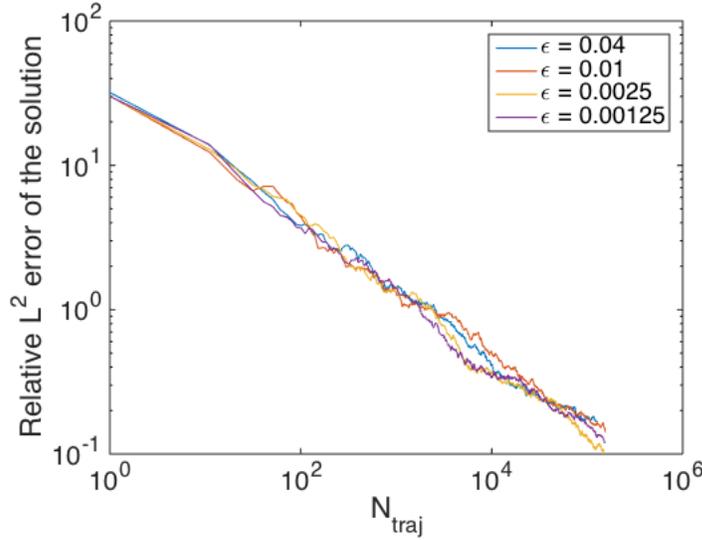

FIGURE 8. Relative L^2 error of the wave function versus the number of trajectories for various ε in the log-log scale. It can be seen that as ε decreases, the number of trajectories does not increase noticeably for the wave packets to pass the transition zone.

4. DERIVATION OF THE ALGORITHM

4.1. **A brief review of the frozen Gaussian approximation.** Recall that our algorithm is based on semi-classical approximation of the propagation on a single diabatic surface, together with the surface hopping to stochastically evaluate the series expansion. In the FGA-SH method, the semiclassical approximation is based on the frozen Gaussian approximation (*aka* Herman-Kluk propagator [22]), which is a convergent semiclassical propagator of scalar Schrödinger equation. The Herman-Kluk propagator is justified in [7, 8, 24] and further applied to high frequency wave propagations [14, 15]. The idea of FGA is to decompose the initial data into a group of localized Gaussian wave packets with fixed width, and then to propagate those Gaussians along classical trajectories with evolving weights. To be more specific, let us recall here the FGA integral representation for the scalar Schrödinger equation

$$i\varepsilon\partial_t u(t, x) = -\frac{\varepsilon^2}{2}\Delta_x u(t, x) + W(x) u(t, x), \ u(0, x) = u_{\text{in}}(x), \qquad (t, x) \in \mathbb{R}^+ \times \mathbb{R}^m$$

where W is the potential (to distinguish notations from V reserved for the matrix potential). The FGA ansatz is given by

$$u_F(t, x) = \frac{1}{(2\pi\varepsilon)^{3m/2}} \int_{\mathbb{R}^{2m}} A_t(q_0, p_0) \exp\left(\frac{i}{\varepsilon}\Theta_t(q_0, p_0, x)\right) dq_0 dp_0,$$

where

$$(4.1) \qquad \Theta_t(q_0, p_0, x) = S_t(q_0, p_0) + P_t(q_0, p_0)(x - Q_t(q_0, p_0)) + \frac{i}{2}|x - Q_t(q_0, p_0)|^2.$$

Notice u_F can be viewed as a summation of localized Gaussian wave packets. The FGA coefficients (Q_t, P_t, S_t, A_t) satisfy the following ODEs:

$$(4.2) \qquad \frac{\mathrm{d}}{\mathrm{d}t}Q = P,$$

$$(4.3) \qquad \frac{\mathrm{d}}{\mathrm{d}t}P = -\nabla W(Q),$$

$$(4.4) \qquad \frac{\mathrm{d}}{\mathrm{d}t}S = \frac{1}{2}|P|^2 - W(Q),$$

$$(4.5) \qquad \frac{\mathrm{d}}{\mathrm{d}t}A = \frac{1}{2}A\operatorname{tr}\left(Z^{-1}(\partial_z P - i\partial_z Q\nabla_Q^2 W(Q))\right),$$

where we have used the shorthands $\partial_z = \partial_q - i\partial_p$ and $Z = \partial_z(Q + iP)$. To assign the initial data of the FGA coefficients, we quote [14, Lemma 3.1]:

Lemma 4.1. *For $u \in L^2(\mathbb{R}^m)$, it holds*

$$(4.6) \qquad u(x) = \frac{1}{(2\pi\varepsilon)^{3m/2}} \int_{\mathbb{R}^{3m}} 2^{m/2} u(y) \exp\left(\frac{i}{\varepsilon}\left(-p(y - q) + \frac{i}{2}|y - q|^2 + p(x - q) + \frac{i}{2}|x - q|^2\right)\right) dy\,dq\,dp.$$

As a direct consequence, the initial conditions of the action and weight of each trajectory can be given as (2.15). The L^2 error of this FGA ansatz compared to the exact solution of the scalar Schrödinger equation is proved to be of first order in ε (see, e.g., [24]).

4.2. **A generalization of FGA to diabatic surface hopping.** In this section, we "translate" the terms of the series (2.6)–(2.8) into a deterministic asymptotic representation of the wave function of (2.1) by using the frozen Gaussian approximation as the semiclassical approximation.

It is clear that the first term of (2.8) needs no "translation". In other words, $u_0^{(0)}(t, x) = \mathcal{U}_0(t, 0) u_0(0, x)$ is exactly the same as the scalar case with the potential $V_{00}(x)$, so we can directly apply the frozen Gaussian approximation as in the previous subsection.

Now consider

$$(4.7) \qquad u_0^{(1)}(t, x) = \frac{-i\delta}{\varepsilon} \int_0^t dt_1 \mathcal{U}_1(t, t_1) V_{10}(x)\mathcal{U}_0(t_1, 0) u_0(0, x)$$

and focus on the integrand at a specific t_1. If we view $\mathcal{U}_1(t, t_1)$ as the propagator on the second diabatic surface and

$$(4.8) \qquad V_{10}(x)\mathcal{U}_0(t_1, 0) u_0(0, x)$$

as the initial data at time t_1, the propagation from time t_1 to t can then be written in a FGA form

$$(4.9) \qquad \frac{1}{(2\pi\varepsilon)^{3m/2}} \int_{\mathbb{R}^{2m}} \hat{A}_t(q_0, p_0) \exp\left(\frac{i}{\varepsilon}\tilde{\Theta}_t(q_0, p_0, x)\right) dq_0 dp_0 + O(\varepsilon),$$

where $\tilde{\Theta}_t$ is defined as (4.1) and the trajectory coefficients $\tilde{Q}_t, \tilde{P}_t, \tilde{S}_t$ and $\hat{A}_t$ evlove according to the ODE system (4.2) with potential energy $W = V_{11}$ from time t_1 to t. To determine the value of those at time t_1 (the initial time of the propagator $\mathscr{U}(t, t_1)$), let us represent (4.8) in a FGA form by evolving the data at time $t = 0$ to time t_1. In particular, the last two terms $\mathscr{U}_0(t_1, 0)u_0(0, x)$ is nothing but the scalar case with potential $W = V_{00}$, and hence can be approximated using the scalar FGA ansatz:

$$\mathscr{U}_0(t_1, 0)\, u_0(0, x) = \frac{1}{(2\pi\varepsilon)^{3m/2}} \int_{\mathbb{R}^{2m}} A_{t_1}(q_0, p_0) \exp\left(\frac{i}{\varepsilon}\Theta_{t_1}(q_0, p_0, x)\right) dq_0 dp_0 + O(\varepsilon).$$

For $V_{10}(x)$, expanding around $x = Q_{t_1}(q_0, p_0)$, we get

$$V_{10}(x)\mathscr{U}_0(t_1, 0)\, u_0(0, x) =$$

$$= \frac{1}{(2\pi\varepsilon)^{3m/2}} \int_{\mathbb{R}^{2m}} V_{10}(Q_{t_1}(q_0, p_0)) A_{t_1}(q_0, p_0) \exp\left(\frac{i}{\varepsilon}\Theta_{t_1}(q_0, p_0, x)\right), dq_0 dp_0$$

$$+ \frac{1}{(2\pi\varepsilon)^{3m/2}} \int_{\mathbb{R}^{2m}} (x - Q_{t_1}(q_0, p_0)) \cdot \nabla V_{10}(Q_{t_1}(q_0, p_0)) A_{t_1}(z_0) \exp\left(\frac{i}{\varepsilon}\Theta_{t_1}(q_0, p_0, x)\right) dq_0 dp_0$$

$$+ \frac{1}{2(2\pi\varepsilon)^{3m/2}} \int_{\mathbb{R}^{2m}} (x - Q_{t_1}(q_0, p_0)) \cdot \nabla^2 V_{10}(Q_{t_1}(q_0, p_0))(x - Q_{t_1}(q_0, p_0)) \times$$

$$\times A_{t_1}(q_0, p_0) \exp\left(\frac{i}{\varepsilon}\Theta_{t_1}(q_0, p_0, x)\right) dq_0 dp_0 + O(\varepsilon^2).$$

According to [14, lemma 3.2], the second and third terms are both $O(\varepsilon)$. Hence,

$$V_{10}(x)\mathscr{U}_0(t_1, 0)\, u_0(0, x) =$$

$$= \frac{1}{(2\pi\varepsilon)^{3m/2}} \int_{\mathbb{R}^{2m}} V_{10}(Q_{t_1}(q_0, p_0)) A_{t_1}(q_0, p_0) \exp\left(\frac{i}{\varepsilon}\Theta_{t_1}(q_0, p_0, x)\right), dq_0 dp_0 + O(\varepsilon).$$

We now match the above formula with (4.9) at time t_1, which leads to the transition conditions of the trajectory coefficients as

$$(4.10) \qquad \tilde{Q}_{t_1}(q_0, p_0) = Q_{t_1}(q_0, p_0),$$

$$(4.11) \qquad \tilde{P}_{t_1}(q_0, p_0) = P_{t_1}(q_0, p_0),$$

$$(4.12) \qquad \tilde{S}_{t_1}(q_0, p_0) = S_{t_1}(q_0, p_0),$$

$$(4.13) \qquad \hat{A}_{t_1}(q_0, p_0) = V_{10}(Q_{t_1}(q_0, p_0)) A_{t_1}(q_0, p_0).$$

To simplify the final expression, let us define $\tilde{A}_t$ such that

$$(4.14) \qquad \hat{A}_t(q_0, p_0) = V_{10}(Q_{t_1}(q_0, p_0)) \tilde{A}_t(q_0, p_0).$$

Since given q_0, p_0 and t_1, $\tilde{A}_t$ only differs with $\hat{A}_t$ by a constant factor (independent of t), $\tilde{A}_t$ follows the same ODE as $\hat{A}_t$ with initial condition

$$(4.15) \qquad \tilde{A}_{t_1}(q_0, p_0) = A_{t_1}(q_0, p_0).$$

Hence, one could drop the "~" and view (Q_t, P_t, S_t, A_t) as a continuous trajectory at hopping time $t = t_1$. In sum, the FGA ansatz of $u_0^{(1)}$ reads

$$(4.16) \quad u_F^{(1)}(t, t_1, q_0, p_0) = \frac{-i\delta}{\varepsilon} \frac{1}{(2\pi\varepsilon)^{3m/2}} \int_0^t dt_1 \int_{\mathbb{R}^{2m}} dq_0 dp_0 V_{10}(Q_{t_1}(q_0, p_0)) \times$$

$$\times A_t(q_0, p_0) \exp\left(\frac{i}{\varepsilon}\Theta_t(q_0, p_0, x)\right),$$

where the quantities in the integrand are defined as above (solving the ODE systems of (4.2)–(4.5) with the potential $W = V_{00}$ from time 0 to t_1 and with the potential $W = V_{11}$ from time t_1 to t, with gluing conditions (4.10)–(4.12) and (4.15).

Similarly, the other terms can be represented as FGA ansatz, by applying consecutively the semiclassical approximation to each propagators. In sum, the FGA surface hopping ansatz reads

$$
(4.17) \qquad u_{\mathrm{F}}(t, x) = \begin{pmatrix} u_{\mathrm{F}}^{(0)}(t, x) + u_{\mathrm{F}}^{(2)}(t, x) + u_{\mathrm{F}}^{(4)}(t, x) + \cdots \\ u_{\mathrm{F}}^{(1)}(t, x) + u_{\mathrm{F}}^{(3)}(t, x) + u_{\mathrm{F}}^{(5)}(t, x) + \cdots \end{pmatrix},
$$

with the FGA ansatz for $u_{\mathrm{F}}^{(n)}$ generalized as

$$
(4.18) \qquad u_{\mathrm{F}}^{(n)}(t, x) = \left(\frac{-i\delta}{\varepsilon} \right)^n \frac{1}{(2\pi\varepsilon)^{3m/2}} \int_{\mathbb{R}^{2m}} dq_0 dp_0 \int_{0 < t_1 < \cdots < t_n < t} dT_{n:1} \prod_{j=1}^{n} V_{l_{t_j} l_{t_{j-1}}}(Q_{t_j}) A_t \exp\left(\frac{i}{\varepsilon} \Theta_t \right),
$$

where $T_{n:1} := (t_n, \cdots, t_1)$ and the phase function Θ_t is given by (2.18).

In this asymptotic derivation, the terms of order higher than $O(\varepsilon)$ in the integrands are ignored, and hence the asymptotic error reads

$$
(4.19) \qquad \sum_{n=0}^{\infty} \left(\frac{\delta t}{\varepsilon} \right)^n \frac{1}{n!} \varepsilon = \exp\left(\frac{\delta t}{\varepsilon} \right) \varepsilon.
$$

This shows the challenge of the algorithm when δ is much larger than ε or for very long time.

Remark 4.2. A rigorous estimate on the asymptotic error of the surface hopping ansatz (4.17)–(4.18) is possible following a similar analysis as in [16], we will omit the details here.

4.3. **Stochastic Interpretation.** As is mentioned, the surface hopping ansatz involves integrals with respect to all of the hopping times, and hence motivates a stochastic representation of the integrals. In this section, we shall interpret the surface hopping ansartz (4.17) via the stochastic process proposed in Section 2.3.

Given a time interval $[0, t]$ and s is some time within this interval, the stochastic process is labelled as (Q_s, P_s, l_s). l_s is piecewise constant with almost surely finite many jumps. Consider a realization with initial condition $(Q_0, P_0, l_0) = (q_0, p_0, 0)$ and the number of jumps n, and the discontinuity set of l_s is denoted as $\{t_1, \cdots t_n\}$. By the properties of the stochastic process, the probability of no jump ($n = 0$) reads

$$
(4.20) \qquad \mathbb{P}(n = 0) = \exp\left(\int_0^t \lambda_{00}(Q_s) ds \right) = \exp\left(-\int_0^t \frac{\delta}{\varepsilon} |V_{01}(Q_s)| ds \right).
$$

Similarly the probability with one jump ($n = 1$) is given by

$$
(4.21) \qquad \mathbb{P}(n = 1) = \int_0^t dt_1 \lambda_{01}(Q_{t_1}) \exp\left(-\int_0^t \lambda_{00}(Q_s) ds \right)
$$

$$
(4.22) \qquad = \int_0^t dt_1 \frac{\delta}{\varepsilon} |V_{01}(Q_{t_1})| \exp\left(-\int_0^t \frac{\delta}{\varepsilon} |V_{01}(Q_s)| ds \right)
$$

Generally,

$$
P(n = k) = \int_{0 < t_1 < \cdots < t_k < t} dT_{k:1} \left(\frac{\delta}{\varepsilon} \right)^k \prod_{j=1}^{k} |V_{01}(Q_{t_j})| \exp\left(-\int_0^t \frac{\delta}{\varepsilon} |V_{01}(Q_s)| ds \right),
$$

and for k jumps in total, the probability density of $(t_1, \cdots, t_n)$ is given by

$$
(4.23) \qquad \rho_k(t_1, \cdots, t_k) = \frac{1}{P(n = k)} \left(\frac{\delta}{\varepsilon} \right)^k \prod_{j=1}^{k} |V_{01}(Q_{t_j})| \exp\left(-\int_0^t \frac{\delta}{\varepsilon} |V_{01}(Q_s)| ds \right) \mathbb{1}_{t_1 \le t_2 \le \cdots \le t_k},
$$

where $\mathbb{1}_{t_1 \leq t_2 \leq \cdots \leq t_k}$ is the indicator function of the set $\{(t_1, \cdots, t_k) : t_1 \leq t_2 \leq \cdots \leq t_k\}$.

Now we consider a trajectory average (2.16) and denote everything inside the square bracket of (2.16) as Υ,

$$(4.24) \qquad \Upsilon = (-i)^n \prod_{j=1}^n \frac{V_{l_{t_j} l_{t_{j-1}}}(Q_{t_j})}{|V_{01}(Q_{t_j})|} \frac{A_t}{|A_0|} \exp\left(\frac{i}{\varepsilon}\Theta_t\right) \exp\left(\int_0^t \frac{\delta}{\varepsilon}|V_{01}(Q_s)|ds\right) \begin{pmatrix} \mathbb{1}_{n \text{ is even}} \\ \mathbb{1}_{n \text{ is odd}} \end{pmatrix}$$

and (Q_t, P_t, l_t) as $\tilde{z}_t$ for short. We want to show that (2.16) is a stochastic representation of the FGA ansatz (4.17). The idea is quite straightforward. One first converts the trajectory average into an integral with respect to $z_0 := (q_0, p_0)$, which could be easily achieved by the law of total expectation,

$$\tilde{u}(t, x) = C_{\mathcal{N}} \mathbb{E}_{\tilde{z}_0}\left(\mathbb{E}_{\tilde{z}_t}\left[\Upsilon \mid \tilde{z}_0 = (z_0, 0)\right]\right)$$

$$= \frac{1}{(2\pi\varepsilon)^{3m/2}} \int_{\mathbb{R}^{2m}} dz_0 \mathbb{E}_{\tilde{z}_t}\left[|A_0(z_0)|\Upsilon \mid \tilde{z}_0 = (z_0, 0)\right],$$

where the normalization factor $C_{\mathcal{N}}$ is given as (2.17), and the probability measure $\mathbb{P}_0$ of the single surface reads

$$\mathbb{P}_0(\Omega) = C_{\mathcal{N}}^{-1} \frac{1}{(2\pi\varepsilon)^{3m/2}} \int_{\Omega} dz_0 |A_0(z_0)|.$$

Next, one rewrites the condition expectation. Note that given the initial condition, the randomness only comes from the hopping times.

$$\tilde{u}(t, x) = \frac{1}{(2\pi\varepsilon)^{3m/2}} \int_{\mathbb{R}^{2m}} dz_0 \sum_{n=0}^{\infty} \mathbb{P}(n) \int_{[0,t]^n} dT_{n:1} \, \rho_n(t_1, \cdots, t_n) |A_0(z_0)|\Upsilon$$

$$= \frac{1}{(2\pi\varepsilon)^{3m/2}} \int_{\mathbb{R}^{2m}} dz_0 \sum_{n=0}^{\infty} \int_{0<t_1<\cdots<t_n<t} dT_{n:1} \left(\frac{-i\delta}{\varepsilon}\right)^n \prod_{j=1}^n V_{l_{t_j} l_{t_{j-1}}}(Q_{t_j}) A_t \exp\left(\frac{i}{\varepsilon}\Theta_t\right) \begin{pmatrix} \mathbb{1}_{n \text{ even}} \\ \mathbb{1}_{n \text{ odd}} \end{pmatrix}$$

$$= \begin{pmatrix} u_{\mathrm{F}}^{(0)} + u_{\mathrm{F}}^{(2)} + \cdots \\ u_{\mathrm{F}}^{(1)} + u_{\mathrm{F}}^{(3)} + \cdots \end{pmatrix} = u_{\mathrm{F}}(t, x).$$

This shows that the stochastic path integral representation can capture the asymptotically correct solutions of the matrix Schrödinger equation (2.1).

Remark 4.3. We remark that our previous transition rate (2.12) depends only on the off-diagnal term V_{01}. However, there are some cases when the off-diagnal term V_{01} is large while the gap between two diabatic potential energy surfaces are even larger. Then according to previous transition rate, we would expect excessive amount of hops happen, which, though still gives the correct wave function, may lead to large variance and hence low convergence rate. As one of the main issues in Monte Carlo sampling, how to reduce the variance is always an interesting problem. But the main purpose of this paper is to provide a general framework of the diabatic surface hopping algorithm while efficient ways to control the variance is beyond the scope of this paper, and will remain as further study. Nevertheless, we remark that for the special case mentioned above, one could choose the generator matrix as

$$(4.25) \qquad \lambda(q) = \begin{pmatrix} \lambda_{00}(q) & \lambda_{01}(q) \\ \lambda_{10}(q) & \lambda_{11}(q) \end{pmatrix}.$$

where

$$\lambda_{01}(q) = \lambda_{10}(q) = \begin{cases} \frac{\delta}{\varepsilon}\left|\frac{V_{01}(q)}{V_{00}(q) - V_{11}(q)}\right|, & \text{if } |V_{00}(q) - V_{11}(q)| > 1 \\ \frac{\delta}{\varepsilon}|V_{01}(q)|, & \text{otherwise.} \end{cases},$$

and
$$\lambda_{00}(q) = \lambda_{11}(q) = -\lambda_{01}(q).$$

Note that when the gap $|V_{00}(q) - V_{11}(q)|$ is large (bigger than 1), the transition rate is modified so that the hops happen less frequently.

5. Conclusion

A diabatic surface hopping algorithm is proposed based on time dependent perturbation theory and semiclassical analysis. It captures the correct asymptotic limit in the weak coupling regime. The validity of the scheme is also verified in the avoided crossing regime for small ε near the transition zone. The diabatic potential energy surfaces remain smooth in the transition zone comparing to the adiabatic energy surfaces. It seems to us a good strategy for application would be a mix of both adiabatic and diabatic representations, i.e., to apply adiabatic representation away from the transition zone and switch to diabatic picture near it. For future works, rigorous asymptotic analysis of the algorithm would be interesting and better Monte Carlo sampling algorithm is also worth exploring.

References

[1] W. Bao, S. Jin, and P. A. Markowich. On time-Splitting spectral approximations for the Schrödinger equation in the semiclassical regime. *Journal of Computational Physics*, 175(2):487–524, 2002.

[2] Z. Cai and J. Lu. A surface hopping Gaussian beam method for high-dimensional transport systems, 2017. preprint, arXiv:1703.06116.

[3] W. Domcke and G. Stock. Theory of ultrafast nonadiabatic excited-state processes and their spectroscopic detection in real time. *Advances in Chemical Physics, Volume 100*, pages 1–169, 1997.

[4] E. Fabiano, G. Groenhof, and W. Thiel. Approximate switching algorithms for trajectory surface hopping. *Chemical Physics*, 351(1):111–116, 2008.

[5] G. A. Hagedorn and A. Joye. Landau-Zener transitions through small electronic eigenvalue gaps in the Born-Oppenheimer approximation. In *Annales de l'IHP Physique théorique*, volume 68, pages 85–134, 1998.

[6] C. F. Kammerer and C. Lasser. Propagation through generic level crossings: A surface hopping semigroup. *SIAM Journal on Mathematical Analysis*, 40(1):103–133, 2008.

[7] K. G. Kay. Integral expressions for the semiclassical time-dependent propagator. *The Journal of Chemical Physics*, 100(6):4377–4392, 1994.

[8] K. G. Kay. The Herman–Kluk approximation: Derivation and semiclassical corrections. *Chemical Physics*, 322(1):3–12, 2006.

[9] H. Köuppel, W. Domcke, and L. Cederbaum. Multimode molecular dynamics beyond the Born–Oppenheimer approximation. *Advances in Chemical Physics, Volume 57*, pages 59–246, 1984.

[10] L. Landau. Zur Theorie der Energieübertragung. II. *Physikalische Zeitschrift der Sowjetunion*, 2(46):1–13, 1932.

[11] B. R. Landry and J. E. Subotnik. Communication: Standard surface hopping predicts incorrect scaling for Marcus' golden-rule rate: The decoherence problem cannot be ignored. *The Journal of Chemical Physics*, 135:191101, 2016.

[12] C. Lasser, T. Swart, and S. Teufel. Construction and validation of a rigorous surface hopping algorithm for conical crossings. *Commun. Math. Sci.*, 5(4):789–814, 12 2007.

[13] C. Lasser and S. Teufel. Propagation through conical crossings: An asymptotic semigroup. *Communications on Pure and Applied Mathematics*, 58(9):1188–1230, 2005.

[14] J. Lu and X. Yang. Frozen Gaussian approximation for high frequency wave propagation. *Communications in Mathematical Sciences*, 9:663–683, 2010.

[15] J. Lu and X. Yang. Convergence of frozen Gaussian approximation for high-frequency wave propagation. *Communications on Pure and Applied Mathematics*, 65(6):759–789, 2012.

[16] J. Lu and Z. Zhou. Frozen Gaussian approximation with surface hopping for mixed quantum-classical dynamics: A mathematical justification of fewest switches surface hopping algorithms. *arXiv preprint arXiv:1602.06459*, 2016.

[17] J. Lu and Z. Zhou. Improved sampling and validation of frozen Gaussian approximation with surface hopping algorithm for nonadiabatic dynamics. *The Journal of Chemical Physics*, 145(12):124109, 2016.

[18] R. A. Marcus. On the theory of oxidation-reduction reactions involving electron transfer. I. *The Journal of Chemical Physics*, 24(5):966–978, 1956.

[19] U. Müller and G. Stock. Surface-hopping modeling of photoinduced relaxation dynamics on coupled potential-energy surfaces. *The Journal of Chemical Physics*, 107(16):6230–6245, 1997.

[20] E. Neria and A. Nitzan. Semiclassical evaluation of nonadiabatic rates in condensed phases. *The Journal of Chemical Physics*, 99(2):1109–1123, 1993.

[21] F. Plasser, G. Granucci, J. Pittner, M. Barbatti, M. Persico, and H. Lischka. Surface hopping dynamics using a locally diabatic formalism: Charge transfer in the ethylene dimer cation and excited state dynamics in the 2-pyridone dimer. *The Journal of Chemical Physics*, 137(22):22A514, 2012.

[22] D. Robert. On the Herman–Kluk semiclassical approximation. *Reviews in Mathematical Physics*, 22(10):1123–1145, 2010.

[23] T. Romero, A. Aguilar, and F. Gadea. Towards the Ab initio determination of strictly diabatic states, study for (NaRb)+. *The Journal of Chemical Physics*, 110(13):6219–6228, 1999.

[24] T. Swart and V. Rousse. A mathematical justification for the Herman-Kluk propagator. *Communications in Mathematical Physics*, 286(2):725–750, 2009.

[25] J. C. Tully. Molecular dynamics with electronic transitions. *The Journal of Chemical Physics*, 93(2):1061–1071, 1990.

[26] J. C. Tully. Mixed quantum–classical dynamics. *Faraday Discussions*, 110:407–419, 1998.

[27] J. C. Tully. Nonadiabatic dynamics. In D. L. Thompson, editor, *Modern Methods for Multidimensional Dynamics Computations in Chemistry*, page 61. World Scientific, 1998.

[28] J. C. Tully and R. K. Preston. Trajectory surface hopping approach to nonadiabatic molecular collisions: the reaction of H+ with D2. *The Journal of Chemical Physics*, 55(2):562–572, 1971.

[29] R. Valero and D. G. Truhlar. A diabatic representation including both valence nonadiabatic interactions and Spin–Orbit effects for reaction dynamics. *The Journal of Physical Chemistry A*, 111(35):8536–8551, 2007.

[30] L. Wang, A. E. Sifain, and O. V. Prezhdo. Communication: Global flux surface hopping in Liouville space. *The Journal of Chemical Physics*, 143(19):191102, 2015.

[31] L. Wang, A. E. Sifain, and O. V. Prezhdo. Fewest switches surface hopping in Liouville space. *The Journal of Physical Chemistry Letters*, 6(19):3827–3833, 2015.

[32] C. Zener. Non-adiabatic crossing of energy levels. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 137(833):696–702, 1932.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN-MADISON, MADISON, WI 53706, USA
E-mail address: di@math.wisc.edu

DEPARTMENTS OF MATHEMATICS, PHYSICS, AND CHEMISTRY, DUKE UNIVERSITY, DURHAM, NC 27708, USA
E-mail address: jianfeng@math.duke.edu